

α -Olefin Biosynthesis in Cyanobacteria

by

Daniel Mendez-Perez

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Chemical and Biological Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2013

Date of final oral examination: 12/04/2013

The dissertation is approved by the following members of the Final Oral Committee:

Brian F. Pfleger, Associate Professor, Chemical and Biological Engineering

Tim S. Bugni, Assistant Professor, School of Pharmacy

Sean P. Palecek, Professor, Chemical and Biological Engineering

Eric V. Shusta, Professor, Chemical and Biological Engineering

John Yin, Professor, Chemical and Biological Engineering

α -Olefin Biosynthesis in Cyanobacteria

Daniel Mendez-Perez

Under the supervision of Professor Brian F. Pfleger

At the University of Wisconsin-Madison

Abstract

Limited fossil fuel resources and the increasing demand for energy are motivating the development of economically feasible alternatives for energy and renewable means of producing chemical feedstocks that are currently supplied by petroleum manufacturing. A recent development in the production of renewable fuels is the use of cyanobacteria as production hosts. Cyanobacteria are photosynthetic organisms that are able to utilize carbon dioxide as their sole carbon source and light as their energy source. Although cyanobacteria naturally produce hydrocarbons, the native yields from wild type strains are insufficient to be cost competitive with petroleum-derived chemicals. Therefore, a deeper understanding of the pathways involved in hydrocarbon biosynthesis in cyanobacteria is needed in order to rationally engineer and enhance this trait in modified organisms.

In this thesis we analyzed the hydrocarbon profile of the cyanobacterium *Synechococcus* sp. PCC 7002 (PCC 7002) and demonstrated the involvement of a gene (*ols* gene for olefin synthase) with modular organization, similar to a polyketide synthase, in the synthesis of medium chain olefins (1-nonadecene and 1, 14-nonadecadiene). Feeding studies suggested that

the putative enzyme used an elongation decarboxylation-mechanism to convert fatty acyl-acyl carrier proteins (fatty acyl-ACPs) to the α -olefins. We also studied PCC 7002's growth under various temperatures and demonstrated the involvement of a desaturase gene (*desE*) in the formation of the internal double bond in 1, 14-nonadecadiene in response to changes in temperature; suggesting that the olefins might play a role in responding to cold stress and maintaining membrane fluidity. After improving CO₂ delivery, increasing expression of the *ols* gene and expressing a thioesterase from *Geobacillus sp.* Y412MC10, a significant increase in olefin production was observed. The discovery and characterization of pathways involved in hydrocarbon biosynthesis, like the Ols pathway describe in this thesis, will be an important contribution to the development of sustainable routes for hydrocarbon production and rationally engineer their production in modified organisms.

Additionally, we developed a method for dissecting failed heterologous expression experiments in *E. coli* by designing a DNA cassette that couples translation of a target gene to a response gene that generates an easily monitored phenotype *in vivo*, such as antibiotic resistance or fluorescence. The translational coupling system was used to optimize conditions for expression of Ols in *E. coli*.

Acknowledgments

First, I would like to thank my advisor Brian Pflieger for giving me the opportunity to work in a very interesting and challenging research field like synthetic biology, despite the fact that I didn't have any microbiology training. I really appreciate the guidance, support and patience during my graduate studies, I was really fortunate to join a lab where I always exposed to new ideas and new directions, but still had the freedom to make my own decisions.

Also to the current and past members of the Pflieger lab: Dan Agnew, Matt Begemann, Jeff Cameron, Ryan Clark, Matt Copeland, Spencer Hoover, Chris Jones, Travis Korosh, Rebecca Lennen, Andrew Markley, Mark Politz, Thomas Raines, Jackie Rand, Josh Thiede and Tyler Youngquist for making our lab a great place to work and learn.

I would like to thank the talented undergraduate students who have worked with me over the years: Nico Herman, Victor Orlor, Suman Gunasekaran, Charles Olmsted, Carlos Mates-Martinez, Joe Villanueva and Cedric Kovacs-Johnson. Your motivation and hard work made my research a lot easier and more enjoyable.

Also, Ron Raines and Ben Caes from the Biochemistry department, Helen E. Blackwell and Aaron Crapster from the Chemistry department, Mick McGee from GLBRC and Jackie Cooper were incredibly helpful with the characterization of cyanobacterial hydrocarbons presented in Chapter 2 and the GC-MS analyses presented in this thesis.

I would also like to thank the CBE department and the members of my thesis committee Tim Bugni, Sean Palecek, Eric Shusta and John Yin.

To friends and housemates: Sarah Forzley, Collin Timm, Andrea Timm, Joe Samaniuk, Tyler Youngquist, Jackie Rand, Becky Carlton, Mark Politz, Ryan Clark, Josh Selekman, Grace Shrader and Patrick Flanagan for all the frisbee games, tv nights and great parties at 727 W. Main.

Finally, I would like to thank my family for their constant love and support. I wouldn't have been able to get through grad school, or life, without you.

Table of contents

Abstract.....	i
Acknowledgments.....	iv
Table of contents.....	vi
List of Figures.....	ix
List of Tables.....	xi
Chapter 1 Introduction.....	1
1.1 Motivation.....	1
1.2 Cyanobacteria.....	2
1.2.1 Photosynthesis and carbon fixation.....	3
1.2.2 Cyanobacterial hydrocarbons.....	5
1.2.3 Mechanisms for hydrocarbon biosynthesis.....	6
1.3 Overview of Thesis.....	9
1.4 References.....	10
Chapter 2 Alkene biosynthesis in the cyanobacterium <i>Synechococcus sp.</i> PCC 7002.....	11
2.1 Introduction.....	11
2.2 Results.....	14
2.2.1 Alkene characterization.....	14
2.2.2 Fatty acid profile in <i>Synechococcus sp.</i> PCC 7002.....	20
2.2.3 Identification of a gene involved in alkene biosynthesis.....	21
2.2.4 Genetic studies.....	26
2.2.5 Elongation of fatty acids is required for the formation of the alkenes.....	27
2.2.6 Acyl-ACPs as precursors for olefin biosynthesis.....	29
2.3 Conclusions.....	31
2.4 Methods.....	32
2.4.1 Gene inactivation/promoter replacement.....	32
2.4.2 Hydrocarbon extraction and GC-MS analysis.....	36
2.4.3 DMDS derivatization.....	37
2.4.4 Permanganate/periodate oxidation.....	37
2. 5 References.....	38
Chapter 3: A desaturase gene involved in the formation of 1,14 nonadecadiene produced by <i>Synechococcus sp.</i> PCC 7002.....	43
3.1 Introduction.....	43
3.2 Results.....	44
3.2.1 Identification of a desaturase gene involved in alkene unsaturation.....	44
3.2.2 Precursors for 1, 14-nonadecadiene biosynthesis.....	46
3.2.3 Growth in the presence of H ₂ O ₂ and under nutrient limitation conditions.....	50
3.2.4 Effect of temperature in alkene unsaturation.....	52

3.2.5 Hydrocarbon localization	54
3.3 Conclusions.....	56
3.4 Methods.....	56
3.4.1 Media and growth conditions.....	56
3.4.2 Strain construction.....	57
3.4.3 Lipid analysis (GC/MS)	59
3.5 References.....	60
Chapter 4: A translational-coupling DNA cassette for monitoring protein translation in <i>Escherichia coli</i>	63
4.1 Introduction.....	63
4.2 Results.....	65
4.2.1 Construction of a translation-coupling DNA cassette for monitoring protein translation in <i>Escherichia coli</i>	65
4.2.2 Test for RFP translation using antibiotic resistance as the response signal	67
4.2.2.1 Optimizing growth conditions using antibiotic resistance as the response signal to test for translation	69
4.2.3 Detecting differences in expression of the target gene.....	71
4.3 Conclusions.....	74
4.4 Methods.....	76
4.4.1 Plasmid construction and oligonucleotides	76
4.4.2 Microtiter plate growth assays	77
4.4.3 Protein isolation.....	78
4.5 References.....	81
Chapter 5: <i>E. coli</i> and <i>Synechococcus sp.</i> PCC 7002 engineering	83
5.1 Introduction.....	83
5.2 Results.....	84
5.2.1 <i>E. coli</i> engineering.....	84
5.2.1.1 Ols expression in <i>E. coli</i>	84
5.2.1.2 Acyl carrier protein (ACP) phosphopantetheinylation	88
5.2.1.3 Identification of a phosphopantetheinyl transferase (PPT) from <i>Synechococcus sp.</i> PCC 7002.....	93
5.2.1.4 Increasing substrate availability in <i>E. coli</i> : acyl-ACP synthetase expression.....	95
5.2.1.5 Expression of <i>cysDNC</i> genes	96
5.2.1.6 Future work and conclusions	97
5.2.2 <i>Synechococcus sp.</i> PCC 7002 engineering.....	100
5.2.2.1 Construction of a chimeric olefin synthase.....	100
5.2.2.2 <i>desE</i> overexpression.....	101
5.2.2.3 Thioesterase (TE) expression.....	102
5.2.2.4 Future work and conclusions	106
5.4 Methods.....	107
5.4.1 <i>E. coli</i> plasmid construction and oligonucleotides.....	107
5.4.2 Cyanobacterial constructs.....	110

5.4.3 Microtiter plate growth assays and protein purifications	110
5.4.4 Preparation of fluorescently labeled CoA and Bodipy-CoA in vitro reactions	112
5.5 References.....	112
Chapter 6: Conclusions	119
6.1 α -olefin biosynthesis in <i>Synechococcus sp.</i> PCC 7002.....	119
6.2 Challenges of engineering cyanobacteria	123
6.3 Translational coupling cassette and <i>E. coli</i> engineering.....	124
6.4 References.....	125

List of Figures

Figure 1.1	Light reactions during photosynthesis.....	3
Figure 1.2	Calvin cycle.....	4
Figure 1.3	Metabolic pathways in cyanobacteria involved in the production of hydrocarbons	6
Figure 2.1	Comparison of hydrocarbon extracts from the wild type and mutant strains of PCC 7002.....	15
Figure 2.2	Mass spectra for the hydrocarbons found in the PCC 7002 lipid extracts.....	16
Figure 2.3	Observed mass spectra of the DMDS derivatives	17
Figure 2.4	Chemical processing of purified 1,14-nonadecadiene and observed mass spectra of the derivative of 1,14-nonadecadiene	18
Figure 2.5	Nuclear magnetic resonance for 1, 14-nonadecadiene	20
Figure 2.6	Final modules involved in Curacin A biosynthesis and terminal double bond formation	23
Figure 2.7	Domain organization and proposed mechanism of a putative olefin synthase encoded by <i>ols</i> . Partial sequence alignments of OIs with consensus polyketide synthase domain motifs.....	25
Figure 2.8	Formation of 1-octadecene after feeding cyanobacterial cultures with heptadecanoic acid	28
Figure 2.9	Comparison of hydrocarbon extracts from wild type (WT), <i>ols</i> deletion mutant (Δols), and Φ (PpsbA- <i>ols</i>) strains of PCC 7002 supplemented with pentadecanoic (C15:0) acid.....	29
Figure 2.10	Fatty acid and 1-octadecene content in the Δaas mutant strains and wild type strain after the addition of pentadecanoic acid.....	31
Figure 2.11	Schematic outlining steps to construct PCC 7002 mutant strains	35
Figure 2.12	Confirmation of PCC 7002 mutants by PCR	36
Figure 3.1	Comparison of the hydrocarbon composition from the wild type (wt), <i>desE</i> knockout ($\Delta desE$), $\Delta desE$ complemented ($\Delta desE^+$) and <i>desE</i> upregulated (<i>desE</i> -up) strains.....	46
Figure 3.2	Hydropathy plot for DesE	48
Figure 3.3	Proposed route for alkene biosynthesis in <i>Synechococcus sp.</i> PCC7002 and comparison of the fatty acid profiles of PCC 7002 and mutants.....	50
Figure 3.4	Comparison of the hydrocarbon composition from cultures grown in Media A and nutrient limitation conditions	52
Figure 3.5	Comparison of the hydrocarbon composition from cultures grown at different temperatures.....	53
Figure 3.6	Growth curves for the wild type and Δols strain at different temperatures	54
Figure 3.7	Hydrocarbons content from a wild type PCC 7002 in different cell fractions	55

Figure 4.1	Schematic of translational coupling cassette.....	67
Figure 4.2	Demonstration of translational coupling using RFP as target gene and antibiotic resistance markers as response genes.....	68
Figure 4.3	Demonstration of translational coupling using pRFP ^{nf} -CM as target gene and antibiotic resistance markers as response genes	70
Figure 4.4	Translational coupling functions at high and low levels of expression	72
Figure 4.5	Translational coupling generates a proportional response.	73
Figure 4.6	Comparison of the relative fluorescence for the RBS library constructs ..	74
Figure 5.1	Use of translational coupling to assay for translation of a foreign gene ...	86
Figure 5.2	Use of translational coupling to optimize expression conditions of the <i>ols</i> gene.....	88
Figure 5.3	Activation of Acyl carrier protein (ACP) by phosphopantetheinyl transferases	89
Figure 5.4	a) SDS-PAGE of Ols containing fluorescently labeled ACP domains scanned on a Typhoon imager.....	91
Figure 5.5	Phosphopantetheinylation of Ols by Sfp <i>in vivo</i> . <i>E. coli</i> cultures expressing Ols and with (+) or without (-) expressing Sfp	92
Figure 5.6	Phosphopantetheinylation of ACP1 by Sfp <i>in vitro</i>	93
Figure 5.7	SDS-PAGE of Ols containing fluorescently labeled ACP domains activated by PPT ₇₀₀₂	94
Figure 5.8	Octadecanoic content in cell pellets after addition of octadecanoic acid and expression of an acyl-ACP synthetase from <i>Synechocystis sp.</i> PCC 6803 or <i>Synechococcus elongatus</i> PCC 7942	96
Figure 5.9	3'-phosphoadenosine-5'-phosphosulfate (PAPS) biosynthetic pathway in <i>E. coli</i>	97
Figure 5.10	Potential olefins that can be produced using Ols by using a different loading module.....	100
Figure 5.11	Olefin production at 38°C, bubbling air.....	102
Figure 5.12	Fatty acid composition after expression of acyl-ACP thioesterases from <i>E. coli</i> (<i>tesA</i>), <i>Umbellularia californica</i> (BTE) and <i>Geobacillus sp.</i> Y412MC10 (Geo).....	105
Figure 5.13	Alkene production in PCC 7002 after expression of thioesterases	105
Figure 6.1	Activities involved in alkene biosynthesis	122
Figure 6.2	α -olefin production in <i>Synechococcus sp.</i> PCC 7002	123

List of Tables

Table 1.1	Mechanisms for fatty acid-derived hydrocarbon biosynthesis in bacteria	7
Table 2.1	Fatty acid profile for wild type PCC 7002 at 38°C	21
Table 2.2	Fatty acid profile for wild type PCC 7002 and Δols at 38°C	26
Table 2.3	Olefin production and olefin synthase expression	27
Table 2.4	Oligonucleotides used in Chapter 2.....	33
Table 3.1	Fatty acid profiles for the wild type and $\Delta desE$ at 22°C and 38°C	47
Table 3.2	Oligonucleotides used in Chapter 3.....	58
Table 3.3	Cyanobacterial strains used in Chapter 3	59
Table 4.1	Plasmids constructed for and used in Chapter 4.....	79
Table 4.2	Oligonucleotides used in Chapter 4.....	80
Table 5.1	<i>E. coli</i> plasmids and strains used in Chapter 5	107
Table 5.2	Oligonucleotides used in Chapter 5.....	108

Chapter 1: Introduction

1.1 Motivation

Limited fossil fuel resources and the increasing demand for energy are motivating the development of economically feasible alternatives for energy and renewable means of producing chemical feedstocks that are currently supplied by petroleum manufacturing (1). Bacterial hydrocarbons are a promising alternative because besides their high energy content (2), they would be compatible with current infrastructure and their separation would not require the use of energy-intensive processes as they would be insoluble in the producing media.

A recent development in the production of renewable fuels is the use of cyanobacteria as production hosts. Cyanobacteria are photosynthetic organisms that are able to utilize carbon dioxide as their sole carbon source and light as their energy source. The use of CO₂-fixing organisms has the advantage that no exogenous carbon source (such as glucose) would be required. Also, sequestration of CO₂ by cyanobacteria would utilize the greenhouse gas for beneficial purposes and close an energy cycle between fuel combustion (CO₂ emission) and subsequent CO₂ fixation (3). Therefore, the development of engineered cyanobacteria offers the promise of renewable and more carbon neutral alternatives for petroleum-based fuels. Although cyanobacteria naturally produce hydrocarbons, it is important to recognize that their extremely low native yields from wild type strains are insufficient to be cost competitive with petroleum-derived chemicals. Metabolic engineering, defined as the redirection of one or more enzymatic reactions to improve the production of a desired compound (4), has the potential to increase yields by transferring product-specific enzymes or entire metabolic pathways from one organism

to another and by rewiring native metabolic pathways. But in spite of its enormous potential, the incorporation of metabolic engineering tools into hydrocarbon production has faltered because limited genetic information is available about their biosynthesis. Therefore, a deeper understanding of the pathways involved in hydrocarbon biosynthesis in cyanobacteria is needed in order to rationally engineer and enhance this trait in modified organisms.

In this thesis we characterize the biosynthesis of 1-alkenes (also known as α -olefins) in the cyanobacterium *Synechococcus sp.* PCC 7002. Since olefins are one of the primary products of petroleum and gas refining (short chain olefins are used as chemical building blocks and medium chain length olefins as diesel or jet fuel additives), it is important to find and develop sustainable routes for their production.

1.2 Cyanobacteria

Cyanobacteria are organisms that harbor, within a prokaryotic cell, a photosynthetic apparatus with remarkable similarities in function, structural and molecular respects to the photosynthetic apparatus contained in the eukaryotic chloroplast (5). They belong to the great subclass of gram-negative bacteria and constitute the largest, most diverse, and most widely distributed group of photosynthetic prokaryotes. Because they are primitive organisms with the ability to grow in environments that favor the preservation of organic matter, such as hypersaline or reducing environments, they are often suggested as source material for the organic matter associated with ancient sediments (6).

1.2.1 Photosynthesis and carbon fixation

During photosynthesis, cyanobacteria use the energy from the sun to generate ATP and NADPH and these molecules are then used to fix carbon dioxide. Photosynthesis includes light-dependent reactions and reactions that are not directly energized by light (also referred as dark reactions).

The light reactions happen in the thylakoid membrane and convert light energy to chemical energy (Figure 1.1). Light energy in the form of photons is captured by thylakoid pigments and transferred to chlorophylls within the reactive core of the photosystems. Photosystem II (PS II) contains an active site that carries out the water splitting reaction, generating oxygen, H^+ and low potential electrons. Electrons help translocation of protons from the cytosol to the thylakoid lumen when they pass to plastoquinone (PQ) and through the cytochrome complex. Additional absorption of solar energy in photosystem I (PS I) lowers the electron potential. Ferredoxin (Fd) accepts the electrons from PSI and donates them for NADPH generation. ATP is generated when protons in the thylakoid lumen are returned to the cytoplasm through ATP synthase.

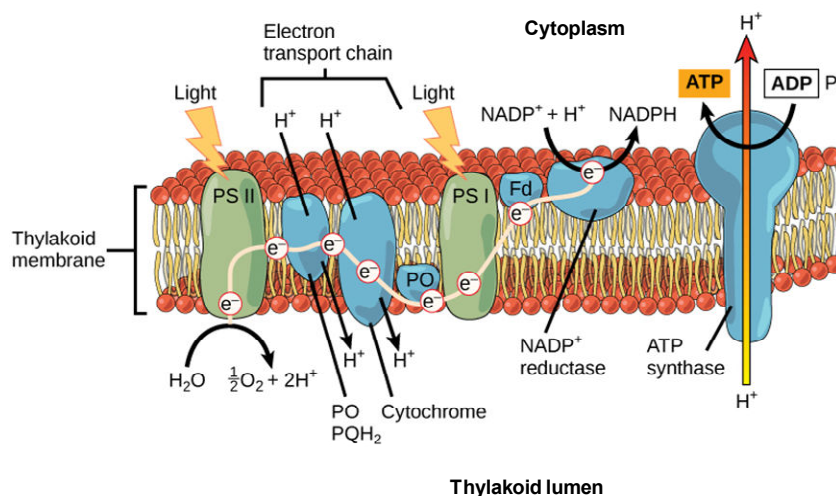


Figure 1.1 Light reactions during photosynthesis. Cyanobacteria use the energy from the sun to generate ATP and NADPH. (Figure taken from <http://cnx.org/content/m45452/latest/?collection=coll11487/latest>)

In the dark reactions ATP and CO₂ are used to generate a three-carbon compound called glyceraldehyde-3-phosphate (G3P), but almost immediately two of these join to form a glucose molecule. This reaction doesn't directly need light in order to occur, but it does need the products of the light reaction (ATP and NADPH). The process is known as the Calvin cycle and there are three phases in the cycle: carbon fixation, reduction and regeneration (Figure 1.2).

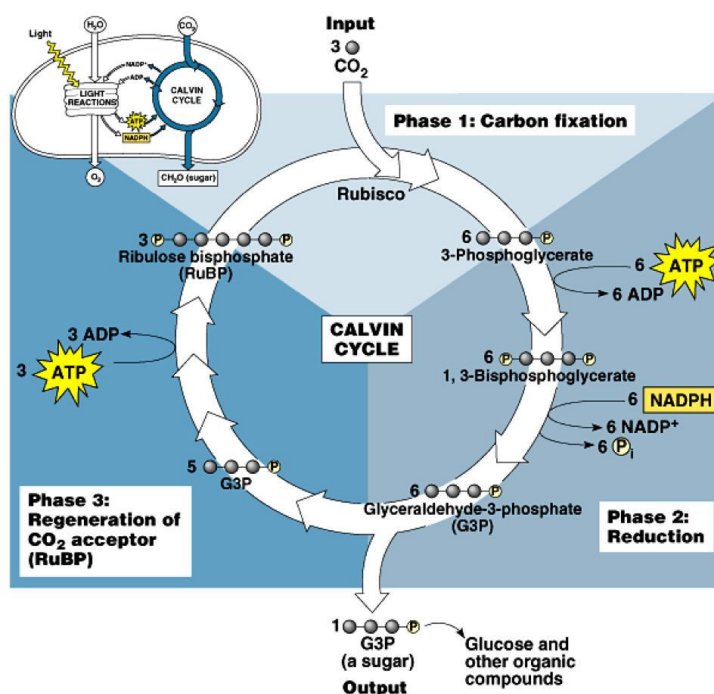


Figure 1.2 Calvin cycle. ATP and CO₂ are used to generate the three-carbon compound glyceraldehyde-3-phosphate. (Figure taken from <http://www.wellsphere.com/heart-health-article/plants-produce-carbs-from-air-calvin-cycle/1163923>)

The first step, catalyzed by RuBisCO (Ribulose-1,5-bisphosphate carboxylase oxygenase), uses CO₂ for the carboxylation of a 5-carbon compound (Ribulose-1,5-bisphosphate, RuBP) to generate a six-carbon intermediate that immediately splits to form two 3-phosphoglycerate (3PGA) molecules. During the reduction step, ATP and NADPH are used to convert 3PGA to

glyceraldehyde-3-phosphate (G3P), the precursor for glucose. In the regeneration step, ATP is used to convert part of the G3P back to RuBP, the acceptor for CO₂, thereby completing the cycle. Intermediates generated within the Calvin cycle enter central metabolic pathways as substrates in the synthesis of carbohydrates, including glucose.

1.2.2 Cyanobacterial hydrocarbons

Currently, the most common feedstocks used for the production of biofuels are derived from carbon present in plant biomass in the form of structural sugar polymers such as cellulose and hemicelluloses, energy storage sugar polymers in the form of starch, aromatic polymers in the form of lignin and oils stored in plant seeds or vegetable tissues. The major constraint of these feedstocks is based on the competition with our food sources for farmland and water. To overcome these limitations, non-food biomass sources for energy supply (including cyanobacteria) have gained increasing attention in the past years.

Cyanobacteria are one of the most promising feedstocks for bioenergy generation; they contain considerable amounts of lipids (mainly present in the thylakoid membranes), they possess higher growth rates and photosynthetic levels compared to algae and higher plants, they grow easily with basic nutritional requirements (CO₂, water, mineral salts and light as the only energy source), their cultivation is therefore relatively simple and inexpensive (7). Another advantage is that cyanobacteria, being prokaryotes, can be genetically engineered more easily than eukaryotic algae.

As shown in Figure 1.3, biosynthesis of hydrocarbons is integrated with the natural metabolism of cyanobacteria, relying upon both primary and secondary metabolic pathways. In particular,

two major metabolic nodes control the carbon flux for their biosynthesis: acetyl-CoA and acyl-ACP nodes (1). These primary metabolic nodes distribute the carbon flux to the secondary pathways essential for cell growth and survival; for example acyl-ACP molecules can be diverted from membrane biosynthesis to produce free fatty acids and hydrocarbons (alkanes or alkenes). By consuming these essential metabolites, hydrocarbon biosynthesis often competes with cell growth, presenting a challenge for metabolic engineering applications to balance cell growth and hydrocarbon production.

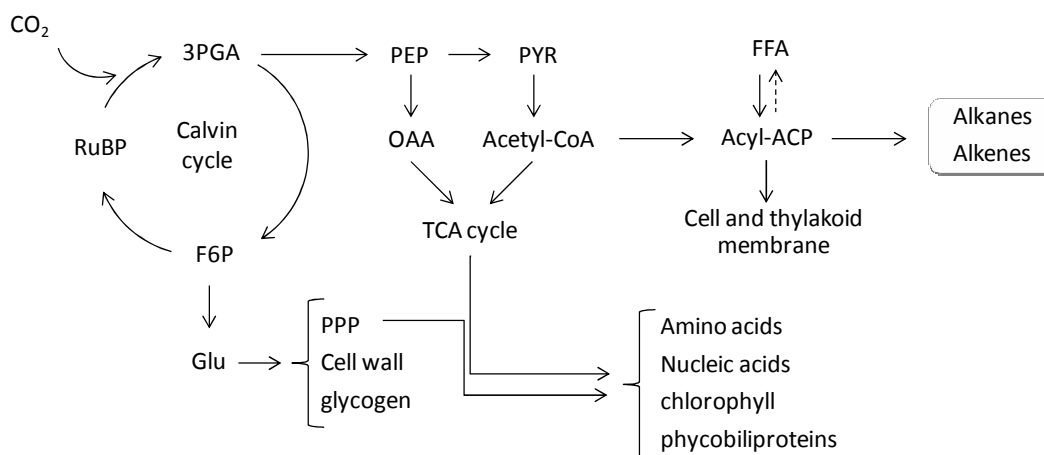


Figure 1.3 Metabolic pathways in cyanobacteria involved in the production of hydrocarbons (8). A dashed line indicates a pathway that is not present in cyanobacteria.

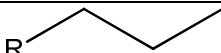
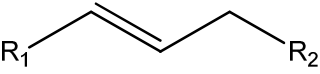
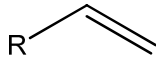
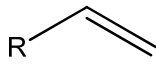
1.2.3 Mechanisms for hydrocarbon biosynthesis

Hydrocarbon accumulation has been observed in both photosynthetic and non-photosynthetic bacteria. The composition of bacterial hydrocarbons is complex, including n-alkanes, alkenes, and branched hydrocarbons with lengths ranging from C₁₅ up to C₃₆ (9). Non-photosynthetic bacteria accumulate long-chain hydrocarbons (C₂₇-C₂₉) and photosynthetic bacteria accumulate

hydrocarbons with shorter chains (C₁₇-C₂₀). Photosynthetic bacteria and anaerobic non-photosynthetic bacteria also produce isoprenic units of pristane and phytane (10).

A few pathways have been described for hydrocarbon biosynthesis via fatty acids in prokaryotes (see Table 1.1), including the decarboxylation of fatty acids via the cytochrome P450 fatty acid decarboxylase to form terminal olefins in *Jeotgalicoccus sp.* (11) and the head-to-head condensation of fatty acids to form long chain alkenes in *Micrococcus luteus* ATCC 4698 (12). Although the resulting hydrocarbons from these pathways are very similar (sometimes even identical) to those produced by cyanobacteria, the biosynthetic steps involved seemed to be distinctively different.

Table 1.1 Mechanisms for fatty acid-derived hydrocarbon biosynthesis in bacteria

<i>Mechanism</i>	<i>Enzymes involved</i>	<i>Products</i>	<i>Reference</i>
Decarbonylation	Fatty acyl-ACP reductase (FAAR) and aldehyde-deformylating oxygenase (ADO)		(13)
Head to head condensation	Thiolase, alpha/beta-hydrolase, AMP-dependent ligase/ synthase, and short-chain dehydrogenase (OleABCD)		(14)
Decarboxylation	P450 fatty acid decarboxylase		(11)
Elongation-decarboxylation	Unknown		n. a.

In cyanobacteria (*Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942) only one pathway has been characterized for the formation of alkanes (13). The pathway involves the conversion of fatty acyl-ACPs to fatty acyl-aldehydes by a fatty acyl-ACP reductase (FAAR), followed by an oxygen dependent conversion of the aldehyde to produce an odd-chain length alkane via an aldehyde-deformylating oxygenase (ADO); an alkene can also be produced if the fatty acid possess preexisting unsaturations (internal double bonds) (15). FAAR can be found in a wide range of organism including plants, eukaryotic algae, bacteria and humans (16) but the step involving the ADO enzyme is unique to cyanobacteria (15).

Besides alkanes and alkenes with internal double bonds, 1-alkenes (α -olefins) have also been detected in various cyanobacterial strains (6, 17). Since the FAAR/ADO pathway would not explain the formation of these compounds (and FAAR/ADO homologs are not present in all cyanobacterial genomes), other pathways must be involved in their biosynthesis. Two distinct mechanisms starting from fatty acids can result in the formation of 1-alkenes: decarboxylation of fatty acids (like the one observed in *Jeotgalicoccus* sp. (11)) or elongation-decarboxylation.

Elongation-decarboxylation of fatty acids to form 1-alkenes has been observed in algae (18) and cyanobacteria (17) and in general, the precursor fatty acid (usually a C₁₆ or C₁₈) is elongated by the addition of C₂ units from malonyl-CoA and the product serve as substrate for the decarboxylation system from where the hydrocarbons are finally released (19). It is considered that the direct loss of the carboxyl carbon of a fatty acid is not mechanistically feasible without an electron-withdrawing group adjacent to the α -carbon (20). Elimination of CO₂ from carboxylic acids requires high energy and, therefore, has to be activated by a β -substituent able to stabilize the negative charge generated by CO₂ release (9). Accordingly, it has been proposed that the activated fatty acid derivatives are the intermediates in the decarboxylation reaction

leading to hydrocarbons (10). Since genes involved in elongation-decarboxylation have not been identified yet, recombinant DNA approaches have not been used to produce the enzyme for characterization.

1.3 Overview of Thesis

In this thesis we analyze the biosynthesis of 1-alkenes (α -olefins) in the unicellular marine cyanobacterium *Synechococcus sp.* strain PCC 7002. This strain was first isolated from the Gulf of Mexico (21) and it is among the fastest growing of all cyanobacteria, with a doubling time under optimal conditions of about 3.5 hours (22). It has been demonstrated that it possesses an efficient, naturally occurring mechanism for the uptake and integration of exogenous DNA (23) and therefore, it is effectively transformable (24). It is facultatively photoheterotrophic with glycerol as substrate and is extremely tolerant of high light intensities. It can be grown to single colonies on solid growth media and therefore some molecular genetic techniques are available.

In Chapter 2 we characterize the structure of the two alkenes synthesized by *Synechococcus sp.* PCC 7002, 1-nonadecene (C19:1) and 1, 14-nonadecadiene (C19:2), and show that they contain a terminal double bond. A gene (referred as *ols* gene for olefin synthase) with modular organization, similar to a polyketide synthase, was identified and genetic studies were used to prove its involvement in alkene biosynthesis. The domain architecture of the *ols* gene and feeding studies suggest that the putative enzyme uses an elongation decarboxylation-mechanism to convert C18 fatty acid acyl-acyl carrier proteins (fatty acyl-ACPs) to the alkenes.

In Chapter 3 we demonstrate the involvement of a desaturase gene (*desE*) in the formation of the internal double bond in the 1, 14-nonadecadiene synthesized by PCC 7002. The amino acid

sequence encoded by the *desE* gene shows a high degree of similarity to $\Delta 9$ desaturases suggesting that its most likely substrate is a C14 fatty acid, which after elongation to a C18:1($\Delta 13$) fatty acid would serve as the precursor for the formation of the hydrocarbon with the internal double bond at position 14th. An increase in C19:2 production at low temperatures suggests that hydrocarbons might play a role in responding to cold to maintain membrane fluidity.

In Chapter 4 we describe a method to quickly determine whether *Escherichia coli* is capable of expressing the product of any target gene by coupling translation of a target gene to a detectable response gene. The translational coupling cassette was designed to encode for a mRNA sequence that forms a secondary structure in the absence of translation and contains the translational start sequence of a detectable response gene. The translational coupling method was tested and only when the target gene was fully translated was the response gene observed. The translational coupling system was used in Chapter 5 to determine that the Ols protein was not actively translated in *E. coli* and to isolate the problem to some of the domains; we then optimized conditions for expressing a codon-optimized sequence variant. Also, some metabolic engineering strategies were used in *Synechococcus sp.* PCC 7002 and *E. coli* for the production of olefins using the OLS pathway

1.4 References

1. **Ruffing AM.** 2011. Engineered cyanobacteria: Teaching an old bug new tricks. *Bioengineered Bugs* **2**:136–149.
2. **Agarwal AK.** 2007. Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. *Progress in Energy and Combustion Science* **33**:233–271.

3. **Rabinovitch-Deere C a, Oliver JWK, Rodriguez GM, Atsumi S.** 2013. Synthetic biology and metabolic engineering approaches to produce biofuels. *Chemical reviews* **113**:4611–32.
4. **Keasling JD.** 2010. Manufacturing molecules through metabolic engineering. *Science (New York, N.Y.)* **330**:1355–8.
5. **Stanier RY, Cohen-Bazire G.** 1977. Phototrophic prokaryotes: the cyanobacteria. *Annual review of microbiology* **31**:225–74.
6. **Winters K, Parker PL, Baalen C Van.** 1969. Hydrocarbons of Blue-Green Algae: Geochemical Significance. *Science* **163**:467–468.
7. **Quintana N, Van der Kooy F, Van de Rhee MD, Voshol GP, Verpoorte R.** 2011. Renewable energy from Cyanobacteria: energy production optimization by metabolic pathway engineering. *Applied microbiology and biotechnology* **91**:471–90.
8. **Ruffing AM, Jones HDT.** 2012. Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnology and bioengineering* **109**:2190–9.
9. **Ladygina N, Dedyukhina EG, Vainshtein MB.** 2006. A review on microbial synthesis of hydrocarbons. *Process Biochemistry* **41**:1001–1014.
10. **Valderrama B.** 2004. Bacterial hydrocarbon biosynthesis revisited. *Studies in Surface Science and Catalysis* **151**:373–384.
11. **Rude M a, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A.** 2011. Terminal olefin (1-alkene) biosynthesis by a novel p450 fatty acid decarboxylase from *Jeotgalicoccus* species. *Applied and environmental microbiology* **77**:1718–27.
12. **Beller HR, Goh E-B, Keasling JD.** 2010. Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Applied and environmental microbiology* **76**:1212–23.
13. **Schirmer A, Rude MA, Li X, Popova E, Del Cardayre SB.** 2010. Microbial biosynthesis of alkanes. *Science* **329**:559–62.
14. **Sukovich DJ, Seffernick JL, Richman JE, Gralnick J a, Wackett LP.** 2010. Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Applied and environmental microbiology* **76**:3850–62.
15. **Li N, Chang W-C, Warui DM, Booker SJ, Krebs C, Bollinger JM.** 2012. Evidence for only oxygenative cleavage of aldehydes to alk(a/e)nes and formate by cyanobacterial aldehyde decarbonylases. *Biochemistry* **51**:7908–16.

