South Dakota State University
Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange

Electronic Theses and Dissertations

2020

Genetic Study of Alkane Production and Oxic Nitrogen Fixation in Anabaena SP. PCC 7120

Jaimie L. Gibbons
South Dakota State University

Follow this and additional works at: https://openprairie.sdstate.edu/etd
Part of the Biology Commons, and the Molecular Biology Commons

Recommended Citation
https://openprairie.sdstate.edu/etd/3958

This Dissertation - Open Access is brought to you for free and open access by Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. For more information, please contact michael.biondo@sdstate.edu.
GENETIC STUDY OF ALKANE PRODUCTION AND OXIC NITROGEN FIXATION

IN ANABAENA SP. PCC 7120

BY

JAIMIE L. GIBBONS

A dissertation submitted in partial fulfillment of the requirements for the

Doctor of Philosophy

Major in Biological Sciences

Specialization in Molecular Biology

South Dakota State University

2020
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 does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.
ACKNOWLEDGEMENTS

I would like to thank my PhD advisor, Dr. Ruanbao Zhou, for supporting and guiding me through my research. Dr. Zhou was kind enough to take me on as an undergraduate researcher and was instrumental in encouraging me to pursue a PhD in his lab. Dr. Zhou, thank you for your trust and confidence in me throughout my years working in your lab. I would also like to thank my committee members Dr. Volker Brozel, Dr. Michael Hildreth, Dr. Senthil Subramanian, and Dr. Charles Dieter for their guidance and feedback throughout my program years. Thank you to Dr. Liping Gu for your advice and help with my research as well as for your loving care.

I have found that working in a lab often brings along with it the feeling of being part of a family. In Dr. Zhou’s lab, I have met some of my best friends and have been lucky enough to benefit from the supportive and encouraging environment that being part of a lab family offers. I would like to thank my lab mates Yeyan Qui, Nate Braselton, Trevor Van Den Top, James Young, Max Jakubiak, Chuck Halfmann, Tylor Johnson, Huilan Zhu, Kangming Chen, Nanfang Wang, Xinyi Xu, and Shengni Tian.

Last but not least, I would like to thank my friends and family for their love and support in all aspects of my life. I am forever thankful for you with gratitude that cannot be expressed in words.

I am so deeply grateful for you all.
TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................................. ix
LIST OF TABLES ...................................................................................................................... xi
ABSTRACT ............................................................................................................................. xii

CHAPTER 1: Literature Review ............................................................................................. 1

1.1 Introduction to cyanobacteria .......................................................................................... 1

1.2 Hydrocarbon production in cyanobacteria ..................................................................... 1

1.2.1 Fatty acid pool and hydrocarbon profiles in cyanobacteria ....................................... 2
1.2.2 Hydrocarbon production pathways ........................................................................... 2
1.2.3 Categorizing cyanobacterial strains by hydrocarbon production ............................... 4
1.2.4 Lipid turnover and metabolism ................................................................................ 4
1.2.5 Alkane/alkene biosynthesis pathways ...................................................................... 6
1.2.6 Biofuel potentials ..................................................................................................... 11
1.2.7 Genetic engineering of cyanobacteria for enhanced alkane/alkene production ........... 13
1.2.8 Physiological role of alkanes/alkenes in cyanobacteria ............................................ 18

1.3 Nitrogen fixation in cyanobacteria .................................................................................. 26

1.3.1 Mechanism of cyanobacterial photosynthetic oxic nitrogen fixation ......................... 27
1.3.2 Filamentous, heterocystous cyanobacteria .................................................................. 31
1.3.3 Filamentous, non-heterocystous cyanobacteria ......................................................... 40
1.3.4 Unicellular cyanobacteria ....................................................................................... 43
1.3.5 Symbiotic nitrogen-fixing cyanobacteria ................................................................... 47
CHAPTER 2: Identification of Two Genes Required for Heptadecane Production in a N₂-fixing Cyanobacterium *Anabaena* Sp. Strain PCC 7120

2.1 Abstract

2.2 Introduction

2.3 Materials and Methods

2.3.1 Bacterial strains and plasmids

2.3.2 Construction of *aar/ado* knockout plasmid pZR935

2.3.3 Knocking out *alr5283-5284* in *Anabaena* sp. PCC 7120

2.3.4 Complementation of Alr5283-Alr5284 knockout mutant

2.3.5 Headspace sample collections and extractions

2.3.6 Gas Chromatography-Mass Spectrometry (GC-MS) analysis of hydrocarbon samples

2.3.7 Chlorophyll content analysis

2.3.8 GenBank access to genes

2.4 Results

2.4.1 Identification of heptadecane emitted from *Anabaena* 7120

2.4.2 Bioinformatics analysis to identify hydrocarbon biosynthesis genes

2.4.3 Construction of DR935

2.4.4 Alr5283 and Alr5284 are responsible for heptadecane production in *Anabaena* 7120

2.4.5 Heptadecane production

2.5 Discussion

2.5.1 *alr5283* and *alr5284* are required for heptadecane production in
CHAPTER 3: Heptadecane Production Enhances Oxic Nitrogen Fixation in

Anabaena sp. PCC 7120

3.1 Abstract

3.2 Introduction

3.3 Materials and Methods

3.3.1 Bacterial strains and culture growth

3.3.2 Heptadecane collection and extraction

3.3.3 Gas Chromatography-Mass Spectrometry (GC-MS) analysis of hydrocarbon samples

3.3.4 AAR/ADO gene expression analysis via fluorescence microscopy

3.3.5 Acetylene reduction assay for measuring nitrogenase activity

3.3.6 GC-MS analysis of acetylene reduction

3.3.7 Heterocyst counts

3.4 Results

3.4.1 Heptadecane production in fixed-nitrogen replete (Bg11), fixed-nitrogen starvation (Bg110), and Bg110-adapted cultures

3.4.2 Growth of Anabaena 7120 and DR935 in fixed-nitrogen starvation (Bg110) and Bg110-adapted cultures
3.4.3 AAR/ADO gene expression in fixed-nitrogen replete (Bg11), fixed nitrogen starvation (Bg110), and Bg110-adapted cultures.................................85
3.4.4 Fragmentation of DR935 following fixed-nitrogen starvation ..........87
3.4.5 Nitrogenase activity in Anabaena 7120 and DR935.........................88
3.4.6 Heterocyst counts following nitrogen stepdown and in long-term nitrogen-fixing cultures.................................................................89

3.5 Discussion...........................................................................................................89

CHAPTER 4: Identification of Two Genes Required for Oxic N₂ Fixation in Anabaena sp. PCC 7120.................................................................................................96

4.1 Abstract............................................................................................................96
4.2 Introduction......................................................................................................97
4.3 Materials and Methods..................................................................................103
4.3.1 Bacterial strains and plasmids.................................................................103
4.3.2 Construction of knockout plasmids for heterocyst-specific genes ......104
4.3.3 Knocking out heterocyst-specific genes via single crossover approach.................................................................108
4.3.4 Screening single-crossover mutants for completely segregated knockout.........................................................................................111
4.3.5 Growth on fixed nitrogen-free medium (Bg110) to screen for Fox genes ..................................................................................................112
4.3.6 Complementation of SR2066..................................................................113
4.3.7 Nitrogenase activity and GC-MS analysis of acetylene reduction ......113

4.4 Results..............................................................................................................114
4.4.1 Identification of Fox genes ................................................................. 114

4.4.2 Verifying requirement of alr0731 for diazotrophic growth via complementation experiment........................................................................ 117

4.4.3 Effect of alr0731 knockdown on nitrogenase activity.............................. 120

4.5 Discussion................................................................................................. 121

CHAPTER 5: Summary and Conclusions............................................................ 126

REFERENCES .................................................................................................. 131
LIST OF FIGURES

Figure 1.1 Alkane/alkene biosynthesis via AAR/ADO pathway in cyanobacteria ..7

Figure 1.2 Four approaches for solar-powered oxic N\textsubscript{2}-fixation by cyanobacteria ..............................................................28

Figure 2.1 Schematic illustration of pZR935 construction for knocking out alr5283 and alr5284 ..............................................................58

Figure 2.2 Identification of heptadecane produced by Anabaena 7120 .........................63

Figure 2.3 Identification of the aldehyde decarboxylase and acyl-ACP reductase Anabaena 7120 ................................................................64

Figure 2.4 Construction of Alr5283-Alr5284 knockout mutant (DR935) and its complemented strains in Anabaena 7120 ........................................66

Figure 2.5 Alr5283 and Alr5284 are responsible for the heptadecane production in Anabaena 7120 ................................................................68

Figure 2.6 Growth phenotypic analysis from wildtype, DR935, and DR935(pZR2239) cultures grown in BG11 medium ........................................................................69

Figure 3.1 Total heptadecane production in wildtype Anabaena 7120 in fixed nitrogen starvation (Bg11\textsubscript{0}), Bg11\textsubscript{0}-adapted, and fixed-nitrogen replete cultures.....83

Figure 3.2 Growth curves comparing wildtype and mutant (DR935) in fixed nitrogen starvation (Bg11\textsubscript{0}), Bg11\textsubscript{0}-adapted, and fixed-nitrogen replete cultures..... 85

Figure 3.3 Alkane gene (aar/ado) expression in fixed nitrogen starvation (Bg11\textsubscript{0}), Bg11\textsubscript{0}-adapted, and fixed-nitrogen replete cultures........................................87

Figure 3.4 Nitrogenase activity in wildtype and mutant (DR935) in fixed-nitrogen starvation (Bg11\textsubscript{0}) and Bg11\textsubscript{0}-adapted cultures.................................88
Figure 4.1 Schematic representation of single crossover recombination to
knockout target gene (Chen et al. 2015) .................................................................110

Figure 4.2 all3132 mutant screening and characterization ....................................116

Figure 4.3 alr0731 mutant screening, characterization, and complementation ......117

Figure 4.4 Strain phenotypes following nitrogen stepdown ....................................119

Figure 4.5 Nitrogenase activity in partial knockout (SR2066-4), complement
[SR2066-1(pZR2316)], and wildtype ........................................................................120
LIST OF TABLES

Table 1.1 Representatives of four categories of N₂-fixing cyanobacteria ................. 29

Table 2.1 Plasmids and bacterial strains used alkane knockout study .................. 56

Table 2.2 Primers used in alkane knockout study ........................................... 57

Table 4.1 Heterocyst-specific protein-coding genes found in Anabaena cylindrica and their homolog in Anabaena 7120 used in heterocyst gene knockout study .......... 99

Table 4.2 Fox genes of Anabaena sp. PCC 7120 .......................................... 100

Table 4.3 Primers used in heterocyst gene knockout study ............................. 104

Table 4.4 Plasmids and bacterial strains created or used in heterocyst gene knockout study ................................................................. 106
ABSTRACT

GENETIC STUDY OF ALKANE PRODUCTION AND OXIC NITROGEN FIXATION IN ANABAENA SP. PCC 7120

JAIMIE GIBBONS

2020

Anabaena sp. PCC 7120 is a filamentous, nitrogen-fixing cyanobacterium that uses spatial separation to perform photosynthesis and nitrogen fixation simultaneously. Under fixed nitrogen limiting conditions, Anabaena 7120 forms specialized cells, called heterocysts, to fix nitrogen. In this research, I sought to add to the knowledge surrounding the process of heterocystous nitrogen fixation in Anabaena 7120.

Cyanobacteria universally produce alkanes, which have been suggested to play a role in helping the organism adapt to abiotic stress, such as diazotrophic conditions. In my first study, I sought to identify the genes required for production of the hydrocarbon heptadecane (C_{17}H_{36}). Through using a double crossover approach and determining loss of function, I identified two genes (alr5283 and alr5284) responsible for heptadecane production in Anabaena 7120. In my second study, I employed the alkane knockout mutant to study the impact of alkanes in nitrogen-fixing conditions. My results showed that heptadecane production spikes immediately following nitrogen starvation. Correspondingly, the mutant, which is unable to produce heptadecane, exhibited a fragmented phenotype with a lower nitrogenase activity in the days immediately following nitrogen starvation. These results suggested that heptadecane may play an important role in maintaining membrane integrity for cell-to-cell connections, especially during the initial response to nitrogen starvation.
In my last study, I knocked out 16 genes whose encoded proteins had been found only in heterocysts (Qiu 2018). Two of these gene mutants (for *all3132* and *alr0731*) were unable to grow under aerobic nitrogen-fixing conditions. I was able to complement the *alr0731* mutant, and it regained its ability to grow in diazotrophic conditions. Therefore, I categorized *alr0731* as a Fox gene (unable to fix nitrogen in the presence of oxygen). This gene encodes an anaerobic ribocucleoside triphosphate reductase activating protein that is needed for conversion of ribonucleoside triphosphates (NTPs) to deoxyribonucleoside triphosphates (dNTPs). Because cell division (requiring DNA replication) and gene rearrangements have been linked to heterocyst differentiation, *Alr0731* may be primarily important in the process of heterocyst differentiation. Future work will continue to add to the understanding of key regulators and contributors to the process of heterocyst differentiation and nitrogen fixation in cyanobacteria.
CHAPTER 1: Literature Review

1.1 Introduction to cyanobacteria

Cyanobacteria are one of the most ancient groups of living organisms and are believed to be the first photosynthetic organisms (Los and Mironov 2015; Ruffing 2011). Furthermore, cyanobacteria are believed to be the ancestors of plant chloroplasts (Falcon et al. 2010). Cyanobacteria, along with eukaryotic algae, are the primary photosynthetic organisms in oceanic waters; as such, they play a large role in global carbon cycling (Ruffing 2011). These microorganisms offer a diverse range of physiological and biochemical abilities. They inhabit fresh and marine waters, may be unicellular or filamentous, and may fix nitrogen or require it from their environment (Ruffing 2011). Nitrogen-fixing cyanobacteria also play an important role in the nitrogen cycle (Falkowski 1997; Ruffing 2011). Diazotrophic cyanobacteria are ideal organisms for studying, understanding, and optimizing nitrogen fixation for applications to replace conventional fertilization methods. Cyanobacteria also produce other substances that can be societally beneficial, such as hydrocarbons for biofuels. Hydrocarbon production and nitrogen fixation by cyanobacteria are thus being investigated for their economic and environmental potential as a drop-in fuel source and biological fertilizer, respectively (Ruffing 2011; Singh et al. 2016).

1.2 Hydrocarbon production in cyanobacteria

All cyanobacteria naturally produce hydrocarbons, such as alkanes and alkenes (Coates et al. 2014). Alkanes are saturated hydrocarbons, containing single bonds between carbon molecules, while alkenes are unsaturated and contain one or more double
bonds between carbon molecules. The presence of double bonds in the fatty acid (FA) precursor determines whether alkanes or alkenes are produced by a biosynthetic pathway.

1.2.1 Fatty acid pool and hydrocarbon profiles in cyanobacteria

FA chain length typically ranges from C14 to C18 in cyanobacteria, with the number of double bonds varying from 0 to 4 (Los and Mironov 2015). Octadecanoic acid and hexadecanoic acid along with various unsaturated derivatives are the most common cyanobacterial FAs. Typically, alkane/alkene production from FAs produces an alkane/alkene with one less carbon than the FA substrate (termed the “n-1” rule) (Schirmer et al. 2010). Correspondingly, heptadecane followed by heptadecene, pentadecane, and 7-methylheptadecane are the most commonly observed cyanobacterial hydrocarbons (Coates et al. 2014). A study by Liu et al. (2013) revealed that branched alkanes/alkenes occur predominately in filamentous cyanobacterial strains and are less commonly found in unicellular strains.

1.2.2 Hydrocarbon production pathways

Cyanobacteria produce hydrocarbons via two metabolic pathways. The first pathway (AAR/ADO pathway) contains a two-step conversion of fatty acids to fatty aldehydes to alkanes and alkenes (Schirmer et al. 2010). In the second pathway, called the olefin synthase (OLS) pathway, α-olefins are produced by fatty acid elongation followed by decarboxylation to produce a terminal alkene. Alpha-olefins uniquely contain a terminal double bond. In cyanobacteria, the two pathways are mutually exclusive (Coates et al. 2014; Yoshino et al. 2015). Though all cyanobacteria produce hydrocarbons, each species only contains one of the hydrocarbon production pathways. The majority of sequenced cyanobacterial strains contain the AAR/ADO pathway (122 of
while the OLS pathway thus far is present in a total of 17 of the sequenced strains (Coates et al. 2014). Cyanobacterial species containing the AAR/ADO pathway produced primarily alkanes, specifically heptadecane and branched alkanes (e.g., 7-methylheptadecane) (Coates et al. 2014). *Lyngbya sp.* PCC 8106 was the only AAR/ADO containing strain that produced alkenes (Coates et al. 2014). Cyanobacterial strains expressing the OLS pathway appeared to only produce alkenes (Coates et al. 2014). Cyanobacteria most commonly produce heptadecane followed by heptadecene, pentadecane, and 7-methylheptadecane (Allen et al. 2010; Coates et al. 2014; Han et al. 1968; Schirmer et al. 2010). This hydrocarbon composition is consistent with cyanobacterial FA profiles, where octadecanoic acid and hexadecanoic acid are the most prevalent FAs and serve as substrates for the AAR/ADO and OLS pathways, respectively (Coates et al. 2014).

The AAR/ADO and OLS strains are mutually exclusive in cyanobacteria. The AAR/ADO pathway, present in the vast majority of cyanobacteria, is hypothesized to be the ancestral hydrocarbon pathway in cyanobacteria (Coates et al. 2014). The OLS pathway is primarily present in a single clade of cyanobacterial species (Clade A). However, *Lyngbya sp.* PCC 6406 (Clade B) and *Moorea* strains (Clade C) also contain the OLS pathway and are outside the major clade of OLS-containing cyanobacteria, suggesting horizontal gene transfer may have occurred (Coates et al. 2014). The demarcation between AAR/ADO and OLS containing strains may indicate an unknown selective pressure that allows only one alkane/alkene biosynthesis pathway to be present in an organism (Coates et al. 2014).
1.2.3 Categorizing cyanobacterial strains by hydrocarbon production

Cyanobacterial hydrocarbon production is widespread. Los and Mironov (2015) have categorized cyanobacteria into four groups based on their fatty acid (FA) composition. Group 1 synthesizes FAs containing only one double bond, typically at the carbon 9 position. Members of group 1 include mesophilic and thermophilic freshwater, unicellular strains, such as *Synechococcus elongatus*; additionally, the filamentous heterocystous thermophilic *Mastigocladus laminosus* is a member. Group 2 contains cyanobacteria that desaturate C16 and C18 lipids at carbon 9 and carbon 12 positions; the marine species *Prochlorococcus* and *Synechococcus* sp. (marine) represent this group. Group 3 produces trienoic FAs and includes *Synechococcus* sp. PCC 7002, *Trichodesmium erythraeum*, and some *Anabaena* species. Group 4 members contain four acyl-lipid FA desaturases and can synthesize tetraenoic stearidonic acid; *Synechocystis* sp. PCC 6803 is a representative of this group (Los and Mironov 2015). These groupings demonstrate that specific hydrocarbons are produced by different strains, and hydrocarbon production can be used as a fingerprint to identify and categorize cyanobacterial strains.

1.2.4 Lipid turnover and metabolism

FAs serve as the precursor molecule for lipid and alkane/alkene biosynthesis in cyanobacteria (Kaczmarzyk and Fulda 2010; Schirmer et al. 2010). Acetyl-CoA carboxylase initiates FA biosynthesis by combining acetyl-CoA and bicarbonate to form malonyl-CoA (Lu 2010). Malonyl-CoA subunits are polymerized in cyanobacteria by a type II fatty acid synthase (FAS) complex (Kaczmarzyk and Fulda 2010; Peralta-Yahya et al. 2012). Acyl carrier proteins (ACPs) elongate the fatty acyl chains by repeated
cycles of decarboxylative condensation, β-keto reduction, dehydration, and enol reduction (Peralta-Yahya et al. 2012). Fatty acyl-ACPs, in turn, can be used directly as substrates for acyl-transferases, which incorporate the FAs into membrane lipids (Kaczmarzyk and Fulda 2010). Alternately, fatty acyl-ACPs may directly enter the alkane/alkene biosynthesis pathway (Kaczmarzyk and Fulda 2010).

Free fatty acids (FFAs) in the cell can serve as another precursor for alkane/alkene biosynthesis. FFAs are released from complex lipids in the cell, such as monoglycosyldiacylglycerol and monogalactosyldiasylglycerol, in the process of lipid remodeling or degradation (Kaczmarzyk and Fulda 2010). Membrane phospholipids can also be a source of FFAs. Cyanobacterial membrane lipids are largely diacylglycerols, present in the forms of monogalactosyl diacylglycerol, digalactosyl diacylglycerol, phosphatidylglycerol, and sulfoquinovosyl diacylglycerol (Wada and Murata 1990). Lipolytic enzymes, such as galactolipase and phospholipase B, release FAs from diacylglycerols through hydrolyzing the carboxylic ester bond (Svendsen 2000). FAs released from membrane lipids transiently appear in the pool of FFAs before becoming reactivated to serve as substrates in lipid synthesis (Kaczmarzyk and Fulda 2010).

FFAs must be activated by the enzyme acyl-ACP synthetase (AAS) before re-entering the membrane as well as before entering the alkane/alkene biosynthesis pathway (Gao et al. 2012; Kaczmarzyk and Fulda 2010). AAS catalyzes the fatty acid esterification to the thiol of an acyl carrier protein (Kaczmarzyk and Fulda 2010). In AAS knockout mutants, cyanobacteria are unable to activate their fatty acids, resulting in loss of fatty acids to the cell exterior as well as an inability to use exogenous fatty acids present in the culture medium (Kaczmarzyk and Fulda 2010).
Membrane lipid turnover is closely tied to FFA activation. Proteomic studies have shown that AAS is localized at the plasma membrane (Pisareva et al. 2007). In this location, AAS fosters its involvement in activating FFAs released from the plasma membrane during lipid turnover. It also is ideally located to activate exogenous FFAs entering the cell (Kaczmarzyk and Fulda 2010). In AAS mutants, FFAs derived from membrane lipids were secreted by the cell into the medium while this secretion did not occur in wild-type cells (Kaczmarzyk and Fulda 2010). Thus, FFA activation is critical to sequestering FFAs inside the cell where they can integrate into the process of lipid turnover for restructuring membranes or entering other biosynthetic pathways (Kaczmarzyk and Fulda 2010). Thus, membrane lipid turnover plays a key role in FFA supply in the cell for lipid synthesis and potentially alkane/alkene production. Though membrane lipid turnover is energetically costly, the process is widespread across diverse genera (Kaczmarzyk and Fulda 2010). Kaczmarzyk and Fulda (2010) propose that one explanation for significant membrane lipid turnover is as an adaptive response to changing environmental conditions.

1.2.5 Alkane/alkene biosynthesis pathways

1.2.5.1 AAR/ADO

The AAR/ADO pathway synthesizes alkanes/alkenes from fatty acids in two steps. First, acyl-ACP reductase (AAR) converts a fatty acyl-ACP into a fatty aldehyde. Next, aldehyde decarbonylase (ADO) converts the fatty aldehyde into an alkane or alkene by removing the carbonyl group (Fig. 1.1). In addition to cyanobacteria, plants, eukaryotic algae, bacteria, and humans contain AAR. ADO, however, is specific to cyanobacteria (Coates et al. 2014; Li et al. 2012). Alkane/alkene production by the
AAR/ADO pathway follows the “n-1” rule; the decarbonylation step produces alkanes/alkenes containing one less carbon than the fatty aldehyde substrate. Since FAs in cyanobacteria are typically even-numbered, odd-numbered alkanes/alkenes are generated (Coates et al. 2014; Schirmer et al. 2010). Heptadecane and pentadecane are the most commonly reported alkanes in cyanobacteria (Schirmer et al. 2010).

**Figure 1.1** Alkane/alkene biosynthesis via AAR/ADO pathway in cyanobacteria. Dashed arrows represent multiple enzymatic reactions. Pyruvate dehydrogenase complex (PCD), acetyl-CoA carboxylase (ACC), acyl-ACP reductase (AAR), aldehyde decarbonylase (ADO).

**AAR activity**

AAR, using NADPH as a cofactor, breaks down the acyl-ACP into a fatty aldehyde and acyl-ACP (Liu et al. 2013; Schirmer et al. 2010). Fatty aldehydes are toxic to cells and must be further processed upon their production to limit their harmful impact on the cell (Kaiser et al. 2013). In *S. elongatus*, the fatty aldehyde has two possible fates...
upon its release from AAR: (1) the fatty aldehyde may be converted into alkanes by ADO, or (2) the fatty aldehyde can be converted into FFAs by AldE (a class 3 aldehyde dehydrogenase) (Kaiser et al. 2013).

**ADO activity**

The ADO-catalyzed reaction produces formate (HCOO⁻) and an alkane/alkene. This reduction reaction requires an endogenous reducing system and oxygen (Zhang et al. 2013). Four electrons are required for ADO’s reducing activity. The aldehyde hydrogen is retained by formate, and one oxygen atom is incorporated into the formate molecule. In *S. elongatus*, ferredoxin (Fd) and ferredoxin-NADP⁺ reductase (FNR) serve as the electron transport system to support ADO reduction activity (Zhang et al. 2013). In the reduction process, electrons are transferred from NADPH to FNR to Fd and finally to ADO, which uses the electrons to catalyze the reduction of the fatty aldehyde to an alkane/alkene (Zhang et al. 2013).

**AAR/ADO gene organization**

Though frequently adjacently located in cyanobacterial chromosomes, *aar* and *ado* are each controlled by individual promoters, and their expression results in the production of monocistronic transcripts in *S. elongatus* and *Anabaena* sp. PCC 7120 (hereafter *Anabaena* 7120), among other cyanobacterial genomes (Klahn et al. 2014). The *ado* gene is located upstream of *aar* (Klahn et al. 2014). Additionally, while *aar* is controlled by a single promoter, *ado* is controlled by a proximal and a distal promoter (Klahn et al. 2014). Gene expression studies indicated that the distal *ado* promoter has stronger activity under normal conditions (Klahn et al. 2014). However, other conditions may exist in which expression for the *ado* proximal promoter is dominant (Klahn et al. 2014).
The distinct transcriptional regulation for \textit{aar} and \textit{ado} enables different and possibly divergent regulation of the genes and indicates that physiological conditions may require or benefit from individual transcription of \textit{aar} and \textit{ado} (Klahn et al. 2014).

Though the AAR/ADO pathway is present in the vast majority of cyanobacterial genomes (~90%), the AAR amino acid sequences vary, with an averaged conserved identity of 67% (Kudo et al. 2016). Corresponding to the sequence difference is a difference in AAR catalytic activity and preferred substrates. Kudo et al. (2016) studied 12 representative AARs from cyanobacteria to compare their activities and found that AAR activity levels differed amongst the species with \textit{S. elongatus} PCC 7942 AAR activity being the highest. The same study also revealed different AAR substrate preferences amongst cyanobacterial species: marine cyanobacteria had a higher specificity toward 16 C fatty acids while freshwater cyanobacteria tended to prefer 18 C substrates (Kudo et al. 2016). Thus, substrate specificity for AAR appears to correlate to host environment; however, the reason for the difference in FA carbon chain length specificity between marine and freshwater cyanobacteria remains unknown (Kudo et al. 2016).

Sequence alignments and phylogenetic trees of \textit{aar} and \textit{ado} from freshwater cyanobacteria revealed that AAR sequences had similar interspecies relationships to those seen in 16S rDNA genes. ADO sequences did not produce similar phylogenetic relationships to those of the 16S rDNA genes (Liu et al. 2013). These results suggest that AAR has an earlier evolutionary origin than ADO (Liu et al. 2013).
1.2.5.2 OLS

Native α-olefins in cyanobacteria are produced by the OLS pathway and are generally one carbon shorter or one carbon longer than the dominant FAs of the cell membrane (Lennen and Pfleger 2013). The OLS pathway is present in over 250 divergent bacteria (Sukovich et al. 2010). In the OLS pathway, fatty acyl-ACP is converted into an odd-chain number α-olefin by olefin synthase (OLS) (Mendez-Perez et al. 2011). This process involves a polyketide synthase (PKS) pathway that elongates the acyl chain; then decarboxylation occurs, producing a terminal alkene (Coates et al. 2014; Mendez-Perez et al. 2011). The terminal double bond formed in the OLS pathway distinctly separates OLS products from AAR/ADO products (Coates et al. 2014).

Olefin synthase is large enzyme complex sharing homology with type I polyketide synthases; it catalyzes the conversion of fatty acyl-ACPs to α-olefins (Coates et al. 2014; Lennen and Pfleger 2013; Mendez-Perez et al. 2011). In the conversion pathway, a fatty acyl-ACP ligase domain transfers the fatty acyl-ACP to the OLS acyl carrier protein (ACP) via an ATP-dependent mechanism. The substrate, bound to ACP, is elongated by an extension module consisting of ketosynthase (KS), acyltransferase (AT), and ketoreductase (KR). During extension, two carbons from malonyl-CoA are added to the substrate, and the β-keto group of the substrate is reduced to a hydroxyl (Coates et al. 2014; Lennen and Pfleger 2013; Mendez-Perez et al. 2011). In the final step in the OLS pathway, a sulfotransferase (ST) activates the β-hydroxyl group by adding sulfate from an activated substrate, such as 3’-phosphoadenosine 5’-phosphosulfate (Lennen and Pfleger 2013). Sulfonation supplies the energy needed for thioesterase (TE)-catalyzed hydrolysis, releasing the substrate from the enzyme’s surface. TE also catalyzes substrate
decarboxylation and desulfonation, yielding the α-olefin product (Coates et al. 2014; Lennen and Pfleger 2013; Mendez-Perez et al. 2011). Other research has suggested that this final step involves an intermediate β-lactone product formation directly before conversion to the α-olefin (Christenson et al. 2017).

1.2.6 Biofuel potentials

In light of diminishing fossil fuel sources with no corresponding decrease in fossil fuel use, the world population will require a renewable and sustainable energy source. The need for renewable fuel has been met in part by biofuel production from land-based biofuel feedstocks (e.g., corn, sugarcane, switchgrass) (Dismukes et al. 2008; Peralta-Yahya et al. 2012). However, this method of renewable fuel production poses its own set of challenges and drawbacks, among them being high production costs, energy-expensive conversion processes, and the use of arable land (Dismukes et al. 2008; Lu 2010; Peralta-Yahya et al. 2012).

Biodiesel production highlights the limits of current biofuel production attempts: production of biodiesel is about two-fold the production cost of normal diesel. The production cost can be broken into two main components: cost of fats and oils as raw material and cost of processing. Raw materials alone account for 60-75% of total production costs (Krawczyk 1996). Inexpensive fats, such as food industry wastes and animal fats, are available. However, they are low quality fats for biodiesel production, primarily because they contain large amounts of FFAs which must undergo the difficult process to transesterification to be converted to biodiesel (Haas 2005). Plant oils, which contain more fatty acid triglycerides, are more readily converted to biodiesel, and for this reason have been used in biodiesel production (Demirbaş 2002). Direct production of
biodiesel by cyanobacteria poses an even greater opportunity, eliminating the need for any external conversion process (Atsumi et al. 2009; Lennen and Pfleger 2013; Lindberg et al. 2010; Ruffing 2011).

