2017

# The Synthetic Biology of N2-Fixing Cyanobacteria for Photosynthetic Terpenoid Production

Charles T. Halfmann
*South Dakota State University*

# THE SYNTHETIC BIOLOGY OF N₂-FIXING CYANOBACTERIA FOR

# PHOTOSYNTHETIC TERPENOID PRODUCTION

BY

CHARLES T. HALFMANN

A dissertation proposal in partial fulfillment of the requirements for the

Doctor of Philosophy

Major in Biological Sciences

Specialization in Microbiology

South Dakota State University

2017

# THE SYNTHETIC BIOLOGY OF N₂-FIXING CYANOBACTERIA FOR

## PHOTOSYNTHETIC TERPENOID PRODUCTION

This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this dissertation does not imply that the conclusions reached by the candidates are necessarily the conclusions of the major department.

Ruanbao Zhou, Ph.D.
Dissertation Advisor                    Date

Volker Brözel, Ph.D.
Head, Department of Microbiology        Date

Kinchel Doerner
Dean, Graduate School                   Date

# ACKNOWLEDGMENTS

CONTENTS

LIST OF FIGURES

## LIST OF TABLES

ABSTRACT

THE SYNTHETIC BIOLOGY OF $N_2$-FIXING CYANOBACTERIA FOR

PHOTOSYNTHETIC TERPENOID PRODUCTION

CHARLES T. HALFMANN

2017

In the last few decades, concerns over global climate change, energy security, and environmental pollution have been rising. To overcome these challenges, the concept of "-n$^{th}$ generation" biofuels has emerged as a strategy to convert solar radiation into fuels and bulk industrial chemicals for societal use, while decreasing our consumption of non-renewable energy sources. Nitrogen-fixing cyanobacteria hold a distinct advantage in biofuel production over plants, given their ability to convert sunlight, air ($CO_2$ and $N_2$), and mineralized water to energy-dense carbon molecules, as well as fix atmospheric nitrogen gas into ammonia for metabolism. Engineered cyanobacteria with re-wired metabolic pathways have recently been designed through synthetic biology, and they possess the ability to synthesize new chemicals and biofuels, which are secreted from their cells. Terpenoids constitute one of the largest classes of organic molecules on Earth, and are attractive candidates as a fourth generation biofuel and industrial chemical. In cyanobacteria, the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway is responsible for building essential metabolites involved in photosynthesis, as well as precursors for terpenoid biosynthesis.

This dissertation encompasses research focused on redirecting MEP flux in the nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120 to engineered terpenoid sinks, namely, linalool ($C_{10}H_{18}O$) and farnesene ($C_{15}H_{24}$). Chapter 1 is a review of literature in the field of biofuels and cyanobacteria, and chapter 2 is an introduction/list of objectives for the research in this dissertation. In chapter 3, we present the genetic engineering of *Anabaena* to synthesize farnesene by expressing a plant farnesene synthase. In chapter 4, we present the genetic engineering of *Anabaena* to synthesize linalool during $N_2$-fixation, and increased linalool production is accomplished by the over-expression of three rate-limiting enzymes in the MEP pathway. In chapter 5, we examine the feasibility of a blocking a native carbon reservoir in the cyanobacterium to increase metabolite and energy availability for terpenoid synthesis, as well as physiological aspects of glycogen-deficiency in the cyanobacterium during diazotrophic growth. In chapter 6, we focus on introducing a synthetic photorespiratory bypass to reduce photorespiration and increase carbon partitioning towards linalool synthesis.

**Chapter 1: Literature Review**

*1.1 Introduction*

Over the last century, advances in medicine, technology, and economic development have raised the quality of life for humans on the planet, causing the global population to skyrocket to unprecedented levels. In 2012, the world population reached 7 billion people, and is projected to increase to roughly 9 billion by 2050 (Cohen 2003). This increase in the global population will undoubtedly strain our world food and energy production, since both economies are heavily dependent on non-renewables for sustainment. Fossil fuels continue to be the integral driver behind energy production, with petroleum production reaching over 27 billion barrels per year in the last several years (Brecha 2013). This finite resource will at some point reach a maximum extraction rate, after which production will begin to decline, a situation known as "peak oil" (Hook and Tang 2013). Declining oil production will likely impose higher costs on modern agriculture, which requires substantial energy use through growing, processing, transportation, and requires arable land and clean water (Stavi and Lal 2013). One must also consider the environmental costs of expanding land for agriculture development, natural gas fracking, and oil drilling. Unregulated expansion of these human activities may take a devastating toll on our environment and contribute to soil erosion, pollution, and loss of natural habitat and biodiversity (Andersen et al. 2013). It is also widely accepted by the scientific community that rising $CO_2$ levels since the industrial revolution are contributing to the observed increase in the average global temperatures and elevated sea levels, which may havoc low-lying coastal areas around the world. The bleak future

described above may transform into reality if humanity cannot adapt to a changing world, and embrace socio-economic strategies that mitigate our impact on the environment.

## 1.2 First-generation biofuels

The concerns regarding fossil fuel consumption has led to biotechnologies involved in sequestering atmospheric $CO_2$ into carbon feedstocks, which can eventually be processed and used for fuel. These strategies are based on the premise that $CO_2$ already present in the atmosphere can be locked into biomass and then combusted for energy, resulting in a carbon negative process that is completely renewable (carbon negative in that plant roots are not harvested). The production of liquid biofuels such as corn bioethanol and soybean biodiesel from plant biomass has gained considerable interest in the last few decades, resulting in the advent of "first generation" biofuels. First generation biofuels are usually created from easily-degradable feedstocks such as starch, sucrose, and vegetable oils, which are extracted from agricultural food crops. These feedstocks are then brought through a chemical or biologically process to create the desired fuel. The process that is used depends on the physical characteristics of the feedstock, as some require more processing than others. Sugar and starch feedstocks are usually fermented by yeast to produce ethanol, while plant oils are transesterfied to produce biodiesel that is compatible with current combustion engines.

### 1.2.1 Sugar & starch feedstocks

The fermentation of sugar and starch feedstocks began its current expansion in the late 1990's, and is most prevalent in Brazil and the United States (Nigam and Singh 2011). Bioethanol fuel use has increased by 56% from 2007-2012, with the United States making up almost 60% of the market in global production (Figure 1). In the U.S., corn is

the major feedstock, with sugar beet, or sweet sorghum also used in smaller amounts.   In

Brazil, sugarcane has become so popular that it accounts for three-quarters of the

feedstock used for ethanol production.  Brazil produced 29 billion liters of ethanol from

sugarcane in 2009, creating an extremely successful bioethanol economy (Caldeira-Pires

et al. 2013).  Currently, bioethanol from sugarcane and sweet sorghum provide the most

sustainable biofuel feedstock, as these crops make the most efficient use of land, water,

nitrogen, and energy resources (de Vries et al. 2010).
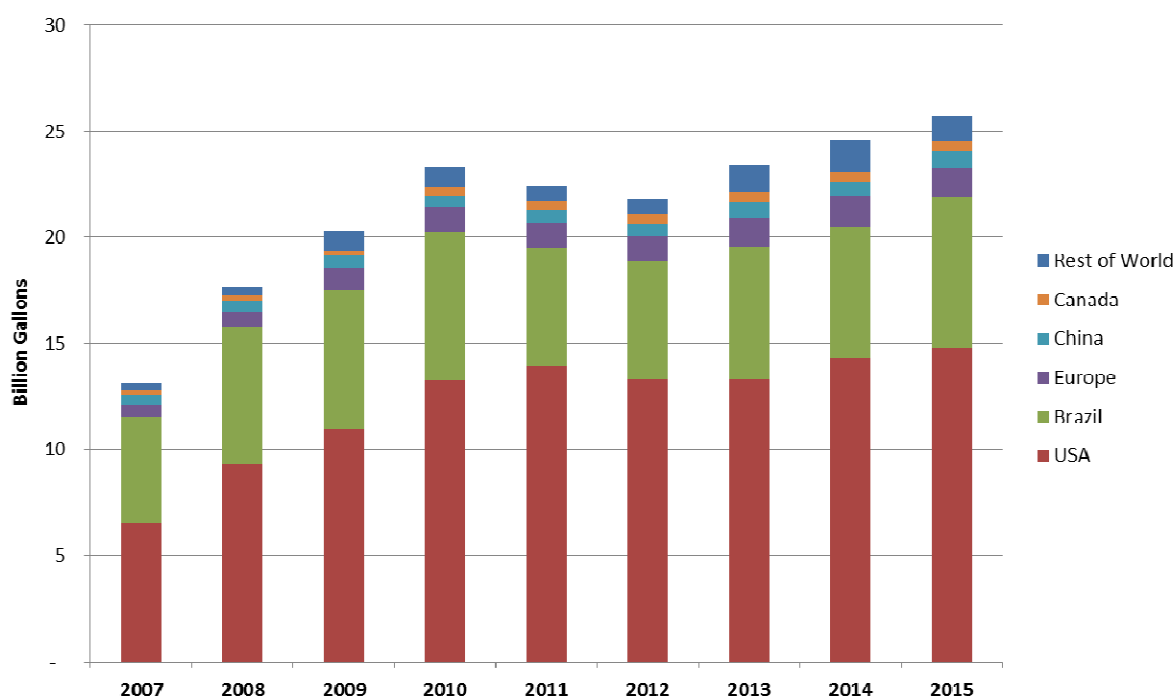


**Figure 1.**  World Ethanol Production 2007-2015.  Source:  www.afdc.energy.gov/dta

Bioethanol production in the United States is largely produced from corn, a starch-

based fermentation feedstock.  Unlike sugarcane, corn starch must first be saccharified to

generate simple sugars (glucose), which can then be fermented into ethanol (Figure 2).  In

dry milling, the kernels are ground to a meal and treated with water and various enzymes

to break down the starch to simple sugars, creating a mash.  Fermentation of the mash

results in an ethanol-stillage mix, which is then distilled to create an azeotropic mixture

of ethanol water (95% ethanol/ 5% water, %w/w) (Martin and Grossmann 2013). This ethanol must be further purified using molecular sieves and dehydrated, to create a fuel product that can be blended with gasoline. Using this process, a 56-lb bushel of corn can be converted to 18-lb ethanol, 18-lb of $CO_2$, and 18-lb of dried distillers grains and solubles (DDGS).
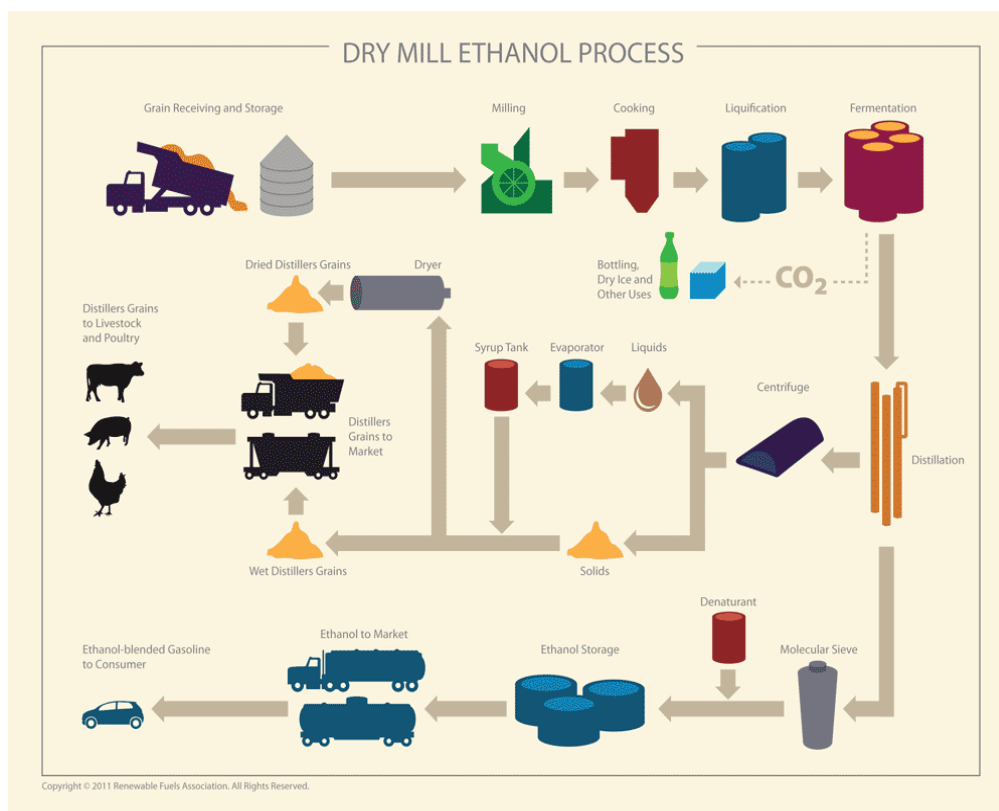


**Figure 2.** Process flow sheet for ethanol manufacture from corn kernels. Source: Renewable Fuels Association.

## 1.2.2 *Oilseed Feedstocks*

Vegetable oils from soybean, sunflower, palm seed, and rapeseed contain high amounts of triglycerides, which can be converted into a biodiesel. This is accomplished through the transesterification of a triglyceride with methanol or alcohol in the presence of a base catalyst to produce a mixture of fatty acid alkyl esters and glycerol (Stephenson

et al. 2008). Biodiesel has gained importance as an environment-friendly diesel fuel substitute, due to its production from renewable biological sources. The biodiesel industry has grown immensely in the United States, with over 1.1 billion gallons produced in 2011, making it the first EPA-designated Advanced Biofuel to reach 1 billion gallons of annual production (Eidman 2007).

### 1.2.3 Disadvantages of First-Generation Biofuels

Though increased production efficiencies have made first-generation biofuels cost-competitive with petroleum, there are still concerns over the sustainability of a biofuel platform that uses foodstuffs for energy. The use of land for biofuels that would otherwise be used for food production creates 'food vs. fuel' competition, resulting in higher food prices (although DDGS are valuable livestock feeds). This concern has been particularly vocalized in the United States (Odling-Smee 2007). Sugar, starch, and oilseed feedstocks also require substantial amounts of fertilizers, herbicides, and pesticides, such that some researchers conclude that corn ethanol and soybean biodiesel production has a net $CO_2$ emission, rather than being carbon neutral (Hill et al. 2006). Corn-based ethanol has been criticized not only for its energy balance, but also because of its high water demand. Currently, the suggested best possible water consumption for processing corn into biofuel is 2.85 gallon water/ gallon ethanol (Martin and Grossmann 2013), making ethanol production inefficient in clean water use. In summary, there are still concerns whether first-generation biofuels can be a sustainable biofuel platform in the long-term future.

*1.3   Second-Generation Biofuels*

In the context of biofuels, the term 'plant biomass' refers mostly to the tough, fibrous, non-food material that makes up a majority of the plant. Up to 75% of plant material is composed of cell wall components, such as lignin, cellulose, and hemicelluloses, which can be broken down from polysaccharides into simple sugars for fermentation (Gomez et al. 2008). This makes plant biomass a very promising source of material for biofuels, considering it is one of the most abundant and underutilized materials on our planet. This plant material often comes from fibrous plants such as wheat and switchgrass, or from the remnants of plant processing, such as wood chips, straw and corn stover. These feedstocks are advantageous over first-generation biofuels, considering their yield from ground to liquid fuel is higher than that using corn, especially switchgrass (Martin and Grossmann 2013). They also can be harvested in areas that do not compete with food crops with minimal input of fertilizers and pesticides, leaving a smaller footprint on the food supply chain.

*1.3.1   Process steps to generate biofuel from lignocelluloses*

There are two main pathways for producing liquid biofuels from lignocellulosic feedstocks. In thermochemical processing, the feedstock is converted into a range of products by thermal decay and chemical reformation, which is performed by heating the biomass in the presence of differing concentrations of oxygen (Antizar-Ladislao and Turrion-Gomez 2008). Heating with low concentrations of air leads to gasification and the production of hydrogen and organic gases, which are transformed into liquid fuels using the Fisher-Tropsch process. In the second process, known as biochemical processing, the lignocellulosic feedstock is treated with either a concentrated acid (e.g.

sulfuric acid), or pretreated with enzymes that are able to deconstruct the feedstock

material into sugars for fermentation (Figure 3).
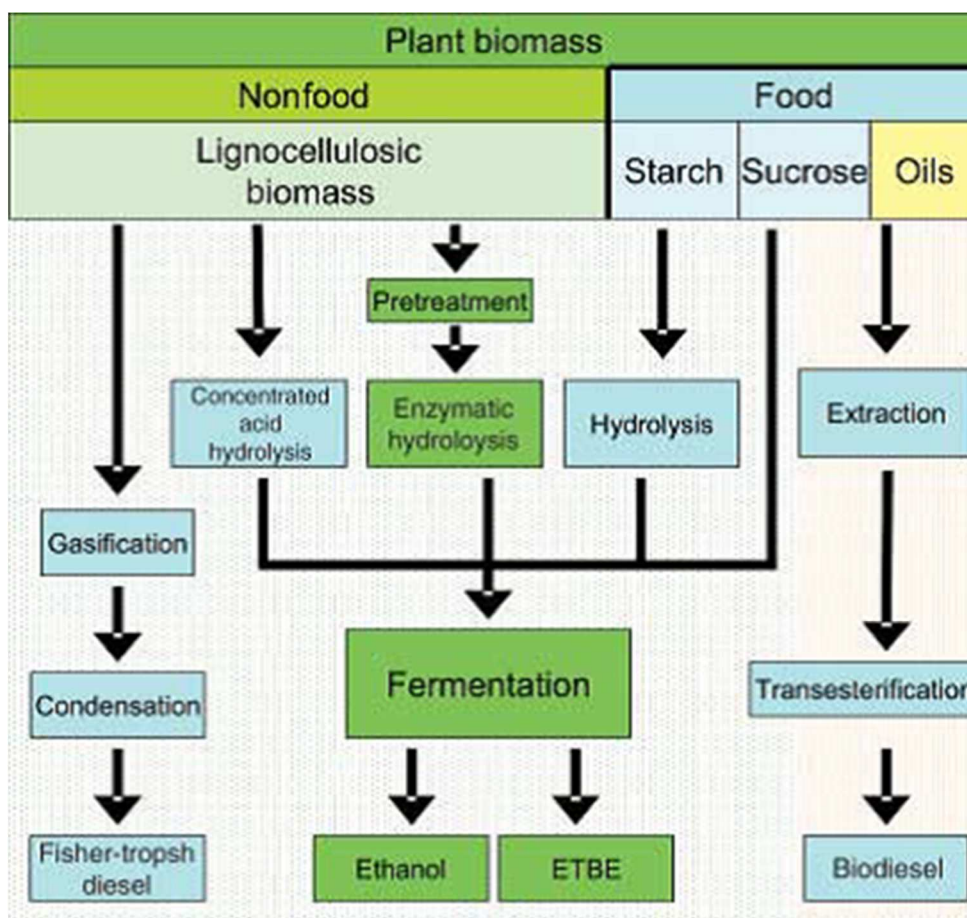


**Figure 3.** Liquid biofuel pathways from plant biomass Adapted from (Gomez et al. 2008)

### 1.3.2   Disadvantages of Second-Generation Biofuels

Though biofuels from lignocellulosic feedstocks have many advantages over food-

crop feedstocks, production of such fuels is not cost-effective because there are a number

of technical barriers that need to be overcome.  The presence of lignin in plant biomass

impedes the hydrolysis of polysaccharides to simple sugars, and can also inhibit the

fermentation process (Simmons et al. 2008).  These factors reduce the biofuel yield and

ultimately drive up the cost for biofuel production.  In the United States, there are

significant challenges in developing large areas of land for biofuel feedstocks, including ensuring sustainability, reducing costs, and creating responsible land-use policies (Hill et al. 2006). Many obstacles need to be overcome until the vision of large-scale biofuel production from lignocellulosic feedstocks can be fully realized.

## 1.4 Third-Generation Biofuels

Though the production of biofuels from plant-based feedstocks (first and second-generation biofuels) has given hope in mitigating the human carbon footprint, these energy strategies are unsustainable in their current forms. Realizing the limitations of higher plants for biofuel production, attention has focused in the last few decades to microscopic photosynthesizers of biofuel feedstocks. Like plants, photosynthetic microorganisms such as microalgae and cyanobacteria engage in oxygenic photosynthesis to harvest light energy and fix $CO_2$ into biomass. They also have the ability to grow and produce valuable byproducts such as lipids, carbohydrates, biogases, and other feedstocks that can be converted into biodiesel, bioethanol, biogases, and biohydrogen (Figure 4). The cultivation of algal biomass for conversion into fuel represents the "third-generation" of biofuels. This emerging sector in biofuel production aims to be a more sustainable method in converting the carbon sequestered in biomass to energy-rich fuels and products.
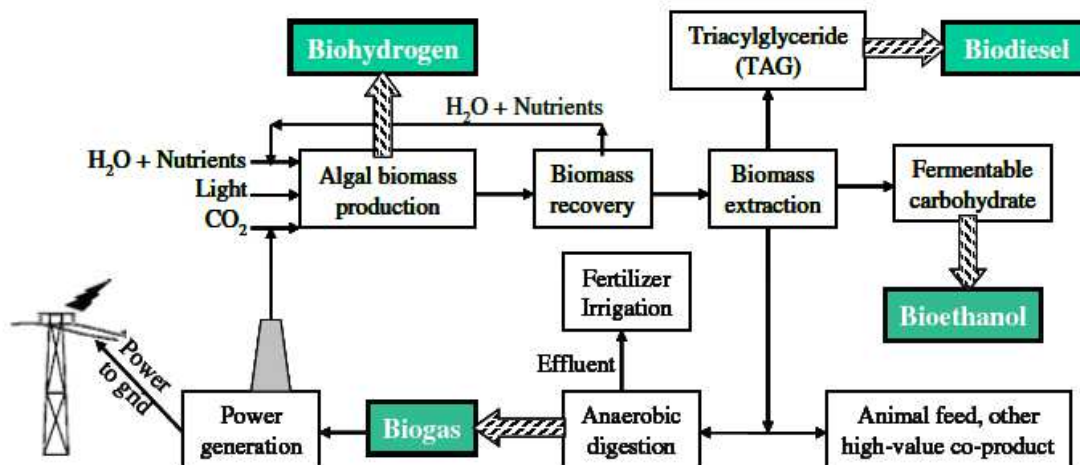
**Figure 4.** A conceptual model for integrated microalgal biomass and biofuel production. Adapted from (Singh and Dhar 2011).

### 1.4.1 Background on Photosynthetic Microorganisms

Aquatic microbial oxygenic photoautotrophs (AMPOS), including microalgae, diatoms, and cyanobacteria, make up an extensive and diverse group of unicellular and multicellular microorganisms that are capable of synthesizing valuable compounds using energy from light reactions, inorganic chemical reactions, and heterotrophic fermentation (Dismukes et al. 2008; Menetrez 2012). Unlike plants, microalgae and cyanobacteria lack many distinct organs and structures such as leaves, roots and stems. However, they do contain chlorophyll to carry out photosynthetic reactions. The most abundant AMOPS are single-cell drifters in plankton, generally called phytoplankton (Singh and Dhar 2011). Microalgae have been used in ancient times in human health food products, feeds for fish and livestock, and have been cultured for their high-value oils (Singh and Dhar 2011). The fossil record of these microorganisms dates back approximately three billion years ago, well into the Precambrian period (Menetrez 2012). They are inventors

of photosynthesis and the architects of our early atmosphere, producing a majority of the breathable oxygen and a large fraction of the coal and petroleum that is currently consumed for energy (Dismukes et al. 2008). Microalgae and cyanobacteria are ubiquitous in the biosphere, and are found in both freshwater and saltwater aquatic environments.

### 1.4.2 Advantages of Third-Generation Biofuels

AMOPS hold several key advantages over plants for biofuel production. First, they have superior photosynthetic capabilities, and are able to convert up to 10% of the harvested light energy into biomass, compared to agricultural plants which possess efficiencies at around 1% (Parmar et al. 2011). They have rapid growth rates (1 to 3 doublings per day), can be cultivated in brackish costal water and seawater, and do not compete with agricultural crops for arable land (Menetrez 2012). They hold abundant amounts of lipids (20-50% dry wt.) and fermentable biomass (starch and glycogen, 20-50% dry wt.), which can be converted directly into biofuels (Table 1). Previous studies have theoretically projected that 1 hectare (2.47 acres) of land could yield up to 200 barrels of oil from cultivated algae (Hu et al. 2008). This yield is 100 times greater than that for soybeans, a commonly used feedstock for biodiesel, and is greater than any demonstrated yield currently (Hu et al. 2008). AMPOS are also particularly attractive for their ability to utilize nitrogen and phosphorous from wastewater and $CO_2$ from industrial sources, which limits competition with clean water and recycles $CO_2$ that would otherwise be released into the atmosphere. Research into the mass of algae that could be supported by wastewater nutrients found that an excess of 71 metric tons per hectare per year of algal biomass can be produced from this process. This equals a total potential

volume of 270,000 gallons of biodiesel, enough fuel to sustain approximately 450 cars per year on average (Dalrymple et al. 2013).

**Table 1.** Constituents of microalgae (% dry matter) Adapted from Menetrez et al., 2012

| Algae | lipids | protein | carbohydrate |
|---|---|---|---|
| *Anabaena cylindrica* | 4−7 | 43−56 | 25−30 |
| *Aphanizomenon f los-aquae* | 3 | 62 | 23 |
| *Arthrospira maxima* | 6−7 | 60−71 | 13−16 |
| *Botryococcus braunii* | 86 | 4 | 20 |
| *Chlamydomonas rheinhardii* | 21 | 48 | 17 |
| *Chlorella ellipsoidea* | 84 | 5 | 16 |
| *Chlorella pyrenoidosa* | 2 | 57 | 26 |
| *Chlorella vulgaris* | 14−22 | 51−58 | 12−17 |
| *Dunaliella salina* | 6 | 57 | 32 |
| *Euglena gracilis* | 14−20 | 39−61 | 14−18 |
| *Prymnesium parvum* | 22−38 | 30−45 | 25−33 |
| *Porphyridium cruentum* | 9−14 | 28−39 | 40−57 |
| *Scenedesmus obliquus* | 12−14 | 50−56 | 10−17 |
| *Spirulina maxima* | 6−7 | 60−71 | 13−16 |
| *Spirogyra sp.* | 11−21 | 6−20 | 33−64 |
| *Spirulina platensis* | 4−9 | 46−63 | 8−14 |

*1.4.3   Environmental Conditions for Cultivating Photosynthetic Microbes*

In order for microalgal cultivation to be successful, specific environmental conditions need to be met, and these can vary from species to species.  The major parameters that influence biomass production include proper light intensity and wavelength, $CO_2$ concentration, nutrient composition, temperature, salinity, and aeration.

*1.4.3.1   Light*

Light is the primary energy source for microalgae, and therefore is a key component in the conversion of $CO_2$ into biomass.  Microalgae utilize photosynthetically active radiation (PAR) in the 400-700 nm light range to strip electrons from water molecules, which are eventually used to form NADPH and ATP for anabolism.  Light-harvesting antenna structures along the thylakoid membrane are extremely efficient at absorbing light; however, most cells are light saturated at about 20% of solar light intensities (Torzillo et al. 2003).  Oversaturation of light can result in photo-inhibition from excess $O_2$ concentrations in the culture, and the formation of reactive oxygen species (ROS) which damage many cellular processes.  While cells in the top layer need to deal with excess light, cell growth at lower depths in the water column can be limited by low light levels from self-shading.  These problems can be solved by decreasing the light path length of culturing systems to ensure that light is uniformly distributed throughout the culture.  Flat-paneled and tubular bioreactor designs have been used to increase the surface area of the culture, thus increasing the efficiency of light-to-cell.  Proper mixing also ensures that cells are not stationary in light or dark zones.

*1.4.3.2    Gas Exchange*

Considering that roughly 45-50% of algal biomass is composed of carbon, maintaining proper $CO_2$ concentrations in the culture is paramount (Doucha et al. 2005). Atmospheric air contains relatively low amounts of $CO_2$ (0.033%), therefore carbon must be supplemented to the culture to maximize growth (Singh and Dhar 2011).Generally, $CO_2$ is blended with air and injected into cultures by gas exchange vessels in photobioreactors or sumps in open raceways.  It is important to not only maintain proper $CO_2$ levels, but to ensure that $O_2$ levels do not exceed the levels in ambient air, as high oxygen levels can cause photo-oxidative damage to chlorophyll reaction centers, inhibiting photosynthesis and reducing productivity (Molina et al. 2001).  This is not usually a problem in open algal systems, but closed bioreactors need to be facilitated with components that allow proper gas exchange between the culture and atmosphere.

*1.4.3.3    Nutrients*

Nutrients in microalgae growth environments include an assortment of macronutrients, inorganic elements, and trace metals (Singh and Dhar 2011).  Though reports for required amounts of nutrients vary from species to species, the main macronutrients needed for biomass growth are nitrogen and phosphorous (N:P = 16:1). Nitrogen supply is especially critical for the production of algal triacylglycerides (TAGs) produced for biodiesel feedstocks, as it was found that 0.36 kg of nitrogen (0.46 kg $NH_4^+$ or 1.6 kg $NO_3^-$) is required for every 1 kg of TAG feedstock (Peccia et al. 2013).  Trace metals such as cobalt, copper, zinc, magnesium, iron, and molybdenum are essential in many photosynthetic processes, as well as co-factors for enzymatic reactions.

*1.4.3.4 Temperature*

Generally, rises in temperature promote cell division and lead to an exponential increase in culture growth until an optimum level is reached. If cells exceed the optimum temperature, they can become stressed, resulting in a culture die-off. Temperatures below the optimum will not necessarily kill the cells (unless it is in freezing conditions), but will inhibit cell growth and ultimately reduce biomass production. The rate of photosynthetic reactions is also affected by temperature, as most photosynthetically crucial enzymes have optimal activity between 25-35°C, and are denatured above 50°C. Changing air temperatures can effect biomass production, especially in open pond systems that are subject to fluctuating temperatures during diurnal cycles (Olaizola 2000).

*1.4.3.5 Mixing*

Constant agitation of the culture prevents sedimentation and keeps cells uniformly distributed in the culture fluid, allowing for proper $CO_2$ and $O_2$ distribution, and evenly dispersal of cells between light and dark zones. This can come from either mechanical agitation or through bubbling. A partial pressure of at least 0.15 kPa $CO_2$ needs to be maintained to prevent carbon limitation, and to allow a stoichiometric demand of 1.7 g $CO_2$ $g^{-1}$ biomass (Singh and Dhar 2011). An external supply of $CO_2$ (e.g., and ethanol plant or an industrial power plant) from flue gas would be a beneficial carbon source to an algal system.

*1.4.4 Cultivation Method*

The cultivation strategy used for microalgae is an extremely important aspect for third-generation biofuel production, and will highly influence the efficiency and productivity of the biomass production (Norsker et al. 2011). There are a wide variety of

strategies used for algal biomass production.  However, the two most common light-dependent methods are described below.

*1.4.4.1   Open Ponds*

Outdoor cultivation of algae and cyanobacteria in open ponds has been well documented (Borowitzka 2005; Green et al. 1996; Nurdogan and Oswald 1996).  These systems rely on a pond that is either excavated and lined or unlined with impermeable materials, or are built above ground and enclosed with walls.  These cultivation platforms are usually only suitable for fast-growing species of algae that can tolerate sub-optimal environmental conditions, such as *Dunaliella*, *Spirulina*, and *Chlorella* (Chisti, 2007).  There are four main types of pond systems that are currently utilized:  1) unmixed ponds, 2) raceway ponds, 3) circular ponds, and 4) thin layer, inclined ponds.  Unmixed ponds are used for the stationary growth of microalgae, have the lowest productivity rates, and are usually unsuitable for most strains (Borowitzka and Moheimani 2013).  Raceway ponds are the most widely used, and allow for a directional flow of culture along a track that is powered by a paddle or pump (Figure 5A).  Circular ponds consist of a centralized, rotating arm that mixes the culture fluid (much like the designs used in wastewater treatment plants) and are commonly used for the production of *Chlorella*.  Circular ponds have shown to have productivities ranging from 8.5 $g \cdot m^{-2} \cdot d^{-1}$ to 21 $g \cdot m^{-2} \cdot d^{-1}$, while thin layer, inclined ponds consist of a slightly inclined shallow trays, and may achieve productivities up to 31 $g \cdot m^{-2} \cdot d^{-1}$ (Singh and Dhar 2011).

*1.4.4.2    Closed Photobioreactors*

Due to the openness of most algal ponds, they often encounter problems with unstable environment, contamination, water evaporation, and low productivity.   This has led to the design of closed photobioreactor (PBR) systems that allow for more control over growth conditions and cultivation.   The design, efficiency, and cost-effectiveness of closed photobioreactors have been reported (Chisti 2007; Reichert et al. 2006; Suh and Lee 2001; Woods et al. 2010).   They can range in a wide variety of designs, e.g., tubular reactors (Figure 5B), vertical bubble columns, helical reactors, flat plate reactors, or a combination of designs (Fuentes et al. 1999; Singh and Sharma 2012; Tamburic et al. 2011; Tredici and Zittelli 1998).
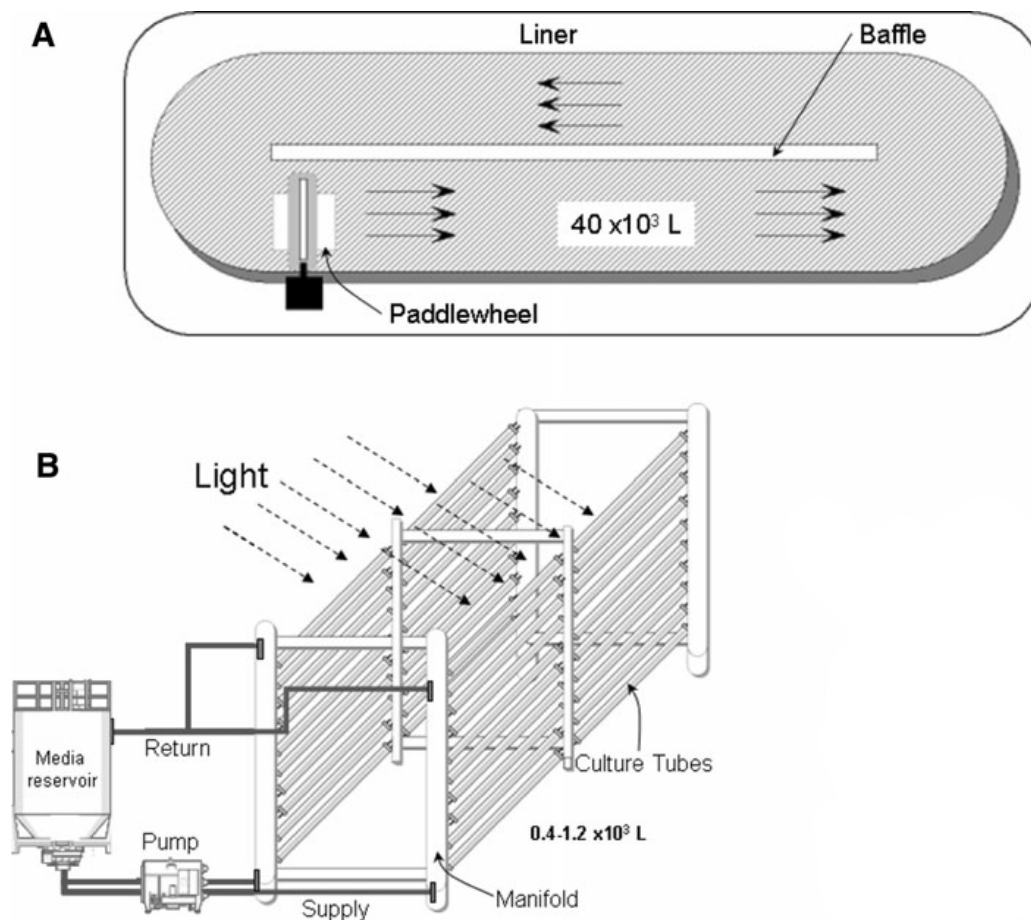
**Figure 5.** Schematic representation of large-scale microalgal cultivation systems currently in use for whole biomass production or production of commodity chemicals. (A) Aerial view of a lined open raceway pond with paddlewheel mixing with volumetric capacity of 40,000 liters. (B) Modular commercial tubular PBR with media reservoir and circulation pump with volumetric capacity of illuminated stage ranging from $0.4-1.2 \times 10^3$. Adapted from (McGinn et al. 2011).

## 1.4.5   *Disadvantages to third generation biofuels*

Although biofuel production from the cultivation of photosynthetic microorganisms is a theoretically viable energy option, there are still issues that need to be addressed before algae fuel can become a sustainable biofuel platform. The biggest hurdles in commercialization of algal fuel continue to be high capital costs and energy requirements. Agitation, harvesting, and drying of the biomass require substantial energy inputs that make these systems inefficient for large-scale production , and non-

competitive with non-renewable sources (Singh and Dhar 2011).  The efficiency of oil production in microalgae is highly dependent on environmental factors, e.g., light intensity, temperature, $CO_2$ concentration, pH, and nutrient levels.   These conditions will need to be fully optimized in order for efficient production of feedstocks for biofuels.

*1.5    Fourth-Generation Biofuels*

The emergence of first, second, third-generation biofuel strategies have currently focused on the production of renewable energy feedstocks to replace petroleum, coal, and natural gas.   Though these approaches are different in their strategies to produce energy feedstocks, they all rely on additional biological or chemical processes to convert them into the final energy product, e.g., fermentation, transesterification, hydrolysis, pyrolysis, gasification, etc.   These additional steps require major energy inputs, making them cost-ineffective and unsustainable in their current forms.   A system that would allow for the direct conversion of $CO_2$ into a "drop-in fuel" (a completely interchangeable substitute for conventional fuel) would have great advantages over current, multi-step, biofuel platforms.    This fuel could be excreted from the microorganism and harvested continuously, without having to destroy the culture or interfere with its growth cycle, substantially reducing production costs.

Recently, photosynthetic microorganisms have been targeted as potential vehicles for the synthesis and secretion of chemicals and high-energy molecules, ushering in the "fourth-generation" of biofuel production.  Microalgae and cyanobacteria already possess the biosynthetic pathways needed to produce biofuels, but are missing crucial components in their "genetic toolbox" to synthesize specific biofuel molecules or produce them efficiently.   Recent advancements in genetic engineering and systems

biology, including transcriptomics, proteomics, and metabolomics can be used to enhance these pathways for biofuel production.

### 1.5.1    *Cyanobacteria:  a model host for biofuel production*

Of all the aquatic photosynthetic microorganisms, cyanobacteria possess the most advantages for biofuel production.  The phylum *cyanobacteria* is a group of photosynthetic bacteria that engage in oxygenic photosynthesis.  Although often referred to as "blue-green algae", the name "cyanobacteria" is more appropriate, since cyanobacteria are prokaryotes and algae are eukaryotes (cyano is the Greek-derived word for "blue").  Cyanobacteria played an important role in the development of Earth's atmosphere, for they are believed to be the first organisms to develop the capacity of oxygenic photosynthesis.  They are also crucial players in the carbon cycle and nitrogen cycle.  Like microalgae, they have simple nutrient requirements, fast growths rates, and are able to synthesize high-energy compounds from $CO_2$, $H_2O$, and light.  Since they are eubacteria, they have circular chromosomes that are much easier to genetically manipulate than eukaryotic microalgae.  This important advantage has given a great deal of attention to cyanobacteria as biofuel-producing organisms through metabolic engineering.

**Table 2. Model strains of cyanobacteria for synthetic biology**

| Strain | Transformation Method | Growth Temp (C°) | Doubling Time (hr) | Notes | References |
|---|---|---|---|---|---|
| *Synechocystis* sp. PCC 6803 | Conjugation, natural, electroporation | 30 | 6-12 | Extensive systems biology datasets are available | Heidorn et al. 2010 |
| *Synechococcus elongatus* PCC 7942 | Conjugation, natural | 38 | 12-24 | A model strain for the study of circadian clocks | Holtman et al., 2005 |
| *Synechococcus* sp. PCC 7002 | Conjugation, natural | 38 | 3.5 | Among the fastest-growing strains known | Xu et al., 2011 |
| *Anabaena* sp. PCC 7120 | Conjugation, electroporation | 30 | >24 | Filamentous, nitrogen-fixing | Ehira et al., 2011 |
| *Anabaena* ATCC 29413 | Conjugation, electroporation | 30 | 14-16 | Produces akinetes under nutrient stress | Mannan & Pakrasi, 1993 |
| *Nostoc punctiforme* ATCC 29133 | Conjugation, electroporation | 30 | 16-20 | Has the largest cyanobacterial genome (~10Mb) | Meeks et al., 2001 |
| *Leptolyngbya* sp. Strain BL0902 | Conjugation | 30 | ~20 | Filamentous, Grows well in outdoor photo-bioreactors in a broad range of conditions | Taton et al., 2012 |
| *Cyanothece* sp. ATCC 51142 | Conjugation | 30 | 16-20 | Separates nitrogen-fixation during light/dark cycles | Alagesan et al., 2013 |

In the last few decades, the revolution of genome sequencing has unveiled the complete chromosome and plasmid sequences of many different cyanobacteria species (which can be viewed at http://genome.microbedb.jp/cyanobase/). A list of model cyanobacteria species commonly used for biofuel production is listed in Table 2. The sequencing of cyanobacteria genomes has allowed researchers to understand the genes and proteins involved in carbohydrate and lipid metabolism, as well as the biosynthesis of amino acids, purines, pyrimidines, cell wall components, and chlorophylls.

Advancements in synthetic biology and biotechnology have also allowed for the development of techniques and tools that enable scientists to synthetically engineer new genes into cyanobacteria, endowing them with the ability to produce biofuels.

*1.5.2   Genetic Engineering*

The introduction of genes that divert fixed carbon into a desired molecule has been the main approach in engineering cyanobacteria for biofuel production.  Ever since the first report of exogenous DNA transfers into a cyanobacterium (Shestakov and Khyen 1970) and the development of recombinant DNA technologies in the 1970's, our knowledge on genetic engineering in cyanobacteria has expanded greatly.  Descriptions of recombinant DNA engineering in cyanobacteria are widespread throughout the literature (Elhai 1994; Golden et al. 1987; Koksharova and Wolk 2002; Vioque 2007; Wolk et al. 1984).  Most genetic manipulations of cyanobacteria were originally performed to gain knowledge into metabolic, genetic, and photosynthetic processes of these cyanobacteria.  These studies are now the backbone of research focused on the genetic manipulation of cyanobacteria for applied purposes.

The primary challenge regarding the insertion of DNA into a cyanobacterium is the efficiency of DNA transfer into the host organism.  Several species of unicellular cyanobacteria, such as *Synechocystis* sp. PCC 6803, *Synechococcus elongatus* PCC 7942, and *Synechococcus* sp. PCC 7002 are naturally competent, and can transport DNA across the cell membrane with moderate to high efficiency. Subsequently, this defining feature makes these cyanobacteria species the most heavily studied for biofuel production.  Other cyanobacterial species, such as *Nostoc punctiforme* ATCC 29133, *Anabaena variabilis* ATCC 29413, and *Anabaena* sp. PCC 7120 have attributes that make them advantageous

hosts, due to their ability to produce hydrogen or fix nitrogen. However, DNA insertion into these strains is problematic, due to their production of restriction enzymes that degrade foreign DNA that enters the cell. Consequently, other transformation techniques, including conjugation and electroporation, have been developed for these strains (Elhai and Wolk 1988). Often, selectable markers are incorporated with the inserted gene, which helps screen for successful transformants. These are often antibiotic resistant markers, which are dependent upon the antibiotic sensitivity of the host as well as the ability of the host to produce the functional protein product of the antibiotic cassette.

Gene expression plays a crucial role in the development of an engineered organism, and several factors must be considered when optimizing gene expression, including promoter strength, copy number, codon usage, and inducible vs. constitutive gene expression. The expression of a gene(s) in a bacterial host can be accomplished using two methods: either by integration of the gene(s) into the host chromosome using homologous flanking regions, or by expressing the gene outside of the chromosome on a self-replicating plasmid (often referred to as a shuttle vector). Each one of these approaches has its advantages and disadvantages. Although genes that are integrated into the chromosome are more stable in the progeny of transformed lines, the known chromosomal loci that can be disrupted without detrimental effects to the cell are limited. Chromosomal "neutral sites" that allow for gene integration and disruption without harmful effects to the cell have been identified in *Synechococcus* sp. PCC 7942, *Synechocystis* sp. PCC 6803, and *Anabaena* sp. PCC 7120. Plasmid-based gene expression outside of the chromosome maintains chromosomal integrity in the host cell. However, these plasmids replicate independently of cell division, creating daughter cells

with inconsistent plasmid copy numbers. Most self-replicating plasmids are derived from native cyanobacterial plasmids with uncharacterized plasmid copy numbers. This makes control of gene expression inside cyanobacteria unpredictable.

### 1.5.3 Biofuels from cyanobacteria

Recently, cyanobacteria have been genetically engineered to produce an assortment of chemicals and fuels. The following sections discuss the wide range of fuels which can be potentially obtained from cyanobacteria.

### 1.5.3.1 Bio-Hydrogen

Hydrogen gas ($H_2$) is an attractive choice for biofuel production, since it is eco-friendly, efficient, and renewable. It can be produced from many strains of cyanobacteria, and can be formed through two major reactions. One is the reversible reaction by the enzyme hydrogenase:

$$2H^+ + 2e\text{-} \leftrightarrow H_2$$
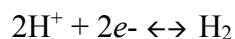
Hydrogen is also formed as a byproduct of nitrogen-fixation by the enzyme complex nitrogenase:

$$8H^+ + 8e^- + N_2 + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$$
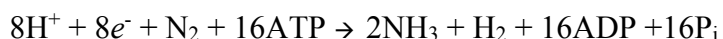
Both of these chemical reactions are closely coupled, as hydrogenases recycle the energy from $H_2$ back into nitrogen-fixation, in order to minimize the energy loss from this process (Bothe et al., 2010). Under $N_2$-limiting conditions, several species of filamentous cyanobacteria are capable of forming specialized cells along the filament, known as heterocysts. These cells protect nitrogenases from irreversible damage by $O_2$ produced from oxygenic photosynthesis, and are also the areas of $H_2$ production. Non-

heterocystous cyanobacteria are also capable of producing hydrogen, although less efficiently. Although the reaction of nitrogenase requires a large amount of ATP, the reaction is irreversible, and allows for $H_2$ to accumulate as high as 20-30% of the total atmosphere (Sakurai and Masukawa 2007). Several reports have reviewed cyanobacteria capable of producing $H_2$, including the genera *Anabaena, Oscillatoria, Calorthrix, Cyanothece, Synechococcus, Synechocystis*, and *Microcystis* (Dutta et al. 2005). Among these strains, it was shown that species of *Anabaena* were able to produce the highest amount of $H_2$, up to 68 $\mu$mol$\cdot$mg chl$^{-1}\cdot$hr$^{-1}$ (Quintana et al. 2011). Many strategies have been employed to increase the production yield of $H_2$ in cyanobacteria, including protein engineering to create nitrogenases that are unaffected by $O_2$ , and increasing the pool of NADPH and ferredoxin for $H_2$ synthesis, as these energy reductants are competed for by other metabolic pathways in the cyanobacteria (Angermayr et al. 2009). For example, deletion of a lactase dehydrogenase in the unicellular cyanobacteria *Synechococcus* 7002 resulted in a five-fold increase in $H_2$ production, compared to the wild-type strain (McNeely et al. 2010).

### 1.5.3.2   *Alcohols and Aldehydes*

Currently, bioethanol is produced by the fermentation of agricultural crops, mainly sugarcane in Brazil and corn in the United States. Cyanobacteria have a potential advantage in ethanol production over traditional crops, in that they ferment naturally without the need to add yeast cultures. The first experiments employed to study the fermenting ability of cyanobacteria was done by Heyer and collaborators (Heyer and Krumbein 1991), where 37 strains of cyanobacteria were studied for the ability to ferment and excrete fermentation products. Of these strains, 16 were able to produce and secrete

ethanol as the main fermentation product. Normally, fermentation is a minor energy pathway in cyanobacteria, and operates only at a minimal level for them to survive. In order to overcome these problems and increase the production of ethanol by cyanobacteria, genetic modification has been applied. The first strain to be genetically engineered for ethanol production was *Synechococcus* sp. PCC 7942, which was transformed with genes coding for pyruvate decarboxylase and alcohol dehydrogenase II from *Zymomonas mobilis* under the regulation of an *rbcLS* operon promoter, alone and in combination with an *Escherichia coli lac* promoter. This resulted in an ethanol production rate of 54 nmol $OD_{730}$ unit$^{-1}$·liter$^{-1}$·day$^{-1}$ (Deng and Coleman 1999). A more recent study showed that these same genes expressed in *Synechocysitis* sp. PCC 6803 under a strong *psbAII* promoter increased the rate of ethanol production to 5.2 mmol $OD_{730}$ unit$^{-1}$·liter$^{-1}$·day$^{-1}$ (Dexter and Fu 2009). A strain engineered by Joule Unlimited (www.jouleunlimited.com), has been reported in the patent literature to secrete ethanol at a rate of approximately 10 mg·L$^{-1}$·h$^{-1}$ (Ducat et al. 2011). This rate greatly outpaces any reports found in the academic literature.

Although ethanol is the world's most common biofuel, longer-chained alcohols and aldehydes are of a particular interest for microbial production, due to their higher energy density and attributes as precursors for liquid fuel additives and industrial solvents. The production of 2,3-butanediol was demonstrated in *Synechococcus* 7942 by over-expressing three enzymes in the cyanobacteria: 1) acetolactate synthase (ALS), which converts two pyruvate molecules from the Calvin Cycle into 2-acetolocatate 2) 2-acetolactate decarboxylase (ALDC), which decarboxylates 2-acetolocatate to produce acetoin, and 3) secondary alchohol dehydrogenase (sADH), which reduces acetoin to

2,3-butanediol. Expressing these genes in *Synechococcus* sp. PCC 7942 resulted in a production titer of 2.38 g/L, which is a significant increase for chemical production from exogenous pathways in cyanobacteria (Oliver et al. 2013). Isobutyraldehyde production was also demonstrated in this strain of cyanobacteria, through genetic modification of the valine biosynthesis pathway (Atsumi et al. 2009). Here, enzymes were used to divert internal pyruvate to the precursor 2-ketoisovalerate (KIV). This intracellular KIV was then converted to isobutyraldehyde by expression of ketoacid decarboxylase (kivD). This resulted in an isobutyraldehyde production rate of 723 mg/L/12 days. The production of 1-butanol by the fermentative coenzyme A (CoA)-dependent pathway was also demonstrated in *Synechococcus* sp. PCC 7942. Here, an ATP-driven malonyl-CoA synthesis pathway was introduced into the cyanobacteria, along with an NADPH-dependent alcohol dehydrogenase. This engineered increase in ATP and recognizable cofactors increased 1-butanol production by 4-fold (Lan and Liao 2012).

### 1.5.3.3  *Fatty Acids*

Efforts into biodiesel production by cyanobacteria have focused on the photosynthetic production of free fatty acids (FFA), which can be converted to biodiesel. Unlike microalgae culturing methods that rely on harvesting, drying and lipid extraction, cyanobacteria can be engineered to secrete FFA into the culture media. In this scheme, the cyanobacteria cells are not converted into biodiesel. Instead, they are the solar factories that synthesize biodiesel precursors. This avoids the costly biomass recovery process, and is one of the defining features of a fourth-generation biofuel.

A study in 2009 experimented with the concept of FFA production in cyanobacteria by genetically modifying *Synechocysitis* sp. PCC 6803 to produce and secrete fatty acids.

An acyl-acyl carrier protein thioesterase gene (*tesA*) with a missing signal sequence peptide was inserted into the cyanobacteria chromosome. Although TesA is normally localized in the periplasmic space, truncating the signal sequence peptide in the TesA protein was found to redirect fatty acid synthesis to FFA secretion in the culture medium. This resulted in an FFA production rate of $197 \pm 14$ mg/L (Liu et al. 2011). To facilitate the passage of FFA across the cell membrane, FFA-producing strains were further modified by knockout of a peptidoglycan assembly protein (PBP2), which weakened the cell wall. The intracellular FFA amount decreased with peptidoglycan weakening, indicating that the disruption of the hydrophilic cell wall facilitates FFA production by decreasing feedback inhibition of enzymes involved in FFA synthesis.

The production of FFA in cyanobacteria is, apparently, not without deleterious effects on the cell. Strains engineered for FFA-production showed reduced photosynthetic yields, chlorophyll degradation, and changes in the cellular localization of the light-harvesting pigments, phycocyanin and allophycocyanin (Ruffing and Jones 2012). These physiological changes are thought to be caused by the production of long-chained, unsaturated fatty acids, which oxidize into a variety of products known to be toxic. These effects may also be caused by an alteration of the chemical composition of thylakoid and cell membranes, as an increase in saturation was observed in FFA-producing strains. These detrimental effects on cell physiology must be analyzed and addressed before additional metabolic engineering strategies are applied to enhance FFA production.
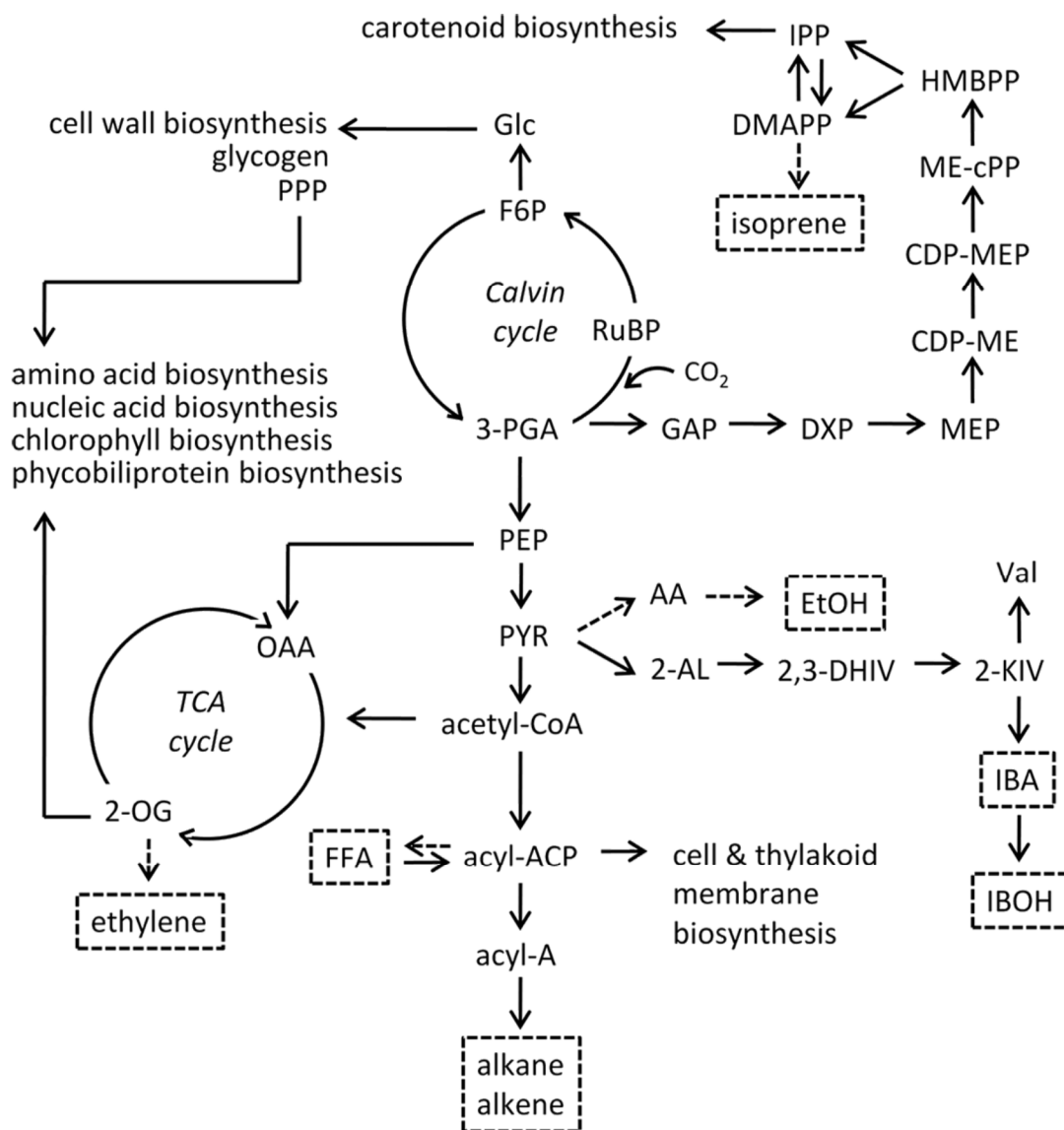
**Figure 6.** Pathways for fuel production in cyanobacteria. Solid arrows indicate natural enzymes or pathways; dashed arrows indicate a required recombinant enzyme. The target hydrocarbon fuels are indicated by dashed boxes. Adapted from Ruffing, 2011. Abbreviations: 2-AL, 2-acetolactate; 2-KIV, 2-ketoisovalerate; 2-OG, 2-oxoglutarate; 2,3-DHIV, 2,3-dihydroxy-isovalerate; 3-PGA, 3-phosphoglycerate; AA, acetaldehyde; ACP, acyl carrier protein; Acyl-A, acyl-aldehyde; CDP-ME, diphosphocytidylyl methylerythritol; CDP-MEP, diphosphocytidylyl methylerythritol 2-phosphate; CO2, carbon dioxide; CoA, coenzyme A; COX, cytochrome oxidase; DMAPP,dimethylallyl diphosphate; DXP, deoxyxylulose-5- phosphate; EtOH, ethanol; F6P, fructose-6- phosphate; FFA, free fatty acid; GAP, glyceraldehyde-3-phosphate; Glc, glucose; HMBPP, hydroxymethylbutenyl diphosphate; IBA, isobutyraldehyde; IBOH, isobutanol; IPP, isopentenyl diphosphate; ME-cPP, ethylerythritol; MEP, methylerythritol-4-phosphate; OAA, oxaloacetate; PEP, phosphoenolpyruvate; PPP, pentose phosphate pathway; PYR, pyruvate; RuBP, ribulose-1,5-bisphosphate; TCA, tricarboxylic acid. Adapted from Ruffing 2011.

*1.5.3.4   Sugars*

Using sugars such as glucose and sucrose as a fermentable feedstock is a particularly attractive biofuel option, and is most commonly employed in Brazil using sugarcane. The production of sucrose using photosynthetic microbes has gained interest in the past few years, with reports of engineered *Synechococcus* sp. PCC 7942 able to secrete sucrose (Ducat et al. 2012), fructose/glucose and lactate (Niederholtmeyer et al. 2010). However, these hydrophilic compounds do not pass through cell membranes efficiently, and therefore need membrane transporters to facilitate movement outside of the cell. Most cyanobacteria lack membrane transporters for hydrophilic organic molecules, therefore, it is crucial for them to be heterogously expressed.

Most heterotrophic bacteria pump protons out across the inner cell membrane into the periplasmic space, making this compartment more acidic than the cytoplasm. Protons then diffuse back into the bacterial cytoplasm to drive ATP synthesis via proton-ATPase. The transport of small molecules in or out of the cell is also coupled to proton transport. In cyanobacteria, a proton gradient is created by the movement of protons across thylakoid membranes instead of the plasma membrane. This, along with the reduction of nitrate and other processes, makes the outside environment more alkaline than the cellular cytoplasm. This reversal in the proton gradient in cyanobacteria compared to other heterotrophic bacteria has been exploited by researchers by expressing sugar and lactate importers from *E.coli* that function as exporters in cyanobacteria (Ducat et al. 2012; Niederholtmeyer et al. 2010). Using numbers from meta-analyses of microalgal production models to calculate a high (volumetric) and low (areal) theoretical yield of

scaled sucrose production from cyanobacteria, Ducat *et al*. theorized a yield of approximately 15-55 metric tons of sucrose per hectare per year could be produced.

### 1.5.3.5   Terpenes

Terpenes are a large and diverse class of long-chained hydrocarbon molecules that have a wide range of applications, from cosmetics, medicines, chemical precursors, and biofuels.  Their naturally volatility also makes them attractive for cyanobacterial production, since an efficient separation from culture biomass could bring down production costs in a large-scale platform.  To this date, only a handful of different terpenes have been produced from engineered cyanobacteria.  The first was isoprene, a volatile $C_5H_8$ precursor that can be polymerized by chemical processes to generate synthetic compounds.  Isoprene has been produced in *Synechocystis* sp. PCC 6803 by expression of isoprene synthase (IspS) from *Pueraria montana* (kudzu) (Lindberg et al. 2010).  Further improvements to product recovery, including a diffusion-based process for isoprene collection, resulted in a production rate of $2\mu g \cdot L^{-1} \cdot hr^{-1}$ (Bentley and Melis 2012).  The monoterpene β-phellandrene has also been produced from *Synechocystis* sp. PCC 6803, by chromosomal integration of a β-phellandrene synthase (β-PHLS), resulting in a production rate of $1\mu g \cdot L^{-1} \cdot hr^{-1}$ (Bentley et al. 2013).

There are two biosynthetic pathways leading to the formation of terpenes; the mevalonic acid pathway (MVA), which functions in the cytosol of eukaryotes and archaea, and the methyl-erythritol-4-phosphate (MEP) pathway, which operates in the bacterial cytosol and plant plastids. Cyanobacteria possess a native MEP pathway, which begins from the condensation of pyruvate and glyceraldehyde-3-phosphate (GAP) and proceeds with a series of reactions, ultimately forming the two main precursors of all

terpenes, isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). These molecules can then be combined to form a wide arrangement of terpene compounds. To bypass the native regulatory components of the MEP pathway in *Synechocystis* sp. PCC 6803, Bentley *et al.* heterologously expressed a foreign MVA pathway in cyanobacteria endowed with IspS, resulting in a 2.5 fold increase in isoprene production (Bentley et al. 2014). Previous studies on *Synechocystis* sp. PCC 6803 have shown that substrates for the MEP pathway may not rely on pyruvate and G3P, but photosynthetic substrates from the pentose phosphate pathway (PPP), which are involved in the formation of IPP and DMAPP (Ershov et al. 2002; Poliquin et al. 2004). Furthermore, reduced ferrodoxin was found to be a critical component to enzyme activity in the MEP pathway, providing yet another direct link to photosynthesis and terpene synthesis (Okada and Hase 2005).

*1.5.4    Metabolic Profiling*

Analysis of carbon partitioning in the metabolic networks of cyanobacteria are a helpful step in evaluating how genetic engineering can impact biofuel production performance, as well identifying pathway bottlenecks and elucidating network regulation in biological systems. It also offers a fascinating insight into the systemic properties of phototrophic metabolism. Once a metabolic reconstruction has been produced (commonly through isotope tracers), the vast array of methods developed by computational systems biology allows for dissecting the interplay between possible metabolic routes and biochemical interconversions. Metabolic profiling has been extensively studied on *Synechosystis* sp. PCC 6803, with a detailed reconstruction of the

comprehensive set of enzyme-catalyzed reactions required to support cellular growth and maintenance (Knoop et al. 2013; Knoop et al. 2010).

Metabolic flux analysis has shown that under photoautotrophic conditions, *Synechocystis* cells exhibit suboptimal carbon efficiency, with a significant amount of carbon lost to the oxidative pentose phosphate pathway (OPP).  Non-stationary labeling experiments confirmed that 100 C-moles of biomass are produced for every 127±2 moles of $CO_2$ fixed, with 16±2 moles of fixed $CO_2$ released via the OPP pathway (Young 2014).  This leads to a significant $CO_2$ loss amounting to 13% of total fixed carbon. The OPP pathway serves a predominantly anabolic role by providing NADPH reductants and intermediates for biosynthetic processes.  Cyanobacteria use NADPH as a key source of electrons for oxidative phosphorylation in the dark, mostly through the oxidation of stored glycogen.  Metabolic flux analysis has shown that residual levels of OPP pathway flux exist even under photoautrophic growth conditions.  This might indicate incomplete suppression of the OPP pathway by regulatory circuits that control the metabolic transition from light to dark.  Photosynthetic carbon partitioning among sugar, terpenoid and fatty acid biosynthetic pathways are not equal, with most of the carbon going towards sugar biosynthesis (80-85%), while fatty acid biosynthesis (~10%) and terpenoid biosynthesis (3-5%), lag far behind that of sugar (Lindberg et al. 2010).