16. **Doan TTP, Carlsson AS, Hamberg M, Bülow L, Stymne S, Olsson P.** 2009. Functional expression of five Arabidopsis fatty acyl-CoA reductase genes in Escherichia coli. *Journal of plant physiology* **166**:787–96.
17. **Goodloe RS, Light RJ.** 1982. Structure and composition of hydrocarbons and fatty acids from a marine blue-green alga, *Synechococcus* sp. *Biochimica et biophysica acta* **710**:485–492.
18. **Templier J, Largeau C, Casadevall E.** 1984. Mechanism of non-isoprenoid hydrocarbon biosynthesis in *Botryococcus braunii*. *Phytochemistry* **23**:1017–1028.
19. **Yong T, Largeau C, Casadevall E.** 1986. Biosynthesis Of Non-isoprenoid Hydrocarbons By The Microalga *Botryococcus-braunii* - Evidences For An Elongation-decarboxylation Mechanism - Activation Of Decarboxylation. *Nouveau Journal De Chimie-new Journal Of Chemistry* **10**:701–707.
20. **Dennis M, Kolattukudy PE.** 1992. A cobalt-porphyrin enzyme converts a fatty aldehyde to a hydrocarbon and CO. *Proceedings of the National Academy of Sciences of the United States of America* **89**:5306–10.
21. **Van Baalen C.** 1984. Studies on marine blue-green algae. *Botanica marina* **4**:129–139.
22. **Sakamoto T, Bryant D a.** 1998. Growth at low temperature causes nitrogen limitation in the cyanobacterium *Synechococcus* sp. PCC 7002. *Archives of microbiology* **169**:10–9.
23. **Stevens SE, Porter RD.** 1980. Transformation in *Agmenellum quadruplicatum*. *Proceedings of the National Academy of Sciences of the United States of America* **77**:6052–6.
24. **Akiyama H, Kanai S, Hirano M, Miyasaka H.** 1998. Nucleotide Sequence of Plasmid pAQ1 of Marine cyanobacterium *Synechococcus* sp . PCC7002. *DNA Research* **5**:127–129.

Chapter 2: Alkene biosynthesis in the cyanobacterium *Synechococcus sp.* PCC 7002

*Portions of this chapter were published in Applied and Environmental Microbiology (1)

2.1 Introduction

The increasing global demand for transportation fuel, which is projected to nearly double in the next two decades (2), in conjunction with the rapid accumulation of greenhouse gases in the environment (3) is encouraging the development of renewable fuels as alternatives to fossil fuels. At present, two types of biofuels are commercially available: ethanol from corn or sugarcane, and mixtures of fatty acid methylesters from vegetable oils (also known as biodiesel). Whereas ethanol might be an alternative for gasoline, biodiesel has had limited use because vegetable oils are also used for human consumption. Therefore, it would be of great interest to have alternatives for biodiesel that do not involve the usage of vegetable oils. Bacterial hydrocarbons are a promising alternative for biodiesel because besides its higher energy content (4), they would be compatible with current infrastructure and their production could yield crude oil without the contaminating sulfur that much petroleum out of the ground contains. Also, unlike ethanol, whose product separation is energy-intensive, hydrocarbons would be insoluble in the producing media and therefore would spontaneously separate during production. In particular, cyanobacterial hydrocarbons have received increasing attention as they can be directly produced from carbon dioxide and sunlight, bypassing the need for the use of biomass-derived sugars.

Although it has been known for decades that some species of cyanobacteria, including *Synechococcus sp.* PCC 7002 (formerly known as *Agmenellum quadruplicatum*), can synthesize hydrocarbons (5), until recently, very little was known about their biosynthesis. Enzymes responsible for producing three types of hydrocarbons (alkanes, internal olefins, and α -olefins) have recently been identified. Cyanobacterial alkanes can be derived from fatty acids by decarbonylation of the corresponding aldehydes (6). Alkenes with internal double bonds can be generated from the head-to-head condensation of fatty acids; genes involved in this pathway have been described for *Micrococcus luteus* (7). Recently, a P450 fatty acid decarboxylase was reported to be involved in α -olefin biosynthesis in *Jeotgalicoccus sp* (8).

This Chapter describes the results of the analysis of the hydrocarbon content from the cyanobacterium *Synechococcus sp.* PCC 7002 (PCC 7002) and demonstration of the involvement of a gene with modular organization, similar to a polyketide synthase, in the synthesis of medium chain α -olefins. Feeding studies suggested that the putative enzyme used an elongation-decarboxylation mechanism to convert fatty acyl-ACPs to α -olefins.

2.2 Results

2.2.1 Alkene characterization

It has been reported that PCC 7002 synthesizes two alkenes with 19 carbons (C₁₉), but nothing was known about their structure (5). In order to characterize the structure of these compounds, cultures were grown photosynthetically (140 μ E/m²/s) at 35°C in 100 ml of medium A to an optical density at 730 nm (OD₇₃₀) of 0.2, and the cell pellets were subjected to lipid extraction and analysis by gas chromatography-mass spectrometry (GC-MS). Two major peaks were

observed at 14.87 min and 15.15 min (Figure 2.1) whose mass spectra were consistent with a 19:2 (two unsaturations) and 19:1 (one unsaturation) hydrocarbon, respectively (Figure 2.2).

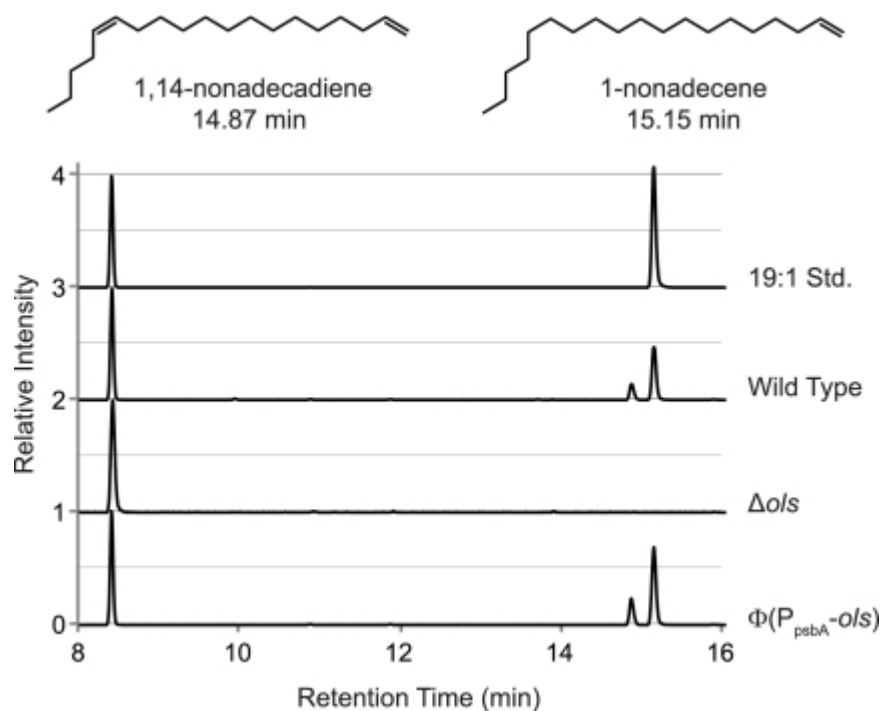


Figure 2.1. Comparison of hydrocarbon extracts from the wild type and mutant strains of PCC 7002, GC-MS signal was normalized to the height of an internal standard peak (hexadecane, 8.3 min). Two hydrocarbons, 1, 14-nonadecadiene and 1-nonadecene, were identified in extracts of the wild type and promoter replacement $\Phi(P_{psbA-ols})$ mutant but not in the extracts of the Δols mutant.

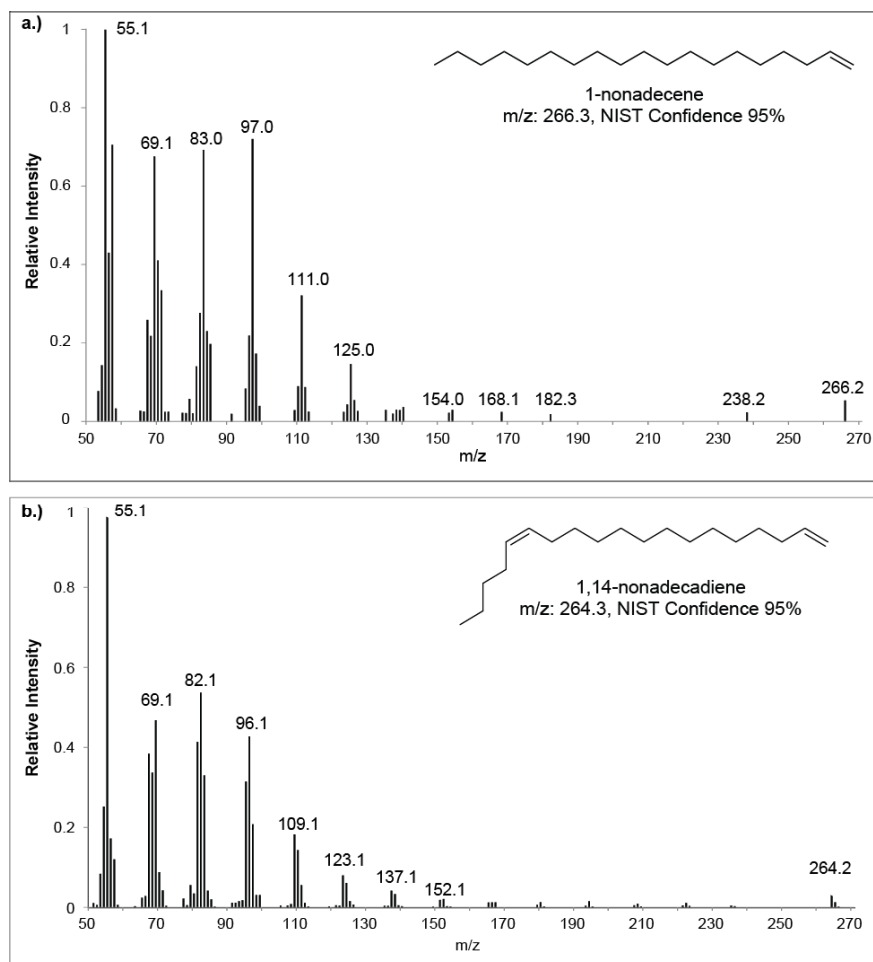


Figure 2.2. Mass spectra for the hydrocarbons found in the PCC 7002 lipid extracts, a) C19:1 hydrocarbon, b) C19:2 hydrocarbon

In order to determine the mechanism involve in the biosynthesis of these compounds, it is important to determine the location of the double bonds. As mentioned before, terminal double bonds are characteristic of decarboxylation mechanisms and internal double bonds are characteristic of head to head condensation mechanisms. Formation of dimethyl disulfide (DMDS) adducts has been one of the most useful reactions for exactly locating the double bonds in monounsaturated and diunsaturated fatty acid methyl esters and hydrocarbons by gas chromatography and mass spectrometry (9–11). The compounds of interest are reacted with

DMDS in I_2 , methylsulphenyl iodide is generated *in situ* and adds to the double bonds via the episulfonium intermediate (12). The generated thiiranium ion is then trapped by dimethyl disulfide to generate the corresponding DMDS adduct (see Figure 2.3). In order to determine the position of the double bonds in the hydrocarbons synthesized by PCC 7002, the hydrocarbon mixture was derivatized with dimethyl disulfide (DMDS) and analyzed by GC-MS. The spectrum for the DMDS adduct of the C19:1 hydrocarbon showed molecular ions corresponding to a terminal double bond (Figure 2.3a). The compound was confirmed as 1-nonadecene by comparison with a commercial standard.

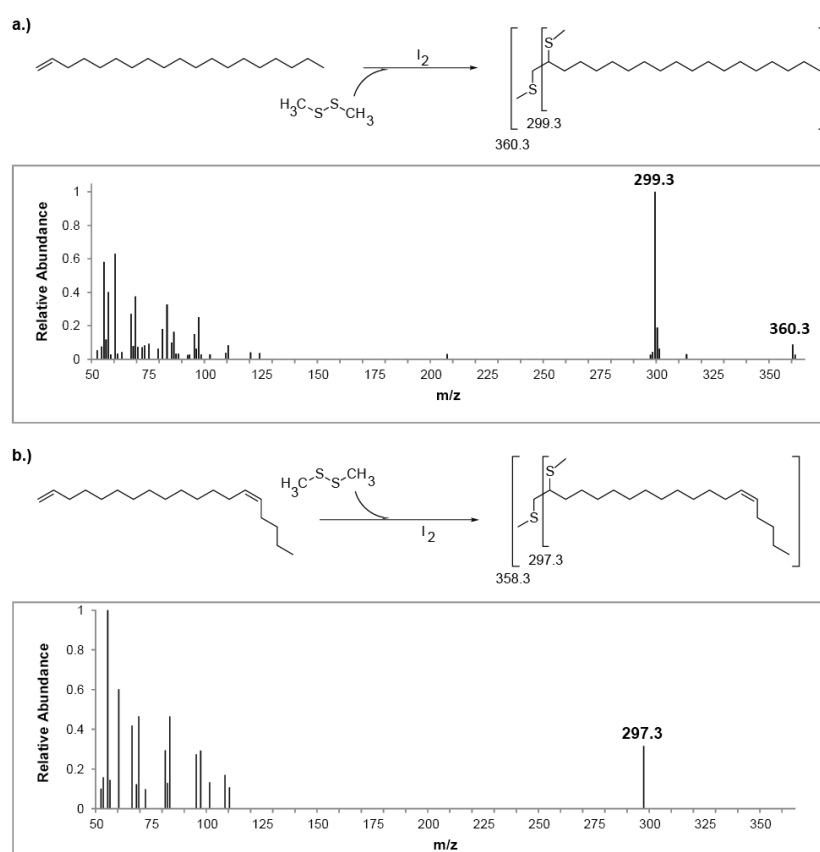


Figure 2.3. Observed mass spectra of the DMDS derivatives of a) 1-nonadecene and b.) 1,14-nonadecadiene.

The spectrum for the DMDS adduct of the C19:2 hydrocarbon also showed ions consistent with a terminal double bond (Figure 2.3b). However, it was not possible to identify the position of the internal double bond. Another approach that can be used to identify the location of internal double bonds is by chemical oxidation and cleavage of the double bond, followed by analysis of the products by GC-MS. Oxidation with a mixture of permanganate and periodate (von Rudloff oxidation) yields monocarboxylic and dicarboxylic acid products (Figure 2.4a) that can easily be identified (13). To identify the position of the internal double bond, the C19:2 hydrocarbon was separated and purified over silica gel and subjected to permanganate/periodate oxidation. The mass spectra of the major resulting peak were consistent with a 13-carbon dicarboxylic, dimethyl ester (Figure 2.4b), suggesting that the C19:2 species is 1, 14-nonadecadiene.

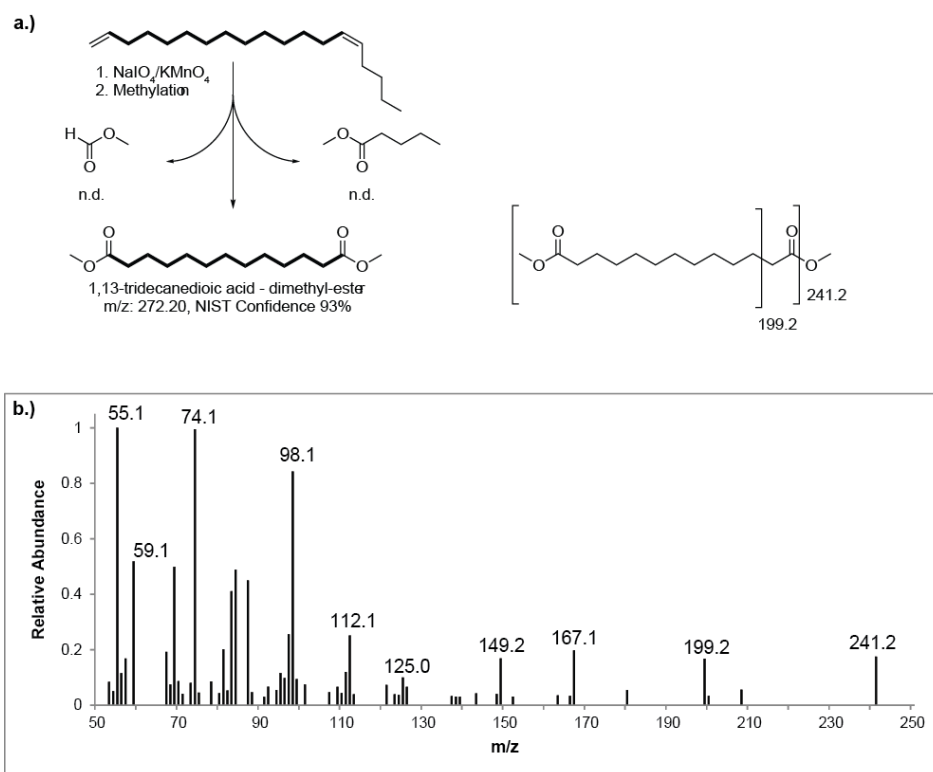


Figure 2.4. a.) Chemical processing of purified 1,14-nonadecadiene. b.) Observed mass spectra of the derivative of 1,14-nonadecadiene. The mass spectra is consistent with the NIST entry for 1,13-tridecanedioic acid – dimethyl ester.

To gain more information about the internal double bond in 1, 14-nonadecadiene, proton nuclear magnetic resonance spectra was taken on a 300 MHz instrument (Bruker AC-300 spectrometer). The sample was dissolved in deuteriochloroform along with tetramethylsilane (TMS) as the internal standard. From Figure 2.5a, 14-nonadecadiene showed signals for the terminal double bond at δ 2.03 ppm, δ 4.89 ppm and δ 5.92 ppm. Also, there was another vinyl proton signal at δ 5.35 ppm integrated as two protons and the allylic proton signal at δ 2.03 ppm integrated as a total of six protons. These additional signals correspond to an internal double bond. Irradiation with the allylic proton frequency caused collapse of the signal at δ 5.35 ppm to a broad peak (Figure 2.5b), indicating a doublet with a coupling constant $J < 9$ Hz ($J = \delta \nu_0$, where δ is the separation of lines in ppm and ν_0 is the resonance frequency of spectrometer in MHz). Since coupling constant between vinylic protons in the *trans* configuration are in the range of 12-18 Hz, and those between *cis* vinylic protons are in the range of 6-12 Hz (14), the internal double bond in 1, 14-nonadecadiene is *cis*.

the possible substrates for hydrocarbon biosynthesis. After extraction and methylation following previously described protocols (15) lipid analysis was performed by gas chromatography-mass spectrometry. As can be seen in Table 2.1, PCC 7002 synthesizes fatty acids with 14, 16 and 18 carbons with different degrees of unsaturation. The major components were C16 and C18:1 fatty acids. It is important to note that no fatty acids larger than C18 were detected.

Table 2.1 Fatty acid profile for wild type PCC 7002 at 38°C

Fatty acid	% mole	
	Experimental	Reported (16)
14:0	0.53 ± 0.40	n. d.
16:0	58.64 ± 2.21	55 ± 2
16:1 (Δ9)	7.05 ± 2.19	6 ± 1
18:0	1.89 ± 0.92	1
18:1 (Δ9)	17.12 ± 1.31	17 ± 3
18:2 (Δ9, Δ12)	11.34 ± 0.60	18 ± 2
18:3 (Δ9, Δ12, Δ15)	2.88 ± 0.66	1

2.2.3 Identification of a gene involved in alkene biosynthesis

Analysis of the fatty acid profile from PCC 7002 indicated that the largest species were C18 fatty acids, therefore it is unlikely that a decarboxylase like the one described for *Jeotgalicoccus sp.* (8) is involved in the biosynthesis of the C19 alkenes observed in PCC 7002 because it would require fatty acids with 20 carbons. As mentioned before, α -olefins can also be formed from fatty acids by an elongation-decarboxylation mechanism (17), resulting in α -olefins one carbon larger than the fatty acid substrate. This type of mechanism seems more likely as C18 fatty acids would

result in the formation of the C19 alkenes. However, the enzymes involved in this conversion were unknown before the research presented in this chapter was done.

Even though no genes involved in formation of alkenes via an elongation-decarboxylation mechanism have been identified, there are genes that have been linked to the formation of terminal double bonds in other molecules. For example, in the mevalonate pathway for isoprenoid biosynthesis (18) mevalonate-5-diphosphate decarboxylase (MDD) catalyzes a decarboxylative elimination reaction by converting a β -hydroxyl group into a phosphate leaving group. Also, in *Streptomyces sp.* CK4412 introduction of a terminal olefin as the final step in tautomycin biosynthesis has been observed (19). The best characterized example of terminal olefin biosynthesis corresponds to curacin A, a metabolite isolated from the marine cyanobacterium *Lyngbya majuscula*. Curacin A is potent cancer cell toxin that contains an unusual hydrophobic terminal olefin instead of a typical terminal carboxyl, aldehyde or alcohol group. It is made by a 64 kb gene cluster comprising nine PKS and one NRPS module, the last of which (*curM*) is responsible for the terminal double bond formation (20).

Analysis of the pathway revealed that the termination module consists of a putative sulfotransferase (ST) domain that is flanked by an acyl carrier protein (ACP) domain and thioesterase (TE) domain (21) (Figure 2.6). *curM*'s ST was predicted to bind to PAPS (adenosine 3'-phosphate 5'-phosphosulfate) and transfer a sulfonate moiety to the β -hydroxyl group of the penultimate chain elongation intermediate tethered to *curM* ACP, forming an excellent leaving group that is positioned chemically to facilitate decarboxylative elimination in the presence of the terminal carboxylate following TE-mediated hydrolysis of the acyl-thioester (Figure 2.6b).

On the basis of these observations, we hypothesized that enzymes with homology to CurM could be involved in biosynthesis of the α -olefins observed in PCC 7002. We used the basic local alignment search tool (BLAST) from NCBI to look for homologs to CurM in PCC 7002. This search identified an open reading frame encoding a protein with 45% amino acid sequence identity to CurM (SYNPCC7002_A1173, here referred to as the *ols* protein for olefin synthase).

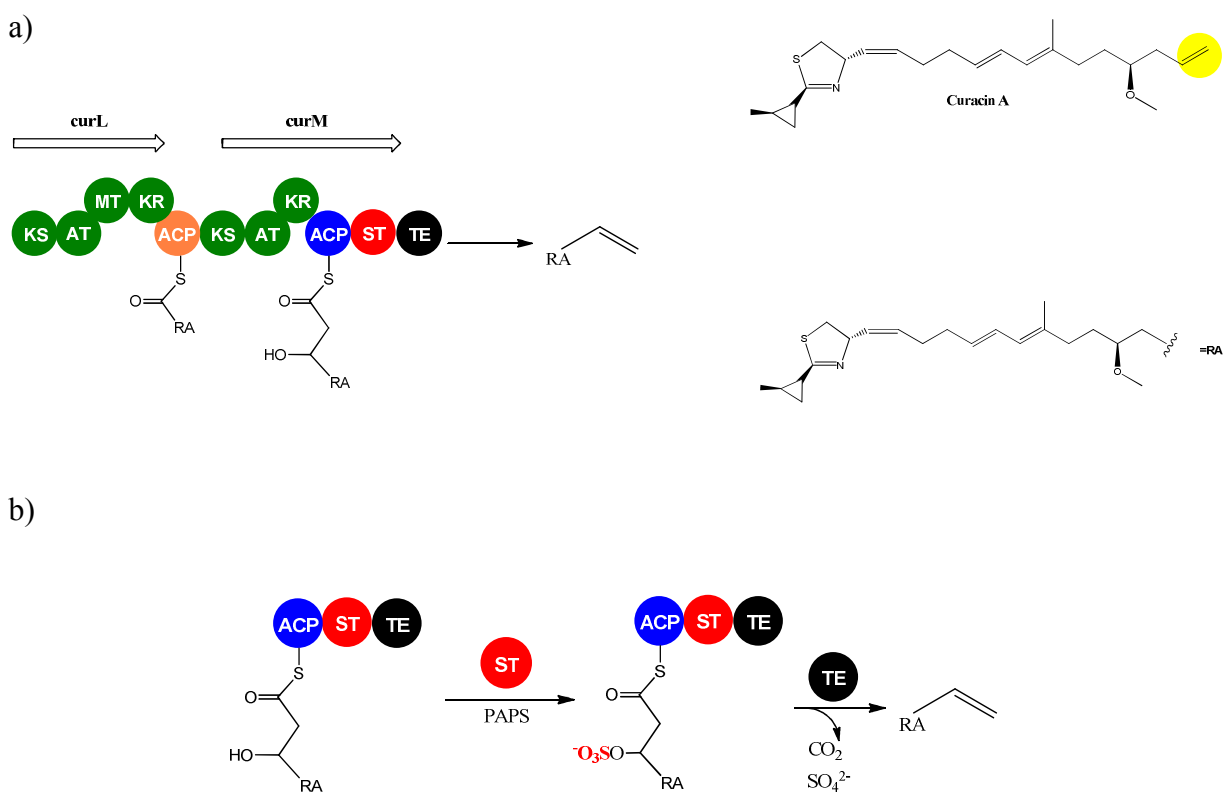
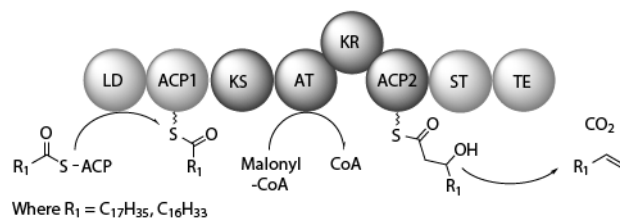


Figure 2.6 a) Final modules involved in Curacin A biosynthesis, b) Terminal double bond formation (Figures adapted from (20, 21))

Several motif sequences commonly found in polyketide synthases were identified within *ols* (Figure 2.7). Overall, *ols* encodes a protein with C-terminal domain architecture that is highly similar to CurM, including the polyketide elongation module and terminal olefin forming domains. Unlike for CurM, the N terminus of Ols contains two additional domains, which are predicted to comprise a loading module (Figure 2.7).

The LD domain contains the highly conserved sequence element that comprises the ATP/AMP signature motif (22) (Figure 2.7b). This domain might be responsible for the loading of the fatty acid to the ACP through the hydrolysis of ATP to yield pyrophosphate. The activation involves the linking of the carboxylic group of the fatty acid through an acyl bond to the phosphoryl group of AMP. Subsequently the activated fatty acid is acylated onto the ACP releasing AMP. The ACP domain contains the 4'-phosphopantetheinyl (Ppant)-binding cofactor box as defined by analysis of 198 ACP sequences (23). Where the serine residue is highly conserved and is required for activation of the apo-ACP to the holo-ACP. The central KS-AT-KR-ACP domains are similar to many polyketide synthase extension modules where two carbons from malonyl-CoA are added to the growing chain and the β -keto group is reduced to a hydroxyl. The putative AT domain contains the amino acid motif specific for malonyl-CoA recognition (24, 25).

a)



b)

LD		ACP1		KS		Domain
ATP/AMP motif (22)		Ppant binding (26)		KS motif (26, 27)		Function
YTSGETTGXPKGV-----GYGXTE		GXDSL		TVDTGCSSSLV		Consensus
YTSGETTGDPKGV-----CYGMAE		GLDSV		SIDTACSSSLV		Ols sequence
168	320	634		841		

AT		KR		ACP2		Domain
Malonyl-CoA recognition (28)		NADP(H) binding site (29)		Ppant binding (26)		Function
ETGYAQXXXXXXXXQXAXFGLL-----GHSXG		GXGXXGXXXA		GXDSL		Consensus
ETVYTQPLLFAYEYAIQLW-----GHSVG		GLGAIQRKIA		GMDSL		Ols sequence
1256	1287	1783		2061		

Figure 2.7 (a) Domain organization and proposed mechanism of a putative olefin synthase encoded by *ols*. (b) Partial sequence alignments of Ols with consensus polyketide synthase domain motifs. Number correspond to the amino acid positions in Ols. References are in parenthesis. Ppant=phosphopantetheine.

Based on the domain architecture, we hypothesize that substrates are loaded onto the ACP1 domain by the ATP consuming loading domain (LD). Once loaded, the central extension module (ketosynthase [KS], acyltransferase [AT], ketoreductase [KR], ACP) would add two carbons from malonyl-coenzyme A (CoA) (a malonyl-CoA recognition motif was found within the AT domain) (Fig. 2a) to the acyl-substrate and reduce the β -keto group to a β -hydroxyl. The presence of a sulfotransferase (ST) domain adjacent to ACP2 suggests that it activates the β -hydroxyl group via sulfation. Activation is required to drive subsequent dehydration and decarboxylation reactions that could be catalyzed by the C-terminal thioesterase (TE) domain analogous to the final reactions performed by CurM.

2.2.4 Genetic studies

To confirm the involvement of *Ols* in α -olefin biosynthesis, a fully segregated null mutant of the *ols* gene was made by homologous recombination of a linear DNA fragment containing a streptomycin resistance cassette flanked by 1,000 bases homologous to the regions flanking (30).. The mutant strain (Δols) was grown under conditions identical to those used for the wild-type strain and subjected to hydrocarbon analysis. Hydrocarbon extracts of Δols did not contain any detectable olefins (Figure 2.1). Moreover, a deletion of only 1,000 bp at the 5' end of the gene corresponding to the putative loading domain gave identical results (not shown). No significant differences were observed when the fatty acid profiles of the wild type and the mutant strain were compared (Table 2.2).

Table 2.2 Fatty acid profile for wild type PCC 7002 and Δols at 38°C

Fatty acid	% mole	
	wt	Δols
14:0	0.53 \pm 0.40	0.20 \pm 0.35
16:0	58.64 \pm 2.21	59.45 \pm 2.61
16:1 ($\Delta 9$)	7.05 \pm 2.19	8.35 \pm 0.53
18:0	1.89 \pm 0.92	1.18 \pm 0.43
18:1 ($\Delta 9$)	17.12 \pm 1.31	13.89 \pm 0.87
18:2 ($\Delta 9, \Delta 12$)	11.34 \pm 0.60	8.99 \pm 0.96
18:3 ($\Delta 9, \Delta 12, \Delta 15$)	2.88 \pm 0.66	7.41 \pm 0.57

To demonstrate a positive correlation between the *ols* gene and olefin production, the 250 bases immediately upstream of the *ols* coding sequence were replaced with the sequence that controls transcription of *psbA* in *Amaranthus hybridus* (31). A fully segregated mutant harboring the promoter replacement, strain $\Phi(P_{psbA}-ols)$, was obtained and verified by PCR. Hydrocarbon

extracts from cultures of the mutant strains contained significantly increased titers of each olefin (Fig. 2.1). A 2-fold increase in 1-nonadecene production and a 5-fold increase in 1,14-nonadecadiene were observed in cultures grown at 35°C in medium A (Table 2.3). mRNA was extracted from each culture using the Trizol 95 method (32). Quantitative PCR of *ols* mRNA, using primers that amplified a short 104-bp segment at the 3' end of *ols*, confirmed a 2.2-fold increase in mRNA in the promoter replacement mutant relative to the level for wild-type PCC 7002 (Table 2.3).

Table 2.3 Olefin production and olefin synthase expression

Description or genotype	C19:2 concentration (µg/ml/OD ₇₃₀)	C19:2 concentration (µg/ml/OD ₇₃₀)	RNA level relative to WT level
Wild type (WT)	0.15 ± 0.06	1.60 ± 0.24	1.00 ± 0.10
Δols	ND	ND	NA
Δols -LD	ND	ND	0.07 ± 0.01
$\Phi(P_{psbA-ols})$	0.75 ± 0.13	3.45 ± 0.71	2.20 ± 0.30

ND, not detected; NA, not applicable; Data represent averages of results from three biological replicates ± standard deviations. Concentrations of 1, 14-nonadecadiene were estimated from a dilution series of 1-nonadecene analytical standards. *ols* RNA levels determined by quantitative PCR (qPCR) were normalized to the amount of *petB* mRNA in each sample and compared to the wild type PCC 7002 ratio. Δols -LD is the loading domain disruption mutant.

2.2.5 Elongation of fatty acids is required for the formation of the alkenes

Further analysis of hydrocarbon extracts of the wild-type and mutant strains revealed trace amounts of two additional hydrocarbons (at 12.6 and 12.8 min) that also disappeared in extracts of the Δols strain and increased in extracts of the $\Phi(P_{psbA-ols})$ mutant. We identified the latter of the two compounds as 1-octadecene by its mass spectrum and by comparison with a

commercial standard. The mass spectrum of the first compound was consistent with octadecadiene, but the compound was present in insufficient quantities to confirm its structure. On the basis of the proposed mechanism for α -olefin formation, we hypothesized that heptadecanoic acid (C17:0) might be the substrate for 1-octadecene formation. In order to test this hypothesis, we fed C17:0 to each of the three strains (final concentration, 0.1 mM) and cultures were grown to an OD₇₃₀ of 0.3. We observed increases in the peak areas for 1-octadecene in both the wild-type and the Φ (PpsbA-*ols*) strains (Figure 2.8). These results, combined with the fact that fatty acids no larger than C18 have been observed in PCC 7002, suggest that elongation of fatty acids is required for α -olefin formation.

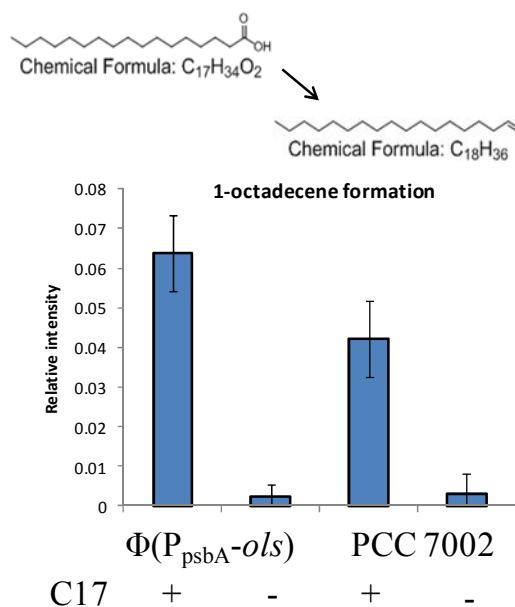


Figure 2.8 Formation of 1-octadecene after feeding cyanobacterial cultures with heptadecanoic acid (C17). Y-axis represents the area of the peaks normalized to the area of the internal standard. Error bars represent the standard deviation from three biological samples.

2.2.6 Acyl-ACPs as precursors for olefin biosynthesis

As mentioned before fatty acids are the precursors for olefin biosynthesis and in principle, free fatty acids or thioester derivatives (acyl-ACPs or acyl-CoAs) can be the substrates for Ols. In order to determine if free fatty acids were the substrates for alkene biosynthesis, we fed pentadecanoic acid (final concentration, 0.1 mM) to the wild type and mutant strains and looked for the formation of the corresponding 1-hexadecene hydrocarbon. Interestingly, feeding of pentadecanoic acid (C15:0) did not result in the formation of 1-hexadecene but in an increase of 1-octadecene as well (Figure 2.9).