Cyanobacteria are one of few organisms reported to directly produce hydrocarbons (Coates et al. 2014). Hydrocarbons are economically and environmentally important because they can be used directly as a drop-in fuel source (Posewitz 2014). Cetane values indicate diesel and jet fuel quality, with the fuel quality being largely dependent on its hydrocarbon constituents (Coates et al. 2014). Long, saturated hydrocarbons have higher cetane ratings (Balaban 1983; Creton et al. 2010). Heptadecane, the most abundant, naturally occurring hydrocarbon in cyanobacteria, has a centane rating of 105 while another commonly observed cyanobacterial alkane, methylheptadecane, has a centane rating of 66. These centane values make cyanobacterial hydrocarbons ideal candidates for diesel fuel, which only has a cetane rating of 40-55 (Balaban 1983; Creton et al. 2010; Schirmer et al. 2010).

Cyanobacteria will store large amounts of carbohydrates or fats to serve as energy reserves under nutrient-limited conditions (Perez et al. 2016). However, under nutrient-replete conditions, these carbohydrate and fat stores are not present. Though the maximum lipid production actually occurs under the fastest growth rate during nutrient-replete conditions, lipids are directed towards biomass requirements for cell reproduction, namely for membrane synthesis (Kaiser et al. 2013; Posewitz 2014; Singh et al. 2002). Thus, biofuel production from cyanobacteria requires genetic engineering to drive lipid accumulation under nutrient-replete conditions such that lipids are produced beyond what
is required for cellular growth and reproduction, and exist as excess stores inside the cell or to be secreted by the cell (Kaiser et al. 2013).

Cyanobacteria are an ideal organism to use as “biofuel-producing factories.” As previously noted, cyanobacteria naturally produce hydrocarbons. Notably, alkanes are the main component of diesel fuel (Peralta-Yahya et al. 2012) and can thus be used directly as a drop-in fuel source without requiring further energy- and cost-expensive processing. In addition to their natural ability to produce certain high-value chemicals, such as alkanes/alkenes, cyanobacteria are readily genetically manipulated (Ruffing 2011) and are photosynthetic. Thus, cyanobacteria provide a combination of the favorable factors seen in both prokaryotic and plant biofuel-producing organisms. Like other prokaryotes used for biofuel production (e.g., *E. coli*, *Z. mobilis*, *S. cerevisiae*, among others), cyanobacteria can be genetically engineered to directly produce fuels in contrast to plants grown to produce polysaccharides (e.g., cellulose and hemicellulose) which are converted to biofuels through fermentation. Like plants, cyanobacteria are photosynthetic and can harvest solar energy along with CO₂ and water to produce fuel compounds without added energy sources (Lu 2010; Ruffing 2011). Importantly, cyanobacteria have a higher solar energy conversion efficiency and growth rate compared to plants (Dismukes et al. 2008). Additionally, cyanobacteria are grown on non-arable land and thus do not limit the land available for food crops (Dismukes et al. 2008).

1.2.7 Genetic engineering of cyanobacteria for enhanced alkane/alkene production

Native cyanobacterial alkane/alkene levels are between 0.025-0.12% dry biomass (Han et al. 1968; Winters et al. 1969). Higher hydrocarbon yields (0.173%±0.032) have been reported for the OLS pathway than for the AAR/ADO pathway (0.070%±0.008).
Genetic engineering promises to be an important approach to increase cyanobacterial hydrocarbon yields to a level which may ultimately be used as a drop-in fuel source (Posewitz 2014; Ruffing 2011; Savakis and Hellingwerf 2015).

1.2.7.1 Increasing FA production

As a precursor to alkane/alkene biosynthesis, FA production optimization is a target for improving alkane/alkene yields. The first committed step in FA production is the conversion of acetyl-CoA (from the Calvin cycle) plus bicarbonate to make malonyl-CoA by the enzyme acetyl-CoA carboxylase (ACC) (Davis et al. 2000). Overproduction of ACC in *E. coli* was shown to lead to overproduction of malonyl-CoA, with malonyl-CoA constituting 17.7% of the CoA compounds pool in ACC-overproducing strains compared to 0.01% of the pool in wildtype strains (Davis et al. 2000). When the FA-production inhibitor cerulenin was added to the ACC overproduction strain, malonyl-CoA increased to the extent of constituting 55.1% of the CoA metabolite pool. Thus, the majority of malonyl-CoA produced in the ACC-overproducing strain was channeled to FA production (Davis et al. 2000). Indeed, measurements of FA production in the ACC-overproducing strain showed an increased production rate of $5.7 \pm 1.2$-fold over wildtype (Davis et al. 2000). Thus, ACC overproduction effectively increased the flux through the FA biosynthesis pathway. When applied in cyanobacteria, overexpression of ACC in the *Synechocystis* PCC 6803 led to an increase in hydrocarbon production. An 11% and 56% increase in hydrocarbon production occurred in a mutant strains in which *acc* expression was driven by the promoter P\textsubscript{petE} and P\textsubscript{rbcl}, respectively (Tan et al. 2011).

In wildtype cyanobacterial strains, FFAs detectable in the medium are negligible (Kaczmarzyk and Fulda 2010). However, genetic engineering attempts have led to FFA
overproduction in cyanobacterial strains. Overproduction of FFAs in bacteria is channeled into an external metabolic sink (Fell 1997). Introducing a thioesterase has been shown to release the feedback inhibition of FAS through hydrolyzing the fatty-acyl ACPs (Davis et al. 2000; Jiang and Cronan 1994; Liu et al. 2011; Lu 2010; Magnuson et al. 1993). In combination with deletion of genes for FA degradation pathways, hydrolysis of acyl-ACP leads to large FA precursor pool that can feed into alkane/alkene production (Peralta-Yahya et al. 2012).

In an *E. coli* strain overexpressing ACC and thioesterase and lacking *fadD* (a gene encoding fatty acyl-CoA synthase involved in FA biodegradation), FFAs reached titers of approximately 2.5 g/L/d (Lu et al. 2008). In *Synechocystis* sp. PCC 6803, deletion of FA activation genes along with insertion of thioesterase genes led to increased FFA production and secretion (Liu et al. 2011). Additionally, FFA production was enhanced through redirecting carbon flux in the cell by deleting carbon sinks, such as polyhydroxybutyrate, S-layer, cyanophycin, and acetate biosynthesis (Liu et al. 2011). Liu et al. (2011) further increased the secretion of FAs by engineering cyanobacteria to degrade cellular membranes through lipase activity initiated under CO$_2$-limiting conditions. This approach further facilitates FA recovery for downstream biofuel production steps (Liu et al. 2011).

Chain-length specific acyl-ACP thioesterases can also be applied to control fatty-acyl chain length to further enhance FA biosynthesis for biofuel production (Lu 2010). Yuan et al. (1995) were able to modify the substrate specificity of an acyl-ACP thioesterase through protein engineering. By controlling the chain length of the FA, the chain length of the final alkane/alkene product could be specified, which is useful for
producing fuels for varying platforms: jet fuels (requiring 10-14 carbon fatty alkanes) or diesel fuels (requiring 14-18 carbon fatty alkanes) (Huang et al. 2010; Lu 2010).

In cyanobacteria, FFAs must be activated by acyl-ACP synthetase (AAS) before entering the alkane/alkene biosynthetic pathway (Gao et al. 2012). When AAS is inactivated, FFA levels in *Synechocystis* sp. PCC 6803 decreased up to 90% (Hu et al. 2013). Interestingly, while AAS is necessary for alkane production, its overexpression in *Synechocystis* sp. PCC 6803 had no effect on alkane production (Gao et al. 2012).

Fatty acyl-ACP is the end product of fatty acid elongation (Lai and Lan 2015). Following its synthesis, fatty acyl-ACP can directly enter the alkane/alkene biosynthesis pathway where it is first reduced by AAR to a fatty aldehyde, which is then converted to an alkane/alkene (Lai and Lan 2015). Though genetic engineering has been able to increase fatty acyl-ACP levels, alkane/alkene production remains low, suggesting a bottleneck may exist following FA production (Lai and Lan 2015). AAR has been suggested as a potential rate-limiting step in alkane/alkene biosynthesis (Lai and Lan 2015). However, AAR overexpression in *S. elongatus* led to only a low level of alkane production while FFA production dominated, at 50-100-fold the alkane yields (Kaiser et al. 2013). Further research must be done to identify and understand bottlenecks and carbon flow through the fatty acid and alkane/alkene biosynthesis pathways. Particularly, understanding is required of how these pathways, their enzyme kinetics and regulations, work together to affect alkane/alkene yields.

**1.2.7.2 Directly increasing alkane/alkene production**

Attempts to increase hydrocarbon production have targeted overexpression of AAR/ADO and OLS pathway genes for increased alkane/alkene and ω-olefin synthesis,
respectively (Hu et al. 2013; Mendez-Perez et al. 2011; Wang et al. 2013). Overexpression of aar/ado has led to doubled alkane production in Synechocystis sp. PCC 6803 (Wang et al. 2013). However, essentially no change in alkane production was seen in strains overexpressing aar or ado individually, suggesting that the genes work tightly in concert for alkane synthesis (Wang et al. 2013). In the α-olefin producing strain Synechococcus sp. PCC 7002 overexpression of ols under control of the constitutive psbA promoter (P_{psbA}) led to a 2-fold increase in 1-nonadecene and a 5-fold increase in 1,14-nonadecadiene (Mendez-Perez et al. 2011).

Though the AAR/ADO and OLS pathways appear to be mutually exclusive in cyanobacterial species, research to combine the pathways within one cyanobacterial species has been done (Yoshino et al. 2015). Synechococcus sp. NKBG15041c, which naturally produces α-olefins, was engineered to also produce alkanes/alkenes by introduction of the AAR/ADO pathway genes from Synechococcus elongatus PCC 7942. The engineered strain had low production of heptadecane even through expression levels of AAR and ADO were high. The decrease in native α-olefin production indicated that substrate flow of acyl-ACP was redirected from the OLS pathway to the AAR/ADO pathway. However, low heptadecane yields suggest inefficient enzyme activation or an unknown regulatory mechanism stunting production (Yoshino et al. 2015). Since ADO requires a reducing system, the low heptadecane yields may be due to an inefficient heterologous reducing system compared to the preferred endogenous system (Yoshino et al. 2015; Zhang et al. 2013). Thus, enzyme expression, enzyme activation, and reducing systems all pose targets for enhancing alkane production (Yoshino et al. 2015).
For $\alpha$-olefin synthesis, it may be possible to engineer or replace the loading domain (ACP) to accept a wider variety of substrates. The domain is highly substrate-specific, but through engineering, it may be able to accept different substrates, thus generating a larger variety of $\alpha$-olefin products, with regard to chain length (Lennen and Pfleger 2013). Improving alkane/alkene and olefin production rates while producing hydrocarbons with desired carbon-chain lengths through genetic engineering promises to generate biofuel sources able to successfully replace traditional fossil fuels.

Genetic engineering approaches have improved hydrocarbon production in cyanobacteria. Additionally, an understanding of the native process and physiological purpose of cyanobacterial hydrocarbon production may provide insights leading to further enhancement of hydrocarbon production.

1.2.8 Physiological role of alkanes/alkenes in cyanobacteria

Though biofuels compete with cell growth for carbon sources, biofuel molecules may benefit cell health and survival (Ruffing 2011). The ethanol biosynthesis pathway regenerates NAD$^+$ needed for glycolysis (Ruffing 2011). Ethylene acts as a hormone in plants to regulate growth and stress response (Pratt and Goeschl 1969; Yang and Hoffman 1984). Isoprene production fosters plant heat tolerance under heat stress (Lindberg et al. 2010). Biofuel products are often toxic to cyanobacteria in high concentrations, despite their physiological importance (Ruffing 2011). An understanding of the physiological role of hydrocarbons may provide insight for conditions to enhance their production while minimizing detrimental cellular effects.

Under normal growth conditions, cyanobacterial hydrocarbon production is very low. One study of hydrocarbon production revealed an average of only $0.11 \pm 0.015\%$
hydrocarbon content per percent biomass across the cyanobacterial species investigated (Coates et al. 2014). Though hydrocarbon production is universal among cyanobacteria, alkane/alkene production does not appear to be essential to survival under normal growth conditions: Schirmer et al. (2010) deleted ado/aaar in *Synechocystis* sp. PCC 6803 without an impact on cell growth and survival.

In cyanobacteria, lipids are generally localized exclusively to membranes (Singh et al. 2002). Membranes, as a structural foundation as well as a gateway between the cell interior and the external environment, are key targets for cell adaptation to environmental conditions (Singh et al. 2002). The wide distribution of cyanobacteria suggests they are able to adapt to many different environmental stressors, such as heat, cold, desiccation, salinity, nitrogen-starvation, high light, etc. (Singh et al. 2002). These factors are moreover hypothesized to influence FA and lipid content in cyanobacteria (Chen and Johns 1991; Huang et al. 2010; Jiang and Chen 2000; Singh et al. 2002). Changes in lipid structure and synthesis to maintain cellular integrity and viability provide an important acclimation response. Cyanobacteria naturally produce many compounds that aid in stress acclimation. Examples include photoprotective compounds such as mycosporine-like amino acids and scytonemin as well as DNA repair enzymes and heat shock proteins (Bhagwat and Apte 1989; Borbely and Suranyi 1988; Singh et al. 2002; Sinha and Häder 1996; Sinha and Hader 2002; Sinha et al. 1998; Sinha et al. 1999; Sinha et al. 2001). Due to their aid in membrane stability, structure, and fluidity, hydrocarbons have been added to this list as an adaptive mechanism to help cyanobacteria cope with environmental stress (Singh et al. 2002).
1.2.8.1 Temperature adaptation

FAs, important lipid constituents of membranes, are found in the cytoplasmic and thylakoid membranes of cyanobacteria. The level of FA unsaturation affects the functions of membrane-bound proteins and thus impacts the photochemical and electron-transport reactions occurring in the thylakoid and cytoplasmic cyanobacterial membranes (Los and Murata 2004). FA composition can also be related to temperature adaptation: thermophilic unicellular cyanobacteria tend to have monoenoic FAs while mesophilic or psychrophilic unicellular species will produce polyunsaturated FAs. Double bonds in FAs helps the latter adjust membrane fluidity for colder temperatures (Los and Murata 2004). FA length is also an important factor in temperature adaptation. For example, *Anabaena variabilis* inhibits 16:0 FA elongation while increasing monoenoic FA synthesis to increase membrane fluidity in response to a decrease in temperature (Maslova et al. 2004).

Membrane physiological properties are attributed largely to the FA composition of the lipids constituting the membrane (Singh et al. 2002). At lower temperatures, unsaturated fatty acids are key components of polar glycerolipids in the membrane (Singh et al. 2002). By increasing FA unsaturation levels under cold temperatures, cyanobacteria maintain membrane fluidity (Murata and Nishida 1987).

The intracellular thylakoid membrane systems of cyanobacteria universally contain C15-C19 hydrocarbons (Berla et al. 2015). Unlike algae, which accumulate lipids as storage materials under stress, cyanobacteria accumulate lipids in thylakoid membranes, likely in order to sustain photosynthesis and growth rate levels (Karataş and Dönmez 2011). The phase transition of thylakoid membranes that occurs as
cyanobacteria are transitioned to cold temperature is accompanied by a loss of photosynthetic activity (Murata and Omata 1988). An engineered mutant strain of *Synechocystis* sp. PCC 6803 that produced no alkanes grew poorly at cold temperatures (Berla et al. 2015). Upon investigation, researchers concluded that the lack of membrane alkanes led to higher levels of cyclic electron flow that reduced growth by forcing the cell to use energy-inefficient pathways and lowering the level of photosynthesis (Berla et al. 2015).

### 1.2.8.2 High light

Though light is required for photosynthesis, high-light (HL) can have harmful effects on cyanobacteria. Excess light energy captured by light-harvesting systems that is unable to be consumed in cellular processes creates reactive oxygen species (ROS) able to damage cellular components (Muramatsu and Hihara 2012; Ruffing 2011). Therefore, in response to HL, cyanobacteria downregulate the amount of photosystems and phycobilisomes (light-harvesting antenna complexes), thereby limiting the absorption of excess light (Muramatsu and Hihara 2012). Simultaneously, CO₂ fixation and other metabolic pathways are upregulated to consume excess light energy (Muramatsu and Hihara 2012).

In *Synechocystis* sp. PCC 6803, carbon assimilation is positively correlated with HL while nitrogen assimilation is negatively correlated. Carbon and nitrogen assimilation are a sink for the reducing power generated by the photosynthetic electron transport chain. Thus, the balance of carbon and nitrogen assimilation shifts in response to HL. An increase in carbon fixation is linked to increased activity of the CO₂ concentrating mechanism (CCM), which involves increased uptake of inorganic carbon sources
In *S. 6803*, inorganic carbon uptake by the CCM system uses light energy to drive two CO$_2$ and three HCO$_3^-$ transporters (Kaplan et al. 2008; Muramatsu and Hihara 2012; Price et al. 2008). Under HL, increased inorganic carbon uptake by the CCM results in an increased CO$_2$ concentration around Rubisco, thus enhancing carbon fixation (Kaplan et al. 2008; Muramatsu and Hihara 2012; Price et al. 2008).

Under HL conditions, cyanobacteria decrease the number of photosystems to avoid excess light energy absorption (Hihara et al. 1998; Murakami and Fujita 1991). In this situation, PSI is the photosystem selectively downregulated rather than PSII (Hihara et al. 1998; Murakami and Fujita 1991; Muramatsu and Hihara 2012). In contrast to the downregulation of PSI under HL, PSII turnover is accelerated in response to HL (Muramatsu and Hihara 2012). Under HL conditions, the reaction center of PSII, consisting of the D1 and D2 protein heterodimer and cytochrome $b_{559}$ protein, is damaged (Muramatsu and Hihara 2012; Nishiyama and Murata 2014). As part of the heightened PSII turnover under HL, PSII repair genes are upregulated to respond to the photodamage of PSII (Muramatsu and Hihara 2012). These repair genes primarily target the repair and replacement of the D1 protein (Muramatsu and Hihara 2012).

In addition to the photodamage PSII experiences under HL, PSII also becomes destabilized within the membrane, likely due to limitations on membrane lipid turnover (Takatani et al. 2015). Membrane lipid remodeling has been shown to be essential to adapting to HL in *Synechococcus elongatus* sp. PCC 7942 (Takatani et al. 2015). A mutant of *S. elongatus*, lacking the *aas* gene, exhibited higher photodamage of PSII under HL conditions compared to the wildtype containing the functional *aas* gene (Takatani et
AAS activates FFAs released from the membrane during membrane remodeling (Kaczmarzyk and Fulda 2010; Takatani et al. 2015). This activation step allows the FFAs to enter cellular pathways for lipid turnover or metabolism (Kaczmarzyk and Fulda 2010). In the aas-deficient mutant, FFAs accumulated within the cells and the cellular medium, indicating their lack of reentry into the membrane lipid recycling process (Takatani et al. 2015). In the mutant strain, PSII had accelerated photodamage due to reduced stability within the membrane, which resulted in hypersensitivity of the mutant strain to photoinhibition (Takatani et al. 2015).

1.2.8.3 Fixed nitrogen-deficient and high salt conditions

Research has shown that under fixed nitrogen-limiting conditions, microorganisms will accumulate lipids (Ratledge 2004). In the microalgae *Haematococcus pluvialis*, cellular lipid content reached 32.99% under fixed nitrogen-deplete conditions combined with high light, compared to a total lipid content of 15.61% in control cells (Damiani et al. 2010). In cyanobacteria, Kageyama et al. (2015) showed that fixed nitrogen-deficiency doubled intracellular alkane accumulation in diazotrophic *Anabaena* 7120. However, in the freshwater, non-diazotrophic cyanobacterium *S. elongatus* sp. PCC 7942, fixed nitrogen depletion resulted in a decrease in alkanes (Kageyama et al. 2015). Nonetheless, non-diazotrophic cyanobacteria have been shown to accumulate lipids when fixed nitrogen is present but limiting in the medium (Karatay and Dönmez 2011). NaNO₃ has been reported as the preferred nitrogen source for cellular lipid accumulation in microalgae (Damiani et al. 2010; Widjaja et al. 2009). When NaNO₃ was decreased to 0.25 g/L in the growth medium, the cyanobacteria *Synechococcus* sp., *Cyanobacterium aponinum*, and *Phormidium* sp. achieved their
highest lipid contents of 44.4%, 45%, and 38.2%, respectively (Karatay and Dönmez 2011).

Kageyama et al. (2015) studied the effects of fixed nitrogen deficiency and salt stress in three species of cyanobacteria. Fixed nitrogen-deficient conditions resulted in increased alkane production in the diazotrophic cyanobacterium *Anabaena* 7120; however alkane levels under these conditions were decreased in non-diazotrophic *Synechococcus elongatus* PCC 7942 (Kageyama et al. 2015). In *Anabaena* 7120 and the halotolerant cyanobacterium *Aphanothece halophytica*, salt stress enhanced alkane production. Moreover, when the alkane genes (*aar/ado*) from *A. halophytica* were transferred to *Anabaena* 7120, alkane production by *Anabaena* 7120 was further enhanced (Kageyama et al. 2015). Combined fixed nitrogen-deficient and high-salt conditions accelerated alkane doubling times in *Anabaena* 7120 (Kageyama et al. 2015). However, the combined stress upon *Anabaena* 7120 slowed the growth rate (Kageyama et al. 2015).

Though *Anabaena* 7120 demonstrated increased alkane accumulation under nitrogen and salt stress conditions, semi-quantitative RT-PCR results indicated that expression levels of the *AAR/ADO* alkane biosynthetic genes were not increased under the stress conditions (Kageyama et al. 2015). These results suggest that another mechanism other than transcriptional regulation of the alkane genes is responsible for alkane accumulation under abiotic stress conditions. Kageyama et al. (2015) propose that post-transcriptional regulation or mRNA stability of the alkane genes may have an important role. Alternatively, a precursor molecule to alkane biosynthesis, such as acyl-ACP, may be rate limiting (Kageyama et al. 2015).
The composition of alkanes (i.e., short-chain C₉-C₁₇ and long-chain C₁₇-C₃₁ alkanes) has been shown to vary in *Anabaena cylindrica* in response to salt stress (Bhadauriya et al. 2008). In *A. cylindrica* grown without salt stress, long-chain alkanes in the range of C₂₃-C₃₁ comprise the majority (~94%) of total alkanes (Bhadauriya et al. 2008). However, under salt stress, *A. cylindrica* shifts to production of short chain alkanes (C₉-C₁₇), which rise to a level of 43% of the total alkane pool, compared to a level of 0.2% in unstressed cells (Bhadauriya et al. 2008). This change in alkane composition may be involved in maintaining important cellular functions under salt stress in *A. cylindrica* (Bhadauriya et al. 2008). However, the change in alkane composition in response to stress is not universal amongst cyanobacteria. *Anabaena* 7120 placed under salt stress does not demonstrate a change in alkane composition: the main alkane produced remains to be heptadecane regardless of the presence or absence of the stress (Kageyama et al. 2015).

**1.2.8.4 Hydrocarbon cycle in natural ecosystems: Beneficial and antimicrobial effects**

Cyanobacteria in marine, freshwater, and terrestrial environments share conserved hydrocarbon biosynthetic pathways (Coates et al. 2014), indicating the role of hydrocarbons in cyanobacteria is independent of the ecosystem inhabited (Lea-Smith et al. 2015). In the marine environment, *Prochlorococcus* and *Synechococcus* are estimated to produce 308-771 million tons of hydrocarbons annually, which exceeds oceanic hydrocarbon input from natural oil seepages and anthropogenic sources (Lea-Smith et al. 2015). The predicted marine cyanobacterial hydrocarbon output of 2-540 pg·ml⁻¹·d⁻¹ is balanced in small part by atmospheric exchange, but largely through biological
degradation (Lea-Smith et al. 2015). Lea-Smith et al. (2015) showed that obligate and facultative hydrocarbon-degrading bacteria native to marine environments are able to metabolize cyanobacterial hydrocarbons. Moreover, cyanobacterial hydrocarbons were sufficient to sustain the growth of one species, *Alcanivorax borkumensis* SK2, an obligate-hydrocarbon degrading bacterium (Lea-Smith et al. 2015). These results reveal the interplay between cyanobacteria and hydrocarbon-degrading bacteria in a marine hydrocarbon cycle.

Hydrocarbons produced by cyanobacteria may be used as a carbon source by other bacteria sharing the habitat. However, they have also been shown to have antimicrobial properties. Volatile components, predominant among which is the alkane heptadecane (39.7%), produced by the cyanobacterium *Spirulina platensis* showed antimicrobial activity, especially against *Streptococcus faecalis*, *Staphylococcus epidermidis*, and *Candida albicans* (Ozdemir et al. 2004). Thus, effects of cyanobacterial alkanes on other organisms in the environment range from beneficial to harmful.

### 1.3 Nitrogen fixation in cyanobacteria

Nitrogen fixation is a globally important process. Nitrogen is a necessary element to all living things due to its presence in nucleic acids and proteins, which are needed to transfer genetic information and carry out life-giving functions of cells. While gaseous nitrogen (N\(_2\)) accounts for 70% of our atmosphere, it is in a non-bioavailable form for most living organisms. Thus, integration of atmospheric nitrogen into cellular building blocks requires an intermediate transfer molecule, such as nitrate or ammonia. Cyanobacteria can use different sources of nitrogen, including ammonium, nitrate, urea, and N\(_2\). Ammonium is the preferred nitrogen source while the least preferred is N\(_2\).
(Ohmori and Hattori 1974; Ohmori et al. 1977). Nitrogen-fixing organisms, such as some species of cyanobacteria, convert N\textsubscript{2} into ammonia, thus supplying a bioavailable form of nitrogen that can be used by non-diazotrophic organisms.

1.3.1 Mechanism of cyanobacterial photosynthetic oxic nitrogen fixation

Diverse bacteria and Archaea, but not Eukaryotes, are able to catalyze biological N\textsubscript{2} fixation (Böhme 1998; Zehr and Bombar 2015). Cyanobacteria are a heterogeneous group of organisms able to perform oxygenic photosynthesis (Giovannoni et al. 1988; Zehr and Bombar 2015). During photosynthesis, two photosystems, PSI and PSII, transfer electrons from water to the electron donors, ferredoxin and NADPH (DeRuyter and Fromme 2008). Water splitting occurs near the PSII center, releasing oxygen as a byproduct (Flores et al. 2015). To date, all cyanobacteria except for *Gloebacter violaceus* carry out photosynthesis in intracellular thylakoid membranes (Flores et al. 2015). Thus, photosynthetic cells produce and contain oxygen within the cell. This poses a problem for cyanobacterial species that also fix nitrogen. Nitrogenase, the key enzyme of nitrogen fixation, is extremely oxygen-sensitive (Fay 1992; Gallon 1981).

The nitrogenase reaction is as follows:

\[
N_2 + 8e^- + 8H^+ + 16 ATP \rightarrow H_2 + 2NH_3 + 16(ADP + Pi)
\]

The nitrogenase enzyme catalyzing the reaction is composed of two proteins: catalytic dinitrogenase (MoFe protein) and dinitrogenase reductase (Fe protein). The latter protein supplies electrons from reduced ferredoxin or flavodoxin to the catalytic dinitrogenase via an ATP-dependent mechanism (Masukawa et al. 2007). Homocitrate, synthesized by *nifV*-encoded homocitrate synthase, is normally bound to the catalytic
FeMo cofactor of nitrogenase and is required for efficient N\textsubscript{2} fixation (Masukawa et al. 2007).

1.3.1.1 Separation of photosynthesis and nitrogen fixation

Two main mechanisms exist to separate the incompatible processes of photosynthesis and nitrogen fixation in diazotrophic cyanobacteria: spatial separation and temporal separation (Figure 1.2). From these two mechanisms of separation, four overarching categories of diazotrophic cyanobacteria can be distinguished (Table 1.1). The first group consists of filamentous cyanobacteria, such as *Anabaena*, which form specialized cells, called heterocysts, in which nitrogen fixation occurs. The second group contains filamentous, non-heterocystous cyanobacteria, such as *Trichodesmium* and *Leptolyngbya*. The third group is comprised of unicellular nitrogen-fixing cyanobacteria, such as *Cyanothece* and *Glycothece*. The fourth group contains symbiotic nitrogen-fixing cyanobacteria, such as UCYN-A, *Calothrix*, and *Richelia*.
Figure 1.2 Four approaches for solar-powered oxic N₂-fixation by cyanobacteria (Young 2020). Unicellular free-living cyanobacteria are reported to separate nitrogen fixation and photosynthesis temporally to protect nitrogenase from oxygen (A). Filamentous cyanobacteria form either heterocysts (B) or diazocytes (C) to spatially separate nitrogenase from oxygen. Unicellular symbiotic cyanobacteria have reportedly lost the genes required to produce PSII, thus limiting oxygen production and localizing nitrogenase to a less oxygenic environment (D). In heterocysts PSI is functioning and can produce ATP for N₂-fixation through cyclic photophosphorylation (Magnuson 2019). Left image of D is taken from Harding et al. (2018). Note: The cell sizes in all left panels are not the actual size. Figure 1.2 taken from Young J, Gu, L., Gibbons, W., Zhou, R. (2020) Harnessing Solar-powered N₂-fixing Cyanobacteria for the BioNitrogen Economy In Press.

