Identification of two genes required for heptadecane production in a N₂-fixing cyanobacterium *Anabaena* sp. strain PCC 7120

Jaimie Gibbons¹, Liping Gu¹*, Huilan Zhu¹, William Gibbons¹ and Ruanbao Zhou¹,²*

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 (C₁₇H₃₆). 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 (CO₂ and N₂ gas) and mineralized water.

Keywords: Cyanobacteria, Nitrogen fixation, Alkane biosynthesis, Hydrocarbons, Biofuels

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 CO₂ 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 CO₂ 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 (Bergman et al. 1997; Berman-Frank et al. 2003; Fay 1992; Zehr 2011) and have minimal requirements for trace nutrients, meaning that their primary requirements are water, sunlight, and CO₂ (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, b; 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 \( (C_{17}H_{36}) \) and pentadecane \( (C_{15}H_{32}) \) are the most commonly observed alkanes produced by cyanobacteria and are hypothesized to be derived from octadecanoic acid \( (C_{18}H_{36}O_2) \) and hexadecanoic acid \( (C_{16}H_{32}O_2) \), 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 aldehydes. Thus, following the \( \text{C}_{n-1} \) rule, odd-numbered alkanes are the typical products.

After the initial discovery of the \( aar/ado \) 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 \( aar/ado \) 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 \( \alpha \)-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 \( \alpha \)-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. showed that \( ado \) (sll0208) and \( aar \) (sll0209) had different expression levels under all conditions (high light, \( \text{CO}_2 \) 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, \( \text{CO}_2 \) 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 \( S. \) elongatus increased alkane production approximately twofold (Kaiser et al. 2013). Overexpression of both native copies of \( aar/ado \) in \( Synechocystis \) sp. PCC 6803 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. PCC 6803 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 \( S. \) elongatus PCC 7942. 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, we identified the \( aar/ado \) genes in a heterocyst-forming cyanobacterium \( Anabaena \) sp. PCC 7120 through BLAST-P alignment with \( S. \) elongatus PCC 7942_orf1593 (\( aar \)) and orf1594 (\( ado \)), the \( aar/ado \) genes initially identified in Schirmer's study (Schirmer et al. 2010). We aimed to directly verify that these genes are required for hydrocarbon production in vivo. Our 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 \( Synechocystis \) PCC 6803 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 our study, we created constructs in which \textit{alr5283–84} were placed under a combination of native and the \textit{glnA} (\textit{alr2328}) promoters. Through varying the promoter systems in the complement study, we hoped to gain further insight into the control mechanisms behind the alkane genes’ transcriptional regulation. Our 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.

**Materials and methods**

**Bacterial strains and plasmids**

\textit{Escherichia coli} strains Top10 (Invitrogen) and NEB10β (New England Biolabs) were used for plasmid construction and maintenance. \textit{E. coli} strains were grown in Luria–Bertani broth. Antibiotic concentrations used for \textit{E. coli} strains were 100 µg mL\(^{-1}\) ampicillin (Ap), 50 µg mL\(^{-1}\) kanamycin (Km), 25 µg mL\(^{-1}\) chloramphenicol (Cm), 100 µg mL\(^{-1}\) spectinomycin (Sp), and 10 µg mL\(^{-1}\) erythromycin (Em). Antibiotic concentrations used for cyanobacterial mutant strains were 100 µg mL\(^{-1}\) neomycin (Nm), 10 µg mL\(^{-1}\) Sp, 10 µg mL\(^{-1}\) Em or except where noted.