*1.6   Anabaena sp. PCC 7120:  A model organism for photosynthetic N$_2$-fixation*

*Anabaena* sp. PCC 7120 is a sub-species of freshwater cyanobacteria, mostly identifiable by its long filaments of green, blue, spherical cells (also known as trichomes) that can range from ten to hundreds of cells long. It is non-motile, although a number of species of filamentous cyanobacteria are capable of gliding motility using a specialized

filaments known as hormongonia.  One of the defining features of *Anabaena* sp. PCC 7120 is that in the absence of a combined nitrogen source, it harbors the ability to differentiate vegetative cells into large, bulbous cells called heterocysts along the filament in a regularly interspaced pattern.  The heterocysts are the site of anoxic $N_2$-fixation, and hold the key enzyme nitrogenase which fixes $N_2$ gas into ammonia in exchange for photosynthates generated vegetative cells.  Heterocyst formation is viewed as an "endpoint" in differentiation, since heterocysts are incapable of cell division and cannot revert back to vegetative cells when combined nitrogen is introduced into the environment.  A normal doubling time is 20-24 hours, in which newly formed daughter cells remain connected to form the long filaments found in exponentially growing cultures.

### 1.6.1   *Taxonomical classification*

The taxonomical history of *Anabaena* sp. PCC 7120 is complicated, due to the constant re-shuffling into different taxonomical hierarchies after experiments aimed to further classify the cyanobacterium based on morphology.  Originally, all filamentous cyanobacteria were placed into the single order 'Hormongales', but later were separated into two separate orders: the 'Stigonematales', which undergo cell division in more than one plane, and 'Nostocales', which only divide along the axial plane of the filament (as is the case with *Anabaena* sp. PCC 7120).  Later, a number of sub-types that were originally classified in the genus *Nostoc* (namely, PCC 6411, 7118, 7119, and 7120) were later moved to the genus *Anabaena,* which are reserved for heterocystous filamentous cyanobacteria that did not exhibit hormogonium formation, and reproduced only by random filament breakage (Colowick et al. 1988).  The separation of *Anabaena* from the

*Nostoc* genus was later verified in DNA hybridization studies. However, parallel investigations demonstrated that co-inhabitants of the *Anabaena* clade, namely PCC 6411, 7118, and 7119 were capable of producing hormongonia under very specific conditions, while repeated studies found that PCC 7120 could not. It is unclear why *Anabaena* sp. PCC 7120 is the only sub-species in the *Anabaena* genus incapable of hormongonia formation; either this property was lost after repeated sub-culturing, or because an essential element for hormongonia formation was absent in the studies.

### 1.6.2   Optimal growth  conditions

   *Anabaena* sp. PCC 7120 grows well in a variety of freshwater media types, the most popular being are BG-11 and Allen & Arnon media, although many more can be used depending on the specific culturing condition needed (a full list of cyanobacteria media can be found at http://www.cyanosite.bio.purdue.edu/media/table/media.html).  These media contain essential nutrients for growth, including nitrate (in the form of $KNO_3$ or $NaNO_3$), phosphate ($K_2HPO_4$), sulfate ($MgSO_4$), chelator (EDTA), salts, and a variety of metals and micronutrients.  $BG11_0$ media lacks nitrate completely, and is routinely used for the cultivating nitrogen-fixing cyanobacteria or for stimulating heterocyst formation. *Anabaena* sp. PCC 7120 fairs well in neutral pH conditions (7.0-7.5), and buffers are commonly added to media, although not needed.  If nitrate is used as a source of combined nitrogen, alkaline byproducts produced from nitrate reduction can raise the pH to high levels (pH 9-10) during the end of exponential phase, although this increased pH does not often kill the cultures.

   Bubbling cultures with $CO_2$ enriched air (1-5% $CO_2$) is often employed to supplement culture growth, although it is not generally not needed.  Sparging liquid cultures with

$CO_2$ results in the lowering of pH, therefore it is recommended that a buffer is used in those conditions. The biological buffer HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)) is the buffer of choice form most growth conditions with cyanobacteria, and can be used at concentrations from 10-25 mM (personal data).

### 1.6.3    Genome

A restriction map of the *Anabaena* sp. PCC 7120 chromosome was first constructed in 1989 by Bancroft, Wolk, and Oren via endonuclease digestion and pulsed-field gel electrophoresis (Bancroft et al. 1989). Through these techniques, they were able to reveal the length of the chromosome, as well as the existence of three large circular plasmids. A complete, annotated genome was reported in 2001 by Keneko et al. using the "bridging shotgun" method (Kaneko et al. 2001). *Anabaena* sp. PCC 7120 genome consists of a single chromosome (6,413,771bp), containing 5368 protein-encoding genes, four sets of rRNA genes, 48 tRNA genes, and 4 genes for small, structural RNAs. More than 60 genes involved in heterocyst development and nitrogen fixation were found on the chromosome, as well as 195 genes coding for two-component signal transduction systems, nearly 2.5 times the amount found in the unicellular cyanobacteria *Synechocystis* sp. PCC 6803. A more recent study (Lechno-Yossef et al. 2011) has reported approximately 85 genes to be involved in heterocyst development and nitrogen fixation.

### 1.7    The Photosynthetic Apparatus in Cyanobacteria

Photosynthesis is arguably the most important biological process on Earth, supplying all of the organic compounds and a vast majority of energy necessary to maintain life. It is also responsible for almost all oxygen production on the planet, which makes aerobic

life possible. Although it has a few variations, the process is rooted in the conversion of solar energy into chemical energy, which is subsequently provided to all higher life forms. The process was invented 2.8 million years ago by the ancient ancestors of cyanobacteria, and to this day, cyanobacteria, microalgae, and phytoplankton play and enormous role in global photosynthesis and oxygen production on Earth. In photosynthesis, light is captured by antennae complexes and transferred to the photosystems, creating a charge separation between the stroma and lumen. The resulting electrochemical gradient created from the photosystems drives the production of ATP from membrane-embedded ATP synthase, as well as electron carriers such as ferredoxin, NADH, and NADPH. The electron carriers and ATP produced from these reactions drive carbon fixation and other anabolic reactions in the cell. *Overview*

Like plants, cyanobacteria engage in oxygenic photosynthesis, which involves the light-driven oxidation of water molecules and subsequent release of $O_2$ as a waste product. The electrons harvested from water are then driven through the electron transport chain, resulting in the reduction of $CO_2$ into carbohydrate:

$$6CO_2 + 6H_2O \longrightarrow C_6H_{12}O_6 + 6O_2$$
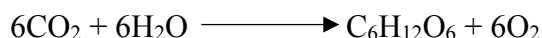
The reaction starts by the capturing light-energy by antenna systems, which transfer excitation energy to the photosynthetic reaction centers in the photosystems. Plants, green algae, and cyanobacteria contain two photosystems, PSI and PSII, which work in a series to create a charge separation across a thylakoid membrane. The two photosystems are connected by another protein complex, cytochrome $b_6f$, as well as a mobile pool of plastoquinones, plastocyanins, and other mobile electron carriers. The excitation energy is passed by the antenna complexes to chlorophyll in the reaction center (RC) of PSII,

which drives electrons from $H_2O$ into the electron transport chain (ETC) while releasing $O_2$ into the stroma and protons into the thylakoid lumen.

Electrons are transferred through a series of reduction-oxidation reactions by electron carriers from PSII through PSI, and are eventually donated to ferredoxin and finally $NADP^+$ to create NADPH.  Protons released from the water-splitting reaction of PSII, as well as the pumping of protons by the cytochrome $b_6f$ complex, creates a positive charge on the lumenal side of the thylakoid membrane.  The electrochemical gradient produced from this charge separation drives the movement of protons out of the lumen through ATP synthase, which utilizes the potential energy of the gradient to phosphorylate ADP to ATP in a process called photophosphorylation.  The reductants (NADPH) and energy (ATP) created from these light-dependent reactions are used to drive almost all anabolic processes in the cell.
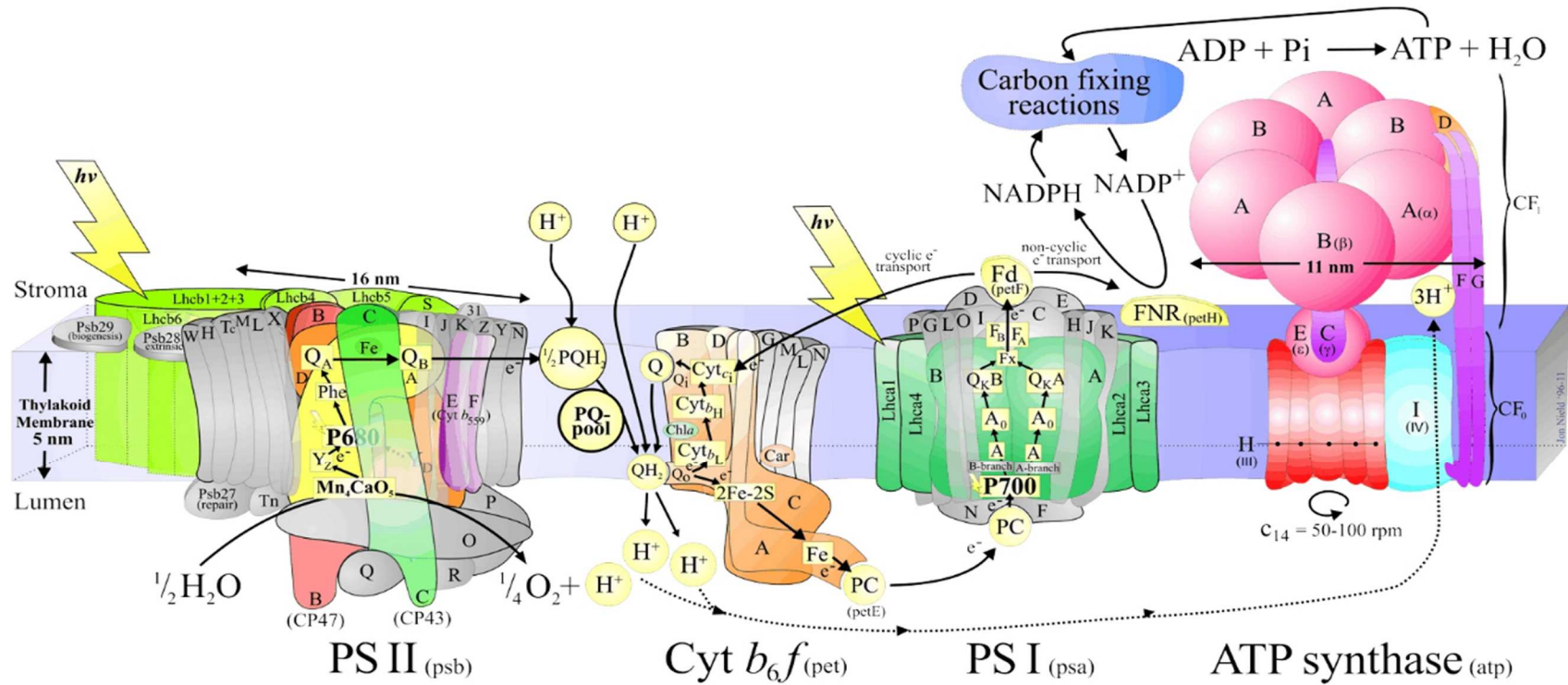
**Figure 7.** The major protein complexes in the photosynthetic electron transport chain. A detailed legend can be found at http://www.queenmaryphotosynthesis.org/nield/.

The next section gives a review of this process in a step-by-step manner, starting with the capture of light energy by phycobilisomes, the oxidation of water and movement of electrons through the ETC to NADPH, and the resulting electrochemical gradient which drives the production of ATP through ATP synthase.

### 1.7.1   Light absorption by chlorophyll

The most common pigments utilized for light capture is the chlorophylls, a class of cyclic tetrapyrroles similar to the heme groups of globins and cytochromes.  Chlorophyll is a highly conjugated molecule, and strongly absorbs visible light that reaches Earth's surface.  Chlorophyll *a* (Chl*a*) is a primary component of intrinsic antennae complexes and the photosynthetic reaction center (RC) of PSII, with absorbance bands at approximately 430-440nm and 680nm (blue and red in the electromagnetic spectrum).  Chl*a* is a very effective photoreceptor, due to its alternating network of single and double bonds among the tetrapyrolle rings that surrounds the Mg center.  It is this unique structure that gives Chl*a* the ability to absorb photons in the visible spectrum and participate in a number of quantum activities.

When a chlorophyll absorbs a photon, the energy excites an electron on the chlorin head group, raising it from its ground state to an excited state.  The excited electron can have several fates:

1)   *Internal conversion*- most organic molecules that absorb light relax back to their ground state, with the electronic energy converted to kinetic energy and dissipated

as heat. Chlorophyll molecules, however, usually relax only to their first excited
state, leaving extra energy to participate photosynthetic reactions.

2) *Fluoresence*- the excited electron decays to its ground state, while releasing a
photon in the process. The emitted photon usually has a longer wavelength than
the photon absorbed, making chlorophyll fluoresce red.

3) *Resonance energy transfer*- the excited molecule directly transfers its excitation
energy to a nearby molecule. Chl*a* molecules packed in the photosynthetic
antenna complex of PSII transfer the excitation energy of photons to the RC,
where it is used to initiate the ETC.

4) *Photoxidation*- a light-excited donor molecule transfers an electron to an acceptor
molecule. Specialized RC chlorophylls absorb excitation energy from resonance
energy transfer and donate their electrons to the ETC. Once oxidized, the donor
oxidizes another molecule and returns to the ground state.

Resonance energy transfer and photoxidation allows for the transfer of light energy to
the RC and the subsequent donation of electrons into the ETC. To make this energy
transfer efficient, Chl*a* is densely packed into photosynthetic antenna complexes in PSII
with appropriate spacing and orientations to each other. Photons absorbed by Chl*a* are
converted into excitation energy and migrate between Chl*a* molecules in the complex
until they are eventually trapped by an RC chlorophyll. *The RC chlorophyll is able to
trap the excitation energy because its lowest excited state has a lower energy than those
of the other antenna chlorophylls in the complex.*

*1.7.2   Phycobilisomes*

The amount of energy absorbed from Chl*a* in the antennae complex of PSII alone is not enough to meet the energy demands of the cell.  Although some proteinaceous pigments aid in the harvesting of light, most photosynthetic organisms are effectively "blind" to wide sections of the electromagnetic spectrum.  Higher plants use light-harvesting complexes (LHCI and LHCII), made up of protein domains packed with Chl*a* and Chl*b* to absorb the needed light energy, while marine photosynthesizers such as diatoms and brown algae use the alternative pigments xanthophyll and Chl*c*. Cyanobacteria and red algae, on the other hand, have evolved a very unique auxillary antenna system that absorbs light missed by Chl*a*, called the phycobilisome (PBS).

The PBS is a water-soluble protein complex anchored to PSII on the stromal side of the thylakoid membrane, and acts as the primary light-harvesting antenna to the photosystem.  The phycobilisome structure consists of a core sub-structure attached to the thylakoid membrane, with a set of peripheral rods protruding from the core in a fan-like pattern, giving the structure a semi-hemidiscoidal shape (Figure 8).  The core base is composed of three cylindrical subassemblies, each of these being a trimer or a double-stacked trimer.  The rods protruding from the base hold a series of stacked protein disks, made up of the pigments phycoerythrin (PE), and phycocyanin (PC), or phyoerythrocyanin (PEC).  These proteins, collectively, are known as the phycobiliproteins (PBPs).  The phycobiliproteins resemble flattened disks that stack like rolls of coins within the core cylinders of the phycobilisome or within the rods of the antennae.  Linker proteins connect the disks together, and aid in energy transfer as well as give the rods structural integrity.
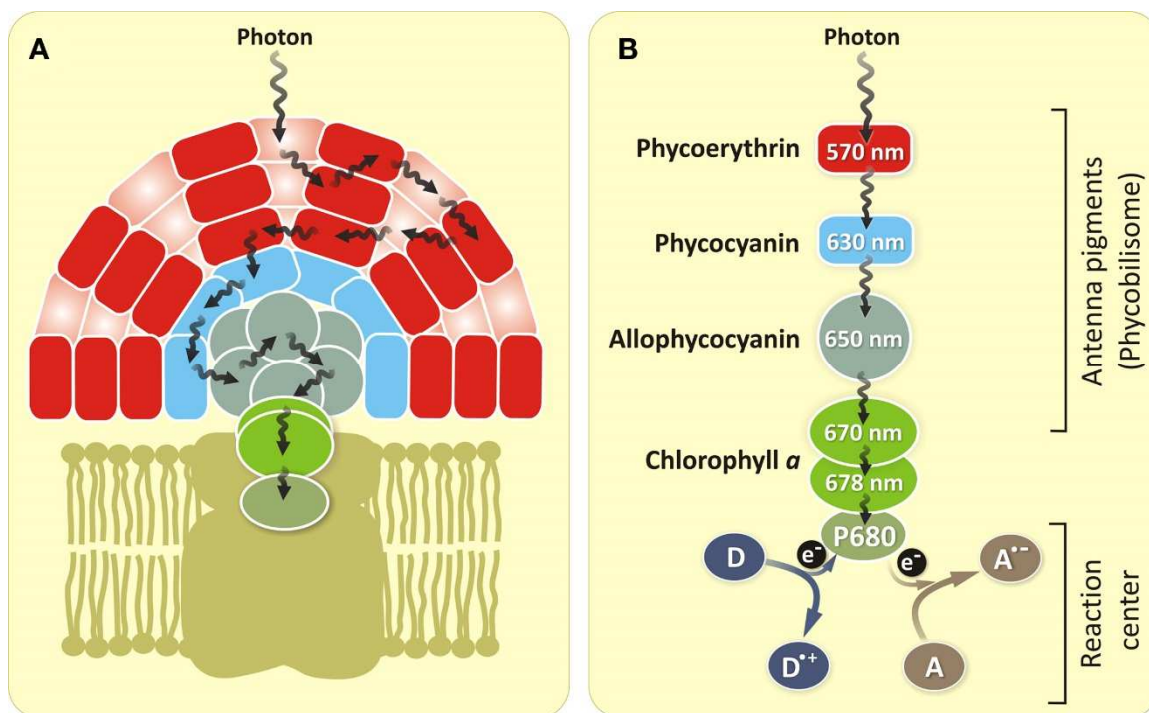
**Figure 8.** Structural organization of the PBS and energy flow from antenna pigments to PSII reaction center. (A) Overall PBS structure, outlining energy transfer from PBS antennas to PSII. (B) The energy absorbed from photons is passed through the PBS proteins PE (red) and PC (blue), until they reach APC in the PBS core (grey). Exited RC chlorophylls (P680) in PSII donate electrons to an electron acceptor A, and the electron vacancy of the RC chlorophylls is filled from an electron donor D. Adapted from (Shevela 2011).

Energy transfer from the antenna pigments to the photosynthetic reaction center is shown in Figure 8B. Each PBP holds covalently attached, open-chain tetrapyrroles called phycobilins, which absorb light in the green, yellow, and orange wavelengths (between 450-660 nm). Phycobilins transfer the energy from the absorbed photons from PE (or PEC) to PC and then to allphycocyanin (APC) in the PBS base. The APC acts as a final energy transmitter, transferring the absorbed energy from the phycobilisome to PSII or PSI reaction centers. Cyanobacteria invest a significant amount of energy into synthesizing PBPs- they make up over 50% of the protein mass in the cyanobacterial cell.

During nitrogen or sulfur starvation, cyanobacteria disassemble phycobilisomes in strict, stepwise-manner as a response to nutrient stress. Disassembly starts begins with the degradation of the terminal rod hexamers, along with the linker peptides, until the entire antenna structure is broken up into individual monomers. The degradation of phycobilisomes takes place within the first 24 hours of nutrient depletion, visually turning cultures from blue-green to yellow-green in a phenomenon known as chlorosis or photobleaching. The newly-released amino acids from the phycobilisomes are then recycled back into metabolism, to aid in cell survival during nutrient stress. Evidence from $NaH^{13}CO_3$ labeling experiments in the cyanobacterium *Arthrospira plantensis* suggests that during nitrate-deprivation, glycogen is synthesized from the carbon skeletons of amino acids derived from protein degradation, including phycobilisomes (Hasunuma et al. 2013). Once favorable conditions are restored, the phycobilisomes are rapidly re-synthesized by the cell.

*1.7.3   Photosystem II: the initiation of electron transport and oxygen evolution*

In oxygenic photosynthesis, excitation energy harvested from phycobilisomes and the antenna complex is used to oxidize water molecules, which serve as the electron donors for the ETC. This oxidation reaction releases protons and $O_2$ in the formula:

$$H_2O \longrightarrow O_2 + 4H^+ + 4e^-$$
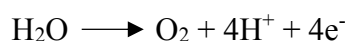
Water oxidation and the ETC is initiated by PSII, a large, oligomeric protein complex embedded in the thylakoid membrane of cyanobacteria (or in higher plants, the lamellae of the chloroplast). Excitation energy is transferred from phycobilisomes and intrinsic antenna complexes into the center of PSII, where it transferred by chlorophylls in the photosynthetic reaction center (RC) until it reaches a specialized set of four RC

chlorophyll molecules, known as P680* (the "*" annotates an excited state P680, while a

"+" annotates a cationic radical). An electron from P680 is then raised to a higher energy

state (P680*) and captured by the electron carrier pheophytin (Pheo), which is associated

with a tightly bound plastoquinone ($Q_A$). Pheophytin then donates the election to the

mobile electron carrier plastoquinone ($Q_B$). After two electron-transfers, the doubly

reduced $PQ_B^{-2}$ picks up two protons from the stroma and leaves the PSII as $PQH_2$.

Electrons are subsequently harvested through the oxidation of water molecules, a process

that is facilitated by the oxygen-evolving complex (OEC) located in PSII. $O_2$ is released

into the stroma as a by-product of this process, while protons are pumped into the

thylakoid lumen and facilitate creation of an electrochemical gradient that serves in ATP

production (discussed in later sections).

### 1.7.3.1    Structure and subunits of PSII

PSII has a homodimeric structure in plants and cyanobacteria, composed of 17

subunits that bind 35-40 chlorophyll and 8-12 carotenoid molecules. Our first glimpse of

the 3D-structure of PSII came from isolated PSII crystals from the thermophilic

cyanobacteria *Thermosynechoccus elongatus*, which were fully active in water oxidation

(Zouni et al. 2001). Similarly, crystallographic structures of PSII from

*Thermosynechococcus vulcanus* have also been reported (Kamiya and Shen 2003).

Recent crystallographic investigations have improved the resolution of the PSII complex

down to 3.0-Å (Loll et al. 2005a), and reprocessing of the same crystals has further

improved the resolution to 2.9-Å (Guskov et al. 2009).

Information gathered from X-ray crystallography experiments from *T.elongatus* have

revealed that PSII consists of 17 subunits, which are named according to their genes:

PsbA to PsbO, PsbU, PsbV, and PsbX. The two major subunits that make up the RC core are PsbA and PsbD (also referred to as D1 and D2, respectively), while the large antennae proteins are from PsbB and PsbC (CP43 and CP47, respectively). Chl*a* molecules bound to PsbB and PsbC subunits absorb light and funnel the resonance energy into the RC. The remaining subunits of note are cytochrome *b-559* with two transmembrane a-helices, PsbE and PsbF. In addition, the subunits PsbO, PsbU and PsbV (cytochrome *b-550*), do not associate with the thylakoid membrane but protrude out from the core complex on the lumenal side. In all, the complex features 14 membrane-intrinsic subunits and 3 membrane-extrinsic subunits.

### 1.7.3.2 Cofactors of the RC

The RC in PSII harbors a number of cofactors that initiate the ETC. They include 6 Chl*a* molecules ($P_{D1}$, $P_{D2}$, $Chl_{D1}$, $Chl_{D2}$, $Chlz_{D1}$, $Chlz_{D2}$), two pheophytins ($Pheo_{D1}$ and $Pheo_{D2}$), two plastoquinones ($Q_A$ and $Q_B$), two tyrosines (TyrZ and TyrD), and the Mn cluster. The chlorophylls $P_{D1}$ and $P_{D2}$ are positioned parallel from one another at a reported Mg-Mg distance of 7.6 Å (Loll et al. 2005b) to 8.2 Å (Ferreira et al. 2004) in *Thermosynechoccus elongatus*. The chlorophylls $Chl_{D1}$ and $Chl_{D2}$ are located on each side of the $P_{D1}$/$P_{D2}$ chlorophylls, with each head group angled ~30° to the membrane plane, and are spaced 10 Å apart from the $P_{D1}$/$P_{D2}$ special pair. These four chlorophylls collectively make up P680, and in its excited state, the electron is delocalized between the four chlorophylls. The other two Chl*a* molecules, named $Chlz_{D1}$ and $Chlz_{D2}$, lie on the periphery of the RC, presumably to transfer excitation energy from the antenna systems to P680. When the excitation energy from the antenna complexes reach the RC, and electron is ejected from P680* and transferred to $Pheo_{D1}$ (presumably through $Chl_{D1}$), and

then to the bound plastoquinone ($Q_A$). Two electrons are sequentially transferred, one at a time, to the mobile plastoquinone ($Q_B$), which then picks up two protons from the stroma and leaves as plastoquinol ($PQH_2$). This electron carrier then passes the electrons to the membrane assembly cytochrome $b_6f$ complex (discussed later).

### 1.7.3.3   *Water-splitting by the oxygen-evolving complex*

The oxidized $P680^+$ cationic radical has an incredibly high redox potential of 1.3 V (Rappaport et al. 2002); the highest biological oxidizing agent found in nature. This characteristic is a prerequisite for water oxidation, orchestrated by the oxygen-evolving complex (OEC) in PSII. The OEC harbors a metallo-oxo co-factor composed 5 metal ions (4 Mn ions and 1 Ca ion), known collectively as the Mn cluster. Two water molecules are bound to the complex and oxidized through the constant cycling of 5 different transition states in the cluster (named $S_0$ through $S_4$). During each transition, an electron is stripped from a water molecule ($S_0$-$S_4$), and an $O_2$ molecule is released during the recovery step ($S_4$-$S_0$). A reactive tyrosine residue (TyrZ) near the cluster serves as an intermediate in transferring electrons from the Mn cluster to $P680^+$, thus reducing the chlorophyll back to its initial state.

The precise structure of the Mn cluster remains an area of speculation. Early crystallography investigations revealed that the four Mn atoms are arranged in a 3+1 organizational model (Zouni et al. 2001), and improved resolution has confirmed this arrangement (Ferreira et al. 2004; Kamiya and Shen 2003; Loll et al. 2005b). Modeling suggests that during the S-state cycle, only one of the Mn atoms binds a water molecule, producing a highly reactive electrophilic intermediate (Ferreira et al. 2004). A tightly bound $Ca^{2+}$ is believed to participate in binding of a second water molecule, which aids in

the formation of the O=O bond in $O_2$ formation. The structure and precise mechanism of the OEC is an area of contentious debate, since all current structures are at medium resolution and are prone to damage by X-ray data collection (Flores and Herrero 2008). Technical improvements in X-ray data collection, along with mutagenesis experiments, will aid in the detailing the structure and mechanism of the OEC.

### 1.7.4   *Electron transport through the cytochrome b6f complex*

Electrons generated from PSII are met by another protein complex during their journey through the ETC. The plastoquinol-cytochrome c553/plastocyanin oxidoreductase (Cyt $b_6f$ complex) is a large, integral protein assembly in the thylakoid membrane that catalyzes the oxidation of $PQH_2$ and subsequent reduction of plastoycyanin (PC) or cytochrome $c_6$. Cyt $b_6f$ therefore acts as an intermediary in electron transfer between PSII and PSI, which drives the pumping of protons from the stroma into the thylakoid lumen. The Cyt $b_6f$ is also a major component of cyclic electron flow from PSI. These reactions result in the conversion of redox potential energy from the oxidation of $PQH_2$ into the transmembrane electrochemical charge gradient needed for ATP synthesis. It is important to note that "between" does not imply a physical location in the thylakoid (although it is in close proximity to both photosystems), but rather a location the overall ETC.

The structure of cyt $b_6f$ shows close resemblance to the bacterial cytochrome $bc_1$ protein (Trumpower 1990) and complex III from the ETC in mitochondria. The complex forms a functional dimer with a total mass of 217 kD, and is composed of 8 subunits and several cofactors (Whitelegge et al. 2002). The four "large subunits" consist of cytochrome *f* (containing a *c*-type heme), cytochrome *b6* (containing two b-type hemes)

the Riesk iron-sulfur protein (containing a 2Fe-2S cluster), and subunit IV.  The remaining four subunits are the small, hydrophobic proteins PetG, PetL, PetM and PetN. Three-dimensional structures of *cyt b6f* have been reported at a resolution of 3.0 Å (Kurisu et al. 2003) revealing the mechanism behind electronic transfer in oxygenic photosynthesis.

### 1.7.4.1    The Q-cycle

PQH$_2$ created from PSII is the primary electron donor in the cyt *b6f* complex, and the liberated free energy from the ETC is used to translocate protons from stromal side of the thylakoid membrane (also referred to as the electronegative (*n*) side) to the lumenal side (electropositive (*p*) side). Interestingly, the electrons released from PQH$_2$ do not follow the same path:  the cyt *b6f* contains a bifurcated ETC that coordinates proton transfer and the oxidation/reduction of PQH$_2$/Q$_B$ in a process known as the "Q-cycle".  This mechanism was proposed by the Nobel-laureate chemist Peter Mitchell in 1975 as a way to balance the stoichiometry between proton transfer and Q$_B$ reduction in the cyt *b6f* complex (Mitchell 1975).

The process is initiated by PQH$_2$ binding to a trans-membrane, plastoquinone binding pocket near the lumenal side of the complex.  PQH$_2$ is oxidized and the electron is sent through a high-potential ETC, facilitated by the Rieske iron-sulfur protein and cytochrome *f*, resulting in the extrusion of a proton into the *p-sid*e of the membrane.  The electron is then transferred to either plastocyanin (PC) or cytochrome $c_6$, the lumenal electron carriers that deliver electrons to PSI.  The second electron is shuttled through a low-potential ETC, facilitated by two b-type hemes on cytochrome $b_6$ and transferred to a second quinone-binding cavity near the stromal side.  The electron is transferred to a Q$_B$

molecule from the quinone pool, along with a proton from the stroma. The result of this process is the uptake of protons into the lumen, driving the electrochemical gradient needed to power ATP synthesis (Cape et al. 2006).

### 1.7.4.2    *Electron carriers of the cyt $b_6f$ complex*

Up to this point, the photosynthetic ETC has been primarily localized to complexes residing in the thylakoid membrane. However, once electrons leave the cyt $b_6f$ complex, they leave the membrane environment and are escorted through the lumen via PC and cyt $c_6$. These small, soluble electron carriers, although structurally different, share almost identical redox potential values (350 mV, at pH 7) and molecular masses (8-10 kD). They are responsible for creating a bridge between cyt $b_6f$ and PSI by shuttling electrons between these complexes and continuing the ETC. Once an electron carrier picks up an electron, it is guided towards PSI through electrostatic or hydrophobic interactions, depending on the species (Flores and Herrero 2008). The electron carrier then docks to PSI on the lumenal side, and transfers electrons to the chlorophyll dimer P700. Once oxidized, the electron carriers re-enters the carrier pool to begin the cycle again.

The utilization of either PC or cyt $c_6$ appears to depend on the species of the organism. Most cyanobacteria and higher plants solely use copper-containing PC, some species of green algae and cyanobacteria utilize the heme-bearing cyt $c_6$, and some use both. The unicellular microalgae *Chlamydomonas reinhardtii* is able to regulate the expression of PC and cyt $c_6$ in response to its environment, with increased expression of PC and lowered levels of cyt $c_6$ when copper is abundant (Merchant and Bogorad 1986). Interestingly, evidence suggests the existence of a third possible carrier in electron

transport between cyt $b_6f$ and PSI, a modified cyt $c_6$ with possible ancient origins previously characterized in *Arabidopsis thaliana* and *C. reinhardtii* (Wastl et al. 2004).

### 1.7.4.3   Plastocyanin

PC belongs to a family of blue copper proteins, known collectively as "cupredoxins", similar to the iron-containing ferrodoxins that also play a role as electron carriers. Although the sequences of PCs from cyanobacteria, algae, and higher plants have diverged, they all have strikingly similar 3-D structures, consisting of an eight-stranded β-barrel made up of two β-sheets.  The PC from *Synechococcus* sp. PCC 7942 contains a Cu(II) ion covalently bonded to a sulfur atom on Cys84, and is further coordinated by two histidines (His87 and His37) and a methionine (Met92) to form a distorted tetrahedron (Inoue et al. 1999).  The copper-containing active site has a large absorption peak at ~600nm, giving the protein a brilliant blue color in its oxidized form. Investigations on the electron-transfer activity of PC in higher plants suggests that two essential surface patches on the protein participate in redox reactions between cyt $b_6f$ and PSI.  One is a hydrophobic patch located around His87, while the other is an acidic patch found near Tyr83 (Gross 1993; Navarro et al. 1997).  These patches are considered essential elements for protein-protein docking and subsequent electron transfer between the PC and the ETC complexes.  Compared to plants, cyanobacterial PCs are smaller and lack the acidic patch, instead harboring a mixture of neutral and positively-charged amino acids (Aitken 1975; Briggs et al. 1990).  There is currently little information on how cyanobacterial PCs engage in electron transfer, although structural analysis of the PsaA and PsaB subunits in PSI reveal a shallow groove which may aid in protein-protein interactions with PC in *Synechococcus elongatus* (Schubert et al. 1997).

*1.7.4.4   Cytochrome c₆*

In environments lacking copper, photosynthetic microorganisms utilize mobile electron carriers containing alternative metal cofactors.  Some cyanobacteria and microalgae utilize the heme-bearing cytochrome $c_6$ to shuttle electrons from cyt $b_6f$ to PSI.  The cyt $c_6$ from the green alga *Monoraphidium braunii* confers two main forms, with the heme-iron taking either Met-His axial coordination (type I) or a Met-His rhombic distortion (type II).  Spectroscopic methods concluded that the two forms are not dependent on the redox state of the heme, but are pH-dependent (Campos et al. 1993). Cyt $c_6$ has also been confirmed in the cyanobacteria *Anabaena* sp. PCC 7119 and *Synechocystis* sp. PCC 6803 (Campos et al. 1993; Hervás et al. 1994; Medina et al. 1997).  The cyanobacterium *T. elongatus* relies solely on cyt $c_6$ to mediate electron transport, as the gene encoding PC is missing in the chromosome (Beißinger et al. 1998). Cyt $c_6$ is believed to be absent from the chloroplasts of higher plants, although a cyt $c_6$-like protein (Atc6) was characterized and proven functional in *Arabidopsis thaliana* (Gupta et al. 2002).

As we have seen, PC and cyt $c_6$ form the primary shuttle of electrons from cyt $b_6f$ to PSI in the ETC.  The evolutionary aspects of these proteins is intriguing, considering they share virtually no sequence or structural similarities, but share identical reduction potentials and molar masses, making them functionally interchangeable (Frazao et al. 1995).  Researchers suggest that cyt $c_6$ was first utilized by the ancestors of photosynthesizing bacteria, considering iron was readily available in the Earth's early atmosphere.

*1.8    Carbon fixation:  the conversion of inorganic carbon into biological material*

Life, though immense and ubiquitous, is dependent on a tiny number of reactions which create all the central precursor metabolites to fuel metabolism.  Fixing inorganic carbon into organic material is an essential starting point for life, and autotrophy (Greek for "self-nourshing") provides a means for creating these building blocks.    All autotrophic cycles rely on a central carboxylation reaction, which siphons inorganic carbon from the environment into the creation complex biological molecules. Heterotrophy consumes the carbon assimilated by autotrophy, oxidizing it for energy and returning it back into the environment to complete the carbon cycle.  The interplay between autotrophy and heterotrophy controls the relative amounts of $CO_2$ and $O_2$ in the atmosphere, as well as stabilizing redox balance on Earth.

Currently, science has illuminated six separate autotrophic pathways, found in organisms in all domains of life.  The Calvin-Benson-Bassham cycle (CBB cycle) is the most significant carbon fixation strategy on Earth, utilized by terrestrial plants, eukaryotic algae, and cyanobacteria during primary production.  Amazingly, nearly half of all global primary production takes place in the photic zones of the ocean, even though total biomass from marine primary producers (cyanobacteria and microalgae) is significantly less than terrestrial plants (Field et al. 1998).  This makes the CBB cycle the most significant carbon fixation strategy on our planet, and its mechanism has been well studied since the mid-20th century (Stitt et al. 2010). Since then, science has uncovered another five carbon fixation pathways, utilized by chemoautrophic bacteria and archaea. Although they are sometimes dubbed "alternative" pathways, this designation underlies our ignorance of their ecological impact and relative carbon contribution, which has only

been studied during the last decade (Raven 2009).  The following sections describe the

biochemical mechanisms behind these autotrophic pathways.

### 1.8.1   Calvin Benson-Bassham cycle: Overview

The pathway architecture behind the CBB cycle was first reported in 1954 by the

research group of Melvin Calvin, and it was long thought that the newly discovered cycle

was the sole carbon fixation pathway on the planet.  In a classical study, Melvin and his

colleagues injected $^{14}CO_2$ into suspensions of the algae *Chlorella* and analyzed samples

at short intervals using 2D-paper chromatography and autoradiography (Bassham et al.

1954).  Detailing the metabolites uncovered by this method lead to the predicted enzymes

that drive the cycle.  The CBB cycle works in joint operation with the photosynthetic

light reactions, which create necessary ATP and reductants to drive the cycle.  The cycle

is mainly composed of three defined parts:  The carboxylation reaction, reduction of

3PGA, and regeneration of ribulose 1,5-bisphosphate. Ribulose 1,5-bisphosphate

carboxylase/oxygenase (RuBisCO) is the primary carboxylating enzyme in the cycle, as it

adds a molecule of $CO_2$ to ribulose-1,5-bisphosphate (RuBP) to form glycerate-3-

phosphate (3PGA), in the first carboxylation step in the cycle.  NADPH and ATP

produced from the light reactions is then used to reduce 3PGA into triose phosphates.

Triose phosphates are then sent through a complex reaction sequence, involving the

rearrangement of the sugar phosphates using enzymes shared between glycolysis

(adolase), gluconeogenesis (fructose-1,6-bisphosphatase, FPBase), the oxidative pentose

phosphate pathway (transketolase, phosphoribose isomerase and phosphoribulose

epimerase), and the enzymes sedoheptulose-1,7-bisphosphatase (SBPase) and

phosphoribulokinase (PRK). This results in the regeneration of RuBP and completion of

the cycle. The triose phosphates created during the CBB cycle serve as the starting material to build sugars, glycogen, nucleotides, and amino acids in the cell.

### 1.8.1.1    Phylogenetic Diversity of RuBisCO

Modern genome sequencing projects have revealed the presence of RuBisCO and RuBisCO-like proteins (RLPs) in organisms in all three domains of life, with some organisms harboring multiple forms (Andersson and Backlund 2008). Phylogenetic analysis of RuBisCO sequences confirm the existence of four separate clades (forms I, II, III, and IV) in nature. Forms I, II, and III perform the carboxylation and oxygenation of RuBP, while RLPs (form IV) appear to lack conserved active-site residues, and therefore cannot participate in those reactions. Two subforms of RLPs are known to perform other functions in the cell, such as the thiosulfate oxidation with a RuBP analog in *Bacillus subtilis*, *Microcystis aeruginosa*, and *Geobacillus kaustophilus* (Saito et al. 2009). Phylogenetic analysis suggests that I, II, III and IV RuBisCO forms are derived from an ancient, methanogenic archaea ancestor (Tabita et al. 2008). As genome sequencing technology continues to improve, more RuBisCOs will be identified from other bacterial lineages, giving us key insights on their functions in those organisms.

### 1.8.1.2    RuBisCO Structure

Three-dimensional structures of RuBisCO is presented in Figure 9. All RuBisCO enzymes from the four major clades are composed of two large (L-) subunits with an approximate molecular weight of 50 kD. All L-subunits share a common structure, composed of an N-terminal domain of approximately 50 amino acids and a larger C-terminal domain with 320 amino acids. Two L-subunits dimerize head-to-tail, forming an $\alpha/\beta$-barrel ($L_2$), with each containing two active sites at the interface of the two L-

subunits. The $L_2$ unit appears to be an evolutionary building block, as diverse organisms use this unit in different arrangements in their specific RuBisCO enzyme. The most abundant form is the form I RuBisCO, which is found in higher plants, algae, and cyanobacteria. In form I RuBisCO, four $L_2$ are arranged to form an $(L_2)_4$ core, with two groups of four small (S-) subunits used to cap the core (each S-subunit weighing approximately 13-17 kD) forming a large $L_8S_8$ molecule. The S-subunits are essential for promoting maximal activity and stability of enzyme complex, but are not strictly required for $CO_2$ fixation (Andersson and Backlund 2008). Interestingly, Forms II and III RuBisCOs found in phototrophic proteobacteria, chemoautotrophs, dinoflagellates, and archaea lack S-subunits altogether, instead adopting only L-subunit complexes arranged from $L_2$ to $(L_2)_5$.

*1.8.1.3   RuBisCO Assembly*

In order to successfully develop strategies to engineer RuBisCO, it will be crucial to understand the complex nature of its assembly and regulation. The mechanism behind the coordination of expression and assembly of the both the L- and S-subunits is still in its infancy. In cyanobacteria and non-green algae, the genes for the L- (*rbcL*) and S- (*rbcS*) subunits are co-expressed in a single operon. In higher plants and green algae, a single *rbcL* gene remains encoded in the chloroplast genome, while multiple copies of *rbcS* can be found in the nucleus. The process behind the regulation and assemblage of RuBisCO S- and L- subunits in higher plants and green algae involves numerous post-translational modifications and interactions with a variety of chaperones (for extensive reviews, see (Gutteridge and Gatenby 1995; Roy 1989; Whitney et al. 2011).
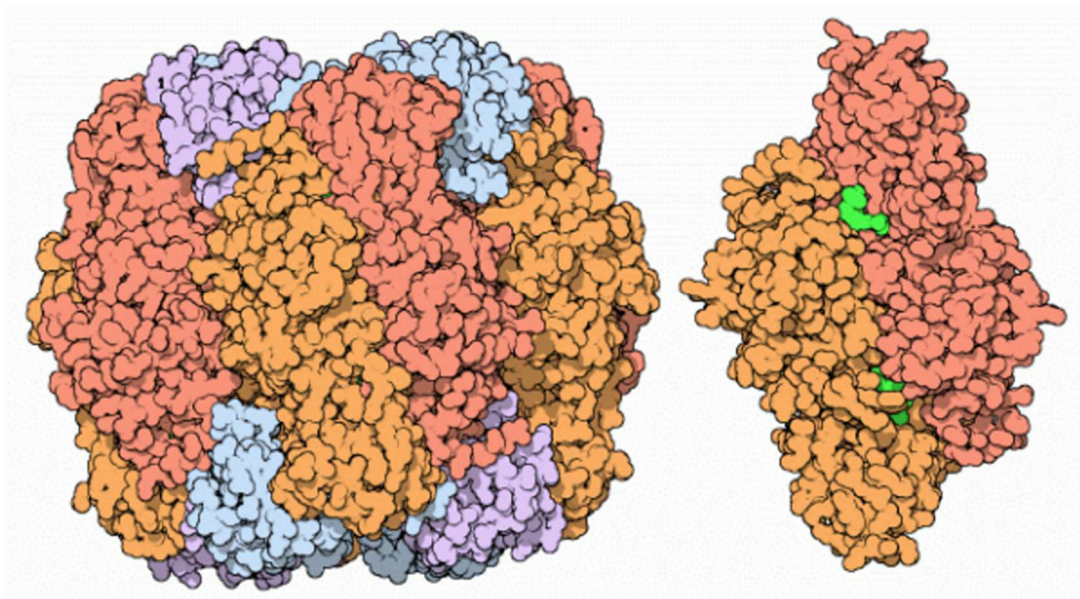
**Figure 9.** Structure of RuBisCO. RuBisCO from spinach (left) and photosynthetic bacteria (right). Image taken from the Protein Data Bank, accessed at pbd101.rcsb.org/motm/11.

In cyanobacteria, GroE and GroL heat-shock proteins were found to play an important role in aiding in the correct assemblage of RuBisCO S- and L- subunits (Gatenby et al. 1985). Another important player in RuBisCO assembly is the gene *rbcX,* found located between the *rbcL* and *rbcS* and co-transcribed along with the RuBisCO subunits in *Anabaena* (Larimer and Soper 1993). Early studies found that RbcX was important for maximal production and functionality of RuBisCO, although its role in RuBisCO activity remained unclear (Li and Tabita 1997). Later, structural evidence from X-ray crystallography showed that RbcX plays the role of a molecular chaperone, facilitating in the formation of the oligomeric $(L_2)_4$ complex (Saschenbrecker et al. 2007). As L-subunits emerge from the ribosome, they are quickly met by GroEL, GroES, and possibly other chaperones assist in folding the newly translated polypeptides. RbcX maximizes the correct assembly of L-subunit octomers, preventing unwanted protein

aggregates in the cell (Tabita 2007). After correct assemblage, the S-subunits cap the L-subunits, forming a complete and functional enzyme.

### 1.8.1.4   *Catalytic Diversity of RuBisCO in Nature*

Characterizing enzymatic characteristics between the RuBisCO families reveal significant catalytic variability, despite sharing the same overall catalytic chemistry. Form I, II, and III RuBisCO parameters from cyanobacteria, green algae, plants, bacteria, and archaea are listed in Table 6. After analyzing the velocity ($V_{CO2}$), affinity for $CO_2$ ($Km_{CO2}$) and specificity between $CO_2$ and $O_2$ between organisms, one trend that becomes apparent is the tradeoff between the velocity rate of carboxylation and affinity for $CO_2$. In other words, organisms that live in high- $CO_2$ environments (such as proteobacteria) or possess $CO_2$-concentrating mechanisms that increase $CO_2$ levels near RuBisCO (C4 plants

| Organism | Form of RuBisCO | $v_{CO2}$ (s$^{-1}$) | $K_m^{CO2}$ (μM) | $S_{C/O}$ | Reference |
|---|---|---|---|---|---|
| **Table 3.** Catalytic properties for different RuBisCO forms. Adapted from sup. data in (Whitney et al. 2011) | | | | | |
| **Cyanobacteria** | | | | | |
| *Anabaena variabilis* | I (green) | n.m. | n.m. | 43 | (Badger 1980) |
| *Synechococcus 7002* | I (green) | 13.4 | 246 | 52 | (Andrews and Lorimer 1985) |
| *Synechococcus 6301* | I (green) | 11.8 | 200 | 42 | (Mueller-Cajar and Whitney 2008b) |
| *Thermosynecococcus BP1* | I (green) | n.m. | n.m. | 51 | Whitney, unpublished |
| **Green algae** | | | | | |
| *Chlamydomonas reinhardtii* | I (green) | 2.1 | 31 | 61 | (Genkov et al. 2010) |
| *Euglena gracilis* | I (green) | n.m. | n.m. | 54 | (Jordan and Ogren 1981) |
| **C4 higher plants** | | | | | |
| *Amaranthus edulis* | I (green) | 4.1 | 18.2 | 77 | (Kubien et al. 2008) |
| *Amaranthus hybridus* | I (green) | 3.8 | 16.0 | 82 | (Jordan and Ogren 1981) |
| *Cynodon dactylon* | I (green) | 3.7 | 21.0 | 89 | (Carmo-Silva et al. 2010) |
| *Flaveria australasica* | I (green) | 3.8 | 22.0 | 77 | (Kubien et al. 2008) |
| *Flaveria bidentis* | I (green) | 4.2 | 20.2 | 76 | (Kubien et al. 2008) |
| *Flaveria kochiana* | I (green) | 3.7 | 22.7 | 77 | (Kubien et al. 2008) |
| *Flaveria trinervia* | I (green) | 4.4 | 17.9 | 77 | (Kubien et al. 2008) |
| *Paspalum dilatatum* | I (green) | 3.4 | 19.9 | 88 | (Carmo-Silva et al. 2010) |
| *Sorghum bicolor* | I (green) | 5.4 | 30.0 | 70 | (Sage and Seemann 1993) |
| *Zea mays* | I (green) | 4.1 | 21.2 | 75 | (Kubien et al. 2008) |
| *Zoysia japonica* | I (green) | 4.1 | 18.5 | 84 | (Carmo-Silva et al. 2010) |
| **C3 higher plants** | | | | | |
| *Arabidopsis thaliana* | I (green) | 3.2 | 14.4 | 81 | Whitney, unpublished |
| *Atriplex glabriuscula* | I (green) | n.m. | n.m. | 87 | (Badger and Collatz 1977) |
| *Chenopodium alba* | I (green) | 2.9 | 11.2 | 79 | (Kubien et al. 2008) |
| *Flaveria cronquistii* | I (green) | 3.1 | 10.8 | 81 | (Kubien et al. 2008) |
| *Flaveria pringlei* | I (green) | 3.1 | 12.0 | 81 | (Kubien et al. 2008) |
| *Helianthus annuus* | I (green) | 2.9 | n.m. | 84 | (Sharwood et al. 2008) |
| *Limonium gibertii* | I (green) | 4.6 | 8.2 | 110 | (Parry et al. 2007) |
| *Nicotiana tabacum* | I (green) | 3.4 | 11.0 | 82 | (Whitney et al. 1999) |
| *Oryza sativa* | I (green) | n.m. | n.m. | 85 | (Kane et al. 1994) |
| *Spinacia oleracea* | I (green) | 3.2 | 12.1 | 80 | (Kubien et al. 2008) |
| *Triticum aestivum* | I (green) | 2.7 | 10.9 | 100 | (Carmo-Silva et al. 2010) |
| **Non-green algae** | | | | | |
| *Cylindrotheca fusiformis* | I (red) | 2.0 | 36.0 | 111 | (Read and Tabita 1994) |
| *Cylindrotheca N1* | I (red) | 0.8 | 31.0 | 106 | (Read and Tabita 1994) |
| *Galdieria sulfuraria* | I (red) | 1.2 | 3.3 | 166 | (Whitney et al. 2001) |
| *Griffithsia monilis* | I (red) | 2.6 | 9.3 | 167 | (Whitney et al. 2001) |
| *Olisthodiscus* | I (red) | 0.8 | 59.0 | 100 | (Read and Tabita 1994) |
| *Phaeodactylum tricornutum* | I (red) | 3.4 | 28.0 | 113 | (Whitney et al. 2001) |
| *Porphyridium* | I (red) | 1.6 | 22.0 | 129 | (Read and Tabita 1994) |
| **Bacteria** | | | | | |
| *Chromatium vinosum* | II | 6.7 | 37 | 41 | (Jordan and Chollet 1985) |
| *Rhodospirillum rubrum* | II | 7.3 | 67 | 12 | (Morell et al. 1990) |
| *Riftia pachyptila* symbiont | II | 1 | 240 | 9 | (Robinson et al. 2003) |
| **Archaea** | | | | | |
| *Methanococcus burtonii* | III | 2 | 130 | 1.2 | (Alonso et al. 2009) |
| *Methanococcus* | III | n.m. | n.m. | 0.5 | (Watson et al. 1999) |

and cyanobacteria) tend to have RuBisCOs with a faster carboxylation rate ($V_{CO2}$) but lower affinity for $CO_2$ (higher $Km_{CO2}$). In contrast, organisms that lack $CO_2$-concentrating mechanisms (C3 crops and most algae) have evolved RuBisCOs with a higher affinity for $CO_2$, but slower carboxylation rates.

*1.8.1.5   Optimizing the kinetic properties of RuBisCO*

RuBisCO is arguably the most important enzyme on Earth, considering almost all carbon in the biosphere is fixed by this enzyme from atmospheric $CO_2$. However, RuBisCO is also one of the most inefficient enzymes in the natural world. Nature has dealt with this dilemma by expressing it in high levels in the cell. RuBisCO is the predominant protein in plant leaves, contributing up to 50% of soluble protein and 20-30% total nitrogen in C3 plants (Evans 1989; Spreitzer and Salvucci 2002). This places a large investment into only a single enzyme, and substantially raises the nitrogen requirements for agricultural crops. Because of this, one of the "holy grails" of engineering photosynthesis is seeking an optimized RuBisCO enzyme Conceptual breakthroughs in the understanding of RuBisCO, as well as advancements in genetic engineering have provoked the idea of re-engineering RuBisCO to enhance specificity and increase its catalytic rate. Though the benefits would be monumental, enhancing the specificity and rate of carbon fixation by engineering RuBisCO has been a challenging endeavor. The next sections review strategies undertaken to optimize this complex and fascinating enzyme.

*1.8.1.6   Optimizing RuBisCO through directed evolution*

The practice accelerating evolution through the use of laboratory-directed protein evolution has been a versatile tool for creating protein variants for useful purposes. The

technique relies on generating a library of protein mutants and screening for mutants harboring a desired phenotype (Arnold and Volkov 1999; Cherry and Fidantsef 2003). Various methodologies to induce random errors in a target gene, including error-prone PCR (Cirino et al. 2003), and gene shuffling (Stemmer 1994), can be employed in this technique. Using directed evolution to screen for RuBisCO variants using expression hosts such as *Escherichia coli*, *Rhodobacter capsulatus*, and the green alga *C. reinhardtii* have been reported (Mueller-Cajar and Whitney 2008a).

*C. reinhardtii* is an attractive host for screening RuBisCO mutants, since the RuBisCO Form I structure ($L_8S_8$) in this green alga is shared by higher plants. Growing mutagenized *C. reinhardtii* cells on media supplemented with acetate (alternative C-source) has proven a reliable way to recover mutants harboring catalytically impaired RuBisCO enzymes (Mueller-Cajar and Whitney 2008a). The identification of catalytically-important residues in RuBisCO has proved invaluable in finding the structural basis for the varying RuBisCO catalytic constants found in nature, and can guide engineering efforts to improve the catalytics of the enzyme. For example, Spreitzer and his colleagues replaced residues of the *C. reinhardtii* S-subunit βA-βB loop with residues from spinach, resulting in a 17% increase in $CO_2$ specificity but lowering the catalytic rate of the *C. reinhardtii* enzyme by almost 50% (Spreitzer et al. 2005). Likewise, amino acid substitutions in the L-subunit of the *Synechococcus* PCC6301 RuBisCO resulted in a 4.9-fold improvement in *in-vivo* folding in *E.coli*, a 12% increase in carboxylation turnover and a 15% improvement in $Km_{CO2}$ (Greene et al. 2007).

Further strategies to improve the reaction kinetics of RuBisCO include the creation of "hybrid" enzymes composed of L- and S- subunits from different organisms.

S-subunits from spinach, *Arabidopsis*, and sunflower were assembled with the L-subunits from *C. reinhardtii* mutants lacking the *rbcS* gene family. Hybrid enzymes showed a 3-11% higher specificity to CO2, while retaining the normal $V_{max}$ for the *C. reinhardtii* RuBisCO. However, despite having adequate levels of this hybrid RuBisCO, the mutants displayed reduced photosynthetic growth and lacked chloroplast pyrenoids (Genkov et al. 2010). Pyrenoids are intracellular mico-compartments associated with elevated $CO_2$ levels, and act as a carbon-concentrating mechanism analogous to cyanobacteria and C4 plants (Giordano et al. 2005). The S-subunit of *C. reinhardtii* is therefore assumed to play a role in targeting RuBisCO for pyrenoid formation (Rawat et al. 1996).

### *1.8.1.7  Is RuBisCO already naturally optimized?*

The trade-off between acquiring higher catalysis and maintaining specificity to $CO_2$ has thwarted most engineering efforts to improve the reaction kinetics of the enzyme. The internal conflict between catalytic rate and specificity is likely a consequence of the enzyme's inability to discriminate between $CO_2$ and $O_2$ as a substrate. It is hypothesized that during the last 25 million years, a decrease in atmospheric $CO_2$:$O_2$ ratios has selected for a carboxylation transition state in which the $CO_2$ moiety closely resembles a carboxylate group. This maximizes the structural differences between the carboxylation/oxygenation transition states, allowing for better differentiation between them. However, increasing structural similarity between the carboxylation transition state and its carboxyketone product strengthens the binding of the intermediate to the active site, slowing down the cleavage and release of the product. In light of this evidence, some argue that attempts to re-engineer RuBisCO to increase productivity are

ultimately futile, since RuBisCO may be already naturally optimized (Tcherkez et al. 2006).

### 1.8.2 Increasing the activity of CBB cycle enzymes

Besides RuBisCO, there are other CBB cycle enzymes that play critical roles in replenishing metabolites to sustain a constant rate of $CO_2$ fixation. Although most CBB cycle enzymes are expressed at high levels, the levels of fructose-1,6-bisphosphatase (FBPase) and sedoheptulose-1,7-bisphosphatase (SBPase) are low compared to other CBB cycle enzymes (Woodrow and Mott 1993). FBPase converts fructose 1,6-bisophosphate (F1,6BP) to fructose-6-phosphate (F6P), creating an exit point for sugars in the CBB cycle to start biosynthesis. In a similar manner, SBPase, hydrolyzes a phosphate from sedoheptulose-1,7-bisphosphate (S1,7BP) to produce sedoheptulose-7-phosphate (S7P), an intermediate in the pentose phosphate pathway (PPP) or into the regeneration of the CBB cycle. Decreasing the activities of these two enzymes *in vivo* compromises the photosynthetic abilities of transgenic tomatoes, suggesting the enzymes are an amenable control point in carbon fixation (Koßmann et al. 1994).

SBP/FBPases are not unique to land plants, but are shared amongst cyanobacteria and algae as well. The SBP/FBPase from the unicellular cyanobacterium *Synechococcus* sp. PCC 7942 is especially intriguing, since it is has the dual capability of hydrolyzing both CBB cycle sugars (Tamoi et al. 1996). Differences in amino acid sequences makes the cyanobacterial SBP/FBPase invisible to plant gene regulation, and an attractive replacement of the native FBPase and SBPase enzymes in the plant. Miyagawa and others demonstrated this by overexpressing the cyanobacterial enzyme in tomato, increasing biomass accumulation and photosynthetic activity (Miyagawa et al. 2001).

Interestingly, they point out that an accumulation of RuBP in transgenic plants increases the activity of RuBisCO through its helper protein, RuBisCO activase (RA), thereby increasing the overall photosynthetic capacity of the plants.

### 1.8.3   3-hydroxypropionate bicycle

The 3-hydroxypropionate bi-cycle (3-HPB) is an alternative carbon fixation pathway, notably utilized by the phototrophic green non-sulfur bacterium *Chloroflexus aurantiacus*. This photoautotrophic bacteria resides in thermal pools and alkaline hot springs, oxidizing hydrogen sulfide ($H_2S$) through anoxic phototrophy. The 3-HPB has a unique overall structure: the cycle is composed of two separate cycles (cycle 1 and cycle 2) conjoined through a set of three core reactions, starting with the carboxylation of two acetyl-CoA from $HCO_3^-$ into two malonyl-CoA, followed by the transformation of two 3-hydroxypropionate to two propionyl-CoA molecules. Each cycle sends one of the two propionyl-CoA molecules down a different metabolic route, ultimately resulting in the production of pyruvate and regeneration of two acetyl-CoA for re-entry into the cycle.

The first step in the series of core reactions is the carboxylation of two acetyl-CoA to malonyl-CoA by a biotin-dependent acetyl-CoA carboxylase (ACC). Each malonyl-CoA is then reduced to 3-hydroxypropionyl-CoA by malonyl-CoA reductase (MCR), and further reduced to propionyl-CoA by propionly-CoA synthase (PCS). At this point, each molecule of propionyl-CoA takes different routes in the bi-cycle. In cycle 1, propionyl-CoA is carboxylated to (*S*)-methylmalonyl-CoA by PCS, followed by isomerization to succinyl-CoA through the successive steps by the enzymes methylmalonyl-CoA epimerase (MCE) and methylmalonyl-CoA mutase (MCM). Succinyl CoA is then converted to (*S*)-malyl-CoA through a series of reaction steps: first, the CoA enzyme

from Succinyl-CoA is transferred to a molecule of (*S*)-malate by the enzyme succinyl-CoA:(*S*)-malate-CoA transferase (MCT), creating (*S*)-malyl-CoA and succinate.  The (*S*)-malate originates from the oxidation of the succinate via the enzymes succinate dehydrogenase (SDH) and fumarate hydratase (FHT).  In the last step of cycle 1, (*S*)-malyl-CoA is split into acetyl-CoA and glyoxylate by (*S*)-malyl-CoA lyase (MCL).  The acetyl-CoA molecule then re-enters for another round of assimilation, while the newly created glyoxylate re-enters cycle two (hence the term bicycle).

In cycle 2, glyoxylate and propionyl-CoA (both created in the cycle 1) are condensed to form a molecule of β-methylmalyl-CoA, catalyzed by the trifunctional MCL.  This is followed by a series of transformation reactions that produce mesaconyl-C1-CoA (with the release of $H_2O$), followed by mesaconyl-C4-CoA, and finally (*S*)-citramalyl-CoA (performed by the enzymes mesaconyl-C1-CoA hydratase, mesaconyl-CoA C1-C4 CoA transferase, and mesaconyl-C4-CoA hydratase, respectively.   In the last step of cycle 2, (*S*)-citramalyl-CoA is cleaved to form acetyl CoA (which re-enters the core reaction steps of the bi-cycle) and a pyruvate, which is released into metabolism.  In all, the entire bi-cycle forms one molecule of pyruvate from three molecules of $^-HCO_3$.

### 1.8.4   *Designing artificial carbon fixation pathways*

As described early, the CBB cycle is used by a vast majority of autotrophs for carbon fixation, and is often limited by the slow catalysis of RuBisCO.   Since engineering strategies to improve RuBisCO function have been marginal, and with evidence suggesting that the enzyme is already naturally optimized, our research efforts may need to expand to other areas involving $CO_2$ capture.   To date, science has identified five separate metabolic pathways apart from the CBB cycle that perform carbon fixation

including the reductive tricarboxylic acid (rTCA) cycle (Evans et al. 1966), the reductive acetyl-CoA (rAcCoA) pathway (Ragsdale and Pierce 2008), the 3-hydroxypropionate cycle (Herter et al. 2002), the 3-hydroxypropionate/4-hydroxybutyrate cycle (Berg et al. 2007), and the recently found dicarboxylate/4-hydroxybutyrate cycle in an anaerobic, autotrophic hyperthermophilic archaeum (Huber et al. 2008). The diversity of carbon-fixation strategies employed by different organisms provides evidence of possible alternative pathways that have yet to be invented by nature. Using a set of approximately 5000 metabolic enzymes, Even et. al created a computational carbon fixation cycle that utilized phosphoenolpyruvate (PEP) synthase and six other enzymes to fix $CO_2$ into glyoxylate (Bar-Even et al. 2010). *In vivo* implementation of an artificial carbon-fixation pathway, as humorously stated by the authors, is analogous to a "metabolic heart transplant" as many immanent problems may arise. Realistically a number of criterion would likely need to be met (optimal gene expression, enzyme activity and stability, regulation, etc.) before the pathway could be truly superior to $CO_2$ pathways arisen through evolution.

## 1.9. *Photorespiration*

RuBisCO is an excellent example of an organic machine that has been unable to keep pace evolutionarily with a rapidly-changing environment. This ancient enzyme first evolved in a highly reduced world, when gases like $H_2$, $H_2S$, and $CO_2$ were in higher atmospheric concentrations than they are today. Aided by the advent of photosynthesis, rising $O_2$ levels in the atmosphere presented a problem for RuBisCO, since it cannot adequately discriminate between $CO_2$ and $O_2$, creating the natural phenomenon of photorespiration (literally, "respiration during photosynthesis). In the photorespiratory

$C_2$ cycle, RuBP and $O_2$ are catalyzed to produce a molecule of 3PGA and the toxic intermediate 2-phosphoglycolate (2-PGL). The 2-PGL is eventually converted back to 3PGA in the $C_2$ cycle, but the energy cost for the cell is significant: it leads to a 25% loss of the carbon in 2-PGL, involves a complex reaction sequence involving over a dozen enzymes and transporters, and releases ammonia in a net loss of nitrogen for the cell.