<table>
<thead>
<tr>
<th>Genus</th>
<th>Representative strains</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I-Filamentous Heterocystous</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anabaena</td>
<td>Anabaena sp. ATCC33047</td>
<td>(Stacey et al. 1977)</td>
</tr>
<tr>
<td></td>
<td><em>A. cylindrica</em> PCC 7122</td>
<td>(Attridge and Rowell 1997)</td>
</tr>
<tr>
<td></td>
<td><em>A. cylindrica</em> ATCC 29414</td>
<td>(Qiu et al. 2019; Zhu et al. 2013)</td>
</tr>
<tr>
<td></td>
<td><em>Anabaena variabilis</em> ATCC 29413</td>
<td>(Thiel et al. 1995)</td>
</tr>
<tr>
<td>Anabaenopsis</td>
<td>Anabaenopsis circularis</td>
<td>(Sarma 2013)</td>
</tr>
<tr>
<td>Nostoc</td>
<td>Nostoc sp. PCC7120</td>
<td>(Chen et al. 2016b; Wolk et al. 1984)</td>
</tr>
<tr>
<td></td>
<td><em>Nostoc punctiforme</em> ATCC 29133</td>
<td>(Meeks et al. 2001; Wong and Meeks 2001)</td>
</tr>
<tr>
<td>Calothrix</td>
<td>Calothrix <em>rhizosoleniae</em> SC01</td>
<td>(Hilton et al. 2016; Rippka et al. 1979)</td>
</tr>
<tr>
<td></td>
<td>Calothrix sp. PCC 7103</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calothrix sp. PCC 6303</td>
<td></td>
</tr>
<tr>
<td>Chlorogloeopsis</td>
<td>Chlorogloeopsis <em>fritschii</em> PCC 6912</td>
<td>(Hilton et al. 2016)</td>
</tr>
<tr>
<td></td>
<td>Chlorogloeopsis sp. PCC 9212</td>
<td></td>
</tr>
<tr>
<td>Cylindrospermopsis</td>
<td>Cylindrospermopsis <em>raciborskii</em> CS-505</td>
<td>(Hilton et al. 2016)</td>
</tr>
<tr>
<td></td>
<td>Cylindrospermum stagnale* PCC 7417</td>
<td></td>
</tr>
<tr>
<td>Fischerella</td>
<td>Fischerella <em>muscicola</em> PCC 73103</td>
<td>(Hilton et al. 2016)</td>
</tr>
<tr>
<td></td>
<td>Fischerella sp. PCC 9339</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fischerella <em>thermalis</em> PCC 7521</td>
<td></td>
</tr>
<tr>
<td>Mastigocladopsis</td>
<td>Mastigocladopsis <em>repens</em> PCC 10914</td>
<td>(Hilton et al. 2016)</td>
</tr>
<tr>
<td>Mastigocoleus</td>
<td>Mastigocoleus <em>testarum</em> BC008</td>
<td>(Hilton et al. 2016)</td>
</tr>
<tr>
<td>Microchaete</td>
<td>Microchaete sp. PCC 7126</td>
<td>(Hilton et al. 2016)</td>
</tr>
<tr>
<td>Organism</td>
<td>Scientific Name</td>
<td>Reference</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Nodularia</td>
<td>Nodularia spumigena CCY9414</td>
<td>Hilton et al. 2016</td>
</tr>
<tr>
<td>Richelia</td>
<td>Richelia intracellularis HH01</td>
<td>Hilton et al. 2016</td>
</tr>
<tr>
<td>Rivularia</td>
<td>Rivularia sp. PCC 7116</td>
<td>Hilton et al. 2016</td>
</tr>
<tr>
<td>Scytonema</td>
<td>Scytonema hofmanni PCC 7110</td>
<td>Hilton et al. 2016</td>
</tr>
<tr>
<td></td>
<td>Scytonema hofmanni UTEX 2349</td>
<td></td>
</tr>
<tr>
<td><strong>II-Filamentous Non-Heterocystous</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leptolyngbya</td>
<td>Leptolyngbya boryana dg5</td>
<td>Tsujimoto et al. 2014</td>
</tr>
<tr>
<td>Microcoleus</td>
<td>Microcoleus sp. PCC7113</td>
<td>Thiel 2019</td>
</tr>
<tr>
<td>Oscillatoria</td>
<td>Oscillatoria sp. strain 23</td>
<td>Rippka et al. 1979; Stal and Krumbein 1987a; Stal and Krumbein 1987b; Thiel 2019</td>
</tr>
<tr>
<td></td>
<td>Oscillatoria limosa</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oscillatoria sp. PCC1082</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oscillatoria sp. PCC6506</td>
<td></td>
</tr>
<tr>
<td>Plectonema</td>
<td>Plectonema boryanum strain 594</td>
<td>Bergman et al. 1997; Stewart and Lex 1970</td>
</tr>
<tr>
<td></td>
<td>Plectonema boryanum PCC73110</td>
<td></td>
</tr>
<tr>
<td>Pseudanabaena</td>
<td>Pseudanabaena sp. PCC6802</td>
<td>Bergman et al. 1997; Thiel 2019</td>
</tr>
<tr>
<td></td>
<td>Pseudanabaena sp. PCC7409</td>
<td></td>
</tr>
<tr>
<td>Symploca</td>
<td>Microcoleus chthonoplastes</td>
<td>Bergman et al. 1997</td>
</tr>
<tr>
<td>Trichodesmium</td>
<td>Trichodesmium erythraeum IMS101</td>
<td>Bergman et al. 2013</td>
</tr>
<tr>
<td><strong>III-Unicellular</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aphanocapsa</td>
<td>Aphanocapsa montana BDHKU 210001</td>
<td>Bhattacharyya et al. 2015</td>
</tr>
<tr>
<td>Chroococcidiopsis</td>
<td>Chroococcidiopsis thermalis PCC7203</td>
<td>Thiel 2019</td>
</tr>
<tr>
<td>cyanathece</td>
<td>Cyanothece ATCC 51142</td>
<td>Liberton et al. 2019</td>
</tr>
<tr>
<td></td>
<td>Cyanothece sp. PCC7424</td>
<td>Thiel 2019</td>
</tr>
<tr>
<td></td>
<td>Cyanothece sp. PCC7425</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cyanothece sp. PCC7822</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cyanothece sp. PCC8801</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cyanothece sp. WH 8904</td>
<td>Mazard et al. 2004</td>
</tr>
<tr>
<td>Gloeocapsa</td>
<td>Gloeocapsa sp. PCC7428</td>
<td>Thiel 2019</td>
</tr>
<tr>
<td>Gloeocapsa</td>
<td>Gloeocapsa sp. PCC7428</td>
<td>Thiel 2019</td>
</tr>
<tr>
<td>Crocosphaera</td>
<td>C. watsonii WH8501</td>
<td>Bench et al. 2011</td>
</tr>
<tr>
<td>Pleurocapsa</td>
<td>Pleurocapsa sp. PCC7327</td>
<td>Thiel 2019</td>
</tr>
<tr>
<td><strong>IV-Symbiotic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCYN-A</td>
<td>UCYN-A (unicellular)</td>
<td>Thompson et al. 2012</td>
</tr>
<tr>
<td>Nostoc</td>
<td>N. azollae 0708 (heterocystous)</td>
<td>Ran et al. 2010</td>
</tr>
</tbody>
</table>
1.3.2 Filamentous, heterocystous cyanobacteria [key organisms: *Anabaena 7120, Nostoc punctiforme, Anabaena variabilis, Anabaena cylindrical*]

Certain cyanobacteria use spatial separation in the form of specially differentiated cells, called heterocysts, whose sole purpose is to carry out nitrogen fixation. Heterocyst-forming cyanobacteria are contained within a single phylogeny, which suggests that evolutionarily, heterocysts only developed once (Flores and Herrero 2010). Research has shown that heterocysts contain most, even all, of the nitrogenase enzyme within the cells of the diazotrophic filament (Herrero et al. 2013).

In nitrate-limiting conditions, 5-10% of cells in the filament differentiate to form heterocyst (Omairi-Nasser et al. 2014). Heterocysts are terminally differentiated, specialized cells for nitrogen fixation, where photosynthesis does not occur (Maldener and Forchhammer 2015). Commitment to heterocyst differentiation occurs approximately 8 h after the start of nitrogen deprivation, and differentiation is complete at about 24 h following nitrogen depletion (Yoon and Golden 2001). The differentiation process involves the cessation of photosynthetic oxygen evolution and the expression of N\textsubscript{2} fixation machinery in the cells that ultimately form heterocysts (Flores et al. 2015). The remaining vegetative cells in the filament retain their ability to perform photosynthesis.

1.3.2.1 Carbon-nitrogen transport within the filament: cell connections and communication

Heterocysts and vegetative cells must communicate and exchange materials within the filament: namely, vegetative cells require fixed nitrogen from heterocysts.
while heterocysts obtain fixed carbon from the vegetative cells (Flores et al. 2015; Omairi-Nasser et al. 2014). Fixed nitrogen is likely transported in the form of amino acids such as glutamine as well as cyanophycin-related metabolites (e.g., aspartate, arginine, and the aspartyl-arginine dipeptide) (Burnat et al. 2014; Herrero et al. 2013; Wolk et al. 1994). Vegetative cells also use metabolites to transfer fixed carbon to heterocysts. Sucrose (Lopez-Igual et al. 2010; Vargas et al. 2011) and alanine (Pernil et al. 2010) are likely used to transport reductant and energy sources. Sucrose and carbohydrates transported to heterocysts are metabolized via the oxidative pentose phosphate (OPP) pathway (Magnuson and Cardona 2016). The OPP pathway in heterocysts has been shown to generate reductants required for nitrogenase activity (Magnuson and Cardona 2016; Winkenbach and Wolk 1973).

Two potential mechanisms of transport between heterocysts and vegetative cells have been identified. The first possibility is transport through the periplasmic space. Cyanobacteria contain an outer membrane beyond the cytoplasmic membrane and peptidoglycan layer. This outer membrane is continuous along all cells of the filament, and creates a periplasmic space that may represent a communication channel amongst the cells (Mariscal et al. 2007). The second mechanism is through channels that directly connect cells in the filament (Omairi-Nasser et al. 2014). Channels have been identified between filamentous cells, and they are believed to allow cells within the filament to communicate directly through intercellular septa (Mullineaux et al. 2008; Omairi-Nasser et al. 2014). In the heterocystous cyanobacterium, *Anabaena* 7120, proteins have been identified that appear to be part of these intercellular septa. These proteins are SepJ
(Mariscal et al. 2011) and FraCD (Merino-Puerto et al. 2011), and each is proposed to form a different type of channel between cells.

The specialized junctions forming these channels develop in the very early stages of heterocyst development (Omairi-Nasser et al. 2014). These communication channels between heterocysts and vegetative cells are formed, in part, by Fra proteins (Omairi-Nasser et al. 2015). FraC, FraD, and FraG have all been located in the septum between cells in the *Anabaena* filament (Flores et al. 2007; Merino-Puerto et al. 2010). In *fraC* and *fraG* mutants, channels are still able to form; however, there are fewer of them. These results suggest that FraC and FraG are part of the channel structure or affect channel assembly; FraC and FraG may also be involved in the assembly of different channels (Omairi-Nasser et al. 2015). FraD has a smaller impact on channel formation and is postulated to play a role in maintaining stable cell-to-cell contact in the septum (Omairi-Nasser et al. 2015).

Though both the periplasmic space as well as intercellular channels are have been proposed as mechanisms of nutrient transport and communication between cells in the heterocystous filament, it still remains to be determined which molecules are transported through the periplasmic space or through the SepJ and FraCD proteinaceous channels (Flores et al. 2015).

**1.3.2.2 Heterocyst differentiation**

The differentiation process producing heterocysts results in morphological and biochemical changes to create micro-oxic conditions within heterocysts for N₂ fixation to occur (Flores et al. 2015). Heterocysts gain a cell envelope to act as a barrier to keep oxygen from entering the cytoplasm (Böhme 1998). The inner layer of this envelope
consists of specific glycolipids and the outer layer consists of specific polysaccharides (Flores and Herrero 2010; Walsby 2007). In mutants in which the synthesis or deposition of either of these layers is impaired, N₂ fixation is inhibited, though it is regained under anoxic conditions (Herrero et al. 2013).

In the very early stages of development, thylakoid membranes reorganize and an additional membrane develops, forming honeycomb-like structures located close to the poles of the cell (Magnuson and Cardona 2016; Muro-Pastor and Hess 2012). Respiratory enzymes, such as heterocyst-specific terminal oxidases, are found in the honeycomb regions at the polar ends of the heterocyst to reduce oxygen within the heterocyst (Magnuson and Cardona 2016; Muro-Pastor and Hess 2012). Thus, the honeycomb membranes are presumed to have the function of consuming molecular O₂ that enters heterocysts near their polar regions (Magnuson and Cardona 2016).

### 1.3.2.3 Heterocyst envelope

In addition to limiting oxygen production by not housing photosynthetic reactions, heterocysts also undergo structural remodeling to support the maintenance of the micro-oxic environment required for N₂ fixation. Specifically, the heterocyst cell wall undergoes reconstruction to form a special envelope required by heterocysts (Maldener and Forchhammer 2015). The heterocyst envelope consists of an inner glycolipid layer and an outer polysaccharide layer, both of which act as a barrier to keep oxygen outside of the cell (Wolk et al. 1994). The envelope polysaccharide layer is required for nitrogen fixation when oxygen is present in the external environment (Huang et al. 2005). The extracellular layers of glycolipids and polysaccharides accumulated during heterocyst
differentiation provide a barrier to prevent O$_2$ entry into the cell (Murry and Wolk 1989; Walsby 1985).

### 1.3.2.4 Oxygen scavenging within heterocysts

Heterocysts lack the oxygen-evolving photosynthetic activity, exhibited by lower levels of PSII (Ferimazova et al. 2013; Flores et al. 2015). Heterocysts do maintain photosynthetic activity based on PSI, which does not produce oxygen. Heterocysts also lack ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), and thus do not perform photosynthetic CO$_2$ fixation. In addition, expression of Calvin cycle genes, such as $rbcLX$ and $prk$, are down-regulated in heterocysts, which is further consistent with the loss of CO$_2$ fixation in heterocysts (Wolk et al. 1994). As a result, the energy in the form of reductants and ATP that would be directed toward carbon fixation is able to be shunted into nitrogen fixation (Flores et al. 2015).

Heterocysts also have increased amounts of respiratory activity, which further lowers the O$_2$ levels in the cells (Böhme 1998; Flores et al. 2015). Increased respiration in heterocysts is related to sugar metabolism within the heterocysts, such as metabolism of sucrose and other sugars obtained from vegetative cells (Flores et al. 2015). Increased sugar metabolism in heterocysts is supported by increased expression of catabolic sugar enzymes, such as the invertase InvB (Lopez-Igual et al. 2010; Vargas et al. 2011) as well as OPP pathway enzymes (Herrero et al. 2013).

Heterocysts contain oxidases, which consume oxygen, thus providing further protection for the nitrogenase complex against oxygen (Wolk et al. 1994). While the presence of oxygen does not seem to affect the expression of $nif$ genes in *Anabaena* 7120, nitrogenase activity under oxic conditions is greatly impaired (Valladares et al. 2007).
*Anabaena* 7120 contains three gene clusters presumed to encode cytochrome \( c \) oxidases (Valladares et al. 2003). Two *cox* operons, *cox2* and *cox3*, are expressed following nitrogen stepdown in a NtcA- and HetR-dependent manner; these two operons encode heterocyst-specific terminal oxidases (Valladares et al. 2003; Valladares et al. 2007). In mutants of *cox2* and/or *cox3* genes, nitrogenase activity was impaired in the single *cox2* mutant and completely abolished in the double *cox2cox3* mutant (Jones and Haselkorn 2002; Valladares et al. 2007). Thus, oxidases are required for normal heterocyst function through protecting \( N_2 \) machinery from oxygen. In addition to maintaining a micro-oxic environment within heterocysts, cytochrome \( c \) oxidases generate ATP, which supports nitrogenase activity (Golden and Yoon 2003).

**1.3.2.5 Heterocyst patterning signals**

Heterocyst pattern formation in *Anabaena* is regulated by the peptide-derived signals, PatS and HetN (Videau et al. 2015). PatS and HetN act by preventing cells adjacent to heterocysts from differentiating into heterocysts (Videau et al. 2015). PatS is responsible for developing the initial heterocyst pattern; once this pattern is established, HetN acts in a similar manner to PatS to maintain patterning (Omairi-Nasser et al. 2014).

The genes *patS* and *hetN* encode putative peptides containing the sequence RGSGR, which is required for the peptides’ inhibitory activity (Corrales-Guerrero et al. 2013; Higa et al. 2012; Yoon and Golden 1998). The PatS and HetN peptides move laterally from source cell to neighboring cells where they bind to HetR through the conserved RGSGR sequence, thereby preventing HetR from binding to DNA and promoting HetR’s degradation (Corrales-Guerrero et al. 2013; Feldmann et al. 2011; Higa et al. 2012; Huang et al. 2004; Risser and Callahan 2009; Wu et al. 2004; Yoon and
Golden 1998). HetR is a key transcription factor that promotes heterocyst differentiation. This lateral inhibition of differentiation via the transfer of PatS and HetR between cells helps to maintain the pattern of a single, terminally differentiated heterocyst separated by approximately 10 vegetative cells (Videau et al. 2015). As the vegetative filament lengthens by growth and addition of new vegetative cells, heterocyst differentiation must also occur, placing new heterocysts between the preexisting heterocysts and maintaining the 1 heterocyst : 10 vegetative cells ratio (Videau et al. 2015).

The full-length 17 amino acid peptide encoded by *patS* is not believed to be transferred from cell to cell; rather, a smaller C-terminal peptide cleavage product of PatS is believed to be transferred along the filament (Corrales-Guerrero et al. 2013; Feldmann et al. 2012; Higa et al. 2012). Indeed, the size of the channels connecting cells within the filament limit the size of the peptide transferred. The entire PatS and HetN proteins are not likely to make it through these channels, thus a smaller peptide for both, containing the RGSGR sequence, is the likely vehicle to transport the signal inhibiting heterocyst development (Corrales-Guerrero et al. 2014; Omairi-Nasser et al. 2014). PatS and HetN transport from heterocysts to vegetative cells through these channels requires HetC, an ATP-binding cassette transporter (Videau et al. 2015). Interestingly, HetC is not required for transport of PatS and HetN between vegetative cells. These results suggest that once PatS and HetN are transported from heterocyst to adjacent vegetative cell via HetC, the inhibitors then move along the vegetative cells in the filament via the periplasm (Videau et al. 2015). How PatS and HetN pass from the periplasm across the cytoplasmic membrane into vegetative cells where they bind to and inhibit HetR remains unknown.
The membrane protein PatN has also been observed to affect the pattern of heterocyst differentiation (Risser et al. 2012). In *Nostoc punctiforme*, deletion of *patN* resulted in increased expression of *patA*, another gene known to affect heterocyst patterning, while also increasing the frequency of heterocysts within the filament (Risser et al. 2012). Thus, PatN is proposed to function by modulating PatA levels, and through doing so prevent vegetative cells from undergoing heterocyst differentiation (Risser et al. 2012).

### 1.3.2.6 Gene expression during heterocyst differentiation

Heterocyst differentiation, including the conferred structural and metabolic changes, is due to changes in gene expression during heterocyst development (Flores et al. 2015). The changes in gene expression are orchestrated in a time-specific manner in the differentiating cells (Herrero et al. 2013). The serial changes in gene expression are controlled by two key transcriptional regulators: NtcA, the global regulator (Herrero et al. 2004; Zhao et al. 2010), and HetR (Buikema and Haselkorn 1991; Kim et al. 2011). The differentiation process is triggered by combined nitrogen scarcity.

Upon combined nitrogen depletion, a series of steps must occur for heterocysts to develop within the filament at a spacing of approximately 1 heterocyst for every 10 cells (Golden and Yoon 2003). Research has indicated that cells destined to become heterocysts are selected early in the during the induction process or even before nitrogen deprivation (Toyoshima et al. 2010). HetR is the master regulator of heterocyst differentiation (Black et al. 1993; Buikema and Haselkorn 1991), and HetR positively regulates its own expression by binding to its promoter (Black et al. 1993). NtcA, a global transcriptional regulator, (Flores and Herrero 2010; Frias et al. 1994; Wei et al.
1994) and NrrA (Cai and Wolk 1997; Ehira and Ohmori 2006b) are also important components of the signaling pathway leading to heterocyst development.

In the absence of combined nitrogen, the carbon-to-nitrogen balance in the cells shifts to a high C:N ratio. When fixed nitrogen is low, cyanobacteria perceive the high C:N status through an increase in 2-oxoglutarate concentration (Muro-Pastor et al. 2001).

NtcA perceives a high C:N ratio, signaled by increased levels of 2-oxoglutarate (Valladares et al. 2008; Zhao et al. 2010). Though the mechanism is not yet clear, 2-oxoglutarate is believed to enhance the activity of NtcA, thus increasing expression of NtcA-activated nitrogen fixation genes (Herrero et al. 2001; Muro-Pastor et al. 2001). NtcA increases the expression of hetR through the transcriptional factor NrrA (Ehira and Ohmori 2006a). Expression of hetR further increases expression of ntcA. By 12 h post combined nitrogen deficiency, ntcA expression is localized to differentiating cells (Olmedo-Verd et al. 2006). HetR, whose expression is increased by itself and NtcA, binds to the promoter regions of heterocyst differentiation genes hepA, patS, hetP, and hetZ (Du et al. 2012; Higa and Callahan 2010; Huang et al. 2004).

Oxygen levels within heterocysts are also a potential regulator of gene expression since nitrogenase activity requires a low partial pressure of O₂ within the heterocyst (Bothe et al. 2010a; Ermakova et al. 2014). The flavodiiron protein, Flv3B, is responsible for light-induced O₂ uptake in heterocysts (Ermakova et al. 2014). In a Flv3B mutant strain of A. 7120, growth rate and the nitrogenase subunit NifH levels decreased significantly compared to wildtype (Ermakova et al. 2014). In addition, the transcript and protein levels of various genes in the mutant, including uptake hydrogenase, phycobilisome components, and heterocyst-specific ferredoxin, also differed from those
in the wildtype, suggesting that oxygen levels may be related to gene expression control in heterocysts (Ermakova et al. 2014).

In the final steps of heterocyst development, gene rearrangements must occur within the genes *nifD*, *hupL*, and *fdxN*, which encode a nitrogenase subunit, a hydrogenase uptake subunit, and a ferredoxin, respectively (Haselkorn 1992; Muro-Pastor and Hess 2012). These gene rearrangements allow the genes downstream of the interrupting elements to be transcribed, and thus expressed in heterocysts (Haselkorn 1992).

1.3.3 Filamentous, non-heterocystous cyanobacteria [key organisms: *Trichodesmium, Leptolyngbya*]

1.3.3.1 *Trichodesmium*

*Trichodesmium* is non-heterocyst-forming, filamentous cyanobacterium. Surprisingly, it performs N₂ fixation only in the light (Capone et al. 1997). The restriction of nitrogen fixation to the light period is elucidated by the degradation of nitrogenase at the end of the light period (Capone et al. 1990). Although *Trichodesmium* does not form heterocysts, physiological and immunological studies suggest spatial variability in nitrogenase distribution and activity in cells. Within the *Trichodesmium* filament, nitrogenase expression is localized within central zones where photosynthetic activity and CO₂ fixation appear to be limited (Bergman et al. 2013). The zones of nitrogen fixation have been termed “diazocytes” (Bergman et al. 1993). Diazocytes have lighter pigmentation compared to other cells in the filament. Additionally, NifH (nitrogenase protein) and cytochrome c oxidase (respiratory protein) have been immunolocalized at increased levels in diazocytes (Bergman et al. 1993). The frequency of diazocytes
changes within diel cycles, where it is lower at dawn and increases toward noon (Bergman et al. 2013; Sandh et al. 2009). However, it remains uncertain whether diazocytes are differentiated cells or simply cells that are transiently used for N₂ fixation (Bergman et al. 2013; Herrero et al. 2008).

Diazocyte patterning mirrors the beginning developmental steps of heterocysts in some filamentous cyanobacteria, such as *Anabaena* sp. (Bergman et al. 2013). In early states of heterocyst development, often subsets of adjacent pro-heterocysts form; however, ultimately, only the central cell develops into a fully differentiated heterocyst. Diazocytes, which occur in adjacent cells within small zones of the filament, reflect the patterning of pro-heterocysts. Thus, diazocytes have been described as representing a “frozen” earlier stage of heterocyst development, perhaps reflecting that diazocytes developed at an earlier evolutionary stage predating heterocysts (Bergman et al. 2013). In contrast to heterocysts, which are terminally differentiated, diazocytes retain the ability to divide (Bergman et al. 2013).

While *Trichodesmium* uses spatial separation to perform photosynthesis and nitrogen fixation, it also temporally segregates the two processes during the day. Short temporal variations in PSII activity have been reported that may help minimize the damaging effects of O₂ produced from photosynthesis (Berman-Frank et al. 2001; Küpper et al. 2004). Other research has also suggested a brief temporal segregation of the two processes, with photosynthesis peaking in the morning and N₂ fixation in the afternoon (Finzi-Hart et al. 2009). During the midday period, nitrogenase is protected from oxygen due to decreased oxygen production as well as increased oxygen scavenging, facilitated by respiration and the Mehler reaction (Bergman et al. 2013). The temporal regulation of
nitrogen fixation is further corroborated by the circadian control observed for nitrogenase transcript abundance, which follows a 24-h oscillation pattern (Chen et al. 1998)

1.3.3.2 *Leptolyngbya boryana*

*Leptolyngbya boryana* (formerly *Plectonema boryanum*) is a filamentous, non-heterocystous cyanobacterium capable of nitrogen fixation under micro-oxic conditions (Stewart and Lex 1970). Both micro-oxic conditions and nitrogen starvation appear to be necessary for nitrogen fixation in *L. boryana* (Finzi-Hart et al. 2009; Nagatani and Haselkorn 1978; Weare and Benemann 1974). Nitrogen fixation and oxygen-evolving photosynthesis are separated temporally (Misra and Tuli 2000). The alternating cycles of nitrogen fixation and photosynthesis observed in *L. boryana* when cultures are grown under continuous light suggest that circadian control plays a role in regulating these processes (Misra and Tuli 1993; Misra and Tuli 2000). Prior to the onset of nitrogen fixation, photosynthetic rates decrease significantly, which helps establish the micro-oxic environment required for nitrogenase activity (Misra and Tuli 2000). The decrease in photosynthetic activity is accompanied by decreased transcript levels of several photosynthesis-related proteins, such as *cpcBA*, *psbA*, and *psbC* (Misra and Tuli 2000). The nitrogen fixation stage is characterized by a decrease in dissolved oxygen; an increase in the expression of nitrogen fixation genes, namely *nifH*; and nitrogenase activity (Misra and Tuli 2000).

A 50-kb nitrogen fixation (*nif*) gene cluster in *L. boryana* contains many of the key nitrogen fixation genes (Tsujimoto et al. 2014). Fourteen of the 50 genes in this cluster are *nif* genes (*nifBSUHDKVZT* and *nifPENXW*) while the remaining genes encode *nif*-related proteins, transcriptional regulators, cytochrome oxidase subunits, and products
of unknown function (Tsujimoto et al. 2014). Two transcriptional regulators within the cluster were identified as ChlR and CnfR, both of which are essential for diazotrophic growth (Tsujimoto et al. 2014). ChlR activates genes responsible for anaerobic chlorophyll biosynthesis. Through doing so, ChlR supports energy production required for nitrogen fixation (Tsujimoto et al. 2014). CnfR appears to be a key nitrogen fixation regulator. It activates nif gene expression when oxygen concentrations are low (Tsujimoto et al. 2014). In a CnfR mutant, transcripts of the nif genes were entirely absent and nitrogen fixation did not occur (Tsujimoto et al. 2014). Thus, CnfR is proposed to be a master transcriptional activator for nitrogen fixation (Tsujimoto et al. 2014; Tsujimoto et al. 2016). CnfR recognizes a conserved cis-acting element, which is conserved in regions upstream of the nif genes in all nitrogen-fixing cyanobacteria (Tsujimoto et al. 2016). This finding suggests that CnfR has a widespread role in regulating nif gene expression in diazotrophic cyanobacteria (Tsujimoto et al. 2016).

1.3.4 Unicellular cyanobacteria [key organisms: *Cyanothece, Glycothece*]

1.3.4.1 *Cyanothece*

*Cyanothece* is an aerobic, unicellular, diazotrophic cyanobacterium. It separates nitrogen fixation from oxygen-producing photosynthesis via a diurnal cycle: photosynthesis occurs in the day (light period) while nitrogen fixation occurs in the night (dark period) (Aryal et al. 2011; Colon-Lopez et al. 1997; Flores et al. 2015; Reddy et al. 1993; Toepel et al. 2008). Gene expression studies have helped solidify the concept of diurnal cycling in *Cyanothece*. Studies have shown that photosynthesis, CO$_2$ fixation, ATP synthase, and glycogen biosynthesis genes are expressed in the light period while glycogen degradation, respiration, and nif genes are expressed upon entering the dark
period or within the dark period (Aryal et al. 2011; Colon-Lopez et al. 1997; Toepel et al. 2008). The \textit{nifHDK} operon, which encodes components of the nitrogenase complex, is under especially tight regulation, being transcribed within the first four hours of the dark period (Sherman et al. 1998). By the end of the dark period, nitrogenase is proteolytically degraded, which further ensures that nitrogen fixation is limited to the dark, non-photosynthetic period (Sherman et al. 1998).

Interestingly, slight variations exist in the expression of nitrogenase amongst different species of \textit{Cyanothece} in a study where five members of the \textit{Cyanothece} genus were grown under 12-h light/12-h dark cycles (Bandyopadhyay et al. 2013). In \textit{Cyanothece} sp. ATCC 51142, nitrogenase activity was induced in the later part of the light period, approximately 2-6 h before the beginning of the dark period (Bandyopadhyay et al. 2013). In the four other strains of \textit{Cyanothece} (\textit{Cyanothece} sp. PCC 7427, 7822, 8801, and 8802) nitrogenase activity was induced at the beginning or middle of the dark phase. Nevertheless, regardless of whether induction occurred prior to or during the dark phase, nitrogenase activity peaked during the dark period for all strains observed (Bandyopadhyay et al. 2013).

When \textit{Cyanothece} is grown under continuous light, the rhythmicity of photosynthetic and nitrogen fixation genes appears to be maintained (Colon-Lopez et al. 1997; Huang et al. 1999; Mitsui et al. 1986). When the length of light and dark periods was shortened to 6 h for each period, nitrogen fixation occurred once every second dark period and only once in per 24 h timespan (Toepel et al. 2009). Temporal separation of photosynthesis and nitrogen fixation has been correlated to control by the circadian rhythm (Cerveny and Nedbal 2009; Flores et al. 2015). In particular, the consumption of
glycogen granules toward the end of the light period along with a burst in respiratory activity is proposed to be controlled by the circadian clock (Cerveny and Nedbal 2009). The catabolic activity from glycogen degradation, which occurs during the respiratory peak contributes to the microoxic environment needed for nitrogen fixation (Cerveny and Nedbal 2009). Thus, respiration is believed to be controlled by the circadian clock (Cerveny and Nedbal 2009). This respiratory burst likely affects the onset of nitrogen fixation (Cerveny and Nedbal 2009).

Due to the separation of nitrogen fixation and photosynthesis in separate stages, the products of each must be stored for the cell to access them during the stage in which synthesis is not occurring. That is, fixed carbon produced during the light phase must be accessible to the cell during the dark stage when nitrogen fixation occurs and photosynthesis is temporarily halted. The same holds true for the availability of fixed nitrogen produced during the dark phase and still needed during the light phase.

_Cyanothece_ resolves this problem through storage granules. While the carbon stores produced from photosynthesis are used to support N\textsubscript{2} fixation in the dark, fixed nitrogen produced during the dark period is stored as cyanophycin, which is subsequently used as the nitrogen source during the light period (Li et al. 2001; Reddy et al. 1993; Sherman et al. 1998). Cyanophycin accumulates in the cell soon after nitrogen fixation peaks during the dark period (Li et al. 2001). _Cyanothece_ creates storage inclusion bodies to hold the products of nitrogen fixation and photosynthesis (Welsh et al. 2008). These inclusion bodies allow products of photosynthesis and nitrogen fixation to be stored during the stage of their respective synthesis, ultimately to be degraded and used during the stage in which their synthesis is inactive.
Nitrogen fixation in the dark period relies on the energy status of the cells. Glycogen granules formed from fixed carbon during the photosynthetic stage is broken down and used as substrate for the respiratory burst that occurs prior to and during the dark phase (Toepel et al. 2009). Glycogen degradation occurs via glycolysis, the OPP pathway, and the TCA cycle. The upregulation of these cycles during the dark period, through increased transcription of the pathway genes involved, produces ATP, pyridine nucleotides, and cellular intermediates to support nitrogen fixation (Stockel et al. 2008). High respiration rates are needed to produce the energy required for N\textsubscript{2} fixation and create an anoxic environment for nitrogenase (Toepel et al. 2009). Approximately one hour before the onset of the dark phase, \textit{Cyanothece} shifts from photosynthesis to respiration. The respiratory burst has been observed to last about 3-4 h, and provides the anoxic environment and energy to support nitrogen fixation (Krishnakumar et al. 2013).