**Construction of aar/ado knockout plasmid pZR935**

To disrupt the \textit{alr5283–5284} sequence in \textit{Anabaena} 7120 genome, plasmid pZR935 was created (details see Table 1) and transferred to \textit{Anabaena} 7120 via conjugal transformation to replace chromosomal \textit{alr5283–84} with the disrupted \textit{alr5283–84} sequence via double recombination. Briefly, \textit{alr5283–5284} and flanking sequences were PCR amplified with specific primers ZR241, ZR242 (Table 2) from \textit{Anabaena} 7120 genomic DNA, cloned into pCR\textsuperscript{2.1}-TOPO\textsuperscript{®} vector (TOPO TA Cloning\textsuperscript{®} kit, Invitrogen), creating pZR932. Next, site directed mutagenesis using primers ZR243, 244 introduced a NotI site into \textit{alr5283} within pZR932, creating pZR933. Then, the mutated \textit{alr5283–84} sequence was excised from pZR933 and transferred to the vector pZR824, creating pZR934. Finally, the 3’ end of \textit{alr5283} and the 5’ end of \textit{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 Additional file 1: Fig. S1).

**Knocking out 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 days until the culture reached an OD\textsubscript{700} of 0.5. Ten millilitre of the culture was harvested (4000×g, 10 min), and the pellet was washed with 1 mL BG11. Following a second centrifugation (12,000×g, 1 min), the pellet was resuspended in 100 µL BG11. This was added to the \textit{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 \textit{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 days. Next, the membrane was transferred to a BG11 plate containing the antibiotic to select for transformed \textit{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. More details were described in Target Gene Inactivation in Cyanobacterium \textit{Anabaena} sp. PCC 7120 (Chen et al. 2016).

Conjugal transformation of pZR935 to \textit{Anabaena} 7120 to achieve a double crossover was first verified by distinguishing single crossover colonies via colony PCR (primers listed in Table 2). Single crossover recombination colonies named SR935 were then grown in 10 mL BG11 plus spectinomycin for 1 week. Then, 1 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 \textit{Anabaena} 7120 over wildtype while sucrose selected against single crossovers due to pZR935 serving as a \textit{SacB}-based 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 \textit{alr5283–alr5284} knockout mutant named DR935.

**Complementation of Alr5283–Alr5284 knockout mutant**

To complement the knockout mutant (DR935), three plasmids (pZR2239, pZR2248, and pZR2243) were constructed containing \textit{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 \textit{glnA} (\textit{alr2328}) promoter, though each gene contained its own ribosomal binding site. Plasmid pZR2243 contained \textit{alr5283} under control of the constitutive
glutamine synthetase (glnA) promoter while *alr5284* remained under control of its native promoter. Table 1 provides further details of the plasmids and bacterial strains used in this study. QS-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.
**Table 2** Primers used in this study