During the 1970's, the biological function of photorespiration was a topic of contentious debate. Early gas-monitoring experiments showed significant amounts of $O_2$ uptake and $CO_2$ release during photosynthesis in higher plants when grown in ambient $CO_2$, this phenomenon later named 'photorespiration'. However, $O_2$ uptake declined in plants when the concentration of available $CO_2$ increased. These findings suggested that $O_2$ was an alternative acceptor of RuBisCO, along with $CO_2$. Interestingly, this revelation coincided with a mystery that had long plagued researchers of photosynthesis. Radio-labeling experiments with $^{14}CO_2$ had previously shown that high amounts of glycolate and glycine were produced during photosynthesis at ambient $CO_2$ levels. It was suggested that the release of $CO_2$ observed during photosynthesis was the result of the formation of glycolate, as well as other metabolites such as serine and glycine. However, no one at the time knew how glycolate formed when photosynthesis operated in ambient $CO_2$. The finding that RuBisCO reacted with $O_2$ to produce 2-PGL under relatively low $CO_2$:$O_2$ levels suggested that photorespiratory 2-PGL was the source of glycolate. This claim was substantiated using mutant screens of *Arabidopsis* that were viable 2% $CO_2$ but lost viability when transferred to air (high $CO_2$-requiring mutant; *hcr*). Forward genetics approaches revealed that *hcr* mutants harbored mutations in key enzymes in photorespiratory and ammonium re-assimilation pathways. Eventually, the identification

of a mutant that accumulated 2-PGL and was deficient in 2-PGL phosphatase activity effectively ended the debate, heavily substantiating the idea that photorespiration acted to recycle the wasteful side reaction of 2-PGL into useable metabolites for the cell.

For cyanobacteria, the negative effects of photorespiration may be inconsequential to overall photosynthetic efficiency. Cyanobacteria may contain multiple C2 pathways, with *Synechocystis* sp. PCC 6803 containing three distinct pathways (Eisenhut et al. 2008b). It is widely assumed that the CCM prevents most photorespiration, making the 2-PGL recycling an unnecessary task in cyanobacteria. However, a triple-mutant harboring a defect in all three C2 pathways exhibited an *hcr* phenotype, and eliminating total glyoxylate production (the primary metabolite in all three pathways) in these mutants resulted in lethality. The authors argue that under ambient CO2 conditions, photorespiration is still an active mechanism and is not completely suppressed in organisms with a CCM (Eisenhut et al. 2008b).

### 1.9.1   Introducing a photorespiratory bypass in cyanobacteria

Photorespiration is a costly price for doing business with RuBisCO. It is estimated that C3 plants lose 25% of carbon during this process, and can reduce photosynthetic efficiency by 49% (Busch 2013; Zhu et al. 2008). The effects on photosynthesis in cyanobacteria is less understood, but it is assumed that photorespiration is mostly inhibited through the CCM mechanism of carboxysomes, although not completely suppressed (Eisenhut et al. 2008b). This raises the question whether the deleterious energy loss of photorespiration can be mitigated by engineering alternative C2 recycling routes.

The glyoxylate fed into the native C2 cycle in plants, algae, and cyanobacteria shares a common pathway utilized for biosynthesis in the thermophilic anoxygenic phototroph *Chloroflexus aurantiacus*. The 3-hydroxypropionate (3HP) bicycle utilized in this bacteria is also utilized in this bacteria as an alternative $CO_2$ fixation pathway. A series of enzymes from the 3HP bicycle were expressed in the cyanobacteria *Synechococcus* sp. PCC 7942 as a more efficient mechanism to recycle 2-PGL into metabolites for biosynthesis. However, mutants harboring the introduced bypass exhibited no changes in biomass accumulation from the wildtype. The authors argue that insufficient enzyme activity of the introduced 3HP enzymes and photorespiratory suppression by the CCM may be the root cause of the null observations (Shih et al. 2014).

*1.10    Energy conversion during photosynthesis*

Photosynthesis is the underlying mechanism for life on our planet, and is the benefactor of food, fibers, and fuels that support human civilization. The amount of solar power that reaches the Earth is staggering: more solar energy reaches the Earth's surface every hour ($4.3 \times 10^{20}$ J) than is consumed on our planet in one year ($4.1 \times 10^{20}$ J). Despite this overwhelming quantity, solar energy readily diffuses as soon as it reaches the Earth's surface, making the conversion efficiency of solar radiation into chemical energy a crucial component of photosynthesis. To capitalize on solar conversion, photosynthetic organisms have evolved specialized mechanisms to minimize energy loss during the conversion of photons into chemical energy. This includes the development of C4 photosynthesis, phycobilisome antennae systems, light harvesting complexes, and C2 photorespiratory pathways. However, these mechanisms utilized by photoautotrophs to minimize energy loss may soon fail to meet the energy demands of a growing human

population. Diminishing land resources coupled with an exponentially increasing human population, modern agriculture will need to design and employ specialized strategies to increase the efficiency of photosynthesis in crop plants, microalgae, and cyanobacteria.

The next sections overview the efficiencies of the photosynthetic apparatus at different levels of energy conversion, identify major areas of energy loss, and offer strategies to increase efficiency in those areas.

### 1.10.1   Light saturation during photosynthesis

The Z-scheme of linear electron transport in oxygenic photosynthesis requires the cooperation of two photosystems (PSI and PSII) to drive the transfer of electrons from water molecules into molecule of fixed $CO_2$. Theoretically, this involves a minimum of 8 photons per molecule of $O_2$ generated, or carbon dioxide converted into biomass. Experimental measurements with the unicellular alga *Chlorella vulgaris* have found the actual minimum requirement is equal to 9.7 photons absorbed per $O_2$ evolved (equivalent to 0.103 $O_2$ evolved per photon absorbed). Similar results are found in higher plants, with about 9.3 photons utilized per $O_2$ evolved. Averaging the number of photons absorbed per $O_2$ evolved between plants and algae (9.5 photons per $O_2$ evolved), we find that the absorbance efficiency is relatively high, around 84%. This high absorbance efficiency is a testament to the photosynthetic architecture's ability to efficiently funnel light energy into the electron transport chain.

It is important to note that the observed high absorbance efficiencies of photosynthesis remain high only under light-limiting conditions, i.e., when incident light intensities and the light reactions are the rate-limiting step in overall productivity, rather than the rate of carbon fixation in the CBB cycle. In the natural world, ambient sunlight

can reach in excess of 2500 $\mu E \cdot m^{-2} \cdot s^{-1}$ when the sun has reached peak apex in the sky. This amount of radiation is exceedingly higher than the requirement for photosynthesis, resulting in the dissipation of energy by either fluorescence or heat. The "light saturation of photosynthesis" can be easily observed by measuring the amount of $O_2$ evolved (or $CO_2$ fixed) as a function of incident light exposure (Figure 12). As irradiance increases from 0-500 $\mu E \cdot m^{-2} \cdot s^{-}$, $O_2$ evolution increases in a linear fashion. It is during this linear range that photon absorption is at its highest efficiency. At higher light intensities, however, the rate of $O_2$ generation begins to plateau, reflecting that the light-harvesting apparatus has become so saturated with photons that there is no continued increase in $O_2$ generation (hence, $CO_2$ fixation). This is known as the light saturation point of photosynthesis, or *Pmax*. The light saturation point reveals other processes that ultimately limit productivity in the photosynthetic cell, e.g., the slow rate of carbon fixation by RuBisCO, bottlenecks in the ETC, and slow turnover rate of the Mn-containing oxygen-evolving complex.
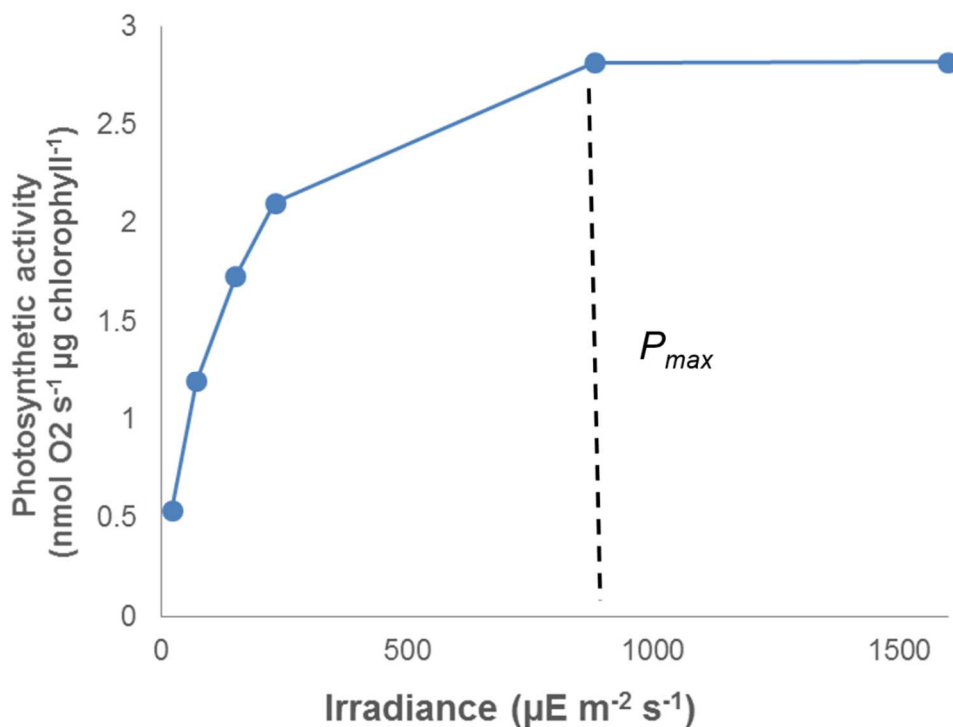
**Figure 12.** The light saturation curve of photosynthesis for the cyanobacterium *Anabaena* sp. PCC 7120. The lowest light intensity at which no increase in $O_2$ evolution is witnessed is known as the light saturation point, or Pmax (figure from personal data).

### 1.10.2   *Photosynthetically active radiation*

Although a large amount of solar radiation reaches the Earth's surface annually, not all of it can be captured by photoautotrophs. As solar radiation reaches the planet's surface, large portions of the spectrum are filtered out by the Earth's atmosphere. Gas molecules such as $CO_2$, $H_2O$, methane, and nitrous oxides absorb light energy in the infrared wavelength ($\lambda$) ranges 900-950, 1100-1150, 1350-1450, and 1800-1950 nm, respectively. Short wavelength radiation (under 400nm) is absorbed by ozone and $O_2$, while wavelengths over 740nm contain insufficient energy to drive photosynthesis. The resulting wavelength range of 400-700nm, referred as photosynthetically active radiation (PAR) is defined as the wavelength spectrum of solar radiation that can be utilized by

photosynthesis to convert photons into chemical energy. PAR consists of only about 48.7% of the electromagnetic spectrum, meaning that 51.3% of the spectrum is effectively useless for photosynthesis.

### 1.10.2.1 *Energy loss from reflectance*

Phototrophic organisms cannot absorb all PAR wavelengths, due to the poor absorbance of chlorophyll molecules in the green spectrum (500-540nm). This wavelength range is therefore reflected, and is why most plants, algae, and cyanobacteria appear green and not black. This reflection accounts for ~10% loss of PAR, therefore ~90% of PAR can be utilized by photoautotrophs to engage in photosynthesis.

### 10.2.2 *Energy loss from relaxation states of chlorophyll*

The Planck-Einstein relation connects photon energy $E$ with its wavelength $\lambda$ in the classic formula $E = hc/\lambda$ (with $h$ denoting Planck's constant and $c$ denoting the speed of light in a vacuum). The equation reveals that photon energy is proportional to photon wavelength, i.e., the shorter the wavelength of a photon, the higher it's energy. This means that the energy of a blue photon (400nm) is 75% higher than the energy of a red photon (700nm). However, higher excitation states of chlorophyll quickly relax, and the resulting resonance energy transfer from chlorophyll to the reaction center is equal to the energy of a red photon, regardless of the initial absorbance energy of the photon. In other words, photosynthesis cannot store the extra energy of a blue photon. This accounts for a ~6.6% energy loss during energy transmission from chlorophyll to the reaction center, which is mainly released as heat.

*1.10.3   Energy loss from carbon fixation in the CBB cycle*

The energy loss from carbon fixation in the CBB cycle must also be accounted for

when totaling the energy losses of photosynthesis.  Assimilating one $CO_2$ molecule to

two 3-PGA (for simplicity, referred to as a 6-C molecule) while regenerating one 5-C

molecule of ribulose-1,5-bisphosphate (RuBP) requires the input of three ATP and two

NADPH.  During electron transport, 4 photons are needed to produce 1 NADPH, while

releasing 6 protons into the thylakoid lumen to create an electrochemical gradient.  Since

2 NADPH are needed to fix one 6-C molecule, a total of 8 photons can drive the

reduction of 2 NADPH, while releasing 12 protons into the thylakoid lumen.  Since 4

protons are needed by ATP synthase for every 1 ATP molecule synthesized, the 8

photons absorbed are enough to convert 1 $CO_2$ into one 6-C molecule.  Eight red photons

represent 1388 kJ energy, while a single unit of carbohydrate (1C; or one-sixth mole of

glucose) contains 477 kJ energy.  This means that the maximum energy capture from 8

red photons to 1 $CO_2$ fixed is 65.6%, reflecting a 34.3% energy loss from photon capture

by the RC to the synthesis of carbohydrate in the CBB cycle.

*1.10.4   Energy loss through photorespiration*

RuBisCO's ability to oxygenate RuBP during photorespiration is another level of

energy loss during photosynthesis.  The oxygenation of RuBP by RuBisCO produces 2-

PGA, a toxic byproduct that is converted into glycolate by the $C_2$ photorespiratory cycle.

In this cycle, two molecules of 2-PGA are converted into two glyoxylate, which feeds

into the $C_2$ cycle to regenerate 3-PGA into the CBB cycle.  From an energy standpoint,

photorespiration is a costly process, requiring both ATP and reductants as well as

releasing newly fixed $CO_2$.  Organisms that engage in $C_3$ photosynthesis, including

cyanobacteria, algae, and some plant species must contend with the energy losses of photorespiration, while C4 organisms eliminate almost all photorespiration by concentrating $CO_2$ around RuBisCO in bundle sheath shells . Here we consider the energy costs of photorespiration from to C3 photosynthesis.

To calculate this energy cost, Zhu et al. defined parameters governing photorespiration and developed a model to quantitate energy loss, defined as $d_{pr}$. The frequency of carboxylation/oxygenation events by RuBisCO is dependent on two primary variables: temperature and relative concentrations of $CO_2$ and $O_2$ around RuBisCO. We can define the ratio of RuBP oxygenation to carboxylation as $\phi$, and compute the amount of ATP associated with one carboxylation and $\phi$ oxygenation as $3 + 3.5\phi$, where 3.5 represents the net ATP use of the CBB cycle and the C2 cycle during one oxygenation event. The number of ATP needed per carboxylation is $(3 + 3.5\phi)/(1 - 0.5\phi)$, with 0.5 representing that one $CO_2$ is produced for every two oxygenation events. The decrease in photosynthetic efficiency from photorespiration can be calculated through the equation:

$$d_{pr} = 1 - \frac{3(1 - 0.5\phi)}{3 + 3.5\phi}$$

with the ratio of oxygenations to carboxylations ($\phi$) defined as:

$$\phi = \frac{O}{C \times \tau}$$

with $O$ and $C$ representing the intracellular levels of $O_2$ and $CO_2$, respectively, and $\tau$ the specificity of RuBisCO to $CO_2$ (more precisely, the ratio of probabilities of carboxylation to oxygenation). As temperature increases, the solubility of $CO_2$ decreases relative to $O_2$. Likewise, the specificity of RuBisCO to $CO_2$ also declines as temperature increases. The

value of τ can be derived from the calculations of Bernacchi et al., demonstrating that the value of τ at 30°C is 70. Given the current concentrations of atmospheric $CO_2$ and $O_2$ at 380ppm and 210000ppm, respectively, photosynthetic efficiency is decreased by 49%, or about 6.1% of the original intercepted solar radiation.

It is important to note that losses from photorespiration are probably lower for microalgae and cyanobacteria, considering these aquatic photosynthetic microorganisms possess a carbon concentrating mechanism (CCM) to increase $CO_2$ levels around RuBisCO (Melis 2009).

### 1.11   Glycogen metabolism in cyanobacteria

Cyanobacteria have evolved to live in rapidly fluctuating nutrient environments. Long periods of nutrient deprivation can be punctuated with momentary conditions where nutrients are in excess. To deal with this adversity, cyanobacteria utilize mechanisms to store carbon, nitrogen, and phosphate when the environment is plentiful in these elements. Fixed carbon is an essential molecule that is stored in reserve polymers in the cytosol of the cell. Storing carbon reserves in a polymers gives the cell an accessible energy and carbon source, while minimizing interference with critical cell processes. Carbon reserves are often called a "metabolic sink" referring to a cellular reservoir where a nutrient is stored until it is degraded by some other metabolic process.

Glycogen is a water-soluble polyglucan composed of linear chains of repeating α-D-glucose units, with by (1-6)-α-D-glucosidic branches. It is the major carbon storage material utilized by all eubacteria (including cyanobacteria), archaea, and animals, and is analogous to starch in algae and plants. The binary, symmetrical branching of a single glycogen polymer can lead to the formation of highly complex spherical structures,

observed as dark inclusion bodies embedded in the thylakoid membrane of the cell (Figure 10).
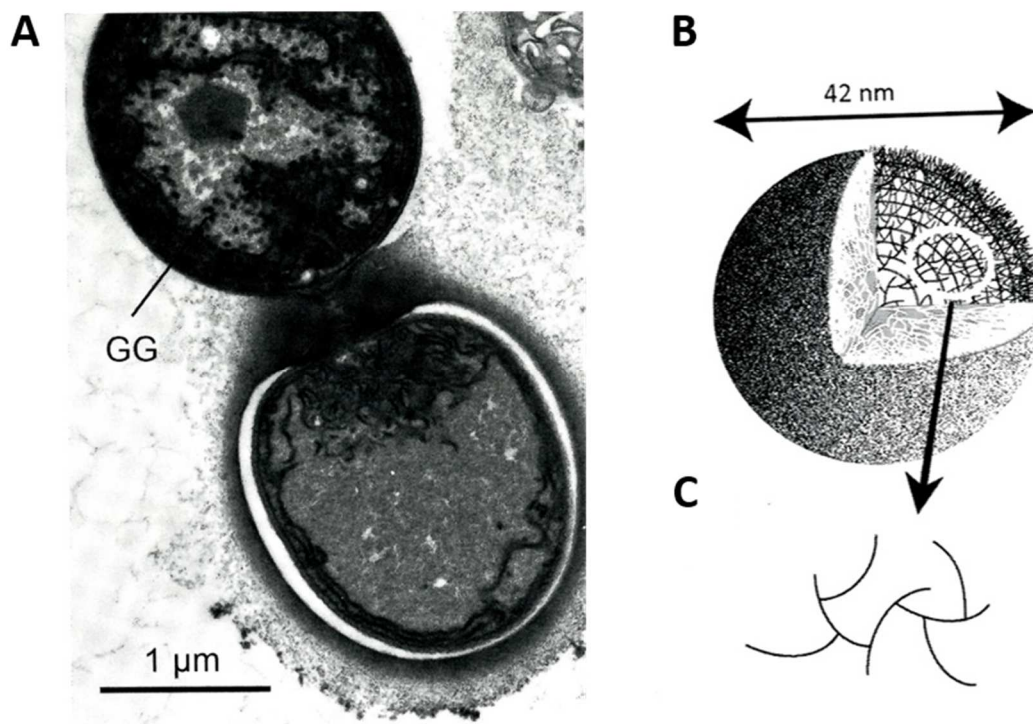


**Figure 10.** Transmission electron micrograph (A) of a filament of *Anabaena* sp. PCC 7120, with glycogen granules (GG) shown in the vegetative cell. Granules are absent in the heterocyst. Drawing of spherical glycogen granule, (B), composed of numerous branching glycogen polymers, which limit the size of the granule to a maximum of ~42 nm. Adapted from (Flores and Herrero 2014).

In cyanobacteria, glycogen metabolism is a regulated process, and is tightly coupled to the light and dark reactions of photosynthesis. During lighted conditions in the presence of abundant carbon, cyanobacteria readily synthesize carbon, which is subsequently respired in the dark when the light reactions of photosynthesis are inoperable. Glycogen is not only used as an internal carbon store, for evidence suggests it plays a larger role in maintaining homeostasis when other key elements (nitrogen, phosphate) are lacking in the surrounding environment. The following sections will

review the process behind glycogen synthesis and degradation, regulation, and roles in

maintaining cellular homeostasis during unbalanced growth conditions.

*1.11.1 Glycogen synthesis*

The first committed step in glycogen synthesis is the addition of an ADP moiety to a

molecule of glucose-1-phosphate (G1P) derived from glucose biosynthesis to produce

ADP-glucose. This glycosyl donor is synthesized by the enzyme ADP-glucose

pyrophosphorylase (AGPase, GlgC) (EC 2.7.7.27), and is sometimes referred to as

glucose-1-phosphate adenylyltransferase. Protein fractions isolated from the glycogen

granules of *Cyanothece* sp. ATCC 51142 show that this enzyme is weakly associated

with the granules, and is therefore likely localized in the cytoplasm (Schneegurt et al.

1997). AGPase catalyzes the formation of ADP-glucose through the following reaction:

$$ATP + \alpha\text{-D-glucose-1-phosphate} \longrightarrow ADP\text{-glucose} + PP_i$$

AGPase activity has been verified through mutagenizing the *agp* ORF in a number of

cyanobacteria strains, including *Synechocystis* sp. PCC 6803 (Gründel et al. 2012; Wu et

al. 2000) *Synechococcus* sp. PCC 7942 (Suzuki et al. 2010), and *Synechococcus* sp. PCC

7002 (Davies et al. 2014). Furthermore, in vitro assays have been performed on the

AGPase from *Anabaena* sp. PCC 7120, revealing its activity in polysaccharide synthesis

(Iglesias et al. 1991; Kawano et al. 2014). Enzymes besides AGPase have been reported

to synthesize ADP-glucose in cyanobacteria. Sucrose synthase (SusA) catalyzes the

cleavage of sucrose to ADP/UDP-glucose and fructose in *Anabaena* sp. PCC 7120

(Cumino et al. 2007; Curatti et al. 2008). Besides glycogen synthesis, ADP-glucose

participates in the production of glucosylglycerol, which acts as a compatible solute during osmotic stress (Hagemann and Erdmann 1994).

In the next step of glycogen synthesis, the ADP-glucose donor is attached to a growing glycogen polymer via the enzyme glycogen synthase (GlgA) (EC 2.4.1.21). Specifically, an ADP-glucose is added to the non-reducing end of an existing $\alpha$-(1,4)-glucan chain through an $\alpha$-(1,4)-glycosidic bond, which expands the glycogen chain linearly.  Genomic sequencing studies show most cyanobacteria contain at least one GlgA, with many strains harboring two unique isoenzymes (GlgA1 and GlgA2). Comparative phylogenetic analyses also reveal that GlgA1 is unique to most eubacteria, while the GlgA2 isoform holds a closer sequence similarity to plant and algal starch synthases (SS3 and SS4) that are predominantly localized in plastids.  Single knockouts of each isoenzyme were still able to accumulate similar amounts of glycogen as the wild-type in *Synechocystis* sp. PCC 6803, demonstrating that each enzyme is active in glycogen synthesis (Gründel et al. 2012).  However, inactivating both isoenzymes results in the abolishment of glycogen synthesis in *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002 (Gründel et al. 2012; Xu et al. 2013).  Therefore, the activity of both GlgA and AGPase is critical for glycogen synthesis in these cyanobacteria.

Glycogen confers a branching confirmation, performed by the activity of the branching enzyme GlgB (EC 2.4.1.18), on the last step of glycogen formation.  This enzyme catalyzes the addition of $\alpha$-(1,6)-linkages to existing $\alpha$-(1,4)-glucan chains, transferring anywhere from 6-8 glucose segments at one time (Boyer and Preiss 1977). Creating $\alpha$-(1,6)-linked branches on a linear glycogen polymer increases the number of

non-reducing ends used for polymerization. Most cyanobacteria harbor only one branching gene (*glgB*), similar to plants in starch formation. The activity of GlgB has been established in *Synechocystis* sp. PCC 6803 by inactivating the gene using homologous recombination, resulting in loss of glycogen branching and an overall decrease in water-soluble glucan (Yoo et al. 2002).

*1.11.2   Glycogen degradation*

Unlike glycogen synthesis, the process of depolymerizing a glycogen polymer into individual sugar nucleotides is less understood in bacteria. Most of our information on glycogen breakdown comes from studies in *Escherichia coli* (Deschamps et al. 2008). Currently, studies have recognized the activities of three separate enzymes, each catalyzing a successive step in depolymerization. The process involves 1) the degradation a of glycogen chain from the non-reducing end, releasing glycosyl nucleotides into metabolism. The chain is degraded until four glycosl nucleotides remain on the polymer, connected by an α-(1-6)-glycosylic linkage to another polymer 2) breaking the α-(1-6)-glycosylic linkage that connects the degraded glucan branch from the main polymer, and 3) transferring the degraded glucan branch to the non-reducing end of another polymer and repeating the degradation process in step one.

The first step involves the sequential breakdown α-(1-4)-glycosylic linkages with orthophosphate on the non-reducing end of a glucan branch, releasing α-D-glucose-1-phosphate. Depolymerization continues until the last four glucose residues remain on each branch. This step is catalyzed by the action of a glycogen phosphorylase (GlgP) (EC 2.4.1.1). Cyanobacteria contain two closely related glycogen phosphorylases, named GlgP and AgpA (Deschamps et al. 2008). Most cyanobacteria species contain either

GlgP (e.g. *Synechocystis* sp. PCC 6803) or AgpA (*Nostoc Punctiforme* PCC 73102), while harboring several homologues of each respective enzyme in their genome. Genomic analyses reveal some cyanobacteria even harbor both GlgP and AgpA homologues (e.g. *Synechococcus* sp. PCC 7002). The reasoning behind this mixed distribution of glycogen phosphorylases across different cyanobacteria species is not completely understood.

In the second step of depolymerization, the remaining four glucosyl residues on the α-(1,6)-glucan branch are further degraded by the action of an isoamylase-type debranching enzyme (DBE) named GlgX (EC3.2.1.68). This enzyme is part of a class of glycosyl-hydrolyases that cleave the α-(1,6)-linkage that connects two glycogen chains, releasing the four-residue glucan polymer (also known as a maltotetraose) previously degraded by the glycogen phosphorylase. Maximum-likelihood phylogeny predicts three different clades of DBE enzymes: a cyanobacterial-type DBE, eubacteria-type DBE, and a plant-like DBE. Similar to glycogen phosphorylases, the distribution of GlgX across cyanobacteria are complex, with some cyanobacteria possessing only the eubacteria-type (e.g. *Nostoc Punctiforme*), the plant-type (e.g. *Synechocystis* sp. PCC 6803), and some lacking a *glgX* gene altogether (e.g. *Anabaena* sp. PCC 7120, *Thermosynechococcus elongatus* BP-1 (Suzuki et al. 2013). The enzyme responsible for α-(1,6)- debranching in these cyanobacteria species has yet to be elucidated.

The last step in depolymerization involves the transfer of the maltotetraose to the end of glycogen polymer by a α-(1,4)-glucanotransferase (α-GTase, MalQ) (E.C. 2.4.1.25), forming an elongated oligosaccharide chain. The newly-formed maltotetraose-oligosaccharide complex is then degraded by either GlgP or AgpA (described above).

All of the annotated cyanobacterial genomes harbor at least one malQ homologue (Nakao et al. 2009).

### 1.11.3  *Light, carbon, and nitrogen availability influences glycogen levels in cyanobacteria*

In cyanobacteria, glycogen constitutes on average 1-5% w/w dry weight in the cell, but can vary as a function of light and nutrient availability (Boyer and Preiss 1977). During lighted conditions when cells are photoactive, a proportion of polysynthates produced from the CBB cycle are allocated into glycogen storage.  As the cell transitions from light to dark conditions, glycogen is subsequently mobilized and used as the main substrate for respiration in the dark.  The observed oscillation in glycogen content is a natural adaption to light availability as the sun rises and sets in a 24-hour period.  Light intensity also plays a role in glycogen accumulation, as it was reported that high light intensity resulted glycogen accumulation up to 34% dry cell weight after 9 hours in *Spirulina maxima* (De Philippis et al. 1992).  In contrast, glycogen synthesis is halted when cells are in dark conditions, as demonstrated by glycogen-deficient *Synechocystis* sp. PCC 6803 mutants that lose viability when grown in diurnal light-dark cycles (Gründel et al. 2012).

Carbon availability greatly influences glycogen accumulation as well.  In the presence of abundant $CO_2$, *Synechocystis* sp. PCC 6803 contain more glycogen granules in the thylakoid membranes than cells grown in low carbon conditions (Eisenhut et al. 2007). As cells transition from a high to low carbon environment, cells mobilize and consume glycogen to compensate for loss of carbon (Eisenhut et al. 2008a).  For instance, intermediates in the oxidative pentose phosphate (OPP) pathway (e.g. 6-P-gluconate) are

shown to increase in low carbon conditions, which replenish CBB cycle intermediates during the early stages of carbon starvation. This carbon appears to be drawn from glycogen, which exhibit decreased permanently decreased pool sizes (Eisenhut et al. 2008a).

When excess carbon is coupled with nitrogen depletion, a rapid and large accumulation of glycogen levels is observed in cyanobacteria, accounting for up to 60-70% w/w dry weight of the cell (Yoo et al. 2007). The absence of an essential macronutrient (e.g. nitrogen, sulfur, phosphate) naturally induces significant glycogen storage, as the cells switch from an active growth state to a dormant, maintenance state. This phenomenon is also reflected in observations of *Anacystis nidulans*, which begins glycogen accumulation when the cells transition from log to stationary phase (Lehmann and Wöber 1976). Inversely, supplementing the missing macronutrient into nutrient starved, glycogen-rich culture leads to rapid degradation of glycogen (Allen and Smith 1969).

### 1.11.4   *Regulating enzyme activity in glycogen metabolism*

Glycogen metabolism in cyanobacteria is orchestrated by conglomerate of metabolic enzymes, whose activities are influenced mainly by i) posttranslational redox-modification ii) allostery by effector molecules. Light availability appears to be a major factor in enzymatic activity, exerting indirect control over glycogen metabolism enzymes through the ferredoxin/thioredoxin system (Buchanan 1991; Schürmann and Buchanan 2008). This system utilizes reduced ferredoxin and thioredoxin (produced from the light reactions of photosynthesis) to reduce and activate various metabolic enzymes in cyanobacteria, thereby regulating their activity. Glycogen metabolic enzymes, including

AGPase (GlgC), glycogen synthase (GlgA), and branching enzyme (GlgB) were identified as thioredoxin targets in *Synechocystis* sp. PCC 6803 when cells were exposed to light (Lindahl and Florencio 2003). This explains why cyanobacteria are able to stimulate glycogen synthesis in the light, and halt production in the dark. The enzyme phosphoglucomutase (PGM, EC 5.4.2.2), which catalyzes the isomerization of glucose-6-phosphate to glucose-1-phosphate, was also shown to be metabolically controlled by the thioredoxin system in this cyanobacteria (Lindahl and Florencio 2003). This observation is extremely interesting, considering PGM represents a central control point in that links glycolysis, OPP pathway, and glycogen metabolism together. Therefore, the light-dependent reactions of photosynthesis play a pivotal role in controlling carbon storage (glycogen synthesis) and respiration (OPP pathway, glycolysis).

AGPase, the enzyme responsible for the first committed step in glycogen synthesis, can be allosterically modulated by effectors derived from the dominant product of carbon metabolism in their respective organism (Iglesias and Preiss 1992). Photoautorophs that utilize the CBB cycle as the main carbon fixation route (plants, cyanobacteria, and green algae) contain AGPases that are activated by 3-phosphoglycerate (3-PGA), the primary product of $CO_2$ fixation. Thus, the production of fixed carbon also signals glycogen synthesis in cyanobacteria (Iglesias et al. 1991). In contrast, inorganic phosphate ($P_i$), which signals low energy levels in the cell, acts as an inhibitor of AGPase and prevents allocation of fixed carbon into glycogen (Iglesias et al. 1994). Although cyanobacterial AGPases are structurally different from plant and algae (cyanobacteria contain homotetrameric AGPase, while plant and algae possess a heterotetrameric AGPase), they

are nonetheless strictly regulated by the 3-PGA/P$_i$ ratio in the cell (Iglesias et al. 1994; Iglesias et al. 1991; Iglesias and Preiss 1992).

*1.11.5   Regulating gene expression in glycogen metabolism*

The genes involved in glycogen metabolism must work in tandem to complete their specific role in glycogen synthesis and depolymerization (e.g., *glgC, glgA, glgB* for synthesis; *glgP, glgX* for depolymerization).  Surprisingly, neither of these gene sets are found co-localized in an operon configuration.  Analyzing the current set of annotated cyanobacterial genomes reveals that glycogen metabolism genes are spatially separated across the chromosome (Memon et al. 2013).  This is in contrast to heterotrophic bacteria (e.g., *Bacillus subtilis*, *Escherichia coli*, and *Agrobacterium tumefaciens*) that express their *glg* genes in an operon (Wilson et al. 2010).  In cyanobacteria, glycogen metabolism genes tend to be arranged next to core carbon metabolism genes.  For instance, the gene encoding GlgP in *Synechococcus* sp. PCC 7942 is found downstream of glyceraldehyde-3-phosphate (GAP), a key enzyme in glycolysis (Koksharova et al. 2004).

A number of glycogen metabolism genes are transcriptionally controlled by unique sigma (σ) factors of RNA polymerase.  Northern blot and microarray analyses revealed that *glgP* and *glgX* are regulated by the group 2 sigma factor SigE (Osanai et al. 2007; Osanai et al. 2005).  SigE follows a pattern of circadian oscillation, thereby up-regulating *glgX* and other sugar catabolism genes towards the end of the day (Osanai et al. 2005). The coupling of sugar catabolism with the circadian rhythm provides an elegant mechanism to maintain the energetic and carbon status of the cell during repeated cycles light and dark.  The circadian rhythm is a crucial regulator of carbon catabolism in other photoauttrophs as well, including starch metabolism in plants (Graf et al. 2010).

Other factors, besides light, are able to exert transcriptional control over glycogen degradation. The glycogen phosphorylase *glgP2* is upregulated in *Syenchocystis* sp. PCC 6803 during low $CO_2$ conditions (Eisenhut et al. 2007). The nitrogen status of cell can also determine expression patterns of genes involved in glycogen metabolism. It has been well established glycogen accumulates in cyanobacteria upon the onset of nitrogen deprivation. However, the observed increase in glycogen content is more than likely due to an increase in sugar substrates, rather than an increase expression of glycogen synthesis genes (von Wobeser et al. 2011). Genes involved in the glycogen degradation, rather than synthesis, are up-regulated during nitrogen starvation (Ehira and Ohmori 2006b; Osanai et al. 2005; von Wobeser et al. 2011). Currently, two protein regulators have been noted to orchestrate this process: the nitrogen regulator PII, and the global transcription factor NtcA. PII, which senses the C/N ratio of the cell through allosteric binding of ATP and 2-oxoglutarate, attaches to the NtcA transcription factor (among others), creating a multi-transcriptional complex that binds to the promoter of *sigE*, inducing its expression (Forchhammer 2004). Two other transcriptional regulators, the OmpR-type regulator Rre37 in *Synechocystis* sp. PCC 6803 (Azuma et al. 2011) and NrrA in *Anabaena* sp. PCC 7120 (Ehira and Ohmori 2006b) are also inducted in this process.

*1.11.6    Roles of glycogen metabolism in cyanobacteria*

Glycogen metabolism is a deeply integral component of the metabolic network of cyanobacteria, and is a vital precursor for essential cellular metabolites during transitions from heterotrophic to phototrophic growth. This is reflected in a flux-balance analysis (FBA) studies on *Synechocystis* sp. PCC 6803 during light-dark transitions (Knoop et al.

2013; Knoop et al. 2010). Under conditions of glycogen utilization (heterotrophic growth), flux distribution switches from the phototrophic $O_2$ generation and RuBisCO-dependent carboxylation to $O_2$-uptake and increased flux through the incomplete TCA cycle of *Synechocystis* sp. PCC 6803 (Knoop et al. 2010). Given its deep-rooted connection to both photosynthesis and respiration, it might be assumed that eliminating this dynamic carbon reserve in cyanobacteria would lead to lethality. Interestingly, successful inactivation of glycogen synthesis and degradation has been reported in a number of cyanobacteria strains, notably the photo-heterotrophic strains *Synechocystis* sp. PCC 6803, *Synechococcus* sp. PCC 7002, the obligate photoautrophic *Synechocystis* sp. PCC 7942, and the nitrogen-fixing *Anabaena* sp. PCC 7120 (Table 3). These mutants possess phenotypes similar to that of wild-type strains when grown in balanced, photoautotrophic conditions. However, the consequences of impaired glycogen synthesis/degradation are wide-ranging and numerous when the mutant cells are subjected to variety of environmental changes (Table 3). Mutant cells contain a lower photosynthetic and reduced respiratory capacity, and are sensitive to high light, low inoculum (cell density), dark-light transitions, low $CO_2$, and macronutrient starvation (e.g. nitrogen and phosphate). Furthermore, the mutants possess two very striking features: (i) an overflow metabolism characterized by the secretion of partially-oxidized TCA cycle intermediates (organic acids) and (ii) the inability to undergo chlorosis during nutrient-starvation (i.e., the non-bleaching phenotype). These unusual phenotypes that emerge from glycogen-disruption likely underlie the essential role that this macromolecule plays in sensing the carbon status on the cell, as well as regulating a wide-range of metabolic networks.

**Table 4.** Phenotype of mutants inactivated in glycogen metabolism in cyanobacteria

| Inactivated gene | Phenotype | PCC strain(s) | Reference |
|---|---|---|---|
| *glgC* | Polyglucan-deficient phenotype | 6803, 7942, 7002 | (Carrieri et al. 2012) |
| | Reduced viability in dark-light transitions | 6803 | (Gründel et al. 2012) |
| | Reduced respiratory activity in dark | 6803, 7942, 7002 | (Hickman et al. 2013) |
| | Reduced photosynthetic capacity | 6803, 7942, 7002 | (Miao et al. 2003a; |
| | Salt-sensitive phenotype | 6803, 7942, 7002 | Miao et al. 2003b) |
| | Glucose-sensitive phenotype | 6803 | |
| | Non-bleaching phenotype | 6803, 7942, 7002 | |
| | Overflow phenotype | 6803, 7942 | |
| | High light-sensitive phenotype | 6803 | |
| | Low inoculum-sensitive phenotype | 6803 | |
| | Thylakoid and ribosome retainment | 7002 | |
| *agp (glgC)* | Polyglucan-deficient phenotype | 7120 | This dissertation |
| | Reduced viability in light-dark | | |
| | Reduced photosynthetic capacity | | |
| | Non-bleaching phenotype | | |
| | High light-sensitive phenotype | | |
| | Reduced nitrogenase activity | | |
| | Increase heterocyst frequency (early) | | |
| *glgA$_n$[a]* | Polyglucan-deficient phenotype | 6803, 7942, 7002 | (Gründel et al. 2012) |
| | Reduced viability in dark-light transitions | 6803, 7002 | (Suzuki et al. 2010; Xu |
| | Reduced respiratory activity in dark | 7002, 7942 | et al. 2013) |
| | Reduced photosynthetic capacity | 6803, 7942, 7002 | |
| | Salt-sensitive phenotype | 7002, 7942 | |
| | Glucose-sensitive phenotype | 6803 | |
| | Non-bleaching phenotype | 6803, 7002 | |
| | Overflow phenotype | 6803 | |
| | High light-sensitive phenotype | 6803 | |
| | Low inoculum-sensitive phenotype | 6803 | |
| *glgA1* | Polyglucan synthesized | 6803, 7002 | (Gründel et al. 2012) (Xu et al. 2013) |
| *glgA2* | Polyglucan synthesized | 6803, 7002 | (Gründel et al. 2012; Xu et al. 2013) |
| *glgB* | Changed polyglucan structure | 6803 | (Yoo et al. 2002) |
| *glgP1* | Temperature-sensitive phenotype | 6803 | (Fu and Xu 2006) |
| *glgP2* | Low-carbon phenotype | 6803 | (Fu and Xu 2006) |
| *glgX* | Changed polyglucan structure | 7942 | (Suzuki et al. 2010) |
| | Increased respiratory activity | | |

The next sections will outline three functional aspects of glycogen disruption that have been extensively studied: (i) the effects of glycogen disruption on photosynthetic activity, (ii) the role of glycogen in acclimating to nitrogen-starvation, and (iii) the role of glycogen in acclimating to osmotic stress.

### 1.11.6.1    The effect of glycogen deficiency on photosynthetic efficiency

A drastic reduction in photosynthetic activity (rate of $O_2$ generation per unit chlorophyll) is observed in mutants defective in glycogen synthesis (Gründel et al. 2012; Jackson et al. 2015; Miao et al. 2003a; Suzuki et al. 2010).  Oxygen evolution is the result of the water-splitting reactions of PSII, a major protein complex in the photosynthetic Z-scheme.  While the connection between the glycogen metabolism and PSII activity is mostly unclear, an extensive study of photosynthetic dynamics in a glycogen-deficient $\Delta glgC$ mutant of *Synechococcus* sp. PCC 7002 during nitrogen starvation was performed by Jackson et al. to illuminate this dark area of cyanobacteria biology.

Cyanobacteria naturally depress photosynthetic activity during nitrogen starvation, via degrading their photoactive PSII centers in the thylakoids.  During first 48 hours of nitrogen starvation, the glycogen-deficient $\Delta glgC$ mutant showed lower PSII activity and a reduced number of functional PSII centers per cell than the wild-type.  However, after prolonged nitrogen starvation (after 48 hours), the mutant was able to maintain low levels of PSII activity and whole-chain photosynthetic electron transport activity under conditions in which the wild-type had continued to down regulate (Jackson et al. 2015).

Under nitrogen deprived conditions, the mutant also possessed (i) an elevated abundance of PSII to PSI complexes, (ii) phycobilisomes that remained energetically coupled to the photosynthetic reaction centers, and (iii) a thylakoid ultrastructure that was nearly identical to cells grown in a nitrogen-rich environment (Jackson et al. 2015). These results seemed to indicate that a broader global response to nitrogen deprivation is impaired in the Δ$glgC$ mutant, allowing for photosynthesis to remain active. Why the "chlorosis signal" is interrupted in the Δ$glgC$ mutant as it transitions into nitrogen-deprived conditions is still largely unknown.

### 1.11.6.2    *Glycogen metabolism and the nitrogen stress response (NSR)*

In photosynthetic microorganisms such as cyanobacteria, most stresses related to carbon and nitrogen metabolism are integrated and respond via a network of enzymes and regulatory proteins that maintain cellular homeostasis. When combined nitrogen ($NH_4^+$, $NO_3^-$, urea, amino acids) becomes limiting in the environment, nitrogen-starved cells trigger the nitrogen stress response (NSR) to mitigate cell stress and ultimately provide temporarily relief from dangerously low nitrogen levels. Sensing the carbon/nitrogen status is an important aspect of properly activating the NSR, and glycogen synthesis and mobilization is central part of the network. Upon the onset of nitrogen starvation, two major processes occur simultaneously. Photosynthetically-fixed carbon flux flows to glycogen storage, leading to a massive increase in intracellular glycogen, up to 60-70% w/w per dry weight), while a lack of ammonia in the GS-GOGAT pathway triggers an increase in the 2-oxoglutarate level (2-OG). Here, 2-OG is the molecule that signals that the carbon/nitrogen ratio is too high, and at elevated levels interacts with the PII protein and the transcription factor NtCA to create a transcriptional complex. The complex, in

turn, induces the expression of a number of genes involved in the nitrogen signaling cascade.

The small adaptor protein NblA is a protein induced during the cascade, and it simultaneously binds to phycobiliproteins and Clp-proteases, thereby aiding in the phenomenon of nitrogen chlorosis, also known as photobleaching. The degradation of phycobilisomes is thought to fulfill two roles: to mobilize nitrogen-stores from phycobiliproteins as a temporary nitrogen source during acclimation, and to repress photosynthesis and prevent photo-damage. Findings from $NaH^{13}CO_3$-labeling experiments indicate that glycogen is synthesized from carbon derived from amino acids via gluconeogenesis, which phycobilisomes could be a potential source (Hasunuma et al. 2013). The degradation of phycobilisome ultrastructure is tightly correlated with glycogen synthesis, as the progression and degree of chlorosis correlates with the amount of glycogen accumulated (Gründel et al. 2012). However, non-bleaching (Δ*nblA*) mutants show no significant change in glycogen accumulation, indicating that glycogen synthesis may directly couple/control phycobilisome degradation.

In cyanobacteria, the blocking of glycogen leads to an uncoupling of catabolic and anabolic reactions, leading to a metabolic spillover of partially-oxidized intermediates of the TCA cycle. Organic acids such as 2-OG, pyruvate, succinate, acetate (Carrieri et al. 2012; Gründel et al. 2012) and alpha-ketoisocaproate (Jackson et al. 2015) are excreted into the media from glycogen-deficient cells of cyanobacteria when placed under nitrogen stress. Researchers on ΔglgC cyanobacteria mutants note that this loss of reductants occurs only during active photosynthesis (Carrieri et al. 2012; Gründel et al. 2012; Hickman et al. 2013). It is possible that glycogen synthesis is an important factor

for sensing and balancing steady-state anabolic and catabolic processes in the cell, and that the loss of glycogen accumulation interferes with the cell's ability to transition from growth mode (anabolism) to dormant mode (catabolism). Since this metabolic switch is catabolism is impeded, the cells build up partially-oxidized material that may interfere with NSR-related processes, such as phycobilisome degradation.

## 1.12   Terpenoids:  An introduction

Terpenoids are structurally diverse group of compounds that function as primary and secondary metabolites (Chappell 1995; Rohmer 1999). The roles of terpenoids in living organisms has been researched extensively, mostly in plants, but also in animals, yeasts, bacteria, and green algae. There are many excellent reviews which highlight the diverse roles of terpenoids in cellular function (Lichtenthaler 1999; Vranova et al. 2013) As secondary metabolites, they play a central role in membrane fluidity, respiration, photosynthesis, protein prenylation, and the regulation of growth and development through hormones synthesis. A major proportion of terpenoids synthesized in plants are used to construct photosynthetic pigments that participate in light harvesting, photo-quenching, and transferring resonance energy in photosynthetic reaction centers (PRCs). Chlorophylls, which consist of a magnesium-bound tetrapyrrole ring attached to an terpenoid-derived phytol chain, are essential molecules in the PRC for photon absorption and energy transfer.  Accessory pigments, such as carotene and xanthopylls, are also derived from terpenoids, and are important for quenching excess excitation energy and protecting the light-harvesting complexes from photo-damage. Plant sterols are steroid-like molecules derived from terpenoid precursors, and like cholesterols in vertebrates,

function to regulate the rigidity and integrity of cell membranes. Monoterpenes emitted from plants have a pleasant aroma, and are utilized to attract insect pollinators and seed-dispersing animals. Terpenoids also participate in plant-pathogen interactions, and can aid in protection against herbivory.

### 1.12.1   Metabolic Routes for Terpenoid Synthesis

All terpenoids are formed through the universal precursor isopentenyl pyrophosphate (IPP) and its isomer, dimethylallyl pyrophosphate (DMAPP) (Vranova et al. 2013). The 1:1 condensation of IPP ($C_5$) and DMAPP ($C_5$) through geranylpyrophosphate synthase (GPPS; EC 2.5.5.1) creates geranylpyrophosphate (GPP; $C_{10}$), the precursor for all $C_{10}$ terpenoids (monoterpenes), including linalool. The condensation of IPP to GPP creates farnesyl pyrophosphate (FPP) the precursor for all sesquiterpenes ($C_{15}$), the major constituents of hops in beer (a major energy source for PhD candidates when writing a dissertation). The addition of IPP to FPP creates geranylgeranyl pyrophosphate (GGPP; $C_{20}$), the precursor for photosynthetic pigments (chlorophylls and carotenoids) and quinones involved in the light-dependent reactions (plastoquinones and ubiquinones). Subsequent additions of IPP subunits to longer prenylated-pyrophosphates creates the precursors needed to create triterpenes ($C_{30}$), tetraterpenes, ($C_{40}$), and longer. IPP and DMAPP are synthesized from two separate pathways: the mevalonate (MVA) pathway, and the recently discovered 2-*C*-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate (MEP/DOXP) pathway (Lichtenthaler 1999; Vranova et al. 2013). Animals, archaea, some Gram-positive cocci bacteria, and yeast contain the MVA pathway, while most Gram-negative bacteria, cyanobacteria, and green algae use the MEP pathway exclusively.

While most organisms have either one or the other, higher plants and some marine alga (e.g., the red alga *Cyanidium caldarium*) contain both the MVA pathway and MEP pathway. Genomic analyses have revealed that higher plants compartmentalize these pathways, with the MVA pathway operating in the endoplasmic reticulum and cytosol, and the MEP pathway functioning in the chloroplast. Why plants have retained both pathways is currently unknown, but brings an interesting discussion on evolutionary aspects regarding terpenoid synthesis. By compartmentalizing both pathways in different organelles, plants may overcome the disadvantages of immobility by having a more stringent control on carbon flux to terpenoids. This would allow accurate and faster responses to environmental cues that affect growth, development, and metabolism (Vranova et al. 2013).

### 1.12.1.1 *Terpenoid Synthesis through the MVA Pathway*

Terpenoids synthesized in the cytosol and mitochondria are synthesized through the MVA pathway (Lichtenthaler 1999; Vranova et al. 2013). The pathway starts with the Claisen condensation of two molecules of acetyl-coenzyme A (Ac-CoA) to create acetoacetyl-CoA (AcAc-CoA) in a reversible reaction, via the enzyme AcAc-CoA thiolase (AACT; EC 2.3.1.9). Thiolases are divided into two classes: class I, which are characterized as degradative thiolases and are involved in fatty acid oxidation, and class II, which are biosynthetic thiolases and can catalyze the first reaction of the MVA pathway. Following models of pathway regulation, AACT activity strictly adheres to substrate/product ratios, and was found to be extremely sensitive to free CoA. In an effort to increase *n*-butanol production in *Clostridium acetobutylicum*, the AACT enzyme ThlA was engineered to be less sensitive to feedback inhibition through free CoA by

substituting three amino acids in the protein. This lead to an increase of ethanol and butanol titers by 46% and 18%, respectively (Mann and Lutke-Eversloh 2013).

AcAc-CoA is further condensed to 3-hydroxy-3-metylglutaryl-CoA (HMG-CoA) by HMG synthase (HMGS; EC 2.3.3.10). In yeast and most plants, these enzymes are found in paralogs (a set of genes created from a gene duplication event) within the chromosome. The exception is found in *Arabidopsis thaliana*, where HMGS is encoded by a single gene found to complement *erg11* and *erg13* yeast mutants that lack HMGS (Vranova et al. 2013).

Next, HMG-CoA is converted to mevalonic acid (MVA) in a two-step reduction process, requiring NADPH as a reducing equivalent. This reaction is catalyzed by 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR; EC 1.1.1.34), a major rate-limiting enzyme in the MVA pathway. *Saccharomyces cerevisiae* harbors two paralogs of HMGR (*hmg1p* and *hmg2p*), each consisting of two major domains: an anchoring transmembrane domain associated with the endoplasmic reticulum (ER), and a catalytic domain facing the cytosol. Research has shown that removal of the N-terminal anchoring domain and overexpression of the catalytic domain increased HMGR activity and the accumulation of squalene in yeast (Donald et al. 1997).

In two successive phosphorylations, mevalonate kinase (MVK; EC 2.7.1.36) phosphorylates MVA to create mevalonate-5-phosphate (MVA-P), which is then phosphorylated by phospho-MVA kinase (PMVK; EC2.7.4) to produce MVA-5-pyrophosphate (MVA-PP). These two enzymatic steps were identified as major bottleneck points in the MVA pathway, and overexpression of both genes in *E.coli* lead to improved productivity of desired bio-products (Redding-Johanson et al. 2011).

The last steps of the pathway involve the decarboxylation of MVA-5-pyrophosphate to IPP, followed by the reversible isomerization of IPP to DMAPP. MVA-5-pyrophosphate to IPP is catalyzed by the enzyme MVA-pyrophosphate decarboxylase (MVD; EC 4.1.1.33). This step relies on the input of ATP and results in the decarboxylation of MVA-5-pyrophosphate to IPP, along with the release of $CO_2$. IPP isomerase (IDI; EC 5.3.3.2) then acts to convert IPP to its isomer DMAPP by transposing the position of a C-C double bond on the IPP molecule. This isomerization reaction is a noted rate-limiting step in the pathway, with equilibrium favored towards DMAPP production (Sun et al. 1998).

*1.12.1.2    Terpenoid Synthesis through the MEP Pathway*

The MEP pathway, also referred to as the 1-deoxy-D-xylulose 5-phosphate (DOXP) pathway, is alternative route to the formation of IPP and DMAPP, and is found in bacteria, various apicomplexan parasites, and plant chloroplasts (Lichtenthaler 1999). The pathway starts with the condensation of pyruvate and glyceraldehyde-3-phosphate (GAP) to produce 1-deoxy-D-xylulose 5-phosphate (DXP). This reaction is catalyzed by the enzyme 1-deoxy-D-xylulose-5-phosphate synthase (DXS; EC 2.2.1.7), which requires a divalent cation ($Mg^{2+}$ or $Mn^{2+}$) and thiamine pyrophosphate (TPP) as a cofactor. DXS plays an important regulatory role in the MEP pathway, since excess IPP and DMAPP can inhibit DXS activity by competing with TPP for binding with the enzyme. In this way, excess IPP and DMAPP (created downstream in the pathway) can effectively slow or shut down MEP flux through an inhibitory feedback loop.

Next, DXP is converted to methylerythritol 4-phosphate (MEP) via DXP reductoisomerase (DXR, also known as IspC; EC 1.1.1.267). DXR requires reducing

power in the form of NADPH, gained from the photosynthetic electron transport chain in photosynthetic organisms. Enzyme activity is controlled by phosphorylation of a particular serine residue in bacteria (Ser177 in *Francisella tularensis*; Ser186 in *Escherichia coli*), and acts as an important residue for binding of substrate (Banerjee and Sharkey 2014). The serine residue is conserved within plant DXRs, however, there is no information at present whether this same mechanism plays any role in regulating the MEP pathway.

The enzyme methylerythritol 4-phosphate cytidylyltransferase (MCT, also known as IspD; EC 2.7.7.60) converts MEP to diphosphocytidylyl metherythritol (CDP-ME) in a CTP-dependent reaction. Like DXR, this enzyme is thought to be dependent on a phosphorylation site on a threonine residue (Thr141 in *F. tularensis*; *Thr140 in E.coli*), which could also play a role in substrate binding (Banerjee and Sharkey 2014). CDP-ME is then phosphorylated from ATP to produce diphosphocytidylyl methylerythritol 2-phosphate (CDP-MEP), catalyzed by CDP-ME kinase (CMK; EC 2.7.1.148). These two steps in the MEP pathway both require CTP and ATP, respectively.

The cyclic conversion of CDP-MEP to CDP-ME-cyclo-pyrophosphate (CDP-MEcPP) is performed by MEcPP synthase (MCS, also known as IspF; EC 4.6.1.12), and represents a major regulatory checkpoint in the MEP pathway (Banerjee and Sharkey 2014). Crystal structural analysis of the MCS from *E.coli* has shown that a hydrophobic cavity present in the enzyme may bind to different terpenoids containing pyrophosphate moieties, such as IPP/DMAPP, GPP, and FPP. Furthermore, sequence alignments show strong conservation of this motif among MCSs from various organisms, suggesting a selective pressure for this specific structure (Banerjee and Sharkey 2014). These results

suggest that the motif in MCS might play a specific role in feedback-regulation in the MEP pathway.

The next two steps in the MEP pathway consist of successive redox reactions. The first enzyme, hydroxy-methylbutenyl diphosphate synthase (HDS, also known as IspG or GcpE; EC 1.17.7.1) converts CDP-MEcPP to hydroxyl-methylbutenyl pyrophosphate (HMBPP), which is then reduced to IPP and DMAPP by HMBPP reductase (HDR, also known as IspH or LytB; EC 1.17.1.2). Both enzymes utilize [4Fe-4S]-clusters and involve a double one-electron transfers in their reaction mechanism. It was demonstrated in *A. thaliana* and the thermophilic cyanobacterium *Thermosynechococcus elongatus BP-1* that HDS reduces MEcPP through reduced ferredoxin, generated by the photosynthetic electron transport chain (Okada and Hase 2005; Seemann et al. 2006). In contrast, *E.coli* HDS requires the flavodoxin/flavodoxin reductase/NADPH system to reduce MEcDP (Xiao et al. 2009). The highest reported HDS activity in *E.coli* is remarkably low at 74-99 $nmol \cdot min^{-1} \cdot mg^{-1}$, which is roughly 300-fold lower than HDR (30.4 $\mu mol \cdot min^{-1} \cdot mg^{-1}$) and 100 to 5,000-fold lower than all of the other MEP enzymes (10-500 $\mu mol \cdot min^{-1} \cdot mg^{-1}$). HDS and HDR both require a very negative reducing power to catalyze their reactions, which are naturally supplied by ferredoxin (plants and cyanobacteria) or the NADPH/flavodoxin system (bacteria). Reducing power is often a limiting factor for HDS and HDR activity. Exchanging NADPH/flavodoxin for an artificial redox partner with a lower reducing potential (methyl viologen; -446 mV) increased the activity of the *E.coli* HDS to 550 $nmol \cdot min^{-1} \cdot mg^{-1}$ (Xiao et al. 2009).
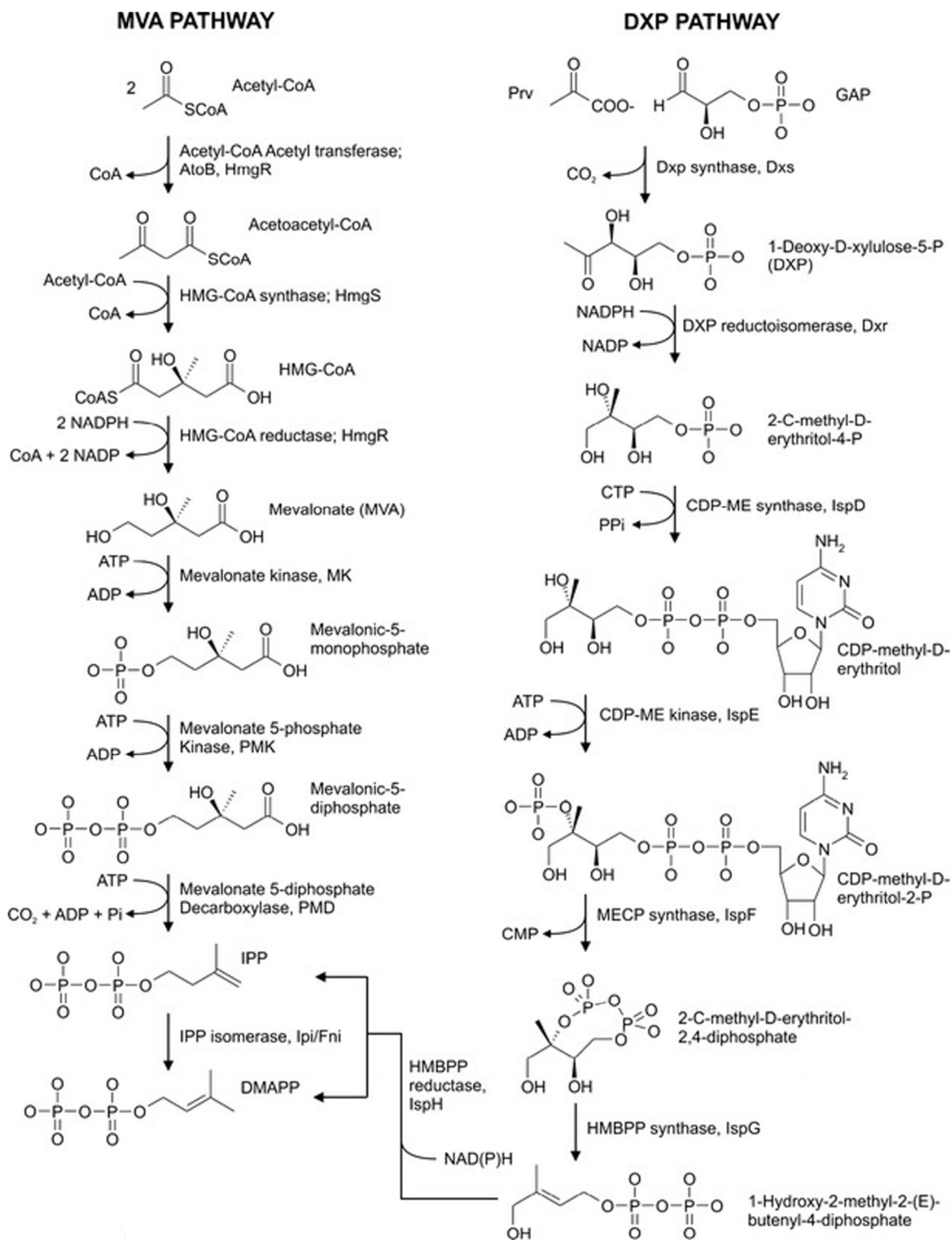
**Figure 11.** Pathways for terpenoid synthesis, using the MVA and MEP pathways. Acronyms are described in the text. Adapted from (Zurbriggen et al. 2012).

*1.12.2   Convergence of MEP and MVA pathway through IPP/DMAPP formation*

In contrast to the MVA pathway, which creates IPP and isomerizes it to DMAPP through IDI, the enzyme HDR catalyzes the simultaneous formation of both IPP and DMAPP in an approximate 5:1 proportion.  The synthesis of IPP and DMAPP links both the MVA and MEP pathway, although the pathways are separated by different organisms, or in the case of plants, different organelles.  Two classes of IDI enzymes have been discovered which show no sequence similarities, display different reaction mechanisms, and require different cofactors for proper catalysis (Perez-Gil and Rodriguez-Concepcion 2013).  The type I enzyme (IDI-I) is found in many bacteria (including *E.coli*) and is similar to IDIs found in fungi, plants, and animals.  The type II enzyme (IDI-II) is found in archaea and some bacteria (*Streptomyces*), but is absent from plants and animals.  Interestingly, genome sequencing has revealed some bacteria contain either IDI-I or II enzymes, some that contain both, and some that contain neither.  Since IDI activity is essential to produce DMAPP in organisms only containing the MVA pathway, it may not be surprising to find its absence in organisms harboring the MEP pathway, since they can form IPP and DMAPP through HDR.   Some species of cyanobacteria, including *Synechocystis*, *Synechococcus*, and *Cyanothece* contain orthologues of IDI-II, yet these enzymes are deemed non-essential for terpenoid synthesis (Perez-Gil and Rodriguez-Concepcion 2013).

As stated earlier, the condensation of IPP and DMAPP creates the prenyl precursors involved in monoterpene ($C_{10}$) and sesquiterpene ($C_{15}$) biosynthesis.  Prenyltransferase-catalzyed condensation of IPP and DMAPP in a 1:1 ratio creates geranylpyrophosphate (GPP), and the subsequent 1'-4 addition of IPP to GPP generates farnesylpyrophosphate

(FPP). Terpenoid synthases that utilize GPP as a substrate are referred to as monoterpene synthases, while synthases that catalyze FPP are known as sesquiterpene synthases. In the following sections, we overview some general features of terpenoid synthases, and describe the reaction mechanisms involved in generating the diverse array of monoterpene and sesquiterpene skeletal structures. For brevity, we will limit our focus to monoterpene and sesquiterpene synthases only.

### 1.12.3   Terpenoid Synthases in Nature

Terpenoid synthases encompass a large class of enzymes that convert acyclic prenyl diphosophates into a multitude of cyclic and acyclic carbon structures. A great number of terepene synthases have been isolated and characterized from a variety of plant species, including angiosperms, gymnosperms, and bryophytes, and they all have similar characteristics. They have a molecular mass from 50-100 kDa, require a divalent metal ion as a cofactor ($Mg^{2+}$ or $Mn^{2+}$), a pI value near 5.0, and a neutral pH optimum. The activities of monoterpene and sesquiterpene synthases have been elucidated primarily by studying crude plant extracts and preparing purified enzymes. For the scope of this dissertation, Table 4 lists a select set of currently characterized terpenoid synthases involved in synthesizing linalool and farnesene, an acyclic monoterpene and sesquiterpene, respectively. It is important for the reader to understand that linalool and farnesene make up a tiny fraction of currently known terpenoid structures. The skeletal diversity of terpenoids arises not only from the sheer number of terpenoid synthases in nature, but from the ability of these synthases to form multiple products from a single prenyl diphosphate substrate.