\textbf{1.3.4.2 Gloeothece spp.}

In \textit{Gloeothece}, nitrogen fixation follows a diurnal cycle; nitrogenase is synthesized at the end of each light phase and degraded at the end of each dark phase (Du and Gallon 1993; Gallon et al. 1988). The circadian rhythm is the central mechanism regulating nitrogen fixation (Taniuchi et al. 2008b). When \textit{Gloeothece} is grown under anoxic conditions, however, it is able to fix nitrogen during the light phase (Du and Gallon 1993; Gallon and Stal 1992). Nitrogen fixation in \textit{Gloeothece} is dependent upon ATP and reductant pools, which support nitrogenase activity (Bergman et al. 1997). ATP and reductant are generated through catabolism of carbohydrate granules, most significant among which are glycogen granules (Bergman et al. 1997). Carbohydrate degradation is believed to occur via the OPP pathway, similar to the carbohydrate...
metabolism mechanisms used to support nitrogen fixation in heterocystous cyanobacteria (Bergman et al. 1997). Also like heterocystous strains, Gloeothecce cells are surrounded by an extracellular polysaccharide sheath; this coat may serve as a barrier to limit the influx of oxygen into the cell, thereby further protecting nitrogenase from O₂ (Bergman et al. 1997).

In non-heterocystous diazotrophic cyanobacteria, such as Gloeothecce and Cyanothece, cellular respiration plays an important role in limiting the amount of O₂ in the cell. To this end, respiration often follows a similar pattern that parallels the activity of nitrogen fixation (Bergman et al. 1997; R. Gallon and Falah Hamadi 2013)

1.3.5 Symbiotic nitrogen-fixing cyanobacteria [key organisms: UCYN-A, Calothrix, Richelia]

1.3.5.1 UCYN-A

UCYN-A is a unicellular, diazotrophic cyanobacterium closely related to Cyanothece (Zehr and Bombar 2015). Nitrogen fixation by UCYN-A makes a major contribution to marine nitrogen input (Bothe et al. 2010b). UCYN-A presumably fix nitrogen in the light as the $nifH$ gene transcript levels peak during the daytime (Zehr et al. 2007).

These diazotrophic marine cyanobacteria contain a small, stream-lined genome of only 1.44 Mb. Their downsized genome lacks many metabolic pathways, such as CO₂ fixation and several amino acid and purine synthesis pathways (Bothe et al. 2010b). These unicellular cyanobacteria have also lost PSII activity (Bothe et al. 2010b). The absence of functional PSII in UCYN-A means that these organisms are not exposed to photosynthetically produced oxygen, which protects nitrogenase from O₂ inactivation.
Because UCYN-A is unable to synthesize basic cellular building blocks and energy substrates, UCYN-A is expected to be an obligate symbiont (Zehr and Bombar 2015). Indeed, UCYN-A forms a symbiotic relationship with a haptophyte unicellular algae (Hagino et al. 2013; Zehr and Bombar 2015).

Three distinct clades of UCYN-A exist (UCYN-A1, UCYN-A2, and UCYN-A3). These clades have overlapping as well as distinct geographic distributions in open ocean and coastal environments (Thompson et al. 2014). Studies of UCYN-A1 and UCYN-A2 have revealed that though both lack the same major pathways and proteins and can thus be classified together, they share only 86% amino acid sequence identity (Bombar et al. 2014). In addition, UCYN-A1 and UCYN-A2 associate with distinct hosts of Braarudosphaera bigelowii, which come from separate clades (Thompson et al. 2014). UCYN-A host specificity, which is dependent upon the UCYN-A clade, may reflect co-evolution of symbiont and host that allowed for adaptation to different niches (Bombar et al. 2014; Thompson et al. 2014).

Studies with B. bigelowii and UCYN-A have indicated that UCYN-A was always present in cells of the algal species, which may indicate that the association is obligate for the host (Cabello et al. 2016). However, UCYN-A may not necessarily require a host for its survival. Despite its known association with a host, UCYN-A has also been observed as presumably free cells (Krupke et al. 2013). The UCYN-A genome does encode transporters for sugars and dicarboxylic acids (Bothe et al. 2010b). Thus these organisms may be able to import sugars and amino acids from ocean waters and metabolize and incorporate the products into cellular components (Bothe et al. 2010b).
1.3.5.2 *Calothrix* and *Richelia*: Symbiotic cyanobacteria with plants

Some N\textsubscript{2}-fixing cyanobacteria are known to exist as diatom symbionts. *Calothrix* and *Richelia* are diazotrophic cyanobacteria that form symbioses with plants in marine environments (Zehr and Bombar 2015). *Calothrix* associates on the spines of *Chaetoceros* diatoms while *Richelia* associates inside the fustule of *Hemiaulus* and *Rhizosolenia* (Zehr and Bombar 2015). Thus, whereas *Calothrix* exists as an epibiont, *Richelia* is an endosymbiont (Zehr and Bombar, 2015). The difference in host-symbiont association between the two cyanobacterial species is reflected in their genome size: *Calothrix* has a genome size of 6.0 Mb while the genome size of *Richelia* is nearly half that at 3.2 Mb (Hilton et al. 2013). Interestingly, the endosymbiont (*Richelia*) lacks many core N metabolism genes, such as ammonium transporters and glutamine oxoglutarate aminotransferase (GOGAT) (Hilton et al. 2013). Genome reduction in the endosymbiont compared to the epibiont evidence a streamlining that occurs due to the extent of association between symbiont and host (Zehr and Bombar 2015).
CHAPTER 2: Identification of Two Genes Required for Heptadecane Production in a N₂-fixing Cyanobacterium *Anabaena* Sp. Strain PCC 7120

2.1 Abstract

Cyanobacteria photosynthetically produce long-chain hydrocarbons, which are considered as infrastructure-compatible biofuels. However, native cyanobacteria do not produce these hydrocarbons at sufficient rates or yields to warrant commercial deployment. This research sought to identify specific genes required for photosynthetic production of alkanes to enable future metabolic engineering for commercially viable production of alkanes. The two putative genes (*alr5283* and *alr5284*) required for long-chain hydrocarbon production in *Anabaena* sp. PCC 7120 were knocked out through a double crossover approach. The knockout mutant abolished the production of heptadecane (C\(_{17}\)H\(_{36}\)). The mutant is able to be complemented by a plasmid bearing the two genes along with their native promoters only. The complemented mutant restored photosynthetic production of heptadecane. This combined genetic and metabolite (alkanes) profiling approach may be broadly applicable to characterization of knockout mutants, using N₂-fixing cyanobacteria as a cellular factory driven by solar energy to produce a wide range of commodity chemicals and drop-in-fuels from atmospheric gases (CO\(_2\) and N\(_2\) gas) and mineralized water.

2.2 Introduction

Oil reserves worldwide are limited, and as prices have risen, renewable fuels have become increasingly important. Is there a biofactory that can convert carbon dioxide, water, and sunlight into fuels? Yes, several species of cyanobacteria are known to
produce and secrete low levels of alkanes and alkenes using carbon dioxide, water, and sunlight (Schirmer et al. 2010).

Cyanobacteria provide numerous advantages as living “biofuel factories”. As photosynthetic organisms, they remove CO$_2$ from the atmosphere to form usable carbon products (sugars, isoprenoids, fatty acids, amino acids, etc.) that support cell growth and maintenance. Fatty acids have been shown to participate in an active recycling process within the cell membrane (Kaczmarzyk and Fulda 2010). Fatty acids activated by acyl-CoA synthetase are incorporated in the membrane and are subsequently degraded and released into the pool of free intracellular fatty acids before becoming re-activated. This continual membrane-lipid recycling may serve a vital role in cyanobacterial adaptation to diverse environments and changing conditions (Kaczmarzyk and Fulda 2010). It has also been observed that cyanobacteria convert fatty acids to alkanes and alkenes via a reduction-decarbonylation pathway (Schirmer et al. 2010). Thus, cyanobacteria provide an autotrophic platform to produce petroleum-replacing chemicals that could be harnessed to also reduce greenhouse gas (GHG) emissions from CO$_2$ emitting facilities, such as ethanol plants or coal-fired power plants.

Another advantage of cyanobacteria as a fuel producing platform is their minimal nutrient requirements. Many species can fix nitrogen aerobically and have minimal requirements for trace nutrients, meaning that the primary requirements are water, sunlight, and CO$_2$ (Ruffing 2011). In addition, the genomes and biochemical pathways of these autotrophic cyanobacteria are widely understood. Combined with their ease of genetic manipulation, this knowledge provides a firm foundation for genetically modifying cyanobacteria to direct carbon sources towards fuel production (Halfmann et
al. 2014a; Halfmann et al. 2014b; Lu 2010; Peralta-Yahya et al. 2012; Posewitz 2014; Ruffing 2011; Savakis and Hellingwerf 2015). Moreover, cyanobacteria have greater biomass production and photosynthetic efficiency compared to terrestrial biofuel crops, which also require use of arable land, thus impacting the food supply (Huang et al. 2010).

Heptadecane and pentadecane are the most commonly observed alkanes produced by cyanobacteria and are hypothesized to be derived from octadecanoic and hexadecanoic acids, respectively (Coates et al., 2014; Schirmer et al., 2010). Schirmer et al. (2010) reported that the pathway consists of an acyl-acyl protein reductase (AAR) and an aldehyde decarbonylase (ADO) to form alkanes and alkenes from generally even-numbered fatty aldehyde. Thus, following the “C_{n-1}” rule, odd-numbered alkanes are the typical products.

After the initial discovery of the aarlado genes in *Synechococcus elongatus* and their involvement in hydrocarbon production (Schirmer et al. 2010), studies have used phylogenetic analysis combined with hydrocarbon profiling to identify orthologs of aarlado in other cyanobacterial species (Coates et al. 2014; Liu et al. 2013). The AAR/ADO alkane biosynthesis pathway described by Schirmer is one of two hydrocarbon biosynthesis pathways operating in cyanobacteria (Mendez-Perez et al. 2011; Schirmer et al. 2010). The second pathway is the α-olefin biosynthesis (OLS) pathway (Zhu et al. 2018), which converts fatty acids into hydrocarbons via an elongation decarboxylation mechanism (Mendez-Perez et al. 2011). The AAR/ADO and OLS pathways are readily distinguishable by their products: alkane/alkenes or α-olefins, respectively (Coates et al. 2014; Mendez-Perez et al. 2011; Schirmer et al. 2010). Coates et al. (2014) reported that while all cyanobacterial species appear to be able to produce
alkanes, a strain’s possession of the AAR/ADO or OLS pathway to produce alkanes is mutually exclusive in available cyanobacterial genomes, suggesting an unknown selective pressure for retaining either pathway, but not both. The AAR/ADO pathway is most prevalent among sequenced cyanobacteria (122 of 139) (Coates et al. 2014). Cyanobacterial species containing the AAR/ADO pathway predominantly produce heptadecane and branched alkanes (e.g., 7-methylheptadecane) (Coates et al. 2014; Schirmer et al. 2010).

Further investigations into hydrocarbon biosynthesis by cyanobacteria have used computational analysis such as microarray data and RNAseq to understand expression of $aar/ado$ (Mitschke et al. 2011). Mitschke et al. (2011) showed that $ado$ (sll0208) and $aar$ (sll0209) had different expression levels under all conditions (high light, CO$_2$ depletion, normal, and darkness). The $ado$ expression did not vary significantly while $aar$ had clear response to the conditions (with expression levels from highest to lowest being high light, CO$_2$ depletion, normal, and darkness). Other research has overexpressed or introduced these genes ($aar/ado$) into other species with the aim of enhancing alkane production. Overexpression of $aar$ in *Synechococcus elongatus* was able to increase alkane production approximately 2-fold (Kaiser et al. 2013). Overexpression of both native copies of $aar/ado$ in *Synechocystis* sp. PCC6803 doubled alkane production compared to the parent strain while overexpression of only one of the genes (either $aar$ or $ado$) resulted in no significant changes in alkane production (Wang et al. 2013). Heterologous overexpression of $aar/ado$ in *Synechocystis* sp. PCC6803 also doubled alkane production when the genes were overexpressed simultaneously (Wang et al. 2013). Yoshino et al. (2015) demonstrated that though the AAR/ADO and OLS pathways are not natively
observed to exist together in a single cyanobacterial species, the OLS-containing strain *Synechococcus* sp. NKBG15041c was able to produce heptadecane by expressing the *aar*/*ado* genes from *Synechococcus elongatus* PCC7942. This research showed that heptadecane production levels in *Synechococcus* sp. NKBG15041c varied according to the expression levels of *aar*/*ado*, with production being the highest when the transformant carried a homologous promoter to the native *aar*/*ado* genes (Yoshino et al. 2015).

In this work, I identified the *aar*/*ado* genes in a heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120 through BLAST-P alignment with *S. elongatus* PCC7942_orf1593 (*aar*) and orf1594 (*ado*), the *aar*/*ado* genes initially identified in Schirmer’s study (Schirmer et al. 2010). I aimed to directly verify that these genes are required for hydrocarbon production in vivo. My approach was to simultaneously knock out both genes to determine the knockout mutant’s phenotype, and then re-insert the functional genes back to the knockout mutant for testing complementation. When the alkane genes were initially identified, they were believed to be part of an operon (Schirmer et al. 2010). Later research using differential RNA sequencing for genome-wide mapping of transcriptional start sites (TSS) in *Synechocystis* PCC6803 revealed that *aar* and *ado* possess their own TSS (Mitschke et al. 2011). Subsequent research identified three promoters involved in controlling expression of *aar* and *ado* (Klahn et al. 2014). One promoter controls *aar* while two promoters (a proximal and distal promoter) control *ado* (Klahn et al. 2014).

For the complement plasmids in my study, I created constructs in which *alr5283-84* were placed under a combination of native and the glnA (*alr2328*) promoters.
Through varying the promoter systems in the complement study, I hoped to gain further insight into the control mechanisms behind the alkane genes’ transcriptional regulation. My results demonstrated that the native promoter system was the only one able to complement the knockout mutant, highlighting the importance and underlying complexity of the native three-promoter system.

2.3 Materials and Methods

2.3.1 Bacterial strains and plasmids

*Escherichia coli* strains Top10 (Invitrogen) and NEB10β (New England Biolabs) were used for plasmid construction and maintenance. *E. coli* strains were grown in Luria-Bertani broth. Antibiotic concentrations used throughout were 100 µg mL\(^{-1}\) ampicillin, 50 µg mL\(^{-1}\) kanamycin, 10 µg mL\(^{-1}\) spectinomycin, and 10 µg mL\(^{-1}\) erythromycin.

2.3.2 Construction of *aar/ado* knockout plasmid pZR935

To disrupt the *alr5283-5284* sequence in *Anabaena* 7120 genome, plasmid pZR935 was created (details see Table 2.1) and transferred to *Anabaena* 7120 via conjugative transformation to replace chromosomal *alr5283-84* with the disrupted *alr5283-84* sequence via double recombination. Briefly, *alr5283-5284* and flanking sequences were PCR amplified with specific primers ZR241, ZR242 (Table 2.2) from *Anabaena* 7120 genomic DNA, cloned into pCR®TOPO2.1® vector (TOPO TA Cloning® kit, Invitrogen), creating pZR932. Next, site directed mutagenesis using primers ZR243,244 introduced a NotI site into *alr5283* within pZR932, creating pZR933. Then, the mutated *alr5283-84* sequence was excised from pZR933 and transferred to the vector pZR824, creating pZR934. Finally, the 3’ end of *alr5283* and the 5’ end of *alr5284* were excised from pZR934 using restriction enzymes NotI and NheI; a promoter-less
GFP-Spec cassette from pZR666 was inserted into NotI and XbaI digested pZR934, creating pZR935 (details see Fig. 2.1).

Table 2.1 Plasmids and bacterial strains used in this study

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAM1956</td>
<td>Promoter-less GFPmut2 cloning vector</td>
<td>Yoon and Golden (1998)</td>
</tr>
<tr>
<td>pRL271</td>
<td>Cmr/Emr. Integration vector</td>
<td>Cai and Wolk (1990)</td>
</tr>
<tr>
<td>pRL278</td>
<td>Km/Nmr; integration vector</td>
<td>Cai and Wolk (1990)</td>
</tr>
<tr>
<td>pRL443</td>
<td>Ap; Conjugal plasmid</td>
<td>Elhai et al. (1997)</td>
</tr>
<tr>
<td>pRL623</td>
<td>Cm; Helper plasmid</td>
<td>Elhai et al. (1997)</td>
</tr>
<tr>
<td>pZR606</td>
<td>Km/Spr, integration vector</td>
<td>Chen et al. (2015)</td>
</tr>
<tr>
<td>pZR618</td>
<td>Ap; T7-MCS-F1-H6</td>
<td>Chen et al. (2012)</td>
</tr>
<tr>
<td>pZR666</td>
<td>Km/Sp, MCS-gfp-MCS cassette, annealed oligonucleotides</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>LK2406/LK2407 ligated to AflII digested pZR934 to disrupt pZR666</td>
<td></td>
</tr>
<tr>
<td>pZR670</td>
<td>Cm/Ems; expression vector for Anabaena</td>
<td>(Chen et al. (2015); Xu et al. (2015))</td>
</tr>
<tr>
<td>pZR824</td>
<td>Km; integration vector, annealed oligonucleotides ZR165/ZR166 digested to BglII-SpeI</td>
<td>This study</td>
</tr>
<tr>
<td>pZR932</td>
<td>Km/A Gor/Or/Or ORF amplified by PCR with primers ZR241,242 from Anabaena 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector</td>
<td>This study</td>
</tr>
<tr>
<td>pZR933</td>
<td>Km/A Gor; site-directed mutagenesis using primers ZR243,244 to introduce NotI site into alr5283-84 within pZR932</td>
<td>This study</td>
</tr>
<tr>
<td>pZR934</td>
<td>Km; BamHI and AvrII digested alr5283-84 ORF from pZR933 ligated to BglII and SpeI digested pZR824</td>
<td>This study</td>
</tr>
<tr>
<td>pZR935</td>
<td>Kmr/Spr; NotI and XbaI promotor-less GFP-Spr cassette from pZR666 ligated to NotI and NheI digested pZR934 to disrupt alr5283-84</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2222</td>
<td>Kmr/ Gor; Primers ZR1584 and ZR1585 PCR amplified 1921 bp Cmr/Emr cassette (NsiI/Ndel-Nhel/BamHI-Emr/Cmr cassette- Ecor/RV/BglII/Xbal/XmaI/XmaI) from pRL271 ligated to pCR2.1-TOPO to produce pZR2222</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2223</td>
<td>Cmr/Emr; expression vector, BamHI/XmaI cut out 1.9 kb Cmr/Emr cassette from pZR2222 and ligated to BglIII/SgrAI-digested pAM1956</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2238</td>
<td>Kmr/ Gor; AarII/SalI-P-alr5283-P-alr5284-KpoI/XmaI PCR amplified by ZR1602 and 1603 from A.7120 ligated to pCR2.1-TOPO</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2239</td>
<td>Cmr/Emr; SalI-P-alr5283-P-alr5284-XmaI from pZR2238 ligated to SalI/XmaI cut pZR2223</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2242</td>
<td>Kmr/ Gor; NsiI/Ndel-alr5283 orf-P-alr5284 orf-BamHI PCR amplified from A. 7120 ligated to pCR2.1-TOPO</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2243</td>
<td>Cmr/Emr; NsiI-alr5283 orf-P-alr5284 orf-BamHI from pZR2242 ligated to NsiI-BamHI cut pZR670</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2244</td>
<td>Kmr/ Gor; NsiI/Ndel-alr5283 orf PCR amplified by ZR1606 and 1608 from A. 7120; RBS-alr5284 orf-BamHI PCR amplified by ZR1609 and 1607 from A. 7120; PCR overlap of NsiI/Ndel-alr5283 orf and RBS-alr5284 orf-BamHI with primers ZR1606 and 1607; PCR overlap product ligated to pCR2.1-TOPO</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2248</td>
<td>Cmr/Emr; NsiI-alr5283 orf-RBS-alr5284 orf-BamHI from pZR2244 ligated to NsiI-BamHI cut pZR670</td>
<td>This study</td>
</tr>
</tbody>
</table>

**Bacterial Strains**

| TOP10  | E. coli cloning host | Invitrogen |
| NEB10β | E. coli cloning host | New England Biolabs |
| WTT7120 | Anabaena sp. PCC 7120 wild-type strain | This study |
| DR935  | Sp; alr5283-84 double knockout mutant | This study |
DR935(pZR2239)  Sp<sup>+</sup>,Em<sup>+</sup>; DR935 containing pZR2239 for complementation study  This study
DR935(pZR2248)  Sp<sup>+</sup>,Em<sup>+</sup>; DR935 containing pZR2248 for complementation study  This study
DR935(pZR2243)  Sp<sup>+</sup>,Em<sup>+</sup>; DR935 containing pZR2243 for complementation study  This study

*Ap<sup>+</sup>, ampicillin resistance; Sp<sup>+</sup>, spectinomycin resistance; Nm<sup>+</sup>/Km<sup>+</sup>, neomycin-kanamycin resistance; Cm<sup>+</sup>/Em<sup>+</sup>, chloramphenicol-erythromycin resistance; F<sub>2</sub>, two FLAG epitopes, MCS: multiple cloning sites

---

### Table 2.2 Primers used in this study

<table>
<thead>
<tr>
<th>Primers</th>
<th>Oligonucleotide sequences (5'→3')</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR165</td>
<td>GATCTCCGGCTAGCGGCCGCAATTGACGTCTCGA</td>
<td>Annealed oligonucleotides ZR165/ZR166 ligated to BglII-SpeI digested pRL278 to produce pZR824</td>
</tr>
<tr>
<td>ZR166</td>
<td>ctagTCTCGAGACGTCAATTGCCCGGCCTAGCCCAGGA</td>
<td></td>
</tr>
<tr>
<td>ZR241</td>
<td>tgggTCACAATCTACAGAATTTGCTG</td>
<td>ZR241,242 primer pair amplifying alr5283-84 ORF (2.7 kb); in knockout mutant with GFP-Spec cassette, primers amplify 4.7 kb</td>
</tr>
<tr>
<td>ZR242</td>
<td>tctAGGAATTTTGATTTGGGGATTG</td>
<td></td>
</tr>
<tr>
<td>ZR243</td>
<td>CAAAAAGCGGCcGCTGAAGTTTCACGGACATCATCACGATCCGGATTTAG</td>
<td>ZR243,244 primer pair site-directed mutagenesis to introduce NotI site at 997 bp of alr5283-84 region</td>
</tr>
<tr>
<td>ZR244</td>
<td>TTTACCTTTTCAGCTACATCATATCATTAAATATATTTATTAAATGTTTGTTT</td>
<td>Primers ZR243/ZR244 PCR amplified 1921 bp Cmr/Emr cassette (NsiI/NdeI-NheI/BglII-XhoI/XmaI/SmaI) from pRL271 ligated to pCR2.1-TOPO to produce pZR2222</td>
</tr>
<tr>
<td>ZR261</td>
<td>CAAGAATTGGGCAACATCCAGTG</td>
<td>ZR261,1602 primer pair for verifying insertion of P&lt;sup&gt;-&lt;/sup&gt; alr5283-P&lt;sup&gt;-&lt;/sup&gt;alr5284 in pZR2223 shuttle vector (construction pZR2239)</td>
</tr>
<tr>
<td>ZR1585</td>
<td>tccCGGGAAGTATCCGCTCGAGATC</td>
<td>ZR1584,1585 primer pair amplifying P&lt;sup&gt;-&lt;/sup&gt; alr5283-P&lt;sup&gt;-&lt;/sup&gt;alr5284 (2.2 kb); amplification retains promoters for both genes</td>
</tr>
<tr>
<td>ZR1606</td>
<td>atgcatATGCAGCACGGTGAGCCCTTAGT</td>
<td>ZR1606,1607 primer pair amplifying alr5283-84 (2 kb); amplification includes promoter for alr5284 only</td>
</tr>
<tr>
<td>ZR1607</td>
<td>ggggCTAACCAGCATCAGATTTCGT</td>
<td>ZR1606-1608-1609-1607 overlap PCR (1.7 kb) (1) primer pair ZR1606,1608 amplify alr5283 ORF (2) primer pair ZR1609,1607 amplify alr5284 ORF (3) primer pair ZR1606,1607 using template fragments from steps 1 and 2 for PCR overlap, combining the fragments (1.7 kb)</td>
</tr>
<tr>
<td>ZR1608</td>
<td>CATGGtatatctctttTTAAGACTAGTACCGGTCTAGATCTACCTAA</td>
<td>ZR1606,1607 primer pair amplifying alr5283-84 (2 kb); amplification includes promoter for alr5284 only</td>
</tr>
<tr>
<td>ZR1609</td>
<td>GACTTACAGCAGCTAaagaaggagatatCACATGGTATTGGT</td>
<td>ZR1606,1607 primer pair amplifying alr5283-84 (2 kb); amplification includes promoter for alr5284 only</td>
</tr>
<tr>
<td>LK2406</td>
<td>TTAAGGCCCCGAGATCTACAGGATTTTAGT</td>
<td>Annealed oligonucleotides LK2406/LK2407 ligated to AffIII digested pZR666 to produce pZR666, Ap&lt;sup&gt;+&lt;/sup&gt;-XmaI-BglII-XhoI-AgeI-SpeI multiple cloning site (MCS)</td>
</tr>
<tr>
<td>LK2407</td>
<td>TTAAGACTAGTGCGTTCTAGACGCTAGCTACGGGCGACC</td>
<td></td>
</tr>
</tbody>
</table>

Sp: spectinomycin resistance; ORF: open reading frame; GFP: green fluorescent protein; MCS: multiple cloning sites, RBS: ribosome-binding site, P<sup>-</sup>: promoter, PCR: polymerase chain reaction
Figure 2.1 Schematic illustration of pZR935 construction for knocking out \textit{alr5283} and \textit{alr5284}. Step 1: Amplified 2.7 kb fragment from \textit{Anabaena} 7120 containing \textit{alr5283-84} using primers ZR241,242. Cloned PCR product to pTOPO 2.1 vector, creating pZR932. Step 2: Site directed mutagenesis using primers ZR243,244 to introduce NotI site into \textit{alr5283} within pZR932, creating pZR933. Step 3: Digestion of pZR933 with BamHI-AvrII to obtain 2.7 kb \textit{alr5283-84} mutated sequence. Ligated sequence to BglII-Spel cut pZR824 vector, creating pZR934. Step 4: Digestion of pZR934 with NotI-NheI and insertion of NotI-XbaI promoter-less GFP-Spec cassette from pZR666 into deleted site of pZR934, creating pZR935.

2.3.3 Knocking out \textit{alr5283-5284} in \textit{Anabaena} sp. PCC 7120

Tri-parental mating was initiated by mixing HB101 [pRL623+443] with \textit{E. coli} 10β containing the cargo plasmid pZR935. The \textit{E. coli} strains were combined in a single 1.5 mL tube and set at room temperature for 30 min to allow the strains to mate.

Following mating, \textit{Anabaena} 7120 was added to the mating mixture. Cultures of \textit{Anabaena} 7120 were grown in Bg11 medium and incubated at 30°C, 120 rpm, under continuous light (60 µmol E m\(^{-2}\) s\(^{-1}\)) for 7 d until the culture reached an OD\(_{700}\) of 0.5. Ten mL of the culture was harvested (4000 xg, 10 min), and the pellet was washed with 1 mL
Following a second centrifugation (12,000 ×g, 1 min), the pellet was resuspended in 100 µL Bg11. This was added to the E. coli mixture containing the cargo plasmid for transformation, the helper plasmid pRL623, and the conjugal plasmid pRL443 (Elhai et al. 1997). The cyanobacteria and E. coli were allowed to mate for 1 h. Then, the solution was plated on a nitrocellulose membrane on Bg11 agar supplemented with 5% LB and incubated at 30°C under light for 2 d. Next, the membrane was transferred to a Bg11 plate containing the antibiotic to select for transformed Anabaena 7120. Plates were incubated at 30°C under light until single colonies formed. On a weekly basis, membranes were transferred to new Bg11 antibiotic plates.

Conjugal transformation of pZR935 to Anabaena 7120 to achieve a double crossover was first verified by distinguishing single crossover colonies via colony PCR (primers listed in Table 2.2). Single crossover recombination colonies named SR935 were then grown in 10 mL BG11 plus spectinomycin for 1 week. Then, one mL of SR935 was sonicated for 10 min until all filaments were separated into single cells (visualized under a microscope). The sonicated culture was harvested (13,000 ×g, 1 min), and the pellet was resuspended in 100 µL Bg11. The suspension was plated on Bg11 plus spectinomycin and 5% sucrose. Spectinomycin selected for transformed Anabaena 7120 over wildtype while sucrose selected against single crossovers due to the sacB gene in pZR935 serving as a suicide vector. After 1 week’s growth under light at 30°C, single colonies appeared and were verified to be double recombinants via colony PCR; double crossover recombinants are henceforth referred to as the alr5283-alr5284 knockout mutant named DR935.
2.3.4 Complementation of Alr5283-Alr5284 knockout mutant

To complement the knockout mutant (DR935), three plasmids (pZR2239, pZR2248, and pZR2243) were constructed containing alr5283-84 under different promoter systems. Briefly, pZR2239 contained the genes both under their native promoters. In plasmid pZR2248, both genes were under control of the glnA (alr2328) promoter, though each gene contained its own ribosomal binding site. Plasmid pZR2243 contained alr5283 under control of the constitutive glutamine synthetase (glnA) promoter while alr5284 remained under control of its native promoter. Table 2.1 provides further details of the plasmids and bacterial strains used in this study. Q5-High-Fidelity DNA Polymerase (NEB) was used for all PCR amplifications. Cloned PCR products and mutated genes were verified by Sanger DNA sequencing at GenScript. All cloning enzymes such as restriction endonucleases were purchased from NEB. All PCR primers synthesized at Integrated DNA Technologies (IDT) are listed in Table 2.2.

Conjugal transformation of individual complement plasmids pZR2239, pZR2248, or pZR2243 to DR935 was performed using the same conjugation method detailed above. After obtaining individual colonies on the antibiotic selection plate, successful complements were verified by performing colony PCR with individual colonies. Verified complements carrying plasmids pZR2239, pZR2248, and pZR2243 were named as DR935(pZR2239), DR935(pZR2248), and DR935(pZR2243), respectively.

2.3.5 Headspace sample collections and extractions

Wildtype Anabaena 7120; DR935, and the complementing strains, DR935(pZR2239), DR935(pZR2248), and DR935(pZR2243), were grown in 100 mL Bg11 for 17 d with an initial OD\textsubscript{700} of 0.03. Cultures were incubated at 30°C, 120 rpm at
an aeration rate of 100 mL min$^{-1}$. A resin column inserted into the rubber stopper sealing each flask was used to capture hydrocarbons in the headspace (Halfmann et al. 2014a); each resin column contained 0.12 g Supelpak-2SV resin (Sigma-Aldrich). During the incubation period, hydrocarbons were extracted from the resin columns on day 5.

To extract hydrocarbons bound to the resin Supelpak-2SV, the resin was transferred to 1.5 mL eppie tubes, and 1 mL pentane containing 5 µg mL$^{-1}$ tetracosane as an internal standard was added to the resin. The resin and pentane mixture was vortexed for 1 min. and left to sit for 10 min. After 10 min., each sample was vortexed for 10 sec. and pentane was removed and stored in a 2 mL glass vial for GC-MS analysis. Ten µg mL$^{-1}$ standards of pentadecane, 1-pentadecene, heptadecane, and 1-heptadecene (TCI) were analyzed by GC-MS for comparison purposes.