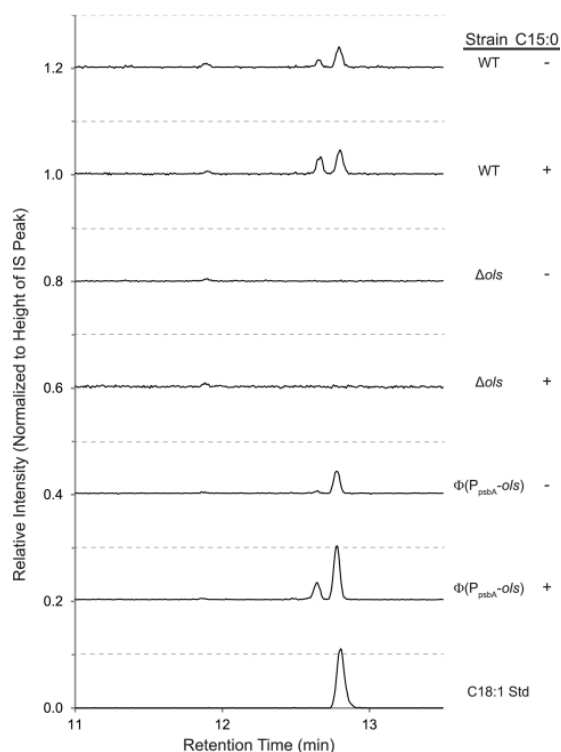


Figure 2.9 Comparison of hydrocarbon extracts from wild type (WT), *ols* deletion mutant (Δols), and $\Phi(PpsbA-ols)$ strains of PCC 7002 supplemented with pentadecanoic (C15:0) acid. The GC-MS signal was normalized to the height of an internal standard peak (hexadecane, 8.3 min). Supplementation of cultures with heptadecanoic (C17:0) resulted in similar traces (not shown). Supplementation of odd chain fatty acids resulted in increased production of 1-octadecene (C18:1) and a compound consistent with a doubly unsaturated 18-carbon hydrocarbon (C18:2). No peaks corresponding to a 1-hexadecene analytical standard were observed in any of the extracts.

The fact that feeding of pentadecanoic acid (C15) resulted in the formation of 1-octadecene (and not 1-hexadecene) suggests that the C15 fatty acid goes through one round of fatty acid elongation (to form a C17 fatty acid) before it is utilized by Ols to synthesize the alkene. Since endogenous fatty acids are synthesized as derivatives of ACPs it seems unlikely that free fatty acids are the substrates for Ols; these intermediates can undergo further elongation or serve as acyl donors depending on the chain lengths of the acyl-ACPs. This is also supported by the fact that homologous proteins to acyl-ACP thioesterases, enzymes that hydrolyze the acyl-ACP thioester to liberate the free fatty acid, haven't been found in cyanobacteria (33).

In cyanobacteria, exogenous free fatty acids that are transported across the outer membrane need to be activated by an acyl-ACP synthetase (Aas) to form acyl-ACPs before they can be incorporated into lipids (33); these intermediates do not dissociate from the enzyme *in vivo* nor are they exchanged with the acyl-ACP pool (34). The acyl-ACP synthetase (Aas) involved in this activation has been characterized in *Synechococcus elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803. They showed that the activation of free fatty acids by Aas is essential for the incorporation of exogenously supplied free fatty acids into cellular lipid metabolism, but it is believed that the most important role of the activation performed by Aas is the recycling of endogenous free fatty acids released from membrane lipids.

Although no Aas has been characterized in *Synechococcus* sp. PCC 7002, there are three enzymes (YP_001733936, YP_001735220 and YP_001735222) with high degree of homology to the Aas's from PCC 7942 and PCC 6803. In order to test if any of these enzymes were involved in the activation of free fatty acids, knockout mutant strains of their encoding genes were constructed and segregation was verified by PCR. Pentadecanoic acid (C15) was then fed to the cultures from the mutant strains. When C15 is fed to the wild type strain, we detected the

formation of the C17 fatty acid and 1-octadecene. The same was observed for the Δaas_{5220} and Δaas_{5222} strains (Figure 2.10). However, for the Δaas_{3932} strain we didn't detect the formation of the C17 fatty acid or 1-octadecene, suggesting that the enzyme encoded by this gene is responsible for the activation of the exogenous fatty acids to acyl-ACPs and its activity is required for the formation of the alkene from the exogenous fatty acids.

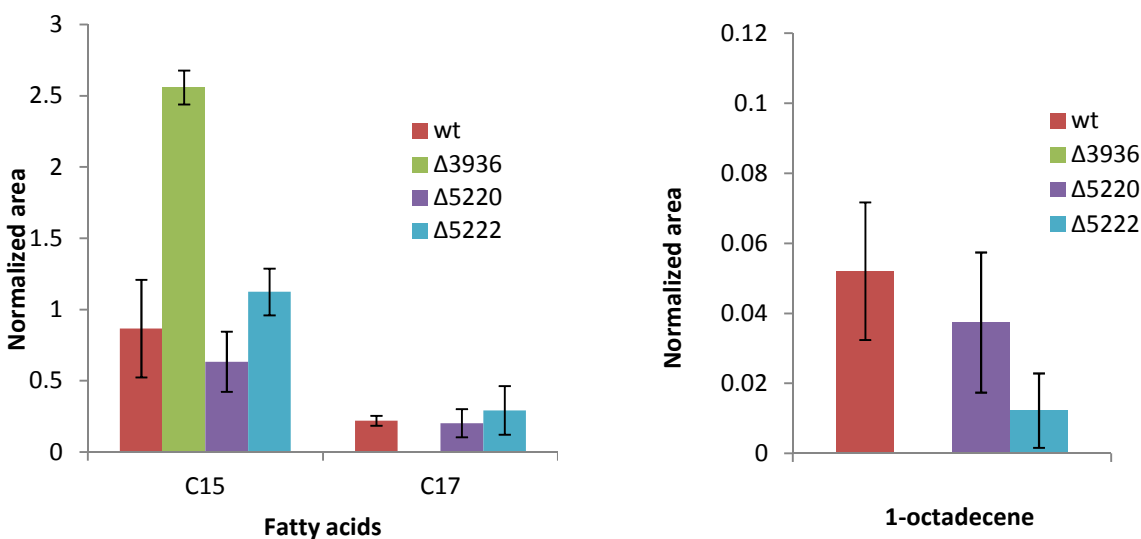


Figure 2.10 Fatty acid and 1-octadecene content in the Δaas mutant strains and wild type strain after the addition of pentadecanoic acid (C15). If the Aas is involved in activation of the free fatty acid, no formation of heptadecanoic (C17) acid should be observed. Areas are normalized to the volume, OD and area of an internal standard. Error bars represent the standard deviation from three biological replicates.

2.3 Conclusions

In summary, the cyanobacterium *Synechococcus sp.* PCC 7002 synthesizes two C19 alkenes containing a terminal double bond: 1-nonadecene and 1, 14-nonadecadiene. A gene (referred as *ols* gene for olefin synthase) with modular organization, similar to a polyketide synthase, was identified and genetic studies were used to prove its involvement in alkene biosynthesis. The

fatty acid profile of PCC 7002, domain architecture of the *ols* gene and feeding studies suggest that the putative enzyme uses an elongation decarboxylation-mechanism to convert C18 fatty acid acyl-acyl carrier proteins (fatty acyl-ACPs) to the alkenes.

Whereas the biosynthesis of the terminal double bond in both hydrocarbons can be explained by the decarboxylation and dehydration reactions performed by Ols domains, the biosynthesis of the internal double bond in 1, 14-nonadecadiene remained unexplained. Chapter 3 will discuss the formation of this internal double bond and the possible functions for the hydrocarbons in PCC 7002.

2.4 Methods

2.4.1 Gene inactivation/promoter replacement

Enzymes and reagents were purchased from New England Biolabs or Fisher Scientific unless otherwise noted. Oligonucleotides used in this study were purchased from Integrated DNA Technologies, Inc. and are listed in Table 2.4. A schematic of the gene inactivation procedure is provided in Figure 2.11. The upstream and downstream flanking sequences of *ols* (SYNPCC7002_A1173) were amplified by PCR (phusion polymerase) from genomic DNA isolated using commercial reagents (Promega). PCR products were digested with *Bam*HI, *Eco*RI or *Pst*I and gel purified using commercial kits (Qiagen). The *aadA* gene (strep^R), was excised from plasmid pSRA81 (35) with *Pst*I and *Bam*HI (for Δ *ols*) or *Pst*I and *Eco*RI (for promoter replacement). The fragments were mixed in a 3:1:3 ratio (left flank: strep^R: right flank) and ligated with T4 DNA ligase. Ligation products were gel extracted and re-amplified to generate sufficient DNA for transformation. DNA was introduced to PCC 7002 using a modification of

the method in reference (30). Briefly, 1 μg of DNA was added to 800 μL of culture ($\text{OD}_{730}=2-3$) and grown at 140 $\mu\text{E}/\text{m}^2/\text{s}$ overnight. The culture was plated on Medium A until a lawn of cells was observed (~ 3 days). The lawn was resuspended in 1-2 ml of Medium A and plated on Medium A containing streptomycin. Single colonies observed after ~ 3 days were screened by PCR using Gotaq (Promega). The upstream region of *ols* was replaced with the ~ 100 bp sequence of the *psbA* promoter from *Amaranthus hybridus* (31) using the recombination method described above. A linear DNA fragment containing ~ 1000 bp from the 5' end of *ols* and the new promoter was constructed by consecutive PCRs. The fragments were digested with EcoRI and PstI, ligated to the *aadA* gene, purified, and transformed as above. Mutants were verified by PCR (Figure 2.12). For the *aas* knockouts, the same procedure was used, except that a kanamycin resistant cassette (*aphII* gene) was used and Gibson (36) cloning was used to construct a plasmid using a pBAD18 backbone and this plasmid was linearized before transformation.

Table 2.4: Oligonucleotides used

Oligonucleotide Name	Sequence
1. OFE7002-Sp-a2	CCGTTCTGCAGCCTGTGAATGGAAATCTGGACTCCGTATCC
2. OFE7002-Sp-a1	CCAACCGGAGGTGAAGCGGACTACA
3. OFE7002-Sp-BamHI-b2	CCGTTGGATCCGCAAAGTGCAGTCCGAAAACCCGTAAATATTAGATCC
4. OFE7002-Sp-BamHI-b1	GCCAAGTCAAAGGGTTTCTGGGCATGG
5. OFE7002-Sp-b2	CCGTTGAATTCCAAGGGACAGAAACAACGGTGACCTTGG
6. OFE7002-Sp-b1	GGGAAAACGACAACCTGAGACCCACCAC
7. Prom-sw-a2	CCAGAATTCCGGAGCTTCATCTGGGGACAATGG
8. Prom-sw-a1	GCTTTCAGCCCACCTGTCCCAATATGC
9. Prom-sw-b1	CCAAGGTCACCGTTGTTTCTGTCCCTTG
10. Prom-Amar-b2-1	GAGACAGGATGAGGATCGTTTCGCATGGTTGGTCAATTTGCAAATTCGTCGATCTGC
11. Prom-Amar-b2-2	CTGTTGAATAACAAGGACGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATG
12. Prom-Amar-b2-3	GTTGACACGGCGGTATAAGACATGTTATACTGTTGAATAACAAGGACGGATCTGATCAAG
13. Prom-Amar-b2-4	CCACTGCAGGATCTCAATGAATATTGGTTGACACGGCGGTATAAGACATGTTATACTG
14. Gaz7002-Seq2-Rv	CGTTGATCGCCTTTAGCCACC
15. aadA-Rv2	GCAAGATAGCCAGATCAATGTTCGATCGTG
16. Gaz7002-Seq11	CCCAAAGACCTCTCGGCGTTC
17. aadA-Fw2	GACATTCTTGCAGGTATCTTCGAGCCAGC
18. SYNPC7002_A1173-RVT	TTATTGTGTTTTGGGTACAGG
19. petB-RVT	TTACAAAGGACCAGAAATACC

20.	SYNPCC7002_A1173-RTF	TGGCATTAGCAGACGACGTTACCT
21.	SYNPCC7002_A1173-RTR	TGGAGATCAGCAGGGCGGTTAAAT
22.	petB-RT-Fw	GATTCGCAATGACCTTCTAC
23.	petB-RT-Rv	CCAGTAATCCAAGTCAGCTC
24.	KM-Fw	TTAGAAAACTCATCGAGCATCAAATGAACTGC
25.	B3936-fw	GCGATCCGAATGGCGGAATCTTC
26.	B5220-fw	CCTCTAAAGGGTGGCGATGAAAATTGC
27.	A3936-KM-Rv	GCAGTTTCATTTGATGCTCGATGAGTTTTTCTAACAGCCGAAATCATGGCTACAATCCTAC
28.	KM-B3936-Rv	GAAGATTCGCCATTCCGATCGCCTTTGGCAGGATCCGGCTGCTAACAAA
29.	A5220-KM-Rv	GCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAGTTTTTCGCAGAATGGGTCATGGTGG
30.	KM-B5220-Rv	GCAATTTTCATCGCCACCCTTTAGAGGCTTTGGCAGGATCCGGCTGCTAACAAA
31.	pBAD18-Fw	GATAAACAGAAATTTGCCTGGCGGCAG
32.	pBAD18-A3936Rv	GAAATTACTGCCGACGTGATGGAAGGCATATGTGCATAGGAGAAACAGTAGAGAGTTGCGATAAAAAAGCG
33.	pBAD18-A5220Rv	CCTCTAGTAGGGCCTTACAGAACTGCCATATGTGCATAGGAGAAACAGTAGAGAGTTGCGATAAAAAAGCG
34.	A5222-FWD	TATGCACATATGACAACTGGGCAAATGGGTAGCAC
35.	A5222-RVS	CGTTACCACCGCTGCGTTCGGCTAAGGGGACGCAACAGTTGTG
36.	aadA-FWD	CGAACGCAGCGGTGGTAACG
37.	aada-RVS	GTTTTGGTTTTCGTTTTCCACTCTAGAGTCATGGAATTAATCTCCTGTCGAGCGAATTGTTAGACATTATTTG CCGA
38.	BTE-FWD	AGGAGATTAATTCCATGACTCTAGAGTGGAAACCGAAACCAAAC
39.	BTE-RVS	CCGTAATAATTGCGTGTGCGATCTATTCAGCTTAATGATGATGATGATGATGATGATGATGAACACGAGGTT CCGCCGGAATTAC
40.	B5222-FWD	GCTGAATAGATGCGACACGCAATTATTACGG
41.	B5222-RVS	CTGCCGCCAGGCAAATTTCTGTTTTATCCATATGCCTTTACCTGGTGTCTAGTTGGGCATAG
42.	pBAD18-Rv	GTGCTACCCATTTGCCAGTTGTCATATGTGCATAGGAGAAACAGTAGAGAGTTGCGATAAAAAAGCG

Notes: Restriction sites used in cloning are underlined. Oligos 1-4 were used for making *ols* knockout cassette; 1, 2, 5, 6 for making the loading domain knockout cassette; 10-13, for making the *Amaranthus* promoter ($P_{P_{sbA}}$) replacement cassette; 7-9, 13 for promoter replacement; 14-17 for screening of mutant strains; 18-23 for qRT-PCR; 24-42 for *aas* knockouts.

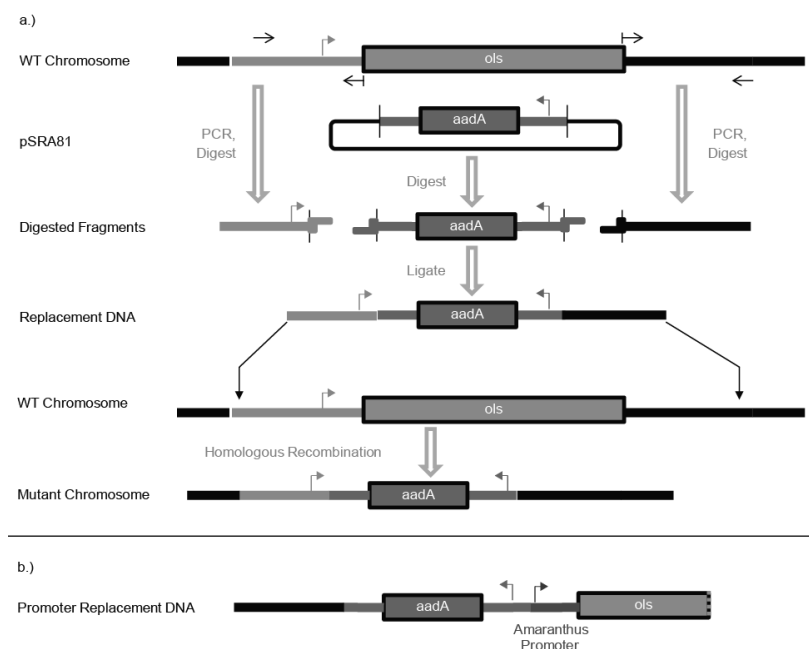


Figure 2.11 Schematic outlining steps to construct PCC 7002 mutant strains. a.) To generate a deletion mutant, three fragments (left homology region, resistance cassette, right homology region) are generated by PCR and restriction digestion. The three DNA fragments are ligated to generate a linear DNA that is introduced to PCC 7002 cells. Once in the cell, the replacement DNA fragment containing the resistance marker is recombined with chromosomal DNA at crossovers located within the left and right homology regions. b.) To replace the native promoter a similar strategy was used. The promoter replacement DNA was assembled from four fragments (left homology region, resistance cassette, heterologous promoter cassette, and right homology region). The promoter cassette was generated by PCR assembly of oligonucleotide primers. The right homology region contains the first ~1000 bp of *ols*. The promoter replacement DNA was introduced to PCC 7002 to initiate homologous recombination.

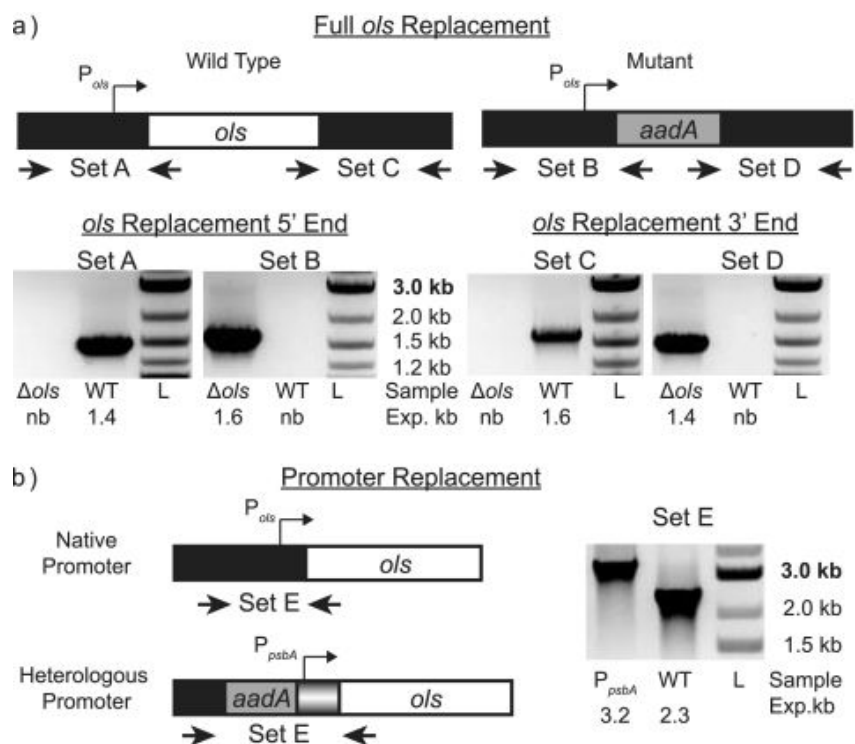


Figure 2.12 Confirmation of PCC 7002 mutants by PCR. a.) Primers amplifying the 5' and 3' junctions of *ols* (sets A&C) and the expected integrated resistance cassette (sets B&D) were used to generate PCR products specific to each strain. Gels of each PCR confirm the replacement of *ols* with a streptomycin resistance cassette (*aadA*). b.) Primers flanking the promoter of *ols* were used to confirm the size of the genomic sequence on both the wild type and $\Phi(P_{psbA}-ols)$ mutant. The larger size of mutant PCR product is due to the presence of an *aadA* expression cassette positioned upstream of P_{psbA} . nb – no band, exp – expected.

2.4.2 Hydrocarbon extraction and GC-MS analysis

Cultures of wild-type and mutant PCC 7002 were grown to an $OD_{730} \sim 0.2$. Cells were harvested by centrifugation, resuspended in 3 mL of H_2O containing 50 μg of hexadecane as an internal standard. Lipids and hydrocarbons were purified and analyzed by GC/MS as in (15). Peak identification was achieved by comparison with internal standards and to the NIST Mass Spectral Database. Quantification was achieved by comparison of integrated peaks with calibration curves of a 1-nonadecene standard (Fluka).

2.4.3 DMDS derivatization

All chemicals were purchased from Sigma-Aldrich unless otherwise noted. To the dried hydrocarbon extract, 100 μ L of dimethyl disulfide (Acros Organics) and 100 μ L of carbon disulfide containing 6 mg/mL of iodine were added (37). The reaction mixture was kept at 60°C and after 40 hr, the reaction was quenched with aqueous $\text{Na}_2\text{S}_2\text{O}_3$ (3×10^{-4} M, Acros Organics) and the organic phase was then evaporated to dryness under a nitrogen stream and dissolved in hexane to be analyzed by GC/MS.

2.4.4 Permanganate/periodate oxidation

Isolation of the hydrocarbon for permanganate/periodate oxidation was performed by flash chromatography using a Pasteur pipette packed with silica gel (Silica flash P60, 40-63 μ m, 60A, Silicycle). The sample was applied after solvating the silica gel column with 8-10 column volumes of hexane. Five fractions were collected and analyzed by thin layer chromatography using commercial glass plates coated with silica gel. Spots were visualized using a permanganate solution. After identifying fractions containing the desired compound (verified by GC/MS), permanganate/periodate oxidation was used to identify the position of the double bond. Permanganate/periodate oxidation was done as described before (14): the sample was dried under nitrogen and 0.3 ml of t-butanol, 0.1 ml of Na_2CO_3 (0.02M) and 0.12 ml of $\text{NaIO}_4/\text{KMnO}_4$ (2.1mg/ml and 1.6 ml/ml respectively) were added to the dried extract. Reaction was maintained at 28°C with shaking for 6 hr. The sample was then decolorized with a few crystals of NaHSO_4 (Flinn Scientific) and acidified with 0.025ml of H_2SO_4 (0.5N, Fisher). 0.5 ml of diethyl ether (Alfa Aesar) was added and vortexed. The top layer was dried under nitrogen

stream and methylation was done essentially as described by (15) and GC/MS was used to identify the compounds (Figure 2.4).

2.5 References

1. **Mendez-Perez D, Begemann MB, Pflieger BF.** 2011. Modular synthase-encoding gene involved in α -olefin biosynthesis in *Synechococcus* sp. strain PCC 7002. *Applied and environmental microbiology* **77**:4264–7.
2. **Herrera S.** 2006. Bonkers about biofuels. *Nature biotechnology* **24**:755–60.
3. **Hoffert MI, Caldeira K, Benford G, Criswell DR, Green C, Herzog H, Jain AK, Kheshgi HS, Lackner KS, Lewis JS, Lightfoot HD, Manheimer W, Mankins JC, Mauel ME, Perkins LJ, Schlesinger ME, Volk T, Wigley TML.** 2002. Advanced technology paths to global climate stability: energy for a greenhouse planet. *Science (New York, N.Y.)* **298**:981–7.
4. **Agarwal AK.** 2007. Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. *Progress in Energy and Combustion Science* **33**:233–271.
5. **Winters K, Parker PL, Baalen C Van.** 1969. Hydrocarbons of Blue-Green Algae: Geochemical Significance. *Science* **163**:467–468.
6. **Schirmer A, Rude MA, Li X, Popova E, Del Cardayre SB.** 2010. Microbial biosynthesis of alkanes. *Science* **329**:559–62.
7. **Beller HR, Goh E-B, Keasling JD.** 2010. Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Applied and environmental microbiology* **76**:1212–23.
8. **Rude M a, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A.** 2011. Terminal olefin (1-alkene) biosynthesis by a novel p450 fatty acid decarboxylase from *Jeotgalicoccus* species. *Applied and environmental microbiology* **77**:1718–27.
9. **Carballeira NM, Shalabi F, Cruz C.** 1994. Thietane, Tetrahydrothiophene and Tetrahydrothiopyran Formation in Reaction of Methylene-interrupted Dienoates with Dimethyl Disulfide. *Tetrahedron Letter* **35**:5575–5578.
10. **Shibahara A, Yamamoto K, Kinoshita A, Anderson BL.** 2008. An Improved Method for Preparing Dimethyl Disulfide Adducts for GC/MS Analysis. *Journal of the American Oil Chemists' Society* **85**:93–94.

11. **Yamamoto K, Shibahara A, Nakayama T, Kajimoto G.** 1991. Determination of double-bond positions in methylene-interrupted dienoic fatty acids by GC-MS as their dimethyl disulfide adducts. *Chemistry and Physics of Lipids* **60**:39–50.
12. **Csizmadia VM, Schmid GH, Mezey PG, Csizmadia IG.** 1977. Ab initio SCF-MO study of the reaction intermediates formed by addition of thiohypochlorous acid to ethylene. *Journal of the Chemical Society, Perkin Transactions 2* 1019.
13. **Longmuir KJ, Rossi ME, Resele-Tiden C.** 1987. Determination of monoenoic fatty acid double bond position by permanganate-periodate oxidation followed by high-performance liquid chromatography of carboxylic acid phenacyl esters. *Analytical Biochemistry* **167**:213–221.
14. **Goodloe RS, Light RJ.** 1982. Structure and composition of hydrocarbons and fatty acids from a marine blue-green alga, *Synechococcus* sp. *Biochimica et biophysica acta* **710**:485–492.
15. **Lennen RM, Braden DJ, West RA, Dumesic JA, Pfleger BF.** 2010. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnology and bioengineering* **106**:193–202.
16. **Sakamoto T, Shen G, Higashi S, Murata N, Bryant D a.** 1998. Alteration of low-temperature susceptibility of the cyanobacterium *Synechococcus* sp. PCC 7002 by genetic manipulation of membrane lipid unsaturation. *Archives of microbiology* **169**:20–8.
17. **Yong T, Largeau C, Casadevall E.** 1986. Biosynthesis Of Non-isoprenoid Hydrocarbons By The Microalga *Botryococcus-braunii* - Evidences For An Elongation-decarboxylation Mechanism - Activation Of Decarboxylation. *Nouveau Journal De Chimie-new Journal Of Chemistry* **10**:701–707.
18. **Bonanno JB, Edo C, Eswar N, Pieper U, Romanowski MJ, Ilyin V, Gerchman SE, Kycia H, Studier FW, Sali A, Burley SK.** 2001. Structural genomics of enzymes involved in sterol/isoprenoid biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **98**:12896–901.
19. **Choi S-S, Hur Y-A, Sherman DH, Kim E-S.** 2007. Isolation of the biosynthetic gene cluster for tautomycetin, a linear polyketide T cell-specific immunomodulator from *Streptomyces* sp. CK4412. *Microbiology (Reading, England)* **153**:1095–102.
20. **Chang Z, Sitachitta N, Rossi J V, Roberts MA, Flatt PM, Jia J, Sherman DH, Gerwick WH.** 2004. Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. *Journal of natural products* **67**:1356–67.
21. **Gu L, Wang B, Kulkarni A, Gehret JJ, Lloyd KR, Gerwick L, Gerwick WH, Wipf P, Håkansson K, Smith JL, Sherman DH.** 2009. Polyketide decarboxylative chain

- termination preceded by o-sulfonation in curacin a biosynthesis. *Journal of the American Chemical Society* **131**:16033–5.
22. **Black PN, Dirusso CC, Metzger AK, Heimert TL.** 1992. Cloning , Sequencing , and Expression of the fadD Gene of Escherichia coli Encoding Acyl Coenzyme A Synthase. *The Journal of Biological Chemistry* **267**:25513–25520.
 23. **Fritzler JM, Zhu G.** 2007. Functional characterization of the acyl-[acyl carrier protein] ligase in the *Cryptosporidium parvum* giant polyketide synthase. *International journal for parasitology* **37**:307–16.
 24. **Silakowski B, Nordsiek G, Kunze B, Blöcker H, Müller R.** 2001. Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the myxalamid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantiaca* Sga1511 This article is dedicated to Prof. Dr. E. Leistner on the occasion of his 60th bi. *Chemistry & Biology* **8**:59–69.
 25. **Ikeda H, Nonomiya T, Usami M, Ohta T, Omura S.** 1999. Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proceedings of the National Academy of Sciences* **96**:9509–9514.
 26. **Donadio S, Katz L.** 1992. Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*. *Gene* **111**:51–60.
 27. **Aparicio JF, Molnár I, Schwecke T, König A, Haydock SF, Ee Khaw L, Staunton J, Leadlay PF.** 1996. Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: Analysis of the enzymatic domains in the modular polyketide synthase. *Gene* **169**:9–16.
 28. **Haydock SF, Aparicio JF, Molnár I, Schwecke T, Khaw LE, König A, Marsden AFA, Galloway IS, Staunton J, Leadlay PF.** 1995. Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Letters* **374**:246–248.
 29. **Scrutton NS, Berry A, Perham RN.** 1990. Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* **343**:38–43.
 30. **Frigaard N-U, Sakuragi Y, Bryant DA.** 2004. Gene inactivation in the cyanobacterium *Synechococcus* sp. PCC 7002 and the green sulfur Bacterium *Chlorobium tepidum* using in vitro-made DNA constructs and natural transformation. *Methods in molecular biology* **274**:325–340.
 31. **Elhai J.** 1993. Strong and regulated promoters in the cyanobacterium *Anabaena* PCC 7120. *FEMS microbiology letters* **114**:179–84.

32. **Pinto FL, Thapper A, Sontheim W, Lindblad P.** 2009. Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC molecular biology* **10**:79.
33. **Kaczmarzyk D, Fulda M.** 2010. Fatty acid activation in cyanobacteria mediated by acyl-acyl carrier protein synthetase enables fatty acid recycling. *Plant physiology* **152**:1598–610.
34. **Gangar A, Karande AA, Rajasekharan R.** 2001. Purification and characterization of acyl-acyl carrier protein synthetase from oleaginous yeast and its role in triacylglycerol biosynthesis. *The Biochemical journal* **360**:471–9.
35. **Frigaard N, Maresca JA, Yunker CE, Jones D, Bryant DA, Jones AD.** 2004. Genetic Manipulation of Carotenoid Biosynthesis in the Green Sulfur Bacterium *Chlorobium tepidum*. *Journal of bacteriology* **186**:5210–5220.
36. **Gibson DG, Young L, Chuang R, Venter JC, Iii CAH, Smith HO, America N.** 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases **6**:12–16.
37. **Vincenti M, Guglielmetti G, Cassani G, Tonini C.** 1987. Determination of Double Bond Position in Diunsaturated Compounds by Mass Spectrometry of Dimethyl Disulfide Derivatives. *Analytical Chemistry* **59**:694–699.

Chapter 3: A desaturase gene involved in the formation of 1,14 nonadecadiene produced by *Synechococcus* sp. PCC 7002

3.1 Introduction

Hydrocarbon biosynthesis is a common trait among both eukaryotes and prokaryotes that has gained significant attention for use in developing sustainable alternatives to petroleum (1). In plants, alkanes are involved in the biosynthesis of epicuticular wax to reduce water loss through the epidermis (2). They are also found on the surface of insects as anti-desiccation agents of the cuticula and are important signaling chemicals in insect communication (3).

All cyanobacteria can synthesize hydrocarbons from fatty acids using two different metabolic pathways: one comprises a two step conversion of fatty acids to fatty aldehydes and then to alkanes that involves a fatty acyl ACP reductase (FAAR) and aldehyde deformylating oxygenase (ADO). The second pathway (presented in Chapter 1) involves a polyketide synthase (OLS pathway) that first elongates the acyl chain followed by decarboxylation to produce a terminal alkene. Analysis of one-hundred-forty-two cyanobacterial genomes showed that there were no instances where both the FAAR/ADO and the OLS pathways were found together in the same genome, suggesting an unknown selective pressure that maintains one or the other pathway, but not both (4). Although it has been known for decades that cyanobacteria accumulate hydrocarbons with medium (C₁₇-C₂₀) chains (5, 6), the physiological or ecological function of hydrocarbon biosynthesis in cyanobacteria remains poorly understood; various possibilities exist including preventing from grazing from herbivores, intra- or inter- species chemical signaling, prevention of desiccation, enhance buoyancy, or membrane fluidity and stability (4).

In the previous chapter we showed that a gene (*ols*) with modular organization, similar to a polyketide synthase, was linked to the biosynthesis of alpha-olefins in the marine cyanobacterium *Synechococcus sp.* PCC 7002 (7). Based on the predicted domain architecture and feeding studies, we postulated that Ols produced alpha-olefins via elongation and decarboxylation of C₁₈ acyl-acyl-carrier proteins (acyl-ACP), intermediates in fatty acid biosynthesis (7). Unlike other cyanobacterial strains that possess a more complex hydrocarbon profile in terms of length and degree of unsaturation (5, 6, 8), PCC 7002 synthesizes only two alkenes: 1-nonadecene (C₁₉:1) and 1, 14-nonadecadiene (C₁₉:2). The only difference between these compounds is the presence of an internal double bond at position 14 in the C₁₉:2 hydrocarbon. Whereas the biosynthesis of the terminal double bond in both hydrocarbons was explained by the decarboxylation and dehydration reactions performed by Ols domains, the biosynthesis of the internal double bond in 1, 14-nonadecadiene remained unexplained. In this chapter, we linked the presence of this internal double bond to a gene predicted to encode a desaturase. Further, we demonstrated an increase in C₁₉:2 abundance as an inverse function of temperature, suggesting that the compound plays a role in responding to cold stress.

3.2 Results

3.2.1 Identification of a desaturase gene involved in alkene unsaturation

Given the lack of functional handles on 1-nonadecadiene, we hypothesized that the internal bond would be present in the substrate of the elongation-decarboxylation mechanism catalyzed by Ols, i.e. an unsaturated C₁₈ acyl-ACP. Unsaturated and polyunsaturated fatty acids are essential constituents of polar glycerolipids and are used to control fluidity of membranes in response to

changes in temperature (9). PCC 7002 synthesizes lipids that incorporate 18-carbon fatty acids with zero, one, two or three double bonds at the $\Delta 9$, $\Delta 12$ and $\Delta 15$ (or $\omega 3$) positions at the *sn*-1 position, and C_{16} fatty acids containing zero or one double bond at the $\Delta 9$ position at the *sn*-2 position. Three acyl-lipid desaturases, encoded by *desA* ($\Delta 12$), *desB* ($\Delta 15$) and *desC* ($\Delta 9$), have been shown to be involved in the biosynthesis of the unsaturated fatty acids observed in PCC 7002 (10–12). Two additional genes are predicted to encode uncharacterized desaturases, *desE* (SYNPCC7002_A2833) and *desF* (SYNPCC7002_A1989). We hypothesized that one of the five desaturases was responsible for the internal double bond in 1, 14-nonadecadiene.