**Table 5.** Linalool and farnesene synthase genes isolated to date.  Adapted from (Degenhardt et al. 2009)

| Gene bank accession no. | Main products | Designation | Species | Reference |
|---|---|---|---|---|
| *Linalool Synthases* | | | | |
| ABR24418 | (+)-(3S)-Linalool | AmNES/LIS-2 | *Antirrhinum majus* | (Nagegowda et al. 2008) |
| AAO85533 | (+)-(3S)-Linalool | At1g61680 | *Arabidopsis thaliana* | (Chen et al. 2003) |
| AAF13357 | (-)-(3R)-Linalool | QH1 | *Artemisia annua* | (Jia et al. 1999) |
| AAF13356 | (-)-(3R)-Linalool | QH5 | *Artemisia annua* | (Jia et al. 1999) |
| AAC49395 | (+)-(3S)-Linalool | LIS | *Clarkia brewerii* | (Dudareva et al. 1996) |
| CAD57081 | (+)-(3S)-Linalool | FaNES1 | *Fragaria ananassa* | (Aharoni et al. 2004) |
| CAD57106 | (+)-(3S)-Linalool | FaNES2 | *Fragaria ananassa* | (Aharoni et al. 2004) |
| ABB73045 | (-)-(3R)-Linalool | LaLINS | *Lavandula angustivolia* | (Landmann et al. 2007) |
| AAX69063 | (-)-(3R)-Linalool | LeMTS1 | *Lycopersicon esculentum* | (van Schie et al. 2007) |
| AAL99381 | (-)-(3R)-Linalool | - | *Mentha citrate* | (Crowell et al. 2002) |
| EU596453 | (+)-(3S)-Linalool | Os02g02930 | *Orya sativa* | (Yuan et al. 2008) |
| AAV63789 | (-)-(3R)-Linalool | LIS | *Ocimum basilicum* | (Iijima et al. 2004) |
| AAS47693 | (-)-(3R)-Linalool | PaTPS-Lin | *Picea abies* | (Martin et al. 2004b) |
| *Farnesene Synthases* | | | | |
| AAX39387 | (*E*)-β-Farnesene | β-FS | *Artemisia annua* | (Picaud et al. 2005) |
| AAK54279 | (*E*)-β-Farnesene | CjFS | *Citrus junos* | (Maruyama et al. 2001) |
| AAU05951 | (*E,E*)-α-Farnesene | CsaFS | *Cucumis sativus* | (Mercke et al. 2004) |
| AAO22848 | (*E,E*)-α-Farnesene | AFS1 | *Malus x domestica* | (Pechous and Whitaker 2004) |
| AAB95209 | (*E*)-β-Farnesene | TSPA11 | *Mentha x piperita* | (Crock et al. 1997) |
| AAS47697 | (*E,E*)-α-Farnesene | PaTPS-Far | *Picea abies* | (Martin et al. 2004) |
| AAO61226 | (*E,E*)-α-Farnesene | Pt5 | *Pinus taeda* | (Phillips et al. 2003) |
| AAX99146 | (*E*)-β-Farnesene | TPS10-B73 | *Zea mays* | (Schnee et al. 2006) |
| AAX07265 | (*E*)-b-Farnesene | PmeTPS4 | *Pseudotsuga menziesii* | (Huber et al. 2005) |

*1.12.3.1    Reaction mechanism of monoterpene synthases*

Acyclic and cyclic monoterpenes arise from a common sequence of molecular events involving GPP.  The reaction mechanism of all monoterpene synthases starts with the divalent-cofactor ionization of GPP, resulting in the removal of oxygenated pyrophosphate (OPP$^-$) from the C1 extension and a positively-charged carbocation on C2. This creates a geranyl cation, which through proton loss forms (*E*)-β-ocimene or myrcene, or addition of water to form geraniol or linalool.  Alternatively, the geranyl cation can proceed through transoid and cisoid isomerizations to form a linalyl cation, which can also generate linalool through water addition, myrcene or (*E*)-β-ocimene through proton loss.  Apart from linear monoterpenes, cyclized monoterpenes take a different route by the 6,1-closure of the linalyl cation to create an α-terpinyl cation.  This intermediate is the precursor for all cyclic monoterpenes, including limonene, β-pinene, β-phellandrene, and α-terpineol.

*1.12.3.2    Reaction mechanism of sesquiterpene synthases*

The formation of sesquiterpenes from FPP follow a similar pattern of carbocation-based reactions utilized by monoterpene synthases.  However, the longer carbon skeleton of FPP and presence of three double bonds instead of two allow for a greater structural diversity of sesquiterpenes.  Similar to monoterpene synthesis, the synthesis of sesquiterpenes starts with the ionization of FPP to create a farnesyl cation, which through isomerization creates linear sesquiterpenes such as farnesene.  Monoyclic sesquiterpenes such as α-bisabolene are formed with by further isomerization of the *trans*-farnesyl cation to nerolidyl diphosphate, followed by ionization-dependent cyclization to form a 6-carbon ring.  However, the size of the farnesyl chain permits alternative cyclizations to

create much larger ring structures, as found in Germacrene C (10 carbon) and γ-Humulene (11 carbon).

### 1.12.4 Conserved Sequences in terpenoid synthases

Despite a large record of annotated terpenoid synthesis sequences, there is very little information on how structural features aid in GPP or FPP substrate binding and catalysis. Interestingly, there is a great deal of sequence similarity between monoterpene and sesquiterpene synthases, but this similarity is based primarily on the taxonomical source of the synthase (e.g. plant, bacteria), rather than the terpene product. Furthermore, it is near impossible to predict the product of a terpene synthase with the protein sequence alone. Comparative genomics have revealed sequence similarities between terpenoid synthases, giving us insight into important structural and catalytic elements of these enzymes (Bohlmann et al. 1998). In general, terpenoid synthases are composed of two distinct structural domains: an N-terminal transit domain, and a C-terminal active site domain. Transit peptides located in the N-terminus target the newly synthesized enzyme to the chloroplast. Post-translational modifications then remove the N-terminal residues upstream of a conserved, tandem arginine repeat (RRx$_8$W), creating a mature protein. The active site of the terpene synthase is composed of an aspartate-rich DDxxD motif, which coordinates two Mg$^{2+}$ ions for aligning the GPP/FPP substrate to the catalytic pocket of the enzyme. Mutagenesis of any of the three aspartates of the DDxxD motif reduces catalytic activity by 1000-fold, proving the absolute necessity of this element for synthase activity (Bohlmann et al. 1998).

*1.13    Introduction to heterocyst development in Anabaena*

One of the hallmark feats of *Anabaena* sp. PCC 7120 is its ability to positively

respond to nitrogen-starvation through the development of heterocysts, specialized $N_2$-

fixing cells that form from vegetative cells at semi-regular intervals along the filament

(Figure 13).  The name is named from the Greek word "*hetero-*" meaning 'other',

(Fritsch 1951).  The interior of the heterocyst is a micro-oxic environment, resulting from

the inactivation of photosystem II, increased respiration, and the expression of

heterocyst-specific terminal oxidases.  Formation of a laminated glycolipid layer and

extra-thick polysaccharide coat around the heterocyst cell wall also prevents the diffusion

of $O_2$ into the cell interior.  These strategies collectively prevent the introduction of $O_2$,

which would otherwise interfere with nitrogenase activity by oxidizing the Fe-S cofactors

that are essential for fixation.  Nitrogenase reduces $N_2$ into ammonia at the expense of

reductants and ATP, which are generated through carbohydrate catabolism (Bryant

1994).  Carbon, in the form of sucrose or alanine, is delivered to the heterocyst to supply

the energy needed to fuel $N_2$ fixation.  Likewise, nitrogen in the form of amino acids is

exported from the heterocysts to the vegetative cells to meet the demands of protein

synthesis.  The exchange of these essential nutrients displays a mutualism between two

cell types, and is a remarkable example of multi-cellularity operating in the bacterial

domain.

*1.13.1    Early experiments on nitrogen-fixation*

Early studies on nitrogen-fixing cyanobacteria made key morphological observations

on *Anabaena* filaments in nitrogen-starved environments, laying the groundwork for

uncovering the regulatory elements responsible for heterocyst differentiation.  In 1941,

the British biologist G.E. Fogg from Cambridge made a series of investigations on nitrogen-fixation in *Anabaena cylindica*. After making painstaking efforts to isolate filaments of *A. cylindrica* and remove all nitrogen contaminants from glassware, he observed that the cyanobacteria could grow vigorously in laboratory media devoid of any combined nitrogen (Fogg 1942). This led him to conclude that without an adequate nitrogen source in the media, the cyanobacteria must be possess the capacity to fix $N_2$ into ammonia. Likewise, he noted that nitrogen fixation was inhibited when *A. cyclindrica* was allowed access to a combined nitrogen source, such as nitrate. Therefore, it appeared that the ability to fix $N_2$ could be controlled, depending on the nitrogen status in the local environment. Still, it was unclear at the time how *A. cylindrica* was able to simultaneously engage in photosynthesis and fix nitrogen, since it was known at the time that oxygen was a potent inhibitor of nitrogenase activity and any $O_2$ released from photosynthesis should disrupt nitrogen-fixation. How were these photosynthetic microorganisms able to engage in $N_2$ fixation in the presence of $O_2$?
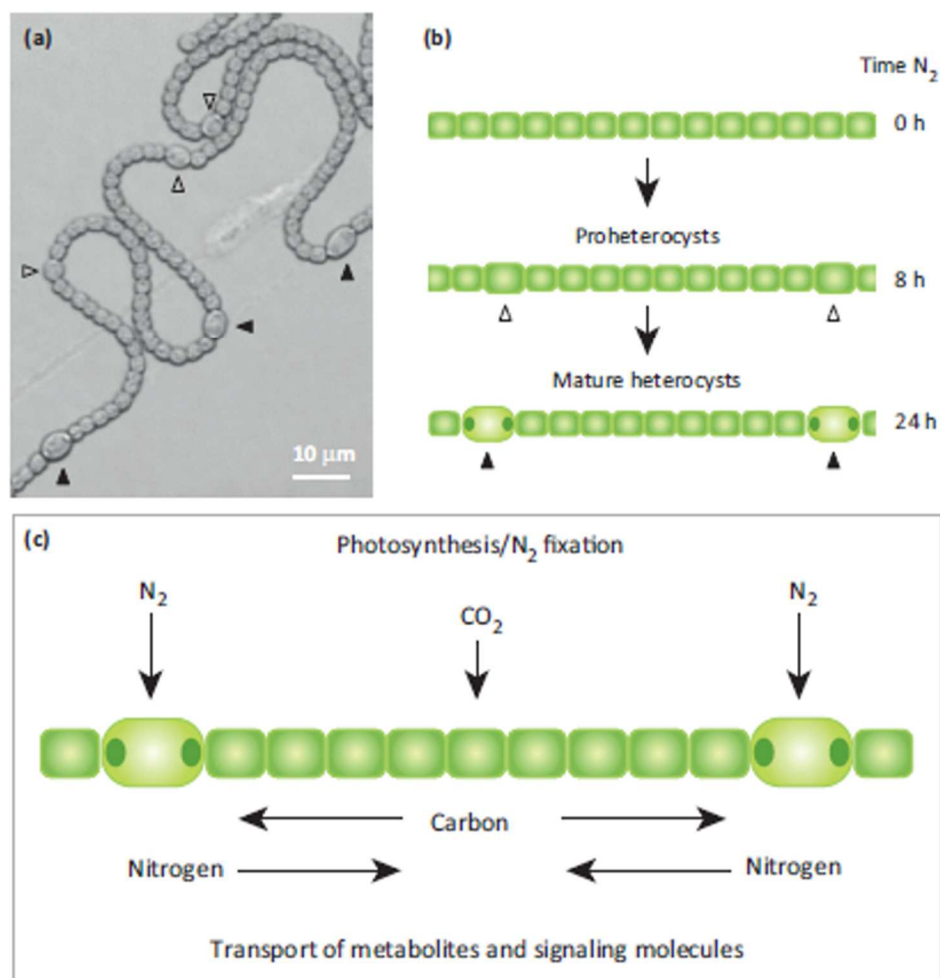
**Figure 13.** Patterned differentiation of vegetative cells into heterocysts, and the relationship between the two cell types. (A). Filaments growing on nitrogen-free solid agar display heterocysts of differing sizes, with large heterocysts (black arrows) presumably older and smaller heterocysts (empty triangles) presumably younger. (B) Heterocysts differentiate at even distances, creating a regular, interspaced pattern. The differentiation program takes 20-24 hours to complete. (C) Products from $CO_2$ fixation are delivered to heterocysts from vegetative cells, while heterocysts deliver nitrogenous compounds to the vegetative cells. Signaling molecules (e.g. PatS) are also transported in this process. Adapted from (Muro-Pastor and Hess 2012).

### 1.13.2    The effects of nitrogen starvation on Anabaena

In the last few decades, advancements in molecular biology tools and techniques have allowed researchers to probe into the inner workings of this unique type of cellular differentiation. In the next few sections, we will review how nitrogen starvation effects nitrogen & carbon metabolism, and also energy production in *Anabaena*. Since most of

our data on the nitrogen stress response in *Anabaena* comes from transcriptional studies, the following sections will be mainly focused on changes on the mRNA level. We will also review the regulatory cascade that guides heterocyst development, and discuss how heterocysts maintain evenly spaced intervals ("pattern formation") along the filament.

### 1.13.2.1    *Nitrogen Metabolism*

When confronted with a sudden loss of combined nitrogen, *Anabaena* responds by rapidly decreasing the transcription levels of genes responsible for assimilating the depleted nitrogen source and increasing transcripts for assimilating an alternative nitrogen source. For example, when encountering a loss of ammonia, an increase in transcripts involved in nitrate assimilation is reported (Cai and Wolk 1997). Likewise, a sudden loss of nitrate reduces the abundance of *nirA-nrtABCD-narB* transcripts involved in nitrate assimilation, while the expression of the *urtABCDE* region for urea transport increased significantly at 3 and 8 hours after nitrate depletion. Cyanophycin, a nitrogen-rich polypeptide composed of aspartic acid-arginine residues, is synthesized early in N-depletion by cyanophycin cynthetase (*cphA1*) and cyanophycinase (*cphB1*). Transcripts of both *cphA1* and *cphB1* increase 2-fold by 3 hours, and are maintained 8 hours into N-depletion. Not all nitrogen assimilation transcript profiles show an increase in expression, however. Expression of genes that encode the glutamine synthetase and glutamate synthase in the $NH_3$-assimilatory GS-GOGAT system do not show any significant change in transcript levels compared to N-replete conditions (Flores and Herrero 2005).

Interestingly, by 24 hours, the gene expression for ammonia transport, urea transport, and cyanophycin metabolism drop to their initial levels (similar to N-replete conditions),

and a new set of genes in nitrogen metabolism begin induction: the structural
components of nitrogenase. The *nifHDK* operon, encoding crucial structural genes for
nitrogenase that are only expressed in heterocysts, see a 100-fold increase in transcript
abundance, averaged over all cells in the filament (Elhai and Wolk 1990). Similarly,
other nitrogenase genes in the *nifB-fdxN-nifS-nifU* and *nifE-nifN-nifX-nifW* operons also
see a lower but significant increase in expression. Genes involved in the construction of
prosthetic groups (such as the Fe-Mo cofactor) are also induced after 8 hours into N-
starvation.

When *Anabaena* becomes deprived of a combined nitrogen source, it must quickly
look to internal sources of nitrogen to maintain protein synthesis. One target for nitrogen
is the phycobilisomes, the large antenna structures that funnel additional light energy to
the photosystems. Upon N-depletion, cyanobacteria breakdown their phycobilisomes and
mobilize the amino acids into metabolism, resulting in the chlorosis of cells from a blue-
green to yellow green culture (also known as "photobleaching"). The color change
occurs because of loss of degradation of phycobiliproteins that absorb light in the yellow-
orange area of the visible light spectrum. With this loss of protein, the abundance of
transcripts involved in the structural components of the phycobiliproteins decreases
(Wealand et al. 1989). At the same time, the gene *nblA*, a protease involved in the
breakdown of the phycobilisomes, is observed to increase in transcriptional abundance to
10-fold levels 8 hours into N-depletion (Baier et al. 2004).

*1.13.2.2   Carbon Metabolism*

A consequence of nitrogen-deprivation is the concomitant decrease in carbon fixation
in the cell (Bradley and Carr 1976). Lowered carbon fixation mainly arises from the

degradation of phycobilisomes, and the repression of transcripts related to PSI proteins and chlorophyll synthesis. The two major subunits of RuBisCO, *rbcL* and *rbcS*, also show a 44% reduction in transcription 8 hours after the onset of nitrogen deprivation (Flores and Herrero 2008). The formation of heterocysts, *de novo* protein synthesis, and subsequent nitrogen-fixing activity requires a substantial amount of carbon and energy. Therefore, the question naturally arises: where does the cyanobacterium acquire these components during nitrogen deprivation if photosynthesis is suppressed?

Cyanobacteria have naturally evolved coping mechanisms to continue carbon and energy utilization, even during nitrogen stress. Cyclic photophosphorylation, which involves the shuttling of electrons by ferrodoxin to the cytochrome b6f complex through PSI to produce a proton-motive force (Bendall and Manasse 1995), is one likely source of ATP in both vegetative cells and heterocysts. Cyclic photophosphorylation is especially crucial in heterocysts, since this process does not produce $O_2$, a potent inhibitor of nitrogenase activity. In order to provide carbon input and energy for $N_2$ reduction in heterocysts, it appears likely that the vegetative cells play an active role in delivering organic metabolites to the heterocysts. Evidence suggest that sucrose is the primary carbon source for heterocysts, (Cumino et al. 2007; Curatti et al. 2002; Wolk et al. 1994). Heterocysts in *Anabaena* 7120 contain a specialized sucrose invertase (*invB*; alr0819) and sucrose synthase (*spsB*; all4376), suggesting an active role in supplying energy for $N_2$-fixation (Cumino et al. 2007). The oxidative pentose phosphate pathway (OPP pathway) is yet another likely source for carbon in heterocysts, considering that transcription for OPP pathway specific genes (glucose-6-phosphate dehydrogenase, 6-phosphogluconolactonase, 6-phosphogluconate dehydrogenase, and transaldolase)

increase by 135% during nitrogen deprivation (Flores and Herrero 2008). Genes involved in glycolysis, including glyceraldehyde-3-phosphate dehydrogenase (*gap1*; all2566) and pyruvate kinase (*pyk1*; all2564) are also upregulated in *Anabaena* 7120 during nitrogen deprivation (Ehira and Ohmori 2006b).

Genes involved in the synthesis of glycogen appear to be downregulated by 59% during nitrogen deprivation, while one gene involved in glycogen breakdown, *glgP1* is upregulated 3-fold after 8 hours into nitrogen deprivation (Ehira and Ohmori 2011). However, the increase in genes related to glycogen catabolism do not actually translate to a decreases in total glycogen levels upon nitrogen deprivation. To the contrary, experimental evidence shows during nitrogen starvation, total glycogen accumulates in filaments, as shown in three strains of *Anabaena* (Ernst and Boger 1985). If glycogen is simultaneously accumulated and broken down in filaments in response to nitrogen deprivation, either a futile cycle of synthesis and breakdown is operating in the cells, or, as suggested by Xu, Elhai, and Wolk (Flores and Herrero 2008), these two contradictory processes are operating in different cell types.

To elaborate, the different cell types in question are cells in the filament that strongly express *hetR* (programmed to differentiate into heterocysts) and cells that weakly express *hetR* (and stay vegetative cells). Within a few hours after nitrogen-deprivation, cells along the filament metabolically differentiate, as directed by *hetR* expression. Cells that express low levels of *hetR* and are destined to remain vegetative cells store and transfer carbon that, due to a lack of combined nitrogen, cannot be used to synthesize new amino acids. Meanwhile, cells expressing high levels of hetR and are destined to become heterocysts do not store carbon, but instead breakdown intracellular carbon stores and

accept sugars from the vegetative cells. This extra carbon and energy is then reserved for synthesizing heterocyst-specific proteins, glycolipids, and the polysaccharide envelope that surrounds the heterocyst (Flores and Herrero 2008).

### 1.13.2.3 Energy Production

Photosynthesis a major source of energy in cyanobacteria. When a combined nitrogen source is present in the culture media, more than half of the top 1% of the most highly expressed genes are related to photosynthesis (Flores and Herrero 2008). Upon the onset of nitrogen deprivation, it is revealed that the protein products from these genes become a substantial source of nitrogen for the cells. Along with decreases in mRNA expression of phycobiliproteins, a general decrease in gene expression for proteins in the photosynthetic electron transport chain are seen. Within 8 hours after nitrogen starvation, a 10-fold reduction in transcripts for genes related to PSII, PSI, chlorophyll, ferredoxin, and cytochrome $b_6f$, and ATP synthase are observed in *Anabaena* 7120 (see Fig. 14.2 in (Flores and Herrero 2008)). In contrast, increases in genes related to respiration are shown to increase within the first 24 hours of nitrogen deprivation. These include genes that encode proteins for cytochrome oxidase and ferredoxin-NADPH reductase (see Fig. 14.2 in (Flores and Herrero 2008)).

### 1.14 The process of heterocyst differentiation in Anabaena 7120

Heterocyst development is a complex, yet fascinating process, and has attracted an abundance of research interest over the last few decades. It is worth noting that our current model of heterocyst formation was developed with help from other studies on differentiation in bacteria, including the sporulation of *Bacillus subtilis* (Errington 1993), *Myxococcus* sp. (Kroos and Kaiser 1987) and *Streptomyces* sp. (Chater 1989), among

others.  Studies on heterocyst development have revealed that the differentiation program is orchestrated by a number of crucial regulatory components that guide the cell in a stepwise manner through early, mid, and late stages of the vegetative cell-heterocyst transition.  Though the genetic events that underlie these components remain largely shrouded in mystery, the information we have on heterocyst differentiation may be organized by referring to three general principles:

1)  *Genes involved in heterocyst development are expressed in an ordered sequence*

A series of *Anabaena* sp. PCC 7120 mutants displaying an arrest in heterocyst development at different stages was produced by Ernst et. al, in an attempt to amplify the effects of mutations at multiple stages of differentiation (Ernst et al. 1992).  The random mutagenesis revealed a number of interesting mutant phenotypes, including Fox$^-$ (unable to fix $N_2$ in the presence of oxygen), Fix$^-$ (unable to fix $N_2$, even in anaerobic conditions), Het$^-$ (unable to form heterocysts after prolonged nitrogen deprivation), Hen$^-$ (aberrant heterocyst envelope), Hep$^-$ (defective polysaccharide layer), and Hgl$^-$ (lacking specific heterocyst glycolipids.   The array of phenotypes shown from the mutants strongly suggests that the transition between a vegetative cell to a heterocyst is governed by a sequential set of biochemical events, controlled by the timing of gene expression for a set of crucial regulatory genes.  The stages of heterocyst development roughly follow a sequential order:  a) pro-heterocyst (intermediate between a vegetative cell and a heterocyst) formation, b) maturation and deposition of heterocyst envelope, c) increased respiration and creation of internal microoxic conditions, and d) nitrogenase activity and $N_2$-fixation.

2) *The expression of genes at later stages of development is dependent upon the expression of genes in earlier stages.*

The idea that cellular differentiation is controlled by a series of interlocking steps, so that one step must be completed before another can be initiated, comes from studies in other bacteria which show genetic dependencies of genes at different stages of development.  As with these bacteria, heterocyst development in *Anabaena* also follows a similar developmental strategy, based on several observations: that the expression of *hepA* and other genes in M2 and P2 mutants of *Anabaena* sp. PCC 7120 depend on the master heterocyst regulator *hetR* (Black et al. 1993), and that proheterocyst development and maturation can be halted in *Anabaena variabilis* by treatment with ammonia and drugs, but only during a specific "time window" in development (Bradley and Carr 1977).  Nitrogen-deprivation appears to be the initiator of the developmental cascade.  However, it is still unclear whether heterocyst differentiation follows a linear pathway in development (each gene product in the regulatory sequence is responsible for inducing the next) or a branched pathway (gene regulatory sequences "branch off", producing parallel but independent sequences of gene products in each sequential step.  For example, the sequence that inducts nitrogenase activity branches off from the sequence that produces morphological changes.

3) *Regulation of expression occurs primarily at the transcriptional level*

There is a growing amount of evidence suggesting that heterocyst development is mainly controlled at the transcriptional level, including the regulation of transcripts responsible for nitrogen assimilation (*glnA*; (Tumer et al. 1983)), early heterocyst differentiation (*hetR*; (Black et al. 1993)), nitrogenase activity (*nifHDK*; (Elhai and Wolk 1990), components of the phycobilisomes (phycobiliproteins phycocyananin and

allophycocyanin; (Wealand et al. 1989)), and the sigma factors *sigB* and *sigC* (Brahamsha and Haselkorn 1992), involved in directing RNA polymerases to specific promoters. Indeed, the evidence for transcriptional control in heterocyst development is substantiated by promoter activity of specific genes between vegetative cells and heterocysts. For instance, the promoter for the *nifHDK* operon ($P_{nifHDK}$) was found to be only active in heterocysts, while promoter activity for the large and small RuBisCO subunits ($P_{rbcLS}$) are confined to heterocysts (Elhai and Wolk 1990).

Strong evidence supports the hypothesis that many genes in the *Anabaena* 7120 genome are regulated by promoters with multiple transcription sites, allowing RNA polymerase to express these genes under different conditions. A detailed analysis of *Anabaena* 7120 promoters from vegetative cells reveals they all contain -10 promoter sequences, similar to those found in *E.coli*. RNA polymerase isolated from vegetative cells was found to recognize and bind to -10 promoters in vegetative cells *in vitro* (Schneider et al. 1987). Interestingly, this RNA polymerase is incapable of recognizing one active promoter, $P_{glnA}$, in filaments deprived of combined nitrogen. A number of genes active in heterocysts contain promoters with multiple transcriptional start sites, including *glnA* (Tumer et al. 1983), the photosystem II protein *psbB* (Lang and Haselkorn 1989), the sigma factor *sigA* (Brahamsha and Haselkorn 1992), and the master heterocyst regulator *hetR* (Black et al. 1993). Using alternative transcriptional start sites under different conditions appears to be a strategy utilized by *Anabaena* 7120 to regulate gene expression in different cell types.

To initiate heterocyst development, the filaments need a mechanism to determine the nitrogen status in the environment and "decide" whether to activate heterocyst-specific genes. In this case, a specific molecule that relays carbon-nitrogen (C-N) status of the cell serves as a useful signal to interpret nitrogen availability. Current studies strongly suggest that 2-oxoglutarate (2-OG), a component of the incomplete TCA cycle in cyanobacteria, participates in the signaling of C-N metabolism. During nitrogen assimilation, 2-OG serves as the carbon backbone in nitrogen assimilation in the GS-GOGAT pathway, catalyzed by the enzymes glutamine synthetase (GS) and glutamate synthase (GOGAT). In this pathway, these enzymes catalyze reactions that incorporate $NH_4^+$ to 2-OG to create glutamate and glutamine. As the availability of $NH_4^+$ (the combined nitrogen source) decreases, a metabolic bottleneck is created in the GS-GOGAT pathway, resulting in an intracellular accumulation of 2-OG. Several lines of study point to 2-OG as being the principle signal of the C-N status for bacteria, including in *E.coli* (Ninfa et al. 2000) and the unicellular cyanobacterium *Synechococcus* sp. strain PCC 7942 (Forchhammer and Hedler 1997; Muro-Pastor et al. 2001). Similarly, 2-OG also stimulates the expression of nitrate reductase and transcription of *nir* and *amt1*, two-nitrogen related genes in *Synechococcus* sp. PCC 7942 (Vasquez-Bermudez et al. 2003). In *Anabaena* 7120, increasing intracellular levels of 2-OG induces heterocyst develop in the presence of nitrate, and increases heterocyst frequency in the absence of combined nitrogen (Li et al. 2003).

As 2-OG accumulates in the cell, it interacts with regulatory proteins that induce the expression of genes involved in heterocyst development. The following sections

describe, in detail, the major regulatory proteins that helps initiate the regulatory cascade of heterocyst development.

### 1.14.1   Key regulatory proteins control the heterocyst differentiation cascade

Research on heterocyst differentiation has sought to unveil the molecular agents responsible for initiating and controlling this process.  Through these studies, it has become evident that the regulators involved are conserved amongst heterocyst-forming cyanobacteria.  After analyzing the sequenced genome of *Anabaena 7120*, it was found that 92% of 26 genes characterized to be involved in heterocyst differentiation have orthologs in two other heterocystous cyanobacteria:  *Anabaena variabilis* and *Nostoc punctiforme* (E-value $>10^{-6}$).  To contrast, a generous 43% of the total genes in *Anabaena* 7120 lack orthologs in at least one of these cyanobacteria (Bryant 1994).  This data emphasizes the highly conserved nature of heterocyst regulatory proteins and their importance in controlling the differentiation process across different cyanobacteria species.  Specifically, three regulatory genes, *ntcA*, *nrrA*, and *hetR* are highly expressed during the early stages of nitrogen starvation.  Details of these key regulatory components are discussed below.

### 1.14.1.1   NtcA

The global nitrogen transcription factor *ntcA* was first identified in mutants of *Synechococcus* sp. strain PCC 7942 that were unable to induce expression of genes involved in nitrogen assimilation when deprived of ammonia (Vega-Palas et al. 1990). Complementation of a 669 nucleotide gene that was different from other N-regulators and could revert the mutant to wild-type function proved the importance of this gene to nitrogen metabolism, thus it was given the name *ntcA*. In *Anabaena* sp. PCC 7120, the

gene encodes transcripts of 1.4, 1.0, and 0.8 kb long (Vega-Palas et al. 1990), with the 0.8

and 1.0 kb transcript being observed in both the presence and in the absence of combined

nitrogen, although they are stronger in the absence. The differing lengths of transcripts

appears to result from the use of different *ntcA* promoters, and may also reflect

differences in overall transcript stability. However, there are no notable changes in

overall total transcripts under different nitrogen conditions (Ramasubramanian et al.

1996).

In *Anabaena*, the *ntcA* transcripts encode a 222 aa protein homologous to members of

the CAP (catabolite activator protein) family. Similar to other CAPs found in bacteria,

NtcA contains a helix-turn-helix motif in its C-terminus for making protein-DNA

interactions (Vega-Palas et al. 1990). In response to increased 2-OG accumulation, NtcA

gains a higher affinity for its binding sites. NtcA appears to have the ability for

autoregulation, and is able to induce the transcription of the *ntcA* gene

(Ramasubramanian et al. 1996). Likewise, NtcA is able to specifically bind upstream to

a number of genes subject to repression by ammonium, and regulates the transcription of

a number of heterocyst development genes, including *hetC* (Muro-Pastor et al. 1999),

*devBCA* (Fiedler et al. 2001), *xisA*, *nifHDK*, and *rbcL* (Ramasubramanian et al. 1996).

Dnase footprinting studies show that NtcA binds to the palindromic sequence signature

$GTAN_8TAC$, a sequence found in a number of CAP-activating promoters (Herrero et al.

2001). In promoters that bear this palindromic sequence, transcription activation operates

in two separate ways: an increase in affinity of the RNA polymerase to the promoter, and

an increase in the rate at which the RNA polymerase transitions from the closed-promoter

complex to the open-promoter complex (Busby and Ebright 1997).

*1.14.1.2   NrrA*

The gene *nrrA* (all4312) is a presumptive two-component response regulator that attracted great interest, due to a predicted NtcA-binding sequence upstream of the *nrrA* ORF.  During this time, it was known that HetR was dependent on NtcA for full induction, but it was unclear how NtcA was able to upregulate HetR, since the gene was devoid of any 5' NtcA-binding sequences.  Muro-Pastor *et. al.* helped shed light on this mystery by showing that NtcA was able to bind upstream from *nrrA*, and that *nrrA* transcriptional activity was low in a NtcA-null mutant (Muro-Pastor et al. 2006).  Later, it was shown that NrrA-null mutants of *Anabaena* 7120 exhibited delayed heterocyst development, and Northern analysis confirmed that NrrA is required for the full induction of HetR (Ehira and Ohmori 2006b).  Furthermore, overexpressing *nrrA* results in an over-abundance of *hetR* transcripts (Ehira and Ohmori 2006a).

The data presented by Muro-Pastor *et. al.* and Ehira & Ohmori suggests that NrrA acts as an intermediary messenger between NtcA and HetR, explaining the phenotype exhibited by the NrrA-null mutant of *Anabaena* 7120.  Interestingly, NrrA not only activates the transcription of *hetR*, but was also found to interact with the promoters of several other genes.  For example, the gene *glgP*1, encoding a glycogen phosphorylase and is involved in glycogen breakdown, saw a significant decrease in transcriptional activity in the *nrrA*-disruptant mutant, resulting in an increase in glycogen content after nitrogen stepdown.  Similarly, the expression of genes involved in glycolysis and the pentose phosphate pathway (the *xfp* operon) was also decreased in the *nrrA*-null mutant (Ehira and Ohmori 2011).  These findings suggest that NrrA not only acts on genes involved in heterocyst differentiation, but also glycogen and sugar catabolism.  However,

to which of the many histidine kinases in *Anabaena* 7120 controls the activity of *nrrA* is unknown.

### 1.14.1.3   HetR

In *Anabaena*, heterocyst development is dependent on a single master regulator, a 299-amino acid protein named HetR.  This regulatory protein was discovered through characterization of a *Anabaena* 7120 mutant (216) that was incapable of forming heterocysts, even after protracted periods of nitrogen deprivation (Buikema and Haselkorn 1991).  It was discovered that this mutant phenotype was caused by a point mutation in a 900 nucleotide ORF, which could be rescued through complementation of the wild-type ORF.  Interestingly, overexpression of *hetR* in the wild-type and *hetR*-null mutant resulted in the development of multiple, contiguous heterocysts, also known as the Mch phenotype (Black et al. 1993; Buikema and Haselkorn 1991; Buikema and Haselkorn 2001).  In *Anabaena* filaments overexpressing HetR, the formation of heterocysts still occur in the presence of combined nitrogen, indicating that it is not only essential for formation, but plays a regulatory mechanism in heterocyst differentiation.

The gene *hetR* shows high expression with 2 hours of nitrogen deprivation, with a 20-fold induction in the first several hours (Black et al. 1993).  HetR activates heterocyst differentiation by homo-dimerization and subsequent DNA-binding to the promoters of several important heterocyst-development genes, including *ntcA* (Huang et al. 2004).  In fact, both HetR and NtcA are mutually dependent on their individual up-regulation.  Auto-regulation appears to also be a critical component of HetR, as the HetR dimer is required to induce expression of the *hetR* gene.  The HetR protein also shows proteolytic activity, showing characteristics of serine-type proteases and is dependent on free

intracellular levels of $Ca^{2+}$ (Zhou et al. 1998).  Similar in eukaryotes, $Ca^{2+}$ ions play an

essential role in cell differentiation, and are also fundamental in heterocyst development

in *Anabaena* 7120.  It was found that CcbpP, a $Ca^{2+}$-binding protein that regulates

intracellular $Ca^{2+}$ levels, is targeted by HetR for degradation (Zhao et al. 2005).  The

proteolytic targeting of CcbpP by HetR appears to be a mechanism of positive auto-

regulation, since increasing levels of free $Ca^{2+}$ promotes heterocyst development and

differentiation (Shi et al. 2006).

The autoregulatory mechanism of HetR and other regulators of differentiation is not

unique to heterocystous cyanobacteria, but appears to be a popular strategy to guide gene

activity in variety of organisms and viruses.  Positive autoregulation is found to control

lysogeny in the bacteriophage lambda gene *cI* (Anderson and Yang 2008), the

development of forespores in *Bacillus sp.* (Sun et al. 1991), and sexual dimorphism in

*Drosophila* (Bell et al. 1991).  Positive autoregulation ensures that gene induction

continues, even after the initial conditions that induced gene expression no longer remain.

*1.14.2   Regulatory Mechanisms in Pattern Formation*

In the mid-20th century, heterocyst-formation in *Anabaena* gained popularity as a

model study of cell differentiation (Fritsch 1951; Wilcox et al. 1973; Wolk 1967; Wolk

and Quine 1975).  As more researchers turned their attention to heterocyst development

in blue-green algae, it became evident that heterocyst patterning was not a random

process, but was coordinated by an unknown mechanism. Observations on heterocyst

spacing suggested a defined set of rules that governed how heterocysts were arranged

spatially on the filament.  Heterocysts either appeared on the ends of a filament (terminal

heterocysts) or within a filament (intercalary heterocysts), and were generally situated

between each other in more or less equal intervals among the other cells within the course of a thread (Fritsch 1951). The latter observation appeared to be a consequence of cell division: as the vegetative cells continued to divide between two heterocysts, a vegetative cell roughly center of the two heterocysts would develop into another heterocyst.

An early model on heterocyst patterning suggested that there was a diffusible substance synthesized by the heterocysts, moving outward and inhibiting nearby vegetative cells from initiating development. As the vegetative cells continue cell division, the distance between two heterocysts increases. This allows for the inhibitory substance to dilute to levels sufficient for vegetative cells to differentiate. This was known as the *simple threshold model*, and was corroborated by experiments showing that heterocyst frequencies increased in filaments exposed to physical fragmentation (Wolk 1967). The model was further improved upon by suggesting that an autoregulatory activator that did could not diffuse between cells worked in concert with the diffusible inhibitor, specifically 1) the synthesis of the inhibitor was dependent on the activator and 2) the inhibitor operated by inhibiting the activator (Wilcox et al. 1973); see also (Gierer and Meinhardt 1972). Future genetic engineering techniques would later show that this model was extremely accurate; a credit to the early pioneers who laid the groundwork for heterocyst-patterning studies. The next sections outline a few of the regulatory proteins that underlie this model.

*1.14.2.1   PatS*

As stated above, early models posited the existence of an inhibitory agent that

diffused from heterocysts and prevented neighboring vegetative cells from

differentiating.  In 1998, a small 54-bp gene named *patS* was demonstrated to be that

very agent.  Overexpression of *patS* in *Anabaena* 7120 blocked heterocyst differentiation,

and a *patS*-null mutant displayed an increased frequency of heterocysts, along with an

abnormal patterning (Yoon and Golden 1998).  The gene was also found to be highly

expressed in both proheterocysts and heterocysts.  How patS inhibited heterocyst

differentiation was postulated by studying a class of phosphate regulator (*phr*) genes from

*Bacillus subtilis,* another bacteria that engages in cell differentiation.  One of these genes,

*phrC*, was found in previous work to encode a quorum-sensing pheromone and a

sporulation stimulating factor (CSF) (Lazazzera et al. 1997; Perego and Hoch 1996).

These proteins were composed of small pentapeptides that were synthesized and secreted

for cell-to-cell communication, and it was believed the *patS* gene product operated in the

same manner.  To test this, a synthetic pentapeptide composed of the last five amino

acids of the unmodified-PatS protein (RGSGR) was constructed.  When introduced to the

culture medium, the protein could effectively prevent heterocyst development in

nitrogen-starved *Anabaena* filaments (Yoon and Golden 1998).  The PatS pentapeptide

prevents differentiation in nearby vegetative cells by interacting directly with HetR, and

lowering the DNA-binding capacity of the HetR dimer to its own promoter (Huang et al.

2004).  Interestingly, expression of the synthetic RGSGR penta-peptide in a patS-null

mutant was unable to recover a normal heterocyst patterning, it was suggested that other

peptides other than the C-terminal end of PatS may be involved in intracellular transport (Wu et al. 2004).

### 1.14.2.2   HetN

Whereas PatS is a crucial regulator of the initial heterocyst pattern, another protein called HetN serves to maintain the pattern after protracted periods of nitrogen starvation. Mutating *hetN* in *Anabaena* 7120 results in normal heterocyst spacing during the initial 24 hours after nitrogen stepdown, but a Mch phenotype emerges after roughly 48 hours (Black and Wolk 1994; Callahan and Buikema 2001).  The gene *hetN* encodes a protein with significant similarity to NAD(P)H-dependent oxidoreductases involved in the synthesis of fatty acids, poly-β-hydroxybutyrate, *nod* factors, and polyketides (Black and Wolk 1994).  HetN is observed to localize on both thylakoid and plasma membranes, and the overexpression of *hetN* prevents the induction of *hetR* during nitrogen starvation. These findings suggest that HetN may be an inhibitor of HetR (Li et al. 2002).  Like the gene *patS*, *hetN* encodes an oligopeptide containing the 'RGSGR' sequence that is involved in HetR inhibition, and when overexpressed, blocks heterocyst formation (Wu et al. 2004).  An *Anabaena* 7120 strain containing a double mutation of both *patS* and *hetN* triggers un-regulated heterocyst formation, where almost all vegetative cells differentiate into heterocysts (Borthakur et al. 2005).  Since the up-regulation of hetN is observed much later than patS (around 12 hours), and the Mch phenotype is only observed around 48 hours, it is believed that HetN role is to maintain a normal pattern once it has been established (Callahan and Buikema 2001).

*1.14.2.2   PatA*

The gene *patA* also contributes to heterocyst patterning in *Anabaena*. The protein

encoded by *patA* is unique: it contains a CheY-like phosphoacceptor at its C-terminus

and a helix-turn-helix (HTH) domain at its N terminus (Makarova et al. 2006). PatA-like

proteins are widespread in cyanobacterial genomes, and are involved in important cellular

activities such as phototaxis in *Synechocystis* sp. PCC 6803 (Yoshihara et al. 2000).

When *patA* is disrupted, the mutant forms terminal heterocysts exclusively and grows

poorly under nitrogen-fixing conditions (Liang et al. 1992). Additionally, an epistasis

analysis suggests that PatA may be involved in attenuating the inhibitory actions of both

PatS and HetN, promoting differentiation independent of the effects of PatS and HetN

activity (Orozco et al. 2006).
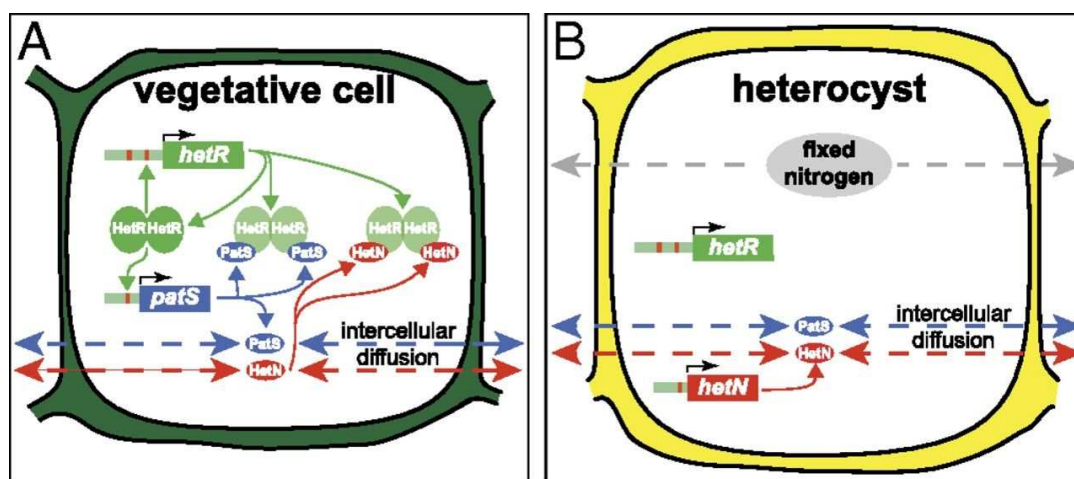


**Figure 14**. A simplistic model of the genetic network involved in heterocyst development. A) In vegetative cells, the HetR regulator activates expression of *hetR* and *patS*. B) In heterocysts, *hetN* is expressed and maintains proper heterocyst spacing, and fixed-nitrogen products are made. The products PatS and HetN can diffuse between cells in the filaments and bind to HetR, preventing it from DNA binding. Fixed-nitrogen products can also diffuse to other cells and contribute weakly to inhibit differentiation. Adapted from (Muñoz-García and Ares 2016).

*1.14.2.3  Concluding remarks on pattern formation*

In response to nitrogen starvation, the differentiation process is not reserved to a single vegetative cell but groups (clusters) of cells along the filament; with only one proheterocyst fully maturing into a heterocyst.  Studies on pattern formation have illuminated key agents that guide this complex process, including PatS, PatA, and HetN (detailed above).  Still, one of the central questions still looming from these studies is how only a single proheterocyst is allowed to fully mature, while the others regress back to into vegetative cells?  In other words, how can a proheterocyst prevent heterocyst formation in neighboring cells without succumbing to its own diffusible inhibitor?  These questions continue to perplex scientists, and as of consequence, a number of stimulating ideas have been put forward on the subject.  One posits that PatS may induce the synthesis of a separate, inhibitory molecule that is produced from neighboring vegetative cells that forms an inhibitory zone, forming areas that inhibit the production of proheterocyst clusters (denoted *Iv*; *I* for inhibitor and *v* for vegetative cell).  The amount of Iv produced is proportional to the number of proheterocysts in a cluster, and as it diffuses, it causes a number of proheterocysts in a cluster to regress.  When there is only one proheterocyst left in a cluster, the concentration of *Iv* may be so low that the remaining proheterocyst is allowed to fully mature (Flores and Herrero 2008).  Models based on the interplay between local autoactivation and early & late long-range inhibition point that the products of nitrogen fixation may also play an essential role in maintaining a normal heterocyst pattern (Muñoz-García and Ares 2016).  Computational modeling that integrates all aspects of cell activity, including internal regulatory process, diffusion

and background noise, may help to unveil the key mechanisms that govern the heterocyst

pattern in filaments (Torres-Sánchez et al. 2015).

## Chapter 2: Introduction & Objectives

**Introduction**

Photosynthesis is the ultimate life-sustaining biological process on the planet, and is the main driver behind the worldwide production of food, fiber, and renewable fuels. With the human population increasing exponentially, innovations in agriculture that will meet the demand for these resources are critically important. More solar energy reaches the Earth's surface every hour ($4.3 \times 10^{14}$ MJ) than is consumed annually ($4.1 \times 10^{14}$ MJ), making solar radiation a plentiful energy resource (Zhu et al. 2008). However, the energy losses of incident solar radiation from cell interception to the formation of chemical energy creates a limit on biomass yields and productivity. This is especially true for most crop plants, which exhibit relatively low photosynthetic efficiencies (2-4%) (Zhu et al. 2008). Furthermore, the extensive energy needed for planting, pesticide and herbicide application, harvesting, and year-round land management means that modern agriculture is a major energy user, and contributes to deforestation, soil erosion, and pollution (Foley et al. 2005; Montgomery 2007). To ensure sustainable agriculture for the future, innovations that can more efficiently fix and sequester carbon without negative effects to the environment are key areas for scientific investigation and economic investment.

Microalgae and cyanobacteria are photosynthetic microorganisms that have been researched extensively as models for carbon fixation (Chisti 2007; Parmar et al. 2011). They exhibit high photosynthetic efficiencies, fast growth rates, and inexpensive nutrient requirements compared to crop plants. These characteristics, along with their abundant composition of lipids and sugars, give them an advantage over terrestrial crops as a

carbon feedstock (Quintana et al. 2011). Over the last few decades, advancements in genomics, recombinant DNA technology, and synthetic biology have provided researchers with a sophisticated toolset for genetically engineering cyanobacteria to over-produce high-value compounds originally harvested from crops. These include naturally-synthesized carbon feedstocks, such as fatty acids and triacylglycerides (TAGs), and fermentable sugars (sucrose, glycogen) that can be used for biodiesel and bioethanol, respectively (Quintana et al. 2011). Furthermore, genetic engineering efforts have created cyanobacteria able to directly synthesize and secrete tailor-made compounds from their cells, eliminating the need for harvesting, processing or fermentation. These new products include fuels, as well as a diverse portfolio of medicines, polymers, food flavorings, fragrances, and industrial solvents (Ducat et al. 2011).

One group of compounds that have gained recent attention as a commercial product from engineered cyanobacteria is the terpenoids (also referred to as terpenoids), a large class of organic molecules that are naturally synthesized by plants, animals, bacteria, and archaea (Vranova et al. 2013). Their energetic composition make them ideal precursors for drop-in diesel and jet fuels (Rude and Schirmer 2009). Given their structural diversity, they also have industrial applications as solvents, nutraceuticals, natural pesticides, and drugs (Harrewijn 2001). During the last fifteen years, a number of genetic engineering projects have made great strides in re-programming different microorganisms to overproduce and secrete terpenoids from their cells (Wang et al. 2015). In time, these achievements may become the foundation of commercial systems that can utilize $CO_2$, $H_2O$, and solar radiation to produce terpenoids and similar molecules for human use.

Although great achievements have been made in this field, the cost of making terpenoids from cyanobacteria and microalgae still outweigh the risk for most commercial investors. This is primarily due to low productivities from currently engineered strains, technicalities of terpenoid collection, and high start-up costs (Parmar et al. 2011). To make large-scale terpenoid production from single-celled photosynthetic microorganisms a reality, future research and ingenuity will be needed to make the "photons-to-bioproduct" strategy economically viable. Increasing our knowledge of cyanobacterial genetics, metabolism, energy distribution, and defining the regulatory networks involved in chemical biosynthesis will expand our ability to predict the outcomes of genetic manipulation. This information, coupled with advancements in gene editing and programming will enable us to design better engineering strategies that increase productivity in large-scale photosynthetic systems.

**Objectives**

The objectives of this study was to:

1)  Genetically engineer *Anabaena* 7120 to synthesize farnesene by expressing of a codon-optimized plant farnesene synthase (Chapter 3).

2) Genetically engineer *Anabaena* 7120 to synthesize linalool by expressing a plant linalool synthase, and to increase productivity by co-expression of a DXP operon in $N_2$-fixing conditions (Chapter 4).

3) Remove glycogen synthesis in *Anabaena* 7120 to increase metabolic flux towards the MEP pathway, as well as analyze the effects of glycogen-deficiency on the nitrogen stress response (Chapter 5).

4) Assemble a synthetic photorespiratory bypass (SPB) in *Anabaena* to recycle products of the $C_2$ cycle into pyruvate for increased linalool production (Chapter 6).

# Chapter 3: Genetically engineering cyanobacteria to convert air (CO₂ and N₂), water, and light into the long-chain hydrocarbon farnesene

## *3.1 Abstract*

Genetically engineered cyanobacteria offer a shortcut to convert $CO_2$ and $H_2O$ directly into biofuels and high value chemicals for societal benefits.  Farnesene, a long-chained hydrocarbon ($C_{15}H_{24}$), has many applications in lubricants, cosmetics, fragrances, and biofuels.  However, a method for the sustainable, photosynthetic production of farnesene has been lacking.  Here, we report the photosynthetic production of farnesene by the filamentous cyanobacterium *Anabaena* sp. PCC 7120 using only $CO_2$, mineralized water, and light.  A codon-optimized farnesene synthase gene was chemically synthesized and then expressed in the cyanobacterium, enabling it to synthesize farnesene through its endogenous non-mevalonate (MEP) pathway.  Farnesene excreted from the engineered cyanobacterium volatilized into the flask head space and was recovered by adsorption in a resin column.  The maximum photosynthetic productivity of farnesene was $69.1 \pm 1.8$ $\mu g \cdot L^{-1} \cdot O.D.^{-1} \cdot d^{-1}$.  Compared to the wild-type, the farnesene-producing cyanobacterium also exhibited a 60% higher PSII activity under high light, suggesting increased farnesene productivity in such conditions. We envision genetically-engineered cyanobacteria as a bio-solar factory for photosynthetic production of a wide range of biofuels and commodity chemicals.

## 3.2 Introduction

Fossil fuel depletion and climate change due to elevated global $CO_2$ emissions are two issues that have far-reaching consequences for the world's future social and economic stability. These problems will be exacerbated by an increasing world population, which reached 7 billion people in 2012 and is projected to increase by another 1 billion by 2025 (Cohen 2003). The concerns regarding fossil fuel consumption and $CO_2$ emissions have led to biotechnologies that convert renewable biomass feedstocks into biofuels. The production of corn ethanol and soybean biodiesel, known as "first generation" biofuels, has gained considerable interest in the last few decades (Nigam and Singh 2011). Advancements in agricultural engineering and biotechnology have further diversified the plant feedstocks which can be used to produce biofuels, such as cellulose and hemicellulose from switchgrass and corn stover (Gomez et al. 2008). However, there are still concerns over the sustainability of these "second generation" biofuels, since they require clean water, fertilizers, pesticides, arable land, and downstream processing steps that drive up production costs and ultimately the price of the fuel.

Volatile, long-chained alkanes and alkenes are of particular interest as biofuels, considering their high energy density, similarities to existing fuels (gasoline, diesel, jet), and propensity to separate from the processing residues. Alkenes, such as terpenes, have recently gained interest as possible biofuels (Gronenberg et al. 2013; Peralta-Yahya and Keasling 2010; Tippmann et al. 2013). Terpenes represent a large and diverse class of organic compounds found in all living organisms They are essential metabolites in bacteria, plants, and humans, and they play important roles in cell wall and membrane

biosynthesis (hopanoids), electron transport (ubiquinones and menaquinones), and the conversion of light into chemical energy (chlorophylls and carotenoids), among other processes (Perez-Gil and Rodriguez-Concepcion 2013).

All terpenes are derived from prenyl diphosphate precursors, synthesized from the mevalonate (MVA) pathway and/or the 2-*C*-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate (MEP) pathway (also known as the non-mevalonate pathway). However, with the exception of plants, most organisms have only one of these two pathways. The MVA pathway is mainly present in archaebacteria, some gram-positive bacteria, yeast and animals (Vranova et al. 2013), while most gram-negative bacteria, cyanobacteria, plant plastids, and green algae have only the MEP pathway (Vranova et al. 2013). In the MEP pathway (Figure 15), pyruvate and glyceraldehyde-3-phosphate (GAP) are condensed and brought through a series of enzymatic reactions, eventually leading to the basic five-carbon precursors for all terpenes: isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). IPP and DMAPP further condense to form geranyl pyrophosphate (GPP), the precursor for all $C_{10}$ monoterpenes and their derivatives. A second condensation step with GPP and IPP forms farnesyl pyrophosphate (FPP), the precursor for all $C_{15}$ sesquiterpenes and their derivatives.

The last decade has seen a surge of research exploring the potentials of using engineered microorganisms as "cellular factories" to produce sesquiterpenes and their derivatives as precursors for biofuels and therapeutic drugs. A majority of this effort has focused on enhancing terpenoid pathways in *Escherichia coli* and *Saccharomyces cerevisiae* through genetic engineering (Buijs et al. 2013; Martin et al. 2003b; Peralta-Yahya et al. 2012; Pitera et al. 2007; Wang et al. 2011). Significant advancements in

pathway engineering have recently allowed for scale up to industrial-sized facilities. For example, Amyris Biotechnologies, a renewable products company headquartered in Emeryville CA, developed an engineered *S. cerevisiae* to convert sugarcane sucrose to farnesene ($C_{15}H_{24}$), a volatile sesquiterpene with commercial applications in lubricants, cosmetics, fragrances, and biofuels (Buijs et al. 2013). This system was able to produce 1 million liters of farnesene during a 45-day period (Velasco 2013). However, using sugar feedstocks for farnesene production is an added cost factor that puts these facilities at an economic disadvantage compared to the direct photosynthetic production of farnesene.

| DXP | 1-deoxy-D-xylulose 5-phosphate | DXS | DXP synthase |
| MEP | methylerythritol-4-phosphate | DXR | DXP reductoisomerase |
| CDP-ME | diphosphocytidylyl methylerythritol | MCT | CDP-ME synthase |
| CDP-MEP | CDP-methylerythritol-2-phosphate | CMK | CDP-ME kinase |
| ME-cPP | methylerythritol-2,4-cyclodiphosphate | MDS | ME-cPP synthase |
| HMBPP | hydroxymethylbutenyl diphosphate | HDS | HMBPP synthase |
| IPP | isopentenyl diphosphate | HDR | HMAPP reductase |
| DMAPP | dimethylallyl-diphosphate | IDI | IPP:DMAPP isomerase |
| GPP | geranyl-diphosphate | GPPS | GPP synthase |
| FPP | farnesyl-diphosphate | FPPS | FPP synthase |
| GGPP | geranylgeranyl-diphosphate | GGPPS | GGPP synthase |
| | | SQS | squalene synthase |

**Figure 15.** Engineering cyanobacteria to produce hydrocarbons ($C_{10}$ and $C_{15}$). Known MEP pathway exists in plants and some bacteria (highlighted in green). The farnesene synthase and farnesene metabolite that are not present in cyanobacteria and targeted for engineering are highlighted in red). *NADPH as cofactor. **ATP dependent

Photosynthetic microorganisms such as algae and cyanobacteria hold a distinct advantage in biofuel production over plants and yeast. They have faster growth rates and higher photosynthetic efficiencies than plants (>10% in cyanobacteria compared to 1-3% in plants), are able to directly utilize $CO_2$ as a carbon source, and are genetically amenable (Ducat et al. 2011; Parmar et al. 2011; Ruffing 2011). Cyanobacteria with re-wired metabolic pathways have recently been created through genetic engineering, and they possess the ability to convert $CO_2$ chemicals and biofuels which can be secreted directly from their cells. Examples include cyanobacteria engineered to produce ethanol (Dexter and Fu 2009), isobutyraldehyde (Atsumi et al. 2009), isoprene (Lindberg et al. 2010), fatty acids (Liu et al. 2011), and hydrogen (Sakurai et al. 2013).

*Anabaena* sp. PCC 7120 (hereinafter referred to as *Anabaena* 7120) is a filamentous, nitrogen-fixing cyanobacterium. We have previously engineered this strain to produce and secrete limonene, a cyclic $C_{10}$ monoterpene that could potentially be used as a biofuel (Halfmann et al. 2014b). Here, we report the viability of using this cyanobacterium to synthesize farnesene, a $C_{15}$ sesquiterpene, from $CO_2$, mineralized water, and light by introducing a codon-optimized farnesene synthase (*FaS*) gene from Norway spruce into the *Anabaena* 7120 genome. We envision this platform of using $CO_2$, water, and sunlight as a sustainable feedstock to be applicable for the production of a wide range of industrial chemicals and drop-in-fuels.

### 3.3 Materials & Methods

*Strains and growth conditions*

   *Anabaena* sp. PCC 7120 was grown in BG11 medium supplemented with 20 mM

HEPES buffer under constant white light (50 $\mu E \cdot m^{-2} \cdot s^{-1}$ at the culture surface) while

bubbling with a 0.22 $\mu$m filtered $CO_2$ and air (1:100 ratio).  Cultures were grown at 30°C

and shaken at 120 rpm in a temperature controlled Innova 44R lighted incubator (New

Brunswick Scientific).  Light intensity of growing cultures was measured using a Li-Cor

250A light meter.  *E. coli* strain NEB 10β (New England Biolabs) was grown in Luria-

Bertani (LB) medium at 37°C, and 250 rpm.  Kanamycin (Km; 50 $\mu g \cdot mL^{-1}$) was used for

antibiotic selection for *E. coli*, and Neomycin (Nm; 100 $\mu g \cdot mL^{-1}$) was used for transgenic

*Anabaena* 7120 harbouring pFaS.

*Plasmid construction*

   A codon-optimized version of  a farnesene synthase gene (GenBank access #:

KJ847279) from Norway spruce (Martin et al. 2004) with a deleted stop codon and an N-

terminal His$_6$ tag was synthesized by DNA 2.0 (Menlo Park, CA) and cloned into the

*Bgl*II-*Sal*I sites of  the shuttle vector pZR1188 (Halfmann et al. 2014b).  The final

construct, pFaS (Figure 16) was transformed into *E. coli* strain NEB-10β, which was used

for conjugal transfer of the plasmid into *Anabaena* 7120.
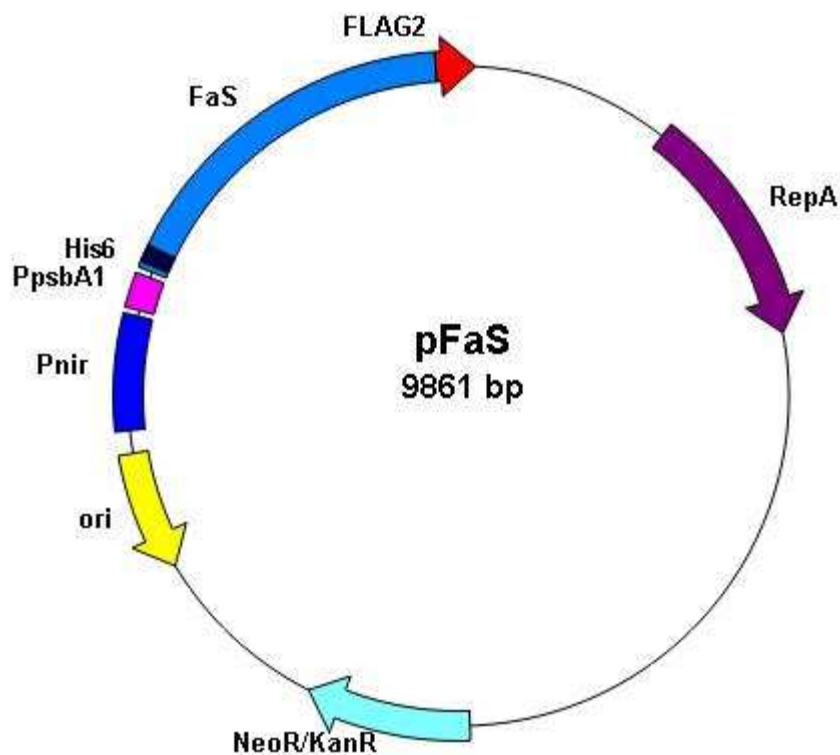
**Figure 16**.  Diagram of pFaS plasmid.  *Pnir*, *nir* promoter of *Anabaena* sp. PCC 7120; *Ppsba1*, *psbA1* promoter of *Anabaena* sp. PCC 7120; His$_6$, 6× histidine tag; Flag$_2$, 2× FLAG tag; FaS, farnesene synthase ORF; *RepA*, replication protein A; *Neo$^R$/Kan$^R$*, neomycin/kanamycin antibiotic gene; *ori*, origin of replication.

*Conjugal transfer of pFaS into Anabaena 7120*

The transfer of pFaS into *Anabaena* 7120 was performed via bacterial conjugation

using a bi-parental *E. coli* mating system (Elhai et al. 1997) with the following

modifications.  *E. coli* HB101 bearing helper (pRL623) and conjugal (pRL443) plasmids

were mated with *E. coli* NEB 10β bearing pFaS, and selected on LB plates containing

triple antibiotic selection for the three plasmids.  Selected colonies were grown overnight

in 2 mL of LB containing appropriate antibiotics, subcultured by adding 200 µL of

overnight culture to 2 mL of fresh LB containing appropriate antibiotics, and grown for

additional 3 h.  Cells were harvested by centrifugation at 3000×g, washed 3 times with 1

mL of LB to completely remove antibiotics, and resuspended with 200 µL LB and subjected for mating with *Anabaena* 7120.

A 10 mL culture of wild-type (WT) *Anabaena* 7120 was grown to early exponential stage (O.D.$_{700 nm}$ of 0.3) and then sonicated (Branson 1510 water bath sonicator) for 60-120 sec to break filaments into 2-4 cell lengths, which were confirmed under light microscopy. Cells were harvested by centrifugation at 5,000×g for 10 min, resuspended in 200 µL of fresh BG11, and then mixed with the above *E. coli* harboring triple plasmids for conjugal transfer. This mixture was placed under lighted conditions at 25°C for 30 min, micro-pipetted on to an autoclaved Millipore Immobilon nitrocellulose filter (HATF08550) placed on a BG-11 with 5% LB plate, and grown under white light for 3 days at 30°C. The filter was then transferred to a BG11 plate containing Nm (100 µg·mL$^{-1}$) to select for positive exconjugants. Individual exconjugant colonies were further purified by restreaking onto fresh BG11 plates containing 100 µg·mL$^{-1}$ Nm and routine colony PCR was used for verification of positive exconjugants.

*Protein extraction and Western blot analysis*

Protein was extracted from both WT and FaS *Anabaena* 7120 cells to determine expression of FaS in the host strains, as described previously (Halfmann et al. 2014b). Five µL protein supernatants were loaded and separated by 12% SDS-PAGE at 200V for 30 min, transferred onto a PVDF membrane, and FaS-FLAG$_2$ was detected by Western blotting using anti-FLAG (Sigma Aldrich) antibodies at a dilution of 1:5000. The membrane was then stained with Coomassie Blue R250 to visualize total protein content.