2.3.6 Gas Chromatography-Mass Spectrometry (GC-MS) analysis of hydrocarbon samples

Hydrocarbon extractions from the cultures were analyzed using GC-MS (Agilent 7890A/5975C) at the Functional Genomics Core Facility of South Dakota State University. One-microliter injected samples were separated by an HP-5MS column (35 m x 250 µm x 0.25 µm), with H$_2$ serving as the carrier gas. The oven temperature was initially held at 145°C for 2 min, increased 5°C min$^{-1}$ until 180°C was reached, and then increased by 40°C min$^{-1}$ to 300°C. Total run time was 14 min. Compounds were identified using the NIST MS library v2.0 and further verified by authentic standards.

The initial GC-MS analysis of the headspace samples from *Anabaena* 7120 used 50 to 500 m/z full scan, which identified heptadecane production. Based on the spectra obtained in the full scan of heptadecane and the internal standard tetracosane, I created a
selected ion monitoring (SIM) method to use for future analysis of wildtype, DR935, and complement strains DR935(pZR2239), DR935(pZR2248), and DR935(pZR2243). SIM was chosen because it is more selective and it provides a better signal for heptadecane. The GC-MS SIM method used selected ion monitoring parameters of (57, 71, 240) and (57, 71, 338) for heptadecane and tetracosane, respectively. Heptadecane had a retention time of 6.31 min; tetracosane had a retention time of 11.8 min.

2.3.7 Chlorophyll content analysis

Chlorophyll content was quantified at OD$_{665}$ following the previously detailed method (Houmard and de Marsac 1988).

2.3.8 GenBank access to genes

The two genes knocked out for the alkane study, *alr5283* and *alr5284*, can be found through GenBank using accession numbers NC_003272.1:6303216-6303911 and NC_003272.1:6304154-6305173, respectively.

2.4 Results

2.4.1 Identification of heptadecane emitted from *Anabaena 7120*

A GC-MS analysis of volatile compounds emitted from wildtype *Anabaena 7120* revealed a prominent peak at retention time 6.31 min (Fig. 2.2a). The peak was identified as heptadecane (C$_{17}$H$_{36}$) by the MS library. To confirm the identity of the hydrocarbon peak, authentic standards for pentadecane, 1-pentadecene, heptadecane, and 1-heptadecene were analyzed by GC-MS. Heptadecane had the same retention time (6.31 min) as the wildtype compound (Fig. 2.2c) and showed the same fragmentation pattern as the wildtype compound (comparing Fig. 2.2d to Fig. 2.2b). Thus, I concluded that the major hydrocarbon peak seen in *Anabaena 7120* is heptadecane (C$_{17}$H$_{36}$).
Figure 2.2 Identification of heptadecane produced by *Anabaena* 7120. **a** GC/MS chromatograph of the volatile metabolites from *Anabaena* 7120 cultures. A peak at the retention time of 6.31 min (black arrow) found in *Anabaena* matches the heptadecane standard (**c**). **b, d** Mass spectra of the 6.31 min peaks display the fragmentation pattern for the compound in *Anabaena* and the heptadecane standard, respectively. Five μg mL⁻¹ tetracosane serves as an internal standard (IS, blue arrow)

### 2.4.2 Bioinformatics analysis to identify hydrocarbon biosynthesis genes

Hydrocarbon biosynthesis genes have been identified in the cyanobacterial species *Synechococcus elongatus* PCC 7942 (Schirmer et al. 2010). Hydrocarbon biosynthesis genes have not yet been verified in *Anabaena* 7120. To identify the genes required for heptadecane production in *Anabaena* 7120, a blastp search using the representative proteins from *S. elongatus* against the publically available *Anabaena* 7120 proteome from GenBank was performed to identify putative alkane biosynthesis proteins
and their respective genes. The protein sequence alignment revealed Alr5283 and Alr5284 from *Anabaena* 7120 were homologous to ADO and AAR from nine cyanobacterial species known to produce heptadecane and pentadecane (Schirmer et al. 2010) (Fig. 2.3a, b, respectively).
Figure 2.3 Identification of the aldehyde decarbonylase and acyl-ACP reductase *Anabaena* 7120. a Multiple sequence alignment of aldehyde decarbonylase in cyanobacteria. *Anabaena* 7120 consists of aldehyde decarbonylase (Alr5283, Accession No. BAB76982) and acyl-ACP reductase (Alr5284, Accession No. BAB76983). Homologs used in the multiple sequence alignment include YP_323043 and YP_323044 in *Trichormus variabilis* ATCC 29413; YP_001865325 and YP_001865324 in *Nostoc punctiforme* PCC 73102; NP_442147 and NP_442146 in *Synechocystis* sp. PCC 6803; YP_001802195 and YP_001802846 in *Cyanothece* sp. ATCC 51142; ZP_03137291, ZP_03139316 and YP_002481152 in *Cyanothece* sp. PCC 7425; YP_400610 (Synpcc7942_1593) and YP_400611 (Synpcc7942_1594) in *Synechococcus elongatus* PCC 7942; YP_170760 and YP_1707601 in *Synechococcus elongatus* PCC 6301; NP_926092 and NP_926091 in *Gloeobacter violaceus* PCC 7421; and NP_892650 and NP_892651 in *Prochlorococcus marinus* subsp. *pastoris* str. CCMP1986, respectively. Sequence alignment was made using an online program MultiAlin (Corpet 1988). The figure was generated by an online program ESPript 3.0 (Robert and Gouet 2014).

2.4.3 Construction of DR935

The two putative genes (*alr5283-alr5284*) whose products may be responsible for long-chain hydrocarbon production were cloned from the *Anabaena* 7120. The genes were rendered nonfunctional through deletions of the 3' end of *alr5283* and 5' end of *alr5284*; the native sequence of these genes was further disrupted by a gfp-spectinomycin cassette inserted between the *alr5283* and *alr5284* gene sequences.

Conjugative transformation of pZR935 to *Anabaena* 7120 resulted in a double crossover, which replaced the functional *alr5283-84* gene sequence with the disrupted gene sequence (Fig. 2.4a). DR935 was verified by colony PCR from the conjugation (Fig. 2.4b). An expected 2.7 kb PCR product using primers ZR241 and ZR242 flanking the *alr5283-84* gene sequence was detected in the wildtype *Anabaena* 7120 (Fig. 2.4b lane 2). In DR935, containing the gfp-spec cassette within the *alr5283-84* gene sequence, the PCR product amplified by primers ZR241 and ZR242 increases to 4.7 kb (Fig. 2.4b lane 3). Thus, pure double recombinants containing only the 4.7 kb amplification were
obtained. Verified double recombinants were used for further analysis. Three plasmids were constructed to complement DR935 (Fig. 2.4c).

**Figure 2.4** Construction of Alr5283-Alr5284 knockout mutant (DR935) and its complemented strains in *Anabaena* 7120. **a** 3' deletion of *ado* (*alr5283*) and 5' deletion of *aar* (*alr5284*) created by inserting a gfp-spec cassette between *alr5283–84* in Anabaena 7120 chromosome via double recombination with knockout plasmid pZR935. **b** PCR verification of DR935. Wildtype *Anabaena* 7120 had the intact *alr5283–84* gene sequence, which has a length of 2.7 kb when amplified by ZR241, 242 (lane 3). DR935 contained the gfp-spec cassette inserted in the *alr5283–84* gene sequence, making the amplified gene sequence 4.7 kb (lane 2). **c** Complementing plasmid constructions: pZR2239 contains *ado* and *aar* both under control of their native promoters, pZR2248 contains the engineered *ado-aar* operon under control of the constitutive *glnA* promoter, a standard ribosome-binding sequence (AAGGAGA) was introduced between *ado* and *aar* in pZR2248, and pZR2243 contains *ado* under control of *PglnA* and *aar* under control of its native promoter.
2.4.4 Alr5283 and Alr5284 are responsible for heptadecane production in *Anabaena* 7120

Volatile compounds emitted from wildtype, DR935, and complement strains were analyzed via GC-MS SIM. In DR935, the heptadecane peak at 6.31 min disappeared, indicating that *alr5283-84* are required for heptadecane production in *Anabaena* 7120 (Fig. 2.5b). Among the three complement plasmids constructed, only DR935(pZR2239), containing *alr5283-84* under control of the native 3-promoter system, was able to rescue heptadecane production in the knockout mutant (Fig. 2.5c). Neither DR935(pZR2248) nor DR935(pZR2243) was able to recover hydrocarbon biosynthesis in the knockout mutant.

2.4.5 Heptadecane production

I analyzed the amount of heptadecane produced from wildtype *Anabaena* 7120, DR935, and DR935(pZR2239) strains on day 5 of the experiment. Overall, heptadecane production was higher in the wildtype compared to the complement and undetectable in the mutant (Fig. 2.5d).
Figure 2.5 Alr5283 and Alr5284 are responsible for the heptadecane production in *Anabaena* 7120. GC/MS SIM chromatographs of the volatile metabolites from wildtype (a), DR935 (b), and DR935(pZR2239) (c) strains, respectively. Heptadecane naturally produced in wildtype *Anabaena* 7120 was not detected in DR935, but reemerged after the complement plasmid pZR2239 was transformed into DR935. Five μg mL\(^{-1}\) tetracosane serves as an internal standard (IS, blue arrow). d Total heptadecane production mg\(^{-1}\) chlorophyll of wildtype *Anabaena* 7120, DR935, and DR935(pZR2239) cultures from days 0 to 5. Heptadecane yield was calculated by measuring the total heptadecane produced from days 0 to 5 and dividing by total chlorophyll content of the 100 mL cultures measured on day 5. The culture density increased over days 0–5 but only day 5 chlorophyll content was used to make the calculation (the chlorophyll content of seed culture at day 0 is negligible). Therefore, the total heptadecane production mg\(^{-1}\) chlorophyll calculation is an underestimate.

Though heptadecane production differed amongst the strains, it did not appear to impact normal growth of the culture as indicated by OD measurements (Fig. 2.6) and visual qualitative analysis of culture growth (data not shown). These results indicate that under the growth conditions used in this study, heptadecane production does not have a significant impact on cell growth or survival.
2.5 Discussion

2.5.1 *alr5283* and *alr5284* are required for heptadecane production in *Anabaena 7120*

In a previous study of alkane production by cyanobacteria, heptadecane and pentadecane were cited as the most commonly produced alkanes (Schirmer et al. 2010). In my work, heptadecane was the predominant volatile metabolite captured by Supelpak 2SV resin and detected in wildtype *Anabaena 7120*. Heptadecane production was lost when *alr5283-84* were disrupted and re-emerged after complemented by intact *alr5283* and *alr5284* driven by their native promoters. Thus, the enzymes Alr5283 and Alr5284 from *Anabaena 7120* are required for heptadecane production. My findings are partially...
supported by previous reports, which detected a range of hydrocarbons produced by the
ado/aaar gene products from various cyanobacteria when the genes were expressed in E. coli (Schirmer et al. 2010) or Synechocystis sp. PCC6803 (Wang et al. 2013). It is possible that pentadecane produced in Anabaena 7120 was trapped inside cells while heptadecane was emitted into headspace. Taken together, I conclude that the AAR/ADO enzymes encoded by alr5283-84 are responsible for heptadecane production from a C18 fatty acid substrate in Anabaena 7120.

2.5.2 Differential expression of ADO and AAR is critical for heptadecane production

A complementation experiment was required to confirm that two genes (ado and aar), rather than a downstream gene alr5285 (acetyl-CoA carboxylase alpha subunit), are responsible for DR935 mutant phenotype (loss of heptadecane production). The replicating plasmid pRL2833a (Wolk et al. 2007) and its derivative pZR670 (Chen et al. 2016a) have been successfully used for complementing many knock-out mutants in Anabaena 7120 (Chen et al. 2016a; Fan et al. 2005). The glnA promoter is a constitutive expression promoter and is upregulated by nitrogen starvation (Flaherty et al. 2011). The glnA promoter was successfully used for complementing an alr4853 mutant in Anabaena 7120 (Xu et al. 2015). DR935(pZR2248) contained both genes in a two gene operon (ado–aar) under the glnA promoter, yet did not recover heptadecane production. Importantly, among the complement plasmids used to recover heptadecane production in the adolaar knockout mutant DR935, only the plasmid pZR2239 containing the genes controlled by their native promoter system was able to complement the mutant. Since the only difference between the complement strains was the promoter system used, the
non-native promoter systems used in DR935(pZR2248) and DR935(pZR2243) are likely responsible for the failure of complementation. It is possible that the constitutive glnA promoter used in the unsuccessful complements did not function as optimally as the native promoters for either aar or ado.

These results indicate the autonomous regulation and perhaps differential expression of aar and ado are necessary for optimal functioning and/or interplay of the enzymes within the alkane biosynthetic pathway. Indeed, comparison of microarray experiments detailing expression of aar and ado in Synechocystis revealed that contrary fold changes occurred in the genes under many conditions (Klahn et al. 2014). My results provide further evidence for the unique and differential expression of aar and ado as well as the importance of the genes’ independent expression for heptadecane production, consistent with the observation that the basal mRNA level of alr5283 was approximately six-fold higher than that of alr5284 in Anabaena 7120 (Flaherty et al. 2011).

Though translational regulation and post-translational regulation may also play important roles in heptadecane production in Anabaena 7120, my results suggest that the native promoter system has a central role in directing hydrocarbon production, and may in fact be vital. As previously suggested, the dual promoter system for ado may indicate different functions for ado, one of those being alkane production (Klahn et al. 2014). One commonality between both of the unsuccessful complement strains used in this study is that both placed ado under control of a non-native promoter (glnA). These results combined with the understanding of transcriptional regulation obtained in Klähn’s study further point to a potential divergence of ADO’s function in more than one cellular
process. In addition, the results suggest a more complex understanding of the native promoter system and its role in directing ADO’s expression and incorporation into cellular metabolism. Future research may be aimed at elucidating the roles of the proximal and distal ado promoters and underscore any differences that exist between ado expressions when driven from either promoter.

2.5.3 Potential role of heptadecane in cyanobacteria and its application in biofuel production

In this research, heptadecane was consistently found in the headspace of the cultures. Another study reported more than 80% of hydrocarbons produced by Nostoc punctiforme PCC 73102 were found outside the cells (Schirmer et al. 2010). In this study, under normal growth conditions (30 °C, normal light) growth of the mutant culture did not appear to be impacted by the loss of heptadecane production. Thus, under these conditions, though heptadecane is produced in small amounts, it is not required for survival and normal growth. However, it is possible that alkanes are required for response to certain stress conditions. Research has indicated that alkane production may be related to stress tolerance under various conditions, such as cold, high salinity, and high light (Berla et al. 2015; Bhadauriya et al. 2008; Kageyama et al. 2015; Takatani et al. 2015). In addition to abiotic stress responses, alkanes have also shown antibacterial activity in the cyanobacterium Spirulina platensis (Ozdemir et al. 2004). It is also possible that heptadecane has a function that is duplicated by another compound in the cell, such that when heptadecane is not produced, its vital function is still carried out.

With its high carbon content, heptadecane confers valuable fuel traits such as higher cetane number and oxidative stability, which are associated with long carbon
chains and saturation (Quintana et al. 2011). Cyanobacteria present themselves as ideal fuel producers given their photosynthetic ability to convert CO₂ to fuel using only the solar energy. Extraction processes often constitute 70–80% of production costs (Liu et al. 2011). Thus, in organisms which do not naturally secrete target compounds, further genetic engineering steps are required to enable product secretion from the cell (Liu et al. 2011). My findings show that the extraction process is bypassed in hydrocarbon production by *Anabaena* 7120 since heptadecane was naturally secreted from the cells and captured from the headspace.

To enable large-scale commercialization of cyanobacteria fuel/chemical production systems, productivity and yield must be increased. As currently understood, alkanes are derived from fatty acids in cyanobacteria via a reduction-decarbonylation pathway. Fatty acids are produced to store energy in the cell. The balance in cells between storage and metabolism/growth is tightly controlled (Greenwell et al. 2010). Since fatty acids are the precursor to the alkanes we seek as fuel, the regulation of this balance must be explored. How do we circumvent this process to attain both growth and the fatty acid production associated with the storage state? Another study attempted to bypass the regulation by genetically engineering microalgae to increase lipid synthesis through over-expression of acetyl-CoA carboxylase. However, the change did not result in greater lipid production (Dunahay et al. 1996). Future work should focus on increasing our understanding of the regulatory mechanisms in lipid storage and cell growth.
CHAPTER 3: Heptadecane Production Enhances Oxic Nitrogen Fixation in

*Anabaena* sp. PCC 7120

3.1 Abstract

In a previous study, I identified two genes responsible for heptadecane (C\(_{17}H_{36}\)) production in *Anabaena* sp. PCC 7120 via a genetic knockout approach, which created a mutant (DR935) unable to produce heptadecane. Though not required for growth under normal conditions, alkanes may be involved in helping *Anabaena* 7120 adapt to abiotic stress, such as nitrogen limitation. To test this hypothesis, *Anabaena* 7120 and DR935 were grown in nitrate-replete (Bg11) and fixed nitrogen-free (Bg11\(_o\)) conditions. Heptadecane production greatly increased in *Anabaena* 7120 in the days following nitrogen stepdown, while the mutant DR935 exhibited a fragmented phenotype compared to the wildtype, which grew in long filaments. The wildtype had significant nitrogenase activity at 48 h post fixed-nitrogen stepdown, but activity was negligible in the mutant. This reflected the mutant’s fragmented phenotype and difficulty in adapting to nitrogen deficient conditions. In long-term steady diazotrophic grown cultures, the wildtype had both higher nitrogenase activity and heterocyst frequency compared to the mutant. These findings suggest that heptadecane is primarily important in helping heterocyst-specific oxic nitrogen fixation by maintaining membrane integrity. After the initial adaptation to diazotrophic conditions, heptadecane’s role becomes less important or may be complemented by some alternate mechanism.

3.2 Introduction

*Anabaena* sp. PCC 7120 (hereafter *Anabaena* 7120) is a filamentous, nitrogen-fixing cyanobacterium. When a source of fixed nitrogen is not present in its environment,
*Anabaena* 7120 will differentiate specialized nitrogen-fixing cells, called heterocysts. Heterocysts terminally differentiate from vegetative cells at a frequency of approximately 1 heterocyst per 15 vegetative cells in a filament (Wolk et al. 1994).

Though not all cyanobacteria are able to fix nitrogen, cyanobacteria are universally able to produce hydrocarbons (Coates et al. 2014). *Anabaena* 7120 specifically produces heptadecane (C\(_{17}\)H\(_{36}\)), a long-chain alkane (Coates et al. 2014; Gibbons et al. 2018). In *Anabaena* 7120, heptadecane is produced via the AAR/ADO pathway in which a C18 fatty acyl-ACP is first converted to a fatty aldehyde by the enzyme fatty acyl-ACP reductase (AAR); then, aldehyde decarbonylase (ADO) removes a carbonyl group from the fatty aldehyde, producing the alkane (Coates et al. 2014; Schirmer et al. 2010). In a previous study, I knocked out two genes (*alr5283*-*alr5284*, encoding ADO and AAR, respectively) in *Anabaena* 7120 and determined that they were responsible for heptadecane production (Gibbons et al. 2018).

Interestingly, though cyanobacteria produce hydrocarbons at constitutively low levels (Coates et al. 2014), hydrocarbons are not required for growth under normal conditions (30°C, 60 µmol E m\(^{-2}\) s\(^{-1}\) light intensity) (Gibbons et al. 2018; Schirmer et al. 2010), and their physiological role is still being elucidated. Since alkanes are derived from fatty acids, which are critical membrane components, it has been hypothesized that alkanes may be important in adaptation to stress conditions via membrane modifications (Klahn et al. 2014). Indeed, subsequent studies have shown that in cyanobacterial strains containing the AAR/ADO pathway, alkanes are involved in response to certain abiotic stress conditions, such as high salinity (Kageyama et al. 2015), cold temperature (Berla et al. 2015), and nitrogen-fixing conditions (Kageyama et al. 2015). As nitrogen fixation is
a central feature in *Anabaena* 7120, I decided to further study the possible connections between alkane production and nitrogen fixation. For though alkane production was observed to increase under nitrogen fixing conditions in *Anabaena* 7120, the exact role of heptadecane in helping the cyanobacterium respond to fixed nitrogen starvation remains unknown.

To understand the possible role alkanes play in helping *Anabaena* 7120 adapt to diazotrophic growth from fixed-nitrogen dependent growth, I employed the alkane knockout mutant, DR935, which is unable to produce heptadecane (Gibbons et al. 2018). By comparing the growth of DR935 to wildtype *Anabaena* 7120 under nitrate replete and fixed-nitrogen deplete conditions, I discovered that the biggest impact to the mutant occurred immediately following fixed-nitrogen starvation. In the days after fixed-nitrogen starvation, the wildtype exhibited the highest production of heptadecane, and congruent with this time frame, the mutant had a fragmented growth phenotype. Over time as the cultures adapted to nitrogen fixing conditions, heptadecane production in the wildtype decreased and remained consistent at very low levels of production. The mutant recovered growth in filaments; however, the heterocyst frequency in the mutant remained lower than that of the wildtype. Consistent with the lower heterocyst frequency, in the long-term diazotrophic cultures, nitrogenase activity in the mutant was lower than that of the wildtype. From these results such as fragmentation and negligible nitrogenase activity of the mutant, I believe that alkanes play an important role in maintaining membrane integrity between cellular junctions, particularly in the days directly following nitrogen stepdown. After this point, when diazotrophic growth enters a steady state, it is possible
that some other cellular mechanism or compound is able to substitute for or replace heptadecane in order to maintain connections at cellular junctions.

3.3 Materials and Methods

3.3.1 Bacterial strains and culture growth

Anabaena 7120 and the alkane knockout mutant DR935 (Gibbons et al. 2018) were used in this study. DR935, a mutant derived from Anabaena 7120, contains a chromosomal knockout of the genes alr5283-5284, which encode aldehyde decarboxylase (ADO) and acyl-acyl protein reductase (AAR), responsible for heptadecane production in Anabaena 7120. A complement strain to DR935, named DR935(pZR2239) (Gibbons et al. 2018) was also used to assess gene expression via a GFP reporter gene.

All cyanobacteria cultures were grown in Bg11 (for nitrate replete conditions) and Bg110 (for fixed-nitrogen deplete conditions). Mutant DR935 contains spectinomycin resistance while complement strain DR935(pZR2239) contains spectinomycin and erythromycin resistance. Antibiotic concentrations used throughout were 10 µg mL\(^{-1}\) spectinomycin and 10 µg mL\(^{-1}\) erythromycin under nitrate replete conditions (Bg11), and 5 µg mL\(^{-1}\) spectinomycin and 5 µg mL\(^{-1}\) erythromycin under fixed-nitrogen deplete conditions (Bg110). Following nitrogen stepdown, antibiotics were not added until 48 h post-stepdown to lessen abiotic stress on the organism. Unless otherwise specified, cyanobacterial cultures were incubated for growth at 30°C, 120 rpm, under 60 µmol E m\(^{-2}\) s\(^{-1}\).
3.3.2 Heptadecane collection and extraction

Alkane production experiments were performed with wildtype *Anabaena* 7120 and DR935. The first trial was performed with *Anabaena* 7120 and DR935 grown in Bg11. The second trial was performed with cultures acclimated to nitrogen-free medium (Bg11₀); these cultures were stably grown in Bg11₀ medium (plus antibiotic for DR935) prior to the start of the experiment. The third trial was performed with wildtype and mutant cultures that underwent nitrogen stepdown 24 h prior to the start of the experiment. For all three experiments, 4 replicates of the wildtype and mutant cultures were grown in 100 mL Bg11 (1st trial) or Bg11₀ (2nd and 3rd trials) (all included antibiotic for DR935) for 17 d with an initial OD₇₀₀ of 0.03. Cultures were incubated at 30°C, 120 rpm at an aeration rate of 100 mL min⁻¹ and light intensity of 60 µE s⁻¹ m⁻². A resin column (containing 0.12 g Supelpak-2SV resin, Sigma-Aldrich) attached to each flask captured hydrocarbons from the headspace (Halfmann et al. 2014a). Hydrocarbons were extracted from the resin columns on days 5, 8, 11, 14, and 17, following the previously described method (Gibbons et al. 2018). Culture samples from these days were also used to measure chlorophyll content. Chlorophyll content was quantified at OD₆₆₅ following the previously detailed method (Houmard and de Marsac 1988).

3.3.3 Gas Chromatography-Mass Spectrometry (GC-MS) analysis of hydrocarbon samples

Hydrocarbon extractions from the cultures were analyzed following the previously detailed method (Gibbons et al. 2018).
3.3.4 AAR/ADO gene expression analysis via fluorescence microscopy

Since DR935 contains a *gfp* reporter gene for *alr5283* (ADO) and DR935(pZR2239) contains a *gfp* reporter gene for both ADO and *alr5284* (AAR), I decided to use fluorescence microscopy for the 3 conditions being studied (nitrate replete, post nitrogen stepdown, and steady state nitrogen fixation adapted) to assess alkane gene expression.

For culture growth and collection for microscopy, *Anabaena* 7120, DR935, and DR935(pZR2239) cultures were grown for 7 d in Bg11 or Bg11\(_0\) [plus antibiotic for DR935 and DR935(pZR2239)] for the nitrate replete and nitrogen fixation adapted conditions, respectively. At 7 d, when the cultures had an OD\(_{700}\) of approximately 0.5, 10 ml of each culture was harvested (4,000 xg, 10 min) and the pellets were resuspended in 300 µL of the respective medium (Bg11 or Bg11\(_0\)). Three µL of the resuspended cell cultures was placed on a microscope slide, and observed using a BX53 microscope at 40x under a GFP filter.

For analyzing *Anabaena* 7120, DR935, and DR935(pZR2239) post nitrogen stepdown, all cultures were transferred from Bg11 to Bg11\(_0\), using the nitrogen stepdown method previously described. The final OD\(_{700}\) for each culture after nitrogen stepdown was 0.3. Following nitrogen stepdown, the cultures were harvested and analyzed under a GFP filter of the BX53 microscope at 48 h and 7 d post stepdown. Since wildtype *Anabaena* 7120 should have no detection of GFP fluorescence signal, this culture grown in Bg11 was used to set the exposure time for the GFP filter. The established exposure time was 175 ms.
### 3.3.5 Acetylene reduction assay for measuring nitrogenase activity

For the acetylene reduction assay with Bg110 adapted cultures, *Anabaena 7120* and DR935 were stably grown in Bg110 (plus antibiotic for DR935) prior to the start of the experiment. Then, the cultures were refreshed and grown in Bg110 (plus antibiotic for DR935) for 7 d, following which, each culture was allocated to a 20 ml sealed vial and injected with 0.5 ml acetylene. Each vial contained 4 ml of culture (OD$_{700}$ 0.8); five replicates were made of each culture. After the culture vials were injected with 0.5 ml acetylene, they were incubated at 120 rpm, 30°C under light for 1.5 h. Following incubation, 5 ml headspace was removed from each vial and was directly injected into the GC-MS for measuring acetylene reduction as an indicator of nitrogenase activity.

For the acetylene reduction assay with Bg110 cultures following nitrogen stepdown, *Anabaena 7120* and DR935 were transferred from Bg11 to Bg110. For nitrogen stepdown, 7-day-old cultures of wildtype and mutant grown in Bg11 (plus antibiotic for DR935) were centrifuged and washed 3 times with Bg110. After the final wash, the pellets were resuspended in 1 ml Bg110 and transferred to flasks containing 100 ml Bg110, with a final culture OD$_{700}$ of 0.1. Cultures were incubated at 120 rpm, 30°C under light for 48 h. Following incubation, the cultures were centrifuged and concentrated to 0.8 OD$_{700}$, and transferred to 20 ml sealed vials, containing 4 ml culture. Again, 5 replicate vials were prepared for both *Anabaena 7120* and DR935. After sealing the vials, 0.5 ml acetylene gas was injected into each vial. The vials were incubated at 120 rpm, 30°C under light for 1.5 h. After incubation, 5 ml headspace was removed from each vial and directly injected into the GC-MS for measuring acetylene reduction. This method was repeated to measure nitrogenase activity at 2.5 weeks and 6 weeks post...
nitrogen stepdown; cultures were refreshed in Bg11₀ on a weekly basis and antibiotic was added to the DR935 culture at 48 h post nitrogen stepdown.

### 3.3.6 GC-MS analysis of acetylene reduction

The GC-MS parameters for the acetylene reduction are as follows: 5 mL of headspace gas sample from the 20-mL culture bottle was administered via a 1 mL GSV Loop to the GC-MS (Agilent 890A/5975C). The volatile compounds were separated by CP7348 column (Agilent Porabond Q 25 m × 250 µm × 3 µm) with Pulsed Split mode at 100:1 ratio at a flow rate of 0.8 mL/min, using hydrogen as a carrier gas. The GC program was initiated at 32°C, held for 4 min, and ramped at 110°C to reach 232°C. The scanning mass range of MSD was 10 to 50 m/z.

### 3.3.7 Heterocyst counts

For heterocyst counts with Bg11₀-adapted cultures, cultures of *Anabaena* 7120 and DR935 adapted to Bg11₀ were refreshed in Bg11₀ medium (plus antibiotic for DR935) and incubated for 7 d until reaching an OD₇₀₀ of 0.5. Then, 10 ml of culture was harvested (4,000 xg, 10 min) and resuspended in 500 µL Bg11₀. Microscope slides were prepared by plating 3 µl for each culture and sealing the cover glass with nail polish. Then, the slides were viewed at 40x resolution with a BX53 microscope under bright field.

For heterocyst counts with Bg11₀ post-stepdown cultures, 10 ml of culture was removed from the wildtype and mutant flasks at 2.5 weeks post-nitrogen stepdown (at the same time as nitrogenase activity was measured for the cultures). The cultures were harvested (4,000 xg, 10 min), and the cells were resuspended in 500 µL Bg11₀. Then, 3 µL of each suspension was plated on a microscope slide beneath a cover glass sealed with...
nail polish. Microscope pictures were taken using a BX53 microscope at 40x under bright field.

Using the microscope images, heterocysts and total cell numbers were counted for both \textit{Anabaena} 7120 and DR935. A total of 10,000 cells was counted for each culture type. Heterocyst frequency was calculated by dividing the number of heterocysts by the total number of cells counted.

### 3.4 Results

#### 3.4.1 Heptadecane production in fixed-nitrogen replete (Bg11), fixed-nitrogen starvation (Bg11$_0$) and Bg11$_0$-adapted cultures

When grown in Bg11 (plus antibiotic for DR935) under normal growth conditions (30°C, 120 rpm, under 60 µmol E m$^{-2}$ s$^{-1}$), heptadecane production tended to increase over the course of the 17 d growth period (note that the first time point for heptadecane extraction was over a 5 d period while other time points were over a 3 d period) (Fig. 3.1, column 3). By looking at the growth curve of the culture throughout the course of the experiment, the chlorophyll content continues to increase, indicating that the culture is still in active growth phase (Fig. 3.2, column 3). These results may indicate that heptadecane production is positively correlated to culture density, potentially due to membrane synthesis as the cells multiply.