To test this hypothesis, a disruption mutant of each desaturase was constructed by homologous recombination of linear pieces of DNA containing an antibiotic resistance cassette. After multiple attempts, we were unable to obtain a $\Delta desC$ mutant, suggesting that such mutation may be lethal to the cells. The same problem was reported when trying to disrupt *desC* in *Synechocystis sp.* PCC 6803 (13). Conversely, fully segregated knockouts of *desA*, *desB*, *desE* and *desF* were obtained after transformation and plating on the appropriate antibiotics. The observed fatty acid profiles of $\Delta desA$ and $\Delta desB$ mutants were consistent with past reports (10); lacking desaturation at the $\Delta 12$ and $\Delta 15$ positions respectively. The hydrocarbon composition of the $\Delta desA$, $\Delta desB$ and $\Delta desF$ were indistinguishable from the wild type PCC 7002 (data not shown). Conversely, the hydrocarbon extract of the $\Delta desE$ mutant contained no detectable $C_{19:2}$ alkene (Figure 3.1). When the deletion was complimented by inserting *desE* under its native promoter, 1, 14-nonadecadiene production was restored. These results suggest that the desaturase encoded by *desE* is responsible for the internal double bond in $C_{19:2}$.

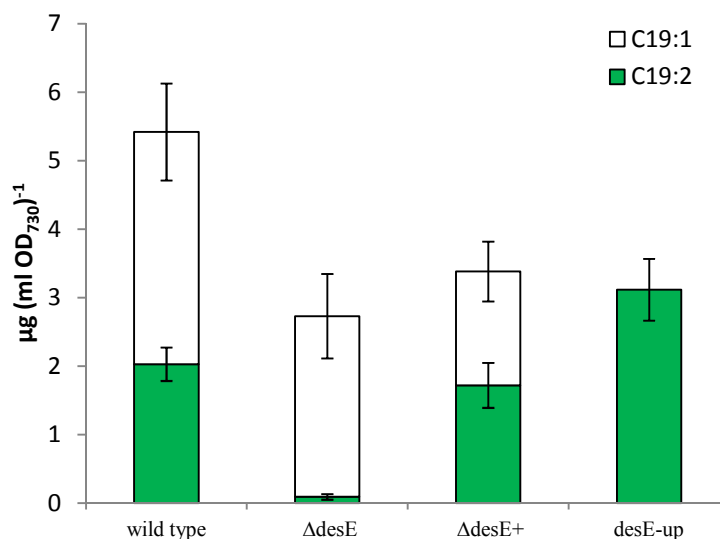


Figure 3.1 Comparison of the hydrocarbon composition from the wild type (wt), *desE* knockout ($\Delta desE$), $\Delta desE$ complemented ($\Delta desE+$) and *desE* upregulated (*desE*-up) strains. Deletion of the *desE* gene eliminated only the production of the hydrocarbon with the internal double bond (C19:2). Cultures were grown at 36°C and error bars represent the standard deviations from three biological replicates.

3.2.2 Precursors for 1, 14-nonadecadiene biosynthesis

The putative Ols mechanism (7) calls for C₁₈ acyl-ACP precursors to be processed via an elongation-decarboxylation mechanism. In the case of the C19:1 hydrocarbon, a fully saturated C₁₈-acyl-ACP (C18:0) would be the precursor. However, for the C19:2 hydrocarbon to have an internal double bond at position 14, a C₁₈ acyl-ACP fatty acid with a double bond at position 13 would be required. This C18:1(Δ 13) acyl-ACP could either be directly synthesized by DesE acting on C₁₈ acyl-ACP or be the elongation product of a shorter unsaturated acyl-ACP. Potential DesE products consistent with this mechanism include C16:1(Δ 11), C14:1(Δ 9), or C12:1(Δ 7) acyl-ACPs. When compared with the $\Delta desE$ extract, the lipid profiles of the wild type strain

contained no additional fatty acids (Table 3.1), providing no assistance in determining the DesE substrate.

Table 3.1 Fatty acid profiles for the wild type and Δ desE at 22°C and 38°C (% mole)

	C14:0	16:0	16:1	18:0	18:1	18:2	18:3
22°C							
Wt	0.97 ± 0.06	37.35 ± 2.12	15.98 ± 0.70	1.10 ± 0.10	4.05 ± 0.14	14.44 ± 1.21	26.03 ± 0.16
Δ desE	n.d.	46.04 ± 1.79	10.41 ± 0.39	0.68 ± 0.35	5.64 ± 0.77	7.45 ± 0.28	26.88 ± 1.63
38°C							
Wt	0.53 ± 0.4	58.64 ± 2.21	7.05 ± 2.19	1.89 ± 0.92	17.12 ± 1.31	11.34 ± 0.60	2.88 ± 0.66
Δ desE	n.d.	61.97 ± 2.32	7.39 ± 0.40	1.12 ± 0.27	12.63 ± 1.14	7.99 ± 0.96	8.53 ± 0.99

The amino acid sequence of DesE shows a high degree of similarity to the Δ 9 desaturases of *Rattus norvegicus* (rat) and *Saccharomyces cerevisiae* (34-38%) (14). In addition, its hydropathy profile has two highly hydrophobic regions (Figure 3.2), suggesting that it is a membrane-integrated protein (14). DesE contains the conserved three histidine-cluster motifs observed in Δ 9 desaturases: HXXXXH (residues 79-84) and the two HXXHH (residues 116-120 and 242-246). These histidine motifs are thought to bind iron atoms and play an important role in the introduction of the double bond in the hydrocarbon chains of fatty acids (15).

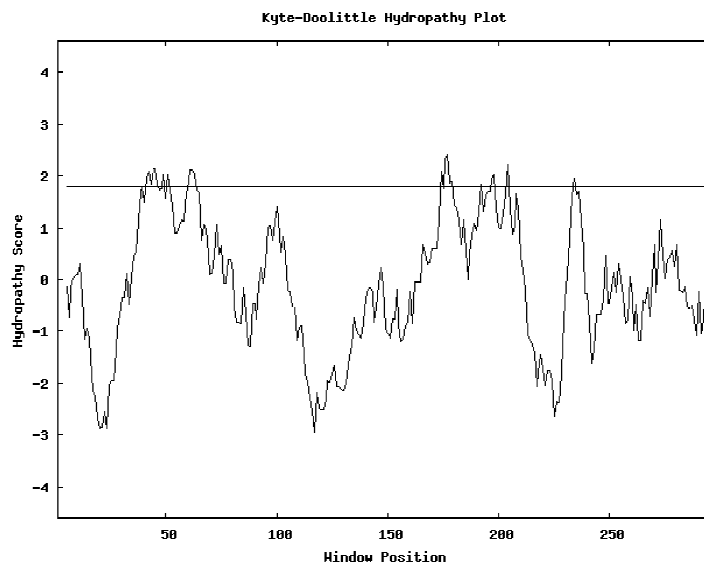


Figure 3.2 Hydropathy plot for DesE. Two clusters of hydrophobic regions, which are putative membrane spanning domains, can be seen from this plot.

The high degree of similarity of DesE to $\Delta 9$ desaturases suggested that a C_{14} acyl-ACP was the most likely substrate to be consistent with the formation of 1, 14-nonadecadiene. If true, the $C_{14}:1(\Delta 9)$ acyl-ACP synthesized by DesE would then be elongated to a $C_{16}:1(\Delta 11)$ and a $C_{18}:1(\Delta 13)$ acyl-ACP that ultimately would be the substrate for Ols (Figure 3.3a). The lack of these intermediates in the lipid extracts of the wild type strain suggested that they might not be accumulating to a detectable extent. Therefore, we decided to increase the expression of *desE* by replacing its promoter with the strong P_{cpcBA} promoter from *Synechocystis sp.* PCC 6803 (16) (strain *desE-up*). The fatty acid profile of this strain contained no additional fatty acids when compared to the wild type. Conversely, the hydrocarbon profile of *desE-up* was completely different from that of the wild type strain. Whereas the $C_{19}:2$ hydrocarbon in the wild type strain represents about 35% of the total hydrocarbons, in *desE-up* it accounts for almost 100% of the total hydrocarbon content (Figure 3.1). These results suggest that the products of DesE are

exclusively used in the formation of olefins and that total olefin content is regulated. If that's the case, knocking out *ols* in the *desE*-up strain should result in the accumulation of the unsaturated acyl-ACPs that ordinarily serve as intermediates in C19:2 biosynthesis. The Δols -*desE*-up strain was constructed and its fatty acid profile was analyzed as before. When compared to the wild type strain the extract of the Δols -*desE*-up mutant contained three additional peaks in its chromatogram at retention times 15.2, 19.6 and 19.9 min. One of these (15.2 min) was identified as 11-hexadecenoic acid, C16:1(Δ 11), by comparison with a commercial standard (Figure 3.3b). The mass spectra for the other two peaks were consistent with C₁₈ unsaturated fatty acids but since no commercial standard was available for the C18:1(Δ 13) fatty acid we could not determine if one was this unsaturated compound. Instead, to corroborate the identity of one of these peaks as the C18:1(Δ 13) fatty acid we fed 9-tetradecenoic acid, C14:1(Δ 9), to wild type PCC 7002. Once activated by an acyl-ACP synthase, its elongation would result in the formation of the C16:1(Δ 11) and C18:1(Δ 13) acyl-ACPs. After analysis, feeding PCC7002 with C14:1(Δ 9) resulted in the formation of two additional peaks with retention times 15.2 min, corresponding to C16:1(Δ 11), and 19.9 min (Figure 3.3b). This suggests that the peak at 19.9 min observed in Δols -*desE*-up is indeed the C18:1(Δ 13) fatty acid.

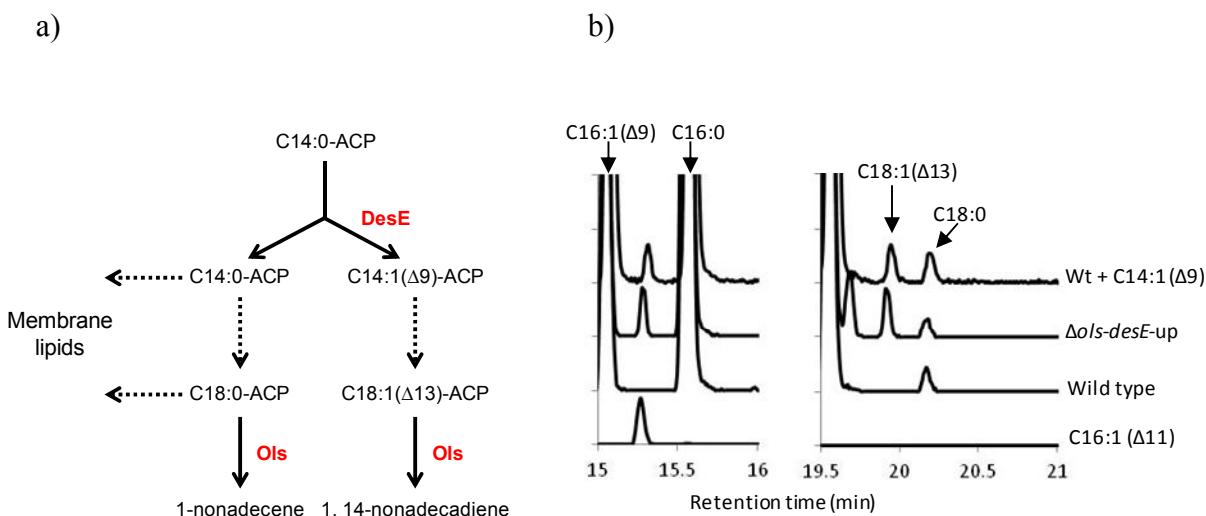


Figure 3.3 a) Proposed route for alkene biosynthesis in *Synechococcus sp.* PCC7002. DesE is required to place a double bond at the $\Delta 9$ position of the C14:0 acyl-ACP substrate. The product is subsequently elongated to a C18:1($\Delta 13$) acyl-ACP that would serve as the precursor for 1, 14-nonadecadiene biosynthesis by Ols. b) Comparison of the fatty acid profiles of PCC 7002 and mutants. Feeding of the C14:1($\Delta 9$) fatty acid to the wild type strain and over expression of *desE* in the Δols strain resulted in the formation of the C16:1($\Delta 11$) fatty acid (retention time 15.2 min) and the C18:1($\Delta 13$) fatty acid (retention time 19.9 min).

3.2.3 Growth in the presence of H_2O_2 and under nutrient limitation conditions

Growth under high light intensities or environmental conditions that slow down anabolism (nonoptimal temperature, nutrient limitation, etc.) may lead to a decrease in photochemical dissipation and oxidative stress (17). During oxidative stress free radicals can attack polyunsaturated fatty acids resulting in a decrease in membrane fluidity and the disruption of membrane bound proteins (18), including enzymes involved in nutrient uptake (10). Transcriptome profiling via RNAseq in *Synechococcus sp.* PCC 7002 has shown that the transcript levels for *desE* increased (more than threefold) in cells grown under nutrient limitation (N, S, P, Fe) (19) and under oxidative stress conditions (20). Only one other desaturase, *desB*, was upregulated under S and Fe limitation. It has been proposed that lipid unsaturation might

play a role in preventing inactivation of enzymes involved in nutrient uptake (10). This would be especially important under nutrient limitation conditions and at low temperatures, where nutrient uptake is reduced (10), as nutrient limitation conditions can result in a decrease in photochemical dissipation and therefore may lead to oxidative stress caused by excess absorbed light.

To test if the unsaturation synthesized by DesE might play a role during oxidative stress, the wild type and $\Delta desE$ strain were exposed to different concentrations of H_2O_2 in liquid cultures for 24 hours at 38°C and 22°C and aliquots were transferred onto solid medium. After 3 days no significant difference was observed between the two strains; both the wild type and the $\Delta desE$ cells were able to grow up to 2 mM H_2O_2 at 38°C and up to 1 mM at 22°C. The strains were also grown under nutrient limitation conditions (nitrogen, sulfate and phosphate) and the hydrocarbon profile was analyzed. As can be seen in Figure 3.4, no difference was observed in the alkene composition for the strains grown under nitrogen or phosphate limitation conditions when compared to Media A growth. However, an increase in the C19:2 levels and a decrease in the C19:1 alkene levels were observed for the wild type strain grown under sulfate limitation conditions. Although changes in expression of *desE* under oxidative stress and nutrient limitation conditions suggest that the unsaturated products from DesE might be important under these conditions, they might not be essential to survive as shown by no differences in growth for wild type and $\Delta desE$ strains exposed to H_2O_2 .

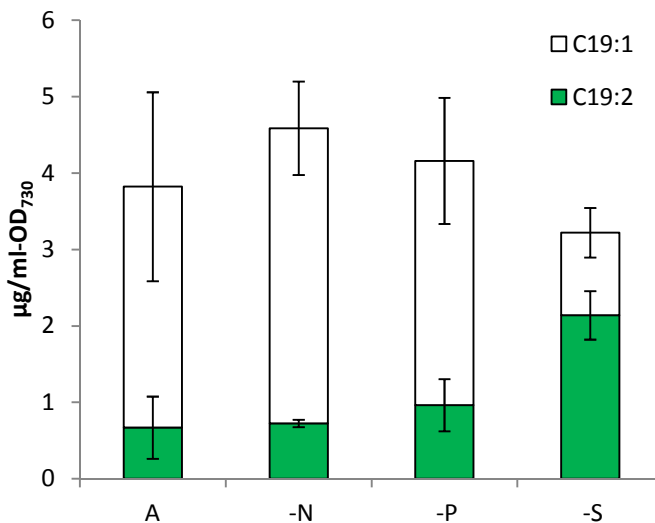


Figure 3.4 Comparison of the hydrocarbon composition from cultures grown in Media A (A) and under nitrogen (-N), phosphate (-P) and sulfate (-S) limitation conditions at 38°C. Under sulfate limitation conditions there is a decrease in the content of the C19:1 hydrocarbon and an increase in the C19:2 hydrocarbon. Error bars represent the standard deviations from three biological replicates.

3.2.4 Effect of temperature in alkene unsaturation

It is known that the expression of some desaturases is upregulated upon a downshift in temperature, resulting in an increase of the unsaturated lipids that are used to maintain membrane fluidity at low temperatures. In the case of *desE*, global transcriptome profiling via RNAseq showed that its transcript levels increased approximately two-fold in cells grown at 22°C compared to cells grown at 38°C (20). Also, it has been shown that its transcripts had an estimated half-life of only 1 min at 38°C and 21 min at 22°C (14) suggesting that *desE* may be regulated to respond to changes in temperature. To test if *desE* upregulation at low temperatures would affect the hydrocarbon composition of PCC 7002, the wild type strain was grown at different temperatures. At 38°C, the C19:2 hydrocarbon corresponds to only 11% of the total hydrocarbon content, whereas at 22°C it is about 96% (Figure 3.5a). Except for 38°C, most of

these differences arise from a decrease in C19:1 hydrocarbon rather than an increase of C19:2. For the $\Delta desE$ strain there was a slight decrease in C19:1 at low temperatures but, on average, hydrocarbon levels remained constant across all temperatures (Figure 3.5b).

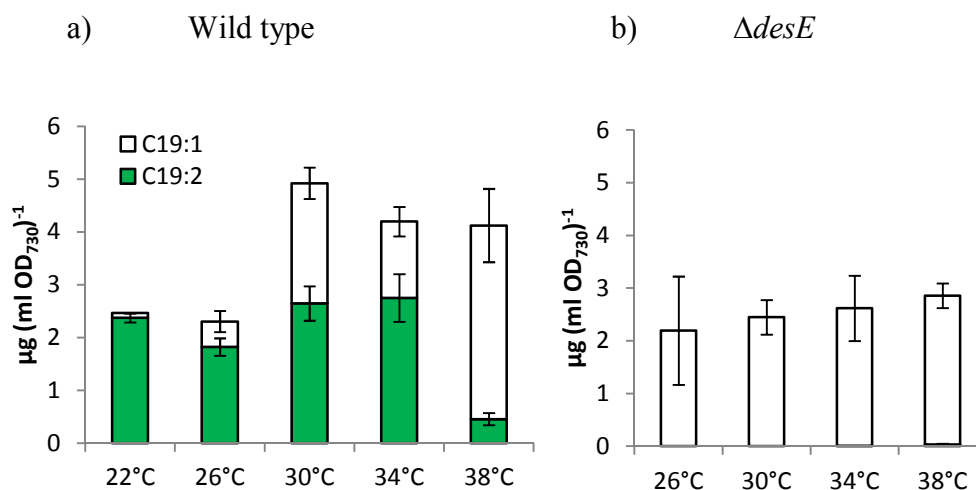


Figure 3.5 Comparison of the hydrocarbon composition from cultures grown at different temperatures. There is a decrease in the content of the less unsaturated hydrocarbon (C19:1) at low temperatures for the wild type strain (a) but not for the $\Delta desE$ strain (b). Error bars represent the standard deviations from three biological replicates.

Interestingly, no C18:1($\Delta 13$) fatty acid was detected at any of the temperatures tested, suggesting that this intermediate is not incorporated into lipids, but rather is exclusively converted to hydrocarbon by Ols. Hydrocarbons also seem to be important for growth, especially at low temperatures. Whereas the wild type strain was able to grow at 22°C after an initial lag period, the Δols strain failed to grow under these conditions. Conversely, the Δols strain grew robustly at 38°C albeit slower than the wild type (Figure 3.6).

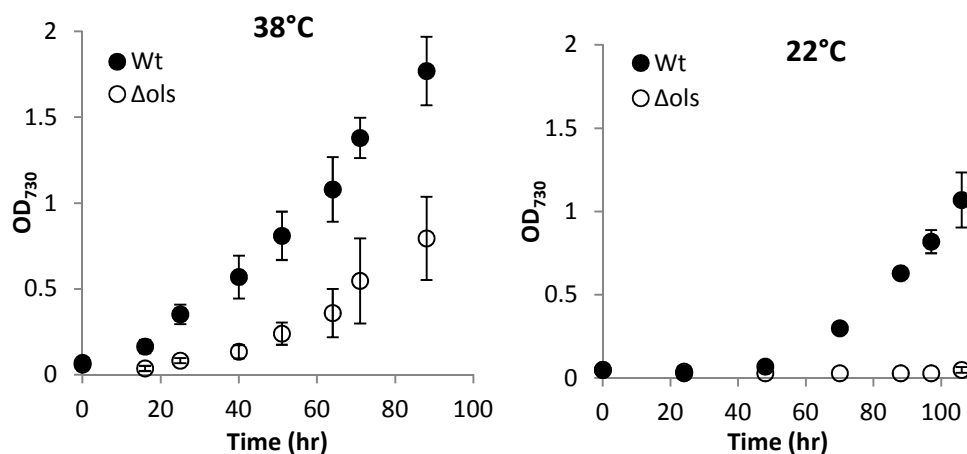


Figure 3.6 Growth curves for the wild type (wt) and Δols strain. The Δols strain was unable to grow at 22°C and grew slower than the wild type at 38°C.

3.2.5 Hydrocarbon localization

Since the hydrocarbon content in PCC 7002 represents about 10-15% of the total lipid content (for example at 36°C is $\sim 5.5 \mu\text{g/ml-OD}_{730}$ and the fatty acid content is $\sim 35 \mu\text{g/ml-OD}_{730}$), changes in alkene composition at low temperatures suggest that they might play a role in maintaining membrane fluidity in response to changes in temperature. However, it has also been shown that bacteria produce extracellular signaling molecules during carbon and energy starvation and that these molecules play an important role in the expression of proteins crucial to the development of starvation- and stress-resistant phenotypes (21). Therefore, we decided to determine whether the alkenes synthesized by PCC 7002 are exported outside the cell. A culture was grown at 33°C and after reaching an OD_{730} of approximately 0.5, it was centrifuged at high speed and the resulting pellet and supernatant were analyzed for alkene content. As can be seen in Figure 3.7, hydrocarbons were found only in the cell pellet; no hydrocarbons were detected in the supernatant. The resulting pellet was then subjected to cell lysis using a bead mill cell

disrupter. The resulting mixture was centrifuged again and the two phases (membranes and lysate supernatant) were analyzed for hydrocarbon content. As can be seen in Figure 3.7 both phases contained similar amounts of hydrocarbons. Since the membranes represent a much smaller fraction of the total cell volume (for example it has been reported that thylakoid membranes represent about 15% of the cell volume (22)), these results suggests that the majority of the hydrocarbons are associated with the membranes. This is also confirmed by the fact that a similar distribution was observed for the fatty acids from the same sample (fatty acids are mainly found as part of the cell membranes).

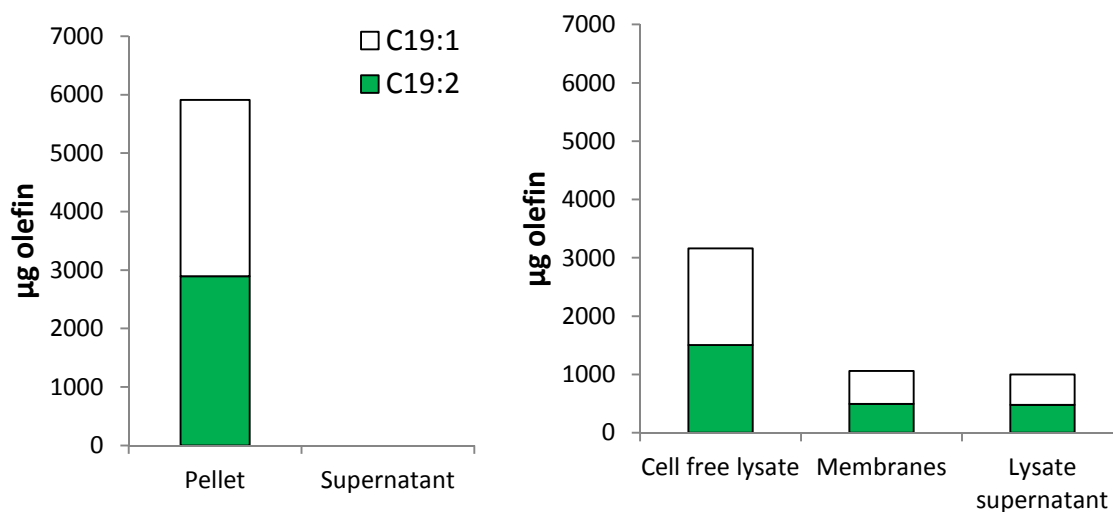


Figure 3.7 Hydrocarbons content from a wilt type PCC 7002 culture grown at 33°C. Left: Hydrocarbon are present only in the pellet and not in the supernatant after centrifugation. Right: analysis of the pellet after lysis suggests that hydrocarbons are part of the membrane.

3.3 Conclusions

In summary, we demonstrated the involvement of a desaturase gene (*desE*) in the formation of the internal double bond in the 1, 14-nonadecadiene synthesized by the cyanobacterium *Synechococcus sp.* PCC 7002. The amino acid sequence encoded by the *desE* gene shows a high degree of similarity to $\Delta 9$ desaturases suggesting that its most likely substrate is a C14 fatty acid, which after elongation to a C18:1($\Delta 13$) fatty acid would serve as the precursor for the formation of the hydrocarbon with the internal double bond at position 14th. When the expression of *desE* was increased by replacing its promoter, we only observed the formation of the C19:2 hydrocarbon (and not the C19:1) suggesting that the products of DesE are earmarked for the hydrocarbon pathway. Moreover, since no hydrocarbons coming from the other unsaturated C18 fatty acids synthesized by PCC 7002 have been detected, such as C18:1($\Delta 9$) or C18:2($\Delta 9, \Delta 11$), it seems like not only the presence of the additional unsaturation is important but also its location. Since the hydrocarbon content in PCC 7002 represents about 10-15% of the total lipid content (for example at 36°C is $\sim 5.5 \mu\text{g/ml-OD}_{730}$ and the fatty acid content is $\sim 35 \mu\text{g/ml-OD}_{730}$), the increase in C19:2 (diunsaturated hydrocarbon) production at low temperatures suggests that hydrocarbons might play a role in responding to cold to maintain membrane fluidity.

3.4 Methods

3.4.1 Media and growth conditions

A wild type strain of *Synechococcus sp.* PCC 7002 obtained from the Pasteur Culture Collection was grown photoautotrophically ($140 \mu\text{E m}^{-2} \text{ s}^{-1}$) in 20 ml Medium A (23) with aeration under

constant illumination from cool-white lamps at the specified temperatures. When required, streptomycin and kanamycin (final concentration 100 $\mu\text{g mL}^{-1}$) was used. Cultures under nutrient limitation conditions were grown as described by Ludwig et al (19). For growth with hydrogen peroxide, the OD_{730} of cells at the exponential growth phase was adjusted to 0.5 and hydrogen peroxide was applied to a range of final concentrations (0.1-10 mM). After incubation for 24 hours under light and aeration, aliquots (5 μL) were placed onto solid medium plates.

3.4.2 Strain construction

Enzymes and reagents were purchased from New England Biolabs or Fisher Scientific unless otherwise noted. Oligonucleotides used in this study were purchased from Integrated DNA Technologies, Inc. and are listed in Table 3.2. For the construction of desaturase knockouts (Table 3.3), the upstream and downstream flanking sequences of the desaturase genes (SYNPCC7002_A2756, SYNPCC7002_A0159, SYNPCC7002_A2198, SYNPCC7002_A2833, SYNPCC7002_A1989) were amplified by PCR (phusion polymerase) from genomic DNA isolated using commercial reagents (Promega). PCR products were digested with BamHI and PstI and gel purified using commercial kits (Qiagen). The *aadA* gene (strep^R), was excised from plasmid pSRA81 (24) with PstI and BamHI. The fragments were mixed in a 3:1:3 ratio (left flank: strep^R: right flank) and ligated with T4 DNA ligase. Ligation products were gel extracted, re-amplified and used for transformation in PCC 7002 as described by Frigaard et al. (24). The upstream region of *desE* was replaced with the *cpcBA* promoter from *Synechocystis sp.* PCC 6803 (16) in the wild type or Δols strain (7) by transformation of a DNA fragment assembled using the method described by Gibson (25) and containing the upstream and downstream

flanking sequences of the wild type promoter, the *aphII* gene that confers resistance to kanamycin (from pJ206 plasmid, DNA 2.0, Menlo Park, CA) and the *cpcBA* promoter. For the construction of the complementation strain ($\Delta desE+$) a DNA fragment assembled using the method described by Gibson and containing the *desE* under its native promoter, the *aphII* gene that confers resistance to kanamycin and the upstream and downstream flanking sequences of locus SYNPC7002_A1838 was used to transformation using a culture from the $\Delta desE$ strain. Complete segregation of the mutants was verified by colony PCR.

Table 3.2 Oligonucleotides used^a

No.	Oligonucleotide	Sequence
1	desA-a2	CCGTTCTGCAGGTTTCTTGCGCAAGGGTTACAGCTTCC
2	desA-a1	GCAACCATGGGAAACCAACGCAAGG
3	desA-b2	CCGTTGGATCCGTACGCTTCCATACCATGTTACCAATATCG
4	desA-b1	CCTCACAGGTTCCGGCTACAGTGG
5	desB-a2	CCGTTCTGCAGCCTTACAAACCCTAATCCGCCTTTATTCATTTC
6	desB-a1	CAGATCGAGGGGAACCTGGTTTGGC
7	desB-b2	CCGTTGGATCCCAACAACGCCTTGCAGAAAATCCCCAGC
8	desB-b1	CCAGTTTTAACAGACCTTGGGTAAAGGCTTC
9	desC-a2	CCGTTCTGCAGCCTTGTACCTACGGCGAAGGTTGG
10	desC-a1	GGAATACACTGACGAATACCGCATGGG
11	desC-b2	CCGTTGGATCCGGTAATTCCAATGCCGCCGGTAATCCAGT
12	desC-b1	GGTTCTGGGCGAATTGCGATCTTGAGG
13	desE-a2	CCGTTCTGCAGATTGTTGTCCCAAAGGGAAATATCTATTCCG
14	desE-a1	GACCCAGAAAACCCCGTAGAATAATCACG
15	desE-b2	CCGTTGGATCCAAATGCTTTCCTAACGAGTTGAGAATATCTTCTATG
16	desE-b1	CGTCGATTTTGCCTCATTAAATTTAGTTAAAGCAGC
17	desF-a2	CCGTTCTGCAGGAATAAAATTTGCGTTTGATCATTACCGCCAATTC
18	desF-a1	GGTTGAAACCATTTAGGGAAACCCATACTG
19	desF-b2	CCGTTGGATCCGAATTCTCAAACAATAGAACAAGACAAAGGGGAATATC
20	desF-b1	CGAGGCAGGTTTTGAGAGCGTCAAC
21	desE-US-Fw	TATGCACATATGCGTCGATTTTGCCTCATTAAATTTAGTTAAAGCAGC
22	desE-US-Rv	GCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAGCTTTCCTAACGAGTTGAGAATATCTTCTATGAAACCG
23	KM-Fw	TTAGAAAAATCATCGAGCATCAAATGAAACTGC
24	KM-Rv	GGACTCTTCTACAGGTGGGTATAGATTTGTTAAGCTTTGGCAGGATCCGGCTGCTAACAAA
25	Cpc-prom-Fw	CTTAACAAATCTATACCCACCTGTAGAGAAGAGTCC
26	Cpc-prom-Rv	GGGTCAAGAACGTTGCTGTAATGCGTCATGGAATTAATCTCCTACTTGACTTTATGAGTTGGG
27	desE-DS-Fw	ACGCATTACAGCAACGTTCTTGACCC
28	desE-DS-Rv	CTGCCCGCAGGCAAATCTGTTTTATCCATATGCTAGGGATTGGCCGCTTTTGTAGATC
29	pBAD18-Fw	GATAAAACAGAATTTGCGTGGCCGCGCAG
30	pBAD18-desE-USRv	GCTGCTTTAACTAAATTAATGAGGCAAAATCGACGCATATGTGCATAGGAGAAACAGTAGAGAGTTGCG ATAAAAAGCG
31	pBAD18-NdeI-Fw	CATATGGATAAAAACAGAATTTGCCTGGCGGCAG
32	pBAD18-A-acsRv	TAGGAGGTTACGGGGAAAAGCCAATAGGCATATGTGCATAGGAGAAAACAGTAGAGAGTTGCGATAAAAAGCG
33	A-acsA-Fw	CCATTGGCTTTTCCCCGTAACCTCCTA
34	A-acsA-Rv	GCAGTTTCATTTGATGCTCGATGAGTTTTTCTAACCTCGGCAGCAAAGTCTGGTG

35	KM-desE-Rv	<u>GTTCGGTGGTGACAGTTTCTGGGCTTTGGCAGGATCCGGCTGCTAACAAA</u>
36	desE-comp-Fw	<u>CCCAGAAACTGTCACCACCGAAAC</u>
37	desE-comp-Rv	<u>GTGTCGCCACAAATTTCTGACCCCCAGGGCATCGTTTTAGCAACG</u>
38	B-acs-Fw	<u>GGTCAGGAAATTGTGGGCGACAC</u>
39	B-acs-Rv	<u>CTGCCCCAGGCAAATTCTGTTTTATCCATATGCCAACAAAGCCTTTGCCGCTGATC</u>

^a Restriction sites are underlined. Oligonucleotides 1 to 20 were used for the desaturase knockouts (a1 and a2 primers were used for upstream region of the gene, b1 and b2 for downstream region), oligonucleotides 21 to 30 for the *desE*-up strain, oligonucleotides 31-39 and 23 for complementation strain

Table 3.3 Cyanobacteria strains used in this study

Strain	Phenotype	Reference
<i>Synechococcus sp.</i> PCC 7002	Wild type PCC 7002	Pasteur culture collection
$\Delta desA$	$\Delta desA :: aadA$	This study
$\Delta desB$	$\Delta desB :: aadA$	This study
$\Delta desE$	$\Delta desE :: aadA$	This study
$\Delta desF$	$\Delta desF :: aadA$	This study
<i>desE</i> -up	$\Phi(P_{cpcAB}-desE)$	This study
$\Delta desE+$	$\Delta desE :: aadA, \Delta acsA :: desE-aphII$	This study
Δols	$\Delta ols :: aadA$	(7)
$\Delta ols-desE$ -up	$\Phi(P_{cpcAB}-desE), \Delta ols :: aadA$	This study

3.4.3 Lipid analysis (GC/MS)

Cultures were grown to an $OD_{730} \sim 1.0$, centrifuged, resuspended in 3ml of water and extracted and analyzed following previously described protocols (7, 26). Samples were analyzed using a Shimadzu GCMP QP2010S gas chromatograph mass spectrometer equipped with an AOC-20i autoinjector and a Restek Rxi®-5ms column (catalog #13423). The temperature program was as follows: 100°C hold for 2 minutes, ramp from 100°C to 150°C at 80°C per min, hold for 4 minutes, ramp from 150°C to 218°C at 4°C per minute, ramp from 218°C to 325°C at 80°C/min and hold at 325°C for 2.5 minutes. A sample injection temperature of 250°C and volume of 1 μ L was used, along with a 1:10 split ration. The MS was operated in scanning mode between 50 and 350 m/z. Quantification was achieved by comparison of integrated peaks with calibration curves of FAME (Sigma) and 1-nonadecene standards (Fluka).