| Primers       | Oligonucleotide sequences (5′ → 3′)                                                                 | Description                                                                                                                                 |
|---------------|--------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| ZR165         | GATCTCCGGCGCTACCGCGCGACATTGAGCTCGTCAAG                                               | Annealed oligonucleotides ZR165/ZR166 ligated to BglII-Spel digested pRL278 to produce pZR824                                               |
| ZR166         | ctagTCGAGACGTCAATTGCGCGCGCTAGCCCGGGA                                                     |                                                                                                                                           |
| ZR241         | tggaTCCAAACTTCAGGAAATTGCTG                                                               | ZR241,242 primer pair amplifying alr5283–84 ORF (2.7 kb), in knockout mutant with GFP-Sp′ cassette, primers amplify 4.7 kb              |
| ZR242         | tccAGAATTGATTGGGAGATG                                                               |                                                                                                                                           |
| ZR243         | CAAAAGGGGCGCTGGAAGTTAATCGGAATCC                                                  | ZR243,244 primer pair for site-directed mutagenesis to introduce NdeI site at 997 bp of alr5283–84 region                               |
| ZR244         | TTACCTTCTAGCgGcccGCTTGTTC                                                      |                                                                                                                                           |
| ZR1584        | atcgcatacatgcagtagctggATCCCTTTACTAATTACATTAAATTG                                       | Primers ZR1584/ZR1585 PCR amplified 1921 bp Cmr/Emr cassette (Nsi/I/Ndel-Nhel/BamHI-Emr/Cmr′ cassette-EcoRV/BglII/Xhol/Xmal/Smal) from pRL271 ligated to pCR1.1-TOPO to produce pZR2222          |
| ZR1585        | tccCGGGAAGATTATCAGCTCGAGATC                                                      | (Nsi/I/Ndel-Nhel/BamHI-Emr/Cmr′ cassette-EcoRV/BglII/Xhol/Xmal/Smal) from pRL271 ligated to pCR1.1-TOPO to produce pZR2222          |
| ZR261         | CAAAATTGGACACACTCGAGT                                                      | ZR261, 1602 primer pair for verification insertion of P-alr5283-P-alr5284 in pZR2223 shuttle vector (construction pZR2239)               |
| ZR1602        | tgacgcGACTCCAAAAACGTCAACGATTTCC                                                  | ZR1602, 1603 primer pair amplifying P-alr5283-P-84 (2 kb); amplification retains promoters for both genes                                   |
| ZR1603        | tccgggactcTAACCCAGCAGTGTCAAAACC                                               |                                                                                                                                           |
| ZR1606        | atcgcATGCCAGACTGCTAGCGGCGATTAG                                                | ZR1606,1607 primer pair amplifying alr5283–84 (2 kb); amplification includes promoter for alr5284 only                                |
| ZR1607        | tggcacaCTCAAACCCAGCATGTTCAAAACC                                               |                                                                                                                                           |
| ZR1608        | CATGgtatctctctcttTAAAGCTGCTGTAAGCTCGTAGGCT                                          | ZR1606-1608-1609-1607 overlap PCR (1.7 kb)                                                                                               |
| ZR1609        | GACTCTACGAGCTTTAAaagaaggagatataCCATGTGGTTGTCAATGAGT                                | (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)            |
| LK2406        | TTAAGGCGCGCGGCAGTAGCTAGCGGAGGTACTAGTGC                                             | Annealed oligonucleotides LK2406/LK2407 digested pZR606 to produce pZR6666, Apal-Xmal-BglII-Xbal-Agel-Spel multiple cloning sites (MCS) |
| LK2407        | TTAAGACTAGTACCGGCTGATCATGTCCTCGGCCGCGCCG                                                |                                                                                                                                           |

Sp: spectinomycin resistance; ORF: open reading frame; GFP: green fluorescent protein; MCS: multiple cloning sites, RBS: ribosome-binding site, P-: promoter, PCR: polymerase chain reaction

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, or pZR2243 to DR935 were named as DR935(pZR2239), DR935(pZR2248), and DR935(pZR2243), respectively.

**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 days with an initial OD700 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 eppendorf 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 10 min and left to sit for 10 min. After 10 min, each sample was vortexed for 10 s 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-heptadecane (TCI) were analyzed by GC/MS for comparison purposes.

**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 × 250 µm × 0.25 µm), with H₂ 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 50–500 m/z full scan, which identified heptadecane production. Based on the spectra obtained in the full scan of heptadecane and the internal standard tetracosane, we created
an 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, and 240) and (57, 71, and 338) for heptadecane and tetracosane, respectively. Heptadecane had a retention time of 6.31 min; tetracosane had a retention time of 11.8 min.

**Chlorophyll content analysis**

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

**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.

**Results**

**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. 1a). The peak was identified as heptadecane (C\(_{17}\)H\(_{36}\)) by the mass spectral 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. 1c) and showed the same fragmentation pattern as the wildtype compound (comparing Fig. 1d to b). Thus, we concluded that the major hydrocarbon peak detected in Anabaena 7120 is heptadecane (C_{17}H_{36}).

**Bioinformatics analysis to identify hydrocarbon biosynthesis genes**

Hydrocarbon biosynthesis genes have been identified in the cyanobacterial species S. elongatus PCC 7942 (Schirmer et al. 2010). Hydrocarbon biosynthesis genes have not yet been verified in any N_{2}-fixing cyanobacteria such as Anabaena 7120. To identify the genes required for heptadecane production in Anabaena 7120, a BLAST-P search using the representative proteins from S. elongatus against the publically available Anabaena 7120 genome database 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. 2a, b, respectively).

**Construction of DR935**

The two putative genes (alr5283–alr5284) whose products may be responsible for long-chain hydrocarbon production were cloned from 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 (Fig. 3a and Additional file 1: Fig. S1).