Interestingly, the highest heptadecane production occurred in the Bg11$_0$ post nitrogen stepdown culture in days 0-5 immediately following nitrogen stepdown (Fig. 3.1, column 1). Heptadecane production remains higher in the stepdown culture compared to the Bg11 and Bg11$_0$-adapted cultures through day 11. However, after this point, heptadecane production is higher in the Bg11 culture.
Among the three data sets, heptadecane production in the Bg11$_0$-adapted culture is the lowest (Fig. 3.1, column 2). This data is consistent with the decrease in heptadecane production seen in the Bg11$_0$ post nitrogen stepdown culture in days 11-17. Thus, while heptadecane production is the highest in Anabaena 7120 directly following nitrogen stepdown, once the culture becomes adapted to nitrogen fixation conditions, heptadecane production decreases. Furthermore, once production decreases in the Bg11$_0$ cultures, it is lower than that of the Bg11 culture, whose production continued to increase throughout the course of the experiment.

**Fig. 3.1** Total heptadecane production in wildtype Anabaena 7120 in fixed nitrogen starvation (Bg11$_0$), Bg11$_0$-adapted, and fixed-nitrogen replete cultures. The fixed nitrogen starvation culture came from Bg11 medium and underwent nitrogen stepdown to Bg11$_0$ medium at the beginning of the experiment. The Bg11$_0$-adapted culture came from a culture stably grown in Bg11$_0$ prior to the start of the experiment. The fixed-nitrogen replete culture came from a culture grown in Bg11 prior to the start of the experiment. Heptadecane (ng) produced per µg chlorophyll was measured on days 0-5, 6-8, 9-11, 12-14, and 15-17. The first data set shows production from Anabaena 7120 directly following nitrogen stepdown. The middle column depicts production from Anabaena 7120 that has been adapted to nitrogen-fixing conditions. The final column represents heptadecane production from Anabaena 7120 grown in nitrate-replete conditions.
By comparing the growth curves of *Anabaena* 7120 grown under all three conditions, we see that the cultures remain in the active growth phase throughout the experiment (Fig. 3.2). Interestingly, as noted above, the Bg11 culture increased heptadecane production as the culture grew. However, this phenomenon did not occur in the Bg11<sub>0</sub>-grown cultures. In the fixed nitrogen starved culture, heptadecane production increased immediately, after day 5, it decreased even as the culture continued to grow. Moreover, in the Bg11<sub>0</sub>-adapted culture, heptadecane production remained at a consistently low level even as the chlorophyll content increased. Thus, with fixed-nitrogen, a positive correlation exists between culture density and heptadecane production. However, this correlation does not appear to exist when *Anabaena* 7120 is grown in the absence of fixed nitrogen.

### 3.4.2 Growth of *Anabaena* 7120 and DR935 in fixed-nitrogen starvation (Bg11<sub>0</sub>) and Bg11<sub>0</sub>-adapted cultures

Since heptadecane production significantly increased in the wildtype following fixed-nitrogen starvation for 48 h, I was curious if the absence of heptadecane production had any impact on culture growth. By comparing culture chlorophyll content measured throughout the course of the alkane production experiment, I determined that the alkane knockout mutant DR935 actually had higher chlorophyll content than the wildtype in days 0 to 5 for all growth conditions, and after 5 days DR935 had comparable chlorophyll content to the wildtype (Fig. 3.2). Thus, the heptadecane does not appear to impact *Anabaena* 7120’s ability to grow even when fixed nitrogen is unavailable in the environment.
Fig. 3.2 Growth curves comparing wildtype and mutant (DR935) in fixed nitrogen starvation (Bg11), Bg110-adapted, and fixed-nitrogen replete cultures. The fixed nitrogen starvation culture came from Bg11 medium and underwent nitrogen stepdown to Bg110 medium at the beginning of the experiment. The Bg110-adapted culture came from a culture stably grown in Bg110 prior to the start of the experiment. The fixed-nitrogen replete culture came from a culture grown in Bg11 prior to the start of the experiment. Chlorophyll content was measured on days 0, 5, 8, 11, 14, and 17, the same time points that were used to quantify heptadecane production in the cultures.

3.4.3 AAR/ADO gene expression in fixed-nitrogen replete (Bg11), fixed-nitrogen starvation (Bg110), and Bg110-adapted cultures

The increase in alkane production following nitrogen stepdown could correspond to an increase in the alkane gene (aar/ado) expression. Additionally, it is possible to gain insight into the role alkanes play in adapting to nitrogen fixing conditions through evaluating their expression. For example, if the alkane genes were expressed solely in heterocysts, we may hypothesize that alkanes are involved in heterocyst development or function. To assess any differences in aar/ado expression, I compared gfp expression in DR935 (which contains a chromosomal Pado-gfp transcriptional fusion) and complement DR935(pZR2239) (which contains the chromosomal Pado-gfp transcriptional fusion and
a plasmid-based \textit{Paar-gfp} transcriptional fusion) under nitrate replete and nitrogen-fixing conditions (Fig. 3.3).

Though DR935 contains a \textit{gfp} reporter gene whose expression is controlled by the promoter of \textit{alr5283 (ado)}, no \textit{gfp} expression was observed under any of the conditions (Fig. 3.3, row 2). This may be reflective of the very low level expression of \textit{ado}. It also indicates that even though heptadecane production increased directly following nitrogen stepdown, since \textit{ado} expression was not upregulated, the increase in heptadecane production may be controlled at some other level than gene transcription.

Complement DR935(pZR2239) contained the same \textit{gfp} fusion to \textit{ado} in DR935 chromosome as the mutant strain. Additionally, the complement plasmid pZR2239 contained a \textit{gfp} fusion to \textit{Palr5284 (aar)}. Compared to DR935, \textit{gfp} expression in DR935(pZR2239) was much higher under all three conditions (Bg11, Bg11$_0$ post stepdown, and Bg11$_0$ adapted). The higher expression of \textit{gfp} in the complement compared to the mutant is likely due to the multiple plasmid copies in the complement from which \textit{gfp} was expressed, whereas the mutant only contained a chromosomal copy. Therefore, the \textit{gfp} expression seen in the complement is likely attributed to \textit{aar}-directed \textit{gfp} transcription rather than \textit{ado}-directed \textit{gfp} transcription. The levels of \textit{gfp} expression did not vary greatly amongst the different culture conditions (Fig. 3.3, row 3). Notably, at 48 h post nitrogen stepdown, \textit{gfp} expression was seen more prominently in heterocysts compared to vegetative cells. However, at 7 d post stepdown and in nitrogen-fixation adapted conditions, fluorescence was visible in both vegetative cells and heterocysts, with no difference in intensity. Thus, alkane production may selectively aid in heterocyst development or function immediately following fixed nitrogen starvation. But, its role
may become less selective or important to heterocysts once the culture has become acclimated to nitrogen fixation.

**Fig. 3.3** Alkane gene (*aar/ado*) expression in fixed nitrogen starvation (Bg11$_0$) at 48 h and 7 d, Bg11$_0$-adapted, and fixed-nitrogen replete (Bg11) cultures. Alkane gene expression was monitored through GFP fluorescence. Wildtype served as a negative control. Mutant DR935 contains a chromosomal copy of *Pado-gfp* transcriptional fusion. Complement DR935(pZR2239) contains the chromosomal copy of the *Pado-gfp* fusion as well as plasmid copies of *Paar-gfp* transcriptional fusion. For each culture under each condition, a bright-field (BF) and GFP image were captured.

### 3.4.4 Fragmentation of DR935 following fixed-nitrogen starvation

Although growth was not negatively impacted (as indicated by chlorophyll content) by the absence of heptadecane under diazotrophic conditions, the phenotype of the mutant DR935 was affected by nitrogen starvation. Directly following nitrogen stepdown, DR935 exhibited fragmentation (Fig. 3.3, column 1 DR935) compared to the
wildtype, which grew in filaments (Fig. 3.3, column 1 WT). The mutant began to recover, forming some filaments following 7 d in nitrogen fixing conditions (Fig. 3, column 2). In the stably grown diazotrophic culture (Fig. 3.3, column 3) the mutant grew in filaments and appeared phenotypically similar to the wildtype.

### 3.4.5 Nitrogenase activity in *Anabaena* 7120 and DR935

Nitrogenase activity in *Anabaena* 7120 and DR935 varied following nitrogen stepdown (Fig. 3.4). In the first 48 h following nitrogen stepdown, nitrogenase activity in the wildtype was significantly higher than in the mutant. However, at 2.5 weeks after nitrogen stepdown, nitrogenase was higher in the mutant than in the wildtype. At 6 weeks post-stepdown, the cultures reached a stage where nitrogenase activity was similar between mutant and wildtype. However, in long-term nitrogen-fixing cultures, nitrogenase activity in the wildtype was consistently higher than that in the mutant.

![Nitrogenase activity in wildtype and mutant (DR935) in fixed-nitrogen starvation (Bg110) and Bg110-adapted cultures.](image)

**Fig. 3.4** Nitrogenase activity in wildtype and mutant (DR935) in fixed-nitrogen starvation (Bg110) and Bg110-adapted cultures. Nitrogenase activity was measured indirectly as indicated by the amount of acetylene converted into ethylene (nmol ethylene/µg chl·h) by the nitrogenase enzyme. Following nitrogen stepdown, nitrogenase activity was measured in the wildtype and DR935 at 48 h, 2.5 weeks, and 6 weeks (antibiotic was added to DR935 at 48 h; cultures refreshed in Bg110 medium on weekly basis). In the – N (Bg110) cultures, wildtype and DR935 stably grown Bg110 (plus antibiotic for DR935)
were refreshed in Bg110 medium (plus antibiotic for DR935) and allowed to grow for 1 week before nitrogenase activity was measured.

3.4.6 Heterocyst counts following fixed-nitrogen starvation and in long-term nitrogen-fixing cultures

Since both the phenotype and nitrogenase activity were affected in DR935, I compared the heterocyst frequency in DR935 with the wildtype to see if it was also impacted by the absence of alkane production. At 2.5 weeks, I observed that nitrogenase activity was higher in the mutant compared to the wildtype. Higher nitrogenase activity in the mutant could be attributed to more heterocysts in the mutant compared to the wildtype. Yet, when I counted 10,000 cells for both the mutant and wildtype cultures, the mutant had a heterocyst frequency of 1.64% while the wildtype had a heterocyst frequency of 1.87%. However, there were more single cells in the mutant culture compared to the wildtype, which may have been detached heterocysts. These cells were not counted since they were not part of a filament, and thus were considered dying cells because neither unicellular heterocysts nor vegetative cells are able to survive on their own in diazotrophic conditions. Alternatively, it is possible that at 2.5 weeks, the mutant had higher nitrogenase activity because the nitrogenase activity on a per heterocyst basis was higher. In long-term Bg110-adapted cultures, nitrogenase activity was consistently higher in the wildtype compared to the mutant. Correspondingly for these cultures the wildtype had a higher heterocyst frequency (3.87%) than the mutant (2.63%).

3.5 Discussion

Though heptadecane production in Anabaena 7120 is not required for growth under normal conditions (Gibbons et al. 2018), other studies have indicated that alkane
production increases in response to certain stress conditions, such as nitrogen limitation (Kageyama et al. 2015). Although alkane production has been reported to increase under diazotrophic conditions, the role the alkanes play in response to fixed-nitrogen starvation or N₂-fixation in *Anabaena* 7120 remains unknown. To gain a better understanding of the role alkanes may play in response to nitrogen limitation, I compared wildtype *Anabaena* 7120’s response to starvation for fixed-nitrogen and long-term diazotrophic growth with that of the alkane knockout mutant DR935.

Firstly, I assessed alkane production in *Anabaena* 7120 grown in Bg11, Bg11₀ post nitrogen stepdown, and Bg11₀-adapted conditions. A previous study reported a 2-fold increase in alkane production in the initial 5 days following nitrogen stepdown in *Anabaena* 7120 (Kageyama et al. 2015). My results revealed a similar increase in alkane production in the first 5 days following nitrogen stepdown. During this period (at 48 h), the wildtype exhibited nitrogenase activity while the mutant did not. Given the increased alkane production following fixed nitrogen starvation and the inability of the mutant to fix nitrogen at this time, I believe heptadecane plays a critical role in helping *Anabaena* 7120 adapt to nitrogen starvation at the early onset.

Interestingly, when *Anabaena* 7120 was grown in the presence of fixed nitrogen, alkane production increased on per cell basis throughout the course of the experiment, suggesting a positive correlation between culture density and alkane production. However, when *Anabaena* 7120 was grown in nitrogen-fixing conditions, after the initially (days 0-5, Fig. 1) high production of alkanes, the production level of alkanes leveled off. This trend can be seen in days 11-17 of the post-stepdown culture as well as in all of the time points for the Bg11₀-adapted culture. By looking at the growth curves
for the Bg11, Bg11_0 post-stepdown and Bg11_0-adapted cultures, we see that all three cultures remain in an active growth phase throughout the experiment. It is possible that since nitrogen fixation requires so much energy, the cell diverts energy away from other anabolic pathways, such as alkane production. The drastic initial increase in alkane production in response to fixed-nitrogen starvation followed by a significant decrease in alkane production also suggests that heptadecane is primarily important in the initial response to starvation for fixed-nitrogen. After this initial response, heptadecane’s role appears to become less important or is replaced by some other compound or mechanism.

Overall growth of DR935 (as indicated by chlorophyll content) was not negatively affected in nitrogen-fixing conditions compared to wildtype. This suggested that the increase in alkane production following fixed-nitrogen starvation may more specifically be affecting heterocyst cells in their development or the process of heterocyst-specific oxic nitrogen fixation. To determine if this was the case, I used DR935, which contains a gfp transcriptional fusion to promoter of ado (Pado-gfp), and the complement strain DR935(pZR2239), which contains a gfp transcriptional fusion to promoter of aar (Paar). Only the complement strain had GFP fluorescence observed (Fig. 3), likely because the plasmid containing the gfp fusion gene was present in multiple copies. The undetectable levels of GFP fluorescence in the mutant, which contained fewer copies of chromosomal fusion gfp, indicate low-level expression of the ado promoter. The low expression level seen for ado was unsurprising as cyanobacteria naturally produce small amounts of alkanes (Coates et al. 2014; Schirmer et al. 2010)

By observing the complement cultures grown in Bg11_0 post-stepdown, Paar-gfp had higher expression in heterocysts compared to vegetative cells at 48 h (Fig. 3).
However, at 7 d following nitrogen stepdown and in the Bg110-acclimated culture, \( Paar-gfp \) expression was similar in both vegetative cells and heterocysts (Fig. 3). These findings indicate that alkane production may be more specifically beneficial to heterocysts immediately following fixed-nitrogen starvation. However, once the culture has adapted to nitrogen-fixing conditions, alkanes may not be as selectively beneficial to heterocysts.

I also observed the cellular morphology of \( Anabaena \) 7120 and DR935 following nitrogen stepdown. Immediately following nitrogen stepdown for 48 h, the mutant grew in single cells or short fragments while the wildtype grew in long filaments. The mutant began to recover and showed some filamentous growth 7 d following nitrogen stepdown. However, it was not until 6+ weeks following nitrogen stepdown that the mutant recovered filamentous growth similar to the wildtype. Combined with the increase in alkane production in the wildtype that corresponds to the initial 11 d period following nitrogen stepdown, I believe that alkanes may play a critical role in helping to maintain membrane integrity following nitrogen stepdown and during the culture’s initial acclimation to nitrogen-fixing conditions.

Since the phenotype of the mutant was affected following nitrogen stepdown, I measured nitrogenase activity in \( Anabaena \) 7120 and DR935 to see if nitrogenase activity was impacted. At 48 h following nitrogen stepdown, nitrogenase activity in the wildtype was much higher than in the mutant. Since the mutant had fragmented growth, with many short filaments and single cells, I was unsurprised that nitrogenase activity was lower as the mutant appeared to be struggling to adapt during this time. However, at 2.5 weeks when the mutant began to grow in filaments, nitrogenase activity in the mutant was
higher than that of the wildtype. When I compared the heterocyst frequencies of the mutant and wildtype cultures at this time point, the mutant had a lower heterocyst frequency (1.64%) compared to the wildtype (1.87%). However, there were many single cells present in the mutant culture, which could have been heterocysts, but were not counted since they were not part of a filament. If these cells were detached heterocysts, they would contribute to the nitrogenase activity measured in the mutant. The detached heterocyst cells present in the mutant culture may also reflect the importance of the role of alkanes in helping to maintain membrane integrity. The junction between heterocysts and vegetative cells may be particularly supported by alkanes. This could explain how in the absence of alkanes, heterocysts are present apart from the filament, where they are unable to survive or replicate on their own.

In long-term nitrogen-fixing cultures, the wildtype exhibits higher nitrogenase activity than the mutant. When heterocyst frequency was calculated for *Anabaena* 7120 and DR935 in Bg110-adapted cultures, more heterocysts were present in the wildtype (3.87%) than in the mutant (2.63%). Thus, the higher nitrogenase activity in the wildtype could be accounted for by the increased number of heterocysts. After long-term adaptation to nitrogen-fixing conditions the mutant grew similarly to the wildtype in long filaments with few single cells or shorter filaments present. I speculate that some other mechanism or compound becomes more important in maintaining membrane integrity after the initial period following nitrogen stepdown. Since alkanes are still produced in later stages following nitrogen stepdown at very low levels, this could still explain the fewer number of heterocysts present in the mutant compared to the wildtype in Bg110-adapted cultures.
Interestingly, ADO requires oxygen in its conversion of the fatty aldehyde to the alkane (Li et al. 2012; Li et al. 2011). Therefore, oxygen consumption by ADO could also contribute to increased nitrogenase activity in the wildtype compared to the mutant. Since nitrogenase is inactivated by oxygen, ADO’s possible involvement to creating a micro-oxic environment in heterocysts could facilitate nitrogenase activity.

In conclusion, under normal growth conditions, alkanes appeared to have no major impact on cell growth or phenotype. However, following nitrogen stepdown, the mutant DR935 became fragmented. This fragmentation coincided with the time period over which the wildtype has a significant increase in alkane production. Combined, these two pieces of evidence suggest that alkanes play an important role in maintaining membrane integrity following nitrogen stepdown. Furthermore, membrane integrity supported by alkanes appears to be more important in the junction between heterocysts and vegetative cells because even once the mutant culture began to recover, detached heterocysts were still present. Moreover, in long-term diazotrophic cultures, the mutant still had a lower heterocyst frequency than the wildtype. This could be attributed to the unstable connection between heterocysts and vegetative cells, making it difficult for heterocysts to remain associated with the filament. Ultimately, since heterocysts are non-dividing cells and rely on vegetative cells for fixed carbon, detached heterocysts are unable to survive, thus decreasing the heterocyst frequency in the mutant in long-term diazotrophic cultures.

Interestingly, since the mutant did recover filamentous growth similar to the wildtype after prolonged exposure to oxic N$_2$-fixation conditions, this suggests that alkanes play a more crucial role immediately following nitrogen stepdown. This
coincides with the decrease in alkane production that occurs following the initial response to alkane production. Thus, while alkanes are important in adaption to nitrogen deficiency, *Anabaena* 7120 is still able to overcome the initial damage incurred due to the absence of alkanes. More research may discover alternate pathways that may help the cell recover and maintain membrane integrity in the absence of alkanes. It may also show how alkanes are involved in strengthening the connection between cells, particularly heterocyst and vegetative cells.
CHAPTER 4: Identification of Two Genes Required for Oxic N\textsubscript{2} Fixation in 

*Anabaena* sp. PCC 7120

4.1 Abstract

*Anabaena* sp. PCC 7120 is a filamentous cyanobacterium that forms heterocysts, specialized cells required for oxic nitrogen fixation. A previous proteomic study identified 57 proteins exclusively or primarily present in heterocysts of *Anabaena cylindrica* ATCC29414 (Qiu 2018). Homologs to 16 of the protein-coding genes were identified in *Anabaena* 7120. Since *Anabaena cylindrica* is not genetically tractable, I knocked out these 16 genes, via a single crossover approach, to determine their roles in heterocysts in genetically tractable *Anabaena* 7120. Of the 16 mutants, two were unable to grow in the absence of fixed nitrogen, suggesting these gene products (All3132 and Alr0731) are necessary for heterocyst development or function. Following nitrogen stepdown, the *alr0731* mutant formed bulging cells that appeared to be dying, and at 16 d post stepdown the culture was mostly dead cells with a few unicellular living cells. This mutant has been successfully complemented, thereby establishing the *alr0731* mutant as a Fox mutant (unable to fix nitrogen in the presence of oxygen). Interestingly, *alr0731* encodes an anaerobic ribocucleoside triphosphate reductase activating protein involved in producing dNTPs from NTPs. Cell division and chromosomal rearrangements are tied to heterocyst differentiation. As both processes require dNTPs, I postulate that Alr0731 is primarily important in heterocyst differentiation and may be expressed in pro-heterocysts more so than in mature heterocysts. Future work using a *gfp* marker tag in the mutant and complement may help elucidate when and where *alr0731* is primarily expressed.
4.2 Introduction

Nitrogen fixation is a globally important process. Organisms require nitrogen in many biomolecules necessary for growth and survival, such as nucleic acids and proteins. Though the earth’s atmosphere is 78% dinitrogen, this form of nitrogen is not bioavailable to most living organisms. Indeed, nitrogen is a limiting resource in many ecosystems (Bernhard 2010). Certain organisms, such as some species of cyanobacteria, are able to fix atmospheric dinitrogen into ammonia, thereby making it available for other organisms to assimilate.

Diazotrophic cyanobacteria are unique in that they must harmonize two incompatible processes: oxygen-producing photosynthesis and oxygen-labile nitrogen fixation. Two main solutions are employed to solve the dilemma: temporal or spatial separation of photosynthesis and nitrogen fixation. Anabaena sp. PCC 7120 (hereafter Anabaena 7120) is a filamentous cyanobacterium, which is capable of fixing dinitrogen to ammonia when fixed nitrogen is unavailable in the environment. This species of cyanobacteria uses spatial separation to operate photosynthesis and nitrogen fixation simultaneously. When fixed nitrogen is limiting, Anabaena 7120 differentiates specialized cells called heterocysts. Heterocysts are terminally differentiated cells whose sole function is to reduce atmospheric dinitrogen to ammonia (i.e., nitrogen fixation) (Wolk 1982). These cells differentiate from vegetative cells at approximately 1 heterocyst per 15 vegetative cells in a filament (Wolk et al. 1994). Under this system, vegetative cells perform photosynthesis while heterocysts perform nitrogen fixation. In this symbiotic arrangement, vegetative cells provide heterocysts with fixed carbon while heterocysts provide vegetative cells with fixed nitrogen.
Because nitrogenase, the enzyme complex for nitrogen fixation, can be deactivated by oxygen in seconds or minutes (Dixon and Wheeler 1986; Fay 1992), heterocyst differentiation must create an environment that houses and protects nitrogenase from oxygen. To accomplish this, as heterocysts differentiate from vegetative cells, they undergo systematic changes in gene expression. Through differentiation, heterocysts develop a thicker outer envelope to prevent oxygen from entering the cell, they degrade PSII (the O$_2$-evolving complex in photosynthesis), and they increase their respiratory capacity to remove oxygen within the cell (Wolk et al. 1994). Once a heterocyst is formed, its gene expression remains distinct from that of a vegetative cell as the two cell types are responsible for different metabolic outputs. In *Anabaena* sp. it is estimated that approximately 15-25% of the DNA sense strand is transcribed solely in heterocysts (Lynn et al. 1986). However, research efforts are still deciphering which of the genes expressed in heterocysts are absolutely essential for its development or function.

In a previous proteomic study with the filamentous, heterocystous cyanobacterium *Anabaena cylindrica*, our lab identified 57 heterocyst-specific proteins (Qiu 2018). In this research, I chose 16 of the genes encoding the heterocyst-specific proteins to study in *Anabaena* 7120, primarily because *Anabaena* 7120 is more readily genetically manipulated than *A. cylindrica*. My goal was to determine if any of the proteins are required for heterocyst formation or function since they were specifically found in heterocysts. Table 4.1 lists the 16 genes found to be exclusively or primarily present in heterocysts in *A. cylindrica* alongside their homologs in *Anabaena* 7120.
Many genes required for heterocyst differentiation and function have been identified through creating mutations in specific genes and screening for mutant inability to grow in aerobic nitrogen-fixing conditions. HetR is a master transcription factor specifically required for heterocyst differentiation (Buikema and Haselkorn 1991; Huang et al. 2004; Zhou et al. 1998). Several other regulatory genes such as nrrA (Ehira and Ohmori 2011), ccbP (Hu et al. 2011), hetN (Higa et al. 2012), hetF, patA (Risser and Callahan 2008), patN (Risser et al. 2012), patU (Meeks et al. 2002), hetZ (Zhang et al. 2007), patS (Hu et al. 2015; Yoon and Golden 1998), hepK (Zhou and Wolk 2003) and hetP (Videau et al. 2016) were also found to play critical roles during heterocyst differentiation.

Mutations that are lethal to cells in aerobic nitrogen-fixing conditions signify that the mutated gene is a Fox gene (unable to fix nitrogen in the presence of oxygen). Fox

<table>
<thead>
<tr>
<th>A. cylindrica ATCC29414 gene (orf)</th>
<th>A. cylindrica PCC7122 gene ID</th>
<th>Anabaena 7120 homolog</th>
</tr>
</thead>
<tbody>
<tr>
<td>acy3366</td>
<td>Anacy_4931</td>
<td>alr0483</td>
</tr>
<tr>
<td>acy1692</td>
<td>Anacy_5197</td>
<td>ali1076</td>
</tr>
<tr>
<td>acy6006</td>
<td>Anacy_0028</td>
<td>ali1523</td>
</tr>
<tr>
<td>acy4151</td>
<td>Anacy_4151</td>
<td>ali1681</td>
</tr>
<tr>
<td>acy5248</td>
<td>Anacy_3102</td>
<td>ali3132</td>
</tr>
<tr>
<td>acy2435</td>
<td>Anacy_5577</td>
<td>ali3680</td>
</tr>
<tr>
<td>acy6556</td>
<td>Anacy_0810</td>
<td>ali4106</td>
</tr>
<tr>
<td>acy4599</td>
<td>Anacy_4264</td>
<td>ali5173</td>
</tr>
<tr>
<td>acy6462</td>
<td>Anacy_2072</td>
<td>ali0618</td>
</tr>
<tr>
<td>acy338</td>
<td>Anacy_3112</td>
<td>ali0731</td>
</tr>
<tr>
<td>acy1261</td>
<td>Anacy_4252</td>
<td>ali0765</td>
</tr>
<tr>
<td>acy3023</td>
<td>Anacy_3471</td>
<td>ali2325</td>
</tr>
<tr>
<td>acy4599</td>
<td>Anacy_4264</td>
<td>ali2428</td>
</tr>
<tr>
<td>acy1391</td>
<td>Anacy_1682</td>
<td>ali2494</td>
</tr>
<tr>
<td>acy4894</td>
<td>Anacy_0232</td>
<td>ali3125</td>
</tr>
<tr>
<td>acy3769</td>
<td>Anacy_2849</td>
<td>ali3829</td>
</tr>
</tbody>
</table>
mutants are generally readily distinguished by their yellowing growth phenotype following nitrogen stepdown under aerobic conditions; ultimately, these mutants cease to grow and die when placed in aerobic diazotrophic conditions (Ernst et al. 1992).

Though Fox mutants were originally described as mutants unable to fix nitrogen under aerobic conditions, other research has suggested that the definition be broadened to mutants “requiring fixed nitrogen for growth in the presence of oxygen” (Lechno-Yossef et al. 2011). The latter definition was proposed because one Fox mutant for the gene conR was identified and later found to show some nitrogenase activity for a short period after diazotrophy begins (Fan et al. 2006; Mella-Herrera et al. 2011). In either case, Fox mutants are those that have gene mutations which are lethal in aerobic nitrogen-fixing conditions. Therefore, heterocyst development or function is severely impacted by the absence of the functional gene product in these mutants. So far there are 85 Fox genes identified in *Anabaena* 7120 (Table 4.2) (Lechno-Yossef et al. 2011).