3.5 References

1. **Valderrama B.** 2004. Bacterial hydrocarbon biosynthesis revisited. *Studies in Surface Science and Catalysis* **151**:373–384.
2. **Aarts MGM, Keijzer CJ, Stiekema WJ, Pereira A.** 1995. Molecular Characterization of the CER1 Gene of Arabidopsis Involved in Epicuticular Wax Biosynthesis and Pollen Fertility. *The plant cell* **7**:2115–2127.
3. **Ashburner M.** 2003. *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
4. **Coates RC, Podell S, Korobeynikov A, Lapidus A, Pevzner P, Sherman DH, Allen EE, Gerwick L, Gerwick WH.** Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. (Personal communication).
5. **Winters K, Parker PL, Baalen C Van.** 1969. Hydrocarbons of Blue-Green Algae: Geochemical Significance. *Science* **163**:467–468.
6. **Han J, Calvin M.** 1969. Hydrocarbon distribution of algae and bacteria, and microbiological activity in sediments. *Proceedings of the National Academy of Sciences of the United States of America* **64**:436–43.
7. **Mendez-Perez D, Begemann MB, Pflieger BF.** 2011. Modular synthase-encoding gene involved in α -olefin biosynthesis in *Synechococcus* sp. strain PCC 7002. *Applied and environmental microbiology* **77**:4264–7.
8. **Ladygina N, Dedyukhina EG, Vainshtein MB.** 2006. A review on microbial synthesis of hydrocarbons. *Process Biochemistry* **41**:1001–1014.
9. **Chapman D.** 1975. Phase Transitions and Fluidity Characteristics of Lipids and Cell Membranes. *Quarterly Reviews of Biophysics* **8**:185–235.
10. **Sakamoto T, Shen G, Higashi S, Murata N, Bryant D a.** 1998. Alteration of low-temperature susceptibility of the cyanobacterium *Synechococcus* sp. PCC 7002 by genetic manipulation of membrane lipid unsaturation. *Archives of microbiology* **169**:20–8.
11. **Sakamoto T, Wada H, Nishida I, Ohmori M, Murata N.** 1994. Identification of conserved domains in the delta 12 desaturase of cyanobacteria. *Plant molecular biology* **24**:643–650.
12. **Sakamoto T, Bryant D.** 1997. Temperature-regulated mRNA accumulation and stabilization for fatty acid desaturase genes in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *Molecular microbiology* **23**:1281–92.

13. **Tasaka Y, Gombos Z, Nishiyama Y, Mohanty P, Ohba T, Ohki K, Murata N.** 1996. Targeted mutagenesis of acyl-lipid desaturases in *Synechocystis*: evidence for the important roles of polyunsaturated membrane lipids in growth, respiration and photosynthesis. *The EMBO journal* **15**:6416–25.
14. **Sakamoto T, Stirewalt VL, Bryant DA.** 1997. Two acyl-lipid $\Delta 9$ desaturase genes of the cyanobacterium, *Synechococcus* sp. strain PCC 7002, p. 380–382. *In* *Physiology, biochemistry, and molecular biology of plant lipids*.
15. **Murata N, Wada H.** 1995. Acyl-lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobacteria. *The Biochemical journal* **308**:1–8.
16. **Xu Y, Alvey RM, Byrne PO, Graham JE, Shen G, Bryant DA.** 2011. Expression of Genes in Cyanobacteria: Adaptation of endogenous plasmids as platforms for high-level gene expression in *Synechococcus* sp. PCC 7002. *Photosynthesis Research Protocols* **684**:273–293.
17. **Perelman A, Uzan A, Hacoheh D, Schwarz R.** 2003. Oxidative Stress in *Synechococcus* sp . Strain PCC 7942 : Various Mechanisms for H₂O₂ Detoxification with Different Physiological Roles. *Journal of bacteriology* **185**:3654–3660.
18. **Cabiscol E, Tamarit J, Ros J.** 2000. Oxidative stress in bacteria and protein damage by reactive oxygen species. *International microbiology : the official journal of the Spanish Society for Microbiology* **3**:3–8.
19. **Ludwig M, Bryant D a.** 2012. Acclimation of the Global Transcriptome of the Cyanobacterium *Synechococcus* sp. Strain PCC 7002 to Nutrient Limitations and Different Nitrogen Sources. *Frontiers in microbiology* **3**:145.
20. **Ludwig M, Bryant DA.** 2012. *Synechococcus* sp. Strain PCC 7002 Transcriptome: Acclimation to Temperature, Salinity, Oxidative Stress, and Mixotrophic Growth Conditions. *Frontiers in microbiology* **3**:354.
21. **Srinivasan S, Ostling J, Charlton T, De Nys R, Takayama K, Kjelleberg S.** 1998. Extracellular signal molecule(s) involved in the carbon starvation response of marine *Vibrio* sp. strain S14. *Journal of bacteriology* **180**:201–9.
22. **Liberton M, Austin JR, Berg RH, Pakrasi HB.** 2011. Unique thylakoid membrane architecture of a unicellular N₂-fixing cyanobacterium revealed by electron tomography. *Plant physiology* **155**:1656–66.
23. **Stevens SE, Patterson CP, Myers J.** 1973. The production of hydrogen peroxide by blue-green algae: a survey. *Journal of Phycology* **9**:427–430.
24. **Frigaard N-U, Sakuragi Y, Bryant DA.** 2004. Gene inactivation in the cyanobacterium *Synechococcus* sp. PCC 7002 and the green sulfur Bacterium *Chlorobium tepidum* using

in vitro-made DNA constructs and natural transformation. *Methods in molecular biology* **274**:325–340.

25. **Gibson DG, Young L, Chuang R, Venter JC, Iii CAH, Smith HO, America N.** 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases **6**:12–16.
26. **Lennen RM, Braden DJ, West RA, Dumesic JA, Pfleger BF.** 2010. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnology and bioengineering* **106**:193–202.

Chapter 4: A translational-coupling DNA cassette for monitoring protein translation in *Escherichia coli*

*Portions of this chapter were published in *Metabolic Engineering* (1)

4.1 Introduction

For metabolic engineers, heterologous gene expression has become a preferred tool for imparting novel biochemistries, by-passing native pathways, or enhancing metabolic activity in model hosts (2–5). Of the model microorganisms used in metabolic engineering studies, *Escherichia coli* remains a popular choice because of the availability of a large knowledge base, a wide variety of engineering tools (e.g. vectors, promoters, construction methodologies), and well characterized genetic systems (6–8). Despite these advantages, heterologous gene expression in *E. coli* is not always straightforward. The development of algorithms for codon optimization (9) and inexpensive gene synthesis (10) methods have dramatically increased success rates but have not solved all problems. The major challenge to using heterologous expression in metabolic engineering experiments lies in the inability to quickly dissect experiments that have failed. If an expected phenotype is not observed, a researcher must determine which step in the path from gene to activity is the cause of failure.

While facile assays for the presence of any target mRNA can be developed (e.g. qPCR, hybridization), assays for monitoring protein activity are highly dependent on the properties of the target gene product. Conventional assays for detecting the presence of target protein, such as

SDS-PAGE, Western blot, mass spectrometry, are not always feasible (for both technical and economic reasons) or compatible with producing active protein *in vivo*. For example, a target protein must be found in high concentration in or purified from a crude cell lysate in order for a specific band to be detected by SDS-PAGE analysis. Western blots offer a lower limit of detection in complex mixtures but require the existence of an antibody that specifically recognizes the target protein. Fusion tags (e.g. 6X-histidine, maltose binding protein, FLAG, fluorescent proteins) can facilitate both purification and Western blot detection but can also interfere with protein folding and activity. For these reasons, the ability to monitor protein expression directly in hosts without significant sequence modification is limited. Therefore, a tool that can be used in *E. coli* to test for translation *in vivo* without affecting the activity of the target protein is of great interest. Such a method could be used to simplify and speed the process of optimizing heterologous expression of novel genes in both protein production and metabolic engineering applications.

To address this challenge, we have developed a synthetic biology method of indirectly detecting protein translation that is based on a microbial method of gene regulation. Often, each cistron of a polycistronic mRNA has its own Shine-Delgarno sequence, which in principle initiates translation independent of other cistrons. However, it has been observed that translation of some genes in a polycistronic mRNA is dependent on the translation of a contiguous upstream gene. This interdependence of translation efficiency of neighboring genes, called translational coupling, is used in phage and bacteria to regulate expression of some proteins (11). In *E. coli*, translational coupling has been observed for genes encoding ribosomal proteins (12, 13) and genes *trpBA* (14), *trpDE* (15) and *galKT* (16). Translational coupling is mediated by mRNA secondary structures that outcompete formation of the mRNA:rRNA base pairs that lead to

translation initiation. When the gene in the 5' position is not translated, the secondary structure forms and prevents translation of the gene in the 3' position. Conversely, when the gene in the 5' position is translated, the helicase activity of the ribosome (17) melts the secondary structure, thereby promoting binding of the ribosome and translation initiation of the gene in the 3' position.

Here, we describe a method to quickly determine whether *E. coli* is capable of expressing the product of any target gene by coupling translation of a target gene to a detectable response gene. A translational coupling cassette was designed to encode a mRNA sequence that forms a secondary structure in the absence of translation and contains the translational start sequence of a detectable response gene. The translational coupling method was successfully tested with fluorescent proteins and antibiotic resistance markers. Only when the target gene was fully translated was the response observed. Further characterization demonstrated that translational coupling functions at both low and high levels of expression and that the response signal is proportional to the amount of target gene product. We anticipate that the translational coupling system can be used to detect complete translation, select for conditions and/or sequences that maximize expression, and tightly regulate multiple genes at desired ratios.

4.2 Results

4.2.1 Construction of a translation-coupling DNA cassette for monitoring protein translation in *Escherichia coli*

To test translation of a heterologous gene in *E. coli*, we designed a cassette that couples the full translation of a target gene to that of a detectable response gene. Translational coupling is

achieved by occluding the initiation site (ribosome binding site (RBS)) of the response gene by formation of a mRNA secondary structure comprised of the 3' end of the target gene's coding region and the response gene's translation initiation site. In the absence of a ribosome actively translating the target gene, formation of a mRNA hairpin comprising the RBS of the response gene will outcompete base-pairing between the 16S rRNA and the RBS that would initiate translation of the response gene. Therefore, active translation of the response gene requires a means of disrupting the inhibitory mRNA hairpin. The ribosome is known to have RNA helicase activity and can translate mRNA with significant secondary structure (17). In the case of the coupling cassette, active translation of the target gene would lead to denaturation of the hairpin, revealing the SD-AUG motif of the response gene, and allowing its translation (Figure 4.1). Selection of appropriate response genes would enable screens based on fluorescence or other detectable signals, selection using resistance markers, or coordinated expression of multiple genes.

To implement our design we constructed a reporter operon containing the translational coupling cassette in a pTrc99A derivative (native RBS and multi-cloning site removed). The plasmid includes the ampicillin resistant gene, an isopropyl- β -D-thiogalactoside (IPTG)-inducible promoter (Ptrc) (18), the *rrnB* anti-termination region (19), a ColEI origin of replication, the gene encoding the Lac repressor (*lacIq*) and restriction sites for cloning a target gene. The translational coupling sequence consists of three elements: a C-terminal 6x-histidine tag for isolation of target protein, a stop codon for the target gene, and the ribosome binding site (RBS) for the response gene. The sequence was designed so that it forms a strong hairpin ($\Delta G = -16.1$ kcal/mol) that base-pairs with the Shine-Delgarno sequence upstream of the response gene (Figure 4.1). To use the cassette, a target gene is cloned into the restriction sites without a stop

codon (ligation independent cloning can also be used in some versions of the system (20)). After induction with IPTG, translation of the target gene can be monitored using the response gene signal.

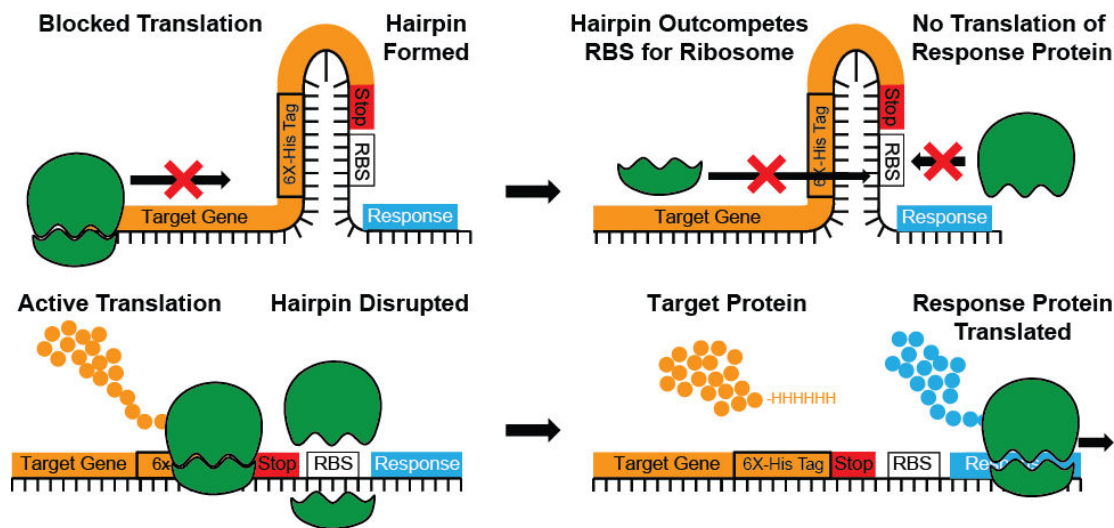


Figure 4.1 Schematic of translational coupling cassette. Top: When complete translation of a target gene is blocked, a hairpin consisting of a C-terminal His-Tag and the RBS of the downstream response gene forms; thereby preventing ribosome recruitment and translation of the response gene. Bottom: When full translation of a target gene occurs, disruption of the secondary structure by ribosomal helicase activity allows initiation of translation of the response gene.

4.2.2 Test for RFP translation using antibiotic resistance as the response signal

To test the system, translation of a target gene encoding a red fluorescent protein (RFP^{EC}) (21) was coupled to the *aphII* gene that confers resistance to kanamycin (response gene). Therefore, if translation of RFP^{EC} occurred, cells were expected to be kanamycin resistant. After cloning RFP^{EC} and *aphII* into the system (resulting in plasmid pRFP-KM), an overnight culture of *E. coli* DH10B pRFP-KM was grown in LB media containing IPTG (0.25 mM) and kanamycin (50

$\mu\text{g/ml}$). Optical density (OD_{600}) and fluorescence were recorded over time and plotted in Figure 4.2. Cells harboring pRFP-KM were able to grow in the presence of kanamycin (up to 200 $\mu\text{g/ml}$) and red fluorescence was detected. As a control for translation, we constructed an *rff*^{EC} variant containing a premature stop codon that prevents full translation of the gene (plasmid pRFP^{nf}-KM). The mutant was grown under the same conditions and no fluorescence or increase in OD_{600} was observed after 16 h (Figure 4.2). RFP^{EC} protein was purified from cultures harboring each construct using immobilized Ni-NTA affinity chromatography. Samples were analyzed by SDS-PAGE and a band of the expected size (26.7 kDa) was observed for pRFP-KM and no band was observed for pRFP^{nf}-KM (data not shown).

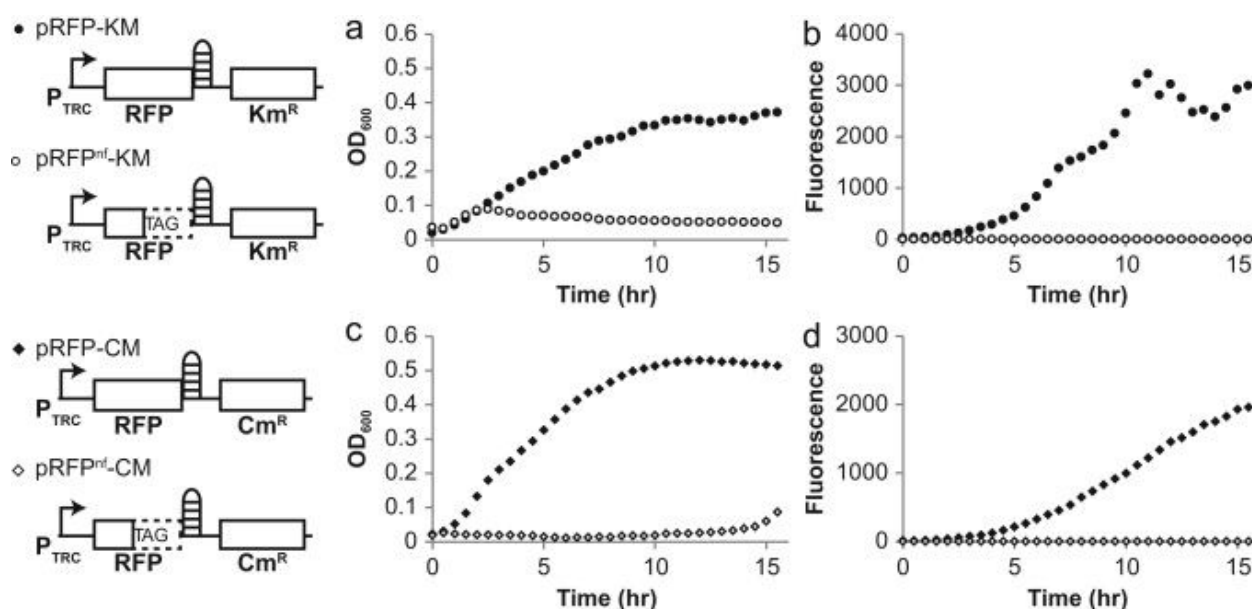


Figure 4.2 Demonstration of translational coupling using RFP as target gene and antibiotic resistance markers as response genes. Translation of RFP was blocked in samples harboring RFP^{nf} by introducing a stop codon midway through the gene. Optical density (a) and (c) and fluorescence (b) and (d) timecourses of *E. coli* harboring translational coupling cassettes (left). Cells were grown at 37°C in microtiter plates in the presence of 50 $\mu\text{g/ml}$ kanamycin (a) and (b) or 34 $\mu\text{g/ml}$ chloramphenicol (c) and (d).

The experiments were repeated using *cat*, which encodes a chloramphenicol resistance marker, as the response gene in plasmids pRFP-CM and pRFP^{nf}-CM. As expected growth and fluorescence in the presence of different concentrations of chloramphenicol were only observed for cultures harboring pRFP-CM (Figure 4.2). These results confirm that the translational coupling cassette can be used as an indirect measure of full translation of a target gene in *E. coli* as well as a platform for producing and purifying target proteins.

4.2.2.1 Optimizing growth conditions using antibiotic resistance as the response signal to test for translation

As can be seen in Figure 4.2, when CM^R was used as a response signal to test for translation, a long lag was observed. We argue that the culturing conditions (96 well plate, incubated in a plate reader) are a major contributor to growth after long lags. We ordinarily ignore data collected after 12 h of incubation in a plate reader due to evaporation from outer wells, cell clumping caused by imperfect mixing, and poor aeration. A similar phenomenon was observed when KM^R was used as the response gene. In approximately 50% of the experiments where a target gene was not fully translated (like RFP^{nf}), growth was observed after a lag of 12-20 hours in a plate reader. Growth may be due to leaky expression, mutations, or possibly nonspecific antibiotic degradation. In order to test what the cause was for this growth, we repeated we used an experiment where 2 out of 6 induced wells grew after a lag of 17 hrs and we subcultured these wells into fresh media containing IPTG, Amp, and KM. In this case no lag was observed, indicating that either a mutant or a contaminant was responsible for growth. Also, when cultures harboring pRFP^{nf}-CM or pRFP^{nf}-KM were grown in 5 mL tubes in an incubator shaker, no

growth was observed in cultures containing the antibiotic corresponding to the response gene after 24 hours suggesting that culturing conditions might be the cause for the growth. Also, the lag in growth was dependent on CM concentration and IPTG concentration with longer lags correlating with higher CM levels (Figure 4.3a). We also performed a similar experiment using solid media. Figure 4.3b shows the results from these experiments, as can be seen from the figure the constructs harboring the premature stop version of RFP (pRFP^{nf}-CM) do not generate CM^R colonies.

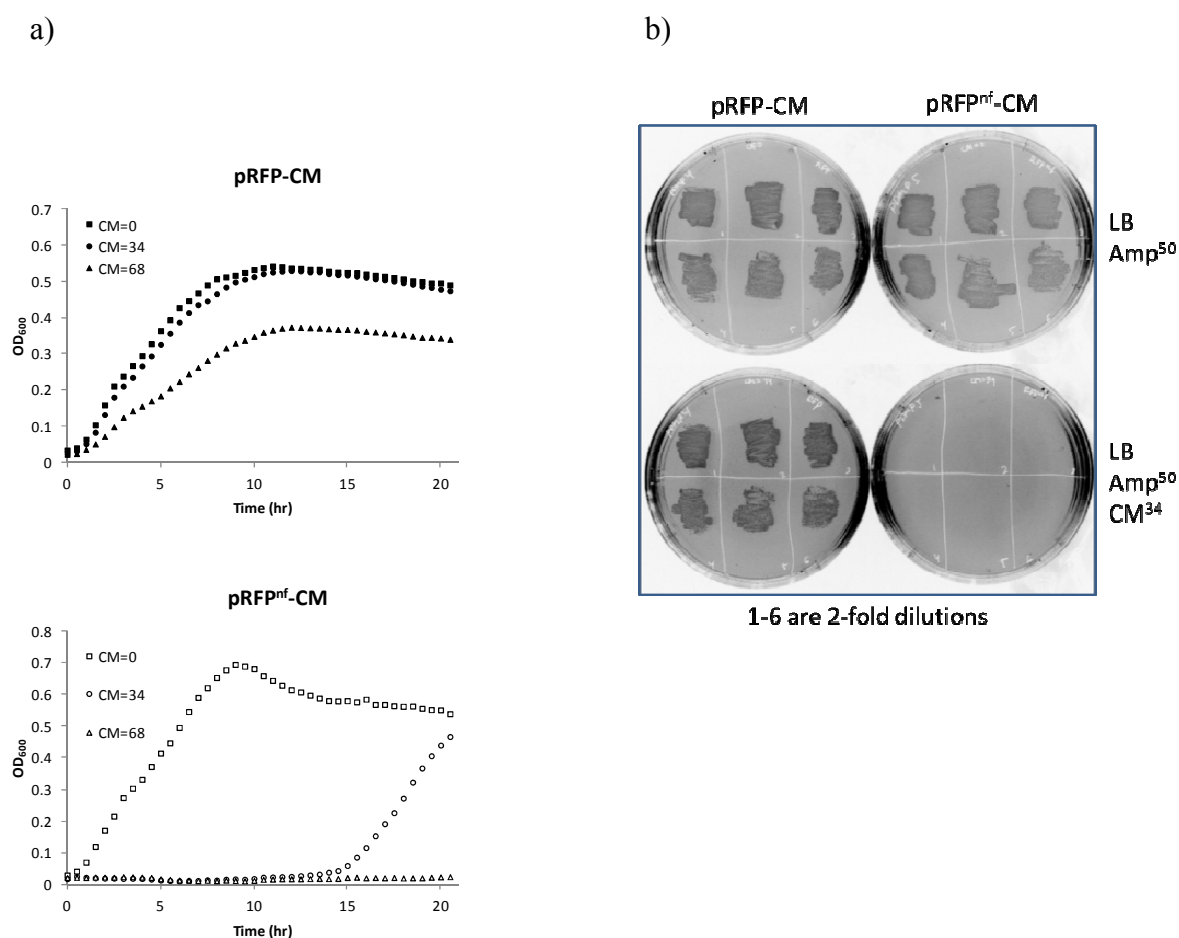


Figure 4.3 a) Demonstration of translational coupling using pRFP^{nf}-CM as target gene and antibiotic resistance markers as response genes. Translation of RFP was blocked in samples harboring RFP^{nf} by introducing a stop codon. b) Testing for translational coupling on solid media. Different concentrations of chloramphenicol CM=0 $\mu\text{g/ml}$ (●), CM=34 $\mu\text{g/ml}$ (▲), CM=68 $\mu\text{g/ml}$ (○) IPTG=0.2 mM, 37°C

To prevent false positives when testing expression of heterologous proteins using this system, we recommend use of pLIC-CM and comparison of growth over a period no more than 12 h in the presence of high concentrations of chloramphenicol (>50 $\mu\text{g/ml}$).

4.2.3 Detecting differences in expression of the target gene

The data presented above demonstrates the ability to use a translational coupling cassette as a qualitative means of confirming full-length translation of a target gene in *E. coli*. Next, we examined the ability of the cassette to function at various levels of transcription and translation. Our goal was to determine if the system could be used to determine the relative abundance of a heterologously expressed target gene. We hypothesized that the frequency of translation of a target gene would be coupled to the rate of translation initiation of a response gene. To test this hypothesis, a new operon was constructed using the RFP^{EC} as the response signal and a gene that encodes for a yellow fluorescent protein (YFP) (22) as the target gene (resulting in plasmid pYFP-RFP).

As a first test, the rate of transcription was varied by titrating IPTG from 0 to 0.8 mM. The yellow fluorescence intensity of each well (from target gene) was normalized to the highest yellow value observed. Similarly, the red fluorescence intensity of each well (from response gene) was normalized to the highest red value observed. As seen in Figure 4.4a, levels of YFP fluorescence increased when exposed to increasing concentrations of IPTG from 0 to 0.2 mM and declined when exposed to concentrations above 0.2 mM. Interestingly, a nearly identical trend in red fluorescence was observed for the response gene, RFP^{EC} (Figure 4.4b). When plotted against each other (Figure 4.4d), the yellow and red fluorescence values trended linearly. The linear relationship shifted at low IPTG concentrations where significantly less red fluorescence

(response gene) was observed. The data indicates that translational coupling functions at both low and high levels of expression caused by changes in transcription initiation but is proportional only above a minimum expression threshold.

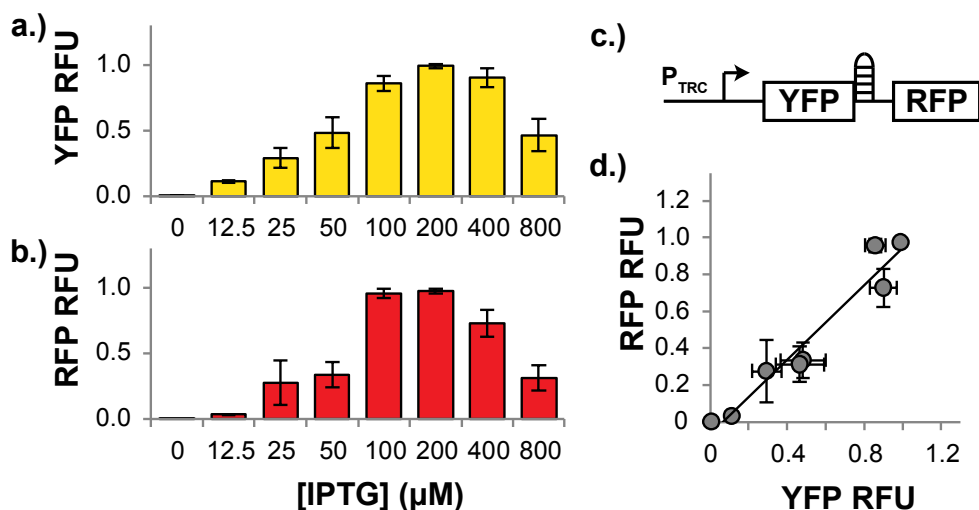


Figure 4.4 Translational coupling functions at high and low levels of expression. Fluorescence of cultures harboring pYFP-RFP grown in the presence of different IPTG concentrations were normalized to the highest observed value: (a) YFP fluorescence, (b) RFP fluorescence and (c) gene arrangement in pYFP-RFP and (d) YFP-RFP correlation. Error bars represent biological replicates.

In a second test we maintained a constant IPTG concentration (0.2 mM) and changed the translation initiation rates of the target gene (YFP) by cloning a library of ribosome binding sites of different strength. A ribosome binding site calculator (23) was used to design a 34 base pair ribosome binding site containing 5 degenerate base pairs that, when changed, were predicted to alter translation initiation rates from 4 to 10,000 (arbitrary units). A degenerate PCR primer containing random nucleotides in these positions was used to amplify *yfp* and clone it into a

plasmid containing the translational coupling cassette upstream of *rfp*^{EC}. After transformation of the RBS library, colony PCR was performed on 96 colonies to determine which contained a library insert. From the clones containing an insert (~50), 11 clones covering the range of observed fluorescence intensities were grown in triplicate 200 μ L cultures using a 96-well plate. Data was normalized as described above (Figure 5a and b). When plotted against each other (Figure 5d), the relative levels of YFP fluorescence correlate with the RFP signal (correlation coefficient $R^2=0.95$) above a minimum expression threshold suggesting that the system is capable of detecting different levels of target gene translation by measuring the level of the response signal.

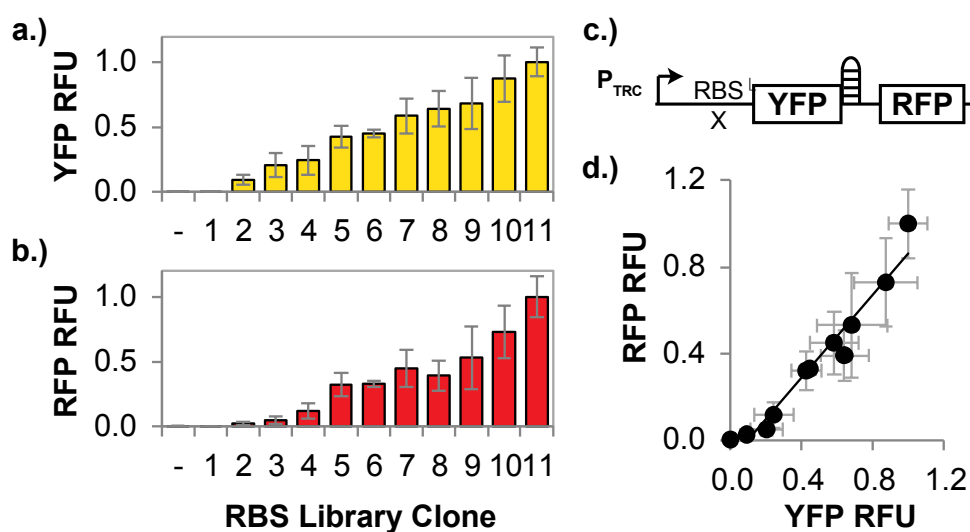


Figure 4.5 Translational coupling generates a proportional response. Fluorescence of cultures harboring individual members of the pYFP-RFP RBS library, grown in the presence of 0.2 mM IPTG, were normalized to the highest observed value: (a) YFP fluorescence, (b) RFP fluorescence, (c) gene arrangement in pYFP-RFP1-11 and (d) YFP-RFP correlation. Error bars represent biological replicates.

The eleven constructs were sequenced and their RBS are shown in Table 4.1 (Methods section). For comparison, the sequences of these eleven constructs were used to predict the translation initiation rate (TIR) of the YFP gene for each of the RBS using the ribosome binding site calculator. As can be seen in Figure 4.6, the predicted TIR did not correlate with the observed RFP fluorescence.

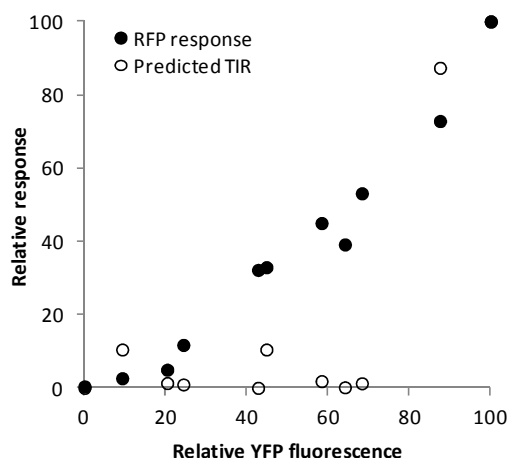


Figure 4.6 Comparison of the relative fluorescence for the RBS library constructs. Values presented as percentage of the maximum value observed (TIR=translation initiation rate). From the figure, the relative levels of YFP fluorescence correlate with the RFP signal (correlation coefficient $R^2=0.95$) whereas the predicted TIR using the RBS calculator does not correlate with the experimental YFP fluorescence values.

4.3 Conclusions

The range of products that can be produced in bacteria has been dramatically enhanced by the use of heterologous expression. Furthermore, heterologous expression is a powerful metabolic engineering tool for improving titers, increasing fitness, and maximizing yield. Here, we presented a method for dissecting failed heterologous expression experiments in *E. coli*. We designed a DNA cassette that couples translation of a target gene to a response gene that

generates an easily monitored phenotype *in vivo*, such as antibiotic resistance or fluorescence. The cassette is a powerful tool that could potentially be used to easily screen/select a large number of mutants or conditions to identify improvements in heterologous expression in *E. coli*. The cassette does not require expensive equipment, is compatible with any gene that can be cloned, and does not require the presence of a fusion tag that could alter the structure of the target gene product. Therefore, the coupling cassette could be useful in a wide range of applications. For testing expression of heterologous proteins, we recommend use of pLIC-CM and comparison of growth over a period no more than 12 h in the presence of high concentrations of chloramphenicol (>50 µg/ml).

We demonstrated the function of system by coupling the expression of either a functional RFP gene or a variant containing a premature stop codon that prevents full translation to an antibiotic resistance marker. As expected, cells that generated red fluorescence were also resistant to the corresponding antibiotics, whereas cells that did not produce red fluorescence were unable to grow in the presence of antibiotics. In the course of this work we tested two translational coupling sequences. The first, plasmid pBP18, contained a smaller (11 vs. 15 base pair) stem and a stronger Shine-Delgarno sequence (AGGAGG vs. UGGUGG) but failed to prevent chloramphenicol resistance in the absence of RFP translation (data not shown). This suggests that translational coupling is sensitive to the strength of the hairpin and that further sequence modulation could alter the degree of coupling.

The coupling cassette functions at a wide range of expression levels and generates a proportional response above a minimal expression threshold. We hypothesize that below this threshold, low initiation rates lead to increased ribosome spacing, and decreased frequency of hairpin melting. In this situation, the lack of consistent helicase activity may not provide sufficient time for

recruitment of the ribosomal 16S rRNA prior to refolding of the coupling hairpin. The proportional response suggests that the coupling cassette could be used to quantitatively monitor expression of target genes *in vivo*. This approach would be useful in metabolic engineering projects for confirming protein levels from heterologous metabolic pathways or for comparing the amount of a desired protein between biological samples, e.g. between two clones of a mutagenized or designed library. In each case, the experiments could be completed without modifying the target protein, time-consuming protein preparation, or expensive analytical equipment.

Beyond challenges in heterologous expression, we hypothesize that the coupling concept could be used to control expression of genes in an operon at a precise ratio. Future work will examine the effect of the hairpin sequence on the degree of coupling and ability to alter expression ratios. A practical application for the system will be presented in the next chapter, where the translational coupling system was used to optimize conditions for expressing a large natural product megasynthase in *E. coli* and to identify regions of the gene in which translation failed.

4.4 Methods

4.4.1 Plasmid construction and oligonucleotides

Plasmid construction and maintenance was performed using *E. coli* DH10B (Invitrogen, Carlsbad, CA). Plasmids were prepared by alkaline lysis (Qiagen, Valencia, CA). Oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, Iowa). Cloning enzymes, including Phusion® DNA polymerase, restriction enzymes, and T4 DNA ligase were purchased from New England Biolabs (NEB, Ipswich, MA) and used according to the

manufacturer's instructions. Plasmids used in this paper are listed in Table 1. New plasmid sequences are deposited in the NCBI nucleotide collection. All cloning procedures were confirmed by DNA sequencing.