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. 3a). DR935 was verified by colony PCR from the conjugation (Fig. 3b). 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. 3b lane 3). 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. 3b lane 2), while the 2.7 kb PCR product in wild-type Anabaena 7120 was under detectable. Thus, the DR935 was completely segregated, the pure double recombinant DR935 containing only the 4.7 kb amplification were obtained. Verified DR935 double mutant was used for further analysis. Three plasmids were constructed (Fig. 3c) to test complementation of DR935.

**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. 4b). Among the three complement plasmids constructed, only pZR2239, containing alr5283–84 under control of the native 3-promoter system, was able to rescue heptadecane production in the knockout mutant DR935 (Fig. 4c). Neither pZR2248 nor pZR2243 was able to recover heptadecane production in the knockout mutant DR935 (data not shown).

**Quantification of heptadecane production in wild-type and complemented strain**

We analyzed the amount of heptadecane produced from wildtype Anabaena 7120, DR935, and DR935(pZR2239) strains during the first 5 days of the experiment. Overall, heptadecane production was higher in the wildtype (WT 7120) compared to the complemented strain DR935(pZR2239) and undetectable in the DR935 (Fig. 4d).

Though heptadecane production differed amongst the strains, it did not appear to impact normal growth of the culture as indicated by OD measurements (Fig. 5) 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.
Discussion

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 our work, heptadecane was the predominant volatile alkane captured by Supelpack 2SV resin (Sigma-aldrich) and detected in headspace of wildtype Ana- baena 7120 culture. 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. Our findings are partially supported by previous reports, which detected a range of hydrocarbons produced by the *ado/aar* gene products from various cyanobacteria when the genes were expressed in *E. coli* (Schirmer et al. 2010) or in *Synechocystis* sp. PCC 6803 (Wang et al. 2013). Taken together, we conclude that the ADO/AAR enzymes encoded by *alr5283–84* are responsible for heptadecane production from a C\textsubscript{18} fatty acid substrate in *Anabaena* 7120. **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.
2016) had been successfully used for complementing many knock-out mutants in *Anabaena* 7120 (Chen et al. 2016; Fan et al. 2005; Xu et al. 2015). The *glnA* promoter is a constitutive expression promoter and is up-regulated by nitrogen starvation (Flaherty et al. 2011). The *glnA* promoter was successfully used for complementing *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 *ado/aar* 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). Our 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 sixfold 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, our 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.

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; Bhaduriya et al. 2008; Kageyama et al. 2015; Takatanii 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). Our 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.
Additional file

Additional file 1: Fig. S1. Schematic illustration of pZR935 construction for knocking out alr5283 and alr5284 in Anabaena sp. strain PCC 7120.

Authors’ contributions
JG carried out the experiments and drafted the paper; LG, RZ designed, supervised and provided technical expertise to the project; WG provided both scientific and funding support; HG, LG and RZ helped with some experiments. All authors have revised the manuscript. All authors read and approved the final manuscript.

Author details
1 Department of Biology and Microbiology, South Dakota State University, Brookings, SD 57007, USA. 2 BioSNTR, South Dakota State University, Brookings, SD 57007, USA.

Acknowledgements
Heartfelt thanks to Charles Halfmann, Kangming Chen, and the other members of Dr. Zhou’s lab for their help in this research.

Competing interests
The authors declare that they have no competing of interests.