### Table 4.2 Fox genes of *Anabaena* sp. PCC 7120

<table>
<thead>
<tr>
<th>Fox Gene</th>
<th>ORF</th>
<th>References and Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>hcvA</td>
<td>all0093</td>
<td>(Zhu et al. 2001)</td>
</tr>
<tr>
<td>hepN</td>
<td>alr0117</td>
<td>(Fan et al. 2006; Ning and Xu 2004)</td>
</tr>
<tr>
<td>conR</td>
<td>all0187</td>
<td>(Fan et al. 2006; Mella-Herrera et al. 2011)</td>
</tr>
<tr>
<td>anaerobic ribonucleoside-triphosphate reductase activating protein</td>
<td>alr0731</td>
<td>This study</td>
</tr>
<tr>
<td>devR, devR&lt;sub&gt;a&lt;/sub&gt;</td>
<td>alr0442</td>
<td>(Campbell et al. 1996) in <em>Nostoc punctiforme</em>; (Zhou and Wolk 2003)</td>
</tr>
<tr>
<td>hglK</td>
<td>all0813</td>
<td>(Black et al. 1995) and M8 of (Ernst et al. 1992)</td>
</tr>
<tr>
<td>invB</td>
<td>alr0819</td>
<td>(Lopez-Igual et al. 2010; Vargas et al. 2011)</td>
</tr>
<tr>
<td>putative glycosyl transferase</td>
<td>alr1000</td>
<td>(Videau et al. 2014)</td>
</tr>
<tr>
<td>henR</td>
<td>alr1086</td>
<td>(Fan et al. 2006)</td>
</tr>
<tr>
<td>asp-glu-rich product</td>
<td>all1338</td>
<td>(Lechno-Yossef et al. 2011)</td>
</tr>
<tr>
<td>nifN</td>
<td>all1437</td>
<td>(Haselkorn 1992; Haselkorn and Buikema 1992)</td>
</tr>
<tr>
<td>nifE</td>
<td>all1438</td>
<td>(Haselkorn 1992; Haselkorn and Buikema 1992)</td>
</tr>
<tr>
<td>nifK</td>
<td>all1440</td>
<td>(Mazur and Chui 1982)</td>
</tr>
<tr>
<td>xisA</td>
<td>alr1442</td>
<td>(Golden and Wiest 1988)</td>
</tr>
<tr>
<td>nifD</td>
<td>all1454</td>
<td>(Lammers and Haselkorn 1983)</td>
</tr>
<tr>
<td>nifH</td>
<td>all1455</td>
<td>(Mevarech et al. 1980)</td>
</tr>
<tr>
<td>xisF</td>
<td>alr1459</td>
<td>(Carrasco et al. 1994; Golden et al. 1992; Kuritz et al. 1993)</td>
</tr>
</tbody>
</table>


- **nifB**: all1517 (Lyons and Thiel 1995) in *Anabaena variabilis* ATCC 29413
- **one of two hisDs**: all1591 (Lechno-Yossef et al. 2011)
- **fraH**: alr1603 (Merino-Puerto et al. 2010)
- **putative RND efflux transporter**: alr1656 (Hahn et al. 2013)
- **RIP**: alr1728 (Lechno-Yossef et al. 2011)
- **prpJ1**: all1731 (Jang et al. 2009; Jang et al. 2007)
- **abp2**: all1939 (Koksharova and Wolk 2002)
- **fraG, sepJ**: alr2338 (Flores et al. 2007; Nayar et al. 2007)
- **hetR**: alr2339 (Buikema and Haselkorn 1991)
- **fraC**: alr2392 (Bauer et al. 1995; Merino-Puerto et al. 2010)
- **fraD**: alr2393 (Merino-Puerto et al. 2010)
- **fraE**: alr2394 (Merino-Puerto et al. 2010)
- **patB**: all2512 (Jones et al. 2003; Liang et al. 1993)
- **hepS**: all2760 (Fan et al. 2006)
- **hetC**: alr2817 (Khudyakov and Wolk 1997)
- **hetP**: alr2818 (Fernández-Piñas et al. 1994; Higa and Callahan 2010)
- **Hep region**: alr2825 (Huang et al. 2005)
- **Hep region**: alr2827 (Huang et al. 2005)
- **Hep region**: alr2831 (Huang et al. 2005)
- **Hep region**: all2833 (Huang et al. 2005)
- **hepC**: alr2834 (Zhu et al. 1998)
- **hepA (earlier, hetA)**: alr2835 (Holland and Wolk 1990; Wolk 2000; Wolk et al. 1988)
- **Hep region**: alr2837 (Huang et al. 2005)
- **Hep region**: alr2839 (Huang et al. 2005)
- **Hep region**: alr2841 (Huang et al. 2005)
- **hgdD**: alr2887 (Moslavac et al. 2007; Wolk et al. 2007)
- **PBP6**: all2981 (Leganés et al. 2005)
- **ABC transporter ATP-binding protein**: all3132 This study; awaiting complementation
- **putative RND efflux transporter**: alr3143 (Hahn et al. 2013)
- **putative cell wall organization**: all3278 (Lechno-Yossef et al. 2011)
- **heterocysts divide internally**: all3520 (Lechno-Yossef et al. 2011)
- **hetF; hetF_α**: alr3546 (Wong and Meeks 2001) in *N. punctiforme*; (Wolk et al. 2007)
- **conserved, vacuolate**: all3582 (Lechno-Yossef et al. 2011)
- **abp3**: alr3608 (Koksharova and Wolk 2002)
- **hepB**: alr3698 (Wolk et al. 1988; Wolk et al. 1999); polar effect not disproven, but unlikely
- **glycosyl transferase**: alr3699 (Wang et al. 2007)
- **devB**: alr3710 (Fiedler et al. 1998); polar effect not disproven, but unlikely
- **devC**: alr3711 (Fiedler et al. 1998)
- **devA**: alr3712 (Maldener et al. 1994)
- **prpA**: alr3731 (Zhang et al. 2005); see also (Zhang et al. 1998)
- **pknE**: alr3732 (Zhang et al. 2005); see also (Zhang et al. 1998)
In this research, I used a single crossover approach to knockout the heterocyst-specific protein-coding genes to identify Fox genes in Anabaena 7120. The single crossover method previously established in Anabaena 7120 (Chen et al. 2015) requires cloning an internal fragment of the target gene into an integrative cargo plasmid vector. By introducing the resulting cargo plasmid to Anabaena 7120 through conjugative transformation (Chen et al. 2016a), the internal fragment in the plasmid homologously recombines with its matching sequence of the target gene in the chromosome (Chen et al. 2015).
2015; Vermaas 1996). Via homologous recombination, the integrative plasmid disrupts the gene, leaving two truncated genes, the 5' and 3' portions of the gene (Fig. 4.1) within the mutated chromosome (Carrasco et al. 1994).

Through mutating and screening the 16 genes, I identified two potential Fox genes: *all3132* and *alr0731*. The mutants for both gene knockouts were unable to grow in the absence of fixed nitrogen. By complementing the *alr0731* mutant, the strain regained its ability to grow under diazotrophic conditions. These results indicate that *alr0731* is a Fox gene. Complementation of the *all3132* mutant is still ongoing, but preliminary results suggest that *all3132* is a potential Fox gene as well.

### 4.3 Materials and Methods

#### 4.3.1 Bacterial strains and plasmids

Plasmid construction and maintenance were performed using *Escherichia coli* strains Top10 (Invitrogen) and NEB10β (New England Biolabs). Strains were grown in Luria-Bertani broth, using the following antibiotic concentrations: 100 µg mL⁻¹ ampicillin, 50 µg mL⁻¹ kanamycin, 100 µg mL⁻¹ spectinomycin, and 25 µg mL⁻¹ chloramphenicol. For studies in cyanobacteria, *Anabaena* 7120 was used as the wildtype strain into which mutations were introduced. *Anabaena* 7120 and its mutant derivatives were grown in Bg11 (+N) or Bg11₀ (-N) media. In Bg11 medium, the following antibiotic concentrations were used: 100 µg mL⁻¹ neomycin, 10 µg mL⁻¹ spectinomycin, 10 µg mL⁻¹ erythromycin. In Bg11₀ medium the same antibiotics were used at lower concentrations (25 µg mL⁻¹ neomycin, 5 µg mL⁻¹ spectinomycin, 5 µg mL⁻¹ erythromycin).
4.3.2 Construction of knockout plasmids for heterocyst-specific genes

A single crossover approach was chosen to knockout the heterocyst-specific protein-coding genes in *Anabaena* 7120. For this method, an internal fragment (IF) of each gene was PCR amplified (for primers, see Table 4.3) and cloned into pCR®TOPO2.1® vector (TOPO TA Cloning® kit, Invitrogen). Primers with restriction enzyme recognized sequences were used to amplify each IF. After the IFs were cloned to into pCR®TOPO2.1® vector, the respective restriction enzymes were used to excise the IFs and insert them into integration vector pZR606 (see Table 4.3 for details of plasmid construction). Immediately downstream of the IF insertion into the multiple cloning site (MCS) of pZR606 is a *gfp* gene that serves as a transcriptional reporter for the inactivated gene (Fig. 4.1). The final cargo (integration) plasmids were pZR2035, pZR2036, pZR2051, pZR2052, pZR2055, pZR2057, pZR2062, pZR2063, pZR2065, pZR2066, pZR2073, pZR2074, pZR2085, pZR2086, pZR2133, and pZR2139 (for genes *alr0483*, *all1076, all1523, all1681, all3132, all3680, all4106, all5173, alr0618, alr0731, alr0765, alr2325, alr2428, alr2494, alr3125, and alr3829*, respectively) (Table 4.4).

Table 4.3 Primers used in this study

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primers</th>
<th>Oligonucleotide sequences (5'→3')</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR90</td>
<td>AAGTTCTTCTCCTTTGCTAGC</td>
<td>Primer based on <em>gfp</em> of integration vector pZR606 used for colony PCR verifying single crossover knockouts</td>
<td></td>
</tr>
<tr>
<td><em>alr0483</em></td>
<td>ZR1256A</td>
<td>agctATGGGACTCCAATTGTCAAATCT</td>
<td>Upstream sequence of ZR1256; pair ZR1256A with ZR90 to verify single crossover in SR2035 (980 bp)</td>
</tr>
<tr>
<td></td>
<td>ZR1256</td>
<td>acTATGCCTTTAGATGACACAGAATTTC</td>
<td>ZR1256,1257 to amplify internal fragment (IF) of <em>alr0483</em> (738 bp)</td>
</tr>
<tr>
<td></td>
<td>ZR1257</td>
<td>tccgGAGTTCTCTCGACAGCCATCCC</td>
<td>Pair ZR1442,ZR1256A to verify complete knockout in SR2035 (1374 bp in WT)</td>
</tr>
<tr>
<td></td>
<td>ZR1442</td>
<td>agatctTTAAACGCGAAGTTGCCCAGTT</td>
<td></td>
</tr>
<tr>
<td><em>all1076</em></td>
<td>ZR1258A</td>
<td>atgcATGATCTCAATTCGTATTTCAGCC</td>
<td>Upstream sequence ofZR1258; pair ZR1258A with ZR90 to verify single crossover in SR2036 (850 bp)</td>
</tr>
<tr>
<td></td>
<td>ZR1258</td>
<td>TCctAGTTTTGCGATGCTTTCTCGGC</td>
<td>ZR1258,1259 to amplify IF of <em>all1076</em> (723 bp)</td>
</tr>
<tr>
<td></td>
<td>ZR1259</td>
<td>tccgGGGATGATCTGAATATAACTCG</td>
<td>Pair ZR1443,ZR1258A to verify complete knockout in SR2036 (997 bp in WT)</td>
</tr>
<tr>
<td></td>
<td>ZR1443</td>
<td>tgtgatcTTACCCATACTCTGTCGT</td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Sequence Details</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>all1523</td>
<td>ZR1260A</td>
<td>Upstream sequence of ZR1260; pair ZR1260A with ZR90 to verify single crossover in SR2051 (873 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1260</td>
<td>ACTAGTTTACTGATTACTTCTATATTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1261</td>
<td>tccGGGACTAATGGCATAAGCCTTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1444</td>
<td>tgtatctCATTITCCAACAAGCGCAATTA</td>
<td></td>
</tr>
<tr>
<td>all1687</td>
<td>ZR1262A</td>
<td>Upstream of ZR1262; pair ZR1262A with ZR90 to verify single crossover in SR2052 (1003 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1262</td>
<td>actAGTTTCAAGAGGTTATTATTTCTCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1263</td>
<td>tcccgGTTCTAATTGGGCAATGGAATAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1445</td>
<td>tgtatctCAGTCTTACCATACCTATT</td>
<td></td>
</tr>
<tr>
<td>all3132</td>
<td>ZR1264A</td>
<td>Upstream sequence of ZR1264; pair ZR1264A with ZR90 to verify single crossover in SR2055 (1020 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1264</td>
<td>ActAGTTTCAAGAGGTTATTATTTCTCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1265</td>
<td>tcccgGTTCTAATTGGGCAATGGAATAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1446</td>
<td>tgtatctCAGTCTTACCATACCTATT</td>
<td></td>
</tr>
<tr>
<td>all3680</td>
<td>ZR1266A</td>
<td>Upstream sequence of ZR1266; pair ZR1266A with ZR90 to verify single crossover in SR2057 (820 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1266</td>
<td>ActAGTTTCAAGAGGTTATTATTTCTCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1267</td>
<td>tcccgGTTCTAATTGGGCAATGGAATAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1447</td>
<td>tgtatctCAGTCTTACCATACCTATT</td>
<td></td>
</tr>
<tr>
<td>all4106</td>
<td>ZR1268A</td>
<td>Upstream sequence of ZR1268; pair ZR1268A with ZR90 to verify single crossover in SR2062 (950 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1268</td>
<td>ActAGTTTCAAGAGGTTATTATTTCTCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1269</td>
<td>tcccgGTTCTAATTGGGCAATGGAATAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1448</td>
<td>tgtatctCAGTCTTACCATACCTATT</td>
<td></td>
</tr>
<tr>
<td>all5173</td>
<td>ZR1270A</td>
<td>Upstream sequence of ZR1270; pair ZR1270A with ZR90 to verify single crossover in SR2063 (1510 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1270</td>
<td>ActAGTTTCAAGAGGTTATTATTTCTCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1271</td>
<td>tcccgGTTCTAATTGGGCAATGGAATAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1449</td>
<td>tgtatctCAGTCTTACCATACCTATT</td>
<td></td>
</tr>
<tr>
<td>alr0618</td>
<td>ZR1272A</td>
<td>Upstream sequence of ZR1272; pair ZR1272A with ZR90 to verify single crossover in SR2065 (690 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1272</td>
<td>actagTTTGGGTTATAGAAACCTCAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1273</td>
<td>tcccgGTTCTAATTGGGCAATGGAATAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1450</td>
<td>tgtatctCAGTCTTACCATACCTATT</td>
<td></td>
</tr>
<tr>
<td>alr0731</td>
<td>ZR1274A</td>
<td>Upstream sequence of ZR1274; pair ZR1274A with ZR90 to verify single crossover in SR2066 (781 bp)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1274</td>
<td>tctAGAAACAATGGAACCTCAGTCAACC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1275</td>
<td>tcccgGTTCTAATTGGGCAATGGAATAT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1451</td>
<td>tgtatctCAGTCTTACCATACCTATT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1768</td>
<td>tgatctGACACAACTTATGCAACCAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZR1769</td>
<td>tggatcggGTCAAGTCAATTCAACACCC</td>
<td></td>
</tr>
</tbody>
</table>

*ZR1768* and *ZR1769* were used to amplify and verify the *Aanbenaen* paralogs *orf-7120* to *orf-7121* from *Anabaena* 7120 to complement the SR2066 mutant.
alr0765

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR1276A</td>
<td>TcTAGAACCACGTCAAGGCAGAAC</td>
<td>Upstream sequence of ZR1276; pair ZR1276A with ZR90 to verify single crossover in SR2073 (730 bp)</td>
</tr>
<tr>
<td>ZR1276</td>
<td>tCctAGGGTAATTATGTAAGACATCAG</td>
<td>ZR1276,1277 to amplify IF of alr0765 (615 bp)</td>
</tr>
<tr>
<td>ZR1277</td>
<td>tcccgGGTTGCTATTCGCTGCAATTTC</td>
<td>Pair ZR1452,ZR1276A to verify complete knockout in SR2073 (766 bp in WT)</td>
</tr>
<tr>
<td>ZR1452</td>
<td>ttgatcCTACAATTGCAAAGGGCAGCT</td>
<td></td>
</tr>
</tbody>
</table>

alr2325

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR1278A</td>
<td>actAGTTAAAAACACTGACGGCAAATAA</td>
<td>Upstream sequence of ZR1278; pair ZR1278A with ZR90 to verify single crossover in SR2074 (1160 bp)</td>
</tr>
<tr>
<td>ZR1278</td>
<td>actAGTAGATATTCATGCAACCTGAGTTT</td>
<td>ZR1278,1279 to amplify IF of alr2325 (544 bp)</td>
</tr>
<tr>
<td>ZR1279</td>
<td>tccgGGATATAAAAAATTTCCCTACC</td>
<td>Pair ZR1453,ZR1278A to verify complete knockout in SR2074 (1191 bp in WT)</td>
</tr>
<tr>
<td>ZR1453</td>
<td>ttgatcTTAGACTTTCCCGCATATTC</td>
<td></td>
</tr>
</tbody>
</table>

alr2428

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR1280A</td>
<td>atgcatATGAAAATTTTAATGTGAGGAGTAG</td>
<td>Upstream sequence of ZR1280; pair ZR1280A with ZR90 to verify single crossover in SR2085 (1510 bp)</td>
</tr>
<tr>
<td>ZR1280</td>
<td>tctcAGGGGATTGTTAGATGCAGGCC</td>
<td>ZR1280,1281 to amplify IF of alr2428 (1195 bp)</td>
</tr>
<tr>
<td>ZR1281</td>
<td>tccggGGTTTGGCTAACAAAGTCGAGCT</td>
<td>Pair ZR1454,ZR1280A to verify complete knockout in SR2085 (4896 bp in WT)</td>
</tr>
<tr>
<td>ZR1454</td>
<td>ttgatcCTAATCCCTAACATTAGTCGGT</td>
<td></td>
</tr>
</tbody>
</table>

alr2494

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR1282A</td>
<td>atgcatATGACGATGCAAGTTTCTCTTAG</td>
<td>Upstream sequence of ZR1282; pair ZR1282A with ZR90 to verify single crossover in SR2086 (1160 bp)</td>
</tr>
<tr>
<td>ZR1282</td>
<td>ActAGTGCGGCGAATTGGGTGCGTC</td>
<td>ZR1282,1283 to amplify IF of alr2494 (951 bp)</td>
</tr>
<tr>
<td>ZR1283</td>
<td>tccggGGCTGAGCGCGATCGCCTACA</td>
<td>Pair ZR1455,ZR1282A to verify complete knockout in SR2086 (1375 bp in WT)</td>
</tr>
<tr>
<td>ZR1455</td>
<td>ttgatcTTAGTCATTAGTAGATGATTA</td>
<td></td>
</tr>
</tbody>
</table>

alr3125

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR1284A</td>
<td>atgCATTTGCAATAGCCCTCCTTTCCG</td>
<td>Upstream sequence of ZR1284; pair ZR1284A with ZR90 to verify single crossover in SR2133 (1182 bp)</td>
</tr>
<tr>
<td>ZR1284</td>
<td>tctcAGGAAGCAAAAAATGTAGTCAATCTAG</td>
<td>ZR1284,1285 to amplify IF of alr3125 (613 bp)</td>
</tr>
<tr>
<td>ZR1285</td>
<td>tccgGGAAAGCTATTTTGCTCTGCTCG</td>
<td>Pair ZR1456,ZR1284A to verify complete knockout in SR2133 (1251 bp in WT)</td>
</tr>
<tr>
<td>ZR1456</td>
<td>ttgatcTTAGTCATTAGTAGATGATTA</td>
<td></td>
</tr>
</tbody>
</table>

alr3829

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZR1286A</td>
<td>atgcatATGTCAATAGGGTATATTTACCAG</td>
<td>Upstream sequence of ZR1286; pair ZR1286A with ZR90 to verify single crossover in SR2139 (1023 bp)</td>
</tr>
<tr>
<td>ZR1286</td>
<td>ACTAGTTTGATGAAGATTATTTTGATAAC</td>
<td>ZR1286,1287 to amplify IF of alr3829 (815 bp)</td>
</tr>
<tr>
<td>ZR1287</td>
<td>tCCGGGCTGAGGGCGATCGCAATATAG</td>
<td>Pair ZR1457,ZR1286A to verify complete knockout in SR2139 (1291 bp in WT)</td>
</tr>
<tr>
<td>ZR1457</td>
<td>ttgatcTTACTTCTCATACTGGTGTTCA</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4 Plasmids and bacterial strains created or used in this study

<table>
<thead>
<tr>
<th>Plasmid or strain</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRL271</td>
<td>(Cai and Wolk 1990)</td>
</tr>
<tr>
<td>pRL443</td>
<td>(Elhai et al. 1997)</td>
</tr>
<tr>
<td>pRL623</td>
<td>(Elhai et al. 1997)</td>
</tr>
<tr>
<td>pZR606</td>
<td>(Chen et al. 2015)</td>
</tr>
<tr>
<td>pZR2222</td>
<td>This study</td>
</tr>
<tr>
<td>pZR2223</td>
<td>This study</td>
</tr>
</tbody>
</table>
pZR2013  Km\(/	opm{Ag}\); *alr0483* internal fragment (IF) amplified by PCR with primers *Z*R1256,1257 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2035  Km\(/	opm{Sp}\); SpeI/XmaI digested *alr0483* IF from pZR2013 ligated to SpeI/XmaI cut pZR606

This study

pZR2014  Km\(/	opm{Ag}\); *all1076* IF amplified by PCR with primers *Z*R1258,1259 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2036  Km\(/	opm{Sp}\); AvrII/XmaI digested *all1076* IF from pZR2014 ligated to SpeI/XmaI cut pZR606

This study

pZR2015  Km\(/	opm{Ag}\); *all1523* IF amplified by PCR with primers *Z*R1260,1261 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2051  Km\(/	opm{Sp}\); SpeI/XmaI digested *all1523* IF from pZR2015 ligated to SpeI/XmaI cut pZR606

This study

pZR2016  Km\(/	opm{Ag}\); *all1681* IF amplified by PCR with primers *Z*R1262,1263 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2052  Km\(/	opm{Sp}\); SpeI/XmaI digested *all1681* IF from pZR2016 ligated to SpeI/XmaI cut pZR606

This study

pZR2017  Km\(/	opm{Ag}\); *alr3132* IF amplified by PCR with primers *Z*R1264,1265 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2055  Km\(/	opm{Sp}\); SpeI/XmaI digested *alr3132* IF from pZR2017 ligated to SpeI/XmaI cut pZR606

This study

pZR2018  Km\(/	opm{Ag}\); *all3680* IF amplified by PCR with primers *Z*R1266,1267 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2057  Km\(/	opm{Sp}\); SpeI/XmaI digested *all3680* IF from pZR2018 ligated to SpeI/XmaI cut pZR606

This study

pZR2019  Km\(/	opm{Ag}\); *all4106* IF amplified by PCR with primers *Z*R1268,1269 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2062  Km\(/	opm{Sp}\); SpeI/XmaI digested *all4106* IF from pZR2019 ligated to SpeI/XmaI cut pZR606

This study

pZR2020  Km\(/	opm{Ag}\); *alr5173* IF amplified by PCR with primers *Z*R1270,1271 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2063  Km\(/	opm{Sp}\); SpeI/XmaI digested *alr5173* IF from pZR2020 ligated to SpeI/XmaI cut pZR606

This study

pZR2021  Km\(/	opm{Ag}\); *alr0618* IF amplified by PCR with primers *Z*R1272,1273 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2065  Km\(/	opm{Sp}\); SpeI/XmaI digested *alr0618* IF from pZR2021 ligated to SpeI/XmaI cut pZR606

This study

pZR2022  Km\(/	opm{Ag}\); *alr0731* IF amplified by PCR with primers *Z*R1274,1275 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2066  Km\(/	opm{Sp}\); Bsal/XmaI digested *alr0731* IF from pZR2022 ligated to SpeI/XmaI cut pZR606

This study

pZR2023  Km\(/	opm{Ag}\); *alr0765* IF amplified by PCR with primers *Z*R1276,1277 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2073  Km\(/	opm{Sp}\); AvrII/XmaI digested *alr0765* IF from pZR2023 ligated to SpeI/XmaI cut pZR606

This study

pZR2024  Km\(/	opm{Ag}\); *alr2325* IF amplified by PCR with primers *Z*R1278,1279 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2074  Km\(/	opm{Sp}\); SpeI/XmaI digested *alr2325* IF from pZR2024 ligated to SpeI/XmaI cut pZR606

This study

pZR2025  Km\(/	opm{Ag}\); *alr2428* IF amplified by PCR with primers *Z*R1280,1281 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2085  Km\(/	opm{Sp}\); AvrII/XmaI digested *alr2428* IF from pZR2025 ligated to SpeI/XmaI cut pZR606

This study

pZR2026  Km\(/	opm{Ag}\); *alr2494* IF amplified by PCR with primers *Z*R1282,1283 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2086  Km\(/	opm{Sp}\); SpeI/XmaI digested *alr2494* IF from pZR2026 ligated to SpeI/XmaI cut pZR606

This study

pZR2027  Km\(/	opm{Ag}\); *alr3125* IF amplified by PCR with primers *Z*R1284,1285 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector

This study

pZR2133  Km\(/	opm{Sp}\); AvrII/XmaI digested *alr3125* IF from pZR2027 ligated to TOPO vector

This study
Temperature for 30 minutes for mating. Simultaneously, 10 ml of

**4.3.3 Knocking out heterocyst-specific genes via single crossover approach**

The final integrative cargo plasmids containing the IFs of all 16 heterocyst-specific protein-coding genes were transferred to *Anabaena* 7120 via conjugative transformation. For conjugation, tri-parental mating was initiated by mixing HB101 [pRL623+443] (Elhai et al. 1997) with *E. coli* NEB10β containing the cargo plasmid. The two *E. coli* strains were combined in a 1.5 mL tube and incubated at room temperature for 30 minutes for mating. Simultaneously, 10 ml of *Anabaena* 7120 (OD700

---

**Spel/XmaI cut pZR606**

- pZR2028: Km\(^r\)/Ap\(^r\); alr3829 IF amplified by PCR with primers ZR1286,1287 from *Anabaena* 7120 chromosomal DNA and cloned into pCR2.1-TOPO vector
- pZR2139: Km\(^r\)/Sp\(^r\); Spel/XmaI digested alr3829 IF from pZR2028 ligated to Spel/XmaI cut pZR606
- pZR2315: Km\(^r\)/Ap\(^r\); AarI/SalI-P-arl0731-orf-TAG-XmaI/BamHI PCR amplified by ZR1768,1769 from *Anabaena*7120 ligated to pCR2.1-TOPO
- pZR2316: Cm\(^r\)/Em\(^r\); SalI-arl0731-orf-TAG-XmaI from pZR2315 ligated to SalI/XmaI cut pZR2223

**Bacterial Strains**

<table>
<thead>
<tr>
<th>Strain Code</th>
<th>Strain Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOP10</td>
<td><em>E. coli</em> cloning host</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>NEB10β</td>
<td><em>E. coli</em> cloning host</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>WT7120</td>
<td><em>Anabaena</em> sp. PCC 7120 wild-type strain</td>
<td>This study</td>
</tr>
<tr>
<td>SR2035</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr0483 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2036</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 all1076 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2051</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 all1523 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2052</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 all1681 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2055</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr3132 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2057</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 all3680 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2062</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 all4106 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2063</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 all5173 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2065</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr0618 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2066-1</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr0731 complete single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2066-4</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr0731 intermediate single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2073</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr0765 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2074</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr2325 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2085</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr2428 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2086</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr2494 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2133</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr3125 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2139</td>
<td>Nm(^r)/Sp(^r); <em>Anabaena</em> 7120 alr3829 single crossover knockout mutant</td>
<td>This study</td>
</tr>
<tr>
<td>SR2066</td>
<td>Nm(^r)/Sp(^r)/Em(^r); SR2066 containing pZR2316 for complementation study</td>
<td>This study</td>
</tr>
<tr>
<td>(pZR2316)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{\text{a}}\)Ap\(^r\), ampicillin resistance; Sp\(^r\), spectinomycin resistance; Nm\(^r\)/Km\(^r\), neomycin-kanamycin resistance; Cm\(^r\)/Em\(^r\), chloramphenicol-erythromycin resistance; IF, internal fragment
0.5) grown in Bg11 was harvested (4000 xg, 10 min), and the pellet was washed with 1 mL Bg11. Following a second centrifugation (12,000 xg, 1 min), the pellet was resuspended in 100 µL Bg11. Then, the *Anabaena* 7120 resuspension was added to the *E. coli* mixture following mating. *Anabaena* 7120 and the mated *E. coli* were allowed to mate for 1 h, following which, the solution was plated on a nitrocellulose membrane on Bg11 agar supplemented with 5% LB. The conjugation plate was incubated 30°C under light for 2 d. Next, the membrane was transferred to a Bg11 plate containing Nm<sub>100</sub>Sp<sub>10</sub> to select for transformed *Anabaena* 7120. This conjugative transformation method was used to transfer the integrative cargo plasmids for all 16 genes to *Anabaena* 7120. The conjugation plates were incubated at 30°C under light (ca. 50-150 μE m<sup>-2</sup> s<sup>-1</sup>) until single colonies formed. On a weekly basis, membranes were transferred to new Bg11 antibiotic plates.

Plasmids successfully transferred to *Anabaena* 7120 via conjugation were expected to undergo a single crossover event with the wildtype chromosome (Fig. 4.1). Through a single crossover, the IF of the target gene contained in the integration plasmid homologously recombines with the matching portion of the gene in the *Anabaena* chromosome. Following the recombination event, the entire plasmid is integrated into the chromosome and disrupts the target gene in the wildtype (Fig. 4.1), thereby rendering it non-functional (Chen et al. 2015).
Fig. 4.1 Schematic representation of single crossover recombination to knockout target gene (Chen et al. 2015). A Representation of integrative cargo plasmid pZR606, including its multiple cloning site (MCS). B Depiction of the single crossover event. The internal fragment of the target gene in the cargo plasmid homologously recombines with the matching region of the gene in the *Anabaena* 7120 chromosome. After recombination, the cargo plasmid sequence is integrated into the *Anabaena* chromosome and disrupts (knocks out) the target gene, leaving two truncated versions of the target gene (the 5’ and the 3’ portions of the target gene). The interrupted gene is unable to encode a functional gene product. Primer pair FP (forward primer) and ZR90 was used to verify single crossover mutants. Primer pair FP and RP (reverse primer) was used to verify the wildtype gene.

Single crossovers from the conjugative transformations were verified by colony PCR with individual colonies that appeared on the conjugation selection plates. To verify single crossovers, a specific primer (Fig. 4.1) matching the chromosomal sequence upstream of the target gene’s internal fragment was used in conjunction with primer ZR90, a primer matching the *gfp* gene downstream of the MCS in the integration plasmid (Table 4.3). Single crossover knockout mutants were verified for all of the 16 genes. The
knockout mutants were named SR2035, SR2036, SR2051, SR2052, SR2055, SR2057, SR2062, SR2063, SR2065, SR2066, SR2073, SR2074, SR2085, SR2086, SR2133, and SR2139 (for genes alr0483, all1076, all1523, all1681, all3132, all3680, all4106, all5173, alr0618, alr0731, alr0765, alr2325, alr2428, alr2494, alr3125, and alr3829, respectively) (Table 4.4).

4.3.4 Screening single-crossover mutants for completely segregated knockout

*Anabaena* 7120 is oligoploid, with approximately 8 chromosome copies per cell (Hu et al. 2007). Therefore, to obtain a true knockout mutant, a single crossover had to occur for each chromosomal copy of the gene. Thus, in addition to verifying that a single crossover had occurred in the mutants, additional confirmation was needed to determine if the knockout mutant was completely segregated, or if the gene was inactivated in all 8 chromosomes. Colony PCR was performed using primers specifically matching the 5' and 3' ends of the intact gene located in the chromosome (see Table 4.3 for primers), not matching the internal fragment. With these primers, a product was only amplified when the intact gene was present. Genes disrupted by the single crossover of a cargo plasmid were unable to be amplified as the region between the primers (disrupted by the entire length of the cargo plasmid) was too large to be replicated during the specific PCR reaction. Therefore, completely segregated knockouts were identified by the absence of amplified DNA, whereas only the incompletely segregated knockouts and wildtype had DNA amplification.

Initially, all of the mutants verified as single crossovers were incompletely segregated knockouts (i.e., heterozygous mutants). To obtain completely segregated knockouts, the mutants were grown in Bg11 Nm<sub>100</sub>Sp<sub>10</sub> for 7 d. After reaching an OD<sub>700</sub>
of 0.5, the cultures were sonicated with a Branson 150 sonicator for 90 s. Sonication was performed to break up the filaments into single cells, so that completely segregated mutants could be more easily isolated due to its ploidy. Following sonication, $10^{-2}$, $10^{-4}$, and $10^{-6}$ dilutions were made of the culture, and 100 µL was plated on Bg11 Nm100Sp10 agar. Plates were incubated at 30°C under light until colonies appeared on the plates. Then, PCR was performed with the colonies to check if the intact gene was still present, or if it was knocked out, thereby indicating a completely segregated knockout. This process of growth, sonication, plating, colony selection, and PCR screening was repeated until the completely segregated mutants were obtained.