Reporter plasmids (Table 1) were constructed to test the ability to couple translation of genes in an operon. All primers used for cloning are listed in Table 2. To construct pLIC-KM, pLIC-RFP, and pLIC-CM, the corresponding response gene was amplified from pJ206 (DNA 2.0, Menlo Park, CA), pRFP^{EC} (21) in a two-step PCR process that added ~60 base pairs encoding the translational-coupling hairpin to the 5' end of each response gene. Each PCR product was digested with SpeI-KpnI or BglIII-KpnI and ligated with a correspondingly digested vector that was constructed by amplifying pTRC99A by PCR. Target genes (e.g. *rfp*^{EC}, *yfp*, *ols*, *ols* domains, *ols*^{opt}) were amplified by PCR using primers that added the appropriate restriction sites (EcoRI-BglIII or Mfe-SpeI), a strong ribosome binding site (AGGAGG) and eliminated the stop codon of each target gene. pRFP^{nf}-KM was constructed from pRFP-KM by PCR using primers that introduced a point mutation in RFP resulting in a stop codon. PCR products were cloned into pLIC-KM, pLIC-RFP, or pLIC-CM by restriction digestion and ligation.

4.4.2 Microtiter plate growth assays

Overnight cultures grown from a single colony were diluted to an OD₆₀₀~0.05–0.1 in Luria-Bertani (LB) medium containing 100 µg/ml ampicillin to maintain plasmids, varying concentrations of isopropyl β-D-1-thiogalactopyranoside (IPTG) to induce transcription, and varying concentrations of kanamycin or chloramphenicol when required. All growth experiments were performed in triplicate. For each assay, 200 µL of culture was added to an individual well

of a black, clear-bottomed 96-well polystyrene plate (Costar, Corning, NY) and incubated for 24 h with shaking at 37°C or 30°C using a Tecan (Männedorf Switzerland) M1000 microplate reader. Fluorescence was measured at 558 nm excitation with emission at 583 for RFP^{EC} and 514 nm excitation with emission at 527 nm for YFP.

4.4.3 Protein isolation

To produce and purify proteins, a single colony of *E. coli* DH10B harboring the plasmid containing the target gene was inoculated into 10 mL of LB broth containing 100 µg/ml ampicillin, and was incubated at 37 °C overnight. The culture was diluted with fresh medium (100 mL) and grown at 30°C to an OD₆₀₀ of ~0.4 at which point IPTG was added to a final concentration of 1 mM. Induced cultures were incubated for an additional 20 h at 15°C. After finding the best conditions for translation, expression of the pOLS-Opt construct was performed in a 50 mL culture grown at 30 °C overnight with 100 µg/ml ampicillin and 0.01 mM IPTG. For each protein preparation, cells were collected by centrifugation and lysed via sonication. The cleared lysate was mixed with Ni-NTA agarose (Qiagen) for 1 h at 4°C with gentle agitation. The resin was washed (30 mM imidazole, 50 mM NaH₂PO₄, 300 mM NaCl) prior to elution of the 6x-his-tagged target protein with elution buffer (250 mM imidazole, 50 mM NaH₂PO₄, 300 mM NaCl). Purified proteins were separated by 8–12% SDS-PAGE and identified by size in comparison to bands from a Precision Plus Protein dual color standard ladder (BIO-RAD, Hercules, CA).

Table 4.1 Plasmids constructed for and used in this chapter

Plasmid	Genotype/Features	Source
pTRC99A	Backbone for translational coupling cassette	(18)
pJ206	Used to amplify <i>aphII</i> and <i>cat</i>	DNA 2.0
pRFP ^{EC}	Used to amplify RFP gene	(21)
pAQ1Ex-PcpcBA::YFP	Used to amplify YFP gene	(22)
pLIC-KM	Translational coupling cassette using <i>aphII</i> as response gene	This report
pLIC-CM	Translational coupling cassette using <i>cat</i> as response gene	This report
pRFP-KM	pLIC-KM with RFP as the target gene	This report
pRFP ^{nf} -KM	pLIC-KM with RFP with premature stop codon as response gene	This report
pRFP-CM	pLIC-CM with RFP as the target gene	This report
pRFP ^{nf} -CM	pLIC-CM with RFP with premature stop codon as response gene	This report
pLIC-RFP	Translational coupling cassette using RFP as response gene	This report
pYFP-RFP	pLIC-RFP with YFP as the target gene	This report
pYFP-RFP-library	pYFP-RFP with different ribosome binding sites upstream YFP	This report
	RBS Sequence (5'-3') of characterized clones (Start codon in bold)	
pYFP-RFP1	TGTTAGCGCCGAGGGAATTAAGTAGCTAATTAGA ATG	
pYFP-RFP 2	TGTGAGCGGCGAGGGAATTAAGGAGTTAATTAGA ATG	
pYFP-RFP 3	TGTTAGCGTTAGGTAGGTAATTAGA ATG	
pYFP-RFP 4	TGTGAGCGCCGAGGGAATTATGGAGATAAATTAGA ATG	
pYFP-RFP 5	TGTTAGCGTCGAGGGAATTATGCAGTTAATTAGA ATG	
pYFP-RFP 6	TGTTAGCGGCGAGGGAATTAAGGAGATAAATTAGA ATG	
pYFP-RFP 7	TGTCAGCGGCGAGGGAATTATGGAGTTAATTAGA ATG	
pYFP-RFP 8	TGTCAGCGGCGAGGGAATTATGAAGATAAATTAGA ATG	
pYFP-RFP 9	TGTGAGCGTCGAGGGAATTATGGAGATAAATTAGA ATG	
pYFP-RFP 10	TGTAAGCGACGAGGGAATTAGGGAGGTAATTAGA ATG	
pYFP-RFP 11	TGTCAGCGACGAGGGAATTAAGGAGGTAATTAGA ATG	
pOLS-WT	pLIC-KM with wilt type <i>ols</i> as the target gene	This report
pLD-ACP1	pLIC-KM with LD and ACP1 as target gene	This report
pKS-AT-KR	pLIC-KM with KS, AT, KR as target gene	This report
pLD-ACP1-KS-AT-KR	pLIC-KM with LD, ACP1, KS, AT, KR as target gene	This report
pACP2-ST-TE	pLIC-KM with ACP2, ST, TE as target gene	This report
pACP2-ST-TE-His	pACP2-ST-TE including a N-terminal 6X-histidine tag	This report
pACP2-ST-TE-Opt	pLIC-KM with codon optimized ACP2, ST, TE as target gene	This report
pOLS-Opt	pLIC-KM with <i>ols</i> gene with codon optimized ACP2-ST-TE domains as target gene	This report

Table 4.2 Oligonucleotides used. Restriction sites are underlined**Construction of pLIC-KM**

LICKm Fw CCACCATCATTAGGATGGTGGTGATGATAATGAGCCATATCAACGGGAAACGTCGAG
 LIC_Km Fw2b GTTCAATTGAAATAGGAGGAAACTAGTCATCATCACCACCATCATTAGGATGGTGGTGAT
 LICKm Rv CCCGGTACCTTTGGATCCTTAGAAAACTCATCGAGCATCAAATGAAACTGC
 MSB6_Fw GGGCAATTGAAATTGTTATCCGCTCACAATTCCACAC
 MSB6_Rv GGGCAATTGAAATTGTTATCCGCTCACAATTCCACAC

Construction of pRFP-KM and pRFP^{nf}KM

Red-SpeI-Rv CCCACTAGTCAGGAACAGGTGGTGGCGGC
 Red-MfeI-Fw CCCCAATTGAGGAGGAAAAAAAAATGGCGAGCAGTGAGGACATCATCAAG
 DMP2mut-Fw GGACTCCTCCCTGTAGGACGGCTCCTTC
 DMP2mut-Rv GAAGGAGCCGCTCTACAGGGAGGAGTCC
 CCCAGATCTAAATAGGAGGAGAATACATGGCGAGCAGTGAGGACATCATC
 DM1-Fw2-BglII
 DM1-Rv-EcoRI CCGAATTCCAGGAACAGGTGGTGGCGGC

Construction of pLIC-GFP, pLIC-RFP, pLIC-YFP

YFP_Fwd1 CCACCATCATTAGGATGGTGGTGATGATAATGGTGAGCAAGGGCGAGGAGCTG
 LIC_Fwd2 GTTAGATCTAAATAGGAGGAGAATTCCATCATCACCACCATCATTAGGATGGTGGTGAT
 YFP-rev CCCGGTACCTTTGGATCCTTAGTACAGCTCGTCCATGCCGAGAG
 MSB6_Fw2 CCAAAGGTACCTCTAGAGTCGACCTG
 MSB6_Rv2 GGGAGATCTGAAATGTTATCCGCTCACAATTCCACAC
 RFP_Fwd1 CCACCATCATTAGGATGGTGGTGATGATAATGGCGAGCAGTGAGGACATCATCAAG
 LIC_Fwd2 GTTAGATCTAAATAGGAGGAGAATTCCATCATCACCACCATCATTAGGATGGTGGTGAT
 RFP-rev CCCGGTACCTTTGGATCCTTACAGGAACAGGTGGTGGCGGC
 GFP_Fwd1 CCACCATCATTAGGATGGTGGTGATGATAATGAGTAAAGGAGAAGAAGCTTTTCACTGGAGTTG
 LIC_Fwd2 GTTAGATCTAAATAGGAGGAGAATTCCATCATCACCACCATCATTAGGATGGTGGTGAT
 GFP-rev CCCGGTACCTTTGGATCCTTATTGTAGAGCTCATCCATGCCATGTGTAATCC

Construction of RBS library

YFP-BglII CACGAGATCTAGGAGGAAAAAAAAATGGTGAGCAAGGGCGAGGAGCTG
 YFP-EcoRI CACGAATTCCTTGTACAGCTCGTCCATGCCGAGAG
 YFP- Degen CACGAGATCTTGTNAGCGNCGAGGGAATTANGNAGNTAATTAGAATGGTGAGCAAGGGCGAGGAGCTG
 YFP-EcoRI CACGAATTCCTTGTACAGCTCGTCCATGCCGAGAG

Construction of ols clones

ols-Fw1 CCCCAATTGAGGAGGAAAAAAAAATGGTGGTCAATTTGCAAATTCGTCGATCTGCTC
 ols-ACP0-Rv CCCACTAGTGCCCTGAATTTTCGGCACCTGTGG
 ols-ACP1-Fw CCCCAATTGAGGAGGAAAAAAAAATGATCTGCGGCGGTCTGTGGCGCAAAT
 ols-Rv3 CCCACTAGTTTGTGTTTTGGGTACAGGGGTCTGGAGTTG
 ols-KS-Fw CCCCAATTGAGGAGGAAAAAAAAATGAAAGAAATTGCCGTGGTGGTCTCAGTTGTC
 ols-KR-Rv CCCACTAGTATCCAGTAGATAGTCAGCCCGCTGAGAAG
 ols-ACP1-Fw-His CCCCAATTGAGGAGGAAAAAAAAATGCATCATCACCACCATCATTATCTGCGGCGGTCTGTGGCGCAAAT
 CodOp-SpeI-Fw CCTACTAGTTACCTGCGCCGCTCTGTGCCCCAAA
 CodOp-SpeI-Rv CCTACTAGTTTGGGTTTTTCGGAACCGGCGTC

Construction of pLIC-KM-PPT₇₀₀₂

PPT-7002-Fw CAATCAATTGAGGAGGAAAAAAAAATGGCAGTGTCGGTGGAATATTGGTTAATTTCC
 PPT-7002-Rv CACTACTAGTAGCTGCCCCATCATATAGCATCCATTCC

4.5 References

1. **Mendez-Perez D, Gunasekaran S, Orlor VJ, Pfleger BF.** 2012. A translation-coupling DNA cassette for monitoring protein translation in *Escherichia coli*. *Metabolic engineering* **14**:298–305.
2. **Keasling JD.** 2008. Synthetic biology for synthetic chemistry. *ACS chemical biology* **3**:64–76.
3. **Keasling JD.** 2010. Manufacturing molecules through metabolic engineering. *Science (New York, N.Y.)* **330**:1355–8.
4. **Pfeifer BA.** 2001. Biosynthesis of Polyketides in Heterologous Hosts **65**:106–118.
5. **Tyo KE, Alper HS, Stephanopoulos GN.** 2007. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends in biotechnology* **25**:132–7.
6. **Boyle PM, Silver PA.** 2012. Parts plus pipes: synthetic biology approaches to metabolic engineering. *Metabolic engineering* **14**:223–32.
7. **Keasling JD.** 2012. Synthetic biology and the development of tools for metabolic engineering. *Metabolic engineering* **14**:189–95.
8. **Michener JK, Thodey K, Liang JC, Smolke CD.** 2012. Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metabolic engineering* **14**:212–22.
9. **Angov E, Hillier CJ, Kincaid RL, Lyon JA.** 2008. Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS one* **3**:e2189.
10. **Carlson R.** 2009. The changing economics of DNA synthesis. *Nature biotechnology* **27**:1091–4.
11. **Ivey-Hoyle M, Steege DA.** 1989. Translation of phage ϕ 1 gene VII occurs from an inherently defective initiation site made functional by coupling. *Journal of molecular biology* **208**:233–44.
12. **Lindahl L, Zengel JM.** 1986. Ribosomal genes in *Escherichia coli*. *Annual review of genetics* **20**:297–326.
13. **Lindahl L, Archer RH, McCormick JR, Freedman LP, Zengel JM.** 1989. Translational coupling of the two proximal genes in the S10 ribosomal protein operon of *Escherichia coli*. *Journal of bacteriology* **171**:2639–45.

14. **Aksoy S, Squires CL, Squires C.** 1984. Translational coupling of the *trpB* and *trpA* genes in the *Escherichia coli* tryptophan operon. *Journal of bacteriology* **157**:363–7.
15. **Oppenheim DS, Yanofsky C.** 1980. Translational coupling during expression of the tryptophan operon of *Escherichia coli*. *Genetics* **95**:785–95.
16. **Schümperli D, McKenney K, Sobieski DA, Rosenberg M.** 1982. Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon. *Cell* **30**:865–71.
17. **Takyar S, Hickerson RP, Noller HF.** 2005. mRNA helicase activity of the ribosome. *Cell* **120**:49–58.
18. **Amann E, Ochs B, Abel KJ.** 1988. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**:301–15.
19. **Li SC, Squires CL, Squires C.** 1984. Antitermination of *E. coli* rRNA transcription is caused by a control region segment containing lambda *nut*-like sequences. *Cell* **38**:851–60.
20. **Aslanidis C, De Jong PJ.** 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic acids research* **18**:6069–74.
21. **Pfleger BF, Fawzi NJ, Keasling JD.** 2005. Optimization of DsRed production in *Escherichia coli*: effect of ribosome binding site sequestration on translation efficiency. *Biotechnology and bioengineering* **92**:553–8.
22. **Xu Y, Alvey RM, Byrne PO, Graham JE, Shen G, Bryant DA.** 2011. Expression of Genes in Cyanobacteria: Adaptation of endogenous plasmids as platforms for high-level gene expression in *Synechococcus* sp. PCC 7002. *Photosynthesis Research Protocols* **684**:273–293.
23. **Salis HM, Mirsky E a, Voigt C a.** 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nature biotechnology* **27**:946–50.

Chapter 5: *E. coli* and *Synechococcus sp.* PCC 7002 engineering

*Portions of this chapter were published in *Metabolic Engineering* (1)

5.1 Introduction

Based on the information obtained from previous chapters, the long term goal of this research is to engineer a host for olefin production. As mentioned before, metabolic engineering can be used to produce from simple and inexpensive starting materials a large number of chemicals that are currently derived from nonrenewable resources (2). This can be done by transferring product-specific enzymes or entire metabolic pathways from one organism (rare or genetically intractable) to those that can be readily engineered, or by rewiring native metabolic pathways in the native organism. Once a pathway has been identified for the production of the desired product (and the enzymes involved are expressed if a heterologous host is being used), several strategies can be used to increase production: elimination of competing pathways by gene knockout (or if the pathway is needed for cell growth attenuation of competing pathways can be achieved by replacing promoters or start codons), increasing the precursor availability by overexpression the enzymes involved in their biosynthesis, engineering promoter strength and gene or plasmid copy number to optimize the metabolic pathway and optimization of cofactor utilization (3). In this chapter we use some of these strategies in *Synechococcus sp.* PCC 7002 and *Escherichia coli* (*E. coli*) for the production of olefins using the OLS pathway.

5.2 Results

5.2.1 *E. coli* engineering

In this section we present the results for the expression of the OLS pathway in a heterologous host (*E. coli*). *E. coli* was used for numerous reasons, like its fast growth, simple fermentation, uncomplicated nutritional and sterility requirements, and extensive characterization. Moreover, since the substrates for hydrocarbon biosynthesis are hypothesized to be fatty acids, we would need an organism with a well understood pathway of fatty acid biosynthesis and regulation and *E. coli* is one of the most studied organisms in that regard. But in order to produce olefins in *E. coli* it is necessary to express the Ols enzyme first, and then make sure all the essential metabolites/cofactors required by Ols are available.

5.2.1.1 Ols expression in *E. coli*

The *ols* gene encodes for a large multi-domain protein (302 kDa) similar to a polyketide synthase (PKS) with eight enzymatic and structural domains (Figure 5.1a): loading (LD), acyl-carrier protein (ACP1), ketosynthase (KS), acyltransferase (AT), ketoreductase (KR), ACP2, sulfotransferase (ST) and thioesterase (TE). This multi-domain protein is larger than proteins normally found in *E. coli*. The largest native *E. coli* protein is the large subunit of glutamate synthase (166 kDa) (4). While many PKS have been successfully heterologously expressed in *E. coli*, success rates can be widely variable depending on the gene, its source, and the degree of sequence optimization. Initial attempts to express this gene in *E. coli* failed to produce olefins or any detectable Ols protein. We used the translational coupling system to confirm that the *ols* gene was not being actively translated at various conditions. The full *ols* open reading

frame (without a stop codon) was cloned using MfeI and SpeI sites into pLIC-KM, a construct that uses kanamycin resistance as the response signal, resulting in plasmid pOLS-WT. Cultures of *E. coli* DH10B harboring pOLS-WT were incubated at various temperatures (15°C, 25°C, 30°C, and 37°C) and IPTG concentrations in the presence of kanamycin. Under each condition, the cells failed to grow, suggesting that there was a problem translating the *ols* mRNA (data not shown).

To determine if size or specific domains were responsible for the lack of full length translation, the gene was divided in three sections and the DNA encoding each section was cloned into pLIC-KM (resulting in plasmids pLD-ACP1, pKS-AT-KR and pACP2-ST-TE). Cultures of *E. coli* harboring pLD-ACP1 and pKS-AT-KR were able to grow in the presence of kanamycin and generated Ni-NTA purified proteins of the expected sizes (Figure 5.1 a, b), confirming that the first two sections of *ols* could be actively translated. The two sections were subsequently cloned in tandem to generate pLD-ACP1-KS-AT-KR. *E. coli* DH10B was capable of growing on kanamycin (data not shown) confirming that the size of the protein was not a barrier to translation. Conversely, cultures of *E. coli* DH10B harboring pACP2-ST-TE (encoding the C-terminal third of Ols) were unable to grow in the presence of kanamycin and did not generate Ni-NTA purified proteins (Figure 5.1d).

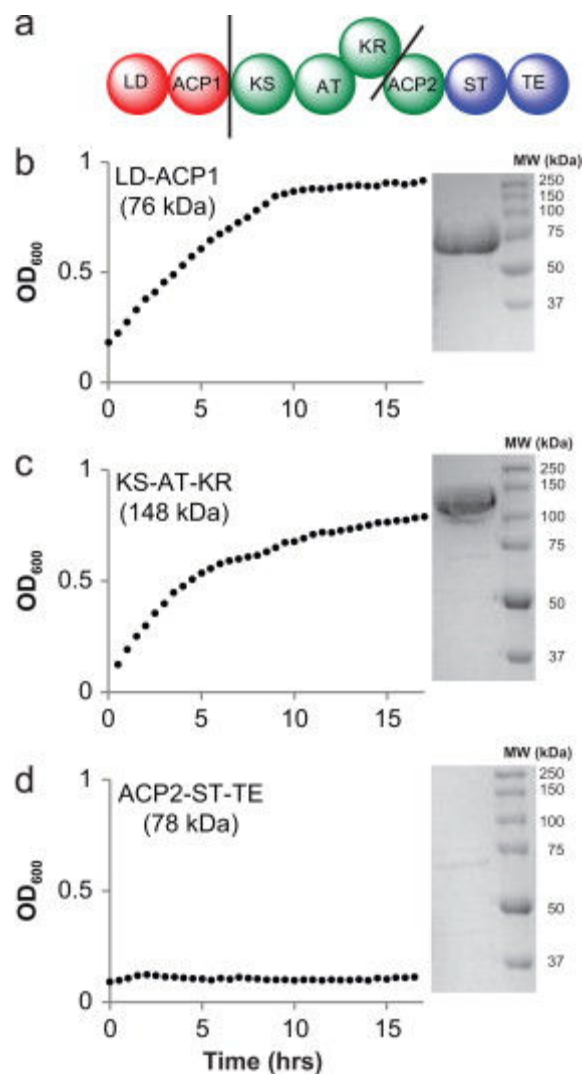


Figure 5.1 Use of translational coupling to assay for translation of a foreign gene: (a) domain organization of olefin synthase encoded by *ols*. Lines delineate sections that were subcloned individually into pLIC-KM which encodes kanamycin resistance as the response. (b)-(d) Cultures of *E. coli* harboring expression vectors for the listed Ols domains were grown in the presence of 50 $\mu\text{g}/\text{ml}$ kanamycin. When cells grew, a band of the appropriate size was observed in SDS-PAGE gels loaded with Ni-NTA purified protein. The results indicate that the C-terminal domains (ACP2-ST-TE) were not expressed.

The ACP2-ST-TE gene was subcloned with an in-frame N-terminal 6X-histidine tag (generating plasmid pACP2-ST-TE-His), expressed in *E. coli*, and the resulting protein was purified using Ni-NTA resin to determine if any region of the protein was expressed. An SDS PAGE analysis

resulted in a distinct band approximately 71 kDa in size (data not shown) that was smaller than the expected size of the full protein (79 kDa) indicating that translation was incomplete. Upon inspection, we identified several rare codons, including two consecutive arginine codons, in the C-terminal region of the protein. As such, we focused sequence optimization efforts on this section of *ols*.

It is well known that differences in codon usage between native and heterologous hosts can prevent successful heterologous expression of genes in *E. coli* (5). Therefore, the third section of the *ols* gene, encoding the ACP2, ST, and TE domains, was codon optimized for expression in *E. coli* (6) and cloned into pLIC-KM, yielding pACP2-ST-TE-Opt. *E. coli* harboring pACP2-ST-TE-Opt was capable of growing in the presence of kanamycin, indicating that the codon-optimized section was being actively translated (data not shown). The codon optimized C-terminal section was fused to the natural gene encoding the N-terminal domains (LD-ACP1-KS-AT-KR), resulting in a hybrid full-length *ols* gene (plasmid pOLS-Opt). Surprisingly, *E. coli* harboring pOLS-Opt was incapable of growing in the presence of 50 µg/ml kanamycin at 37°C when induced with a range of IPTG concentrations (data not shown). Considering that lower temperatures facilitate heterologous expression of PKS genes in *E. coli* (7), we decided to test for growth at 30°C and different IPTG concentrations. Whereas no growth was observed for *E. coli* harboring pOLS-WT (wild type gene) at these conditions, growth was observed for pOLS-Opt at 30°C at low IPTG concentrations and tailed off with increasing IPTG (Figure 5.2a). When *E. coli* harboring pOLS-Opt was cultured in the absence of kanamycin, a similar trend of decreasing growth rates with increasing IPTG concentration was observed. Combined with the lack of growth at 37°C, this data suggests that high rates of Ols expression are detrimental to *E. coli* growth. Conversely, at the optimal conditions, *E. coli* harboring pOLS-Opt generated a Ni-

NTA purifiable protein that migrated slower than a 250 kDa standard (highest in the ladder) using SDS-PAGE (Figure 5.2b).

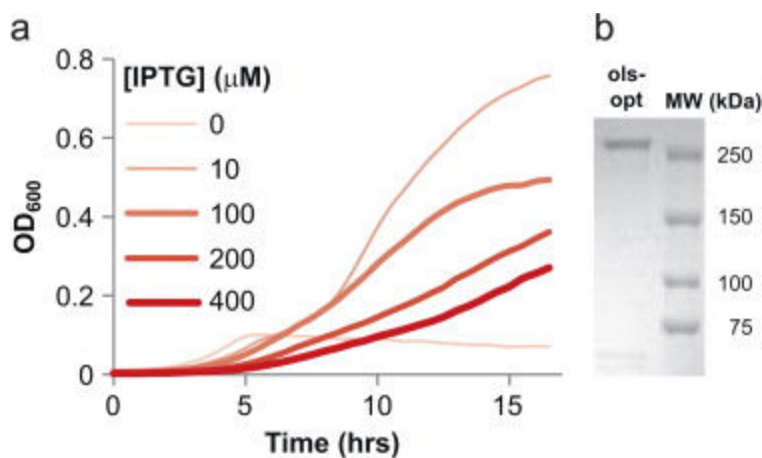


Figure 5.2. Use of translational coupling to optimize expression conditions of the *ols* gene: (a) growth curves for the hybrid full-length *ols* gene (pOLS-Opt) at 30°C, 50 μg/ml kanamycin and various IPTG concentrations and (b) SDS-PAGE gel showing a protein band produced by Ni-NTA purification of *E. coli* pOLS-Opt grown at the optimal temperature and inducer concentration.

5.2.1.2 Acyl carrier protein (ACP) phosphopantetheinylation

The loading domain (LD) in Ols is proposed to load a fatty acid substrate onto the adjacent ACP domain by thioesterification. However, the ACPs need to be modified by 4'-phosphopantetheinyl transferases (PPTs) from the inactive apo-enzymes to their active holo forms by transferring the 4'-phosphopantetheinyl (P-pant) moiety of coenzyme A to a conserved serine residue of the carrier protein (8) (Figure 5.3). When incorporated into an appropriate carrier protein, this 4'-phosphopantetheinyl arm has two main functions: the reactive thiol group of the phosphopantetheine acts as a covalent connection for the pathway intermediates; and second, the

length and flexibility of the moiety assists the relocation of intermediates between the spatially distinct modules of the complex (9–11).

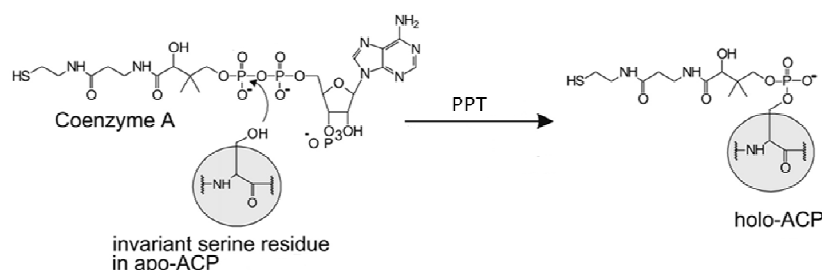


Figure 5.3 Activation of Acyl carrier protein (ACP) by phosphopantetheinyl transferases (PPT), Figure taken from (12)

Incomplete phosphopantetheinylation of ACP domains will reduce the catalytic efficiency of the PKS and could lead to a product-limiting bottleneck. Despite the requirement for PPTs in a wide range of important biosynthetic pathways, these enzymes have remained elusive due to their low sequence identity and lack of proximity to their respective biosynthetic clusters. This has prevented efforts to produce polyketides and nonribosomal peptide products in heterologous hosts systems like *E. coli*, largely due to the inability of *E. coli* PPTs to activate foreign substrates (13, 14).

It has been shown that phosphopantetheinylation modification of ACP domains in several PKS can be achieved by co-expressing the surfactin P-pant transferase (Sfp) from *Bacillus subtilis* in *E. coli* (15–17) due to its wide range of substrate specificity. Therefore, we decided to test if Sfp would be able to activate the apo-ACP domains from Ols. Various methods have been reported to assay for 4'phosphopantetheinylation including radio labeled CoA and liquid scintillation

(10), and HPLC separation and analysis of purified proteins (12). Also, it has been reported that PPTs are able to transfer small molecules of diverse structures from CoA to the conserved serine residue on the ACP domain, which provides an attractive method for one-step site-specific protein labeling with small molecules of diverse structures and functionalities such as biotin, fluorophores, sugars and peptides (18). Given that PPTs accept substrates other than thioesters, easily detectable analogs of CoA can be created and used in an *in vitro* reaction to test for the activation of the ACP by PPT. In this study we used the fluorescent dye N-(2-aminoethyl) maleimide (Bodipy) for the creation of the CoA reporter analog.

In order to test if Sfp was able to activate the ACP domains in Ols, its encoding gene was PCR amplified from genomic DNA and cloned onto a plasmid under control of a P_{trc} promoter with a C-terminal histidine tag (plasmid pLIC-KM-sfp). The protein was purified and after running on a SDS-PAGE gel a band of the right size (26 kDa) was observed. Ols protein extracts were then subjected to an *in vitro* phosphopantetheinylation reaction containing fluorescently labeled CoA and purified Sfp. The reaction was run on an SDS-PAGE gel and analyzed on a Typhoon imager. As can be seen in Figure 5.4, after incubation of Ols with Sfp activation of the ACP domains was observed (a strong band at 302 kDa was observed when Ols, Sfp, Bodipy and CoA were present in the reaction). No phosphopantetheinylation was observed when CoA or Bodipy were not present.

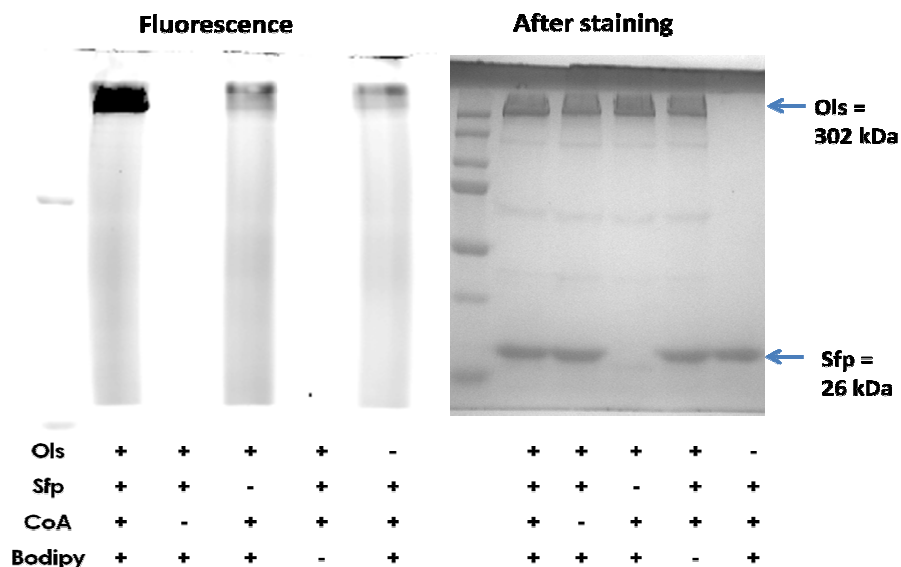


Figure 5.4 a) SDS-PAGE of Ols containing fluorescently labeled ACP domains scanned on a Typhoon imager. b) Image of the same gel after staining for total protein with cromassie dye reagent to confirm the presence of the proteins. Imager parameters were as follows: 488 nm laser, 520 nm bandpass 40 emission filter, 470 V PMT.

We then decided to test for activation *in vivo* by co-expressing Ols and Sfp in *E. coli*. In this case, *E. coli* cells were transformed using a plasmid containing the *ols* gene under control of an IPTG inducible promoter and a plasmid containing *sfp* under control of an arabinose inducible promoter; cells were grown at 30°C (optimized conditions for Ols expression found in previous section) and the Ols protein was purified. Then, the purified Ols was used for an *in vitro* reaction using purified Sfp. Therefore, if the ACP domains in Ols are not phosphopantetheinylated *in vivo* they should be fluorescently labeled *in vitro*. Conversely if Ols is activated, phosphopantetheinylation *in vivo* blocks fluorescent labeling *in vitro* and no band should be observed on a gel. From Figure 5.5, Sfp was able to activate Ols *in vivo* as no *in vitro* fluorescent labeling was observed.

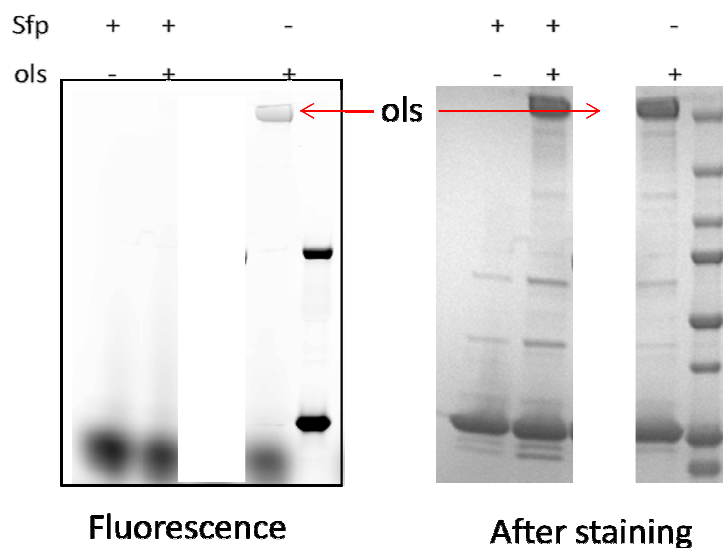


Figure 5.5 Phosphopantetheinylation of Ols by Sfp *in vivo*. *E. coli* cultures expressing Ols and with (+) or without (-) expressing Sfp were grown and the purified Ols was used for an *in vitro* labeling reaction. Ols was only fluorescently labeled in the cultures that didn't express Sfp *in vivo* (-Sfp), showing that Ols is activated (phosphopantetheinylation *in vivo* blocks fluorescent labeling *in vitro* and no fluorescence is observed on a gel).

As mentioned before, there are two ACP domains in the Ols protein that need to be phosphopantetheinylated by Sfp. The results from the previous experiment don't necessarily show that both ACPs are activated in Ols so we decided to test for phosphopantetheinylation of individual domains. The LD-ACP1 and ACP2-ST-TE domains were cloned and expressed in *E. coli* with a C-terminal 6x-histidine tag and protein purification was performed. When cloned into the translational coupling system using antibiotic resistance as the response signal we observed growth in the presence of antibiotics for the ACP2-ST-TE construct, suggesting that these domains were translated; however, we were unable to obtain purified protein so we couldn't test for activation of this second ACP domain. For the LD-ACP1 construct the protein was purified and subjected to *in vitro* activation and as shown in Figure 5.6 Sfp was able to activate this ACP domain.

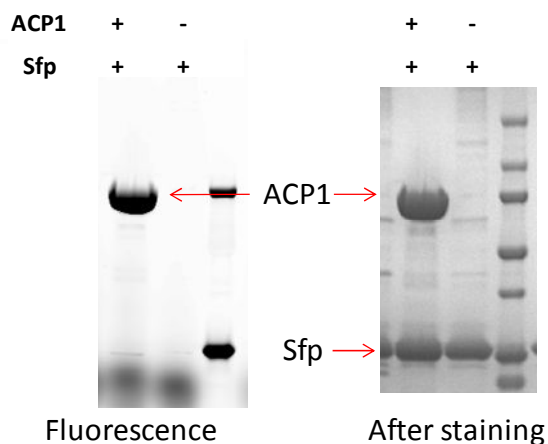


Figure 5.6 Phosphopantetheinylation of ACP1 by Sfp *in vitro*. The LD-ACP1 domain were purified from *E. coli* cultures and tested *in vitro* using purified Sfp. Fluorescence showed that the ACP1 domain is activated by Sfp.

5.2.1.3 Identification of a phosphopantetheinyl transferase (PPT) from *Synechococcus sp.*

PCC 7002

Although Sfp has broad substrate specificity, its activity might not be optimal for metabolic engineering applications using the Ols pathway so it would be of great interest to identify the native PPT from 7002 that is responsible for the activation of the ACP domains in Ols.