*Farnesene collection from culture headspace and GC-MS analysis*

Starter cultures were set at an O.D.$_{700 \text{ nm}}$ of 0.02 using a UV-Vis spectrophotometer (Genesys 10S, Thermo Scientific). WT and FaS *Anabaena* 7120 strains were grown in 250 mL stoppered Erlenmeyer flasks containing 100 mL BG11 supplemented with 20 mM HEPES buffer and 100 µg·mL$^{-1}$ Nm, and bubbled with 0.22µm filtered $CO_2$ and air (1:100 ratio) at a rate of 100 mL·min$^{-1}$ in the conditions described in Section 2.1 for 15 days, while replacing the culture media every 5 days. A small glass column filled with 100 mg of Supelpak 2SV resin (Sigma Aldrich) was attached to the exhaust port of each flask to capture farnesene and other volatile metabolites from the culture headspace (Figure 17). Fresh 2SV resin columns were replaced on the culture flasks every 3 days. To analyse volatiles from the resin, samples were eluted twice with 2.5 mL of pentane containing 1 µg·mL$^{-1}$ tetracosane as an internal standard (IS), pooled in 5 mL total volumes, evaporated to 1 mL using a gentle stream of $N_2$ gas (thus concentrating the IS to 5 µg·mL$^{-1}$), and subjected to gas chromatography-mass spectrometry (GC-MS) analysis (Agilent 7890A/5975C) at the Functional Genomic Core Facility of South Dakota State University. One µL injected samples were separated using a HP-5MS column (30m×250µm×0.25µm), with $H_2$ as the carrier gas. The oven temperature was initially held at 40°C for 1 min, increased 5°C·min$^{-1}$ until 180°C was reached, and then further increased by 20°C·min$^{-1}$ to 320°C. Compounds were identified using the NIST MS library v2.0 and authentic standards.
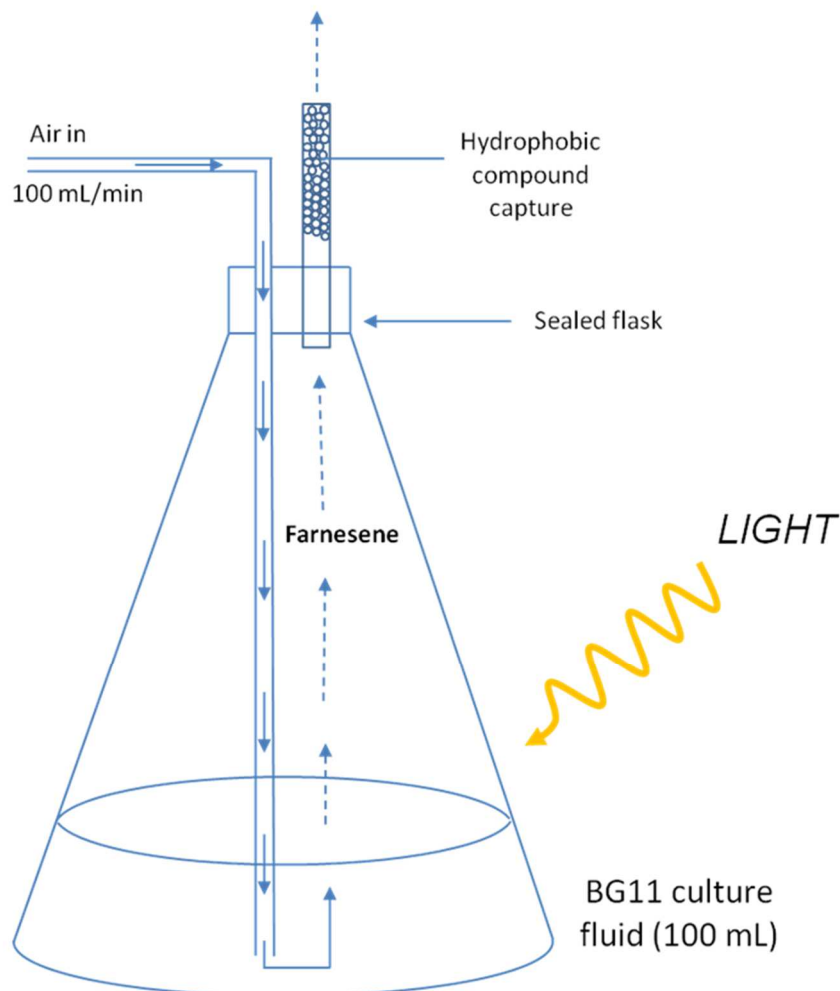
**Figure 17.** Diagram of the flask apparatus for capturing volatile compounds emitted from cyanobacterial cultures by Supelpak 2SV resin column.

*Quantification of farnesene and absorption test of 2SV resin*

A set of trans-β-farnesene standards (Sigma Aldrich no. 73942) at 1, 5, 10, 25, 50, 100, and 200 µg·mL$^{-1}$ and 5 µg·mL$^{-1}$ tetracosane in pentane was used to create a standard curve to quantify farnesene captured by 2SV columns. To test absorption capacity of the 2SV resin, 1, 2, 3, 4, and 5 µL aliquots of trans-β-farnesene standard were pipetted into separate 10-mL sealed glass vials, with 100 mg of 2SV resin inside a glass column attached the exhaust port on each vial. A gentle stream of air was injected into the glass vial for 3-5 minutes, allowing the farnesene to evaporate and pass into the resin. The

resin was then washed with pentane and mixed with the tetracosane IS as described above. These farnesene/tetracosane peak ratios were compared to the standard curve to calculate the recovered amounts of farnesene. An α-farnesene standard peak was generated by grinding approximately 100 mg of Granny Smith apple (*Malus domestica* × *M. sylvestris*) peel with a mortar and pestle, and sonicating the grounds in 2 mL of room-temperature pentane with a Branson Digital Sonifier at 70% amplitude for 30 sec. The extracts in organic phase were then concentrated by evaporating pentane into 0.5 mL using a gentle stream of $N_2$ and analyzed using GC-MS.

*Chlorophyll extraction and analysis of photosynthetic activity*

Chlorophyll was extracted using a 90% methanol extraction procedure described previously (Meeks and Castenholz 1971). Oxygen evolution was quantified in 1 mL culture samples (adjusted to a cell density that has 10 μg chlorophyll $mL^{-1}$) in BG11 with 10 mM $NaHCO_3$ using a Clark-type electrode and DW2 Oxygen Electrode Chamber with $O_2$ View Oxygen Monitoring software (Hansatech). Light was introduced to samples using an LS2/H Tungsten-halogen 100 W light source and adjusted with neutral density filters. Light intensity and sample temperature was monitored using a Quantitherm light-temperature meter during experimentation. All measurements were performed in triplicate.

### 3.4 Results

*Expression of a farnesene synthase (FaS) in Anabaena 7120*

We first sought to utilize the native MEP pathway in *Anabaena* 7120 to produce farnesene via photosynthesis. Although *Anabaena* 7120 possesses an endogenous MEP

pathway, a BLAST search of the genome using a characterized FaS protein sequence (Martin et al. 2004) resulted in no similarities (0.1 E value as cut-off). We then acquired a farnesene synthase gene (PaTPS-Far; Genbank accession no. AY473627.1) from Norway spruce (*Picea abies*), which was previously characterized as an E, E-α-farnesene synthase through an *in vitro* enzymatic activity assay (Martin et al. 2004). PaTPS-Far was then codon-optimized (DNA 2.0, Menlo Park, CA) for optimal expression in the cyanobacterium. We included an N-terminal $His_6$ tag and a C-terminal $FLAG_2$ tag to verify protein expression, and cloned the gene into the replicating shuttle vector pZR1188, a high-copy plasmid with the ability to replicate independently from the chromosome, and maintains stability in the cell without chromosomal integration. The resulting construct was named pFaS (Figure 16). This construct was then transferred into WT *Anabaena* 7120 using conjugation with an *E.coli* donor. *FaS* transcription was driven by a dual *Anabaena* 7120 *nir-psbA1* promoter to maintain maximal protein expression constitutively in the cells. As seen in Figure 18 an expected 70 kDa protein was detected from immunoblotting with anti-FLAG antibodies in FaS (lane 2), but not in WT *Anabaena* 7120 (lane 1), suggesting that FaS was expressed in the transgenic *Anabaena* 7120.
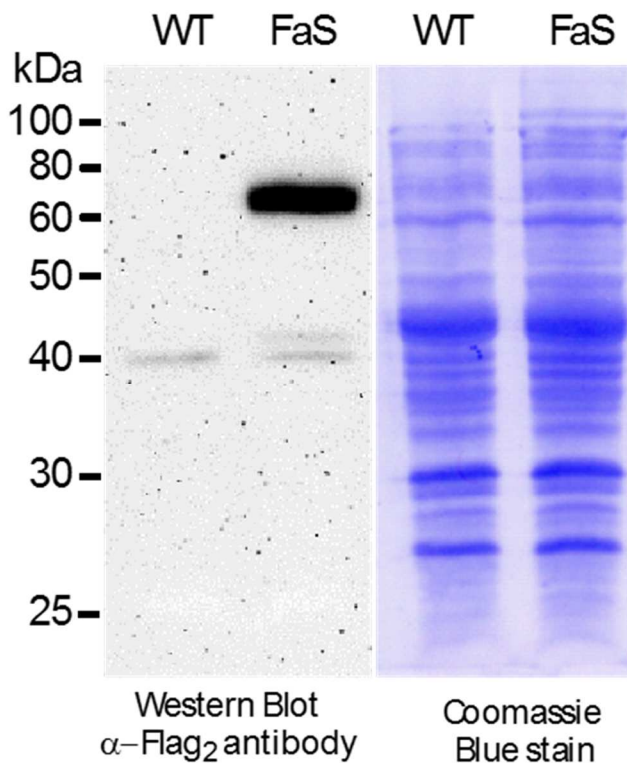
**Figure 18.** Western blot analysis revealed a 70-kDa FaS-FLAG protein in FaS Anabaena bearing pFaS plasmid (lane 2) but absent in the WT sample (lane 1). The same PVDF membrane was stained with Coomassie blue R250 (lane 3 and 4) to show equal protein loading.

*Identification of farnesene emitted from FaS Anabaena 7120 using GC-MS*

Our previous studies with engineered *Anabaena* 7120 strains capable of producing

limonene have shown that volatile compounds are emitted from the cells and volatilized

into the flask headspace (Halfmann et al. 2014b). This volatilization is especially rapid

when the culture is bubbled with air or $CO_2$ enriched air to provide carbon. To collect

these chemicals, we passed the exhaust gas from the headspace of a culture flask through

a resin column filled with Supelpak 2SV resin (Sigma Aldrich) as shown in Figure 17.

Our absorption tests with a trans-β-farnesene standard as described showed that 100 mg

of 2SV resin could absorb at least 1 mg of farnesene. To determine whether the

expressed FaS was functionally active and produce farnesene in *Anabaena* 7120.  We

collected the volatile compounds emitted from the culture fluid through a glass column

filled with 2SV resin, and then analysed the resin pentane extract using GC-MS. Figure

19 shows GC-MS chromatographs of WT *Anabaena* 7120 and FaS *Anabaena* 7120.  A

prominent peak appears at 21.4 minutes in the FaS *Anabaena* 7120 chromatograph

(Figure 19B) but not in WT *Anabaena* 7120 (Figure 19A).  This new peak in FaS

*Anabaena* 7120 was identified as α-farnesene by the MS library (NIST MS Library v2.0).
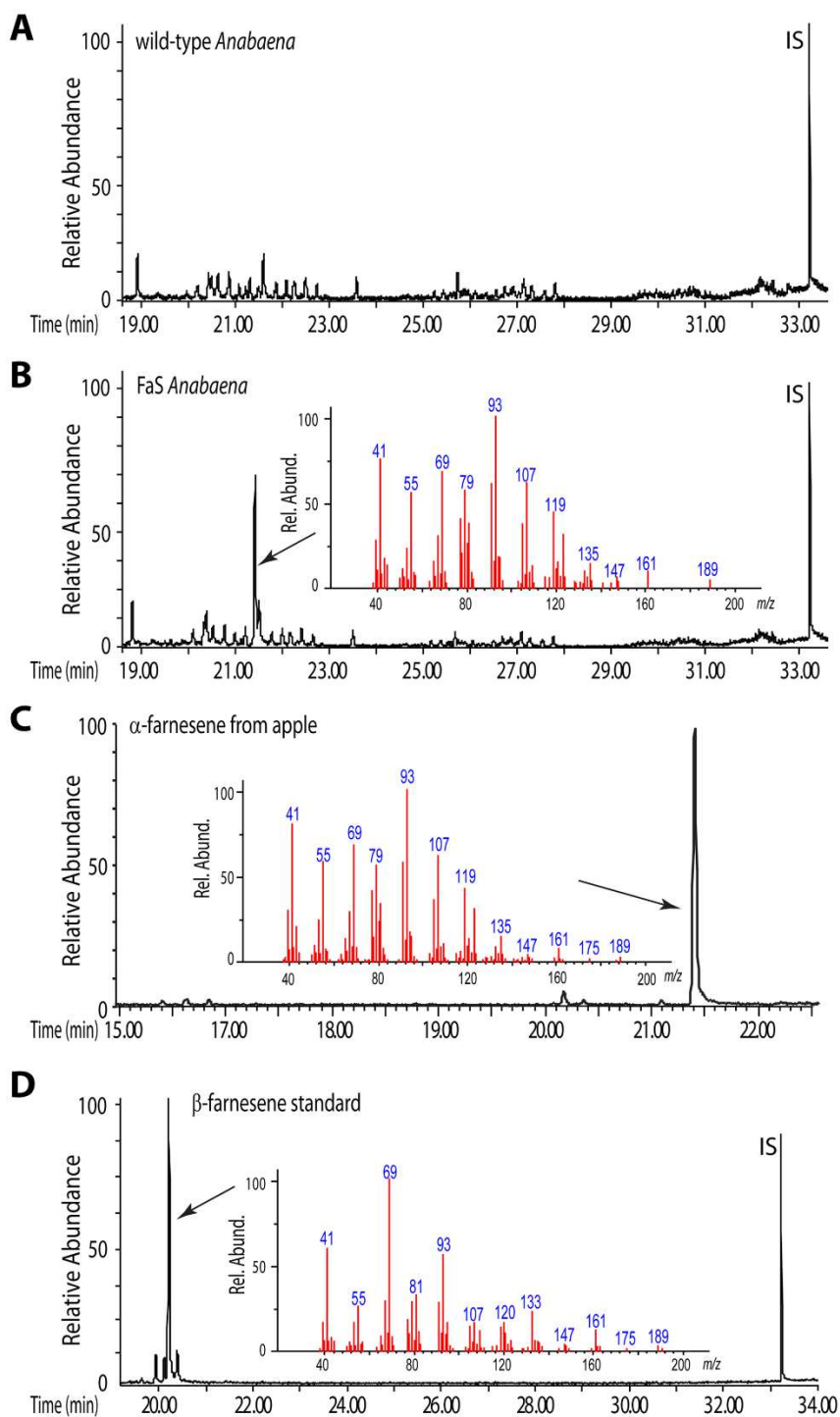
**Figure 19.** GC-MS chromatographs of flask headspace samples from WT (A) and FaS (B) *Anabaena* cultures. A peak at the rentention time 21.4 minutes (arrowed) found in FaS *Anabaena* matches the α-farnesene peak found in rind of green apple (C), but is absent in the WT (A). Mass specra (inserted red) of these peaks also display the expected pattern for farnesene. Trans-β-farnesene (D) was used to create a standard curve for farnesene quantification. Five µg ml$^{-1}$ tetracosane serves as an internal standard (IS**).**

Since α-farnesene is not commercially available, we compared the probable farnesene peak from the transgenic *Anabaena* 7120 to authentic farnesene from the rind of green apple, which is known to be abundant with E,E-α-farnesene (Pechous and Whitaker 2004). GC-MS revealed a major peak in the rind extract (Figure 19C), having the same retention time (21.4 min) and ion fragmentation pattern as the probable farnesene peak from FaS *Anabaena* 7120 (Figure 19B), and identified as α-farnesene by the MS library. This is consistent with the reported enzymatic product being E,E-α-farnesene (Martin et al., 2004). No farnesene was detected in the culture medium extracts or homogenized cell extracts (data not shown), suggesting an efficient transfer of farnesene from the growing culture to the headspace. Taken together, the genetically engineered cyanobacterium bearing the codon-optimized farnesene synthase gene from Norway spruce is capable of producing and emitting α-farnesene.

To quantify farnesene production from Fas *Anabaena* 7120, we used a commercially available trans-β-farnesene as a standard. Although trans-β-farnesene peak gave a shorter retention time (20.2 min, Figure 19D) than the peak observed in the FaS *Anabaena* 7120, as well as α-farnesene peak from apple extract (21.4 min, Figure 19B and C, respectively), the fragmentation pattern was similar to α-farnesene (comparing the inserted figure in Figure 19D to that in 19C). Therefore, trans-β-farnesene standard was used to create a calibration curve with tetracosane serving as an IS.

*Farnesene production in FaS Anabaena 7120 using $CO_2$, light, and mineralized water*

We employed a continuous growth and farnesene recovery system to analyze the amount of farnesene produced from FaS *Anabaena* 7120 over a 15 day growth period. To measure farnesene production, the 2SV columns fitted to each culture flask were replaced every three days, and volatiles from the resin samples were eluted with pentane and quantified by GC-MS using a standard curve created by trans-β-farnesene. Our results showed that FaS *Anabaena* 7120 produced a total of $45.7\pm2.5$ µg·$L^{-1}$ farnesene from days 0-3, and these amounts continued to increase over the remainder of the growth period, to approximately $93.3\pm9.6$ µg·$L^{-1}$ from days 13-15 (Figure 20A). Total farnesene accumulation was measured at $305.4\pm17.7$ µg·$L^{-1}$ during 15 days of growth. We further assessed the specific farnesene production rate by monitoring culture O.D.$_{700nm}$ every 3 days (Figure 20B) and comparing it to the amount of farnesene produced during each three-day collection period (Figure 20C). Water loss due to evaporation was accounted for by refilling culture flasks with sterilized dd$H_2O$ to proper culture volume before culture density analysis. The highest farnesene production rate was observed during the first 3 days, with a rate of $69.1\pm1.8$ µg·$L^{-1}$·O.D.$^{-1}$· day $^{-1}$. The productivity decreased during the remainder of the growth trial, leveling off at 15-20 µg·O.D$^{-1}$·$L^{-1}$·day $^{-1}$.
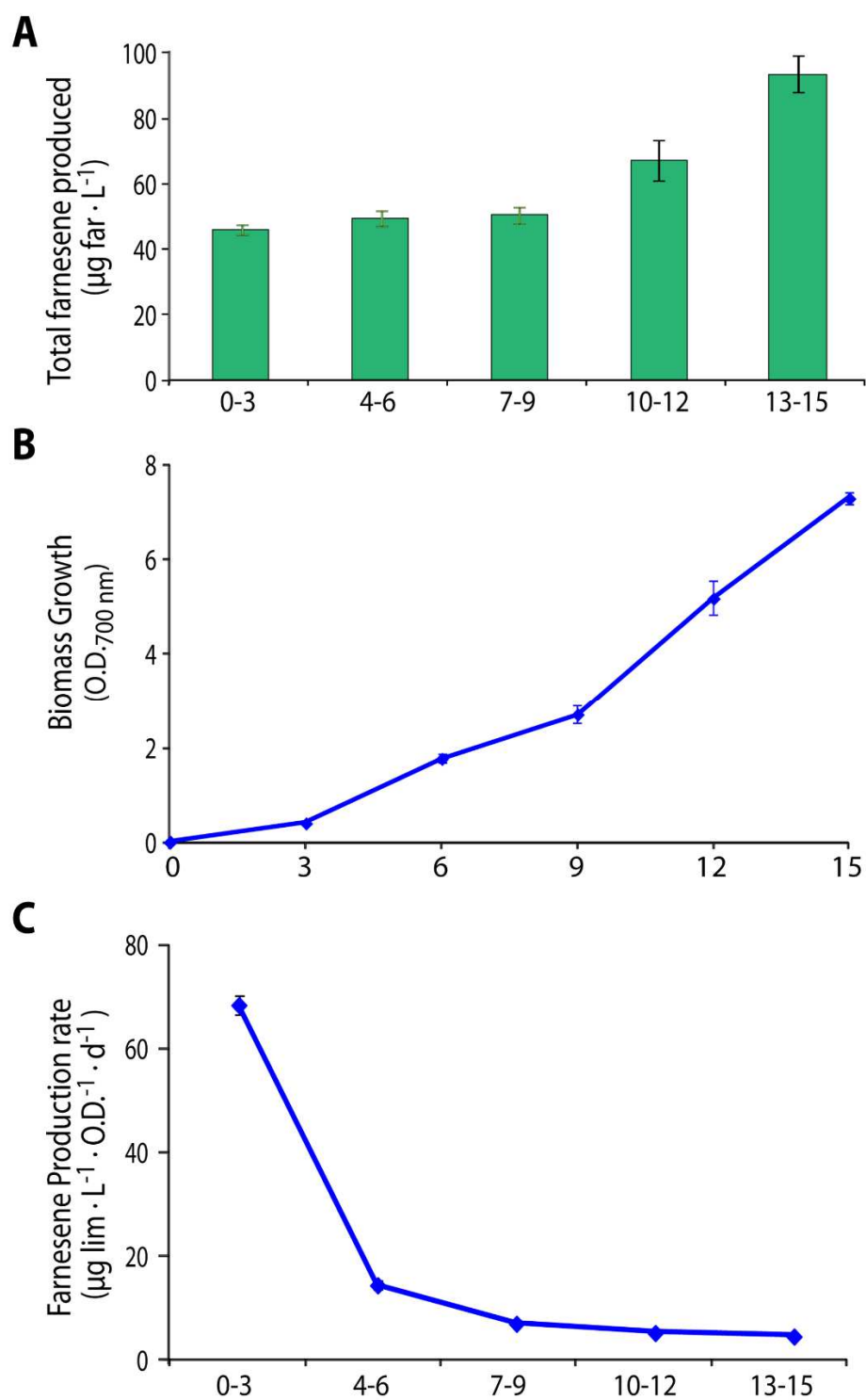
**Figure 20.** Growth and farnesene production characteristics of genetically engineered FaS *Anabaena*. A. Total farnesene captured in 2SV column every 72 hours from FaS *Anabaena*. B Biomass growth of FaS *Anabaena*. C. Rate of farnesene produced from FaS *Anabaena* per optical density. The horizontal axis represents time (day). Results are the average of three biological replicates, with error bars representing standard error of the mean.

*Farnesene production enhances photosynthetic activity*

In order to test the effects of farnesene synthesis on photosynthetic activity and biomass yield in the engineered cyanobacterium, we measured biomass growth and oxygen evolution rates (Figure 21) of the water-splitting reaction of photosystem II (PSII) from WT and FaS *Anabaena* 7120. We found no significant difference in biomass accumulation between WT and FaS *Anabaena* 7120 (data not shown). The rate of oxygen release from the cells of FaS *Anabaena* 7120 was similar to the WT in the light ranges of 50- 500 $\mu E \cdot m^{-2} \cdot s^{-1}$. However, at high light intensities ($\geq 1000$ $\mu E \cdot m^{-2} \cdot s^{-1}$), oxygen evolution from FaS *Anabaena* 7120 was approximately 60% higher compared to the WT. Interestingly, FaS *Anabaena* 7120 maintained similar oxygen evolution rates ($\sim 3$ $\mu mol$ $O_2 \cdot min^{-1} \cdot mg$ $chl^{-1}$) at 1000 and 1800 $\mu E \cdot m^{-2} \cdot s^{-1}$, while the WT exhibited a slight decrease in this light range (Figure 21).
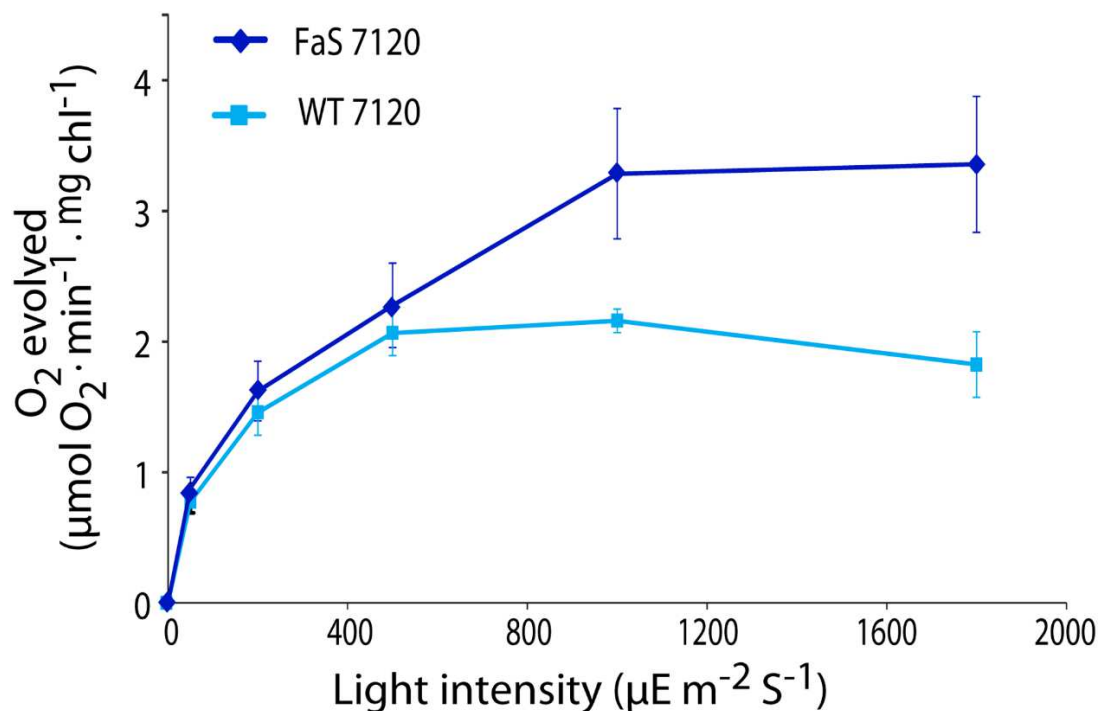
**Figure 21**. Photosynthetic oxygen evolution of WT and FaS *Anabaena* under indicated light intensities. The culture samples adjusted to a cell density of 10 μg chlorophyll ml$^{-1}$ were tested in the presence of 10 mM of NaHCO$_3$ in BG11 media.

### 3.5 Discussion

The conversion of CO$_2$ to terpenes using a cyanobacterium as a bio-catalyst has been

widely researched in recent years.  It was first demonstrated in the unicellular

cyanobacterium *Synechocystis* sp. PCC 6803, with the photosynthetic production of

isoprene (C$_5$H$_8$) by heterologous expression of a codon-optimized isoprene synthase

(IspS) from *Pueraria montana* (Lindberg et al. 2010).  Initial production rates using a

gaseous/aqueous two-phase bioreactor were 2 μg isoprene·L$^{-1}$·hr$^{-1}$ during the first 48

hours of growth (Bentley and Melis 2012).  Our results show slightly lower production

rates of farnesene (approximately 0.625 μg·L$^{-1}$·hr$^{-1}$ during the first 72 hours; 1.29 μg·L$^{-1}$·hr$^{-1}$ during the last 72 hours of growth).  This is likely due from differences in

bioreactor design and terpenoid extraction methods, catalytic efficiency of the expressed

enzyme, or differences in precursor metabolite pools of DMAPP and FPP for isoprene and farnesene respectively. In addition, the expression of FaS on a replication vector can result in multiple copies of *fas* per chromosome, in contrast to chromosomal integration methods which theoretically produce one gene copy per chromosome in purified cultures. Recent studies have found that pDU1-based plasmids can have widely variable copy numbers in *Anabaena* 7120, ranging from 0.53-1812 per chromosome (Yang et al. 2013). These variables together likely underlie the differences in terpene production rates found between research groups.

Compared to the other biosynthetic pathways, allocation of fixed carbon into the MEP pathway is low, constituting only 3-5% of the total biomass in cyanobacteria (Lindberg et al. 2010). In order to increase metabolite flux for enhanced terpenoid production, over-expression of key rate-limiting MEP pathway enzymes was previously shown to improve the production of limonene in *Anabaena* 7120 (Halfmann et al. 2014b). Heterologous expression of the whole MVA pathway in *Synechocystis* 6803 also increased isoprene production 2.5-fold, as it endowed the cyanobacterium to draw more photosynthetic metabolites into isoprene production alongside its native MEP pathway (Bentley et al. 2014). In order to exploit the full potential of designed pathways in cyanobacteria for farnesene production, researchers will need to modulate transcriptional and translational elements of enzyme networks, in order to relieve bottlenecks, suppress feedback inhibition, and ultimately maximize carbon flow through the MEP pathway. These strategies have been researched in cyanobacteria through modulation of ribosomal binding sites (Oliver et al. 2014) and promoters (Huang et al. 2010; Huang and Lindblad 2013; Qi et al. 2013). The suppression of alternate pathways downstream of FPP

synthesis (e.g. squalene sythesis) can also contribute to an increase in carbon flux into sesquiterpenes, previously demonstrated by replacing the native squalene synthesis promoter in engineered *Saccharomyces cerevisiae* with a methionine-repressible promoter, which resulted in a five-time increase in amorphadiene production (Paradise et al. 2008).

Although the total amount of farnesene produced by culture flasks increased every 3 days, we noticed that the specific rate of farnesene production dropped almost 78% from day 3 to 6, until it level off during the remainder of the growth trial. We attribute this decline in farnesene productivity to an elevated cell density (O.D.) in culture flasks towards the end of the growth trial, which shaded the cells and inhibited the amount of photons collected. This could decrease photosynthetic activity and the reducing power needed for carbon flux through the MEP pathway. In cyanobacteria, enzymes in the MEP pathway has been found to interact closely with re-dox partners involved in light-independent reactions, making the MEP pathway tightly regulated by photosynthesis (Okada and Hase 2005). Low light quality could also affect the transcriptional activity of the engineered *psbA1* promoter, which known to be induced by light in cyanobacteria (Golden 1995). Indeed, increasing the light intensity during growth conditions has shown to significantly improve productivity and yield in biofuel-producing cyanobacteria (Halfmann et al. 2014b; Ungerer et al. 2012).

The observed increase in $O_2$ production from the *Anabaena* 7120 strain endowed with the ability to synthesize farnesene is a phenomenon that has been observed in previous studies on cyanobacteria engineered to produce hydrocarbons (Bentley et al. 2013; Oliver et al. 2013). The observed enhancement in photosynthetic activity of FaS

*Anabaena* 7120 under increasing light intensities could be explained by farnesene's role as an added carbon sink in the cyanobacterium. The continuous synthesis of farnesene could help maintain a relatively oxidized electron transport chain and help suppress photoinhibition, caused by imbalances between light absorption and carbon fixation. This opening of a new carbon sink would therefore create a positive feedback in the light reactions for photochemical energy production, resulting in more $O_2$ production compared to the WT. Similar traits such as increased PSII activity, chlorophyll content, and carbon fixation were also found in a strain of *Synechococcus elongatus* sp. PCC 7942 engineered to accumulate and excrete sucrose (Ducat et al. 2012).

The fine tuning of the modular units in biofuel pathways will be a crucial next step in engineering fourth-generation biofuel systems, as the current rates of fuel precursors produced from engineered cyanobacteria will need to be increased to make large-scale systems economically realistic. Previously, we used parameters from algae biomass production models and a hypothetical "greenhouse-ethanol plant" relay system to calculate an ideal limonene productivity using cyanobacteria (Halfmann et al. 2014b). Using the same approach, we calculated that a 6.4 million L PBR capable of converting 1% of the annual $CO_2$ emissions from a single ethanol plant (275000 metric tons) into farnesene could produce roughly 851.3 metric tons (1047117 L) of farnesene annually, about 30% of the annual farnesene yields reported by Amyris in 2013 (Voegele 2014). Assuming rough estimates for capital, labor and operation, a total production costs of $1.30 \cdot L^{-1}$ farnesene is possible, compared to the current $4.00 \cdot L^{-1}$ production cost reported by Amyris. This considerable cost difference emphasizes the advantages of

fourth generation biofuels, compared to biofuel-production systems that utilize land plants as a feedstock.

Cyanobacteria are ideal organisms for biofuel production, due to their fast growth rates, penchant for converting carbon into energy-rich molecules, and ability to secrete these molecules out of the cell. Their genetic amenability also provides for a way to open new metabolic channels, allowing for the creation of a myriad of chemical products. As a proof-of-concept, we show that expressing a chemically synthesized farnesene synthase *(faS)* gene in *Anabaena* 7120 led to photosynthetic production of farnesene directly from $CO_2$ and $H_2O$. This direct "$CO_2$ to fuel approach" completely bypasses the expensive, multi-step processes currently used to produce biofuels from plant biomass. We envision $CO_2$ conversion systems using engineered cyanobacteria to pave the way for a greener, post-petroleum era.

**Chapter 4: Biosolar synthesis of the floral fragrance linalool from the $N_2$-fixing cyanobacterium *Anabaena* sp. PCC 7120**

## *4.1 Abstract*

Linalool ($C_{10}H_{18}O$) is a naturally-occurring terpene alcohol, emitted as an aroma molecule from many flowers and aromatic plants. Its use in fragrances, flavors, and pharmaceuticals have made it a sought-after commodity chemical. Here, we demonstrate an alternative production strategy for this aromatic compound using only light, $CO_2$, $N_2$, and mineralized water. A plant linalool synthase (LinS) from *Picea abies* was expressed in the $N_2$-fixing cyanobacterium *Anabaena* sp. PCC 7120 via a cyanobacterial expression vector, and linalool was quantitated from the flask headspace of engineered cultures. To increase MEP pathway flux, we co-expressed LinS with three additional enzymes in the MEP pathway (DXS-IDI-GPPS) to create a LinSDXP operon. Our results demonstrate synthetic biology strategies to improve linalool production in cyanobacteria, while also showing that $N_2$-gas is a sufficient nitrogen source for *Anabaena d*uring photoautotrophic linalool production.

## *4.2 Introduction*

Terpenoids are a large and diverse family of organic compounds, with over 50,000 distinct members produced naturally by organisms in all domains of life. They exhibit an extraordinary diversity in carbon skeleton length ($C_5$-$C_{50}$), structure (cyclic, linear), and functionality in many cellular processes. In plants, terpenoids are essential in primary and secondary metabolism, with roles in membrane fluidity, light absorption, electron transport, and hormone signaling. Furthermore, their volatile nature and pleasant aroma

in floral plants make them effective signaling molecules in attracting insect pollinators for pollen dispersion.

As with pollinators, the alluring aroma of terpenoids has long attracted the attention of the commodity chemical industry, which has created a +$1-billion dollar industry in producing flavors and fragrances from plant terpenes (Wu et al. 2006).  Linalool ($C_{10}H_{18}O$) is terpene alcohol produced naturally in over 200 species of plants, and is a popular ingredient in soaps, detergents, fragrances, and food flavorings.  An increasing market demand for all-natural products (products containing non-synthetic ingredients and chemicals) has put a premium on manufacturing linalool and other essential oils from plant extracts.  However, this process is usually not economical, as it requires growing and harvesting of plant material, followed by mechanical extraction through grinding and fractional distillation.  Thus, terpenoid harvesting from plants falls short of demand, and as consumer interest in all-natural products grows, it will be imperative to develop new technology that can deliver sufficient quantities of linalool and other essential oils in an economically feasible manner.

In recent years, cyanobacteria have been successfully implemented for photosynthetically generating a wide array of natural products, including terpenoids. Their employment as a microbial "solar-converters" for producing economically-relevant products has been well established, and stems from several desirable characteristics. They exhibit higher biomass productivity and energy-conversion efficiencies than terrestrial plants, are able to be grown year-round on non-arable land, and are easily amenable to genetic engineering. Furthermore, diazotrophic ($N_2$-fixing) cyanobacteria possess the physiological mechanisms needed to convert $N_2$ gas into proteins and other essential, N-

bearing compounds. In the absence of combined nitrogen (e.g. nitrate, urea), the filamentous cyanobacterium *Anabaena* sp. PCC develops specialized, $N_2$-fixing cells called heterocysts, which are spatially organized between vegetative cells in regular intervals along the filament. Heterocysts convert $N_2$ gas to ammonia, which is then distributed to the filament in exchange for sugars synthesized in the vegetative cells. $N_2$-fixation in engineered strains is an appealing attribute for upscaling in bioreactors, considering fertilizer production through the use of the Haber-Bosch process contributes up to 40% of energy inputs in microalgae biofuel systems (Clarens et al. 2010). Therefore, diazotrophic cyanobacteria are extremely attractive cellular chassis for commodity chemical production, as long as the proper engineering efforts can be implemented.

In this study, we demonstrate a series of metabolic engineering strategies that endowed the hetrocystous, $N_2$-fixing cyanobacterium *Anabaena* sp. PCC 7120 (hereafter *Anabaena*) with the ability to synthesize linalool, while utilizing $CO_2$ and $N_2$ gas as the sole carbon and nitrogen inputs, respectively. A *Picea abies* (-)-linalool synthase (LinS) was heterogously expressed in *Anabaena* through a self-replicating expression vector. The engineered cyanobacterium was able to draw substrates from the native 2-C-methyl-D-erythritol 4 phosphate (MEP) pathway (Figure 22) and convert endogenous geranylpyrophosphate (GPP) into linalool, which evaporated into the culture headspace. To increase metabolite flux to linalool, we also introduced and overexpressed three rate-limiting enzymes (DXS, IDI, GPPS) from the MEP pathway to promote substrate flow to linalool. Our experimental findings show that our engineered *Anabaena* strains exhibit continuous, photosynthetic linalool production when these engineering strategies are

employed.  Furthermore, $N_2$ gas serves as an adequate nitrogen source for linalool production in engineered *Anabaena*, with similar productivities compared to media with combined nitrogen.  To our knowledge, these findings represent the first demonstration of linalool production using $N_2$- gas as the sole nitrogen source.
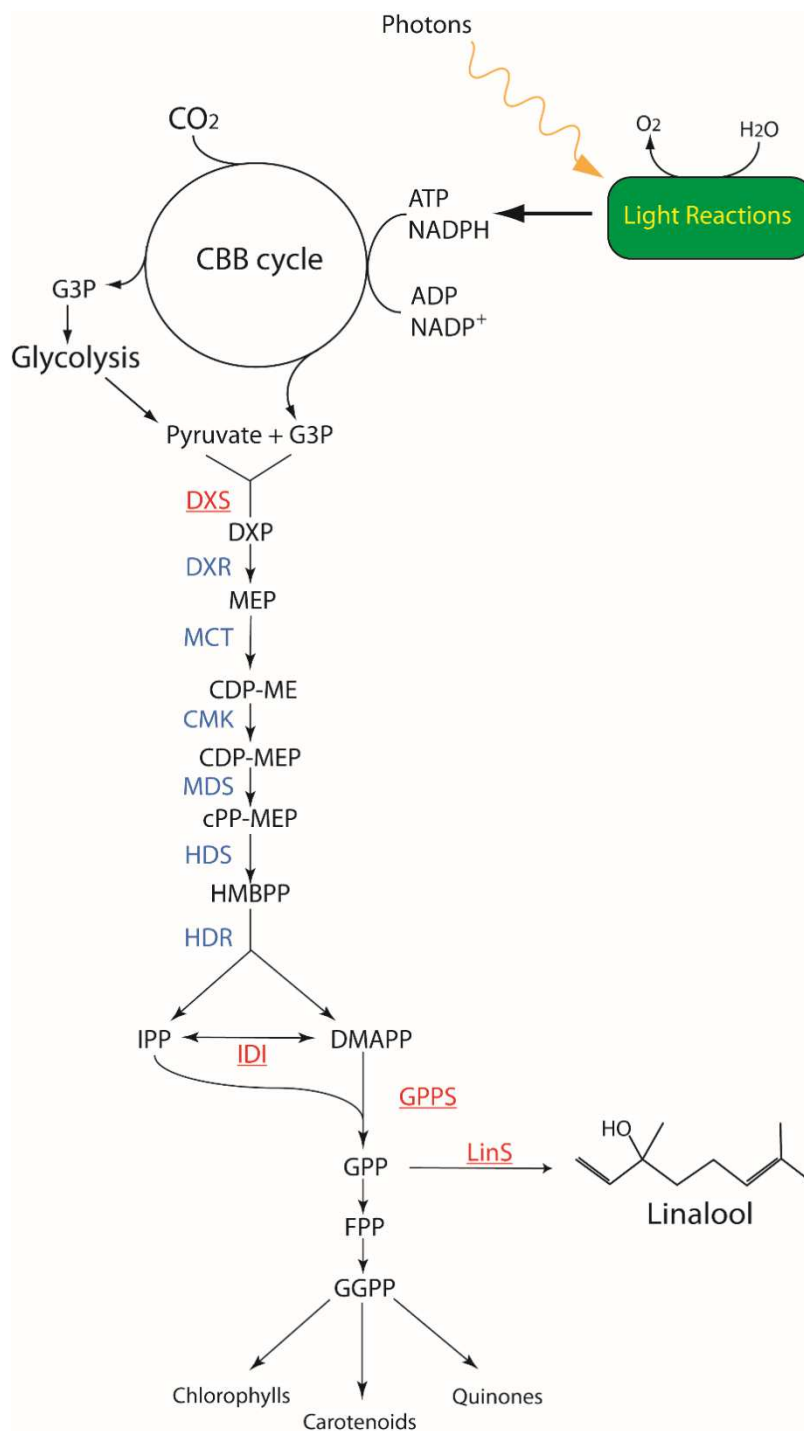
**Figure 22.** The MEP pathway in cyanobacteria. Pathway enzymes in the pathway are highlighted in blue, metabolites highlighted in black. Enzymes targeted for over-expression in this study are highlighted in red and underlined. *Metabolites:* G3P, glyceraldehyde-3-phosphate; DXP, 1-deoxy-D-xylulose 5-phosphate; MEP, methylerythritol-4-phosphate; CDP-ME, diphosphocytidylyl methylerythritol; CDP-MEP, CDP-methlerythritol-2-phosphate; cPP-ME, methylerythritol-2,4-cyclodiphosphate; HMBPP, hydroxymethylbutenyl diphosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl-diphosphate; FPP, farnesyl-diphosphate; GGPP, geranylgeranyl-diphosphate. *Enzymes*: DXS, DXP synthase; DXR, DXP reductoisomerase; MCT, CDP-ME synthase; CMK, CDP-ME kinase; MDS, ME-cPP synthase, HDS, HMPBB synthase; HDR, HMAPP reductase, IDI, IPP:DMAPP isomerase; GPPS, GPP synthase; *Cycles*: CBB cycle, Calvin-Benson-Bassham cycle

### 4.3 Materials & Methods

*Strains, growth conditions, and absorbance measurements*

*Anabaena* strains were grown photoautotrophically at 30°C in either BG11 (containing 17.6 mM $NaNO_3$) or $BG11_0$ (no combined nitrogen) media under constant illumination at 50 µmol $m^{-2}$ $s^{-1}$. Cultures were transitioned to nitrogen deprivation by centrifuging at 4000×g and washing the cell pellets 3× in $BG11_0$ media to remove all traces of combined nitrogen. Cultures were constantly agitated at 120 rpm in an Innova 44 lighted incubator (New Brunswick Scientific). Transgenic *Anabaena* strains were grown in the presence of 100 µg $mL^{-1}$ neomycin (Nm). *E. coli* used for cloning and conjugal transfer were grown in Luria-Bertani (LB) broth and shaken at 37°C, while rotating at 250 rpm. Kanamycin (50 µg $mL^{-1}$), ampicillin (100 µg $mL^{-1}$), and chloramphenicol (25 µg $mL^{-1}$) antibiotics were used to select for strains harboring engineered plasmids. Cyanobacterial growth was determined by transferring 0.15 ml aliquots of culture into a Costar 96-well plate and reading the optical density at 700nm, using a Biotek Synergy 2 plate reader.

*Plasmid Construction*

A previously characterized (-)-linalool synthase (PaTPS-Lin) from *Picea abies* was gifted from Jörg Bohlmann from University of British Columbia, Vancouver (Martin et al. 2004). The PaTPS-Lin cDNA lacking an N-terminal transit sequence was excised from pET100/D-TOPO using AatII-BglII restriction enzymes, and the 2323-bp gene fragment was ligated into AatII-BglII digested pZR807, a cyanobacterial expression vector designed in our laboratory (for construction details, see Table 6).

**Table 6.** Plasmids used for this study

| Name | Relevant Characteristics | Construction | Reference |
|---|---|---|---|
| pRL443 | Amp$^r$, Tc$^r$; Conjugal plasmid, derivative of RP4 | | (Elhai et al. 1997) |
| pRL623 | Cm$^r$; Helper plasmid | | (Elhai et al. 1997) |
| pFPN | Kan$^r$, P$_{psbA1}$ | | (Chaurasia et al. 2008; Elhai 1993) |
| pNIR | Kan$^r$, P$_{nir}$ | | (Desplancq et al. 2005) |
| pZR703 | Kan$^r$, modified pNIR; | Annealed oligonucleotides ZR64/65 containing a TTS ligated to XhoI-NotI cut pNIR | This study |
| pZR749 | Kan$^r$, modified pFPN | Annealed oligonucleotides ZR66/67 containing BglII-RBS-MCS-BamHI ligated to BamHI cut pFPN | This study |
| pZR769 | Kan$^r$, dual P$_{nir}$-P$_{psbA1}$ promoter vector | MfeI-F$_1$-P$_{psbA1}$-MCS-BamHI fragment from pZR749 ligated to EcoRI-BamHI cut pZR703 | This study |
| pZR780 | Kan$^r$, modified pZR769 | pZR769 subjected to SDM with primers ZR116 and ZR117 to delete SmaI-XhoI sites and add SalI site | This study |
| pZR804 | Kan$^r$, deleted a 300bp F1 fragment in pZR780 | pZR780 was subjected to SDM with primers ZR135 and ZR136 to delete a 300-bp F$_1$ fragment | This study |
| pZR807 | Kan$^r$, deleted a BglII site in pZR804 | pZR804 subjected to SDM with primers ZR137 and ZR138 to delete a BglII site at bp 4856 | This study |
| pTPS-Lin | PaTPS-Lin (*lins*) | Linalool synthase cDNA from *Picea abies* in pET100/D-TOPO vector | (Martin et al. 2004) |
| pZR808 (pLinS) | Kan$^r$; P$_{nir}$-P$_{psbA1}$, H$_6$-*lins* | BglII-H$_6$-LinS-AatII from PTPS-LinS ligated to BglII-AatII cut pZR807 | This study |
| pZR1462 | Kan$^r$, Nm$^r$; P$_{nir}$-P$_{psbA1}$, *dxs-ippHp-gpps* | BamHI-NotI fragment excised from pZR1546, subcloned into BamHI-NotI digested pZR1461 | (Halfmann et al. 2014b) |
| pZR1464 (pLinSDXP) | Kan$^r$, Nm$^r$; P$_{nir}$-P$_{psbA1}$, *linS-dxs-ippHp-gpps* | Digested pZR1462 with XhoI-NotI to excise 3872-bp *dxs-ippHp-gpps*, subcloned into SalI-NotI pLinS | This study |

Acronyms: Amp$^r$, ampicillin resistance; Tc$^r$- tetracycline resistance; Cm$^r$- chloramphenicol resistance; Kan$^r$- kanamycin resistance; P$_{psbA1}$, psbA1 promoter from *Anabaena*; P$_{nir}$, nitrate reductase promoter from *A.*7120; TTS, transcriptional termination site; RBS, ribosomal binding site; MCS, multiple cloning site; SDM, site-directed mutagenesis; F$_1$-flanking region from *Anabaena*; H$_6$, Histidine ×6 epitope tag.

**Table 7.** Primers and used in this study

| Primers | Sequence (5' to 3') |
|---------|---------------------|
| ZR64 | TCGAGAATGCCTCCGATTTCTAATCGGAGGCATTGC |
| ZR65 | GGCCGCAATGCCTCCGATTAGAAATCGGAGGCATTC |
| ZR66 | GATCTAAGGAGATATACATATGAGATCTCCCGGGCCCGGTACCCTGCAGGCATGCCTCGAGGACGTCCCTAGGGCTAGCG |
| ZR67 | GATCCGCTAGCCCTAGGGACGTCCTCGAGGCATGCCTGCAGGGTACCGGGCCCGGGAGATCTCATATGTATATCTCCTTA |
| ZR116 | CGGATCCCCCGGTCGAC GAGAATGCCTCC |
| ZR117 | GGAGGCATTCTCGTCGACCGGGGGATCCG |
| ZR135 | ACTTTATGAGAACGCACTGCAGGGATTCCCA |
| ZR136 | TGGGAATCCCTGCAGTGCGTTCTCATAAAG |
| ZR137 | GGGGATCAAGATATGATCAAGAGAC |
| ZR138 | GTCTCTTGATCATATCTTGATCCCC |

**Table 8.** Strains used in this study

| Strains | Relevant characteristics | Source |
|---------|--------------------------|--------|
| One Shot® Top10 *E.coli* | Cloning host | Invitrogen |
| NEB 10-beta *E.coli* | Cloning host | New England Biolabs |
| HB101 *E.coli* | For conjugal transfer of cargo plasmids into *Anabaena* sp. PCC 7120 | This study |
| *Anabaena* sp. PCC 7120 | Wild-type *Anabaena* sp. PCC 7120 | This study |
| LinS *Anabaena* | *Anabaena* sp. PCC 7120 carrying plasmid pLinS | This study |
| LinSDXP *Anabaena* | *Anabaena* sp. PCC 7120 carrying plasmid pLinSDXP | This study |

The resulting construct was designated as plasmid no. 808 (pZR808) and named pLinS. A three-gene DXP operon (*dxs-idi-gpps*) was inserted downstream of the linalool synthase gene in pLinS by excising the 3827-bp fragment from pZR1462 with XhoI-NotI and ligating it to SalI-NotI digested pLinS. The resulting plasmid was designated as plasmid no. 1464 (pZR1464) and named pLinSDXP. See Table 6 for details.

*Insertion of cargo plasmids into Anabaena*

The transfer of shuttle plasmids into *Anabaena* were performed via bacterial conjugation using a tri-parental *E. coli* mating system (Elhai et al. 1997) with the following modifications. *E. coli* HB101 bearing helper (pRL623) and conjugal (pRL443) plasmids were mated with *E. coli* NEB 10β harboring pLinS or pLinSDXP, respectively, and selected on LB plates containing triple antibiotic selection for the three plasmids. Selected colonies were grown overnight in 2 mL of LB containing appropriate antibiotics, subcultured by adding 200 µL of overnight culture to 2 mL of fresh LB containing appropriate antibiotics, and grown for additional 3 hours. Cells were pelleted through centrifugation, washed with 3x 1 mL of LB to remove antibiotics, and resuspended with 200 µL LB and subjected for mating with *Anabaena*. A 10 mL culture of WT *Anabaena* was grown to early exponential stage (O.D.$_{700}$ = 0.3) and then sonicated (Branson 1510 water bath sonicator) for 60-120 seconds to break filaments into 2-4 cell lengths, which were confirmed under microscopy. Cells were harvested by centrifugation at 5,000 x g for 10 minutes, resuspended in 200 µL of fresh BG11, and then mixed with above *E. coli* harboring triple plasmids for conjugal transfer of plasmids from *E. coli* into *Anabaena*. The above cell mixture was placed under lighted conditions at 25°C for 30 minutes, micro-pipetted on to a Immobolin-NC transfer membrane (Millipore, HATF08550) filter placed on a BG-11 + 5% LB plate, and grown under white light for 3 days at 30°C. The filter was then transferred to a BG11 plate containing Nm100 µg/mL to select for positive exconjugants. Individual exconjugant colonies were further purified by restreaking onto fresh BG11 plates containing 100 µg/mL Nm and routine colony PCR was used for verification of positive exconjugants.

*Protein Extraction and Western Blotting*

Protein was extracted from both WT and LinS *Anabaena* cells to determine

expression of LinS in the host strains.  Ten mL cultures were grown in BG11 with

appropriate antibiotics until mid-exponential phase (O.D.$_{700}$ = 0.5), and the cells were

pelleted and washed 3x with sucrose buffer (50mM Tris-HCl, 40mM EDTA, 0.75M

sucrose) resuspended with 50 µL of lysis buffer (0.5 mg/mL lysozyme, 10 ug/mL DNAse

I, and 10 ug/mL RNase A.), and incubated at 37°C for 15 minutes.  After incubation, 50

µL of 2x SDS loading buffer (50mM Tris-Hcl, pH 6.8, 4% SDS, 20% [vol/vol] glycerol,

200 mM dithiothreitol, 0.03% bromophenol blue) was added, and the mixture was boiled

for 5 minutes, and then centrifuged at 13,000 rpm for 10 minutes at 4°C.  5 µL protein

extracts were loaded and separated by 12% SDS-PAGE at 200V for 30 minutes. The total

proteins were transferred onto a PVFD membrane, stained by Comassie blue R-250, and

His$_6$-LinS were detected by Western blot using anti-His antibodies (Qiagen) at a dilution

of 1:5000.

*Linalool identification and quantification from Anabaena*

Wild-type (WT) and linalool-producing *Anabaena* strains were grown in 250-mL

Erlenmeyer flasks with either 100mL BG11 or BG11$_0$, and bubbled with filtered air at a

rate of 100 mL min$^{-1}$.  A small glass column filled with 50 mg of Supelpak 2SV resin

(Supelco, Sigma-Aldrich) was attached to the exhaust port of each flask to capture

linalool and other volatiles from the culture headspace.  Volatiles were washed from the

2SV resin using the following procedure:   1 mL of pentane containing 5 µg mL$^{-1}$

tetracosane as an internal standard (IS) was added to each resin sample vial (sealed) and

vortexed at 3000 rpm for 1 minute, incubated at room temperature for 10 minutes,

followed by a brief 10-sec vortex (3000 rpm). Volatiles extracted from the pentane wash
were then analyzed using gas chromoatography-mass spectrometry (Agilent
7890A/5975C). 1 µL injected samples were separated using a HP-5MS column (35m x
250µm x 0.25µm), with $H_2$ as the carrier gas. The oven temperature was initially held at
60°C for 2 minutes, and increased 20°C $min^{-1}$ until a temperature 300°C was reached.
Analytes were identified by comparing mass spectra from compounds in the
NIST/EPA/NIH Mass spectral library, version 2.0.

A set of linalool standards (Sigma Aldrich) at 0.1, 0.5, 1, 5, 10 $µg·mL^{-1}$ in pentane
were used to create a standard curve to quantify linalool captured by each resin sample.
To test 2SV resin absorption capacity, 5, 50, 500, and 1000 µg of linalool standard was
spiked into 100 mL cultures of wild-type *Anabaena* in exponential phase (O.D.$_{700}$ of 0.5)
and incubated in experimental growth conditions for up to 68 hours, allowing linalool to
volatilize and be collected by the 2SV resin. Resin samples were then washed as
described above, and analyzed by GC-MS to measure linalool recovery. All
measurements were performed in triplicate.

## *4.4 Results*

### *Expression of a linalool synthase in Anabaena for GPP conversion to linalool*

Linalool synthesis in cyanobacteria requires redirecting endogenous GPP from the
MEP pathway into linalool through the action of the enzyme LinS. Since *Anabaena* does
not possess any genes with sequence similarities to *linS* (revealed by a BLAST search; E-
value >0.001), we expressed a foreign plant LinS in *Anabaena* through a self-replicating
pDU1 expression vector, in order to divert MEP pathway substrates to linalool. The
truncated *linS* gene from *Picea abies* (Norway spruce) lacking a native chloroplast transit

peptide was cloned into the multiple-cloning site of the pDU1-based cyanobacterial expression vector pZR804, and fused with a His$_6$ epitope tag on the N-terminus of the protein. To maintain constitutive expression His$_6$-LinS mRNA, a dual *nir-psbA1* promoter was engineered upstream of the His$_6$-LinS gene fusion. The engineered plasmid (named pLinS) was then transferred into wild-type *Anabaena* sp. PCC 7120 using conjugation, and resulting transformants were renamed LinS-*Anabaena*. Western blotting was performed to confirm the expression level of His$_6$-LinS when LinS-*Anabaena* when grown in either BG11 (+N) or BG11$_0$ (-N). The predicted molecular weight of the His$_6$-LinS protein was 69 kD, and anti-His antibodies showed a strong and specific cross-reaction of a protein of this size when raised against LinS-*Anabaena* protein extracts (Figure 23). Strong expression of His$_6$-LinS was confirmed at 0 and 24 hours growth in +N conditions (Figure 23, lane 1 & 2). His$_6$-LinS expression was also confirmed in –N conditions at 0 hours (lane 3), although it decreased after 24 hours (lane 4). As expected, His$_6$-LinS was absent in WT *Anabaena* in either nitrogen condition (lanes 5-8).
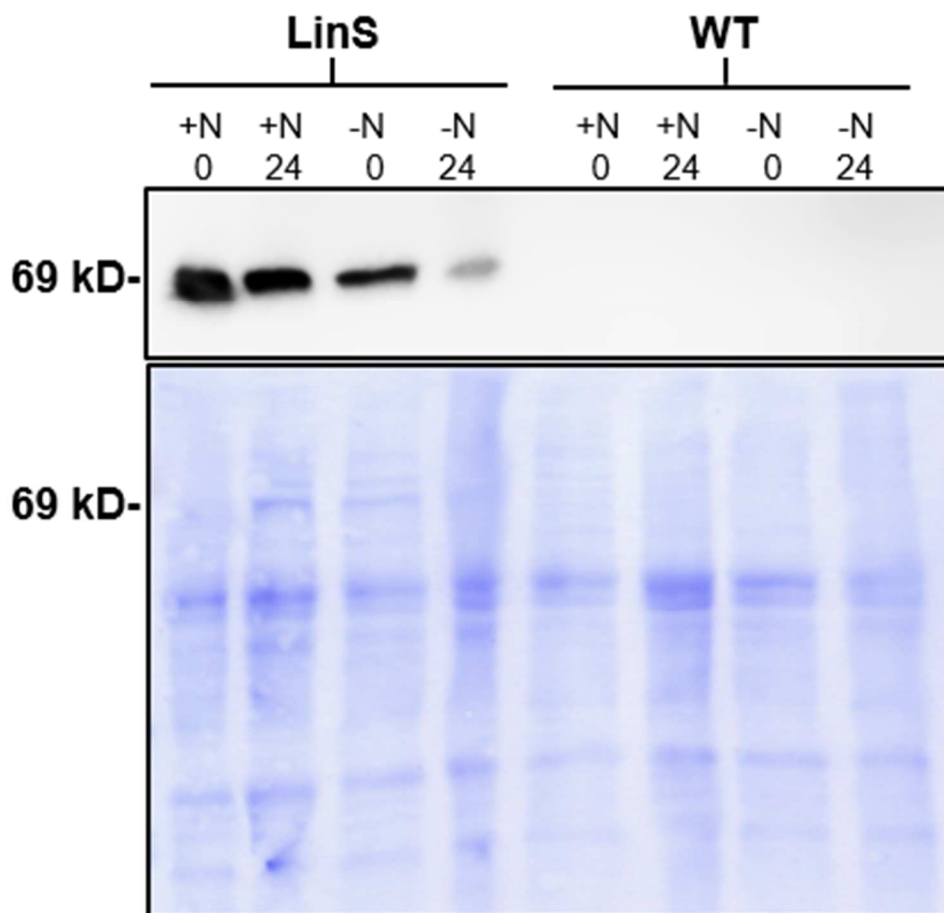
**Figure 23.** Immunodetection of LinS protein in transformed filaments of *Anabaena*. (Upper box) Western blot analysis with anti-His antibodies reveal the 69kD LinS protein in transformed cell lines of LinS-*Anabaena* at different time points (0 and 24 hours) in both nitrogen-replete (+N) and nitrogen-deplete (-N) media. (Lower box) Coomassie-stained SDS-PAGE lanes containing total cell protein to display equal protein loading.

*Determination of linalool emitted from engineered LinS-Anabaena*

Filtering the flask headspace air through styrene-divinylbenzene resin (2SV) has

proven to be a successful method for capturing terpenoids emitted from engineered

cyanobacteria cultures. Linalool, a $C_{10}$-alcohol, readily phase-separates and evaporates

when placed in aqueous solutions, making 2SV-headspace collection an ideal strategy for

linalool detection and quantitation. Our experiments with linalool-spiked cultures of

wild-type *Anabaena* demonstrate that 50 mg of 2SV is capable of collecting up to 500 µg

linalool during 20 hours in our experimental growth conditions with nearly 100%
extraction efficiency (data not shown). We used this strategy to collect volatiles emitted
from wild-type (WT) and LinS-*Anabaena*, and identified the collected eluents through
pentane-washing and subsequent gas-chromatography mass spectrometry (GC-MS)
analysis. To distinguish naturally-emitted volatiles from intrinsic 2SV eluents, we also
included a pentane wash of untreated 2SV resin (not shown). As shown in Figure 24A &
B, photosynthetically-active cultures of WT and LinS-*Anabaena* naturally release a
mixture of acids, alkanes, and alcohols, with the most prominent being 2-ethylhexanol
($C_8H_{18}O$; 4.6 min), heptadecane ($C_{19}H_{40}$; 9.1 min), and hexadecanoic acid ($C_{17}H_{34}O_2$;
10.2 min) However, linalool was absent from the volatile portfolio of WT *Anabaena*,
shown by comparing the retention time of a linalool standard (Figure 24C). Analysis of
volatiles collected from the headspace of LinS-*Anabaena* revealed a new peak with
identical elution time (5.2 min) to the linalool standard. Additionally, the eluted
compound contained a molecular weight (154) and major mass fragments (41, 71, and 93
m/z) identical to the standard. This data supported that the *Picea abies* LinS endowed
*Anabeana* the ability to synthesize linalool from substrates from the endogenous MEP
pathway.
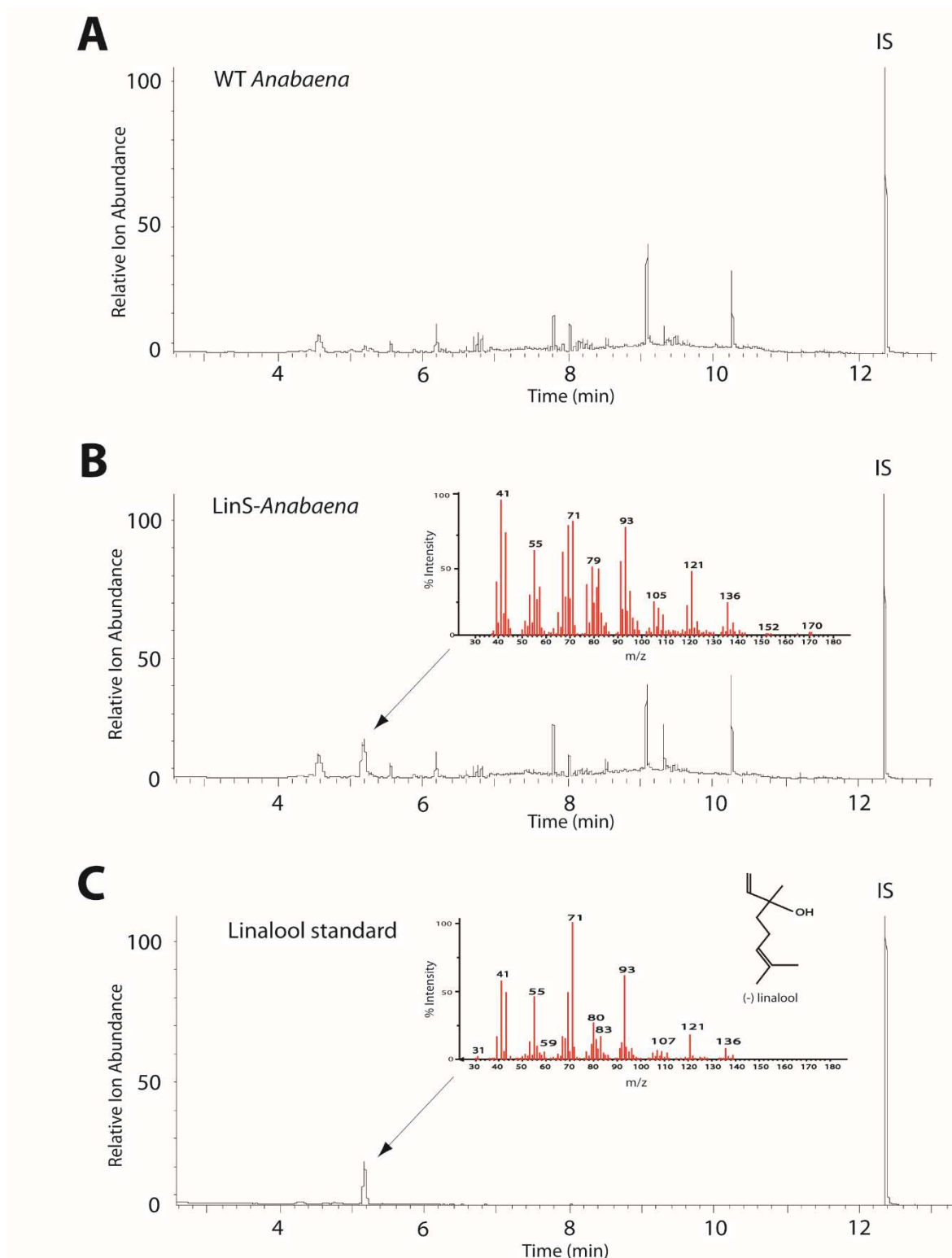
**Figure 24.** GC-MS profiles of headspace volatiles collected from cyanobacteria. A peak at the retention time of 5.2 minutes (arrowed) found in LinS-*Anabaena* matches the linalool standard (C), but is absent in the WT (A). Mass spectra (inserts) of these peaks display the expected fragmentation pattern for linalool. Five µg ml$^{-1}$ tertracosane serves as an internal standard (IS).