4.3.5 Growth on fixed nitrogen-free medium (Bg110) to screen for Fox genes

After completely segregated mutants were verified for the single crossover mutants, the mutants were subjected to nitrogen stepdown and plated on Bg110 agar to determine if they were able to grow in the absence of fixed nitrogen. For nitrogen stepdown, 3 independent mutants for each gene were grown in Bg11 for 7 d until reaching an OD$_{700}$ of approximately 0.5. Then, the cultures were harvested by centrifugation and washed three times with Bg110. After the final wash, the cultures were resuspended in 1 ml B110 at a final concentration of 0.3 at OD$_{700}$. Then, each mutant was spot plated by aliquoting 3 µL of each culture onto a Bg110 agar plate; wildtype Anabaena 7120 was also spotted on the same plate as a control. In addition, the mutant colonies and wildtype were plated on Bg11 agar plates as a positive control to compare the growth of the mutants in nitrogen-fixing (Bg110) vs. non-nitrogen-fixing (Bg11) conditions. The plates were incubated at 30°C under light for 2 weeks while the growth of the mutants and wildtype was monitored.
4.3.6 Complementation of SR2066

Since mutant SR2066-1 was unable to grow in the absence of fixed nitrogen, I constructed a complement plasmid to verify that the loss of function was due to the disruption of *alr0731* in the mutant. To construct the complement plasmid (pZR2316), the promoter region and ORF of *alr0731* were cloned into the replicative plasmid pZR2223 (Table 4.3). Then, pZR2316 was transferred to SR2066-1 via conjugal transformation, according to the method described above. SR2066-1 containing the complement plasmid was selected for on Bg11 Nm<sub>100</sub>Sp<sub>10</sub>Em<sub>10</sub> agar plates. Independent colonies were verified to contain the plasmid via colony PCR.

Following complementation of SR2066-1, nitrogen stepdown was performed with the wildtype, SR2066-1 mutant, and SR2066-1(pZR2316) complement. Following stepdown, cultures were both spotted onto Bg11<sub>0</sub> agar plates and inoculated into Bg11<sub>0</sub> medium (both with antibiotic for the mutant and complement) as described previously. The Bg11<sub>0</sub> plates were incubated at 30°C under light for 2.5 weeks while growth of the cultures was observed. The flasks were also incubated for 2.5 weeks (120 rpm, 30°C, under light) and bright field microscope pictures were taken of the liquid cultures at 48 h, 8 d, and 16 d following nitrogen stepdown.

Complementation for the second Fox mutant, SR2055-1, is still ongoing.

4.3.7 Nitrogenase activity and GC-MS analysis of acetylene reduction

One of the *alr0731* knockout mutants (SR2066-4) was an incompletely segregated, intermediate knockout since it contained the single crossover in some chromosome copies and the intact gene in others. Since SR2066-4 showed an intermediate growth phenotype compared the completely segregated mutant (SR2066-1)
and wildtype *Anabaena* 7120, I performed a nitrogenase activity assay to understand the impact of a heterozygous mutant or a partial knockout (similar to a gene knockdown) on oxic nitrogen fixation. To measure nitrogenase activity, an acetylene reduction assay (ARA) was performed. For the assay, *Anabaena* 7120, SR2066-4, and SR2066-1(pZR2316) stably grown in Bg11₀ medium (with antibiotic for mutant and complement) were refreshed and grown in Bg11₀ (plus antibiotic for mutant and complement) for 7 d until reaching an OD₇₀₀ of 0.5. After incubation, the cultures were harvested and concentrated to 0.8 OD₇₀₀ in Bg11₀ medium ready for ARA. For each culture, 5 replicates of 20-mL sealed vials were prepared, each containing 4 ml culture. The sealed vials were injected with 0.5 ml acetylene gas and incubated at 120 rpm, 30°C under light for 1.5 h. After incubation, 5 ml headspace was removed from each vial and directly injected into the GC-MS to measure acetylene reduction to ethylene.

The GC-MS parameters for the acetylene reduction are as follows: 5 mL of headspace gas sample from the 20-mL culture bottle was administered via a 1 mL GSV Loop to the GC-MS (Agilent 890A/5975C). The volatile compounds were separated by CP7348 column (Agilent PoraBOND Q 25 m × 250 μm × 3 μm) with Pulsed Split mode at 100:1 ratio at a flow rate of 0.8 mL/min, using hydrogen as a carrier gas. The GC program was initiated at 32°C, held for 4 min, and ramped at 110°C to reach 232°C. The scanning mass range of MSD was 10 to 50 m/z.

4.4 Results

4.4.1 Identification of Fox genes

I chose to do a screening of the 16 knockout mutants to determine which of the gene products, if any, were required for heterocyst development or function. By growing
the mutants on Bg11\textsubscript{0} agar, I found that two of the mutants (the \textit{alr\textsubscript{3132}} knockout SR2055 and the \textit{alr\textsubscript{0731}} knockout SR2066) were unable to grow while the wildtype was able to grow (Fig. 4.2 & Fig. 4.3). Moreover, the knockout mutants were able to grow as well as the wildtype on Bg11 agar (Fig. 4.2 & Fig. 4.3). Though in these plate experiments no antibiotic was used so that the wildtype could be tested on the same plates as the mutants, colony PCR was performed before and after the experiment to verify the presence of the single crossover in the mutants. Because SR2055 colony 1 (SR2055-1) and SR2066 colony 1 (SR2066-1) were unable to grow on Bg11\textsubscript{0}, these results indicated that the gene products of \textit{alr\textsubscript{3132}} and \textit{alr\textsubscript{0731}} are required for growth under nitrogen-fixing conditions.

Before performing nitrogen stepdown with the mutants and plating them on Bg11\textsubscript{0} agar, PCR was performed to verify that the mutants were free of the wildtype gene. Through PCR, if the gene was completely knocked out in the mutant, no DNA amplification occurred; however, as expected, amplification occurred in the wildtype where the intact gene was present (Fig. 4.2a & Fig. 4.3a). Interestingly, although no DNA was amplified for the SR2055 and SR2066 mutants, only one mutant isolate for each was unable to grow on Bg11\textsubscript{0}. The growth of SR2055-1 and SR2066-1 was completely abolished on fixed nitrogen-free medium. However, SR2055-2, SR2055-3 and SR2066-3, SR2066-4 were all able to grow on Bg11\textsubscript{0} agar. These results indicate that though PCR did not detect any wildtype gene in SR2055-2,-3 and SR2066-3,-4, a small amount of the intact gene may have still been present, thereby allowing these mutants to still grow on Bg11\textsubscript{0}. Therefore, I conclude that SR2055-1 and SR2066-1 were completely segregated.
mutants while SR2055-2,-3 and SR2066-3,-4 were incompletely segregated mutants or heterozygous mutants.

![A figure showing mutant screening and characterization.](image)

**Fig. 4.2** *all3132* mutant screening and characterization. Conjugation colonies verified to contain the single crossover were subjected to a second screening to detect if the *all3132* gene was completely knocked out. A PCR results of mutant screening. Lane 1 of the gel is the 1 kb ladder, lanes 2-8 are colonies 1-7, and lane 10 is wildtype *Anabaena* 7120. For the PCR to detect if the wildtype gene was still present in the mutant, primers ZR1446 and ZR1264A were used. If the wildtype gene was still present, an 1159 bp DNA segment would be amplified. In the knockout gene, no amplification occurs. The presence of the banding in the wildtype (lane 10) and the absence of banding in the mutants (lanes 2-8), indicate that the gene is successfully knocked out in the mutant colonies. B Characterization of mutant colonies 1, 2, and 3 (labeled C₁, C₂, and C₃) on fixed nitrogen-free medium (Bg11₀) compared to nitrate replete medium (Bg11); both plates were without antibiotic so the wildtype could be grown alongside the mutant. Wildtype *Anabaena* 7120 was used as a control. Cultures were spot plated (3 µL, 0.3 OD₇₀₀) and incubated at 30°C under light for 2 weeks while growth was monitored; picture was taken at 2 weeks.
4.4.2 Verifying requirement of \textit{alr0731} for diazotrophic growth via complementation experiment

Because SR2066-1 was unable to grow on fixed nitrogen-free medium, I needed to confirm that \textit{alr0731} is indeed responsible for the phenotype through a complementation experiment. The mutant SR2066-1 was conjugatively transformed with a replicative plasmid pZR2316 containing intact \textit{alr0731} along with its native promoter. Therefore, if the plasmid successfully complemented the mutant by re-endowing it with the ability to grow on fixed nitrogen-free medium, we can conclude that the gene \textit{alr0731} is responsible for growth in diazotrophic conditions. By performing nitrogen stepdown with the complement SR2066-1(pZR2316) alongside the knockout mutant SR2066-1 and plating both on Bg110 agar, I found that the complement was able to grow whereas the mutant died (Fig. 4.3c).
**Fig. 4.3** *alr0731* mutant screening, characterization, and complementation. Conjugation colonies verified to contain the single crossover were subjected to a second screening to detect if the *alr0731* gene was completely knocked out. A PCR results of mutant screening. Lane 1 of the gel is the 1 kb ladder, lanes 2-8 are colonies 1-7, and lane 10 is wildtype *Anabaena* 7120. For the PCR to detect if the wildtype gene was still present in the mutant, primers ZR1451 and ZR1274A were used. If the wildtype gene was still present, an 843 bp DNA segment would be amplified. In the knocked outs, no amplifications were detected. The presence of the 843 bp band in the wildtype (lane 10) and its absence in the mutants (lanes 2-8), suggesting that the gene is successfully knocked out in the mutant colonies. B Characterization of mutant colonies 1, 3, and 4 (labeled C1, C3, and C4) on fixed nitrogen-free medium (Bg110) compared to nitrate replete medium (Bg11); both plates were without antibiotic so the wildtype could be grown alongside the mutant. Wildtype *Anabaena* 7120 was used as a control. Cultures were spot plated (3 µL, 0.3 OD700) and incubated at 30°C under light for 2 weeks while growth was monitored; picture was taken at 2 weeks. C Complementation of *alr0731* knockout mutant SR2066-1. Since colony 1 (named SR2066-1) was unable to grow in the absence of fixed nitrogen, this mutant strain was complemented with plasmid pZR2316, containing the intact *alr0731* gene. Both mutant SR2066-1 and complement SR2066-1(pZR2316) were spot plated (3 µL, 0.3 OD700) on Bg110(Nm25Sp5) and Bg11 (Nm100Sp10) agar; neomycin and spectinomycin were used as both the mutant and complement shared this antibiotic resistance, thereby allowing them to be compared side-by-side. Plates were incubated at 30°C under light for 2 weeks while growth was monitored; picture was taken at 2 weeks. Results showed that the complement was able to grow on fixed nitrogen-free medium while the mutant was unable to grow, indicating successful complementation.

The complete and partial knockout mutants for SR2066 along with the complement strain SR2066-1(pZR2316) were used to study culture morphology following nitrogen stepdown. Microscope pictures of the cultures at 36 h, 8 d, and 16 d following stepdown revealed the cellular response to nitrogen stepdown. For SR2066-1 (the completely segregated mutant), growth immediately deteriorated following nitrogen stepdown (Fig.4.4, column 3, 36 h). At 8 d and 16 d, the culture contained unicellular cells and dead cells (Fig. 4.4, column 3). This was consistent with the death of SR2066-1 grown on Bg110 agar. For SR2066-4 (the incompletely segregated mutant), its growth
was stunted directly following nitrogen stepdown (Fig. 4.4, column 2, 36 h). However, the culture then began to recover and began growing in shorter and then longer filaments at 8 d and 16 d post-nitrogen stepdown, respectively (Fig. 4.4, column 2). By observing the complement phenotypically under the microscope following nitrogen stepdown, I observed that it grew in shorter filaments initially following stepdown; however, over time, it grew in longer filaments more similar to the wildtype (Fig. 4.4, column 4).

Together, these results indicate that *alr0731* is required for growth under nitrogen-fixing conditions.

**Fig. 4.4** Strain phenotypes following nitrogen stepdown. Bright field microscope images at 40x magnification were taken at 36 h, 8 d, and 16 d following nitrogen stepdown. Column 1 shows the wildtype’s response to nitrogen stepdown. Column 2 shows the response of partial *alr0731* knockout SR2066-4. Column 3 shows the fate of *alr0731* knockout SR2066-1 following nitrogen stepdown. Column 4 depicts the growth of complement SR2066-1(pZR2316) following stepdown.
4.4.3 Effect of alr0731 knockdown on nitrogenase activity

Although SR2066-1 was unable to grow under nitrogen-fixing conditions since it was a completely segregated knockout, the mutant SR2066-4 was able to grow although its growth was hindered due to the partial knockout of alr0731. Because SR2066-4 showed an intermediate phenotype of the alr0731 knockout, I performed a nitrogenase activity test to compare the effect knocking down alr0731 had on oxic nitrogen fixation. By comparing nitrogenase activity of SR2066-4, complement SR2066-1 (pZR2316), and Anabaena 7120, I observed that nitrogenase activity was highest in the wildtype, followed by the complement, while SR2066-4 had the lowest nitrogenase activity (Fig. 5). These results confirm again the importance of alr0731 in Anabaena 7120’s growth in nitrogen fixing conditions.

![Graph showing nitrogenase activity](image)

**Fig. 4.5** Nitrogenase activity in partial knockout (SR2066-4), complement [SR2066-1(pZR2316)], and wildtype. Nitrogenase activity was measured indirectly via the acetylene reduction assay. The amount of acetylene reduced to ethylene was indicative of the level of nitrogenase activity. The partial knockout, complement, and wildtype cultures were stably grown in Bg110 (with antibiotic for the partial knockout and complement). Then, the cultures were refreshed and grown in Bg110 (plus antibiotic for the partial knockout and complement) for 7 d before nitrogenase activity was measured. Nitrogenase activity is indicated by nmol ethylene produced per µg chlorophyll per h. Error bars represent standard error of mean.
4.5 Discussion

By knocking out 16 heterocyst-specific protein-coding genes, I found two gene mutants (all3132 mutant and alr0731 mutant) that were unable to grow under aerobic nitrogen-fixing conditions. These results suggested that these genes may be Fox genes. Since it was possible that the mutants’ inability to grow in diazotrophic conditions was due to some polar effect, I performed complementation with the alr0731 mutant SR2066-1. The mutant was complemented by reinserting the functional gene and its native promoter into the mutant via a cargo plasmid (pZR2316). The complement strain SR2066-1(pZR2316) regained its ability to grow in aerobic nitrogen-fixing conditions and the microscopy phenotype is restored to wildtype as well. Therefore, the results indicate that the functional gene product of alr0731 is required for growth under nitrogen-fixing conditions. Because of alr0731’s necessity to aerobic diazotrophic growth, I categorize this as a Fox gene.

Because Anabaena 7120 is a multicellular organism containing approximately 8 chromosome copies per cell, all (or nearly all) of the gene copies must be disrupted in a mutant to display a true knockout effect. Another alr0731 mutant, SR2066-4, was a heterozygous knockout mutant since some of the copies of alr0731 were successfully disrupted while others were not. I considered this heterozygous knockout mutant to represent a gene knockdown of alr0731. I compared the effect of the aerobic nitrogen-fixing conditions on the alr0731 knockout mutant (SR2066-1), heterozygous knockout mutant (SR2066-4), and complement [SR2066-1(pZR2316)]. By observing the cell morphological phenotype under the microscope following nitrogen stepdown, I observed that the knockout mutant SR2066-1 started forming bulging cells 36 h post-nitrogen
stepdown. I have consistently witnessed these types of bulging cells when a culture is unhealthy or dying. Consistent with this, at days 8 and 16 post-stepdown, the culture was almost entirely dead cells with a few unicellular living cells. Because Anabaena 7120 requires the symbiotic relationship between photosynthetic vegetative cells and nitrogen-fixing heterocysts to grow under diazotrophic conditions, the cells must grow in filaments in order to exchange fixed carbon and nitrogen. Therefore, even though single live cells were observed at 8 d and 16 d, these cells must eventually die as they are not part of a filament. These results confirmed the initial screening results of the mutant grown on fixed nitrogen-free medium (Bg110 agar), where the culture died completely in aerobic nitrogen-fixing conditions.

The heterozygous knockout mutant SR2066-4 showed growth on the Bg110 agar plate during the initial screening for Fox mutants. When I observed the culture under the microscope following nitrogen stepdown, I noticed that initially the culture formed bulging cells much like the knockout mutant SR2066-1. However, by 8 d following stepdown, the heterozygous knockout mutant SR2066-4 began to grow in short filaments, and at 16 d had even longer filament lengths, though still shorter than the wildtype. Finally, by observing the complement strain SR2066-1(pZR2316) following nitrogen stepdown, I saw it was able to grow in filaments at all time points witnessed post nitrogen stepdown. Though the complement’s filaments were shorter in general than those seen in the wildtype, the complement never had the bulging dying cells seen in the knockout SR2066-1 and partial knockout SR2066-4. Together, these results further supported my conclusion that alr0731 is required for aerobic nitrogen fixation.
Lastly, I compared the nitrogenase activity of the heterozygous knockout mutant SR2066-4, complement SR2066-1(pZR2316), and wildtype. Again, since SR2066-4 was a heterozygous knockout mutant, I was curious of the impact a knockdown of alr0731 would have on nitrogen fixation. The cultures were stably grown in Bg110 with antibiotics for the mutant and complement prior to the nitrogenase assay. By comparing the nitrogenase activity, the lowest activity occurred in heterozygous knockout mutant SR2066-4, and the complement and wildtype had similar activity levels. The nitrogenase activity in the heterozygous knockout SR2066-4 was likely due to the presence of the wildtype gene, giving the cells a partial-wildtype phenotype. These results further established the importance of alr0731 under aerobic nitrogen-fixing conditions.

The gene product of alr0731 is an anaerobic ribonucleoside triphosphate reductase activating protein. This protein, Alr0731, activates an anaerobically expressed ribonucleoside-triphosphate reductase under anaerobic conditions (Eliasson et al. 1992; Ollagnier et al. 1997; Sun et al. 1995). The ribonucleoside reductase activated by Alr0731 generates deoxyribonucleoside triphosphates (dNTPs) from ribonucleoside triphosphates (NTPs) (Eliasson et al. 1990). Deoxyribonucleoside triphosphates are primarily required for DNA replication. Though heterocysts themselves are non-dividing cells, studies have suggested that heterocyst differentiation is tied to cell division (Adams and Carr 1981; Sakr et al. 2006a; Sakr et al. 2006b). Therefore, DNA replication and repair may be important to heterocyst differentiation. Additionally, heterocyst development involves phage-mediated genomic rearrangements in which DNA elements interrupting nitrogen fixation genes are excised from the heterocyst-specific genome so the uninterrupted intact genes can be reconstituted and transcribed in heterocysts.
(Carrasco et al. 1995; Carrasco and Golden 1995; Golden et al. 1988; Golden et al. 1991; Matveyev et al. 1994; Qiu 2018). DNA repair may be required during this genomic rearrangement process as chromosomal damage is possible during the genomic rearrangements. Moreover, DNA repair mechanisms may require dNTPs to make the repair in such situations.

Since Alr0731 is an enzyme that activates an anaerobic ribonucleoside-triphosphate reductase, it makes sense that it would be expressed in heterocysts, which provide a mostly anaerobic environment for nitrogenase to function. If the function of Alr0731 is more important during heterocyst differentiation when DNA replication and rearrangements are occurring, alr0731 may have higher expression immediately following nitrogen stepdown and may be more prevalent in pro-heterocysts than in mature heterocysts. Since my knockout mutants for alr0731 contain a gfp reporter gene controlled by the alr0731 promoter, it would be interesting to see if alr0731 expression is localized to pro-heterocysts and if its expression is higher immediately following nitrogen stepdown when heterocysts are forming. These results may give further clues as to how Alr0731 plays a role during acclimation to and/or growth in aerobic nitrogen-fixing conditions.

In summary, through target gene inactivation via a single crossover approach, I was able to identify one gene alr0731 as a Fox gene. The protein that this gene encodes had been previously detected only in heterocyst cells of *A. cylindrica*. Interestingly, alr0731 encodes an anaerobic ribonucleoside triphosphate reductase activating protein. This protein aids in the pathway for conversion of NTPs to dNTPs, which are required for DNA synthesis. Events such as DNA replication and repair may be more important
during heterocyst differentiation than once the heterocyst is fully developed since
differentiation may be linked to cell division and requires genome editing. Future
research may uncover the potential role Alr0731 plays in response to aerobic nitrogen-
fixing conditions. Additionally, the integration plasmid used to disrupt the target genes in
this experiment contains a \textit{gfp} transcriptional fusion to the target gene. Therefore, future
studies can use the \textit{gfp} marker to understand the location and timing of the gene’s
transcription. This may provide further insight into the role \textit{alr0731} plays under aerobic
nitrogen-fixing conditions. Finally, since a second mutant, SR2055-1 (\textit{all3132} knockout),
was discovered that was unable to grow in aerobic diazotrophic conditions. Future work
will include the complementation of this mutant to decipher if this is a Fox gene as well.
CHAPTER 5: Summary and Conclusions

Cyanobacteria are a diverse group of photosynthetic prokaryotes. They are found in many diverse habitats, from freshwater to salt water, from temperate to tropical to Polar regions (Gaysina et al. 2019). Their cellular metabolisms must be equally encompassing to allow cyanobacteria to inhabit these varied environments.

Photosynthesis is the predominant metabolic system in cyanobacteria. Just as profound, however, is the ability of certain cyanobacteria to perform nitrogen fixation. Because nitrogenase, the nitrogen fixation complex, is inactivated by oxygen (Pienkos et al. 1983), diazotrophic cyanobacteria are posed with a unique dilemma: separating oxygen-evolving photosynthesis from oxygen-labile nitrogen fixation (Esteves-Ferreira et al. 2017).

Anabaena 7120 is a diazotrophic, filamentous cyanobacterium. This cyanobacterium uses spatial separation to perform photosynthesis and nitrogen fixation simultaneously. When fixed nitrogen becomes limiting in the environment, Anabaena 7120 differentiates specialized nitrogen fixation cells called heterocysts (Wolk 1982). The sole function of heterocysts is to perform nitrogen fixation. Heterocysts differentiate from vegetative cells, whose function is photosynthesis. The differentiation process requires many changes in gene expression as well as chromosomal modifications to create structurally and metabolically distinct heterocysts from progenitor vegetative cells (Ehira et al. 2003; Golden and Wiest 1988; Herrero et al. 2013; Qiu 2018). My primary objective in this study was to understand heptadecane (C_{17}H_{36}) production in Anabaena 7120 and how it impacts nitrogen fixation, and secondly to identify gene products required for oxic nitrogen fixation (Fox genes) in Anabaena 7120.
Cyanobacteria are universally able to produce long-chain hydrocarbons (Coates et al. 2014). *Anabaena* 7120 produces the hydrocarbon heptadecane via the AAR/ADO pathway. In this pathway, a C18 fatty acyl-ACP is converted to a fatty aldehyde by the enzyme fatty acyl-ACP reductase (AAR); then, aldehyde decarbonylase (ADO) removes a carbonyl group from the fatty aldehyde (Coates et al. 2014; Schirmer et al. 2010). In my first study (Ch. 2), I identified the genes required for heptadecane production in *Anabaena* 7120 through using a double crossover knockout approach. By knocking out the genes *alr5283*-*alr5284* (*ado* and *aar*, respectively), which were homologous to known alkane genes identified in *Synechococcus elongatus* PCC 7942 (Schirmer et al. 2010), I determined that heptadecane production was abolished in the mutant DR935. Through a complementation experiment, I reinserted the functional genes into DR935 and saw the re-emergence of heptadecane production. These results verified that *alr5283*-*alr5284* are required for heptadecane production in *Anabaena* 7120.

Similar to another study (Schirmer et al. 2010), the findings of my heptadecane study (Gibbons et al. 2018) indicated that heptadecane does not appear to be required for growth under normal conditions (30°C, 60 µmol E m⁻² s⁻¹ light intensity). Other studies have indicated that alkane production is correlated to certain stress conditions, such as high salinity (Kageyama et al. 2015), cold temperature (Berla et al. 2015), and nitrogen-fixing (Kageyama et al. 2015) conditions. Because *Anabaena* 7120 is a diazotrophic cyanobacterium, I chose to study the alkane knockout mutant DR935 under nitrogen-fixing conditions to elucidate how alkanes may play a role in helping the organism adjust to diazotrophy (Ch. 3).
My results revealed that heptadecane plays an important role in the days immediately following fixed-nitrogen starvation. From days 0-5 following fixed-nitrogen starvation, heptadecane production in the wildtype is significantly higher than in the wildtype grown in nitrate-replete medium. Additionally, mutant DR935, which is unable to produce heptadecane, exhibited a fragmented growth phenotype during this time period initially following fixed-nitrogen starvation. Coinciding with this, the mutant had negligible nitrogenase activity at 48 h, while nitrogenase activity occurred in the wildtype. Furthermore, through observing alkane gene expression using a *gfp* marker, I observed that *aar* expression was more prominent in heterocysts compared to vegetative cells at 48 h following fixed-nitrogen starvation. From these results, I propose that heptadecane helps maintain membrane integrity in *Anabaena* 7120, particularly immediately following fixed-nitrogen starvation. Interestingly, as DR935 acclimated to nitrogen-fixing conditions, it regained filamentous growth similar to the wildtype. This suggests that heptadecane’s function may be complemented by some other cellular compound or mechanism as the culture becomes adapted to diazotrophic conditions.

In my final study, I identified two genes required for oxic nitrogen fixation in *Anabaena* 7120 (Ch. 4). A previous study identified 57 proteins exclusively or primarily localized in heterocysts of *Anabaena cylindrica* (Qiu 2018). I identified homologs to 16 of the genes encoding these proteins in *Anabaena* 7120. Through a single crossover approach, I knocked out the genes in *Anabaena* 7120 and assessed the mutants’ ability to grow in fixed-nitrogen deplete medium (Bg11). I identified two mutants (for genes *all3132* and *alr0731*) that were unable to grow in nitrogen-fixing conditions. The *alr0731* mutant (SR2066-1) was complemented, and the complement [SR2066-1(pZR2316)]
regained its ability to grow in diazotrophic conditions. These results verified the requirement of Alr0731 for oxic nitrogen fixation (i.e., alr0731 is a Fox gene).

Interestingly, alr0731 encodes an anaerobic ribocucleoside triphosphate reductase activating protein, which is involved in the conversion of ribonucleoside triphosphates (NTPs) to deoxyribonucleoside triphosphates (dNTPs) (Eliasson et al. 1990). Cell division has been postulated to be a part of heterocyst differentiation (Adams and Carr 1981; Sakr et al. 2006a; Sakr et al. 2006b). Additionally, genomic rearrangements are known to occur during heterocyst development (Carrasco et al. 1995; Carrasco and Golden 1995; Golden et al. 1988; Golden et al. 1991; Matveyev et al. 1994; Qiu 2018). Both of these processes require dNTPs. Therefore, I believe that Alr0731 plays a critical role in heterocyst development more so than in heterocyst function because once heterocysts are formed, they are terminal cells that do not divide and thus have little obvious requirement for dNTPs. Future research will investigate the expression of alr0731 using a gfp transcriptional marker to determine when and where its expression may be upregulated. I hypothesize that alr0731 expression will be upregulated primarily following fixed nitrogen starvation and will be localized to pro-heterocysts or early developed heterocysts. Additionally, a complementation experiment will be performed with the all3132 mutant to determine if it is also a Fox mutant.

Much research has been done to understand the process of nitrogen fixation in Anabaena 7120. Future research will continue to add to our understanding of how different cellular pathways and metabolites, such as alkane production, may impact heterocyst development and function. Additionally, with 85 Fox genes already identified
in *Anabaena* 7120 (Ch. 4, Table 4.2), future research will likely add to this list as more genes required for oxic nitrogen fixation are discovered.
REFERENCES


Espinosa J, Brunner T, Fiedler N, Forchhammer K, Muro-Pastor AM, Maldener I (2010) DevT (Alr4674), resembling a Ser/Thr protein phosphatase, is essential for heterocyst function in the cyanobacterium Anabaena sp. PCC 7120. Microbiology 156(Pt 12):3544-3555

Espinosa J, Brunner T, Fiedler N, Forchhammer K, Muro-Pastor AM, Maldener I (2010) DevT (Alr4674), resembling a Ser/Thr protein phosphatase, is essential for heterocyst function in the cyanobacterium Anabaena sp. PCC 7120. Microbiology 156(Pt 12):3544-3555


Fell D (1997) Understanding the Control of Metabolism,


Mazur BJ, Chui CF (1982) Sequence of the gene coding for the $\beta$-subunit of dinitrogenase from the blue-green alga Anabaena. Proc Natl Acad Sci USA 79:6782-6786


Murata N, Nishida I (1987) Lipids of Blue-Green Algae (Cyanobacteria),

Murata N, Omata T (1988) Isolation of cyanobacterial plasma membranes, vol 167C,


Murry MA, Wolk CP (1989) Evidence that the barrier to the penetration of oxygen into heterocysts depends upon two layers of the cell envelope. Archives of microbiology 151(6):469-474 doi:10.1007/bf00454860


Sinha R, Häder D (1996) Photobiology and Ecophysiology of Rice Field Cyanobacteria, vol 64,


Stal L, Krumbein W (1987a) Temporal separation of nitrogen fixation and photosynthesis in the filamentous, non-heterocystous cyanobacterium Oscillatoria sp. 149(1):76-80


Toepel JR, McDermott JE, Summerfield TC, Sherman LA (2009) Transcriptional analysis of the unicellular, diazotrophic cyanobacterium Cyanothece sp. ATCC
51142 grown under short day/night cycles. Journal of phycology 45(3):610-20

candidacy for differentiation into heterocysts in the filamentous cyanobacterium
Anabaena sp. PCC 7120. Archives of microbiology 192(1):23-31
doi:10.1007/s00203-009-0525-4

Tsujimoto R, Kamiya N, Fujita Y (2014) Transcriptional regulators ChlR and CnfR are
essential for diazotrophic growth in nonheterocystous cyanobacteria. Proceedings
of the National Academy of Sciences of the United States of America
111(18):6762-7 doi:10.1073/pnas.1323570111

fixation genes recognized by CnfR in the nonheterocystous nitrogen-fixing
doi:10.1111/mmi.13402

Valladares A, Flores E, Herrero A (2008) Transcription activation by NtcA and 2-
oxoglutarate of three genes involved in heterocyst differentiation in the
cyanobacterium Anabaena sp. strain PCC 7120. Journal of bacteriology
190(18):6126-33 doi:10.1128/jb.00787-08

genes required for nitrogenase activity and diazotrophic growth in Anabaena sp.
PCC 7120. Molecular microbiology 47(5):1239-49

development and diazotrophic metabolism in terminal respiratory oxidase mutants
of the cyanobacterium Anabaena sp. strain PCC 7120. Journal of bacteriology

Vargas WA, Nishi CN, Giarrocco LE, Salerno GL (2011) Differential roles of
alkaline/neutral invertases in Nostoc sp. PCC 7120: Inv-B isoform is essential for

Vermaas W (1996) Molecular genetics of the cyanobacteriumSynechocystis sp. PCC
6803: Principles and possible biotechnology applications. Journal of Applied

Callahan SM (2014) Expanding the direct HetR regulon in Anabaena sp. strain

Videau P, Rivers OS, Higa KC, Callahan SM (2015) ABC Transporter Required for
Intercellular Transfer of Developmental Signals in a Heterocystous

Videau P, Rivers OS, Hurd K, Ushijima B, Oshiro RT, Ende RJ, O'hanlon SM, Cozy LM
(2016) The heterocyst regulatory protein HetP and its homologs modulate
heterocyst commitment in Anabaena sp. strain PCC 7120. Proc Natl Acad Sci U S
A doi:10.1073/pnas.1610533113

Wada H, Murata N (1990) Temperature-Induced Changes in the Fatty Acid Composition
of the Cyanobacterium, Synechocystis PCC6803. Plant physiology 92(4):1062-
1069


Anabaena sp. strain PCC 7120. Archives of microbiology 188(6):551-63
doi:10.1007/s00203-007-0276-z
doi:10.1146/annurev.pp.35.060184.001103