PPTs have been classified into two types: a) the AcpS type accepts ACPs of fatty acid synthases and some ACPs of type II polyketide synthase as substrates, b) the Sfp type exhibits extraordinary broad substrate specificity (12). Members of the Sfp family are approximately 230 amino acids in length whereas the AcpS family consists of approximately 115 amino acids. Bioinformatic analysis has revealed the absence of AcpS enzymes in almost all cyanobacterial genomes available (11). In the case of *Synechococcus sp.* PCC 7002, analysis of its genome revealed the presence of a singular Sfp-like PPT (Accession no. NC_003488) of 227 amino acids (11), but it hasn't been characterized.

In order to test if the Sfp-like PPT identified in PCC 7002 (referred as PPT₇₀₀₂) was responsible for the activation of the ACP domains in Ols, its encoding gene was PCR amplified from genomic DNA and cloned onto a plasmid under control of a P_{trc} promoter with a C-terminal histidine tag (plasmid pLIC-KM-PPT₇₀₀₂). The protein was purified from *E. coli* cultures and after running on a SDS-PAGE gel a band of the right size (26kDa) was observed. Ols protein extracts were then subjected to an *in vitro* phosphopantetheinylation reaction containing fluorescently labeled CoA and purified the PPT₇₀₀₂. The reaction was run on an SDS-PAGE gel and analyzed on a Typhoon imager. As can be seen in Figure 5.4, after incubation of Ols with PPT₇₀₀₂ activation of the ACP domains was observed. No phosphopantetheinylation was observed when CoA or Bodipy were not present.

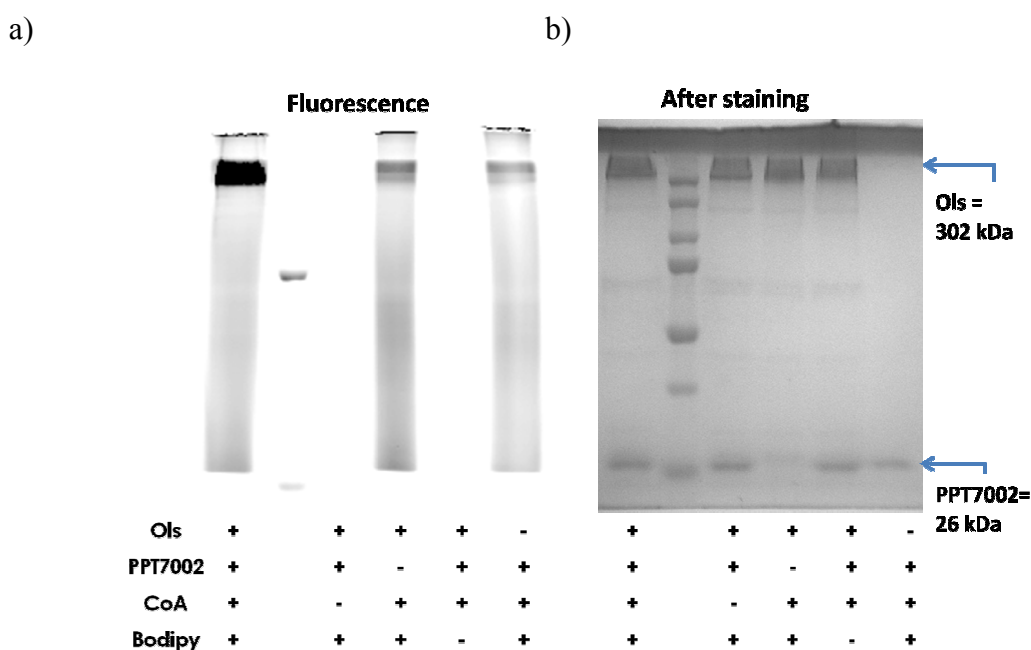


Figure 5.7 a) SDS-PAGE of Ols containing fluorescently labeled ACP domains scanned on a Typhoon imager. **b)** Image of the same gel after staining for total protein with cromassie dye reagent. Imager parameters were as follows: 488 nm laser, 520 nm bandpass 40 emission filter, 470 V PMT.

5.2.1.4 Increasing substrate availability in *E. coli*: acyl-ACP synthetase expression

As demonstrated in Chapters 2 and 3, the precursors for alkene biosynthesis in PCC 7002 are C18:0 and C18:1(Δ 13) acyl-ACP fatty acids. Therefore, in order to produce alkenes in *E. coli* using Ols, these acyl-ACP intermediates need to be either endogenously produced or supplemented to the cultures. Since it has been shown that *E. coli* produces mainly C16:0, C16:1, C18:1(Δ 9) and C14:0 fatty acids (19) (and only trace amounts of C18:0 fatty acids), at first we tested if it was possible to produce alkenes by supplementing C18 fatty acids to the *E. coli* cultures. Although fatty acids added to cultures of *E. coli* can be readily transported into the cell, they are converted to CoA thioesters and degraded to acetyl-CoA by the β -oxidation system (20). In fact, it has been shown that labeled C14 fatty acids are incorporated into glycerophospholipids only as C14 fatty acids, not as a sixteen or eighteen carbon fatty acid (21), indicating that *E. coli* lacks the means to efficiently convert either free fatty acids or acyl-CoA thioesters to the acyl-ACP thioester substrates that would be required by Ols to produce alkenes. Therefore, the expression of an acyl-ACP synthetase (Aas) that can activate the C18 free fatty acids to their acyl-ACP form is required. We decided to use the Aas from the cyanobacterium *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942 since it has been shown that they display broad substrate specificity, accepting fatty acids with chain lengths between C12 and C18 (22). The genes encoding for these Aas were amplified from genomic DNA and cloned into a plasmid under control of an arabinose inducible promoter (pBAD33). After transformation, octadecanoic acid was fed to the cultures and the fatty acid content from cell pellets was analyzed after ~24 hr. As can be seen in Figure 5.8, addition of octadecanoic acid and expression of the Aas resulted in a significant increase of intercellular levels of the C18 fatty acids. These plasmids were then cotransformed with a plasmid containing the *ols* gene under control of an IPTG inducible

promoter in a strain containing a chromosomal copy of the *sfp* gene (BAP1 cells) and cultures were grown at 30°C and tested for alkene production. However, no alkenes were detected.

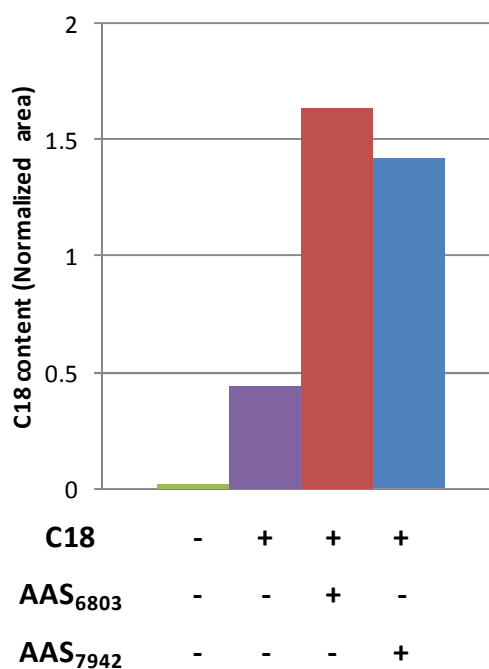


Figure 5.8 Octadecanoic content in cell pellets after addition of octadecanoic acid (C18, ~0.2mM) and expression of an acyl-ACP synthetase from *Synechocystis sp.* PCC 6803 (AAS₆₈₀₃) or *Synechococcus elongatus* PCC 7942 (AAS₇₉₄₂). Plasmids were transformed into BAP1 cells (containing a chromosomal copy of the *sfp* gene) and cultures were grown at 30°C. The area of the peak corresponding to the C18 fatty acid was normalized to the area of an internal standard (C17).

5.2.1.5 Expression of *cysDNC* genes

Based on the known mechanism of sulfotransferase (ST) enzyme function, the ST domain in Ols is predicted to bind to a sulfate donor and transfer a sulfonate moiety to the β -hydroxyl group of the intermediate terminal carboxylic acid that results from the previous ketoreduction tethered to ACP. This activation is required for the subsequent dehydration and decarboxylation reactions

where the thioesterase (TE) is presumed to release the intermediate from the ACP. The donor of sulfate in most sulfation reactions, including the cyanobacterium *Synechococcus 6301* (23), is 3'-phosphoadenosine-5'-phosphosulfate (PAPS) (24). While *E. coli* is capable of synthesizing PAPS, olefin biosynthesis may require elevated levels of this compound. Therefore, an alternative to increase the production of PAPS can be to increase the flux through the sulfate assimilatory pathway. Three genes are known to be involved in the conversion of sulfate from the media to PAPS (25) (Figure 5.9): *cysD*, *cysN* (sulfate adenylyltransferase) and *cysC* (adenylylsulfate kinase). These three genes were cloned into a plasmid containing the Aas (from *Synechocystis sp.* PCC 6803 or *Synechococcus elongatus* PCC 7942) and Sfp; all the genes were under control of an arabinose inducible promoter. This plasmid was used to transform *E. coli* cells (Δ fadD Δ araBAD to prevent degradation of the fatty acids) expressing *ols* and grown at 30°C. After lipid extraction, no alkenes were detected in the cultures.

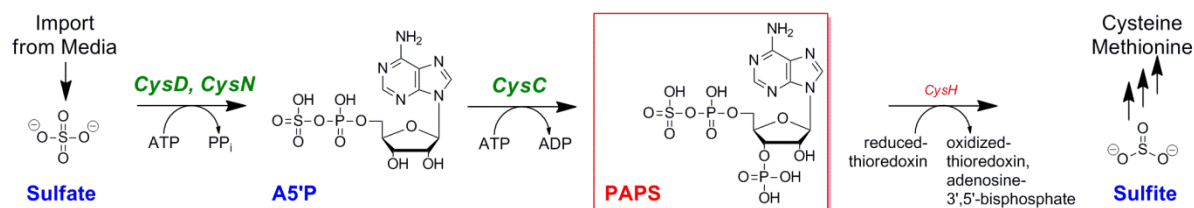


Figure 5.9 3'-phosphoadenosine-5'-phosphosulfate (PAPS) biosynthetic pathway in *E. coli*. To increase PAPS production, three genes (*cysD*, *cysN*, *cysC*, green) were over expressed

5.2.1.6 Future work and conclusions

Although we were able to produce the Ols protein in *E. coli*, we didn't detect production of the alkenes under any of the conditions tested. Functional expression of polyketide synthases (like

Ols) in *E. coli* has been proven to be challenging owing to two key issues: production of active protein and production of substrates.

Problems in phosphopantethenylation of ACP domains and protein folding (26) are the major causes for inactivity of polyketide synthases in *E. coli*. As demonstrated in section 5.2.1.2, the PPT from *Bacillus subtilis* (Sfp) was able to activate Ols both *in vitro* and *in vivo*. However, in the experiments presented in this section we only confirmed the activation of the first ACP domain of Ols due to problems in protein expression of the second ACP domain; since Sfp was able to activate the first ACP domain it seems likely that it would also activate the second one. In order to confirm this, it is necessary to optimize conditions for protein production of this individual ACP domain and test for activation *in vivo*. Although Sfp displays broad specificity towards a variety of heterologous ACP domains, incomplete phosphopantetheinylation of the ACP domains might cause a reduction in the catalytic efficiency of Ols. Therefore, once it is confirmed that the two ACP domains can be activated by Sfp it would be important to determine to what extent these domains are activated by Sfp. A quantitative fluorescent assay to measure the extent of phosphopantetheinylation of ACP domains in *E. coli* (27) can be used for this purpose. If it is determined that the extent of phosphopantetheinylation by Sfp is low, the native PTT from *Synechococcus sp.* PCC 7002 identified in section 5.2.1.3 can be used instead.

As presented in section 5.2.1.1, we optimized the conditions for production of soluble Ols protein in *E. coli*; but soluble proteins are not necessarily biologically active and properly folded. Several methods have been reported to improve folding of polyketide synthases in *E. coli* (28–30), including the fusion of the protein with a more soluble partner (like fusion of maltose binding protein), slowing down translation by lowering temperature, codon harmonization (31),

coexpression of chaperons (32) and even directed evolution can be used to modify the intrinsic sequence-dependent folding characteristics of the protein.

As mentioned before, functional expression of polyketide synthases also depends on the availability of the substrates and cofactors used by the enzyme. In the case of Ols, mainly four components are required: C18 fatty acid acyl-ACPs, PAPS, malonyl-CoA and NADPH. The two strategies presented in sections 5.2.1.4 and 5.2.1.5 to increase C18 acyl-ACPs and PAPS didn't result in production of alkenes; and native *E. coli* metabolism produces malonyl-CoA at levels sufficient to promote polyketide biosynthesis (26). Since NADPH is an essential component for the anabolism inside the cell, efficient regeneration of this cofactor might be a limiting factor that constrains the production of alkenes using the Ols pathway. In order to increase the availability of NADPH, three main approaches can be followed (33): modulation of the pentose phosphate pathway, amplification of the transhydrogenases systems and replacement of native glyceraldehyde-3-phosphate dehydrogenase (GAPDH) with NADP-dependent GAPDH from *Clostridium acetobutylicum* and introduction of NADH kinase catalyzing direct phosphorylation of NADH to NADPH from *Saccharomyces cerevisiae*.

Finally, there are organisms that are better suited for polyketide expression than *E. coli* and they might be better candidates for heterologous alkene production. For example, *Streptomyces coelicolor* naturally produces at least two known polyketides, there is an extensive knowledge about its biology, there are expression systems that can be used to engineer it and its polyketide natural products are exclusively synthesized from malonyl-CoA; this makes *Streptomyces coelicolor* an ideal host for the heterologous production of polyketides (34).

5.2.2 *Synechococcus* sp. PCC 7002 engineering

5.2.2.1 Construction of a chimeric olefin synthase

The modularity of type I polyketide synthases (PKS) like Ols make them attractive targets for protein engineering because desired changes in product structure can be translated back to the sequence of PKS domains (35–37). In principle, the native loading domain (LD) of the PCC 7002 olefin synthase can be replaced with loading modules that have specificity for other substrates and generate olefins with different structures (Figure 5.10). For example, if the loading module of Debs1 is used the chimaeric enzyme would yield 1-butene, fusion of a fatty acid loading module found in MycA (38) would yield an olefin with diesel-like properties and fusion of the soraphen loading domain which uses benzoic acid (39) would yield styrene.

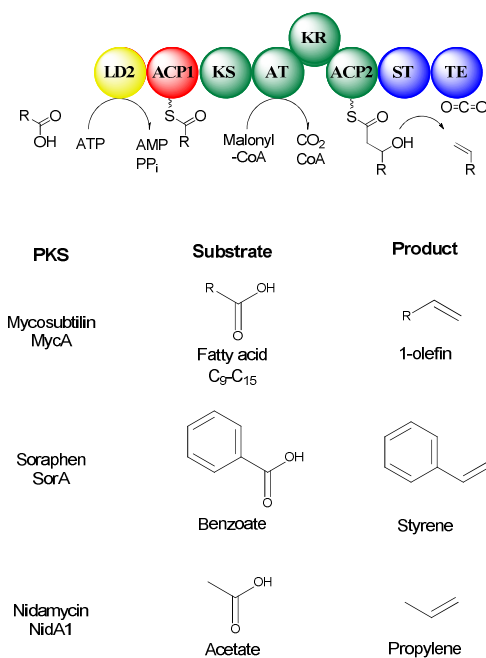


Figure 5.10 Potential olefins that can be produced using Ols by using a different loading module (LD2). The different loading modules, starter units (substrate), and putative olefin products from three known PKS are shown.

We constructed a chimeric Ols by replacing the native loading domain and ACP1 domain of the PCC 7002 olefin synthase with the well studied loading domain from the mycosubtilin pathway in *Bacillus subtilis spizizenii* that is known to load shorter chain saturated and unsaturated fatty acids onto an ACP domain (38). A mutant strain was constructed replacing the LD-ACP1 domains in PCC 7002's chromosome and adding a strong promoter in front of it (P_{psbA}). The fusion point was chosen to be around the ACP domain because of its strongly conserved secondary structure (36, 40) and the crossover point was made at the domain border of the ACP1 and KS domains. The mutant was constructed, segregated and verified by PCR and since this loading domain is known to have preference for C10-C16 free fatty acids, C16 and C14 fatty acids were fed to the cultures. The lipid profile was analyzed for the presence of 1-heptadecene and 1-pentadecene but none of these alkenes were detected.

5.2.2.2 *desE* overexpression

As mentioned in Chapter 3, the unsaturated fatty acids used for the biosynthesis of 1, 14 nonadecadiene don't accumulate in cyanobacterial cells, they are not incorporated into the cell membranes and they are exclusively used for the biosynthesis of the diunsaturated alkene. This suggests that their levels might be limiting and increasing their production might result in an increase of alkene production. To test this hypothesis, the expression of *desE* was increased by replacing the upstream region of *desE* with the *cpcBA* promoter from *Synechocystis sp.* PCC 6803 (41) in the wild type or *ols* upregulated strain. As can be seen in Figure 5.11, over expression of *desE* didn't result in an increase of the alkene total content. Interestingly, the hydrocarbon profile of the *desE* upregulated strains (*desE*-up) was completely different from that

of the wild type strain. Whereas the C19:2 hydrocarbon in the wild type strain represents about 35% of the total hydrocarbons, in desE-up it accounts for 100% of the total hydrocarbon content. The fact that the levels of the C18:0 fatty acid remain relatively constant for these strains (not shown), suggests that Ols might have a preference for the unsaturated fatty acid (C18:1(Δ 13)).

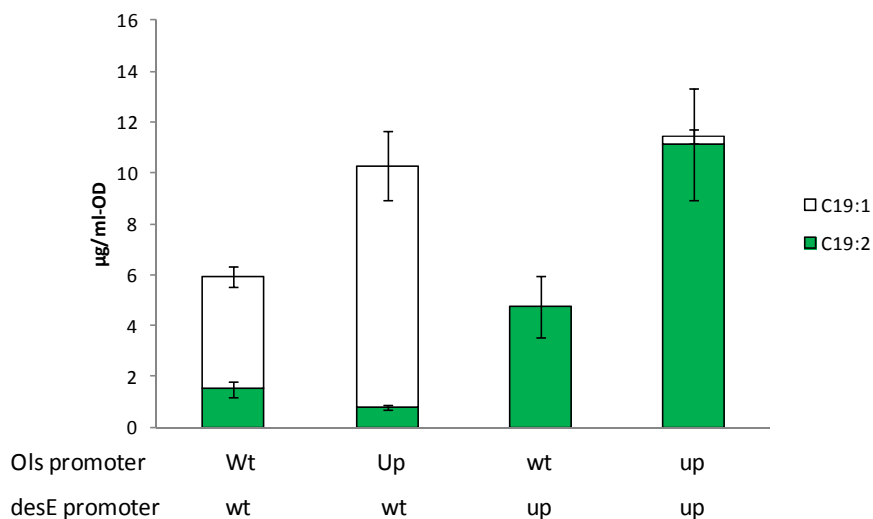


Figure 5.11 Olefin production at 38°C, bubbling air

5.2.2.3 Thioesterase (TE) expression

As mentioned in previous sections, fatty acids are the precursors for olefin biosynthesis; therefore, we wanted to test if increasing their production would result in an increase of alkene biosynthesis in PCC 7002. In bacteria, fatty acid synthesis is accomplished by a multienzyme system (type II system) starting from acetyl-CoA. Acetyl-CoA carboxylase (ACC) catalyzes the first step through the formation of malonyl-CoA from acetyl-CoA; this carboxylation of acetyl-CoA is thought to be a key rate-limiting step in fatty acid biosynthesis in *E. coli* (42). Fatty acyl-ACPs are then synthesized from malonyl-CoA and fatty acids are made via an iterative reduction

cycle that operates on acyl carrier protein thioesterases. In each iteration, two carbons (acetate) are added from malonyl-ACP to a growing acyl-chain and the resulting β -keto group is reduced to a saturated methylene. The process continues until long-chain acyl-ACPs are incorporated into phospholipids by acyltransferases or converted to other metabolites (e.g. olefins) (43). It has been shown that *E. coli*'s fatty acid biosynthesis is feedback-inhibited by long-chain fatty acyl-ACPs, and that high-level expression of acyl- ACP thioesterases, enzymes that hydrolyze the acyl-ACP thioester to liberate the free fatty acid, can relieve this inhibitory mechanism (42, 44). Furthermore, the composition of fatty acids synthesized by *E. coli* has been modified by heterologous expression of plant thioesterases (45, 46).

Since the step catalyzed by ACC might also be a rate-limiting step in fatty acid biosynthesis in PCC 7002, an artificial operon $P_{cpcBA} accB accC accD accA$ using the ACC genes from *E. coli* or *Synechocystis sp.* 6803 and the *cpcBA* promoter from *Synechocystis sp.* 6803 (41) was introduced into the native pAQ1 plasmid of PCC 7002. After analyzing the fatty acid content of this mutant strains no significant difference was observed when compared to the wild type. No increase in fatty acid production after expression of ACC genes has been reported in other cyanobacterial strains (47), suggesting that ACC might not be a rate-limiting step in cyanobacteria. But in order to confirm this hypothesis, more experimental work would be necessary to show that the ACC genes used for this study are active in cyanobacteria.

The accumulation of acyl-ACPs feedback inhibits multiple enzymes in fatty acid biosynthesis, coordinating lipid production to the growth rate. Therefore, another metabolic engineering strategy for overproducing free fatty acids is the expression of thioesterases, enzymes that hydrolyze the acyl-ACP thioester to liberate the free fatty acid, to decouple fatty acid biosynthesis from normal modes of regulation. Moreover, expression of thioesterases that have

preference for substrates of specific length can result in an increase of production of a fatty acid of the desired length. In the case of olefin production in PCC 7002 it would be necessary to identify a thioesterase that has preference for C18 fatty acids.

Three thioesterases that have been shown to have broad substrate specificity were inserted into PCC 7002's genome replacing the acyl-ACP synthetase (*Aas*) identified in chapter 2 under control of an IPTG inducible promoter. These thioesterases are the acyl-ACP thioesterase I (encoded by *tesA* gene) from *E. coli*, a plant acyl-ACP thioesterase from *Umbellularia californica* (BTE) (48) and an acyl-ACP thioesterase from *Geobacillus sp.* Y412MC10 (Geo) (19). The fatty acid composition of the mutant strains containing the thioesterases were analyzed by GC-MS and the results are presented in Figure 5.12. For the wild type strain the C18 content represents about 1% of the total fatty acid content, there is a slight increase after expression of *tesA* (5%) and BTE (3%) but the highest increase is observed in the strain harboring the *Geobacillus* thioesterase where the C18 content represents about 20% of the total fatty acid content. The highest increase in total fatty acid production also corresponds to the strain with the *Geobacillus* thioesterase. No significant difference was observed when alkene composition was analyzed (Figure 5.13a) for cultures grown at 35°C bubbled with air. However, when the cultures were supplemented with carbon dioxide an increase in alkene production was observed (Figure 5.13b), suggesting that carbon dioxide delivery might be a limiting factor.

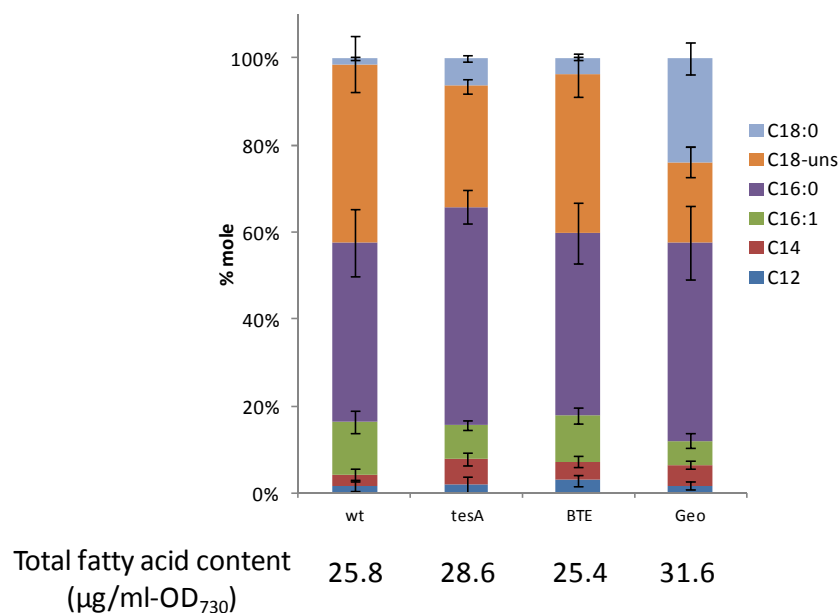


Figure 5.12 Fatty acid composition after expression of acyl-ACP thioesterases from *E. coli* (*tesA*), *Umbellularia californica* (BTE) and *Geobacillus sp.* Y412MC10 (Geo). Cultures were grown at 35°C with aeration and error bars represent standard deviations from three biological replicates. C18-uns represents the C18 unsaturated fatty acids (C18:1, C18:2 and C18:3).

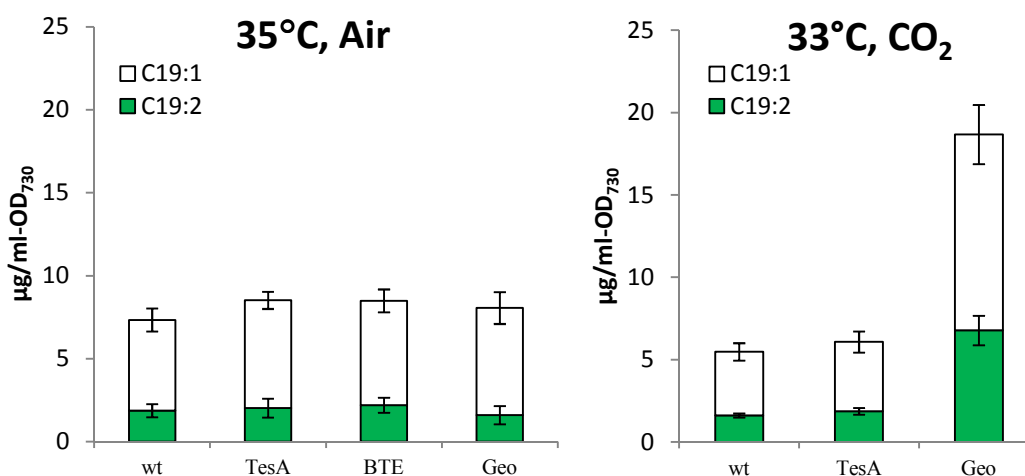


Figure 5.13 Alkene production in PCC 7002 after expression of thioesterases. Cultures were grown using air or CO₂ for bubbling. Error bars represent standard deviations from three biological replicates.

5.2.2.4 Future work and conclusions

Replacing the loading domain in Ols with the loading domain from the mycosubtilin pathway didn't result in the formation of new alkenes in PCC 7002. Although the fusion point was chosen to be at the linking region of the ACP domain, it is possible that the fusion caused structural changes that resulted in problems in solubility or activity of the resulting chimeric protein. A simple method to test if the enzyme is being produced in PCC 7002 would be to run an SDS-PAGE gel after protein extraction for both the soluble and the insoluble fractions using common tags. Other fusion points can be tested within strongly conserved regions of the ACP domain if a problem in solubility is detected. Also, it would be interesting to test if other loading domains that have preference for acyl-ACPs would result in the formation of alkenes, since Ols is believed to load C₁₈ acyl-ACP fatty acids whereas the loading domain from the mycosubtilin pathway loads free fatty acids.

Expression of thioesterases from *Geobacillus* sp. Y412MC10 and *E. coli* resulted in a slight increase in total fatty acid content in PCC7002 when compared to the wild type. Moderate increases in fatty acid production have also been observed when other cyanobacterial strains have been engineered for fatty acid overproduction using thioesterases (47, 49). Although fatty acid biosynthesis is conserved in bacteria, it is possible that its regulation is different in cyanobacteria. Most cyanobacterial genomes, including PCC 7002's, do not have homologs for enzymes involved in beta oxidation/degradation of fatty acids. Instead, free fatty acids released by membrane degradation are recycled for membrane synthesis via an acyl-ACP synthetase (22). Therefore, in order to improve production it is necessary to investigate fatty acid regulation in cyanobacteria and determine if acyl-ACP is the key regulatory signal for controlling fatty acid biosynthesis.

5.4 Methods

5.4.1 *E. coli* plasmid construction and oligonucleotides

Plasmid construction and maintenance was performed using *E. coli* DH10B (Invitrogen, Carlsbad, CA). Plasmids were prepared by alkaline lysis (Qiagen, Valencia, CA). Oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, Iowa). Cloning enzymes, including Phusion® DNA polymerase, restriction enzymes, and T4 DNA ligase were purchased from New England Biolabs (NEB, Ipswich, MA) and used according to the manufacturer's instructions. Plasmids and strains used in this paper are listed in Table 5.1 and the primers used are listed in Table 5.2. All cloning procedures were confirmed by DNA sequencing.

Table 5.1 *E. coli* plasmid/strain used in this chapter

Strain (S)/plasmid (P)	Relevant genotype/property	Reference
pOLS-WT (P)	pLIC-KM with wild type <i>ols</i> as the target gene	This report
pLD-ACP1 (P)	pLIC-KM with LD and ACP1 as target gene	This report
pKS-AT-KR (P)	pLIC-KM with KS, AT, KR as target gene	This report
pLD-ACP1-KS-AT-KR (P)	pLIC-KM with LD, ACP1, KS, AT, KR as target gene	This report
pACP2-ST-TE (P)	pLIC-KM with ACP2, ST, TE as target gene	This report
pACP2-ST-TE-His (P)	pACP2-ST-TE including a N-terminal 6X-histidine tag	This report
pACP2-ST-TE-Opt (P)	pLIC-KM with codon optimized ACP2, ST, TE as target gene	This report
pOLS-Opt (P)	pLIC-KM with <i>ols</i> gene with codon optimized ACP2-ST-TE domains as target gene	This report
BAP1 (S)	BL21 (DE3) with chromosomal copy of Sfp	(50)
pLIC-KM-PPT ₇₀₀₂	pLIC-KM with PTT from PCC 7002	This study
pAAS6803 (P)	pBAD33 containing Aas from <i>Synechocystis sp.</i> PCC 6803	This study
pAAS7942 (P)	pBAD33 containing Aas from <i>Synechococcus longates</i> PCC 7942	This study
pVO02 (P)	pLIC-KM with codon optimized <i>ols</i>	This study
pVO03 (P)	pAAS6803 with a copy of <i>sfp</i>	This study
pVO04 (P)	AAS7942 with a copy of <i>sfp</i>	This study
pVO05 (P)	AASvibrio with a copy of <i>sfp</i>	This study
pVO03-cysDNC (P)	pVO03 and <i>cysDNC</i> genes	This study
pVO04-cysDNC (P)	pVO04 and <i>cysDNC</i> genes	This study
pVO05-cysDNC (P)	pVO05 and <i>cysDNC</i> genes	This study

Ols-LD (S)	<i>Ols</i> loading domain replacement	This study
Ols-up (S)	$\Phi(P_{psbA-ols})$	This study
<i>desE</i> -up (S)	$\Phi(P_{cpcAB-desE})$	This study
<i>Ols-desE</i> -up (S)	$\Phi(P_{psbA-ols}), \Phi(P_{cpcAB-desE})$	This study
pNACC (P)	pAQ1 with <i>E. coli acc</i> operon inserted	This study
pACC6803 (P)	pAQ1 with PCC 6803 <i>acc</i> operon inserted	This study
tesA (S)	$\Delta as :: aadA-tesA$	This study
BTE (S)	$\Delta as :: aadA-BTE$	This study
GeoTE (S)	$\Delta as :: aadA-GeoTE$	This study

Table 5.2 Oligonucleotides used. Restriction sites are underlined

Construction of *ols* clones

ols-Fw1	CCCCAATTGAGGAGGAAAAAAAAAATGGTTGGTCAATTTGCAAATTCGTCGATCTGCTC
ols-ACP0-Rv	CCC <u>ACTAGT</u> GCCTGAATTTTCGGCACCTGTGG
ols-ACP1-Fw	CCCCAATTGAGGAGGAAAAAAAAAATGTATCTGCGGGCGTCTGTGGCGCAAAT
ols-Rv3	CCC <u>ACTAGT</u> TTGTGTTTTGGGTACAGGGGTCTGGAGTTG
ols-KS-Fw	CCCCAATTGAGGAGGAAAAAAAAAATGAAAGAAATTCGCCGTGGTGGTCTCAGTTGTC
ols-KR-Rv	CCC <u>ACTAGT</u> ATCCAGTAGATAGTCAGCCCGCTGAGAAG
ols-ACP1-Fw-His	CCCCAATTGAGGAGGAAAAAAAAAATGCATCATCACCACCATCATTATCTGCGGGCGTCTGTGGCGCAAAT
CodOp-SpeI-Fw	CCT <u>ACTAGT</u> TACCTGCGCCGCTCTGTGCGCCAAA
CodOp-SpeI-Rv	CCT <u>ACTAGT</u> TTGGGTTTTTCGGAACCGGCGTC

Construction of pLIC-KM-PPT₇₀₀₂ and pLIC-sfp

PPT-7002-Fw	CAATCAATTGAGGAGGAAAAAAAAAATGGCAGTGTCCGGTGGAAATATTGGTTAATTTCC
PPT-7002-Rv	CAC <u>TCTAGT</u> AGCTGCCCATATAGCATCCATTCC
Sfp-fw	CAATCCCGGGAGGAGGAAAAAAAAAATGAAGATTTACGGAATTTATATGGACCGCCCGCTTT
Sfp-rv	CAC <u>TCTAGT</u> ATTATAAAAGCTCTTCGTACGAGACCATTGTGATATCC

Construction of pAAS6803 and pAAS7942, pVO03, pVO04, pVO03-cysDNC and pVO04-cysDNC

AAS-7942-kpn-fw	CTAAGGTACCAGGAGGAAAAAAAAAATGACTGGAACCGCCCTCGCGCAAC
AAS-6803-kpn-fw	CTAAGGTACCAGGAGGAAAAAAAAAATGGACAGTGGCCATGGCGCTCAATC
AAS-7942-Rv	CAC <u>TCTAGT</u> ATTAACTCGCCGATTCAAACATCCCGTCC
AAS-6803-Rv	CAC <u>TCTAGT</u> ATTAAAAATTCGTCATTAATGTTGATAAGTTGGGTTACC
Sfp-xbaI-fw	CAATCTAGAAGGAGGAAAAAAAAAATGAAGATTTACGGAATTTATATGGACCGCCCGCTT
Sfp-sphI-Fw	CAATGCATGCAGGAGGAAAAAAAAAATGAAGATTTACGGAATTTATATGGACCGCCCGCTT
Sfp-sphI-Rv	CAC <u>TGCATG</u> CTTATAAAAGCTCTTCGTACGAGACCATTGTGATATCC
cysDNC-sphI-Fw	CAATGCATGCAGGAGGAAAAAAAAAATGGATCAAATACGACTTACTCACCTGCG
cysDNC-sphI-Rv	CAC <u>TGCATG</u> CTCAGGATCTGATAATATCGTTCTGTCTCAACAG
cysDNC-pstI-Fw	CAATCTGCAGAGGAGGAAAAAAAAAATGGATCAAATACGACTTACTCACCTGCG
cysDNC-pstI-Rv	CAC <u>TCTGCAGT</u> CAGGATCTGATAATATCGTTCTGTCTCAACAG