*Growth dynamics and linalool productivity in LinS-Anabaena*

Next, we measured the dynamics of linalool production from LinS-*Anabaena* under continuous light and bubbling when grown in the presence of two different nitrogen sources:  sodium nitrate (supplemented in BG11 media at a concentration of 17.6 mM) or atmospheric $N_2$ (BG11$_0$, no combined nitrogen).  Growth rates between LinS-*Anabaena* and WT *Anabaena* were similar in each nitrogen condition, suggesting that the expression of LinS did not impose any detrimental effects to the rate of cell division (data not shown).   In the presence of combined nitrogen, LinS-*Anabaena* displayed positive growth during a 14-day trial, with an average growth rate of $0.17 \pm 0.01$ O.D. ml$^{-1}$ d$^{-1}$ (1.7 $\pm 0.23$ µg chl$^{-1}$ml$^{-1}$d$^{-1}$).  During the growth trial, linalool productivity was highest during the first four days (~25 µg linalool L$^{-1}$O.D.$^{-1}$d$^{-1}$; ~2.5 µg linalool mg chl$^{-1}$d$^{-1}$), but decreased to ~5 µg L$^{-1}$O.D.$^{-1}$d$^{-1}$ (~0.3 µg linalool mg chlorophyll$^{-1}$day$^{-1}$) during the last six days of the trial (Figure 25B).  The observed reduction in linalool productivity over time is arguably attributed to a loss of photosynthetically-active as cell density increases, a phenomenon known as the self-shading effect.   When grown with $N_2$ as a sole nitrogen source, LinS-*Anabaena* also displayed positive, yet slower growth ($0.14 \pm 0.01$ O.D. ml$^{-1}$ day$^{-1}$; 1.3 $\pm 0.07$ µg chl$^{-1}$ml$^{-1}$d$^{-1}$) compared to conditions with combined nitrogen, with decreasing productivity over the course of the growth period.  Interestingly, we did not observe a significant difference in linalool productivity at each time point when either NaNO$_3$ or $N_2$ was administered as a nitrogen source (measured through a student's t-test). In fact, linalool productivity was shown to be slightly higher during diazotrophic growth on days 9 through 12.

*Increasing MEP pathway flux by overexpressing rate-limiting enzymes*

A key challenge in increasing terpenoid titers in engineered cyanobacteria is controlling the partitioning of carbon in the MEP pathway into the desired terpenoid product. Previously, we have demonstrated the application of a synthetic 'DXP operon' (named from 1-deoxy-D-xylulose 5-phosphate, the first substrate in the MEP pathway) to increased carbon flux through the MEP pathway in *Anabaena* (Halfmann et al. 2014b). The operon consists of the enzymes DXS (DXP synthase, encoded by the gene *dxs* from *E. coli*), IDI (an IPP:DMAPP isomerase, encoded by the gene *ipphp1* from *Haematococcus pluvialis*) and GPPS (a GPP synthase, encoded by the gene *Rv0989c* from *Mycoplasma tuberculosis*). These enzymes control rate-limiting steps in the MEP pathway, and which co-expression with a specific terpenoid synthase resulted in an overall improvement in limonene production in cyanobacteria (Halfmann et al. 2014b) and amorpha-4,11-diene in *E.coli* (Martin et al. 2003). To express the operon with LinS, we linked the multi-cistronic *dxs-idi-gpps* downstream of LinS in pLinS to create a new plasmid, named pLinSDXP, and expressed the construct in *Anabaena* to create a new strain, LinSDXP-*Anabaena*.
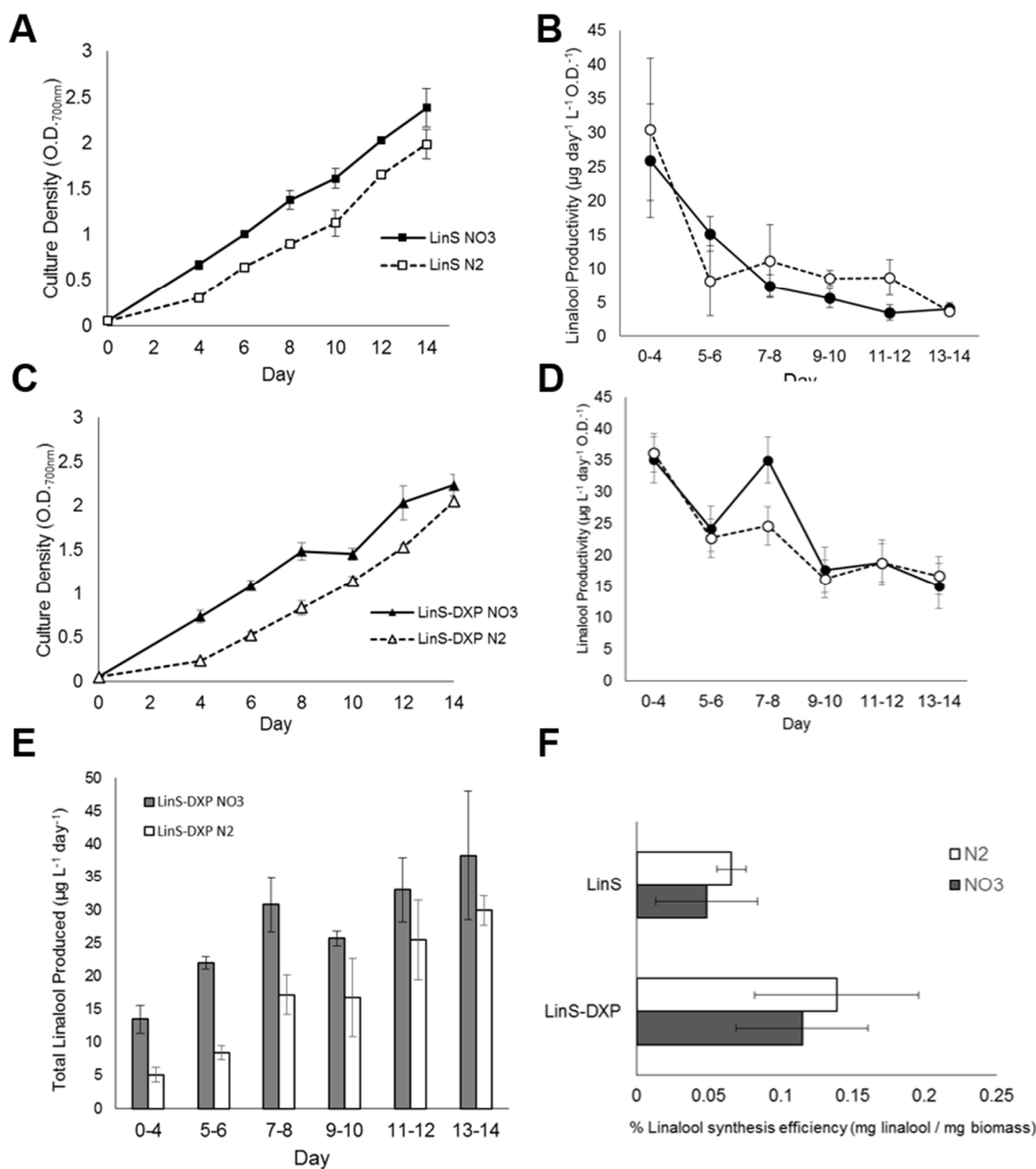
**Figure 25.** Growth kinetics and linalool productivity in engineered strains of *Anabaena*. Growth was determined by measuring optical density (700nm) for (A) LinS-*Anabaena* and (C) LinSDXP-*Anabaena* over a 14 day time span in the presence of either combined nitrogen ($NO_3$; bold lines) or without ($N_2$; dashed lines). Linalool productivity for (B) LinS-*Anabaena* and (D) LinSDXP-*Anabaena* was determined by dividing the total amount of linalool collected per liter per day by the average O.D.700nm of cultures during the shown time period (x-axis). (E) Total linalool produced from LinSDXP-Anabaena per day during each time period (x-axis) in the presence of combined nitrogen ($NO_3$; grey bars) or in $N_2$-fixing conditions ($N_2$; white bars). (F) Linalool synthesis efficiency from LinS and LinSDXP-*Anabaena* with either $NO_3$ or $N_2$ as a nitrogen source (grey bars and white bars, respectively) was determined by dividing the total amount of linalool collected by total DCW at the end of the experiment (shown as percentage). Error is presented as SEM from three independent experiments.

Culture growth and linalool productivity results of DXP operon/LinS co-expression are shown in Figure 25C and 25D, respectively.  Growth rates between LinS and LinSDXP-*Anabaena* were similar in each growth condition, indicating that the added expression of the three gene operon did not impose a heavy metabolic burden on the cells.  Similar to LinS-*Anabaena,* LinSDXP-*Anabaena* displayed a slower growth rate in $N_2$-fixing conditions (0.16 ±0.01 O.D. $ml^{-1}$ $d^{-1}$; 1.23 ±0.04 µg $chl^{-1}ml^{-1}d^{-1}$) than with nitrate (0.14 ±0.004 O.D. $ml^{-1}$ $d^{-1}$; 1.61 ±0.08 µg $chl^{-1}ml^{-1}d^{-1}$), but exhibited higher linalool production than the *Anabaena* strain expressing LinS solely.  During days 0-4, LinSDXP-*Anabaena* produced linalool at a rate of ~35 µg linalool $L^{-1}day^{-1}O.D.^{-1}$ in each nitrogen condition, declining to ~16 µg linalool $L^{-1}day^{-1}O.D.^{-1}$ on days 13-14. Volumetric linalool production between nitrogen sources are shown in Figure 25E. Linalool production per L culture is initially 2.6-fold higher in cultures utilizing nitrate versus $N_2$-fixing cultures, as nitrate is a more accessible nitrogen source than $N_2$ and promotes higher biomass growth.  Over time, actively $N_2$-fixing cultures reach cell densities that are similar to nitrate-grown cells (days 13-14), thus developing similar linalool production rates per volume culture.

While the DXP operon increased yield, linalool synthesis efficiency in both genetically-engineered strains (Figure 25F) was found to be extremely low.  Linalool made up ~ 0.05% of the total photosynthetically-produced biomass when expressing LinS, and was increased to ~0.13% of total biomass when co-expressing the DXP operon. Linalool synthesis efficiency was not dependent on the nitrogen source used ($N_2$ or $NO_3$).

## *4.5 Discussion & Conclusions*

Generating renewable fuels and materials from metabolically-engineered microbes has gained interest in the last decade, as a strategy to divest from fossil fuel usage. Utilizing diazotrophic cyanobacteria cells as solar-driven fuel generators is an extremely attractive strategy for "green chemical" production, considering they can use atmospheric gases ($CO_2$ and $N_2$) as building blocks for engineered chemical synthesis. Incorporating $N_2$-fixation into an industrial linalool production system can lower operational costs, and make 4th-generation linalool production economically viable.

Here, we show that the filamentous, $N_2$-fixing cyanobacterium *Anabaena* confers the ability synthesize the $C_{10}$-alcohol linalool through metabolic engineering. A plant LinS was constitutively expressed in *Anabaena* on a multi-copy expression plasmid, with expression driven by a dual psbA1-pNir promoter. Expression was confirmed using Western blotting, and the foreign plant protein was shown to successfully convert endogenously-produced GPP into linalool, which was emitted from the cyanobacteria cells into the flask headspace. Since *Anabaena* filaments are capable of simultaneous $CO_2$ and $N_2$ fixation, we tested linalool productivity in the engineered strain LinS-*Anabaena* when either supplemented nitrate or atmospheric $N_2$ was employed as a nitrogen source. Under each nitrogen regiment, the engineered filaments maintained similar rates of linalool production, which was measured at ~25 μg linalool $L^{-1}O.D.^{-1}d^{-1}$ at the productivity maximum. These results suggest that atmospheric $N_2$ is a suitable substitute for nitrate during photosynthetic linalool production by genetically engineered *Anabaena*.

Increasing pathway flux to linalool by overexpressing rate-limiting enzymes in the MEP pathway improved linalool productivity in *Anabaena*. Adding additional copies of *dxs*, *idi*, and *gpps* genes from microbial sources into the *Anabaena* genome resulted in an maximum productivity increase of ~35 µg linalool $L^{-1}O.D.^{-1}d^{-1}$, during the midpoint of the experiment. Overall, we measured a 2.8-fold increase in total linalool produced when expressing the DXP operon with nitrate, and a 2.1-fold increase when $N_2$ gas was used as a sole nitrogen source. The productivity results indicate that over-expressing specific enzymes in the MEP pathway, in conjunction with the native cyanobacterial MEP enzymes, increases metabolic flux towards essential terpenoid precursors. Additional over-expression of LinS draws precursor metabolites into linalool synthesis, effectively creating a new carbon sink in the cyanobacterium.

In summary, our research promotes *Anabaena* as a model organism for renewable chemical production. Future engineering research may one day rationalize the employment of $N_2$-fixing cyanobacteria for 4th-generation biofuel production at the industrial scale.

# Chapter 5: Discovering the effects of glycogen-deficiency on the nitrogen-stress response in *Anabaena* sp. PCC 7120

## *5.1 Abstract*

The nitrogen stress response (NSR) is an adaptive mechanism that prepares the cell for nitrogen starvation. In the filamentous cyanobacterium *Anabaena* sp. PCC 7120 (*Anabaena*), the NSR activates a major morphological transformation, converting a proportion of vegetative cells into specialized nitrogen-fixing cells called heterocysts. The internal mechanisms that control the NSR in *Anabaena* are deeply rooted to the carbon/nitrogen status of the cell. Glycogen, a multi-branched glucose polymer, serves as a major carbon sink in *Anabaena*. The interplay between carbon and nitrogen acquisition is an important area of study in cyanobacteria, especially for developing future engineering strategies to divert carbon, nitrogen and energy into biofuels. To understand the interplay between glycogen metabolism and the NSR in $N_2$-fixing cyanobacteria, we have created a glycogen-deficient mutant of *Anabaena* through an inactivation of glucose-1-phosphate adenylyltransferase (*agp*), the enzyme responsible for the first committed step in glycogen synthesis. Our analysis reveals that during nitrogen deprivation, the Δagp mutant displays several physiological changes compared to the wild-type, including 1) decreased growth rate and photosynthetic activity, 2) inability to degrade light-harvesting phycobilisomes, and 3) altered heterocyst frequency and reduced nitrogenase activity during diazotrophic growth. Additionally, the expression of a farnesene synthase in the glycogen-deficient background revealed no significant difference in farnesene production in the cyanobacterium. These findings reveal that disrupting glycogen synthesis alters many characteristics associated with

nitrogen stress in *Anabaena*, and provides insight on future strategies for engineering cyanobacteria for biofuel production.

### *5.2 Introduction*

Cyanobacteria constitute a group of bacteria that are arguably the most successful organisms on our planet, given their unique ability to synthesize polysaccharides and lipids using oxygenic photosynthesis. Several species of cyanobacterial also exhibit the ability to fix dinitrogen ($N_2$) into ammonia, a feat reserved to only to a few unique bacteria groups. *Anabaena* sp. PCC 7120 (here in after as *Anabaena*) is a filamentous cyanobacterium that has long been studied as a model for heterocyst differentiation and nitrogen fixation (Golden and Yoon 2003; Haselkorn 1992; Muro-Pastor and Hess 2012; Wolk 2000). During nitrogen-depleted conditions, *Anabaena* undergoes a physiological transformation, converting some of its vegetative cells into specialized $N_2$-fixing cells called heterocysts, which develop across the filament approximately every 15-20 cells (Bryant 1994). Vegetative cells continue photosynthesis during this transformation, and transport sugars (possibly sucrose) to the heterocysts in exchange for the amino acids needed to support metabolism (Cumino et al. 2007; Curatti et al. 2002). The potential biotechnological applications for *Anabaena* as a platform for renewable biofuels and commodity chemicals has also been researched (Halfmann et al. 2014a; Halfmann et al. 2014b; Johnson et al. 2016). *Anabaena* and other $N_2$-fixing cyanobacteria are attractive cellular hosts for fourth-generation biofuels, since nitrogenous fertilizer makes up a significant energy input to large-scale biofuel production systems (Clarens et al. 2010).

The internal mechanisms that signal the nitrogen-stress response (NSR), which ultimately leads to $N_2$-fixation in heterocystous cyanobacteria, are well studied. Low

ammonia in the glutamine synthetase (GS) and glutamate synthase (GOGAT) system results in increased levels of 2-oxoglutarate (2-OG) in the cell, triggering the induction of the NSR regulator NtcA (Espinosa et al. 2006). NtcA is a global-nitrogen transcription factor, responsible for the upregulation of many genes involved in the NSR, including nrrA, a nitrogen-regulatory transcription factor needed for the full induction of HetR, the master regulator of heterocyst formation in *Anabaena* (Ehira and Ohmori 2006a; Ehira and Ohmori 2006b). Other genes up-regulated during the NSR include *nblA*, (involved in the temporary breakdown of the PSI/PSII-associated antennae known as phycobilisomes), *glgP* (glycogen catabolism), and *sigE* (a group 2 σ factor of RNA polymerase which induces expression of genes involved in glycolysis and the pentose phosphate pathway).

The mechanisms governing the functional operation of the NSR in *Anabaena* have been illuminated through the discovery and characterization of genes upregulated when filaments are introduced to nitrogen deprivation (Buikema and Haselkorn 1991; Ernst et al. 1992; Rice et al. 1982). Interestingly, disruptions in key genes in other metabolic pathways, including carbon assimilation and storage, have also been found to have a profound effect on the NSR in non-diazotrophic cyanobacteria. Glycogen, a multi-branched glucose polymer, serves as the principle carbon sink in cyanobacteria. Glucose-1-phosphate adenylyltransferase (AGPase; also known as glgC) establishes the first committed step in glycogen synthesis through the conversion of glucose-1-phosphate into ADP-glucose, which is added to the end of a growing glycogen polymer. During nitrogen stress, cyanobacteria accumulate glycogen, presumably using it as a reductant sink to mitigate the redox imbalance caused by disrupted nitrogen assimilation and continued operation of the photosynthetic electron transport chain (PETC).

Studies with *ΔglgC* mutants of non-diazotrophic species of cyanobacteria, including

*Synechococcus* sp. PCC 7942 (Hickman et al. 2013; Suzuki et al. 2010), *Synechococcus*

sp. PCC 7002 (Jackson et al. 2015) and *Synechocystis* sp. PCC 6803 are incapable of

accumulating glycogen during N-starvation, and consequentially, exhibit unique

phenotypes during these conditions.  The most visually striking is the deactivation of the

chlorotic response: an inability to degrade phycobilisomes, which maintains a blue-green

color in nitrogen-starved cultures.  Other intriguing phenotypes emerge from glycogen-

deficiency during nitrogen stress, including reduced photosynthetic capacity, organic acid

excretion in lieu of a native carbon sink, and the whole scale retainment of a thylakoid

ultrastructure in the cell.  These results point to the existence of a link between glycogen

synthesis and the NSR in cyanobacteria, in that the inability to form glycogen results in

the disruption of the cyanobacterium's ability to completely activate the NSR and

physiologically transition during nitrogen stress.  *Anabaena* offers a unique

cyanobacterial model to study this phenomenon because in contrast to non-diazotrophic

cyanobacteria, it harbors the ability to properly adapt to nitrogen stress through

heterocyst differentiation and nitrogen-fixation.  Additionally, eliminating glycogen as a

carbon sink may free more energy and metabolites into engineered chemical pathways,

including terpenoid synthesis.  Farnesene ($C_{15}H_{24}$) is a sesquiterpene synthesized through

the MEP pathway, is an attractive candidate as a renewable fuel and commodity

chemical.  Abolishing carbon storage in cyanobacteria may be a key strategy for

increasing biofuel yields in large-scale photosynthetic systems.

To understand the interplay between glycogen availability and $N_2$-fixation in

cyanobacteria, we have constructed an AGPase-deleted mutant of *Anabaena* (Δagp) to

elucidate glycogen's impact on diazotrophic growth (Figure 26). Furthermore, we tested the practicality of removing a major carbon sink in *Anabaena* to redistribute metabolites into the MEP pathway for farnesene production. These results shed light on the interwoven networks that underlie carbon and nitrogen response signals in $N_2$-fixing cyanobacteria, and may prove useful to future engineering strategies for photosynthetic biofuel production.
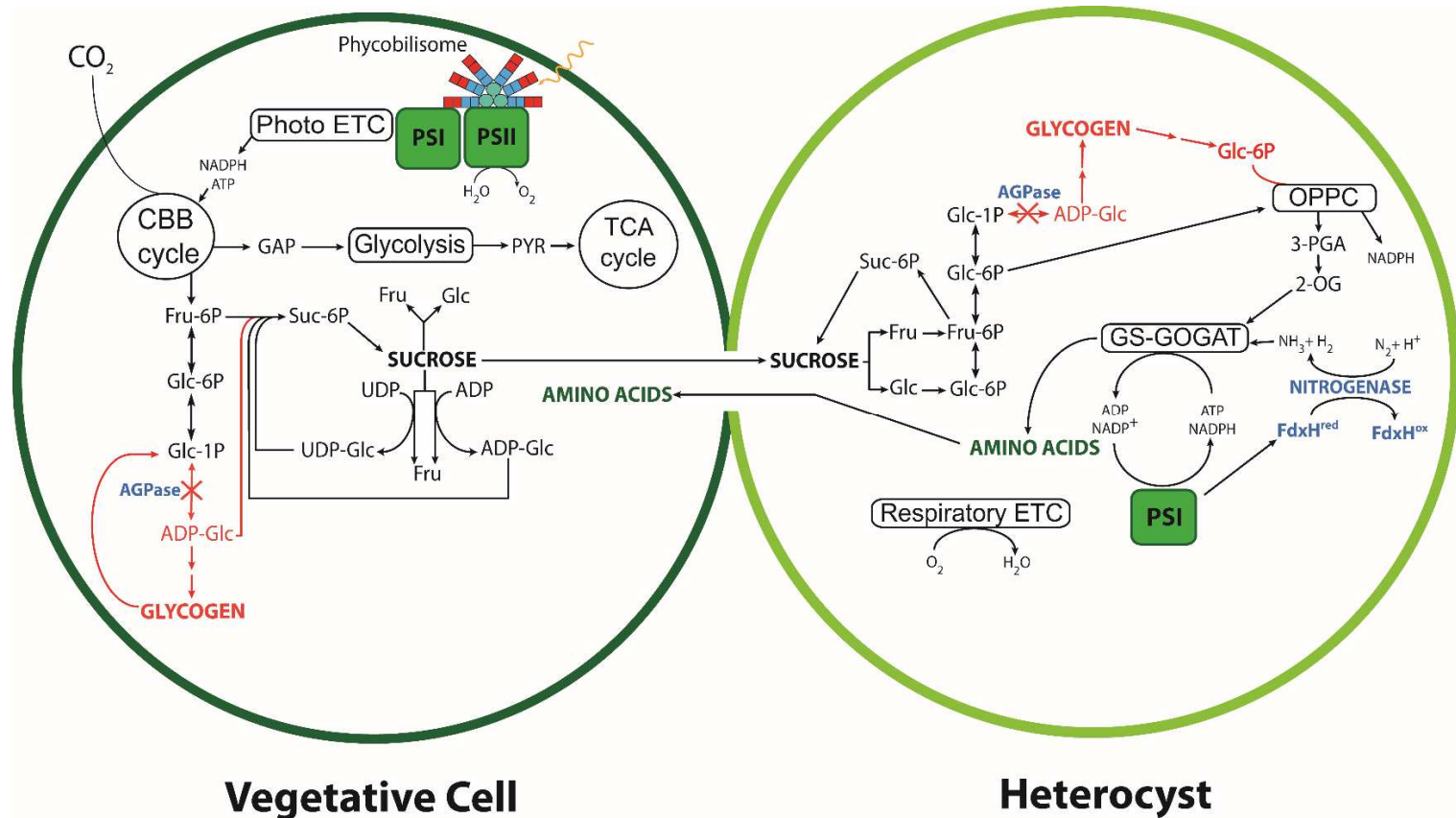
**Figure 26.** A simplistic model of carbon and nitrogen metabolism in vegetative cells and heterocyst cells of *Anabaena*. Phycobilisomes transfer energy from light to PSII and PSI in the vegetative cells to provide ATP and NADPH for carbon fixation, producing sucrose and glycogen as carbon sinks. Sucrose is transported to the vegetative cell, where its catabolism in the oxidative pentose phosphate chain (OPPC) provide reducing energy for nitrogenase. The products of nitrogen-fixation (amino acids) are transported to the vegetative cells to support filament-wide metabolism. The pathways absent in the Δagp mutant are outlined in red. Acronyms: PSI, photosystem I; PSII, photosystem II; Photo ETC, photosynthetic electron-transport chain; CBB cycle, Calvin-Benson-Bassham cycle; GAP, glyceraldehyde-3-phosphate; PYR, pyruvate; TCA cycle, tricarboxylic acid cycle; Fru, fructose; Glc, glucose; Fru-6P, fructose-6-phosphate; Glc-6P, glucose-6-phosphate; Glc-1P, glucose-1-phosphate; Suc-6P, sucrose-6-phosphate; ADP-Glc, ADP-glucose; UDP-Glc, UDP-glucose; 3-PGA, 3-phosphoglycerate; 2-OG, 2-oxoglutarate; GS-GOGAT, glutamine synthetase-2-oxoglutarate aminotransferase; FdxH, heterocyst-specific ferredoxin.

## 5.3 Materials & Methods

### Bacterial strains and growth conditions

*Anabaena* sp. strain PCC 7120 was grown in BG11 (containing 17.6 mM NaNO$_3$) or BG11$_0$ (free of combined nitrogen) containing 20 mM HEPES buffer under constant 50 $\mu$E·m$^{-2}$·s$^{-1}$ white light at 30°C in liquid media (shaken at 120 r.p.m.) or maintained on solid media supplemented with 1% agar. Light intensity was measured using a Li-Cor LI-190 Quantam sensor. Colonies were transplanted from agar plates to liquid starter cultures, with Δagp *Anabaena* media supplemented with 10 $\mu$g mL$^{-1}$ spectinomycin, grown for 5-7 days, and cells were centrifuged and washed ×3 in fresh media (with no antibiotics) before all experiments. *Escherichia coli* strains were grown in Luria-Bertani (LB) medium at 37°C, and 250 rpm. Kanamycin (Km; 50 $\mu$g·mL$^{-1}$), ampicillin (Ap; 100 $\mu$g·mL$^{-1}$), and spectinomycin (Sp; 100 $\mu$g mL$^{-1}$) were used for antibiotic selection for *E. coli* during strain construction.

### Construction of Δagp Anabaena mutant

To inactivate glycogen synthesis in *Anabaena*, an 850-bp internal fragment of *agp* (*all4645*) was amplified from *Anabaena* chromosomal DNA by standard PCR using a high-fidelity Q5 DNA polymerase (NEB) and the following oligonucleotide primers: ZR942 (forward primer) 5'-TGGATCCTCGCCTTTACCCACTAACCAAAC-3'; ZR943 (reverse primer) 5'-TGCTAGCACCGGTGGTAAGTAACGAGCGCGGGTAT-3'. The internal fragment was then cloned into a pCR®2.1-TOPO® vector (TOPO TA Cloning® kit, Invitrogen) to create pZR1607, which was transformed into Oneshot® Top10 *E.coli* using the manufacturer's protocols. Resulting colonies were verified using blue/white colony screening, and plasmid DNA was extracted from verified colonies using a

QIAprep® Spin Miniprep Kit (Qiagen). The internal fragment in pZR1607 was excised using SpeI and NsiI restriction endonucleases and ligated into SpeI/PstI double digested pRL277, an integration vector that can be mobilized in *Anabaena* (Black et al. 1993) . The resulting plasmid, named pZR1839, was transferred to *Anabaena* using conjugation as previously described (Halfmann et al. 2014a). Exconjugates were further purified by streaking on BG11 agar plates with 10 µg·mL$^{-1}$ spectinomycin. Routine PCR was used to verify single-crossover integration of pZR1839 into the *agp* locus of *Anabaena*.

*Quantification of Glycogen*

Glycogen was extracted from cell pellets using the following method: A 2-mL aliquot of culture (O.D.730nm of 0.3) was centrifuged at 4000× g for 5 min., washed ×3 in distilled H$_2$O, and stored in -80˚C for up to 72 hours before extraction. Pellets were resuspended in 50 µL of 30% KOH and incubated in a 95˚C water bath for 2 hours. 150 µL of ice-cold ethanol was then added to each KOH-cell mixture and then placed on ice for an additional 2 hours to precipitate the glycogen. Samples were centrifuged for 15 min., supernatant removed, and the glycogen pellet washed with 70% and 95% ethanol. After the final wash, the pellets were dried in a Savant ISS110 speedvac concentrator (Thermo Scientific). To enzymatically hydrolyze glycogen to glucose, each pellet was treated with 2 mg·mL$^{-1}$ amyloglucosidase (Roche Diagnostics) in 100mM sodium acetate (pH 4.5) and incubated at 50˚C for 1 hour. Glucose equivalents were measured using a glucose reagent hexokinase kit (Thermo Scientific) and a validated calibration curve. Chlorophyll extracts of 2-ml cultures from the same flasks were used for normalization.

*Optical density, chlorophyll extraction, and absorbance spectra*

Culture growth was determined by transferring 0.15 ml aliquots of culture into a 96-well plate (Costar) and reading the optical density at 730nm, using a Biotek Synergy 2 plate reader. Similarly, photosynthetically-activate radiation (PAR) scans were performed on cultures by reading absorbance in 10 nm intervals from 350 to 800 nm. Chlorophyll *a* concentration from 2-ml culture suspensions was determined by from 90% methanol extracts as previously described (Meeks and Castenholz 1971).

*Acetylene reduction assay*

Nitrogenase activity was determined by measuring the reduction of acetylene ($C_2H_2$) to ethylene ($C_2H_4$) via nitrogenase in *Anabaena* filaments with mature heterocysts. Cultures of WT and Δagp *Anabaena* in exponential phase were washed 3× in $BG11_0$ to remove combined N, and resuspended in 25 mL of $BG11_0$ ($O.D._{730}$ of 0.5) in photoautotrophic growth conditions. At 48 and 96 hours after N-stepdown, heterocyst formation was confirmed by microscopy, then 3-ml suspensions were transplanted into a 10-ml glass vial sealed with a rubber stopper and injected with 0.5 ml of acetylene. Vials were placed on a rotary shaker (120 rpm) for 5 hours under constant 50 $\mu m$ $m^{-2}s^{-1}$ white light at 30ºC in liquid BG11, after which 0.5 ml of vial headspace was manually injected (pulsed split 100:1) into an Agilent PoraBOND Q GC-MS column (Model CP7348, 25m×0.25mm×3μm). Column temperature was maintained at 32°C for 3.4 min, followed by 110 to 200°C $min^{-1}$ for 1 min to separate acetylene and ethylene peaks. Ethylene concentration was determined using a standard curve.

*PSII activity measurements*

PSII activity was measured via oxygen evolution in 1 mL culture samples (adjusted to

a cell density of 10 µg chlorophyll mL$^{-1}$) in BG11 or BG11$_0$ with 10 mM NaHCO$_3$ using a Clark-type electrode and DW2 Oxygen Electrode Chamber with O$_2$ View Oxygen Monitoring software (Hansatech).  Light was introduced to samples using an LS2/H Tungsten-halogen 100 W light source and adjusted with neutral density filters.  Light intensity and sample temperature was monitored using a Quantitherm light-temperature meter during experimentation.  All measurements were performed in triplicate.

*Heterocyst patterning and microscopy*

WT and Δagp cultures grown in 25 ml BG11 to a density of 0.5 (O.D.$_{730nm}$) were pelleted and washed 3× in BG11$_0$, and transplanted into 25 ml BG11$_0$ to induce heterocyst formation.  Filaments were photographed using an Olympus BX53 microscope at 48 and 96 hours after nitrogen-stepdown.  Intervals were counted as the number of vegetative cells between two heterocysts, and fifty intervals from each biological replicate (n=3) were pooled to establish distribution and percent heterocyst frequency (heterocyst number / total cell number).  A minimum of 500 total cells were counted for each biological replicate.

*Expression of farnesene synthase and farenesene collection from cyanobacteria*

The plasmid pFaS, containing a farnesene synthase gene codon-optimized for *Anabaena* (GenBank: KJ847279.1) was conjugally transferred into WT and Δagp *Anabaena*, as previously described in Halfmann 2014a.  PCR-verified colonies were grown in BG11 media supplemented with neomycin (100 µg/ml) to an O.D.$_{730nm}$ of 0.8, before being split into separate flasks containing 100 ml BG11 and BG11$_0$ (cultures in BG11$_0$ were pelleted and washed 3× to remove NaNO3).  All cultures were bubbled with

1% $CO_2$ at 100 ml min$^{-1}$ and farnesene was collected using styrene-divinylbenzene resin (2SV) as previously described in Halfmann et al. 2014a.

## 5.4 Results

*Engineering and verifying glycogen-deficiency in Δagp Anabaena*

*Anabaena* possesses a single AGPase homologue (*agp*; *all4645*), located on the chromosome upstream of *all4644* and a *dnaJ*, and downstream of *all4646*. We inactivated *agp* by inserting the integration vector pZR1839, containing an internal fragment of *agp* into the cyanobacterium, which integrated into the *agp* open frame in the chromosome. Colonies that grew on agar plates containing spectinomycin (10 µg/mL) were selected for PCR to verify the purity of *agp*-deleted colonies (Δagp *Anabaena*), using primers upstream and downstream of the *agp* ORF (Figure 27).
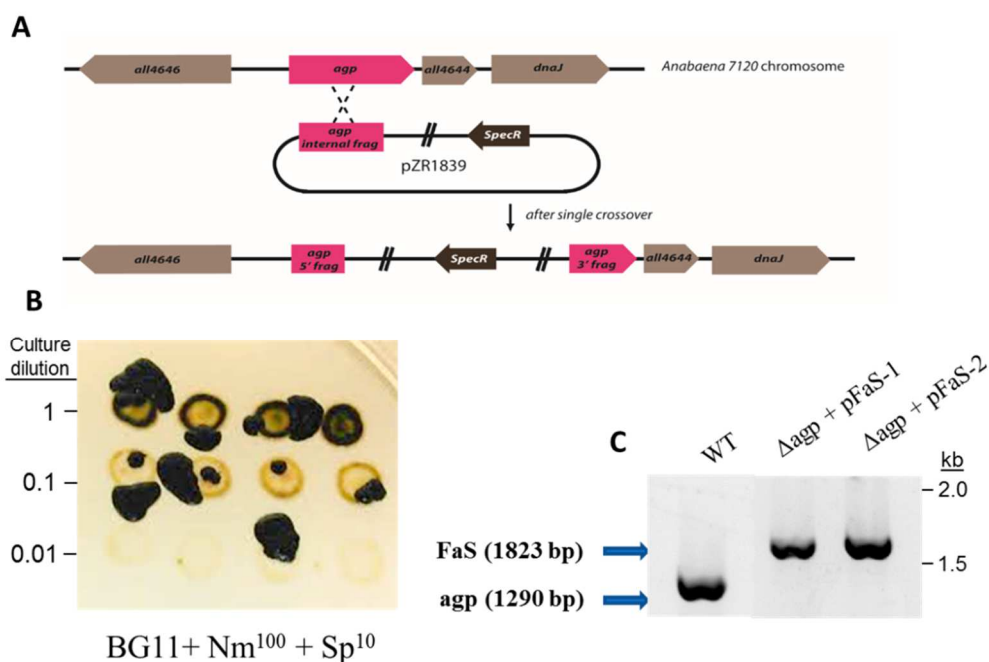


**Figure 27.** Engineering of Δagp *Anabaena.* (A) Schematic of single-crossover recombination on an agp internal fragment in pZR1839 into the *Anabaena* genome. (B) *agp*-deleted *Anabaena* colonies grown on spectinomycin antibiotic approximately 3 weeks after conjugation. (C) PCR verification of Δagp *Anabaena* with pFaS (lane 2 and 3) compared to WT *Anabaena* (lane 1).

To verify the abolishment of glycogen accumulation in the mutant, batch cultures of wild-type (WT) and Δagp *Anabaena* were grown in both BG11 (+N) and $BG11_0$ (-N) under continuous illumination, and glycogen amounts were analyzed at 0, 24, 48, and 72 hours.  Glycogen levels in WT *Anabaena* remained fairly constant when grown in +N, ranging from 3-5 µg glucose equivalents per µg chlorophyll *a* (Figure 28).  When placed under nitrogen-starvation, glycogen levels increased~4-fold within the first 48 hours (up to 20 µg glucose equivalents per µg chl), before declining from 48-72 hour, coinciding with an increase in $N_2$-fixation from mature heterocysts (Figure 28).  The Δagp mutant, however, showed trace to no glycogen in either growth condition (trace amounts detected are likely attributed to other intracellular sugars at basal levels).  These results indicate that the chromosomal disruption of *agp* prevents the synthesis of glycogen in *Anabaena*.
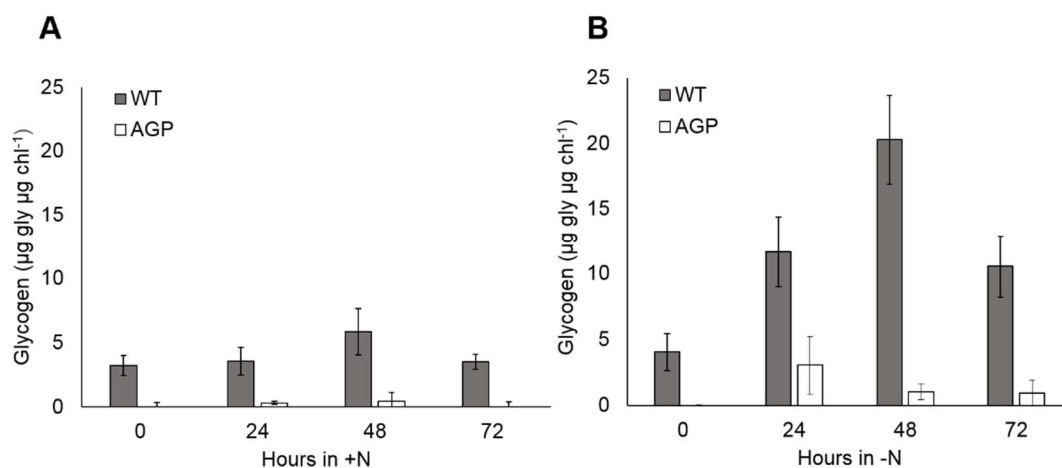


**Figure 28.**  Glycogen content over time in WT and Δagp *Anabaena* filaments when grown in nitrogen-replete BG11 media (A; +N) or BG110 (B; -N).  Glycogen is reported in glucose equivalents from a hexokinase assay.  Error bars are in standard deviation (SD) from three replicated experiments.

*Growth dynamics and phycobilisome retainment*

Next, WT and Δagp *Anabaena* cultures were grown photoautotrophically in +N and –N conditions to reveal differences in growth dynamics and PBS retainment. PAR

absorbance spectra of WT and Δagp *Anabaena* cultures showed a prominent ~630nm

peak after 48 hours in +N conditions (Figure 29A, arrowed), indicating the retainment of

PBSs that give the cultures a brilliant, blue-green color (Figure 29C).  The strains also

displayed similar growth rates of approximately 0.05 O.D.$_{730nm}$ day$^{-1}$ from 24 to 72 hours

in +N media (Figure 30A).   When both strains were introduced to nitrogen-starvation,

the WT filaments exhibited lagged yet positive growth during the first 48 hours of N-

stress (compared to WT +N, Figure 30A).  Phycobilisome absorbance rapidly decreased

during the first 48 hours (Figure 29B), coinciding with a blue-green to yellow-green color

shift in the WT, a characteristic of nitrogen chlorosis (Figure 29D).  However,

phycobilisome absorbance was regained from 48 to 72 hours –N (Figure 29B),

accompanied by a yellow-green to blue-green color shift (data not shown), an increase in

culture density (Figure 29A) and chlorophyll *a* concentration (data not shown).
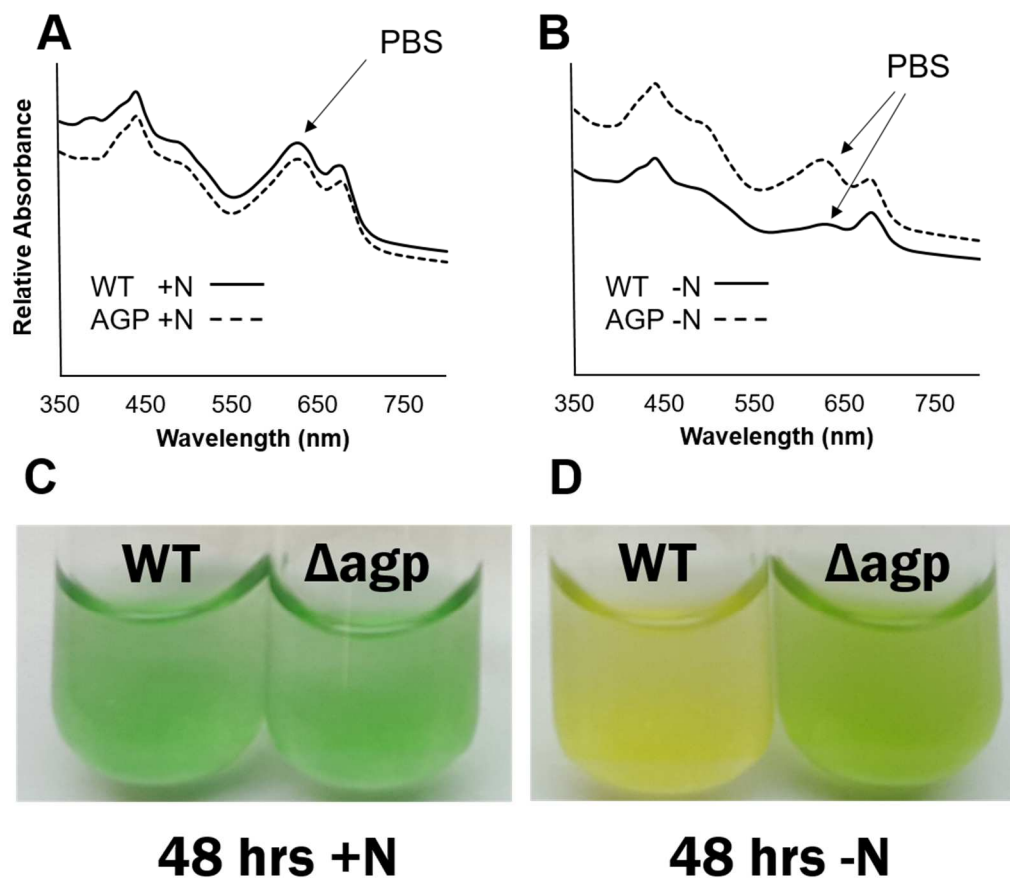
**Figure 29.** The non-bleaching phenotype of Δagp *Anabaena* during N-starvation. Phycobilisomes (PBS), which absorb at 620 nm in whole-cell PAR scans (A), are abundant the in thylakoid membranes of WT and Δagp *Anabaena* when supplied with a combined N-source, giving photoautotrophically-grown cultures a blue-green color (C). Upon N-deprivation, the WT degrades its phycobilisomes and undergoes chlorosis, turning filaments yellow-green (B and D). However, glycogen-deficiency disrupts phycobilisome degradation in the Δagp mutant, resulting in the filaments retaining a green color.

In contrast, Δagp *Anabaena* cell division was completely arrested during the first 72

hours of N-stress, verified by both optical density (Figure 30A) and chlorophyll

measurements (not shown). PAR scans throughout this time period indicated, in contrast

to the WT, no loss in phycobilisome absorbance (Figure 29B), or blue-green color

(Figure 29D), similar to previously characterized Δ*glgC* cyanobacteria mutants. Although

it conferred a lagged growth initially, the Δagp mutant was able to re-establish cell

division after 72 hours in –N conditions, until it reached similar densities to the WT after

around 200 hours post-inoculation (data not shown).  Re-introducing stationary-phase

Δagp cultures to fresh $BG11_0$ resulted again in an extended lag phase (not shown),

suggesting that the adaptation to N-stress during prolonged incubation is not due to a

secondary suppressor mutation .
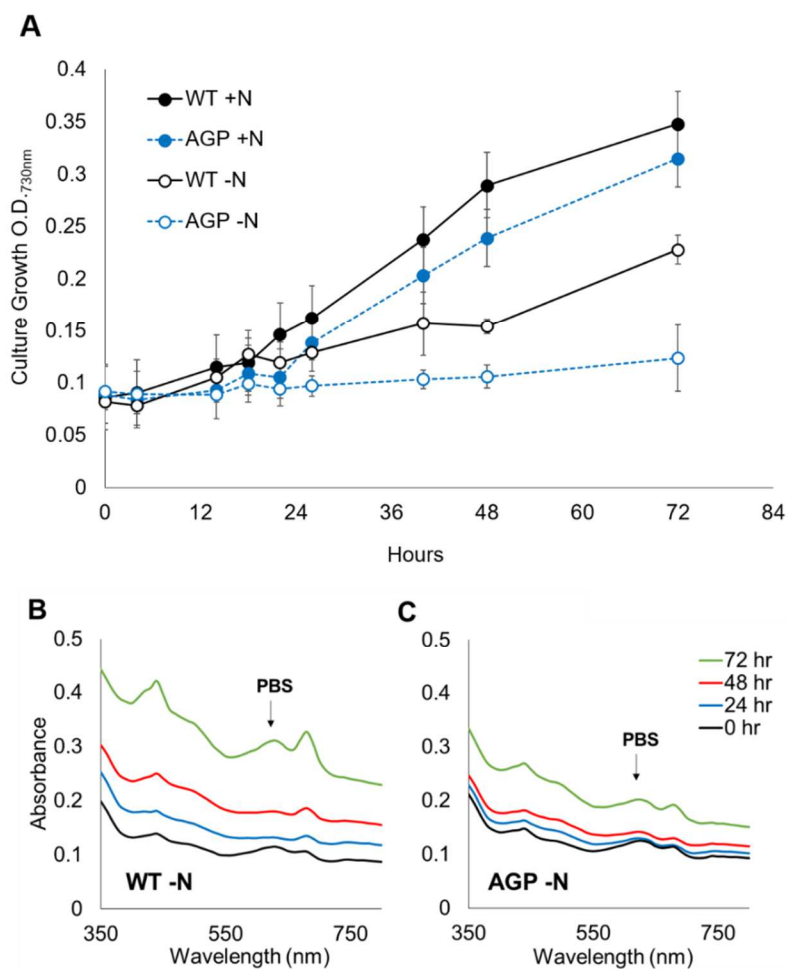


**Figure 30.**  Culture Growth of WT and Δagp *Anabaena* with or without a combined N source (+N and –N, respectively).  During $N_2$-fixing conditions (-N), PAR scans from 350 to 800 nm on WT filament reveal the degradation and reassembly of phycobilisomes over 72 hours (WT), while the Δagp mutant retains its phycobilisomes during this time period (C).

*Glycogen-deficiency effects heterocyst patterning and nitrogenase activity*

When introduced to an environment deprived of a combined nitrogen source, *Anabaena* develops heterocysts at regularly-spaced intervals along the filament, thus enabling $N_2$ fixation. To determine whether glycogen deficiency impacts heterocyst development and functionality, we analyzed morphological differences in heterocyst patterning and $N_2$ fixation between the WT and Δagp *Anabaena* as it transitioned through cell differentiation. The number of vegetative cells between heterocysts were scored at 48 and 96 hours after N-stepdown, as well as an acetylene reduction assay to measure nitrogenase activity. In the presence of combined nitrogen, both WT and Δagp filaments were long and composed solely of vegetative cells. We noted that at around 18-21 hours after nitrogen-stepdown, the Δagp mutant tended to have more filament breakage while the WT filaments remained intact. Over the course of $N_2$-fixing conditions, both strains formed heterocysts with no observable abnormalities at approximately the same time (21-24 hours), yet displayed a difference in the interval length by 48 hours, with an average interval of 26 and 9.2 vegetative cells between two heterocysts in WT and Δagp *Anabaena*, respectively (Figure 31A and 31B). This resulted in an approximate 3-fold higher heterocyst frequency in Δagp *Anabaena* (14.6±0.25% heterocysts) compared to the WT (5.5±1.6% heterocysts) (Figure 31C). The findings were supported by a higher $N_2$-fixing activity in the glycogen-deficient mutant at 48 hours (Figure 31D), while found to be barely detectable in the WT.

From 48 to 96 hours after N-stepdown, the WT interval length decreased on from 26 to 14.5 vegetative cells, as new heterocysts continued to form midway in established intervals. This was accompanied by a significant increase in $N_2$-fixation, up from 0.15 to

12.8 nM ethylene $hr^{-1}O.D^{-1}$ over this time period (Figure 31D). Although the interval

length between heterocysts did not change significantly in Δagp *Anabaena* from 48 to 96

hours, we measured a 50% decline in nitrogenase activity at 96 hours, approximately

one-fifth the rate of ethylene produced from the WT. Further acetylene reduction assays

overtime confirmed a prolonged reduction in nitrogenase activity in the glycogen-

deficient mutant at 214 hours after N-stepdown (Δagp, 8.7 nM total ethylene per $10^8$ cells;

WT, 49.9 nM ethylene $10^8$ cells; data not shown).

To test if the lowered $N_2$-fixing activity at 96 hours was due to lowered expression of

nitrogenase, we expressed a *nifB-gfp* reporter fusion driven by the *nifB* promoter in both

WT and Δagp to gauge nitrogenase expression through fluorescence microscopy. GFP

fluorescence was detected from WT and Δagp heterocysts at 48 and 96 hours after

nitrogen-stepdown, with no discernable difference in intensity (not shown).
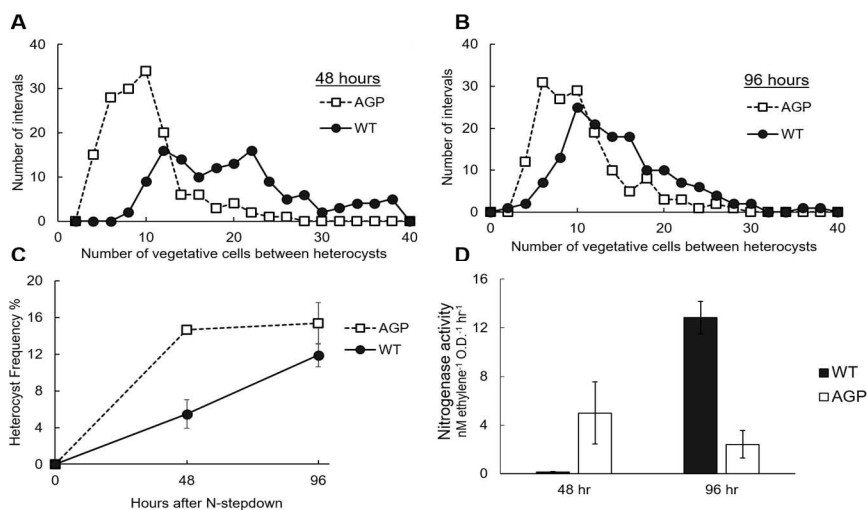


**Figure 31**: Analysis of $N_2$-fixing characteristics of WT and AGP *Anabaena*. Distribution of heterocyst intervals (number of vegetative cells between two heterocysts at 48 hours (A) and 96 hours (B) after N-stepdown. Fifty random intervals were counted from each biological replicate (n=3) of AGP and WT *Anabaena* at 48 and 96 hours. Heterocyst frequency (C) was calculated using the same 50 intervals from each biological replicate at 48 and 96 hours after N-stepdown. $N_2$-fixation was measured using an acetylene reduction assay at 48 and 96 hours (D).

*Increased photoinhibition and reduced PSII activity during diazotrophic growth*

Photosynthetic electron transport rates in WT and Δagp were analyzed by measuring PSII-mediated $O_2$ production during nitrate assimilation or diazotrophic growth.  When combined nitrogen is available, both WT and Δagp *Anabaena* displayed similar $O_2$-evolving rates when subjected to low light intensity (50 $\mu E$ $m^{-2}$ $s^{-1}$), and the rates increased 2-fold under high light conditions (500 $\mu E$ $m^{-2}$ $s^{-1}$) (Figure 31A).  When the strains are in diazotrophic growth conditions, similar PSII activities are observed in filaments actively fixing $N_2$ in 50 $\mu E$ $m^{-2}$ $s^{-1}$ in low light.   However, whereas the $N_2$-fixing WT exhibited increased $O_2$-evolving rates that were only slightly lower that in nitrogen-replete conditions, PSII activity in the glycogen-deficient mutant remained unchanged.

Changes in photosynthetic activity over time in saturated light conditions is presented in Figure 32B.  Initially, oxygen production rates dropped from approximately 3 to 1.2 $\mu mol$ O2 $min^{-1}$ $\mu g$ $chl^{-1}$ in both strains from 14 to 24 hours, likely due to the cells metabolically adjusting to re-supplied nutrients.  Thereafter, oxygen production increased in both strains from 24 to 48 hours, although PSII activity was still lower in the glycogen deficient mutant (2.2 $\mu mol$ $O_2$ $min^{-1}$ $\mu g$ $chl^{-1}$) than the WT (3.6 $\mu mol$ $O_2$ $min^{-1}$ $\mu g$ $chl^{-1}$).
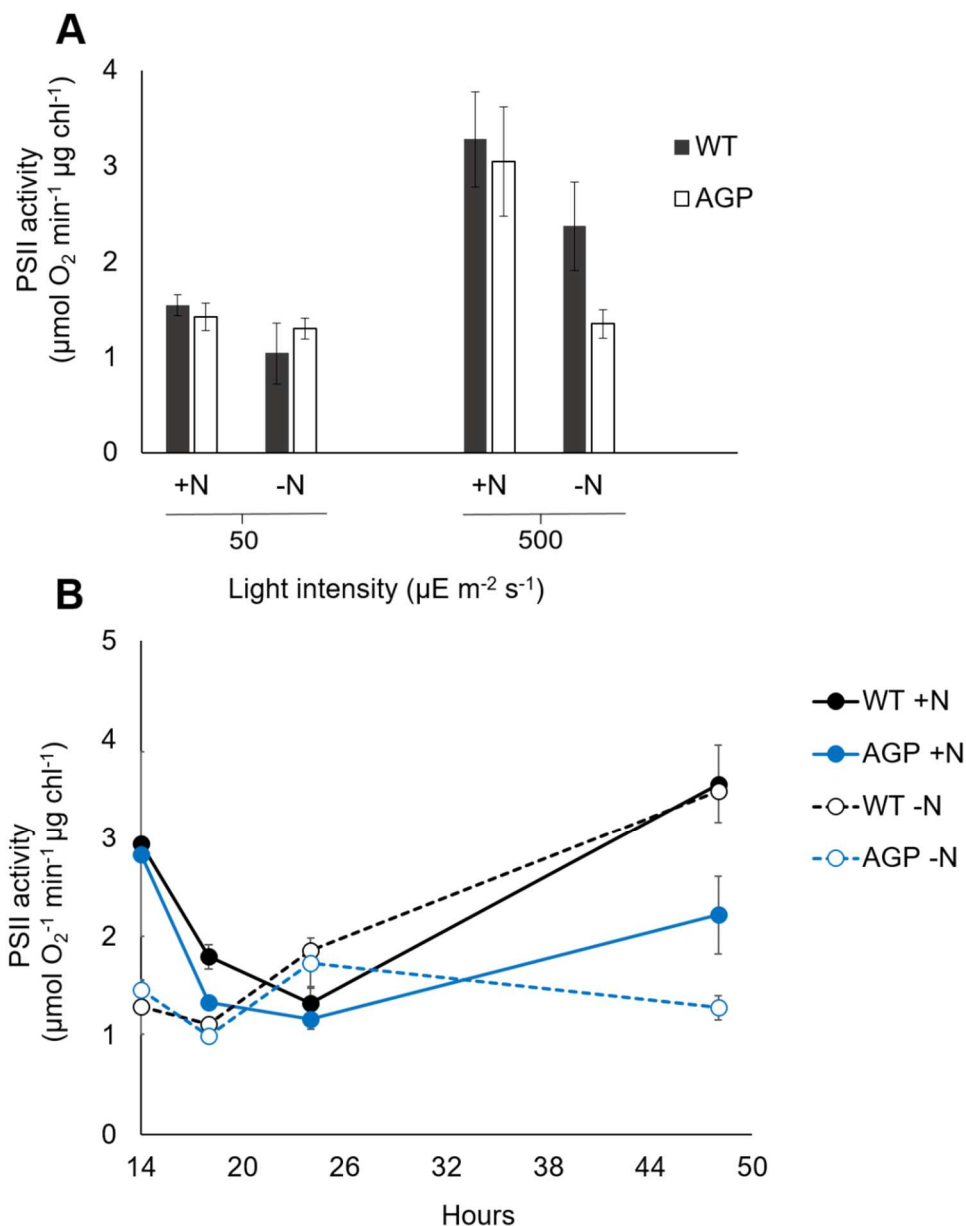
**Figure 32.** PSII-mediated oxygen evolution of WT and Δagp in nitrogen-replete (+N) or $N_2$-fixing (-N) conditions. A) PSII activity of filaments in different irradiance levels (50 and 500 μE $m^{-2}s^{-1}$) after 96 hours in the indicated nitrogen source. B) PSII activity of filaments under saturated light conditions (2000 μE $m^{-2}s^{-1}$) at the indicated time after initially inoculating in each nitrogen source. Error is standard deviation of three independent colonies.

Under $N_2$-fixing conditions, both WT and Δagp *Anabaena* followed a strikingly

similar pattern in PSII activity during the onset of N-starvation, although lower compared

to +N conditions (Figure 32B). Similar to the change in PSII activity in media with

combined N, the WT showed an increase from approximately 2 to 3.5 µmol $O_2$ min$^{-1}$ µg chl$^{-1}$ during 24 to 48 hours after N-stepdown, as heterocysts begin to fix $N_2$ and supply the necessary proteins needed to maintain photosynthetic activity. Unlike the WT, the glycogen-deficient mutant was unable to illicit an increase in PSII activity during this time period, as the rate of oxygen evolution decreased from 1.7 to 1.3 µmol $O_2$ min$^{-1}$ µg chl$^{-1}$. The slow rate of $O_2$-evolution in the $N_2$-fixing mutant was shown to persist up to 170 hours after nitrogen-stepdown (data not shown).

*Farnesene production in Δagp Anabaena*

Glycogen is the principle carbohydrate reserve in cyanobacteria, constituting up to 60% of cell weight under nitrogen-deprived conditions (Allen 1984). Therefore, glycogen represents a major competitor for energetic metabolites, and a possible target for redirecting metabolite flux towards useful, engineered biofuel pathways, such farnesene synthesis through the MEP pathway (Halfmann et al. 2014a). Previous studies of ΔglgC mutants of non-diazotrophic cyanobacteria reported that under nitrogen starvation, the cells excreted a suite of organic acids (2-OG, pyruvate, succinate, and acetate), possibly as an alternative carbon sink in the absence of glycogen (Carrieri et al. 2012; Davies et al. 2014; Gründel et al. 2012). Presumably, introducing terpenoid synthase could re-direct these metabolites into the MEP pathway for terpenoid production, especially since pyruvate is an essential precursor for the pathway. We tested the effects of farnesene production in the glycogen-null background by expressing a codon-optimized farnesene synthase (FaS) in Δagp *Anabaena* under both nitrogen-replete and deplete conditions. Flasks containing 100-mL culture were bubbled with 1% $CO_2$, and farnesene was collected by passing the flask headspace through styrene-

divinylbenzene resin (2SV), which was analyzed through GC-MS for eight consecutive days (Figure 33).

Farnesene productivity was highest for all strains in both nitrogen-replete and depleted conditions during the first two days of growth, and slowly diminished throughout the production experiment.  Overall, the absence of glycogen did not significantly impact  farnesene production in the engineered *Anabaena* either media condition, although we did note that productivity was lower for Δagp in combined-nitrogen (~40 μg farnesene $L^{-1}$O.D.$^{-1}$) than agp-intact in combined-nitrogen (~73 μg farnesene $L^{-1}$O.D.) during the first two days of collection.  This difference was abated over the course of the experiment, as WT productivity decreased to Δagp levels on days 7 to 8.   In media deprived of combined nitrogen, farnesene production was similar between the WT and glycogen-deficient mutant, although the WT productivity was slightly higher than Δagp from days 5 to 8.  In conclusion, the absence of glycogen did not increase farnesene production in *Anabaena* during diazotrophic or non-diazotrophic growth.
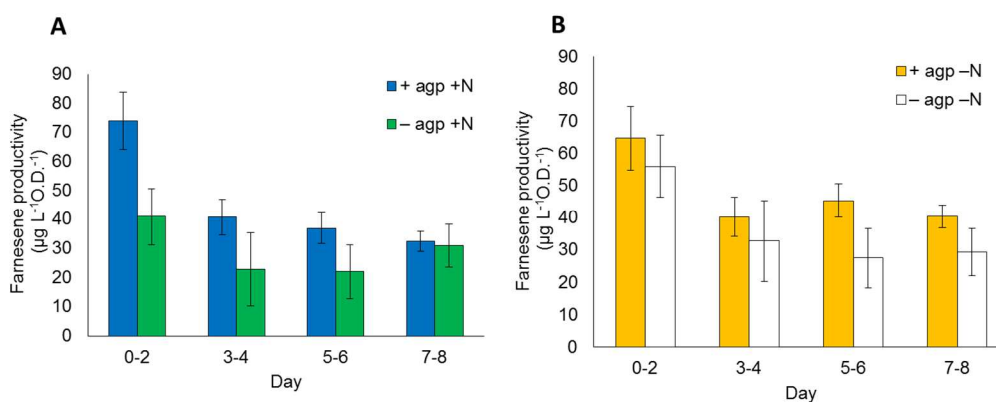


**Figure 33.**   Farnesene productivity of WT and Δagp harboring FaS in nitrogen-replete (A; +N) and $N_2$-fixing conditions (B; -N).   X-axis indicates the productivity during each time span over the course of the growth experiment.  Error is standard deviation of three independent experiments.

## 4.5 Discussion & Conclusions

The NSR system is a dynamic response mechanism, able to detect changes imbalances in the cellular carbon/nitrogen ratio through 2-OG and react by the induction of genes to maintain cell survival during nitrogen stress. In cyanobacteria, these genes are involved in a variety of unique responses, including the degradation of phycobilisome, reduced PSII activity, glycogen accumulation, and in *Anabaena* 7120, inducing cellular differentiation through the formation of heterocysts and $N_2$-fixation. The engineered deletion of glycogen synthesis in non-diazotrophic cyanobacteria has been studied extensively, revealing aspects of carbon-nitrogen metabolism in these microorganisms. Since *Anabaena* engages in photosynthesis as well as heterocyst differentiation and $N_2$-fixation, we found it intriguing to discover if these cellular activities were altered during nitrogen stress in the Δagp background. Additionally, we investigated whether the inability to synthesize glycogen could free carbon and energy into an engineered farnesene sink, through heterologous expression of a farnesene synthase in the Δagp background.