For loading domain replacement in PCC 7002

mycA10-F1	GAGACAGGATGAGGATCGTTTCGCATGTATACCAGTCAATTTCAAACCTTAGTCGATG
Prom-amar-b2-2	CTGTTGAATAACAAGGACGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATG
Prom-amar-b2-3	GTTGACACGGGCGTATAAGACATGTTACTGTTGAATAACAAGGACGGATCTGATCAAG
Prom-amar-b2-4	CC <u>ACTGCA</u> GGATCTCAATGAATATTGGTTGACACGGGCGTATAAGACATGTTACTACTG
mycA10-R1	GTGGTATTTTCTAGCGCTTGATTACCTGGACTTGCTTGATTTCGGAATGGTTTCGACAAGG
mycA10-R2	GCAATTTCTTTGCCCTGAATTTTCGGCACCTGTGGIATTTTCTAGCGCTTGATTACCC
mycA10-R3	GGAAAACGACAACCTGAGACCCACCACGGCAATTTCTTTGCCCTGAATTTTCGGCACCT
mycA10-R4	CCAAAAAGCTTCGGGGTTGTACGCTTGGGGAAAACGACAACCTGAGACCCACCACG
7002mycA10-a2	CCAGAATTCGGGAGCTTCATCCTGGGACAATGG
7002mycA10-a1	GCTTTACGCCACCTGTTCCCAATATGC
7002mycA10-b1	CCG <u>TCTAGT</u> CAATGCGGGGATTGAGC
7002mycA10-b2	CAAGCTGACAACCCCGAAGCTTTTTGGG
7002mycA10-a1-SpeI	CCA <u>ACTAGT</u> GCTTTCAGCCACCTGTTCCCAATATGC

desE promoter replacement in PCC 7002

desE-US-Fw	TATGCACATATGCGTCGATTTTGCCTCATTAATTTAGTTAAAGCAGC
desE-US-Rv	GCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAGCTTTCCTAACGAGTTGAGAATATCTTCTATGAAAACCG
KM-Fw	TTAGAAAACTCATCGAGCATCAAATGAAACTGC
KM-Rv	GGACTCTTCTCTACAGGTGGGTATAGATTTGTAAAGCTTTGGCAGGATCCGGGTGCTAACAAA
Cpc-prom-Fw	CTTAACAAATCTATACCCACCTGTAGAGAAGAGTCC
Cpc-prom-Rv	GGTCAAGAACGTTGCTGTAATGCGTCATGGAATTAATCTCCTACTTGACTTTATGAGTTGGG
desE-DS-Fw	ACGCATTACAGCAACGTTCTTGACCC
desE-DS-Rv	CTGCCGCCAGGCAAATCTGTTTTATCCATATGCTAGGGATTGGCCCGTTTTGTAGATC
pBAD18-Fw	GATAAAACAGAATTTGCCTGGCGGCAG
pBAD18-desE-USRv	GCTGCTTTAACTAAATTAATGAGGCAAAATCGACGCATATGTGCATA
	GGAGAAACAGTAGAGAGTTGCGATAAAAAGCG

Thioesterase expression in PCC 7002

B3936-fw	GCGATCCGAATGGCGGAATCTTC
KM-Prom-Rv	GCTCACAATTCGTTTGTTCGTTTCAATGAAGCTTTGGCAGGATCCGGCTGCTAACAAA
Prom-Fw	CCTTCATTGAAACGAACAAACGAATTGTGAGC
Prom-Rv	GTTTTGGTTTCGGTTTTCCACTCTAGAGTCAT
BTE-Prom-Fw	ATGACTCTAGAGTGGAAACCGAAACCAAAAC
BTE-Prom-Rv	CAGCCAACCTCAGCTTCCTTTTCGGTTAAACACGAGGTTCCGCCGAATTAC
LacI-Fw	CCGAAAAGGAAGCTGAGTTGGCTG
LacI-Rv	GAAGATTCCGCCATTTCGGATCGCGAACCCGGAAGGAGCTGACTGGG
KM-tesA-Rv	CCAGAATCAATAACGTGTCCGCCATGGAATTAATCTCCTACTTGACTTTATGAGTTGGTGTG
tesA-Fw	ATGGCGGACACGTTATTGATTCTGG
tesA-Rv	CAGCCAACCTCAGCTTCCTTTTCGGTTATGAGTCATGATTTACTAAAGGCTGCAACTGC
tesA-colonyRv	GCTTCATTATAACGGCGACCATAGTTTGC
KM-GeoTE-Rv	GGTCCACTTATCGATCATCAGCTCCATGGAATTAATCTCCTACTTGACTTTATGAGTTGGTGTG
GeoTE-Fw	ATGGAGCTGATGATCGATAAGTGGACC
GeoTE-Rv	CAGCCAACCTCAGCTTCCTTTTCGGTTAGTGGTGATGGTGATGATGGCTTTTAC

Plasmids for ACC overexpression in PCC 7002 (pNACC and pACC6803)

ACC-Fw (6803)	CAACATATGTAGGCTGTGGTTCCTTAGGCAACAGT
ACC-Rv (6803)	CAAGGATCCTCATTACACCCCGTTTCTAAAAAATTGACCCAAATG
ACC-Fw2 (6803)	CAACATATGCAGGCAGCAGCCTCGATGGCTATTAACCTTACGGAACCTGCGGGAATTG
ACCD-Fw (<i>E. coli</i>)	ACTCATATGCAGGCAGCAGCCTCG-ATGAGCTGGATTGAACGAATTAAGCAACATTACT
ACCD-Rv (<i>E. coli</i>)	TACGGGATCCTCAGGCCTCAGGTTCTGATCCG
ACCabc-Fw2 (<i>E. coli</i>)	GCTAGGATCCAGGAATACTATGAGTCTGAATTTCTTGATTTGAACAGCC
ACCabc-Rv (<i>E. coli</i>)	GGATTCTAGATTATTTTTCTGAAGACCGAGTTTTTCTCCAGATAGTG

Target genes (e.g. *ols*, *ols* domains, *ols*^{opt}) were amplified by PCR using primers that added the appropriate restriction sites (EcoRI-BglII or Mfe-SpeI), a strong ribosome binding site (AGGAGG) and eliminated the stop codon of each target gene. The *sfp* gene was amplified from genomic DNA extracted from BAP1 cells, the *ppt* from PCC 7002 was amplified from genomic DNA and the *cys*DNA operon was amplified from *E. coli* genomic DNA. Each PCR product was digested with the appropriate restriction enzyme (SpeI-KpnI, BglII-KpnI) and ligated with a correspondingly digested vector (pLIC-KM, pBAD33). The ACP2-ST-TE domains were codon optimized for expression in *E. coli* and synthesized (DNA 2.0). For construction of pOLS-Opt,

the ACP2-ST-TE codon optimized domains were amplified including SpeI sites at the 5' and 3' end and ligated with the SpeI digested pLD-ACP1-KS-AT-KR construct; clones with the correct orientation were screened using colony PCR. Hydrocarbon and fatty acid analysis were performed as described in Chapters 2 and 3.

5.4.2 Cyanobacterial constructs

The upstream and downstream flanking sequences of *ols* (SYNPCC7002_A1173), *desE* (SYNPCC7002_A2833) and *aas* (SYNPCC7002_A0675) were amplified by PCR (phusion polymerase) from genomic DNA isolated using commercial reagents (Promega). The *acc* operons from *Synechocystis sp.* PCC 6803 (47) or *E. coli* (48) were amplified and products were digested with appropriate restriction enzymes and gel purified using commercial kits (Qiagen). The *aadA* gene (strep^R) and *aphII* gene (Km^R) were excised from plasmid pSRA81 or pET28. Thioesterases were amplified from plasmids pBTRCk-tesA (*tesA*), pBAD34-BTE (BTE) and pBAD18-GeoTE (GeoTE) (19) and were inserted into pBAD18 along with the upstream and downstream regions flanking the *aas* gene, *lacI*, the *aphII* gene (Km^R) and an IPTG inducible promoter using Gibson cloning (51). Transformations were performed as described in chapter 2. Hydrocarbon and fatty acid analysis were performed as described in Chapters 2 and 3.

5.4.3 Microtiter plate growth assays and protein purifications

Overnight cultures grown from a single colony were diluted to an OD₆₀₀~0.05–0.1 in Luria-Bertani (LB) medium containing 100 µg/ml ampicillin to maintain plasmids, varying

concentrations of isopropyl β -d-1-thiogalactopyranoside (IPTG) to induce transcription, and varying concentrations of kanamycin or chloramphenicol when required. All growth experiments were performed in triplicate. For each assay, 200 μ L of culture was added to an individual well of a black, clear-bottomed 96-well polystyrene plate (Costar, Corning, NY) and incubated for 24 h with shaking at 37°C or 30°C using a Tecan (Männedorf Switzerland) M1000 microplate reader. Fluorescence was measured at 558 nm excitation with emission at 583 for RFPEC and 514 nm excitation with emission at 527 nm for YFP.

To produce and purify proteins, a single colony of *E. coli* DH10B harboring the plasmid containing the target gene was inoculated into 10 mL of LB broth containing 100 μ g/ml ampicillin, and was incubated at 37 °C overnight. The culture was diluted with fresh medium (100 mL) and grown at 30°C to an OD₆₀₀ of ~0.4 at which point IPTG was added to a final concentration of 1 mM. Induced cultures were incubated for an additional 20 h at 15°C. After finding the best conditions for translation, expression of the pOLS-Opt construct was performed in a 50 mL culture grown at 30 °C overnight with 100 μ g/ml ampicillin and 0.01 mM IPTG. For each protein preparation, cells were collected by centrifugation and lysed via sonication. The cleared lysate was mixed with Ni-NTA agarose (Qiagen) for 1 h at 4°C with gentle agitation. The resin was washed (30 mM imidazole, 50 mM NaH₂PO₄, 300 mM NaCl) prior to elution of the 6x-his-tagged target protein with elution buffer (250 mM imidazole, 50 mM NaH₂PO₄, 300 mM NaCl). Purified proteins were separated by 8–12% SDS-PAGE and identified by size in comparison to bands from a Precision Plus Protein dual color standard ladder (BIO-RAD, Hercules, CA).

5.4.4 Preparation of fluorescently labeled CoA and Bodipy-CoA in vitro reactions

Fluorescently labeled CoA were prepared following previously published protocols (27). Bodipy (Invitrogen) was dissolved in dimethyl sulfoxide (DMSO) to 60 mM and added to a reaction mixture (2ml) containing CoA (Sigma), DMSO and reaction buffer. Final concentrations of reaction components were 1.2 mM Bodipy, 0.5mM CoA and 10% (v/v) DMSO in reaction buffer (137mM NaCl, 10 mM sodium phosphate and 50 mM MgCl₂). The reaction was incubated 85 min at room temperature in the dark. The sample was then extracted with ethylacetate to remove excess Bodipy and the concentration of the Bodipy-CoA derivative was determined by measuring the absorbance at 505 nm using a molar extinction coefficient of 79000.

Purified Ols was added to a reaction mixture (25µL) containing CoA reaction buffer (75mM Tris-HCl, 10mM MgCl₂, 25mM dithiothreitol [DTT]), Bodipy-CoA (9.5 µM) and PPT7002 (0.012 µg/µL). The reactions were incubated in the dark for 45 min at room temperature and then terminated by adding 25 µL of 2x SDS gel loading buffer solution and heating to 100°C for 2 min. Samples were loaded onto a 10% Tris-HCl SDS-PAGE gel. After electrophoresis the SDS-PAGE gel was analyzed in a Typhoon Scanner (GE healthcare) for the Bodipy fluorophores. Following scanning on the Typhoon imager, the gel was stained with cromassie blue.

5.5 References

1. **Mendez-Perez D, Gunasekaran S, Orlor VJ, Pfleger BF.** 2012. A translation-coupling DNA cassette for monitoring protein translation in *Escherichia coli*. *Metabolic engineering* **14**:298–305.
2. **Keasling JD.** 2010. Manufacturing molecules through metabolic engineering. *Science* (New York, N.Y.) **330**:1355–8.

3. **Lee JW, Na D, Park JM, Lee J, Choi S, Lee SY.** 2012. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nature chemical biology* **8**:536–46.
4. **Oliver G, Gosset G, Sanchez-Pescador R, Lozoya E, Ku LM, Flores N, Becerril B, Valle F, Bolivar F.** 1987. Determination of the nucleotide sequence for the glutamate synthase structural genes of *Escherichia coli* K-12. *Gene* **60**:1–11.
5. **Kane JF, Pharmaceuticals SB, Prussia K.** 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli* **8**:494–500.
6. **Gustafsson C, Govindarajan S, Minshull J.** 2004. Codon bias and heterologous protein expression. *Trends in biotechnology* **22**:346–53.
7. **Mutka SC, Carney JR, Liu Y, Kennedy J.** 2006. Heterologous production of epothilone C and D in *Escherichia coli*. *Biochemistry* **45**:1321–30.
8. **Waters NC, Kopydlowski KM, Guszczynski T, Wei L, Sellers P, Ferlan JT, Lee PJ, Li Z, Woodard CL, Shallom S, Gardner MJ, Prigge ST.** 2002. Functional characterization of the acyl carrier protein (PfACP) and beta-ketoacyl ACP synthase III (PfKASIII) from *Plasmodium falciparum*. *Molecular and biochemical parasitology* **123**:85–94.
9. **Marahiel MA, Stachelhaus T, Mootz HD.** 1997. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chemical reviews* **97**:2651–2674.
10. **Quadri LE, Weinreb PH, Lei M, Nakano MM, Zuber P, Walsh CT.** 1998. Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37**:1585–95.
11. **Copp JN, Neilan BA.** 2006. The phosphopantetheinyl transferase superfamily: phylogenetic analysis and functional implications in cyanobacteria. *Applied and environmental microbiology* **72**:2298–305.
12. **Mofid MR, Finking R, Marahiel MA.** 2002. Recognition of hybrid peptidyl carrier proteins/acyl carrier proteins in nonribosomal peptide synthetase modules by the 4'-phosphopantetheinyl transferases AcpS and Sfp. *The Journal of biological chemistry* **277**:17023–31.
13. **Caffrey P, Green B, Packman LC, Rawlings BJ, Staunton J, Leadlay PF.** 1991. An acyl-carrier-protein-thioesterase domain from the 6-deoxyerythronolide B synthase of *Saccharopolyspora erythraea*. High-level production, purification and characterisation in *Escherichia coli*. *European journal of biochemistry / FEBS* **195**:823–30.

14. **Roberts GA, Staunton J, Leadlay PF.** 1993. Heterologous expression in *Escherichia coli* of an intact multienzyme component of the erythromycin-producing polyketide synthase. *European journal of biochemistry / FEBS* **214**:305–11.
15. **Lambaloti R, Gehring AM, Fluge RS, Zuber P, Lacelle M, Marahie MA, Reid R, Khosia C, Walsh CT.** 1996. A new enzyme transferases superfamily - the phosphopantetheinyl. *Chemistry & Biology* **3**:923–936.
16. **Gokhale RS.** 1999. Dissecting and Exploiting Intermodular Communication in Polyketide Synthases. *Science* **284**:482–485.
17. **Pfeifer B a, Admiraal SJ, Gramajo H, Cane DE, Khosla C.** 2001. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science (New York, N.Y.)* **291**:1790–2.
18. **Yin J, Lin AJ, Golan DE, Walsh CT.** 2006. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nature protocols* **1**:280–5.
19. **Lennen RM, Pflieger BF.** 2013. Modulating Membrane Composition Alters Free Fatty Acid Tolerance in *Escherichia coli*. *PloS one* **8**:e54031.
20. **Jiang Y, Morgan-Kiss RM, Campbell JW, Chan CH, Cronan JE.** 2010. Expression of *Vibrio harveyi* acyl-ACP synthetase allows efficient entry of exogenous fatty acids into the *Escherichia coli* fatty acid and lipid A synthetic pathways. *Biochemistry* **49**:718–26.
21. **Pluschke G, Hirota Y, Overath P.** 1978. Function of phospholipids in *Escherichia coli*. Characterization of a mutant deficient in cardiolipin synthesis. *The Journal of biological chemistry* **253**:5048–55.
22. **Kaczmarzyk D, Fulda M.** 2010. Fatty acid activation in cyanobacteria mediated by acyl-acyl carrier protein synthetase enables fatty acid recycling. *Plant physiology* **152**:1598–610.
23. **Schmidt A, Christen U.** 1978. A factor-dependent sulfotransferase specific for 3'-phosphoadenosine-5'-phosphosulfate (PAPS) in the Cyanobacterium *Synechococcus* 6301. *Planta* **140**:239–244.
24. **Negishi M, Pedersen LG, Petrotchenko E, Shevtsov S, Gorokhov A, Kakuta Y, Pedersen LC.** 2001. Structure and function of sulfotransferases. *Archives of biochemistry and biophysics* **390**:149–57.
25. **Leyhl TS, Taylor JC, Markham D.** 1988. The Sulfate Activation Locus of *Escherichia coli* K12: cloning, genetic, and enzymatic characterization. *The Journal of biological chemistry* **263**:2409–2416.

26. **Yuzawa S, Kim W, Katz L, Keasling JD.** 2012. Heterologous production of polyketides by modular type I polyketide synthases in *Escherichia coli*. *Current Opinion in Biotechnology* **23**:727–735.
27. **Lee KKM, Da Silva NA, Kealey JT.** 2009. Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. *Analytical biochemistry* **394**:75–80.
28. **Sørensen HP, Mortensen KK.** 2005. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial cell factories* **4**:1.
29. **Sørensen HP, Mortensen KK.** 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of biotechnology* **115**:113–28.
30. **Makrides SC.** 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiological reviews* **60**:512–38.
31. **Angov E, Hillier CJ, Kincaid RL, Lyon JA.** 2008. Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PloS one* **3**:e2189.
32. **Betancor L, Fernández M-J, Weissman KJ, Leadlay PF.** 2008. Improved catalytic activity of a purified multienzyme from a modular polyketide synthase after coexpression with *Streptomyces* chaperonins in *Escherichia coli*. *Chembiochem : a European journal of chemical biology* **9**:2962–6.
33. **Lee W-H, Kim M-D, Jin Y-S, Seo J-H.** 2013. Engineering of NADPH regenerators in *Escherichia coli* for enhanced biotransformation. *Applied microbiology and biotechnology* **97**:2761–72.
34. **Pfeifer BA, Khosla C.** 2001. Biosynthesis of polyketides in heterologous hosts. *Microbiology and molecular biology reviews : MMBR* **65**:106–18.
35. **Weissman KJ, Leadlay PF.** 2005. Combinatorial biosynthesis of reduced polyketides. *Nature reviews. Microbiology* **3**:925–36.
36. **Menzella HG, Reid R, Carney JR, Chandran SS, Reisinger SJ, Patel KG, Hopwood DA, Santi D V.** 2005. Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nature biotechnology* **23**:1171–6.
37. **Khosla C, Kapur S, Cane DE.** 2009. Revisiting the modularity of modular polyketide synthases. *Current opinion in chemical biology* **13**:135–43.

38. **Hansen DB, Bumpus SB, Aron ZD, Kelleher NL, Walsh CT.** 2007. The loading module of mycosubtilin: an adenylation domain with fatty acid selectivity. *Journal of the American Chemical Society* **129**:6366–7.
39. **Wilkinson CJ, Frost EJ, Staunton J, Leadlay PF.** 2001. Chain initiation on the soraphen-producing modular polyketide synthase from *Sorangium cellulosum*. *Chemistry & biology* **8**:1197–208.
40. **Byers DM, Gong H.** 2007. Acyl carrier protein: structure-function relationships in a conserved multifunctional protein family. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **85**:649–62.
41. **Xu Y, Alvey RM, Byrne PO, Graham JE, Shen G, Bryant DA.** 2011. Expression of Genes in Cyanobacteria: Adaptation of endogenous plasmids as platforms for high-level gene expression in *Synechococcus* sp. PCC 7002. *Photosynthesis Research Protocols* **684**:273–293.
42. **Davis MS, Solbiati J, Cronan JE.** 2000. Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in *Escherichia coli*. *The Journal of biological chemistry* **275**:28593–8.
43. **Lennen RM, Pflieger BF.** 2012. Engineering *Escherichia coli* to synthesize free fatty acids. *Trends in biotechnology* **30**:659–667.
44. **Cho H, Cronan JE.** 1995. Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis. *The Journal of biological chemistry* **270**:4216–9.
45. **Voelker T a, Davies HM.** 1994. Alteration of the specificity and regulation of fatty acid synthesis of *Escherichia coli* by expression of a plant medium-chain acyl-acyl carrier protein thioesterase. *Journal of bacteriology* **176**:7320–7.
46. **Yuan L, Voelker TA, Hawkins DJ.** 1995. Modification of the substrate specificity of an acyl-acyl carrier protein thioesterase by protein engineering. *Proceedings of the National Academy of Sciences of the United States of America* **92**:10639–43.
47. **Liu X, Sheng J, Curtiss R.** 2011. Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **108**:6899–904.
48. **Lennen RM, Braden DJ, West RA, Dumesic JA, Pflieger BF.** 2010. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnology and bioengineering* **106**:193–202.

49. **Ruffing AM, Jones HDT.** 2012. Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnology and bioengineering* **109**:2190–9.
50. **Pfeifer B a, Admiraal SJ, Gramajo H, Cane DE, Khosla C.** 2001. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science (New York, N.Y.)* **291**:1790–2.
51. **Gibson DG, Young L, Chuang R, Venter JC, Iii CAH, Smith HO, America N.** 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases **6**:12–16.

Chapter 6: Conclusions

6.1 α -olefin biosynthesis in *Synechococcus sp.* PCC 7002

Cyanobacteria are photosynthetic organisms that have the potential to produce hydrocarbons and other valuable products from inexpensive substrates. In order to rationally engineer and enhance production of these compounds, a deeper understanding of the pathways involved in their biosynthesis is required. Cyanobacteria produce two types of long chain hydrocarbons from fatty acids: alkanes and 1-alkenes (α -olefins). Prior to this thesis, only the enzymes involved in the biosynthesis of alkanes had been characterized in cyanobacteria. In this thesis we studied the biosynthesis of α -olefins, which are one of the primary products of petroleum and gas refining, and identified the genes involved in their biosynthesis.

After analyzing the hydrocarbon profile from the cyanobacterium *Synechococcus sp.* PCC 7002 (PCC 7002), we demonstrated that this organism synthesizes two α -olefins from fatty acids, 1-nonadecene (C19:1) and 1, 14-nonadecadiene (C19:2), *via* an elongation-decarboxylation mechanism. Since no genes had been linked to the formation of α -olefins via an elongation-decarboxylation mechanism, we searched for homologs of the well characterized CurM domain of the curacin A biosynthetic pathway in *Lyngbya majuscula* that is responsible for the production of the terminal alkene functionality in curacin A. We identified a gene (*ols* gene, for olefin synthase) encoding for a protein with modular organization, similar to a polyketide synthase, and then demonstrated the involvement of this gene in α -olefin biosynthesis using genetic studies.

The biosynthesis of the internal double bond in 1, 14-nonadecadiene was not explained by the decarboxylation and dehydration reactions performed by the Ols domains, so in Chapter 3 we linked the presence of this internal double bond to a gene predicted to encode a desaturase (DesE); desaturases are enzymes that introduce double bonds into the hydrocarbon chains of fatty acids in response to changes in temperature. The amino acid sequence encoded by the *desE* gene showed a high degree of similarity to Δ^9 desaturases, suggesting that it is acting on C14 fatty acids, which after elongation to C18:1(Δ^13) fatty acids would serve as the precursors for the formation of the hydrocarbon with the internal double bond. When PCC 7002 is grown at low temperatures, there is an increase in the expression of the *desE* gene (1), suggesting that the alkenes synthesized by PCC 7002 are important at low temperatures. We studied the growth of the wild type and knockout mutant strains under various temperatures and showed that there is an increase in C19:2 abundance as an inverse function of temperature; suggesting that the compound plays a role in responding to cold stress. This was also supported by the fact that the Δols strain was unable to grow at low temperatures.

Besides *ols* and *desE*, two additional genes that are also important in alkene biosynthesis were identified in PCC 7002: a 4'-phosphopantetheinyl transferase (PPT) and an acyl-ACP synthetase (Aas). PPTs are enzymes that modify acyl carrier proteins (ACP), which are important components in both fatty acid and polyketide biosynthesis, from the inactive apo-enzyme to their active holo-form by transferring the 4'-phosphopantetheinyl moiety from coenzyme A to a conserved serine residue of the carrier protein. By expressing and purifying the PPT from PCC 7002 in *E. coli*, we showed that the identified gene in fact encodes for a PPT that was able to activate the ACP domains in Ols.

As mentioned in Chapter 5, C18 fatty acids are the precursors for olefin biosynthesis and in order to increase the production of olefins it is necessary to make sure these substrates are available; therefore, expression of thioesterases, enzymes that hydrolyze the acyl-ACP thioester to liberate the free fatty acids, is a strategy that might be used to increase olefin production. However, the free fatty acids released by thioesterases need to be activated by an acyl-ACP synthetase (Aas) to form acyl-ACPs before they can serve as substrates for Ols. We identified the gene encoding the Aas from PCC 7002 that is responsible for the activation of free fatty acids by comparing Aas homologs from other cyanobacterial strains (2) and tested it in *E. coli*.

After the identification of the genes involved in α -olefin biosynthesis, the long term goal of this project will be to engineer a host for olefin production. Mainly four enzymes are directly involved in α -olefin biosynthesis in PCC 7002 (Figure 6.1): the olefin synthase (Ols), a 4'-phosphopantetheinyl transferase (Ppt), a desaturase (DesE) and an acyl-ACP synthase (Aas). From the pathway presented in Figure 6.1, two strategies can be used to increase production of α -olefins: increasing expression of the *ols* and *desE* genes, and overexpression of a thioesterase with preference for C18 fatty acids to increase the availability of substrates for Ols.

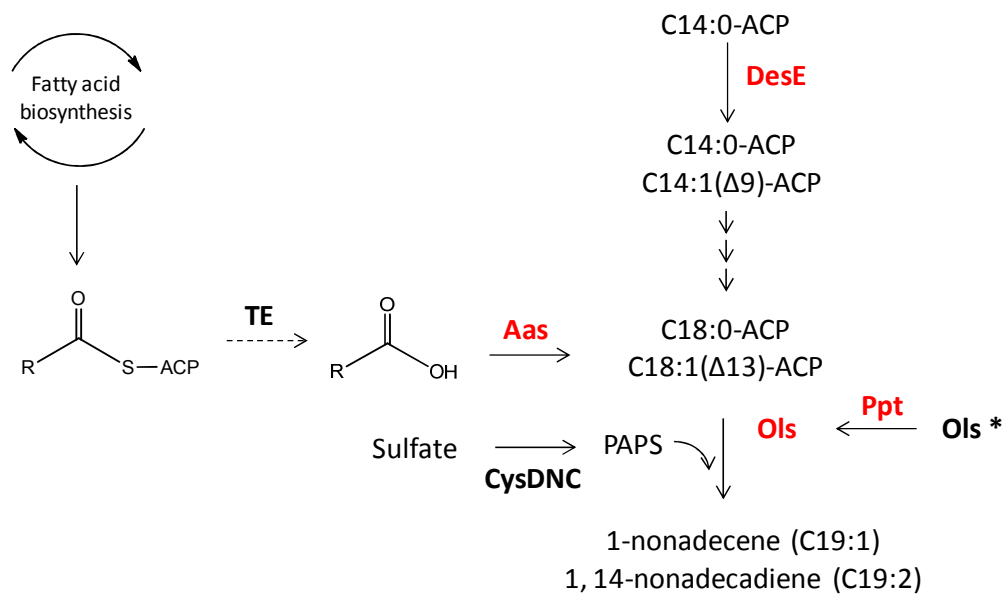


Figure 6.1 Activities involved in alkene biosynthesis, in red are shown the genes identified in this thesis. A dashed line indicates a pathway that is not present in cyanobacteria. R=C₁₇H₃₅, Ols*=inactive Ols holo form, TE=thioesterase, Aas, acyl-ACP synthase, Ols=olefin synthase, DesE=desaturase, PAPS=3'-phosphoadenosine-5'-phosphosulfate, Ppt=4'-phosphopantetheinyl transferase.

As can be seen in Figure 6.2, we were able to increase production of olefins by approximately 10-fold after improving CO₂ delivery (by bubbling cultures with air or CO₂), increasing expression of the *ols* gene and expressing a thioesterase from *Geobacillus sp.* Y412MC10. Interestingly, increasing expression of *desE* didn't result in an increase of total hydrocarbons, but resulted in a change in hydrocarbon composition. These results suggest that Ols might have a preference for the unsaturated fatty acid, but more experiments would be required to confirm this hypothesis.

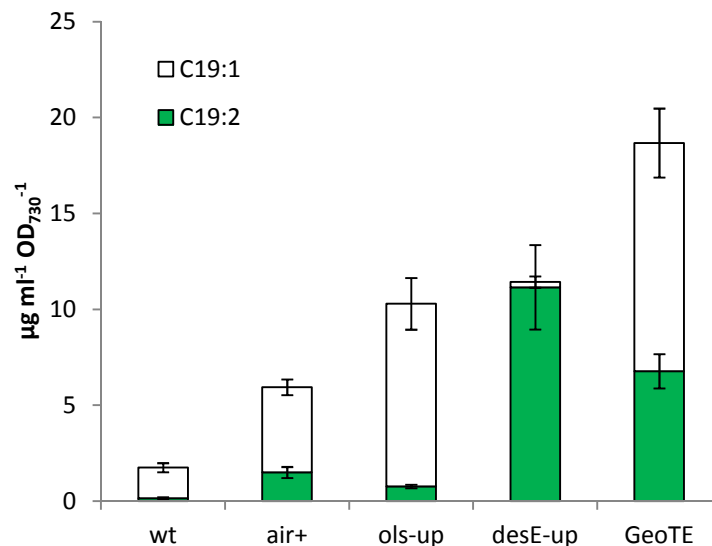


Figure 6.2 α -olefin production in *Synechococcus sp.* PCC 7002. Wt: wild type (No bubbling, 35°C), air+: Wild type (Air bubbling, 37°C), Ols-up: Ols upregulated strain (air bubbling, 37°C), desE-up (ols and desE upregulated strain 37°C), GeoTE: Expression of thioesterase from *Geobacillus sp.* Y412MC10 (CO₂ bubbling, 33°C).

6.2 Challenges of engineering cyanobacteria

Only a few cyanobacterial strains have been investigated as host organisms for chemical production, including *Synechococcus sp.* PCC 7002, *Synechocystis sp.* PCC 6803 and *Synechococcus sp.* PCC 7942. Over the past few years, they have been engineered for the production of a wide variety of fuels and chemicals like ethanol (3), fatty acids (4, 5), isobutanol (6), 1-butanol (7), 2, 3-butanediol (8) and 3-hydroxybutyrate (9). Although the productivities achieved in these studies are promising and have generated interest in cyanobacteria as a production host, the potential of genetically modified cyanobacteria is still in the initial stages of exploration and large-scale production facilities must still be constructed and tested.

The development of well characterized genetic tools and the advancement in understanding of fundamental metabolic pathways will likely be the rate-limiting steps in the development of cyanobacterial production systems. It is essential to investigate the native cyanobacterial genetics, metabolism and regulatory systems in order to predict the effects of targeted genetic manipulations to increase production as well as to enable the synthesis of new products. Also, a well characterized system of genetic tools like inducible promoters, counter selection systems, ribosome binding site libraries and expression plasmids is vital for future engineering efforts. Although some efforts have been made in this regard (10, 11), the development of such tools should be a major focus in the next few years. As the genetic tools, knowledge and engineering of cyanobacteria progresses, industrial interest will continue to expand to cover new applications, paving the way for a greener, post-petroleum era (12).

6.3 Translational coupling cassette and *E. coli* engineering

Heterologous gene expression is a powerful tool for imparting novel biochemistries, by-passing native pathways, or enhancing metabolic activity in model hosts (13–16). However, the major challenge to using heterologous expression in metabolic engineering experiments lies in the inability to quickly dissect experiments that have failed. In Chapter 4, we presented a method for dissecting failed heterologous expression experiments in *E. coli* by designing a DNA cassette that couples translation of a target gene to a response gene that generates an easily monitored phenotype *in vivo*, such as antibiotic resistance or fluorescence. The cassette is a powerful tool that could potentially be used to easily screen a large number of mutants or conditions to identify improvements in heterologous expression in *E. coli*.

We demonstrated the utility of the coupling cassette by optimizing conditions for expression of the *ols* gene, which encodes for a large multi-domain protein (302 kDa) similar to a polyketide synthase (PKS) with eight enzymatic and structural domains. First we used the translation coupling system to confirm that the *ols* gene was not being actively translated and then identified specific domains that were the cause for the lack of translation; after codon optimization of the problematic domains we found the best conditions for expression in *E. coli*. Although we were able to produce the Ols protein in *E. coli*, we didn't detect production of the alkenes under any of the conditions tested, suggesting that further engineering is required. It is also possible that *E. coli* is not the best candidate for heterologous alkene production and other organisms that are better suited for polyketide expression might be better candidates for alkene production using the Ols pathway.

6.4 References

1. **Ludwig M, Bryant DA.** 2012. Synechococcus sp. Strain PCC 7002 Transcriptome: Acclimation to Temperature, Salinity, Oxidative Stress, and Mixotrophic Growth Conditions. *Frontiers in microbiology* **3**:354.
2. **Kaczmarzyk D, Fulda M.** 2010. Fatty acid activation in cyanobacteria mediated by acyl-acyl carrier protein synthetase enables fatty acid recycling. *Plant physiology* **152**:1598–610.
3. **Dexter J, Fu P.** 2009. Metabolic engineering of cyanobacteria for ethanol production. *Energy & Environmental Science* **2**:857.
4. **Liu X, Sheng J, Curtiss R.** 2011. Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **108**:6899–904.
5. **Ruffing AM, Jones HDT.** 2012. Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnology and bioengineering* **109**:2190–9.

6. **Atsumi S, Higashide W, Liao JC.** 2009. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nature biotechnology* **27**:1177–80.
7. **Lan EI, Liao JC.** 2011. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metabolic engineering* **13**:353–63.
8. **Oliver JWK, Machado IMP, Yoneda H, Atsumi S.** 2013. Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proceedings of the National Academy of Sciences*.
9. **Wang B, Pugh S, Nielsen DR, Zhang W, Meldrum DR.** 2013. Engineering cyanobacteria for photosynthetic production of 3-hydroxybutyrate directly from CO₂. *Metabolic Engineering* **In press**:1–10.
10. **Begemann MB, Zess EK, Walters EM, Schmitt EF, Markley AL, Pfeleger BF.** 2013. An organic acid based counter selection system for cyanobacteria. *PloS one* **8**:e76594.
11. **Liu X, Curtiss R.** 2009. Nickel-inducible lysis system in *Synechocystis* sp. PCC 6803. *Proceedings of the National Academy of Sciences of the United States of America* **106**:21550–4.
12. **Ruffing AM.** 2011. Engineered cyanobacteria: Teaching an old bug new tricks. *Bioengineered Bugs* **2**:136–149.
13. **Keasling JD.** 2008. Synthetic biology for synthetic chemistry. *ACS chemical biology* **3**:64–76.
14. **Keasling JD.** 2010. Manufacturing molecules through metabolic engineering. *Science (New York, N.Y.)* **330**:1355–8.
15. **Pfeifer BA.** 2001. Biosynthesis of Polyketides in Heterologous Hosts **65**:106–118.
16. **Tyo KE, Alper HS, Stephanopoulos GN.** 2007. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends in biotechnology* **25**:132–7.