*Glycogen deficiency triggers the non-bleaching phenotype in Anabaena*

As observed in previously characterized ΔglgC cyanobacteria, the absence of glycogen plays an important role in the cell's ability to properly adapt to nitrogen limitation. The ability to synthesize glycogen during nitrogen stress allows WT filaments to catabolize phycobilisomes as a temporary nitrogen source and continue cell division during the onset of nitrogen starvation. As the filaments develop heterocysts, glycogen stores are respired for energy, and the products of $N_2$-fixation allow for the regeneration of phycobilisomes in vegetative cells. In contrast, the Δagp mutant was incapable of

glycogen synthesis and phycobilisome degradation, resulting in a complete cessation of cell division that only commenced after $N_2$-fixation had been established in mature heterocysts. Although Δagp heterocysts exhibited lower $N_2$-fixing activity throughout diazotrophic growth, the filaments continue cell division, ultimately reaching optical density levels similar to the WT approximately 200 hours after nitrogen-stepdown. Therefore, the rate of $N_2$-fixation does not appear to be a limiting factor during long-term diazotrophic growth.

The connection between the absence of glycogen and phycobilisome retainment during nitrogen-starvation is a fascinating yet poorly understood area of cyanobacteria metabolism. Deconstruction of phycobilisome antennae complexes is triggered by NblA, an adaptor protein that binds to phycobilin subunits and targets their destruction by recruiting proteolytic chaperones (Karradt et al. 2008; Nguyen et al. 2017). Since transcriptional levels of *nblA* were upregulated during nitrogen starvation in ΔglgC *Synechocystis* sp. PCC 6803, it is presumed that the disruption in the phycobilisome breakdown pathway occurs downstream of *nblA* induction (Carrieri et al. 2017). Besides maintaining structural integrity, phycobilisomes remain energetically coupled to PSII in ΔglgC *Synechococcus* sp. PCC 7002, and the cells retain an extensive thylakoid membrane with no apparent reduction in ribosome content (Jackson et al. 2015). This suggests that glycogen-deficiency does not disrupt solely phycobilisome breakdown, but extends globally across many nitrogen-deprivation signals. The identity of the "effector agent" involved in disrupting the nitrogen-stress transition is obscure, but many groups have suggested it to be a metabolite synthesized as a by-product from altered carbon

partitioning in the glycogen-deficient background (Carrieri et al. 2017; Hickman et al. 2013; Jackson et al. 2015).

It must be noted that the mechanism preventing phycobilisome degradation during in vegetative cells of Δagp *Anabaena* is not observed in heterocysts. PAR scans of isolated heterocysts from both WT and Δagp *Anabaena* indicated similarly low absorbance of phycobiliproteins (630nm). An independent experiment showing loss of phycobiliprotein auto-florescence in Δagp *Anabaena* heterocysts corroborates this finding. This data suggests that the activity of the posited "effector" agent is possibly absent or in too low in abundance in heterocyst to prevent phycobilisome breakdown, emphasizing a remarkable intracellular difference between the differentiated cell types. It is also possible that heterocysts may develop from daughter vegetative cells with little to no phycobilisomes.

*Glycogen availability effects the rate of N$_2$-fixation but is not essential for heterocyst development*

Analyses of heterocyst frequency and nitrogenase activity in the glycogen-deficient mutant reveals a number of key points on how glycogen's absence affects the NSR. The glycogen-deficient *Anabaena* maintained the ability to develop heterocysts, suggesting that the absence of glycogen did not prevent the mutant from sensing the nitrogen status and inducing heterocyst development through transcriptional regulation (e.g. NtcA, NrrA, HetR). This is in agreement with a transcriptional analysis of nitrogen-starved ΔglgC *Synechocystis* sp. PCC 6803, which showed a NtcA-mediated transcriptional response during nitrogen stress (Carrieri et al. 2017). However, the Δagp mutant exhibited a higher frequency of heterocysts compared to the WT 48 hours after nitrogen-stepdown.

The regulators PatS and HetN, which control early and late stages of heterocyst patterning respectively, diffuse from heterocysts and inhibit nearby vegetative cells from undergoing differentiation (Callahan and Buikema 2001; Liang et al. 1992; Muro-Pastor and Hess 2012).  It is plausible that a by-product of altered carbon-partitioning in Δagp could be inhibiting the activity of one or both of these regulators, akin to the repression in phycobilisome degradation, and that the inhibition is alleviated once matured heterocyst begin supplying the filaments with fixed nitrogen.  A transcriptional and proteomic survey of heterocyst maintenance genes in the Δagp mutant would be useful in uncovering any differences in NSR regulation, and may lead to a probable cause behind this patterning anomaly.

Compared to the WT, nitrogenase activity in matured heterocysts was significantly depressed in the glycogen-deficient mutant during protracted periods of nitrogen-depletion.   Considering that a NifB-GFP reporter fusion revealed no differences between *nifB* expressions in WT or Δagp heterocysts, we posit that this difference lies in the post-translational level.  Nitrogen-fixation is an energy-intensive process requiring both ATP and reductants, and increases exponentially shortly after glycogen begins to become catabolized  (Ernst et al. 1984).  It is possible that the absence of glycogen lowers the overall pool of carbon compounds and reductants needed for nitrogenase to efficiently catalyze.  Findings from an nrrA-deleted mutant of *Anabaena*, which inhibited glycogen catabolism, also resulted in lowered nitrogenase activity in *Anabaena* during diazotrophic growth. (Ehira and Ohmori 2011).

*Loss of a dynamic carbon sink lowers PSII activity during N₂-fixation*

In saturated light, photosynthetic activity is sensitive to changes in combined-nitrogen availability, shown by low PSII activity in both WT and Δagp vegetative cells during the nitrogen stress transition (14-18 hours after nitrogen-stepdown), compared to nitrogen-replete conditions.   In the WT, lowered PSII activity primarily occurs through the degradation of photosynthetic components in the thylakoids, including light-harvesting antennae on PSI/PSII complexes.  In this way, phycobilisome disassembly not only provides a temporary nitrogen source during nitrogen stress, but acts to mitigate excess light absorption and prevent unwanted reactive-oxygen species (ROS) generated by PSII during an absence of downstream metabolic reactions, (e.g. nitrogen assimilation).  Once N₂-fixation begins, WT cells can begin to rebuild their photosynthetic complexes and restore PSII activity to nitrogen-replete conditions.    In contrast, the abundance of phycobilisome antennae in the Δagp vegetative cells possibly prevented the cells from lowering light absorption, and in the absence of a native glycogen sink to balance the cellular redox state, the photosynthetic electron transport chain (PETC) was constantly bottlenecked, thus inhibiting PSII efficiency.  If the observed low PSII activity during diazotrophic growth is due to disrupted electron flow, it is worthy to note that the consumption of PETC products (e.g. ferredoxin, ATP) via nitrogen-fixation in heterocysts did not adequately relieve the PETC, considering photosynthetic activity was low even when nitrogen-fixation was underway.  Indeed, the depressed rate of photosynthetic activity and nitrogen-fixation point to glycogen as being a vital component in both these activities.

*Abolishing native carbon sinks to increase engineered product synthesis*

Redistributing metabolic flux from major carbon sinks to a targeted product is a key strategy to producing commodity chemicals from engineered microbes at quantities that are economically sustainable. Abolishing glycogen as a carbon sink may free up energy and metabolites, and re-direct their flow into engineered biofuel sinks, such as farnesene. Although glycogen biosynthesis was blocked, there was no apparent increase in flux through the MEP pathway, as observed from farnesene yields in Δagp when expressing FaS. Similar results were found with limonene and bisabolene production in ΔglgC *Synechococcus* sp. PCC 7002 (Davies et al. 2014), and may be a reflection of tight regulatory control over MEP enzymes, or low amounts of reductants needed to drive metabolic flux. Other chemicals, such as sucrose and isobutanol, have shown an increase in production in the glycogen-deficient background (Ducat et al. 2012; Li et al. 2014), suggesting that increased metabolic flux imparted by glycogen abolishment may be highly dependent on the pathway of interest. Future research will need to be conducted to explain the "metabolic flux bias" in glycogen-deficient cyanobacteria strains.

**Chapter 6:  Introducing 3-hydroxyproprionate pathway genes to prevent photorespiration and increase linalool production in *Anabaena* sp. PCC 7120**

## 6.1 Abstract

In photosynthetic organisms, the assimilation of $CO_2$ into organic compounds is limited by the slow catalysis and poor substrate specificity of RuBisCO, the major carboxylating enzyme of the CBB cycle.  Oxygenation of 3-PG by RuBisCO leads to photorespiration, a process that results in the loss of fixed carbon and nitrogen for the cell.  Synthetic biology now allows for the expression of foreign metabolic pathways *in vivo*, in order to improve metabolic processes designed by nature.  Here, we borrowed genes from the 3-hydroxypropionate bicycle to create a photorespiratory bypass in the filamentous, $N_2$-fixing cyanobacteria *Anabaena*.  The bypass is designed as an additional carbon fixation pathway, while also recycling glyoxylate from photorespiration into pyruvate.  Furthermore, we expressed a LinSDXP operon alongside the bypass to pull extra pyruvate into the MEP pathway for linalool synthesis.

## 6.2 Introduction

The conversion of sunlight into fixed carbon is an essential starting point for the creation of all central metabolites in nature.   The Calvin-Benson-Bassham (CBB) cycle is the most significant carbon fixation process on Earth, utilized by terrestrial plants, eukaryotic algae, and cyanobacteria for primary production.  Terrestrial primary production from the CBB cycle accounts for approximately 56-57 gigatons $yr^{-1}$ of biomass annually (Geider et al. 2001; Normile 2009).  Modern agriculture depends on the CBB cycle for annual crop yields, as well as a majority of the cultivatable land and fresh

water on our planet.  The rate of carbon fixation by photosynthesizing organisms varies

in nature, and is mostly dependent on the availability of water, light, phosphorous, fixed-

nitrogen, and iron.  However, man-made advancements in irrigation, nutrient

supplementation, and pesticide treatment can make carbon fixation a rate-limiting factor.

RuBisCO is the carboxylating enzyme that assimilates $CO_2$ with ribulose-1,5-

bisphosphate to create 3-phosphoglycerate (3-PG) in the CBB cycle.  Unfortunately, the

enzymatic activity of RuBisCO is extremely slow, only fixing 3-10 $CO_2$ molecules $s^{-1}$.

The competing oxygenase activity of RuBisCO also produces 2-phosphoglycolate (2-

PG), a byproduct that is recycled through the photorespiratory $C_2$ cycle.  This costly cycle

requires ATP and reducing equivalents, and leads to a 25% loss in fixed carbon. Efforts

to engineer RuBisCO with higher affinity for $CO_2$ have been met with limited success.

Evidence has shown that RuBisCO might already be optimized, despite confused

substrate specificity and a slow catalysis rate.

One strategy to stop this carbon and energy loss is to pull the byproducts of

photorespiration back into an engineered RuBisCO-independent carbon fixation pathway,

where they can be used for biosynthesis (Figure 34). The 3-hydroxypropionate (3HP) bi-

cycle is an $O_2$-insensitive, RuBisCO independent carbon fixation pathway utilized by the

green, non-sulfur, thermophilic bacterium *Chloroflexus aurantiacus*. Incorporating one

half of this pathway alongside the CBB cycle could offset the wasteful products of

photorespiration by pulling 2-PG into the engineered bypass, which would eventually

convert it to pyruvate for many useful biosynthetic pathways.  The 2-*C*-methyl-D-

erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate (MEP/DOXP) pathway in

cyanobacteria has gained attention for its use in producing terpenoids in genetically-

engineered cyanobacteria.  Terpenoids are long-chained alkanes and alkenes that are particular interest as biofuels, considering their high energy density, similarities to existing fuels (gasoline, diesel, jet), and propensity to separate from the processing residues.  Pyruvate is a precursor substrate for the MEP/DOXP pathway, and is condensed with 3-PG by the enzyme DXP synthase (DXS) to create DXP in the first committed step of the MEP pathway.  Considering that overexpressing DXS, along with two downstream enzymes in the MEP pathway lead to an increase in linalool productivity in *Anabaena* (Chapter 4), we posited that the addition of a functional, respiratory bypass may have synergistic effects on terpenoid productivity in engineered cyanobacteria.
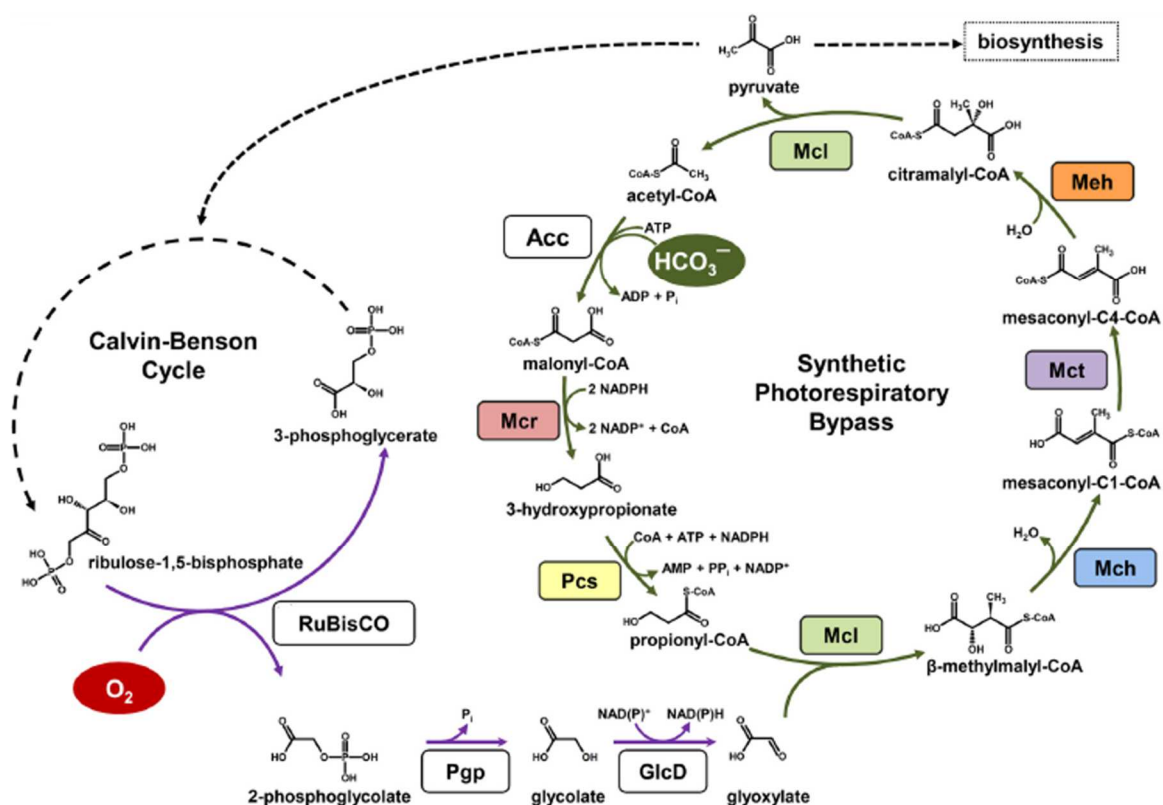
**Figure 34**. Schematic of the engineered photorespiratory bypass in *Anabaena*. Colored boxes are introduced enzymes for the bypass, while non-colored are native to *Anabaena*. Abbreviations: MCL, (*S*)-malyl-CoA/ß-methylmalyl-CoA/(*S*)-citramalyl-CoA (MCC) lyase; MCH, mesaconlyl-C1-CoA hydratase; MCT, mesaconlyl-CoA C1-C4 CoA transferase; MEH, mesaconyl-C4-CoA hydratase; ACC, acetyl-CoA carboxylase; MCR, malonyl-CoA reductase; PCS, propionyl-CoA synthase; Ru-1,5-BP, ribulose-1,5-bisphosphate; 3-PG, 3-phosphoglycerate; 2-PG, 2-phosphoglycerate; Glyc, glycolate; Glyox, glyoxylate; Me-CoA, ß-methylmalyl-CoA; Mes-1-CoA, mesaconyl-C1-CoA; Mes-4-CoA, mesaconyl-C4-CoA; Cit-CoA, citramalyl-CoA; Ac-CoA, acetyl-CoA; Mal-CoA, malonyl-CoA; 3-HPO, 3-hydroxypropionate; Prop-CoA, propionyl-CoA. Adapted from Shih et al. 2014

In this study, we designed a two-plasmid system for the expression of six 3-hydroxyproprionate enzymes from *C. aurantiacus*, as well as four enzymes for photosynthetic linalool production (LinSDXP), in the filamentous $N_2$-fixing cyanobacterium Anabaena sp. PCC 7120 (herein after *Anabaena*). The replicative plasmids were conjugally transferred in *Anabaena*, in order to siphon the products of photorespiration into pyruvate and increase carbon flux to linalool synthesis. Our findings indicate that *Anabaena* strain harboring the two plasmids (named LinSDXP-SPB

*Anabaena*) exhibited lagged growth and lower biomass yield than either the WT (no plasmids) or *Anabaena* strain carrying a single plasmid for linalool synthesis genes only (LinSDXP *Anabaena*). Interestingly, we observed a higher amount of linalool synthase (LinS) expression in LinSDXP-SPB *Anabaena*, correlating with a higher rate of linalool production in LinSDXP-SPB *Anabaena* during early stages of growth. However, linalool synthesis is unsustainable in the two-plasmid mutant, as production drops over time. We discuss possibilities for these observations and offer future aspects for engineering *Anabaena* with a functional photorespiratory bypass.

### 6.3 Methods & Materials

#### Strains and growth conditions

*Anabaena* strains were grown photoautotrophically at 30°C in BG11 (containing 17.6 mM NaNO$_3$ under constant illumination at 50 µmol m$^{-2}$ s$^{-1}$. Cultures were constantly agitated at 120 rpm in an Innova 44 lighted incubator (New Brunswick Scientific). Transgenic *Anabaena* strains were grown with neomycin (50 µg mL$^{-1}$) and erythromycin (10 µg mL$^{-1}$). *E. coli* used for cloning and conjugal transfer were grown in Luria-Bertani (LB) broth and shaken at 37°C, while rotating at 250 rpm. Kanamycin (50 µg mL$^{-1}$), ampicillin (100 µg mL$^{-1}$), and chloramphenicol (25 µg mL$^{-1}$) antibiotics were used to select for *E.coli* harboring engineered plasmids.

#### Plasmid Construction

All plasmids used in this study is listed in Table 9, and were constructed with traditional cloning methods (Sambrook and Russell 2001). The plasmids pCM62-MP and pUC19-10-12-13-11 were gifted by Matthew Mattozzi from the Wyss Institute for Biologically Inspired Engineering, Harvard University, MA (Mattozzi et al. 2013).

Plasmid pCM62-MP, containing *E.coli* codon-optimized genes *mcr* (Caur2614, 3.7 kb) and *pcs* (Caur0613, 5.4 kb) from the *C. aurantiacus* J-10-fl genome (Tang et al. 2011) was cut with EcoRI and XhoI restriction endonucleases, and the released fragment containing both genes was ligated into EcoRI-XhoI cut pNir, a cyanobacteria expression vector (Desplancq et al. 2005).  The resulting plasmid was named pZR2053 (Table 9).  Plasmid pUC19-10-12-13-11, containing the genes *mcl* (Caru174), *mct* (Caur175), *meh* (Caur180), and *mch* (Caur173) was cut with BglII and BamHI restriction endonucleases, and the four-gene fragment was ligated into BamHI-AvrII cut pZR2050 to create pZR2054.  The plasmid pZR2050 is a modified version of the cyanobacteria expression vector pZR670, and was created by ligating an annealed, oligonucleotide duplex from ZR1293 and ZR1294 in BamHI-NsiI cut pZR670 (Table 9).

The linalool synthesis genes used in this study are as follows:  *Picea abies* linalool synthase (*lins*, GenBank: AY473623.1), *E.coli* DXP synthase (*dxs*, GenBank: EFF14228.1), *Haematococcus pluvialis* IPP isomerase (*idi*, GenBank: AF082325.1), and GPP synthase (*gpps*, GenBank: EFO75729.1).  The four genes were arranged on a four-gene operon in pZR1464 (pLinSDXP, see Chapter 4 in this dissertation).  To combine the 4-gene *mcl-mct-meh-mch* operon with *lins-dxs-idi-gpps*, pZR1464 was modified to add an additional multiple-cloning site (MCS) by ligating an annealed, oligonucleotide duplex from ZR1577 and ZR1578 to NotI-AflII cut pZR1464 to create pZR2212.  Then, *mcl-mct-meh-mch* was excised from pZR2054 with KpnI and XhoI, and ligated into KpnI-XhoI cut pZR2212 to create pZR2213.  In order to ensure both pZR2053 and pZR2213 could be maintained in *Anabaena* with different antibiotics, the Nm resistant marker in pZR2213 was replaced with a $Cm^r/Em^r$ cassette by cutting the $Nm^r$

aminoglycoside phosphotransferase ORF with FspI, and blunt-ligating a Cm$^r$/Em$^r$ cassette from pZR2222 (Table 9). *E.coli* colonies harboring pZR2247 were tested for Nm susceptibility as well as Cm/Em resistance.

*Biomass growth*

Culture growth was determined by transferring 0.15 ml aliquots of culture into a Costar 96-well plate and reading the optical density at 700nm, using a Biotek Synergy 2 plate reader. Dried cell weight (DCW) was measured by centrifuging cultures at 4000 xg for 10 minutes, removing supernatant, drying cell pellets for 1 hour at 200°C in a drying oven. Pellets were weighed using an Ohaus Adventurer Pro scale.

*Transferring cargo plasmids into Anabaena*

The insertion of pZR2053 and pZR2247 into *Anabaena* was performed via successive conjugal transfers using a tri-parental *E. coli* mating system (Elhai et al. 1997) with the following modifications. *E. coli* HB101 bearing helper (pRL623) and conjugal (pRL443) plasmids were mated with *E. coli* NEB 10β harboring pZR2053 and pZR2247 respectively, and selected on LB plates containing triple antibiotic selection for the three plasmids. Selected colonies were grown overnight in 2 mL of LB containing appropriate antibiotics, subcultured by adding 200 µL of overnight culture to 2 mL of fresh LB containing appropriate antibiotics, and grown for additional 3 hours. Cells were pelleted through centrifugation, washed with 3x 1 mL of LB to remove antibiotics, and resuspended with 200 µL LB and subjected for mating with *Anabaena*. A 10 mL culture of WT *Anabaena* was grown to early exponential stage (O.D.700 = 0.3) and then sonicated (Branson 1510 water bath sonicator) for 60-120 seconds to break filaments into

2-4 cell lengths, which were confirmed under microscopy. Cells were harvested by centrifugation at 5,000 x g for 10 minutes, resuspended in 200 µL of fresh BG11, and then mixed with above *E. coli* harboring triple plasmids for conjugal transfer of plasmids from E. coli into *Anabaena*. The above cell mixture was placed under lighted conditions at 25°C for 30 minutes, micro-pipetted on to a Millipore Immobolin nitrocellulose filter placed on a BG-11 + 5% LB plate, and grown under white light for 3 days at 30°C. The filter was then transferred to a BG11 plate containing Nm100 µg/mL to select for positive exconjugants. Individual exconjugant colonies were further purified by restreaking onto fresh BG11 plates containing 100 µg/mL Nm and routine colony PCR was used for verification of positive exconjugants. After the successful insertion of pZR2053 was verified in *Anabaena,* a second conjugation was performed to conjugally transfer pZR2247, and colonies were selected for on BG11 + 1% agar plates containing 100 µg ml$^{-1}$ Nm and 15 µg ml$^{-1}$ Em.

*Protein extraction and Western blotting*

Protein was extracted from both WT and LinSDXP and LinSDXP-SPB *Anabaena* cells to determine expression of LinS in the host strains. Ten mL culture was grown in BG11 with appropriate antibiotics until it reached a mid-exponential phase (O.D.$_{700}$ = 0.5), and the cells were pelleted and washed 3x with sucrose buffer (50mM Tris-HCl, 40mM EDTA, 0.75M sucrose) resuspended with 50 µL of lysis buffer (0.5 mg/mL lysozyme, 10 ug/mL DNAse I, and 10 ug/mL RNase A.), and incubated at 37°C for 15 minutes. After incubation, 50 µL of 2x SDS loading buffer (50mM Tris-Hcl, pH 6.8, 4% SDS, 20% [vol/vol] glycerol, 200 mM dithiothreitol, 0.03% bromophenol blue) was added, and the mixture was boiled for 5 minutes, and then centrifuged at 13,000 rpm for

10 minutes at 4°C.  5 µL protein extracts were loaded and separated by 12% SDS-PAGE at 200V for 30 minutes. The total proteins were transferred onto a PVFD membrane, stained by Comassie blue, and $His_6$-LinS were detected by Western blot using anti-His antibodies (Invitrogen) at a dilution of 1:5000.

*Linalool identification and quantification from Anabaena*

Wild-type (WT) and linalool-producing *Anabaena* strains were grown in 250-mL Erlenmeyer flasks with either 100mL BG11 or $BG11_0$, and bubbled with filtered air at a rate of 100 mL $min^{-1}$.  A small glass column filled with 50 mg of Supelpak 2SV resin (Supelco) was attached to the exhaust port of each flask to capture linalool and other volatiles from the culture headspace.  Volatiles were washed from the 2SV resin using the following procedure:   1 mL of pentane containing 5 µg $mL^{-1}$ tetracosane as an internal standard (IS) was added to each resin sample and vortexed at 3000 rpm for 1 minute, incubated at room temperature for 10 minutes, followed by a brief 10-sec sonication (3000 rpm).  Volatiles extracted from the pentane wash were then analyzed using gas chromoatography-mass spectrometry (Agilent 7890A/5975C).  1 µL injected samples were separated using a HP-5MS column (35m x 250µm x 0.25µm), with $H_2$ as the carrier gas.  The oven temperature was initially held at 60°C for 2 minutes, and increased 20°C $min^{-1}$ until a temperature 300°C was reached.  Analytes were identified by comparing mass spectra from compounds in the NIST/EPA/NIH Mass spectral library, version 2.0.

**Table 9.** Plasmids used for this study

| Name | Relevant Characteristics | Construction | Source |
|------|--------------------------|--------------|--------|
| pRL443 | Conjugal plasmid, derivative of RP4; $Amp^r$, $Tc^r$ | | (Elhai et al. 1997) |
| pRL623 | Helper plasmid; $Cm^r$ | | (Elhai et al. 1997) |
| pRL271 | Empty cyanobacteria integration vector; $Em^r$ | | (Black et al. 1993) |
| pZR670 | Empty cyanobacteria replicative vector; $Cm^r$ $P_{glnA}$ | | (Chen et al. 2015) |
| pNir | Empty cyanobacteria replicative vector; $Kan^r$/$Neo^r$ | | (Desplancq et al. 2005) |
| pCR2.1TOPO | Empty TA cloning vector; $Kan^r$/$Amp^r$ | | Invitrogen |
| pCM62-MP | *mcr-pcs* | | (Mattozzi et al. 2013) |
| pUC19-10-12-13-11 | *mcl-mct-meh-mch* | | (Mattozzi et al. 2013) |
| pZR2050 | $P_{glnA}$-MCS; $Cm^r$ | Annealed oligonucleotides ZR1293 and ZR1294 into BamHI-NsiI cut pZR670 | This study |
| pZR2053 | $P_{nir}$-*mcr-pcs*; $Kan^r$/$Neo^r$ | EcoRI-XhoI fragment from pCM62-MP, ligated into EcoRI-XhoI cut pNir | This study |
| pZR2054 | $P_{glnA}$- *mcl-mct-meh-mch*; $Cm^r$ | BglII-BamHI fragment from pUC19-10-12-13-11 ligated into BamHI-AvrII cut pZR2050 | This study |
| pZR2212 | $P_{nir}$/$P_{psbA1}$-$H_6$-*lins-dxs-idi-gpps*; $Kan^r$/$Neo^r$ | Annealed oligonucleotides ZR1577 and ZR1578 into NotI-AflII cut pZR1464 | This study |
| pZR2213 | $P_{nir}$/$P_{psbA1}$-*lins-dxs-idi-gpps*-$P_{glnA}$- *mcl-mct-meh-mch*; $Kan^r$/$Neo^r$ | KpnI-XhoI from pZR2054 into KpnI-XhoI cut pZR2212 | This study |
| pZR2222 | $Cm^r$/$Em^r$; $Kan^r$/$Amp^r$ | Cloned $Cm^r Em^r$ cassette from pRL271 using primers ZR1584 and ZR1585, ligated into pCR2.1TOPO vector | This study |
| pZR2247 | $P_{nir}$/$P_{psbA1}$-$H_6$-*lins-dxs-idi-gpps*-$P_{glnA}$- *mcl-mct-meh-mch*; $Em^r$/$Cm^r$ | ZraI-Eco53kI fragment from pZR2222 ligated into FspI cut pZR2213 | This study |
| pZR1464 (pLinSDXP) | $Kan^r$, $Nm^r$; $P_{nir}$-$P_{psbA1}$, *lins-dxs-idi-gpps* | Digested pZR1462 with XhoI-NotI to excise 3872-bp *dxs-idi-gpps*, ligated into SalI-NotI pLinS | This dissertation, Ch.4 |

Acronyms: $Amp^r$, ampicillin resistance; $Tc^r$- tetracycline resistance; $Cm^r$- chloramphenicol resistance; $Kan^r$-kanamycin resistance; $Neo^r$, nemocyin resistance$P_{psbA1}$, psbA1 promoter from *Anabaena*; $P_{nir}$, nitrate reductase promoter from *Anabaena*; $P_{glnA}$, glutamine synthetase promoter from *Anabaena*; TTS, transcriptional termination site; RBS, ribosomal binding site; MCS, multiple cloning site; SDM, site-directed mutagenesis; $F_1$-flanking region from *Anabaena*; $H_6$, Histidine ×6 epitope tag;

**Table 10.** Oligonucleotides used in this study

| Name | Sequence (5' to 3') |
|---|---|
| ZR1293 | TATGGGATCCTTCGAAAGGCCTAGGGACGTCGTACGCTCGA |
| ZR1294 | GATCTCGAGCGTACGACGTCCCTAGGCCTTTCGAAGGATCCCATATGCA |
| ZR1577 | GGCCATACGTAGGTACCATGCATCTCGAGCGGCCGGTTTAAACC |
| ZR1578 | TTAAGGTTTAAACGCGGCCGCTCGAGATGCATGGTACCTACGTAT |
| ZR1584 | ATGCATATGCTAGCGACGTCGGATCCCTTAACTTACTTATTAAATAATTTATAG |
| ZR1585 | TCCCGGGAAGTATCCAGCTCGAGATC |

**Table 11.** Strains used in this study

| Strains | Relevant characteristics | Source |
|---|---|---|
| One Shot® Top10 *E.coli* | Cloning host | Invitrogen |
| NEB 10-beta *E.coli* | Cloning host | New England Biolabs |
| HB101 *E.coli* | For conjugal transfer of cargo plasmids into *Anabaena* sp. PCC 7120 | This study |
| *Anabaena* sp. PCC 7120 | Wild-type *Anabaena* sp. PCC 7120 | This study |
| LinS *Anabaena* | *Anabaena* sp. PCC 7120 carrying plasmid pZR808 | This study |
| LinSDXP *Anabaena* | *Anabaena* sp. PCC 7120 carrying plasmid pZR1464 | This study |
| LinSDXP-SPB *Anabaena* | *Anabaena* sp. PCC 7120 carrying plasmid pZR2247 and pZR2053 | This study |

## 6.4 Results

### Engineering of LinSDXP-SPB Anabaena

The SPB requires seven separate enzymes to convert glyoxylate produced from $C_2$-photorespiration into pyruvate.  Acetyl-CoA carboxylase (ACC), which adds bicarbonate to a molecule of acetyl-CoA in fatty acid biosynthesis, is already found in *Anabaena* as an alpha-subunit (*alr5285*) and a beta-subunit (*all2364*).  The remaining genes from *C. aurantiacus* used in this study with relevant characteristics are listed in Table 12.  The genes were acquired from researchers from a previous study that tested the validity of expressing sub-pathways of the 3HPA-bicycle in *E.coli* for producing precursors for biomaterials (Mattozzi et al. 2013).  We employed combinatorial plasmid engineering

strategy to combine the six synthetic photorespiratory enzymes (named SPB in this study)

with four linalool synthesis enzymes (LinSDXP) for expression in *Anabaena*.

**Table 12.** Characteristics of synthetic photorespiratory genes used in this study

| Full name | Acronym | Gene length (bp) | Protein molecular weight (kD) | RBS strength[c] |
|---|---|---|---|---|
| L-malyl CoA lyase | *mcl* | 1047 | 38.4 | 1.0 |
| mesaconyl-C1-CoA-C4-CoA transferase | *mct* | 1230 | 44.8 | 0.764 |
| mesaconyl-C4-CoA hydratase | *meh* | 843 | 31 | 0.6 |
| β-methylmalyl-CoA dehydratase[a] | *mch* | 1104 | 40.4 | unknown |
| malonyl-CoA reductase[b] | *mcr* | 3696 | 133.3 | unknown |
| propionyl-CoA synthase[a] | *pcs* | 5499 | 202.5 | unknown |

[a] contains a C-terminal Myc epitope tag; [b] contains a N-terminal Myc epitope tag; [c] RBS strength was obtained from the Registry of Standard Biological parts: http://partsregistry.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Community_Collection

The genes *mcl-mct-meh-mch* were expressed as a single 4.3kb operon by splicing

them downstream of the *Anabaena* glnA promoter ($P_{glnA}$), a strong promoter used

previously for protein expression in *Anabaena* (Chen et al. 2015). The $P_{glnA}$-*mcl-mct-*

*meh-mch* construct was then inserted directly downstream of the LinSDXP operon in

pZR1464 to create a two-operon system in a single plasmid, called pZR2247 (Figure 35).

The genes *mcr* and *pcs* are large genes, with lengths of 3.7 kb and 5.5 kb,

respectively. Because of their size and practical constraints of engineering, we chose to

express these genes together on a separate replicative plasmid in *Anabaena*. For strong

expression, the *mcr-pcs* operon was inserted directly downstream of PglnA promoter in

the plasmid pZR2050 (a derivative of pZR670) to create pZR2053 (Table 9). Both

pZR2247 and pZR2053 could then be utilized for co-expression of *mcr-pcs* with *mcl-mct-meh-mch* and *lins-dxs-idi-gpps*.  The plasmids were successively transferred into WT *Anabaena*, and resulting strain was named LinSDXP-SPB.
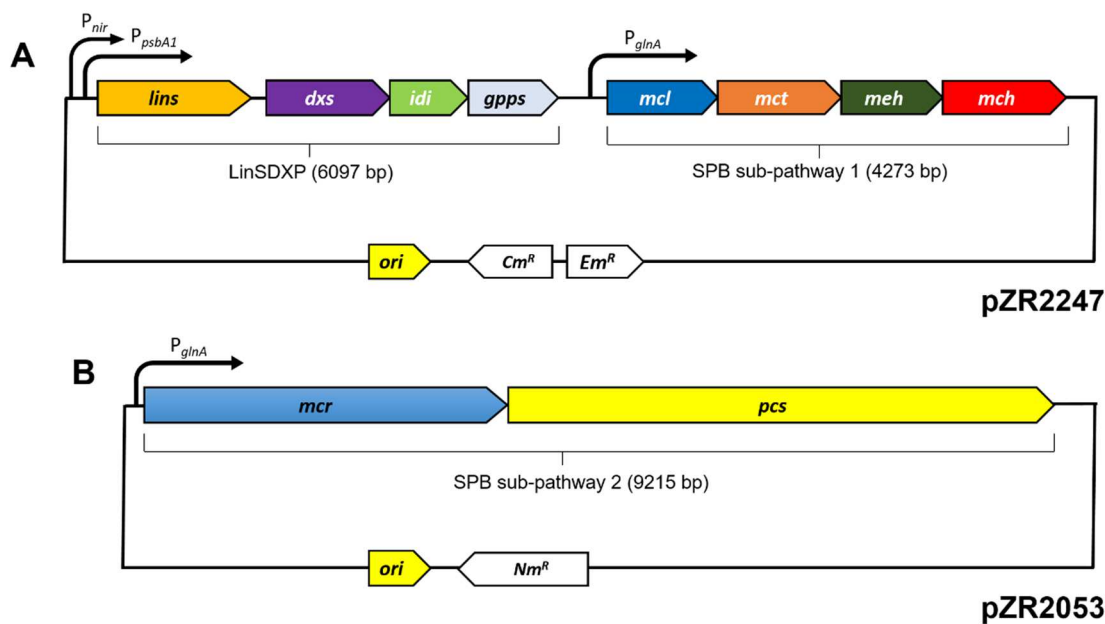


**Figure 35.**  Schematics of plasmids pZR2247 (A) and pZR2053 B) used for co-expressing synthetic photorespiratory bypass genes with linalool synthesis genes in *Anabaena.*

*Linalool production in LinSDXP-SPB Anabaena*

To discover if the expression of SPB genes impart any difference in LinSDXP-driven linalool production, we set up a triplicate linalool production trial between LinSDXP *Anabaena* (containing pZR1464) and LinSDXP-SPB *Anabaena* (containing pZR2247 and pZR2053), using the styrene-divinylbenzene resin (2SV) method as described previously (Halfmann et al. 2014a; Halfmann et al. 2014b) and in Chapters 3, 4, and 5 of this dissertation.  A WT *Anabaena* culture was also included in the experiment, to see if the inserted plasmids affected the growth rates of LinSDXP and LinSDXP-SBP

*Anabaena*.  Small volume starter cultures (25ml) cultures were grown to mid-exponential phase in BG11 with appropriate antibiotics, before transplanting cultures into 100-ml volumes of BG11 while bubbling with 1% $CO_2$.  Culture density was taken every two days, and samples from 2SV resin were analyzed with GC-MS to detect linalool in the flask headspace of all cultures.  While the WT and LinSDXP-*Anabaena* displayed a similar growth during the trial, LinSDXP-SPB growth was repressed (Figure 36A).  This difference was also seen through DCW measurements at the end of the trial, showing that LinSDXP-*Anabaena* had accumulated less biomass (170.6±8.2 mg DCW) than either WT (239±20 mg DCW) or LinSDXP (290.9±5.8 mg DCW) (Figure 36B).

Although cell density and biomass accumulation was lower in LinSDXP-SPB *Anabaena,* we measured a higher production of linalool than LinSDXP *Anabaena* during the first four days of the trial (Figure 36C).  Linalool productivity was over 2-fold higher in LinSDXP-SBP *Anabaena* (145±20.3 $\mu$g $L^{-1}$O.D.$^{-1}$day$^{-1}$) than LinSDXP *Anabaena* (66.3±17.8 $\mu$g $L^{-1}$O.D.$^{-1}$day$^{-1}$) during the first two days of growth, and productivity increased to 182.6±39.9 $\mu$g $L^{-1}$O.D.$^{-1}$day$^{-1}$ during days 3-4.  However this productivity was short-lived, as production dropped from 190.1±4.5 $\mu$g $L^{-1}$ on days 4 to 5 to 24.7±4.3 $\mu$g $L^{-1}$ on days 5-6 in LinSDXP-SPB *Anabaena*, and that production was steadily maintained throughout the remainder of the experiment.  Through days 9-10, LinSDXP-SBP *Anabaena* produced only 25.9±2.6 $\mu$g $L^{-1}$ linalool (compared to 97.4±4.8 $\mu$g $L^{-1}$ in LinSDXP *Anabaena*).  The drop in linalool production did not correlate with a drop in cell density, as optical density measurements showed a slight increase in growth during that time frame (Figure 36A).

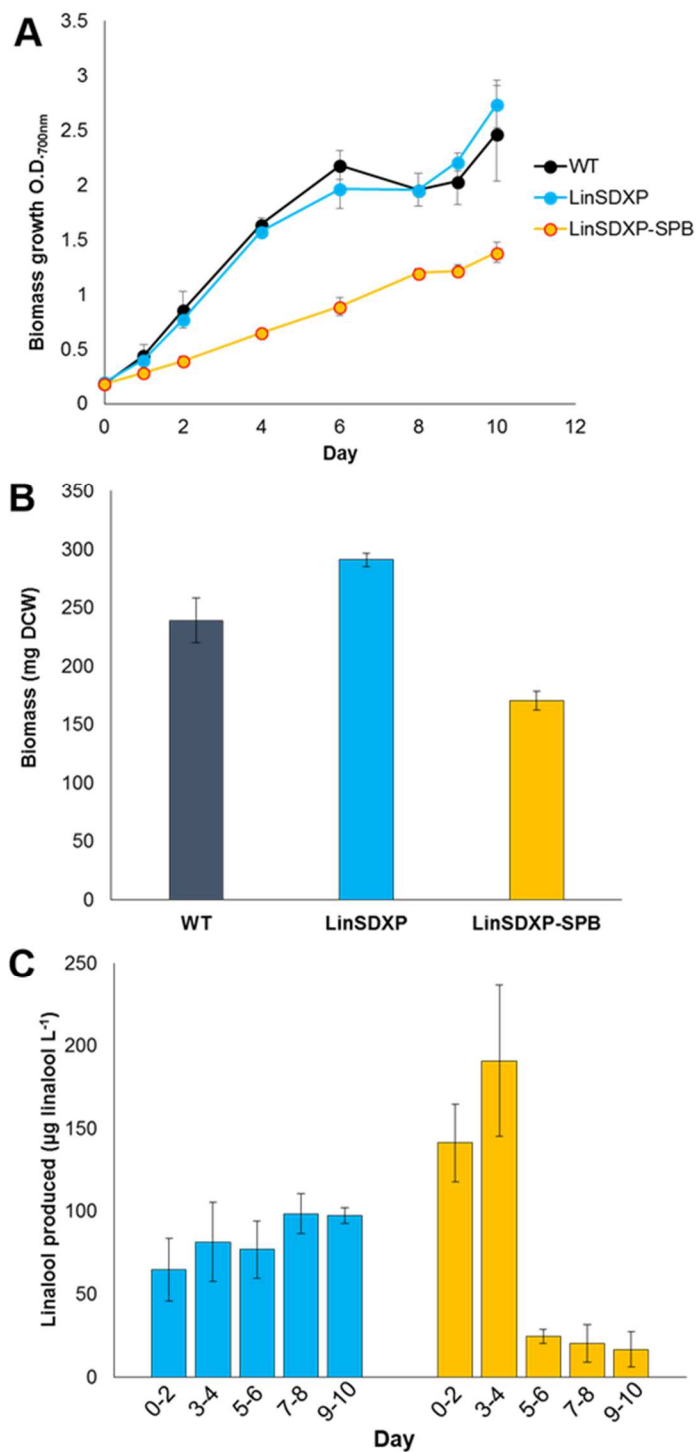**Figure 36.** Growth dynamics and linalool production from LinSDXP and LinSDXP-SPB *Anabaena* during photoautotrophic growth with 1% CO2. A) Optical density measurements B) Dry weight of total biomass at day 10 C) Linalool collected from LinSDXP (blue bars) and LinSDXP-SPB (orange bars) *Anabaena* during the time period indicated on the X-axis. Error is standard deviation of three technical replicates.

*Differences in LinS expression between multiple linalool-synthesis strains*

The sudden decline in linalool production exhibited in LinSDXP-SPB *Anabaena* made us question the stability of the plasmids in the cyanobacteria cells. PCR on LinSDXP-SPB *Anabaena* filaments at the end of the linalool production experiment showed the plasmids were still present (data not shown). We then decided to check the expression levels of LinS protein in LinSDXP-SPB *Anabaena* through Western blotting, in order to detect any possible inhibitory effects of the SPB proteins on *lins* expression. For comparison, we added LinS *Anabaena* (*Anabaena* strain harboring pZR808; *lins* only) and LinSDXP *Anabaena* (*lins-dxs-idi-gpps*) as protein expression benchmarks. We found that the 69kD LinS protein was expressed in LinSDXP-SPB *Anabaena*, and surprisingly, that it conferred an approx. 10-fold higher level of expression than either LinS or LinSDXP *Anabaena* (Figure 37). This can be visualized in the Western blot (Figure 37A) and the Coomassie-stained SDS-PAGE (Figure 37B), which shows a heavy ~70 kD protein band in an equally-loaded protein sample of LinSDXP-SPB *Anabaena*.
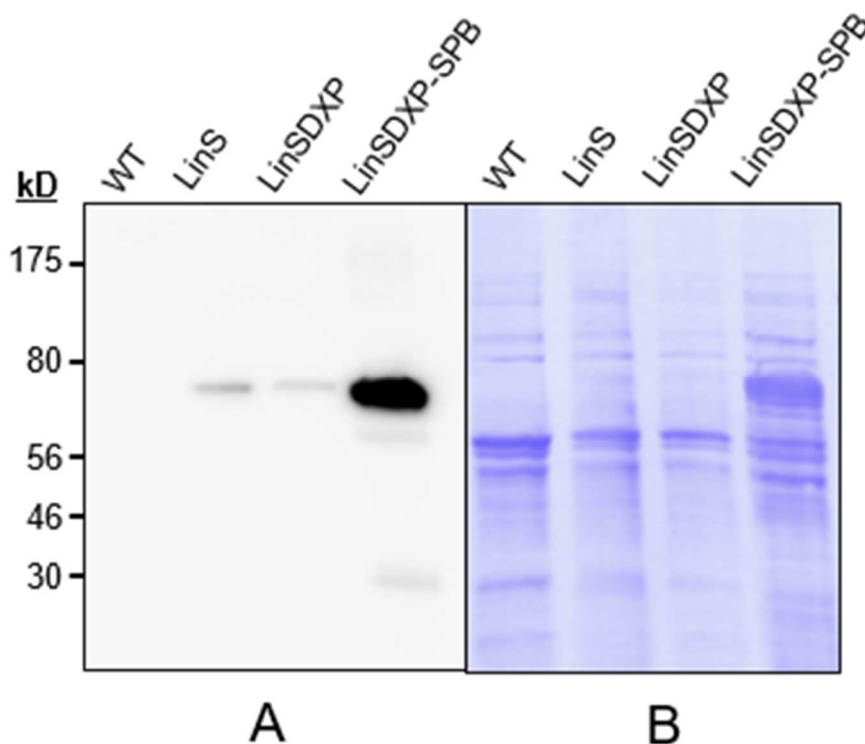
**Figure 37.** LinS expression in different linalool-synthesis strains of *Anabaena*. A) Western blotting with anti-His antibodies shows the 69-kD LinS protein expressed in *Anabaena* strains harboring genes for LinS only (LinS), LinS with DXS-IDI-GPPS enzymes (LinSDXP) and LinSDXP with enzymes for a synthetic photorespiratory bypass (LinSDXP-SPB). B) Coomassie-stained SDS-PAGE lanes show equal protein loading. Observations were made in duplicate experiments.

## 6.5 Discussion & Conclusions

The synthetic photorespiratory bypass for mitigating the wasteful processes of photorespiration was inspired by the 3-HP bicycle, a carbon fixation pathway discovered in the green nonsulfur bacterium *C. aurantiacus* (Ivanovsky et al. 1993). This microorganism lives commensally with cyanobacteria in hot springs, and the 3-HP bicycle enzymes are likely optimized for these thermophilic conditions (55-60°C). The oxygenase reaction of RuBisCO is favored at high temperatures, and may promote high levels of photorespiration from the hot spring cyanobacteria community. It is possible that the oxygen-insensitive 3-HPA bicycle serves two functions: as a carbon fixation

pathway in *C. aurantiacus*, and to salvage the high amounts of $C_2$ photorespiratory products (glyoxylate and glycolate) excreted from cyanobacteria neighbors (Zarzycki and Fuchs 2011). The pathway has also drawn interest from a groups looking to transfer carbon fixation pathways in heterotrophic microbes (Mattozzi et al. 2013) or to reassimilate photorespiratory products into biosynthesis in cyanobacteria (Shih et al. 2014). Introducing foreign metabolic pathways in microbes may be a crucial component in renewable fuel and biomaterial production systems in the future, and the experiments detailed in this study attempted to add to our knowledge in this area.

We assembled the necessary enzymes for a synthetic respiratory bypass in the filamentous, $N_2$-fixing cyanobacterium *Anabaena*, along with genes to direct carbon flux from SPB-driven pyruvate synthesis into the MEP pathway for linalool production. A total of ten genes were organized into three operons on two separate replicative plasmids, with strong promoters used to drive the constitutive transcription of all genetic elements. The engineered *Anabaena* (LinSDXP-SPB *Anabaena*) was able to maintain both large plasmids during the course of our experiments, and we observed that the strain exhibited lagged growth and unsustainable linalool biosynthesis during our production trials. It is possible that the high amount of linalool produced initially during the experiment is the result of higher LinS expression, which was found to be approx.10-fold higher in LinSDXP-SPB *Anabaena* than either LinS or LinSDXP *Anabaena*. The reason behind the elevated LinS protein levels in the SBP mutant is unclear, and further research will need to be conducted to elucidate whether this overexpression is occurring at the DNA (plasmid copy), transcriptional, or translational level. It is tempting to suggest that the SBP elements themselves may directly or indirectly increasing upstream LinS expression,

perhaps by altering the physical structure of the plasmid, interacting with the LinS

promoter, or some unknown mechanism.

Since LinSDXP *Anabaena* manifested a higher linalool productivity with lower

amounts of LinS protein during the middle and end of the trial, we hypothesize that the

activity of the SPB genes might be inhibiting linalool synthesis and delaying cell growth.

The possible deleterious effects caused by the SPB genes are numerous.  Intermediates in

the last two steps to complete the pathway, MCR and PCS, are toxic to cells.  3-OHP, the

product of MCR, is a β-hydroxy acid that may cause organic acid toxicity if it is

accumulated (Lipscomb et al. 2012).  Propionyl-CoA, the product of PCS, was also

shown to inhibit both pyruvate dehydrogenase and citrate synthase activity (Horswill et

al. 2001).  Furthermore, both MCR and PCS are large, multidomain proteins that may

struggle fold correctly in *Anabaena*, creating a toxic protein overload.  Indeed, all the

expressed SBP enzymes might be effected by protein-misfolding and low activity, since

they are from a thermophilic source.  Future efforts may need to focus measuring enzyme

expression and activity *in vivo*, in order to further investigate the physiological impacts of

the SBP pathway in cyanobacteria (Shih et al. 2014).  Once these variables are found, the

cycle can be optimized by finding mesophilic enzyme homologs and regulating their

expression through different promoters (Elhai 1993; Huang et al. 2010; Qi et al. 2013),

ribosomal binding sites (Oliver et al. 2014) and other genetic elements.

## Chapter 7: Summary & future directions

The renewed interest in alternative biofuels derived from agricultural crops, algal

biomass, and photosynthetic substrates in cyanobacteria have been recently triggered by

the fear of a reduction in energy production from non-renewable resources.  Of the

different biofuel energy strategies proposed, cyanobacteria have received significant

consideration, considering these microorganisms are able to manage the problems that

have arisen with preceding biofuel generations (first, second, and third-generation

biofuels).  Cyanobacteria offer a promising biomass feedstock for both organic (ethanol,

isobutanol, biodiesel) and inorganic ($H_2$) biofuels.  Many of the discussed strategies are

still under development, and their energy yield may not yet be feasible for industrial

production levels.  Therefore, further studies into the metabolic profiling and

engineering of cyanobacteria need to be implemented, in order to optimize energy

efficiency into biofuels for economic production on a commercial scale.

Despite the precarious position of cyanobacteria-based biofuels, the future of biofuels

produced by photosynthetic microorganisms looks optimistic. The potential of

genetically-modified cyanobacteria for biofuel production is still in its infancy.

Currently, only a handful of cyanobacterial species have been investigated for their

potential for chemical production.  As new species are discovered and sequenced and

new biotechnology tools and techniques become available, the inherit properties of these

microorganism can be exploited for human applications.  Increasing the knowledge

cyanobaceterial genetics, metabolism, and regulatory systems will improve production,

enable the synthesis of new products, and improve the ability to predict the outcomes of

genetic manipulation.  Furthermore, expanding knowledge into the quickly evolving

omics fields will help bring molecular-level cellular processes into the larger framework of global cellular regulation, genomics, metabolomics, and proteomics. Proof-of-concept demonstrations of 4th generation biofuel and chemical production from cyanobacteria and algae will hopefully catch the attention of industries, bringing the idea of direct $CO_2$ conversion into fuels from the laboratory to pilot-scale operations.  This next step in biofuel sustainability has already progressed, thanks to a number of private companies. Recently, Joule Unlimited announced the development of a 4-acre test plant in Hobbs, New Mexico, which is predicted to produce up to 10,000 gallon diesel/acre/year (almost 20 times higher productivity than corn ethanol).  At full-scale commercialization in ideal locations, the company ultimately targets 25,000 gallons of ethanol and 15,000 gallons of diesel per acre annually, for approximately $1.20/gallon.   Algenol, headquartered in Fort Myers, Florida, confirmed that they had exceeded production rates of 9,000 gallons of ethanol/acre/year from their 4-acre, outdoor Process Development Unit in Lee County, Florida.  This facility had previously achieved continuous ethanol production at 7,000 gallons ethanol/acre/year in September 2013 (up from a projected 6,000 gallons ethanol/acre/year).  Company CEO Paul Woods announced that this "Direct-to-Ethanol" technology allowed the production of ethanol for around $1.00/gallon, using sunlight, carbon dioxide, and seawater.  Many other companies are continuing research into extracting oil from directly from algae biomass (3rd generation biofuels), including Solazyme, Solix Biofuels, and Sapphire Energy.

The research discussed in this dissertation encompasses a series of genetic manipulations to increase photosynthetic farnesene and linalool production in the filamentous, $N_2$-fixing cyanobacterium *Anabaena* sp. PCC 7120. A summary of the implemented engineering changes and production results are listed in Table 13. The highest rate of farnesene productivity by expressing a codon-optimized farnesene synthase (*fas*) was calculated at 2.9 µg $L^{-1}$ $hr^{-1}$ O.D.$^{-1}$. Deleting the gene *agp* in the Anabaena chromosome and blocking glycogen synthesis was intended to re-distribute more fixed carbon into the MEP pathway, but instead resulted in a 72% decrease in productivity. A more successful series of genetic changes were observed for linalool productivity, by 1) expression of a plant linalool synthase gene *lins*, 2) co-expression of DXP operon (containing the genes *dxs-ippHp-gpps* with *lins*, and 3) co-expression of *lins-dxs-ippHp-gpps* with six 3HP genes to create a synthetic photorespiratory bypass (SPB). Expressing the SPB genes (SPB sub-pathway 1 and sub-pathway 2) along with the LinSDXP operon resulted in a productivity of 7.6 µg $L^{-1}$ $hr^{-1}$ O.D.$^{-1}$, demonstrating a 590% improvement in linalool productivity, compared to expressing *lins* alone.

**Table 13.** Summary of genetic engineering in *Anabaena* for terpenoid production

| Farnesene | | | |
|---|---|---|---|
| **Strain** | **Genotype** | **Highest observed production rate (µg $L^{-1}hr^{-1}$O.D.$^{-1}$)** | **Percent change (%)** |
| FaS-*Anabaena* | *fas* | 2.9 | - |
| FaS –glycogen *Anabaena* | *fas*; Δ*agp* | 0.8 | -72 |
| Linalool | | | |
| LinS- *Anabaena* | *lins* | 1.1 | - |
| LinSDXP- *Anabaena* | *lins-dxs-ippHp-gpps* | 2.8 | +154 |
| LinSDXP SPB-*Anabaena* | *lins-dxs-ippHp-gpps*; SPB sub-pathway 1 SPB sub-pathway 2 | 7.6 | +590 |

Applying parameters from algae productivity models (Chisti 2008), we calculated linalool productivity from the 3$^{rd}$ generation strain of *Anabaena* (LinSDXP-SPB) in a large-scale application. An ideal setting for a photosynthetic, linalool production facility in the Midwest would be near an ethanol plant, which could produce 275,000 metric tons of $CO_2$ annually and low grade heat to operate a 10 hectare greenhouse. Linalool productivity from a theoretical greenhouse can be calculated by taking in the following assumptions: a total enclosed photobioreactor (PBR) volume of 6.4 million L, a growing season of 300 days annually, and 8 hours of daily sunlight. Using these inputs, we calculated a productivity of ~117 kg linalool annually from this system, given our highest productivity of 7.6 $\mu$g L$^{-1}$ hr$^{-1}$ O.D.$^{-1}$. This calculated amount of linalool demonstrates that current productivities are too low to rationalize up-scaling at this stage of the engineering.

Recycling $CO_2$ from an ethanol plant into a greenhouse is an attractive scenario for linalool production, since $CO_2$ produced from the fermentation process could provide a feedstock for linalool synthesis. If a ten hectare facility could convert 1% of annual $CO_2$ emissions from one ethanol plant (275000 kg) into linalool, it would theoretically produce 963843 kg of linalool year$^{-1}$ (1116852 L year$^{-1}$, assuming a density of 0.863 g mL$^{-1}$). Assuming rough estimates for capital ($0.50 L$^{-1}$), labor ($50,000 × 6 people = $300,000 year$^{-1}$/1116852 L year$^{-1}$ = $0.27 L$^{-1}$) and operation ($0.50 L$^{-1}$), a total production costs of $1.27 L$^{-1}$ linalool is possible, compared to the current $4-5 L$^{-1}$. This considerable cost difference emphasizes the advantages of fourth generation biofuels, compared to biofuel-production systems that utilize land plants as a feedstock. It also

reveals that cyanobacterial engineering efforts would need to increase productivity to

$\sim$62.8 mg L$^{-1}$hr$^{-1}$O.D.$^{-1}$, in order for the above theoretical scenario to be successful.

The milestones reached by industries to engineer algae and cyanobacteria to convert inorganic carbon, water, and light into biofuels brings hope to the future of 4$^{th}$ generation biofuels. However, more research needs to be conducted to continue exploring the potential of microbial cell factories, as well as to identify specific drawbacks. Some of the most significant aspects that need to be addressed to advance the field need to span into multiple disciplines, including genetics, systems biology, biotechnology, and engineering. Increasing our knowledge on the genetic toolkit of cyanobacteria, designing cultivation strategies, and weighing the ecological risks of genetically-modified organisms on the environment need to be modeled in advanced before commercial sustainability can be truly realized. As these technologies progress, we envision $CO_2$ conversion systems using cyanobacteria to pave the way for a greener, post-petroleum era.

## Chapter 8: References

Aharoni A, Giri AP, Verstappen FW, Bertea CM, Sevenier R, Sun Z, Jongsma MA, Schwab W, Bouwmeester HJ (2004) Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. The Plant Cell 16(11):3110-3131

Aitken A (1975) Prokaryote-eukaryote relationship and the amino acid sequence of plastocyanin from *Anabaena variabilis*. Biochemical Journal 149(3):675-683

Alagesan S, Gaudana SB, Krishnakumar S, Wangikar PP (2013) Model based optimization of high cell density cultivation of nitrogen-fixing cyanobacteria. Bioresource technology 148:228-233

Allen MM (1984) Cyanobacterial cell inclusions. Annual Reviews in Microbiology 38(1):1-25

Allen MM, Smith AJ (1969) Nitrogen chlorosis in blue-green algae. Archiv für Mikrobiologie 69(2):114-120

Alonso H, Blayney MJ, Beck JL, Whitney SM (2009) Substrate-induced Assembly of Methanococcoides burtonii D-Ribulose-1,5-bisphosphate Carboxylase/Oxygenase Dimers into Decamers. J Biol Chem 284(49):33876-33882

Andersen SO, Halberstadt ML, Borgford-Parnell N (2013) Stratospheric ozone, global warming, and the principle of unintended consequences-An ongoing science and policy success story. J Air Waste Manage 63(6):607-647

Anderson LM, Yang H (2008) DNA looping can enhance lysogenic CI transcription in phage lambda. Proceedings of the National Academy of Sciences 105(15):5827-5832

Andersson I, Backlund A (2008) Structure and function of RuBisCO. Plant Physiology and Biochemistry 46(3):275-291

Andrews TJ, Lorimer GH (1985) Catalytic properties of a hybrid between cyanobacterial large subunits and higher plant small subunits of ribulose bisphosphate carboxylase-oxygenase. J Biol Chem 260:4632-4636

Angermayr SA, Hellingwerf KJ, Lindblad P, de Mattos MJT (2009) Energy biotechnology with cyanobacteria. Current opinion in biotechnology 20(3):257-263

Antizar-Ladislao B, Turrion-Gomez JL (2008) Second-generation biofuels and local bioenergy systems. Biofuels, Bioproducts and Biorefining 2(5):455-469

Arnold FH, Volkov AA (1999) Directed evolution of biocatalysts. Current opinion in chemical biology 3(1):54-59

Atsumi S, Higashide W, Liao JC (2009) Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. Nature biotechnology 27(12):1177-1180

Azuma M, Osanai T, Hirai MY, Tanaka K (2011) A response regulator Rre37 and an RNA polymerase sigma factor SigE represent two parallel pathways to activate sugar catabolism in a cyanobacterium Synechocystis sp. PCC 6803. Plant and cell physiology 52(2):404-412

Badger MR (1980) Kinetic-properties of ribulose 1,5-bisphosphate carboxylase-oxygenase from *Anabaena variabilis*. Arch Biochem Biophys 201:247-254

Badger MR, Collatz GJ (1977) Studies on the kinetic mechanism of RuBP carboxylase and oxygenase reactions, with particular reference to the effect of temperature on kinetic papameters. Carnegie YB 76:355-361

Baier K, Lehmann H, Stephan DP, Lockau W (2004) NblA is essential for phycobilisome degradation in *Anabaena* sp strain PCC 7120 but not for development of functional heterocysts. Microbiol-Sgm 150:2739-2749

Bancroft I, Wolk C, Oren E (1989) Physical and genetic maps of the genome of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. Journal of bacteriology 171(11):5940-5948

Banerjee A, Sharkey TD (2014) Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. Nat Prod Rep 31(8):1043-55

Bar-Even A, Noor E, Lewis NE, Milo R (2010) Design and analysis of synthetic carbon fixation pathways. Proceedings of the National Academy of Sciences 107(19):8889-8894

Bassham JA, Benson AA, Kay LD, Harris AZ, Wilson AT, Calvin M (1954) The Path of Carbon in Photosynthesis. XXI. The Cyclic Regeneration of Carbon Dioxide Acceptor1. Journal of the American Chemical Society 76(7):1760-1770

Beißinger M, Sticht H, Sutter M, Ejchart A, Haehnel W, Rösch P (1998) Solution structure of cytochrome c6 from the thermophilic cyanobacterium Synechococcus elongatus. The EMBO journal 17(1):27-36

Bell LR, Horabin JI, Schedl P, Cline TW (1991) Positive autoregulation of Sex-lethal by alternative splicing maintains the female determined state in Drosophila. Cell 65(2):229-239

Bendall DS, Manasse RS (1995) Cyclic photophosphorylation and electron transport. Biochimica et Biophysica Acta (BBA)-Bioenergetics 1229(1):23-38

Bentley FK, García-Cerdán JG, Chen H-C, Melis A (2013) Paradigm of monoterpene (β-phellandrene) hydrocarbons production via photosynthesis in cyanobacteria. BioEnergy Research 6(3):917-929

Bentley FK, Melis A (2012) Diffusion-based process for carbon dioxide uptake and isoprene emission in gaseous/aqueous two-phase photobioreactors by photosynthetic microorganisms. Biotechnol Bioeng 109(1):100-109

Bentley FK, Zurbriggen A, Melis A (2014) Heterologous Expression of the Mevalonic Acid Pathway in Cyanobacteria Enhances Endogenous Carbon Partitioning to Isoprene. Mol Plant 7(1):71-86

Berg IA, Kockelkorn D, Buckel W, Fuchs G (2007) A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. Science 318(5857):1782-1786

Black TA, Cai YP, Wolk CP (1993) Spatial Expression and Autoregulation of Hetr, a Gene Involved in the Control of Heterocyst Development in *Anabaena* (Vol 9, Pg 77, 1993). Mol Microbiol 10(5):1153-1153

Black TA, Wolk CP (1994) Analysis of a Het-mutation in *Anabaena* sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. Journal of bacteriology 176(8):2282-2292

Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proceedings of the National Academy of Sciences of the United States of America 95(8):4126-4133

Borowitzka MA (2005) Cultivation of Microalgae from Village Level to Industrial Scale. Phycologia 44(4):11-12

Borowitzka MA, Moheimani NR (2013) Sustainable biofuels from algae. Mitigation and Adaptation Strategies for Global Change 18(1):13-25

Borthakur PB, Orozco CC, Young-Robbins SS, Haselkorn R, Callahan SM (2005) Inactivation of patS and hetN causes lethal levels of heterocyst differentiation in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. Mol Microbiol 57(1):111-123

Boyer C, Preiss J (1977) Biosynthesis of bacterial glycogen. Purification and properties of the Escherichia coli B α-1, 4-glucan: α-1, 4-glucan 6-glycosyltransferase. Biochemistry 16(16):3693-3699

Bradley S, Carr N (1976) Heterocyst and nitrogenase development in *Anabaena* cylindrica. Microbiology 96(1):175-184

Bradley S, Carr N (1977) Heterocyst development in *Anabaena* cylindrica: the necessity for light as an initial trigger and sequential stages of commitment. Microbiology 101(2):291-297

Brahamsha B, Haselkorn R (1992) Identification of multiple RNA polymerase sigma factor homologs in the cyanobacterium *Anabaena* sp. strain PCC 7120: cloning, expression, and inactivation of the sigB and sigC genes. Journal of bacteriology 174(22):7273-7282

Brecha RJ (2013) Ten Reasons to Take Peak Oil Seriously. Sustainability-Basel 5(2):664-694

Briggs LM, Pecoraro VL, McIntosh L (1990) Copper-induced expression, cloning, and regulatory studies of the plastocyanin gene from the cyanobacterium Synechocystis sp. PCC 6803. Plant molecular biology 15(4):633-642

Bryant DA (1994) The molecular biology of cyanobacteria. Kluwer Academic Publishers, Dordrecht ; Boston

Buchanan BB (1991) Regulation of CO 2 assimilation in oxygenic photosynthesis: the ferredoxin/thioredoxin system: perspective on its discovery, present status, and future development. Archives of Biochemistry and Biophysics 288(1):1-9

Buijs NA, Siewers V, Nielsen J (2013) Advanced biofuel production by the yeast Saccharomyces cerevisiae. Curr Opin Chem Biol 17(3):480-488

Buikema WJ, Haselkorn R (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. Genes & development 5(2):321-330

Buikema WJ, Haselkorn R (2001) Expression of the *Anabaena* hetR gene from a copper-regulated promoter leads to heterocyst differentiation under repressing conditions. Proceedings of the National Academy of Sciences 98(5):2729-2734

Busby S, Ebright RH (1997) Transcription activation at class II CAP-dependent promoters. Mol Microbiol 23(5):853-859

Busch F (2013) Current methods for estimating the rate of photorespiration in leaves. Plant Biology 15(4):648-655

Cai Y, Wolk CP (1997) *Anabaena* sp. strain PCC 7120 responds to nitrogen deprivation with a cascade-like sequence of transcriptional activations. Journal of bacteriology 179(1):267-271

Caldeira-Pires A, da Luz SM, Palma-Rojas S, Rodrigues TO, Silverio VC, Vilela F, Barbosa PC, Alves AM (2013) Sustainability of the Biorefinery Industry for Fuel Production. Energies 6(1):329-350

Callahan SM, Buikema WJ (2001) The role of HetN in maintenance of the heterocyst pattern in Anabaena sp. PCC 7120. Mol Microbiol 40(4):941-950

Campos AP, Aguir AP, Hervas M, Regalla M, Navarro JA, Ortega JM, Xavier AV, Rosa MA, Teixeira M (1993) Cytochrome c6 from Monoraphidium braunii. European journal of biochemistry 216(1):329-341

Cape JL, Bowman MK, Kramer DM (2006) Understanding the cytochrome bc complexes by what they don't do. The Q-cycle at 30. Trends in plant science 11(1):46-55

Carmo-Silva AE, Keys AJ, Andralojc PJ, Powers SJ, Arrabaca MC, Parry MAJ (2010) RuBisCO activities, properties, and regulation in three different C-4 grasses under drought. J Exp Bot 61(9):2355-2366

Carrieri D, Lombardi T, Paddock T, Cano M, Goodney GA, Nag A, Old W, Maness P-C, Seibert M, Ghirardi M (2017) Transcriptome and proteome analysis of nitrogen starvation responses in Synechocystis 6803 ΔglgC, a mutant incapable of glycogen storage. Algal Research 21:64-75

Carrieri D, Paddock T, Maness P-C, Seibert M, Yu J (2012) Photo-catalytic conversion of carbon dioxide to organic acids by a recombinant cyanobacterium incapable of glycogen storage. Energy & Environmental Science 5(11):9457-9461

Chappell J (1995) Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. Annual review of plant biology 46(1):521-547

Chater KF (1989) Multilevel regulation of Streptomyces differentiation. Trends Genet 5(11):372-7

Chaurasia AK, Parasnis A, Apte SK (2008) An integrative expression vector for strain improvement and environmental applications of the nitrogen fixing cyanobacterium, *Anabaena* sp. strain PCC7120. Journal of microbiological methods 73(2):133-141

Chen F, Tholl D, D'Auria JC, Farooq A, Pichersky E, Gershenzon J (2003) Biosynthesis and emission of terpenoid volatiles from Arabidopsis flowers. The Plant Cell 15(2):481-494

Chen K, Xu X, Gu L, Hildreth M, Zhou R (2015) Simultaneous gene inactivation and promoter reporting in cyanobacteria. Applied microbiology and biotechnology 99(4):1779-1793

Cherry JR, Fidantsef AL (2003) Directed evolution of industrial enzymes: an update. Current opinion in biotechnology 14(4):438-443

Chisti Y (2007) Biodiesel from microalgae. Biotechnology advances 25(3):294-306

Chisti Y (2008) Biodiesel from microalgae beats bioethanol. Trends in biotechnology 26(3):126-131

Cirino PC, Mayer KM, Umeno D (2003) Generating mutant libraries using error-prone PCR. Directed Evolution Library Creation: Methods and Protocols:3-9

Clarens AF, Resurreccion EP, White MA, Colosi LM (2010) Environmental life cycle comparison of algae to other bioenergy feedstocks. Environmental science & technology 44(5):1813-1819

Cohen JE (2003) Human population: the next half century. science 302(5648):1172-1175

Colowick SP, Kaplan NO, Packer L, Glazer AN (1988) Methods in Enzymology: Cyanobacteria / edited by Lester Packer, Alexander N. Glazer. Academic Press

Crock J, Wildung M, Croteau R (1997) Isolation and bacterial expression of a sesquiterpene synthase cDNA clone from peppermint (Mentha x piperita, L.) that produces the aphid alarm pheromone (E)-β-farnesene. Proceedings of the National Academy of Sciences 94(24):12833-12838

Crowell AL, Williams DC, Davis EM, Wildung MR, Croteau R (2002) Molecular cloning and characterization of a new linalool synthase. Archives of biochemistry and biophysics 405(1):112-121

Cumino AC, Marcozzi C, Barreiro R, Salerno GL (2007) Carbon cycling in *Anabaena* sp. PCC 7120. Sucrose synthesis in the heterocysts and possible role in nitrogen fixation. Plant physiology 143(3):1385-1397

Curatti L, Flores E, Salerno G (2002) Sucrose is involved in the diazotrophic metabolism of the heterocyst-forming cyanobacterium *Anabaena* sp. FEBS letters 513(2-3):175-178

Curatti L, Giarrocco LE, Cumino AC, Salerno GL (2008) Sucrose synthase is involved in the conversion of sucrose to polysaccharides in filamentous nitrogen-fixing cyanobacteria. Planta 228(4):617-625

Dalrymple OK, Halfhide T, Udom I, Gilles B, Wolan J, Zhang Q, Ergas S (2013) Wastewater use in algae production for generation of renewable resources: a review and preliminary results. Aquatic biosystems 9(1):1

Davies FK, Work VH, Beliaev AS, Posewitz MC (2014) Engineering Limonene and Bisabolene Production in Wild Type and a Glycogen-Deficient Mutant of Synechococcus sp. PCC 7002. Front Bioeng Biotechnol 2:21

De Philippis R, Sili C, Vincenzini M (1992) Glycogen and poly-β-hydroxybutyrate synthesis in Spirulina maxima. Microbiology 138(8):1623-1628

de Vries SC, van de Ven GWJ, van Ittersum MK, Giller KE (2010) Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion techniques. Biomass Bioenerg 34(5):588-601

Degenhardt J, Köllner TG, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70(15):1621-1637

Deng M-D, Coleman JR (1999) Ethanol synthesis by genetic engineering in cyanobacteria. Applied and environmental microbiology 65(2):523-528

Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux J-L, Buléon A, Haebel S, Ritte G, Steup M, Falcón LI (2008) Metabolic symbiosis and the birth of the plant kingdom. Molecular Biology and Evolution 25(3):536-548

Desplancq D, Bernard C, Sibler A, Kieffer B, Miguet L, Potier N, Van Dorsselaer A, Weiss E (2005) Combining inducible protein overexpression with NMR-grade triple isotope labeling in the cyanobacterium *Anabaena* sp. PCC 7120. Biotechniques 39(3):405

Dexter J, Fu P (2009) Metabolic engineering of cyanobacteria for ethanol production. Energy & Environmental Science 2(8):857-864

Dismukes GC, Carrieri D, Bennette N, Ananyev GM, Posewitz MC (2008) Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. Current opinion in biotechnology 19(3):235-240

Donald KAG, Hampton RY, Fritz IB (1997) Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in Saccharomyces cerevisiae. Applied and Environmental Microbiology 63(9):3341-3344

Doucha J, Straka F, Lívanský K (2005) Utilization of flue gas for cultivation of microalgae Chlorella sp.) in an outdoor open thin-layer photobioreactor. Journal of Applied Phycology 17(5):403-412

Ducat DC, Avelar-Rivas JA, Way JC, Silver PA (2012) Rerouting carbon flux to enhance photosynthetic productivity. Applied and environmental microbiology 78(8):2660-2668

Ducat DC, Way JC, Silver PA (2011) Engineering cyanobacteria to generate high-value products. Trends in biotechnology 29(2):95-103

Dudareva N, Cseke L, Blanc VM, Pichersky E (1996) Evolution of floral scent in Clarkia: novel patterns of S-linalool synthase gene expression in the C. breweri flower. The Plant Cell 8(7):1137-1148

Dutta D, De D, Chaudhuri S, Bhattacharya SK (2005) Hydrogen production by cyanobacteria. Microbial Cell Factories 4(1):1

Ehira S, Ohmori M (2006a) NrrA directly regulates expression of hetR during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. Journal of bacteriology 188(24):8520-8525

Ehira S, Ohmori M (2006b) NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol Microbiol 59(6):1692-1703

Ehira S, Ohmori M (2011) NrrA, a nitrogen-regulated response regulator protein, controls glycogen catabolism in the nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. J Biol Chem 286(44):38109-38114

Eidman VR (2007) Economic parameters for corn ethanol and biodiesel production. Journal of Agricultural and Applied Economics 39(02):345-356

Eisenhut M, Huege J, Schwarz D, Bauwe H, Kopka J, Hagemann M (2008a) Metabolome phenotyping of inorganic carbon limitation in cells of the wild type and photorespiratory mutants of the cyanobacterium Synechocystis sp. strain PCC 6803. Plant physiology 148(4):2109-2120

Eisenhut M, Ruth W, Haimovich M, Bauwe H, Kaplan A, Hagemann M (2008b) The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. Proceedings of the National Academy of Sciences 105(44):17199-17204

Eisenhut M, Von Wobeser EA, Jonas L, Schubert H, Ibelings BW, Bauwe H, Matthijs HC, Hagemann M (2007) Long-term response toward inorganic carbon limitation in wild type and glycolate turnover mutants of the cyanobacterium Synechocystis sp. strain PCC 6803. Plant physiology 144(4):1946-1959

Elhai J (1993) Strong and regulated promoters in the cyanobacterium *Anabaena* PCC 7120. Fems Microbiol Lett 114(2):179-184

Elhai J (1994) Genetic techniques appropriate for the biotechnological exploitation of cyanobacteria. Journal of Applied Phycology 6(2):177-186

Elhai J, Vepritskiy A, Muro-Pastor AM, Flores E, Wolk CP (1997) Reduction of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain PCC 7120. Journal of Bacteriology 179(6):1998-2005

Elhai J, Wolk CP (1988) [83] Conjugal transfer of DNA to cyanobacteria. Methods in enzymology 167:747-754

Elhai J, Wolk CP (1990) Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena*. EMBO J 9(10):3379-88

Ernst A, Black T, Cai Y, Panoff J-M, Tiwari D, Wolk C (1992) Synthesis of nitrogenase in mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst development or metabolism. Journal of bacteriology 174(19):6025-6032

Ernst A, Boger P (1985) Glycogen accumulation and the induction of nitrogenase activity in the heterocyst-forming cyanobacterium *Anabaena variabilis*. Microbiology 131(12):3147-3153

Ernst A, Kirschenlohr H, Diez J, Böger P (1984) Glycogen content and nitrogenase activity in *Anabaena variabilis*. Arch Microbiol 140(2):120-125

Errington J (1993) Bacillus-Subtilis Sporulation - Regulation of Gene-Expression and Control of Morphogenesis. Microbiol Rev 57(1):1-33

Ershov YV, Gantt RR, Cunningham FX, Gantt E (2002) Isoprenoid biosynthesis in Synechocystis sp strain PCC6803 is stimulated by compounds of the pentose phosphate cycle but not by pyruvate or deoxyxylulose-5-phosphate. Journal of Bacteriology 184(18):5045-5051

Espinosa J, Forchhammer K, Burillo S, Contreras A (2006) Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. Mol Microbiol 61(2):457-469

Evans JR (1989) The allocation of protein nitrogen in the photosynthetic apparatus: costs, consequences and control. Photosynthesis

Evans M, Buchanan BB, Arnon DI (1966) A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium. Proceedings of the National Academy of Sciences 55(4):928-934

Ferreira KN, Iverson TM, Maghlaoui K, Barber J, Iwata S (2004) Architecture of the photosynthetic oxygen-evolving center. Science 303(5665):1831-1838

Fiedler G, Muro-Pastor AM, Flores E, Maldener I (2001) NtcA-dependent expression of the devBCA operon, encoding a heterocyst-specific ATP-binding Cassette Transporter in *Anabaena* spp. Journal of Bacteriology 183(12):3795-3799

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281(5374):237-240

Flores E, Herrero A (2005) Nitrogen assimilation and nitrogen control in cyanobacteria. Biochem Soc Trans 33(Pt 1):164-167

Flores E, Herrero A (2014) The cell biology of cyanobacteria. Caister Academic Press, Norfolk, UK

Flores FG, Herrero A (2008) The cyanobacteria: molecular biology, genomics, and evolution. Horizon Scientific Press

Fogg G (1942) Studies on nitrogen fixation by blue-green algae. Nitrogen fixation

Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, Chapin FS, Coe MT, Daily GC, Gibbs HK (2005) Global consequences of land use. Science 309(5734):570-574

Forchhammer K (2004) Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. FEMS microbiology reviews 28(3):319-333

Forchhammer K, Hedler A (1997) Phosphoprotein PII from cyanobacteria--analysis of functional conservation with the PII signal-transduction protein from Escherichia coli. Eur J Biochem 244(3):869-75

Frazao C, Soares C, Carrondo M, Pohl E, Dauter Z, Wilson K, Hervés M, Navarro J, De la Rosa M, Sheldrick G (1995) Ab initio determination of the crystal structure of cytochrome c 6 and comparison with plastocyanin. Structure 3(11):1159-1169

Fritsch FE PRESIDENTIAL ADDRESS: The Heterocyst: A Botanical Enigma. In: Proceedings of the Linnean Society of London, 1951. vol 162. Wiley Online Library, p 194-211

Fu J, Xu X (2006) The functional divergence of two glgP homologues in Synechocystis sp. PCC 6803. Fems Microbiol Lett 260(2):201-209

Fuentes MMR, Sanchez JLG, Sevilla JMF, Fernandez FGA, Perez JAS, Grima EM (1999) Outdoor continuous culture of Porphyridium cruentum in a tubular photobioreactor: quantitative analysis of the daily cyclic variation of culture parameters. J Biotechnol 70(1-3):271-288

Gatenby AA, van der Vies SM, Bradley D (1985) Assembly in E. coli of a functional multi-subunit ribulose bisphosphate carboxylase from a blue-green alga. Nature 314(6012):617-620

Geider RJ, Delucia EH, Falkowski PG, Finzi AC, Grime JP, Grace J, Kana TM, La Roche J, Long SP, Osborne BA (2001) Primary productivity of planet earth: biological determinants and physical constraints in terrestrial and aquatic habitats. Global Change Biology 7(8):849-882

Genkov T, Meyer M, Griffiths H, Spreitzer RJ (2010a) Functional hybrid RuBisCO enzymes with plant small subunits and algal large subunits engineered rbcS cDNA for expression in Chlamydomonas. J Biol Chem 285(26):19833-19841

Gierer A, Meinhardt H (1972) A theory of biological pattern formation. Kybernetik 12(1):30-39

Giordano M, Beardall J, Raven JA (2005) CO 2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. Annu Rev Plant Biol 56:99-131

Golden JW, Yoon H-S (2003) Heterocyst development in *Anabaena*. Current opinion in microbiology 6(6):557-563

Golden SS (1995) Light-Responsive Gene-Expression in Cyanobacteria. J Bacteriol 177(7):1651-1654

Golden SS, Brusslan J, Haselkorn R (1987) [12] Genetic engineering of the cyanobacterial chromosome. Methods in enzymology 153:215-231

Gomez LD, Steele-King CG, McQueen-Mason SJ (2008) Sustainable liquid biofuels from biomass: the writing's on the walls. New Phytologist 178(3):473-485

Graf A, Schlereth A, Stitt M, Smith AM (2010) Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. Proceedings of the National Academy of Sciences 107(20):9458-9463

Green FB, Bernstone LS, Lundquist TJ, Oswald WJ (1996) Advanced integrated wastewater pond systems for nitrogen removal. Water Sci Technol 33(7):207-217

Greene DN, Whitney SM, Matsumura I (2007) Artificially evolved Synechococcus PCC6301 RuBisCO variants exhibit improvements in folding and catalytic efficiency. Biochemical Journal 404(3):517-524

Gronenberg LS, Marcheschi RJ, Liao JC (2013) Next generation biofuel engineering in prokaryotes. Curr Opin Chem Biol 17(3):462-471

Gross EL (1993) Plastocyanin: structure and function. Photosynthesis research 37(2):103-116

Gründel M, Scheunemann R, Lockau W, Zilliges Y (2012) Impaired glycogen synthesis causes metabolic overflow reactions and affects stress responses in the cyanobacterium Synechocystis sp. PCC 6803. Microbiology 158(12):3032-3043

Gupta R, He Z, Luan S (2002) Functional relationship of cytochrome c6 and plastocyanin in Arabidopsis. Nature 417(6888):567-571

Guskov A, Kern J, Gabdulkhakov A, Broser M, Zouni A, Saenger W (2009) Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. Nature structural & molecular biology 16(3):334-342

Gutteridge S, Gatenby AA (1995) RuBisCO synthesis, assembly, mechanism, and regulation. The Plant Cell 7(7):809

Hagemann M, Erdmann N (1994) Activation and pathway of glucosylglycerol synthesis in the cyanobacterium Synechocystis sp. PCC 6803. Microbiology 140(6):1427-1431

Halfmann C, Gu L, Gibbons W, Zhou R (2014a) Genetically engineering cyanobacteria to convert CO2, water, and light into the long-chain hydrocarbon farnesene. Applied microbiology and biotechnology 98(23):9869-9877

Halfmann C, Gu LP, Zhou RB (2014b) Engineering cyanobacteria for the production of a cyclic hydrocarbon fuel from CO2 and H2O. Green Chem 16(6):3175-3185

Harrewijn P (2001) Natural terpenoids as messengers: a multidisciplinary study of their production, biological functions, and practical applications. Springer Science & Business Media

Haselkorn R (1992) Nitrogen Fixation in Cyanobacteria Robert Haselkorn and William J. Buikema. Biological nitrogen fixation:1166

Hasunuma T, Kikuyama F, Matsuda M, Aikawa S, Izumi Y, Kondo A (2013) Dynamic metabolic profiling of cyanobacterial glycogen biosynthesis under conditions of nitrate depletion. Journal of experimental botany:ert134

Heidorn T, Camsund D, Huang H-H, Lindberg P, Oliveira P, Stensjö K, Lindblad P (2010) Synthetic biology in cyanobacteria engineering and analyzing novel functions. Methods in enzymology 497:539-579

Herrero A, Muro-Pastor AM, Flores E (2001) Nitrogen control in cyanobacteria. Journal of Bacteriology 183(2):411-425

Herter S, Fuchs G, Bacher A, Eisenreich W (2002) A Bicyclic Autotrophic CO2 Fixation Pathway inChloroflexus aurantiacus. J Biol Chem 277(23):20277-20283

Hervás M, Ortega J, Navarro J, Miguel A, Bottin H (1994) Laser flash kinetic analysis of Synechocystis PCC 6803 cytochrome c6 and plastocyanin oxidation by photosystem I. Biochimica et Biophysica Acta (BBA)-Bioenergetics 1184(2-3):235-241

Heyer H, Krumbein W (1991) Excretion of fermentation products in dark and anaerobically incubated cyanobacteria. Arch Microbiol 155(3):284-287

Hickman JW, Kotovic KM, Miller C, Warrener P, Kaiser B, Jurista T, Budde M, Cross F, Roberts JM, Carleton M (2013) Glycogen synthesis is a required component of the nitrogen stress response in Synechococcus elongatus PCC 7942. Algal Research 2(2):98-106

Hill J, Nelson E, Tilman D, Polasky S, Tiffany D (2006) Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. Proceedings of the National Academy of sciences 103(30):11206-11210

Holtman CK, Chen Y, Sandoval P, Gonzales A, Nalty MS, Thomas TL, Youderian P, Golden SS (2005) High-throughput functional analysis of the Synechococcus elongatus PCC 7942 genome. DNA research 12(2):103-115

Hook M, Tang X (2013) Depletion of fossil fuels and anthropogenic climate change-A review. Energ Policy 52:797-809

Horswill AR, Dudding AR, Escalante-Semerena JC (2001) Studies of propionate toxicity in Salmonella enterica identify 2-methylcitrate as a potent inhibitor of cell growth. J Biol Chem 276(22):19094-19101

Hu Q, Sommerfeld M, Jarvis E, Ghirardi M, Posewitz M, Seibert M, Darzins A (2008) Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. The Plant journal : for cell and molecular biology 54(4):621-39

Huang HH, Camsund D, Lindblad P, Heidorn T (2010) Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. Nucleic Acids Res 38(8):2577-2593

Huang HH, Lindblad P (2013) Wide-dynamic-range promoters engineered for cyanobacteria. J Biol Eng 7(1)

Huang X, Dong Y, Zhao J (2004) HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. Proceedings of the National Academy of Sciences of the United States of America 101(14):4848-4853

Huber DP, Philippe RN, Godard K-A, Sturrock RN, Bohlmann J (2005) Characterization of four terpene synthase cDNAs from methyl jasmonate-induced Douglas-fir, Pseudotsuga menziesii. Phytochemistry 66(12):1427-1439

Huber H, Gallenberger M, Jahn U, Eylert E, Berg IA, Kockelkorn D, Eisenreich W, Fuchs G (2008) A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum Ignicoccus hospitalis. Proceedings of the National Academy of Sciences 105(22):7851-7856

Iglesias AA, Charng Y, Ball S, Preiss J (1994) Characterization of the kinetic, regulatory, and structural properties of ADP-glucose pyrophosphorylase from Chlamydomonas reinhardtii. Plant physiology 104(4):1287-1294

Iglesias AA, Kakefuda G, Preiss J (1991) Regulatory and structural properties of the cyanobacterial ADPglucose pyrophosphorylases. Plant physiology 97(3):1187-1195

Iglesias AA, Preiss J (1992) Bacterial glycogen and plant starch biosynthesis. Biochemical education 20(4):196-203

Iijima Y, Gang DR, Fridman E, Lewinsohn E, Pichersky E (2004) Characterization of geraniol synthase from the peltate glands of sweet basil. Plant Physiology 134(1):370-379

Inoue T, Sugawara H, Hamanaka S, Tsukui H, Suzuki E, Kohzuma T, Kai Y (1999) Crystal structure determinations of oxidized and reduced plastocyanin from the cyanobacterium Synechococcus sp. PCC 7942. Biochemistry 38(19):6063-6069

Ivanovsky RN, Krasilnikova EN, Fal YI (1993) A pathway of the autotrophic $CO_2$ fixation in Chloroflexus aurantiacus. Arch Microbiol 159(3):257-264

Jackson SA, Eaton-Rye JJ, Bryant DA, Posewitz MC, Davies FK (2015) Dynamics of Photosynthesis in a Glycogen-Deficient glgC Mutant of Synechococcus sp. Strain PCC 7002. Applied and environmental microbiology 81(18):6210-6222

Jia J-W, Crock J, Lu S, Croteau R, Chen X-Y (1999) (3R)-Linalool synthase from Artemisia annua L.: cDNA isolation, characterization, and wound induction. Archives of biochemistry and biophysics 372(1):143-149

Johnson TJ, Jahandideh A, Johnson MD, Fields KH, Richardson JW, Muthukumarappan K, Cao Y, Gu Z, Halfmann C, Zhou R (2016) Producing next-generation biofuels from filamentous cyanobacteria: An economic feasibility analysis. Algal Research 20:218-228

Jordan DB, Chollet R (1985) Subunit dissociation and reconstitution of ribulose-1,5-bisphosphate carboxylase from *Chromatium vinosum*. Arch Biochem Biophys 236:487-496

Jordan DB, Ogren WL (1981) Species variation in the specificity of ribulose bisphosphate carboxylase/oxygenase. Nature 291:513-515

Kamiya N, Shen J-R (2003) Crystal structure of oxygen-evolving photosystem II from Thermosynechococcus vulcanus at 3.7-Å resolution. Proceedings of the National Academy of Sciences 100(1):98-103

Kane HJ, Viil J, Entsch B, Paul K, Morell MK, Andrews TJ (1994) An improved method for measuring the $CO_2/O_2$ specificity of ribulosebisphosphate carboxylase-oxygenase. Aust J Plant Physiol 21:449-461

Kaneko T, Nakamura Y, Wolk C, Kuritz T, Sasamoto S, Kaneko T (2001) Complete genomic sequence of the filamentous nitrogen-fixing. DNA Res 8(5):205-213

Karradt A, Sobanski J, Mattow J, Lockau W, Baier K (2008) NblA, a key protein of phycobilisome degradation, interacts with ClpC, a HSP100 chaperone partner of a cyanobacterial Clp protease. J Biol Chem 283(47):32394-403

Kawano Y, Sekine M, Ihara M (2014) Identification and characterization of UDP-glucose pyrophosphorylase in cyanobacteria *Anabaena* sp. PCC 7120. Journal of bioscience and bioengineering 117(5):531-538

Knoop H, Grundel M, Zilliges Y, Lehmann R, Hoffmann S, Lockau W, Steuer R (2013) Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of Synechocystis sp PCC 6803. Plos Computational Biology 9(6)

Knoop H, Zilliges Y, Lockau W, Steuer R (2010) The metabolic network of Synechocystis sp. PCC 6803: systemic properties of autotrophic growth. Plant physiology 154(1):410-422

Koksharova O, Brandt U, Cerff R (2004) The gap1 Operon of the Cyanobacterium SynechococcusPCC 7942 Carries a Gene Encoding Glycogen Phosphorylase and Is Induced under Anaerobic Conditions. Microbiology 73(3):326329

Koksharova O, Wolk C (2002) Genetic tools for cyanobacteria. Applied microbiology and biotechnology 58(2):123-137

Koßmann J, Sonnewald U, Willmitzer L (1994) Reduction of the chloroplastic fructose-1, 6-bisphosphatase in transgenic potato plants impairs photosynthesis and plant growth. The Plant Journal 6(5):637-650

Kroos L, Kaiser D (1987) Expression of many developmentally regulated genes in Myxococcus depends on a sequence of cell interactions. Genes Dev 1(8):840-54

Kubien DS, Whitney SM, Moore PV, Jesson LK (2008) The biochemistry of RuBisCO in *Flaveria*. J Exp Bot 59(7):1767-1777

Kurisu G, Zhang H, Smith JL, Cramer WA (2003) Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. Science 302(5647):1009-1014

Lan EI, Liao JC (2012) ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. Proceedings of the National Academy of Sciences 109(16):6018-6023

Landmann C, Fink B, Festner M, Dregus M, Engel K-H, Schwab W (2007) Cloning and functional characterization of three terpene synthases from lavender (Lavandula angustifolia). Archives of biochemistry and biophysics 465(2):417-429

Lang JD, Haselkorn R (1989) Isolation, sequence and transcription of the gene encoding the photosystem II chlorophyll-binding protein, CP-47, in the cyanobacterium *Anabaena* 7120. Plant molecular biology 13(4):441-456

Larimer FW, Soper TS (1993) Overproduction of *Anabaena* 7120 ribulose-bisphosphate carboxylase/oxygenase in Escherichia coli. Gene 126(1):85-92

Lazazzera BA, Solomon JM, Grossman AD (1997) An exported peptide functions intracellularly to contribute to cell density signaling in B. subtilis. Cell 89(6):917-925

Lechno-Yossef S, Fan Q, Wojciuch E, Wolk CP (2011) Identification of ten *Anabaena* sp. genes that, under aerobic conditions, are required for growth on dinitrogen but not for growth on fixed nitrogen. Journal of bacteriology:JB. 05010-11

Lehmann M, Wöber G (1976) Accumulation, mobilization and turn-over of glycogen in the blue-green bacterium Anacystis nidulans. Arch Microbiol 111(1-2):93-97

Li B, Huang X, Zhao J (2002) Expression of hetN during heterocyst differentiation and its inhibition of hetR up-regulation in the cyanobacterium *Anabaena* sp. PCC 7120. FEBS letters 517(1-3):87-91

Li J-H, Laurent S, Konde V, Bedu S, Zhang C-C (2003) An increase in the level of 2-oxoglutarate promotes heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. Microbiology 149(11):3257-3263

Li L-A, Tabita FR (1997) Maximum activity of recombinant ribulose 1, 5-bisphosphate carboxylase/oxygenase of *Anabaena* sp. strain CA requires the product of the rbcX gene. Journal of bacteriology 179(11):3793-3796

Li X, Shen CR, Liao JC (2014) Isobutanol production as an alternative metabolic sink to rescue the growth deficiency of the glycogen mutant of Synechococcus elongatus PCC 7942. Photosynthesis research 120(3):301-310

Liang J, Scappino L, Haselkorn R (1992) The patA gene product, which contains a region similar to CheY of Escherichia coli, controls heterocyst pattern formation in the cyanobacterium *Anabaena* 7120. Proceedings of the National Academy of Sciences 89(12):5655-5659

Lichtenthaler HK (1999) The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. Annual review of plant biology 50(1):47-65

Lindahl M, Florencio FJ (2003) Thioredoxin-linked processes in cyanobacteria are as numerous as in chloroplasts, but targets are different. Proceedings of the National Academy of Sciences 100(26):16107-16112

Lindberg P, Park S, Melis A (2010) Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism. Metab Eng 12(1):70-79

Lipscomb TW, Lipscomb ML, Gill RT, Lynch MD (2012) Metabolic Engineering of Recombinant E. coli for the Production of 3-Hydroxypropionate. Engineering Complex Phenotypes in Industrial Strains:185-200

Liu X, Sheng J, Curtiss III R (2011) Fatty acid production in genetically modified cyanobacteria. Proceedings of the National Academy of Sciences 108(17):6899-6904

Loll B, Kern J, Saenger W, Zouni A, Biesiadka J (2005a) Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. Nature 438(7070):1040-1044

Loll B, Kern J, Zouni A, Saenger W, Biesiadka J, Irrgang K-D (2005b) The antenna system of photosystem II from Thermosynechococcus elongatus at 3.2 Å resolution. Photosynthesis research 86(1-2):175-184

Makarova KS, Koonin EV, Haselkorn R, Galperin MY (2006) Cyanobacterial response regulator PatA contains a conserved N-terminal domain (PATAN) with an alpha-helical insertion. Bioinformatics 22(11):1297-1301

Mann MS, Lutke-Eversloh T (2013) Thiolase engineering for enhanced butanol production in Clostridium acetobutylicum. Biotechnol Bioeng 110(3):887-97

Mannan RM, Pakrasi HB (1993) Dark heterotrophic growth conditions result in an increase in the content of photosystem II units in the filamentous cyanobacterium *Anabaena variabilis* ATCC 29413. Plant physiology 103(3):971-977

Martin DM, Fäldt J, Bohlmann J (2004) Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. Plant Physiology 135(4):1908-1927

Martin M, Grossmann IE (2013) On the Systematic Synthesis of Sustainable Biorefineries. Ind Eng Chem Res 52(9):3044-3064

Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nature biotechnology 21(7):796-802

Maruyama T, Michiho I, Honda G (2001) Molecular cloning, functional expression and characterization of (E)-β-farnesene synthase from Citrus junos. Biological and Pharmaceutical Bulletin 24(10):1171-1175

Mattozzi M, Ziesack M, Voges MJ, Silver PA, Way JC (2013) Expression of the sub-pathways of the Chloroflexus aurantiacus 3-hydroxypropionate carbon fixation bicycle in E. coli: Toward horizontal transfer of autotrophic growth. Metabolic engineering 16:130-139

McGinn PJ, Dickinson KE, Bhatti S, Frigon J-C, Guiot SR, O'Leary SJ (2011) Integration of microalgae cultivation with industrial waste remediation for biofuel and bioenergy production: opportunities and limitations. Photosynthesis research 109(1-3):231-247

McNeely K, Xu Y, Bennette N, Bryant DA, Dismukes GC (2010) Redirecting reductant flux into hydrogen production via metabolic engineering of fermentative carbon metabolism in a cyanobacterium. Applied and environmental microbiology 76(15):5032-5038

Medina M, Louro RO, Gagnon J, Peleato ML, Mendes J, Gómez-Moreno C, Xavier AV, Teixeira M (1997) Characterization of cytochrome c 6 from the cyanobacterium *Anabaena* PCC 7119. JBIC Journal of Biological Inorganic Chemistry 2(2):225-234

Meeks J, Castenholz R (1971) Growth and photosynthesis in an extreme thermophile, Synechococcus lividus (Cyanophyta). Archiv Mikrobiol 78(1):25-41

Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P, Atlas R (2001) An overview of the genome of Nostoc punctiforme, a multicellular, symbiotic cyanobacterium. Photosynthesis research 70(1):85-106

Melis A (2009) Solar energy conversion efficiencies in photosynthesis: minimizing the chlorophyll antennae to maximize efficiency. Plant science 177(4):272-280

Memon D, Singh AK, Pakrasi HB, Wangikar PP (2013) A global analysis of adaptive evolution of operons in cyanobacteria. Antonie Van Leeuwenhoek 103(2):331-346

Menetrez MY (2012) An overview of algae biofuel production and potential environmental impact. Environmental science & technology 46(13):7073-7085

Merchant S, Bogorad L (1986) Regulation by copper of the expression of plastocyanin and cytochrome c552 in Chlamydomonas reinhardi. Molecular and cellular biology 6(2):462-469

Mercke P, Kappers IF, Verstappen FW, Vorst O, Dicke M, Bouwmeester HJ (2004) Combined transcript and metabolite analysis reveals genes involved in spider mite induced volatile formation in cucumber plants. Plant Physiology 135(4):2012-2024

Miao X, Wu Q, Wu G, Zhao N (2003a) Changes in photosynthesis and pigmentation in an agp deletion mutant of the cyanobacterium Synechocystis sp. Biotechnology letters 25(5):391-396

Miao X, Wu Q, Wu G, Zhao N (2003b) Sucrose accumulation in salt-stressed cells of agp gene deletion-mutant in cyanobacterium Synechocystis sp. PCC 6803. Fems Microbiol Lett 218(1):71-77

Mitchell P (1975) The protonmotive Q cycle: a general formulation. FEBS letters 59(2):137-139

Miyagawa Y, Tamoi M, Shigeoka S (2001) Overexpression of a cyanobacterial fructose-1, 6-/sedoheptulose-1, 7-bisphosphatase in tobacco enhances photosynthesis and growth. Nature biotechnology 19(10):965-969

Molina E, Fernández J, Acién F, Chisti Y (2001) Tubular photobioreactor design for algal cultures. J Biotechnol 92(2):113-131

Montgomery DR (2007) Soil erosion and agricultural sustainability. Proceedings of the National Academy of Sciences 104(33):13268-13272

Morell MK, Kane HJ, Andrews TJ (1990) Carboxylterminal deletion mutants of ribulosebisphosphate carboxylase from *Rhodospirillum rubrum*. FEBS Lett 265:41-45

Mueller-Cajar O, Whitney SM (2008a) Directing the evolution of RuBisCO and RuBisCO activase: first impressions of a new tool for photosynthesis research. Photosynthesis Research 98(1-3):667-675

Mueller-Cajar O, Whitney SM (2008b) Evolving improved *Synechococcus* RuBisCO functional expression in *Escherichia coli.* Biochem J 414(2):205-14

Muñoz-García J, Ares S (2016) Formation and maintenance of nitrogen-fixing cell patterns in filamentous cyanobacteria. Proceedings of the National Academy of Sciences 113(22):6218-6223

Muro-Pastor AM, Hess WR (2012) Heterocyst differentiation: from single mutants to global approaches. Trends in microbiology 20(11):548-557

Muro-Pastor AM, Olmedo-Verd E, Flores E (2006) All4312, an NtcA-regulated two-component response regulator in *Anabaena* sp. strain PCC 7120. Fems Microbiol Lett 256(1):171-177

Muro-Pastor AM, Valladares A, Flores E, Herrero A (1999) The hetC gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development. Journal of Bacteriology 181(21):6664-6669

Muro-Pastor MI, Reyes JC, Florencio FJ (2001) Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. J Biol Chem 276(41):38320-38328

Nagegowda DA, Gutensohn M, Wilkerson CG, Dudareva N (2008) Two nearly identical terpene synthases catalyze the formation of nerolidol and linalool in snapdragon flowers. The Plant Journal 55(2):224-239

Nakao M, Okamoto S, Kohara M, Fujishiro T, Fujisawa T, Sato S, Tabata S, Kaneko T, Nakamura Y (2009) CyanoBase: the cyanobacteria genome database update 2010. Nucleic Acids Res:gkp915

Navarro JA, Hervás M, De la Rosa M (1997) Co-evolution of cytochrome c 6 and plastocyanin, mobile proteins transferring electrons from cytochrome b 6f to photosystem I. JBIC Journal of Biological Inorganic Chemistry 2(1):11-22

Nguyen AY, Bricker WP, Zhang H, Weisz DA, Gross ML, Pakrasi HB (2017) The proteolysis adaptor, NblA, binds to the N-terminus of β-phycocyanin: Implications for the mechanism of phycobilisome degradation. Photosynthesis Research:1-12

Niederholtmeyer H, Wolfstädter BT, Savage DF, Silver PA, Way JC (2010) Engineering cyanobacteria to synthesize and export hydrophilic products. Applied and environmental microbiology 76(11):3462-3466

Nigam PS, Singh A (2011) Production of liquid biofuels from renewable resources. Prog Energ Combust 37(1):52-68

Ninfa AJ, Jiang P, Atkinson MR, Peliska JA (2000) Integration of antagonistic signals in the regulation of nitrogen assimilation in Escherichia coli. Curr Top Cell Regul 36:31

Normile D (2009) Round and round: a guide to the carbon cycle. American Association for the Advancement of Science

Norsker N-H, Barbosa MJ, Vermuë MH, Wijffels RH (2011) Microalgal production—a close look at the economics. Biotechnology advances 29(1):24-27

Nurdogan Y, Oswald WJ (1996) Tube settling of high-rate pond algae. Water Sci Technol 33(7):229-241

Odling-Smee L (2007) Biofuels bandwagon hits a rut. Nature 446(7135):483

Okada K, Hase T (2005) Cyanobacterial non-mevalonate pathway: (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase interacts with ferredoxin in Thermosynechococcus elongatus BP-1. J Biol Chem 280(21):20672-9

Olaizola M (2000) Commercial production of astaxanthin from Haematococcus pluvialis using 25,000-liter outdoor photobioreactors. Journal of Applied Phycology 12(3-5):499-506

Oliver JW, Machado IM, Yoneda H, Atsumi S (2013) Cyanobacterial conversion of carbon dioxide to 2, 3-butanediol. Proceedings of the National Academy of Sciences 110(4):1249-1254

Oliver JWK, Machado IMP, Yoneda H, Atsumi S (2014) Combinatorial optimization of cyanobacterial 2,3-butanediol production. Metab Eng 22:76-82

Orozco CC, Risser DD, Callahan SM (2006) Epistasis analysis of four genes from *Anabaena* sp. strain PCC 7120 suggests a connection between PatA and PatS in heterocyst pattern formation. Journal of bacteriology 188(5):1808-1816

Osanai T, Azuma M, Tanaka K (2007) Sugar catabolism regulated by light-and nitrogen-status in the cyanobacterium Synechocystis sp. PCC 6803. Photochemical & Photobiological Sciences 6(5):508-514

Osanai T, Kanesaki Y, Nakano T, Takahashi H, Asayama M, Shirai M, Kanehisa M, Suzuki I, Murata N, Tanaka K (2005) Positive regulation of sugar catabolic pathways in the cyanobacterium Synechocystis sp. PCC 6803 by the group 2 σ factor SigE. J Biol Chem 280(35):30653-30659

Paradise EM, Kirby J, Chan R, Keasling JD (2008) Redirection of flux through the FPP branch-point in Saccharomyces cerevisiae by down-regulating squalene synthase. Biotechnol Bioeng 100(2):371-378

Parmar A, Singh NK, Pandey A, Gnansounou E, Madamwar D (2011) Cyanobacteria and microalgae: a positive prospect for biofuels. Bioresour Technol 102(22):10163-72

Parry MAJ, Madgwick PJ, Carvalho JFC, Andralojc PJ (2007) Prospects for increasing photosynthesis by overcoming the limitations of RuBisCO. J Ag Sci 145(Part 1):31-43

Peccia J, Haznedaroglu B, Gutierrez J, Zimmerman JB (2013) Nitrogen supply is an important driver of sustainable microalgae biofuel production. Trends in biotechnology 31(3):134-138

Pechous SW, Whitaker BD (2004) Cloning and functional expression of an (E, E)-α-farnesene synthase cDNA from peel tissue of apple fruit. Planta 219(1):84-94

Peralta-Yahya PP, Keasling JD (2010) Advanced biofuel production in microbes. Biotechnol J 5(2):147-162

Peralta-Yahya PP, Zhang F, del Cardayre SB, Keasling JD (2012) Microbial engineering for the production of advanced biofuels. Nature 488(7411):320-328

Perego M, Hoch JA (1996) Cell-cell communication regulates the effects of protein aspartate phosphatases on the phosphorelay controlling development in Bacillus subtilis. Proceedings of the National Academy of Sciences 93(4):1549-1553

Perez-Gil J, Rodriguez-Concepcion M (2013) Metabolic plasticity for isoprenoid biosynthesis in bacteria. Biochemical Journal 452:19-25

Phillips MA, Wildung MR, Williams DC, Hyatt DC, Croteau R (2003) cDNA isolation, functional expression, and characterization of (+)-α-pinene synthase and (−)-α-pinene synthase from loblolly pine (Pinus taeda): Stereocontrol in pinene biosynthesis. Archives of Biochemistry and Biophysics 411(2):267-276

Picaud S, Brodelius M, Brodelius PE (2005) Expression, purification and characterization of recombinant (E)-β-farnesene synthase from Artemisia annua. Phytochemistry 66(9):961-967

Pitera DJ, Paddon CJ, Newman JD, Keasling JD (2007) Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. Metab Eng 9(2):193-207

Poliquin K, Ershov YV, Cunningham FX, Woreta TT, Gantt RR, Gantt E (2004) Inactivation of sll1556 in Synechocystis strain PCC 6803 impairs isoprenoid biosynthesis from pentose phosphate cycle substrates in vitro. Journal of Bacteriology 186(14):4685-4693

Qi FX, Yao L, Tan XM, Lu XF (2013) Construction, characterization and application of molecular tools for metabolic engineering of Synechocystis sp. Biotechnol Lett 35(10):1655-1661

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(3):471-490

Ragsdale SW, Pierce E (2008) Acetogenesis and the Wood–Ljungdahl pathway of CO 2 fixation. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1784(12):1873-1898

Ramasubramanian TS, Wei TF, Oldham AK, Golden JW (1996) Transcription of the *Anabaena* sp strain PCC 7120 ntcA gene: Multiple transcripts and NtcA binding. Journal of Bacteriology 178(3):922-926

Rappaport F, Guergova-Kuras M, Nixon PJ, Diner BA, Lavergne J (2002) Kinetics and pathways of charge recombination in photosystem II. Biochemistry 41(26):8518-8527

Raven JA (2009) Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. Aquatic Microbial Ecology 56(2-3):177-192

Rawat M, Henk MC, Lavigne LL, Moroney JV (1996) Chlamydomonas reinhardtii mutants without ribulose-1, 5-bisphosphate carboxylase-oxygenase lack a detectable pyrenoid. Planta 198(2):263-270

Read BA, Tabita FR (1994) High substrate specificity factor ribulose bisphosphate carboxylase oxygenase from eukaryotic marine algae and properties of recombinant cyanobacterial RuBisCO containing algal residue modifications. Arch Biochem Biophys 312(1):210-218

Redding-Johanson AM, Batth TS, Chan R, Krupa R, Szmidt HL, Adams PD, Keasling JD, Lee TS, Mukhopadhyay A, Petzold CJ (2011) Targeted proteomics for metabolic pathway optimization: application to terpene production. Metab Eng 13(2):194-203

Reichert CdC, Reinehr CO, Costa JAV (2006) Semicontinuous cultivation of the cyanobacterium Spirulina platensis in a closed photobioreactor. Brazilian Journal of Chemical Engineering 23(1):23-28

Rice D, Mazur B, Haselkorn R (1982) Isolation and physical mapping of nitrogen fixation genes from the cyanobacterium *Anabaena* 7120. J Biol Chem 257(21):13157-13163

Robinson JJ, Scott KM, Swanson ST, O'Leary MH, Horken K, Tabita FR, Cavanaugh CM (2003) Kinetic isotope effect and characterization of form II RuBisCO from the chemoautotrophic endosymbionts of the hydrothermal vent tubeworm *Riftia pachyptila*. Limnol and Oceanog 48(1):48-54

Rohmer M (1999) The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants†. Natural product reports 16(5):565-574

Roy H (1989) RuBisCO assembly: a model system for studying the mechanism of chaperonin action. The Plant Cell 1(11):1035

Rude MA, Schirmer A (2009) New microbial fuels: a biotech perspective. Current opinion in microbiology 12(3):274-281

Ruffing AM (2011) Engineered cyanobacteria: teaching an old bug new tricks. Bioengineered bugs 2(3):136-49

Ruffing AM, Jones HD (2012) Physiological effects of free fatty acid production in genetically engineered Synechococcus elongatus PCC 7942. Biotechnol Bioeng 109(9):2190-2199

Sage RF, Seemann JR (1993) Regulation of rribulose-1,5-bisphosphate carboxylase/oxygenase activity in response to reduced light intensity in C4 plants. Plant Physiol 102(1):21-28

Saito Y, Ashida H, Sakiyama T, de Marsac NT, Danchin A, Sekowska A, Yokota A (2009) Structural and functional similarities between a ribulose-1, 5-bisphosphate carboxylase/oxygenase (RuBisCO)-like protein from Bacillus subtilis and photosynthetic RuBisCO. J Biol Chem 284(19):13256-13264

Sakurai H, Masukawa H (2007) Promoting R & D in photobiological hydrogen
    production utilizing mariculture-raised cyanobacteria. Marine Biotechnology
    9(2):128-145

Sakurai H, Masukawa H, Kitashima M, Inoue K (2013) Photobiological hydrogen
    production: Bioenergetics and challenges for its practical application. J Photoch
    Photobio C 17:1-25

Sambrook J, Russell D (2001) Molecular cloning: a laboratory manual. Gold Spring
    Harbor, New York

Saschenbrecker S, Bracher A, Rao KV, Rao BV, Hartl FU, Hayer-Hartl M (2007)
    Structure and function of RbcX, an assembly chaperone for hexadecameric
    RuBisCO. Cell 129(6):1189-1200

Schnee C, Köllner TG, Held M, Turlings TC, Gershenzon J, Degenhardt J (2006) The
    products of a single maize sesquiterpene synthase form a volatile defense signal
    that attracts natural enemies of maize herbivores. Proceedings of the National
    Academy of Sciences of the United States of America 103(4):1129-1134

Schneegurt MA, Sherman DM, Sherman LA (1997) Composition of the carbohydrate
    granules of the cyanobacterium, Cyanothece sp. strain ATCC 51142. Arch
    Microbiol 167(2-3):89-98

Schneider GJ, Tumer NE, Richaud C, Borbely G, Haselkorn R (1987) Purification and
    characterization of RNA polymerase from the cyanobacterium *Anabaena* 7120. J
    Biol Chem 262(30):14633-14639

Schubert W-D, Klukas O, Krauß N, Saenger W, Fromme P, Witt HT (1997) Photosystem
    I of Synechococcus elongatus at 4 Å resolution: comprehensive structure analysis.
    Journal of molecular biology 272(5):741-769

Schürmann P, Buchanan BB (2008) The ferredoxin/thioredoxin system of oxygenic
    photosynthesis. Antioxidants & redox signaling 10(7):1235-1274

Seemann M, Bui BTS, Wolff M, Miginiac-Maslow M, Rohmer M (2006) Isoprenoid
    biosynthesis in plant chloroplasts via the MEP pathway: direct
    thylakoid/ferredoxin-dependent photoreduction of GcpE/IspG. FEBS letters
    580(6):1547-1552

Sharwood RE, von Caemmerer S, Maliga P, Whitney SM (2008) The catalytic properties
    of hybrid RuBisCO comprising tobacco small and sunflower large subunits mirror
    the kinetically equivalent source RuBisCOs and can support tobacco growth.
    Plant Physiol 146(1):83-96

Shestakov S, Khyen NT (1970) Evidence for genetic transformation in blue-green alga
    Anacystis nidulans. Molecular and General Genetics MGG 107(4):372-375

Shevela D (2011) Adventures with cyanobacteria: a personal perspective. Frontiers in plant science 2:28

Shi Y, Zhao W, Zhang W, Ye Z, Zhao J (2006) Regulation of intracellular free calcium concentration during heterocyst differentiation by HetR and NtcA in *Anabaena* sp. PCC 7120. Proceedings of the National Academy of Sciences 103(30):11334-11339

Shih PM, Zarzycki J, Niyogi KK, Kerfeld CA (2014) Introduction of a synthetic CO2-fixing photorespiratory bypass into a cyanobacterium. J Biol Chem 289(14):9493-9500

Simmons BA, Loque D, Blanch HW (2008) Next-generation biomass feedstocks for biofuel production. Genome biology 9(12):1

Singh NK, Dhar DW (2011) Microalgae as second generation biofuel. A review. Agron Sustain Dev 31(4):605-629

Singh RN, Sharma S (2012) Development of suitable photobioreactor for algae production - A review. Renew Sust Energ Rev 16(4):2347-2353

Spreitzer RJ, Peddi SR, Satagopan S (2005) Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal RuBisCO. Proceedings of the National Academy of Sciences of the United States of America 102(47):17225-17230

Spreitzer RJ, Salvucci ME (2002) RuBisCO: structure, regulatory interactions, and possibilities for a better enzyme. Annual review of plant biology 53(1):449-475

Stavi I, Lal R (2013) Agriculture and greenhouse gases, a common tragedy. A review. Agron Sustain Dev 33(2):275-289

Stemmer WP (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proceedings of the National Academy of Sciences 91(22):10747-10751

Stephenson AL, Dennis JS, Scott SA (2008) Improving the sustainability of the production of biodiesel from oilseed rape in the UK. Process Safety and Environmental Protection 86(6):427-440

Stitt M, Lunn J, Usadel B (2010) Arabidopsis and primary photosynthetic metabolism–more than the icing on the cake. The Plant Journal 61(6):1067-1091

Suh IS, Lee SB (2001) Cultivation of a cyanobacterium in an internally radiating air-lift photobioreactor. Journal of Applied Phycology 13(4):381-388

Sun D, Cabrera-Martinez RM, Setlow P (1991) Control of transcription of the Bacillus subtilis spoIIIG gene, which codes for the forespore-specific transcription factor sigma G. Journal of bacteriology 173(9):2977-2984

Sun Z, Cunningham FX, Gantt E (1998) Differential expression of two isopentenyl pyrophosphate isomerases and enhanced carotenoid accumulation in a unicellular chlorophyte. Proceedings of the National Academy of Sciences of the United States of America 95(19):11482-11488

Suzuki E, Ohkawa H, Moriya K, Matsubara T, Nagaike Y, Iwasaki I, Fujiwara S, Tsuzuki M, Nakamura Y (2010) Carbohydrate metabolism in mutants of the cyanobacterium Synechococcus elongatus PCC 7942 defective in glycogen synthesis. Applied and environmental microbiology 76(10):3153-3159

Suzuki E, Onoda M, Colleoni C, Ball S, Fujita N, Nakamura Y (2013) Physicochemical variation of cyanobacterial starch, the insoluble α-glucans in cyanobacteria. Plant and cell physiology 54(4):465-473

Tabita FR (2007) RuBisCO: the enzyme that keeps on giving. Cell 129(6):1039-1040

Tabita FR, Satagopan S, Hanson TE, Kreel NE, Scott SS (2008) Distinct form I, II, III, and IV RuBisCO proteins from the three kingdoms of life provide clues about RuBisCO evolution and structure/function relationships. Journal of experimental botany 59(7):1515-1524

Tamburic B, Zemichael FW, Crudge P, Maitland GC, Hellgardt K (2011) Design of a novel flat-plate photobioreactor system for green algal hydrogen production. Int J Hydrogen Energ 36(11):6578-6591

Tamoi M, Ishikawa T, Takeda T, Shigeoka S (1996) Molecular Characterization and Resistance to Hydrogen Peroxide of Two Fructose-1, 6-bisphosphatases from SynechococcusPCC 7942. Archives of Biochemistry and Biophysics 334(1):27-36

Tang K-H, Barry K, Chertkov O, Dalin E, Han CS, Hauser LJ, Honchak BM, Karbach LE, Land ML, Lapidus A (2011) Complete genome sequence of the filamentous anoxygenic phototrophic bacterium Chloroflexus aurantiacus. BMC genomics 12(1):334

Taton A, Lis E, Adin DM, Dong G, Cookson S, Kay SA, Golden SS, Golden JW (2012) Gene transfer in Leptolyngbya sp. strain BL0902, a cyanobacterium suitable for production of biomass and bioproducts. PloS one 7(1):e30901

Tcherkez GG, Farquhar GD, Andrews TJ (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. Proceedings of the National Academy of Sciences 103(19):7246-7251

Tippmann S, Chen Y, Siewers V, Nielsen J (2013) From flavors and pharmaceuticals to advanced biofuels: Production of isoprenoids in Saccharomyces cerevisiae. Biotechnol J 8(12):1435-1444

Torres-Sánchez A, Gómez-Gardeñes J, Falo F (2015) An integrative approach for modeling and simulation of heterocyst pattern formation in cyanobacteria filaments. PLoS Comput Biol 11(3):e1004129

Torzillo G, Pushparaj B, Masojidek J, Vonshak A (2003) Biological constraints in algal biotechnology. Biotechnology and bioprocess engineering 8(6):338-348

Tredici MR, Zittelli GC (1998) Efficiency of sunlight utilization: Tubular versus flat photobioreactors. Biotechnol Bioeng 57(2):187-197

Trumpower BL (1990) Cytochrome bc1 complexes of microorganisms. Microbiol Rev 54(2):101-129

Tumer NE, Robinson SJ, Haselkorn R (1983) Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. Nature 306:337-342

Ungerer J, Tao L, Davis M, Ghirardi M, Maness P-C, Yu J (2012) Sustained photosynthetic conversion of $CO_2$ to ethylene in recombinant cyanobacterium Synechocystis 6803. Energy & Environmental Science 5(10):8998-9006

van Schie CC, Haring MA, Schuurink RC (2007) Tomato linalool synthase is induced in trichomes by jasmonic acid. Plant molecular biology 64(3):251-263

Vasquez-Bermudez M, Herrero A, Flores E (2003) Carbon supply and 2-oxoglutarate effects on expression of nitrate reductase and nitrogen-regulated genes in Synechococcus sp. strain PCC 7942. Fems Microbiol Lett 221(2):155-159

Vega-Palas MA, Madueno F, Herrero A, Flores E (1990) Identification and cloning of a regulatory gene for nitrogen assimilation in the cyanobacterium Synechococcus sp. strain PCC 7942. J Bacteriol 172(2):643-7

Velasco J (2013) Amyris Meets Major Milestone at Its Farnesene Production Facility. Amyris Incorporated

Vioque A (2007) Transformation of cyanobacteria Transgenic microalgae as green cell factories. Springer, pp 12-22

Voegele E (2014) Amyris reports reduced farnesene production costs in 2013 results. Biomass Magazine

von Wobeser EA, Ibelings BW, Bok J, Krasikov V, Huisman J, Matthijs HC (2011) Concerted changes in gene expression and cell physiology of the cyanobacterium Synechocystis sp. strain PCC 6803 during transitions between nitrogen and light-limited growth. Plant physiology 155(3):1445-1457

Vranova E, Coman D, Gruissem W (2013) Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annu Rev Plant Biol 64:665-700 doi:10.1146/annurev-arplant-050312-120116

Wang C, Yoon SH, Jang HJ, Chung YR, Kim JY, Choi ES, Kim SW (2011) Metabolic engineering of Escherichia coli for alpha-farnesene production. Metab Eng 13(6):648-55

Wang X, Ort DR, Yuan JS (2015) Photosynthetic terpene hydrocarbon production for fuels and chemicals. Plant biotechnology journal 13(2):137-146

Wastl J, Purton S, Bendall DS, Howe CJ (2004) Two forms of cytochrome c6 in a single eukaryote. Trends in Plant Science 9(10):474

Watson GM, Yu JP, Tabita FR (1999) Unusual ribulose 1,5-bisphosphate carboxylase/oxygenase of anoxic Archaea. J Bacteriol 181(5):1569-75

Wealand JL, Myers JA, Hirschberg R (1989) Changes in gene expression during nitrogen starvation in *Anabaena variabilis* ATCC 29413. J Bacteriol 171(3):1309-13

Whitelegge JP, Zhang H, Aguilera R, Taylor RM, Cramer WA (2002) Full Subunit Coverage Liquid Chromatography Electrospray Ionization Mass Spectrometry (LCMS+) of an Oligomeric Membrane Protein Cytochrome b6f Complex From Spinach and the Cyanobacterium Mastigocladus Laminosus. Molecular & Cellular Proteomics 1(10):816-827

Whitney SM, Baldet P, Hudson GS, Andrews TJ (2001) Form I RuBisCOs from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. Plant J 26(5):535-47

Whitney SM, Houtz RL, Alonso H (2011) Advancing our understanding and capacity to engineer nature's $CO_2$-sequestering enzyme, RuBisCO. Plant Physiology 155(1):27-35

Whitney SM, von Caemmerer S, Hudson GS, Andrews TJ (1999) Directed mutation of the RuBisCO large subunit of tobacco influences photorespiration and growth. Plant Physiol 121(2):579-588

Wilcox M, Mitchison GJ, Smith RJ (1973) Pattern formation in the blue-green alga *Anabaena*. II. Controlled proheterocyst regression. J Cell Sci 13(3):637-49

Wilson WA, Roach PJ, Montero M, Baroja-Fernández E, Muñoz FJ, Eydallin G, Viale AM, Pozueta-Romero J (2010) Regulation of glycogen metabolism in yeast and bacteria. FEMS microbiology reviews 34(6):952-985

Wolk CP (1967) Physiological basis of the pattern of vegetative growth of a blue-green alga. Proc Natl Acad Sci U S A 57(5):1246-51

Wolk CP (2000) Heterocyst formation in *Anabaena* Prokaryotic development. American Society of Microbiology, pp 83-104

Wolk CP, Ernst A, Elhai J (1994) Heterocyst metabolism and development The molecular biology of cyanobacteria. Springer, pp 769-823

Wolk CP, Quine MP (1975) Formation of one-dimensional patterns by stochastic processes and by filamentous blue-green algae. Dev Biol 46(2):370-82

Wolk CP, Vonshak A, Kehoe P, Elhai J (1984) Construction of shuttle vectors capable of conjugative transfer from Escherichia coli to nitrogen-fixing filamentous cyanobacteria. Proceedings of the National Academy of Sciences 81(5):1561-1565

Woodrow IE, Mott KA (1993) Modelling C3 photosynthesis: a sensitivity analysis of the photosynthetic carbon-reduction cycle. Planta 191(4):421-432

Woods RP, Legere E, Moll B, Unamunzaga C, Mantecon E (2010) Closed photobioreactor system for continued daily in situ production, separation, collection, and removal of ethanol from genetically enhanced photosynthetic organisms. Google Patents

Wu G, Shen Z, Wu Q, Zhao N (2000) Molecular cloning and construction of agp gene deletion-mutant in cyanobacterium Synechocystis sp. PCC 6803. Acta Botanica Sinica 43(5):512-516

Wu S, Schalk M, Clark A, Miles RB, Coates R, Chappell J (2006) Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. Nature biotechnology 24(11):1441-1447

Wu X, Liu D, Lee MH, Golden JW (2004) patS minigenes inhibit heterocyst development of *Anabaena* sp. strain PCC 7120. Journal of bacteriology 186(19):6422-6429

Xiao YL, Zahariou G, Sanakis Y, Liu PH (2009) IspG Enzyme Activity in the Deoxyxylulose Phosphate Pathway: Roles of the Iron-Sulfur Cluster. Biochemistry 48(44):10483-10485

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:273-293

Xu Y, Guerra LT, Li Z, Ludwig M, Dismukes GC, Bryant DA (2013) Altered carbohydrate metabolism in glycogen synthase mutants of Synechococcus sp. strain PCC 7002: cell factories for soluble sugars. Metabolic engineering 16:56-67

Yang Y, Huang XZ, Wang L, Risoul V, Zhang CC, Chen WL (2013) Phenotypic variation caused by variation in the relative copy number of pDU1-based plasmids expressing the GAF domain of Pkn41 or Pkn42 in *Anabaena* sp PCC 7120. Res Microbiol 164(2):127-135

Yoo S-H, Keppel C, Spalding M, Jane J-l (2007) Effects of growth condition on the structure of glycogen produced in cyanobacterium Synechocystis sp. PCC6803. International journal of biological macromolecules 40(5):498-504

Yoo S-H, Spalding MH, Jane J-l (2002) Characterization of cyanobacterial glycogen isolated from the wild type and from a mutant lacking of branching enzyme. Carbohydrate research 337(21):2195-2203

Yoon H-S, Golden JW (1998) Heterocyst pattern formation controlled by a diffusible peptide. Science 282(5390):935-938

Yoshida S, Atomi H, Imanaka T (2007) Engineering of a type III RuBisCO from a hyperthermophilic archaeon in order to enhance catalytic performance in mesophilic host cells. Appl Environ Microbiol 73(19):6254-61

Yoshihara S, Suzuki F, Fujita H, Geng XX, Ikeuchi M (2000) Novel putative photoreceptor and regulatory genes required for the positive phototactic movement of the unicellular motile cyanobacterium Synechocystis sp. PCC 6803. Plant and Cell Physiology 41(12):1299-1304

Young JD (2014) INCA: a computational platform for isotopically non-stationary metabolic flux analysis. Bioinformatics 30(9):1333-1335

Yuan JS, Köllner TG, Wiggins G, Grant J, Degenhardt J, Chen F (2008) Molecular and genomic basis of volatile-mediated indirect defense against insects in rice. The Plant Journal 55(3):491-503

Zarzycki J, Fuchs G (2011) Coassimilation of organic substrates via the autotrophic 3-hydroxypropionate bi-cycle in Chloroflexus aurantiacus. Applied and environmental microbiology 77(17):6181-6188

Zhao Y, Shi Y, Zhao W, Huang X, Wang D, Brown N, Brand J, Zhao J (2005) CcbP, a calcium-binding protein from *Anabaena* sp. PCC 7120, provides evidence that calcium ions regulate heterocyst differentiation. Proceedings of the National Academy of Sciences of the United States of America 102(16):5744-5748

Zhou R, Wei X, Jiang N, Li H, Dong Y, Hsi K-L, Zhao J (1998) Evidence that HetR protein is an unusual serine-type protease. Proceedings of the National Academy of Sciences 95(9):4959-4963

Zhu X-G, Long SP, Ort DR (2008) What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? Current opinion in biotechnology 19(2):153-159

Zouni A, Witt H-T, Kern J, Fromme P, Krauss N, Saenger W, Orth P (2001) Crystal structure of photosystem II from Synechococcus elongatus at 3.8 Å resolution. Nature 409(6821):739-743

Zurbriggen A, Kirst H, Melis A (2012) Isoprene production via the mevalonic acid pathway in Escherichia coli (bacteria). BioEnergy Research 5(4):814-828