Availability of data and materials
Please contact author for data and material requests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Funding
Seed Funding (to J.G.) for this project was provided by the Joseph Nelson Scholarship program from South Dakota State University. This work was partially supported by the National Science Foundation Grant CBET1133951 (to R.Z.), NASA Grant No. NNX11AM03A (to R.Z. and W.G.), and by the South Dakota Agricultural Experiment Station and the CyanoSun Energy through the USDA-SBIR Grant 2012-33610-19524.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 August 2018 Accepted: 9 October 2018
Published online: 13 October 2018

References
Bergman B, Gallon J, Rai A, Stal L (1997) Nj fixation by non-heterocystous cyanobacteria. FEBS Microbiol Rev 19:139–185
Berla BM, Saha R, Maranas CD, Pakrasi HB (2015) Cyanobacterial alkanes modulate photosynthetic cyclic electron flow to assist growth under cold stress. Sci Rep 5:14894. https://doi.org/10.1038/srep14894
Berman-Frank I, Lundgren P, Falkowski P (2003) Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. Res Microbiol 154:157–164
Bhadouria P, Gupta R, Singh S, Bisen PS (2008) n-Alkanes variability in the diazotrophic cyanobacterium Anabaena cylindrica in response to NaCl stress. World J Microbiol Biotechnol 24:139–141. https://doi.org/10.1007/s11274-007-9439-y
Cai YP, Wolk CP (1990) Use of a conditionally lethal gene in Anabaena sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. J Bacteriol 172:3138–3145
Chen K, Gu L, Xiang X, Lynch M, Zhou R (2012) Identification and characterization of five intramembrane metalloproteases in Anabaena variabilis. J Bacteriol 194:6105–6115. https://doi.org/10.1128/JB.01366-12
Chen K, Xu K, Gu L, Hildreth M, Zhou R (2015) Simultaneous gene inactivation and promoter reporting in cyanobacteria. Appl Microbiol Biotechnol 99:1779–1793. https://doi.org/10.1007/s00253-014-6209-2
Chen K, Zhu H, Gu L, Tian S, Zhou R (2016) Target gene inactivation in cyanobacterium Anabaena sp. PCC 7120. Bio-protocol 6:e1890. https://doi.org/10.21769/bioprotoc.1890
Coates RC, Podell S, Korobeynikov A, Lapidus A, Pevzner P, Sherman DH, Allen EE, Gerwick L, Gerwick WH (2014) Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. PLoS ONE 9:e85140. https://doi.org/10.1371/journal.pone.0085140
Corpet F (1989) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16:10881–10890
Dunahay T, Jarvis E, Dais S, Roessler P (1996) Manipulation of microalgal lipid production using genetic engineering. Appl Biochem Biotechnol 57:223. https://doi.org/10.1007/BF02941703
Elhai J, Veprskis A, Muro-Pastor AM, Flores E, Wolk CP (1997) Reduction of conjugal transfer efficiency by three restriction activities of Anabaena sp. strain PCC 7120. J Bacteriol 179:1998–2005
Fan Q, Huang G, Lechno-Yossef S, Wolk CP, Kaneko T, Tabata S (2005) Clustered genes required for synthesis and deposition of envelope glycolipids in Anabaena sp. strain PCC 7120. Mol Microbiol 58:227–243. https://doi.org/10.1111/j.1365-2958.2005.04818.x
Fay P (1992) Oxygen reactions of nitrogen fixation in cyanobacteria. Microbiol Rev 56:340–373.
Flaherty BL, Van Nieuwenburgh F, Head SR, Golden JW (2011) Directional RNA deep sequencing sheds new light on the transcriptional response of Anabaena sp. strain PCC 7120 to combined-nitrogen deprivation. BMC Genomics 12:332. https://doi.org/10.1186/1471-2164-12-332
Greenwell HC, Laurens LML, Shields RJ, Lovett RW, Flynn KJ (2010) Placing microalgae on the biofuels priority list: a review of the technological challenges. J R Soc Interface 7:703–726. https://doi.org/10.1098/rsif.2009.0322
Halfmann C, Gu L, Gibbons W, Zhou R (2014a) Genetically engineering cyanobacteria to convert CO2, water, and light into the long-chain hydrocarbon farnesane. Appl Microbiol Biotechnol 98:8969–8977. https://doi.org/10.1007/s00253-014-6118-4
Halfmann C, Gu L, Zhou R (2014b) Engineering cyanobacteria for the production of a cyclic hydrocarbon fuel from CO2 and H2O. Green Chem 16:3175–3185. https://doi.org/10.1039/C6GC02591F
Houmarad J, de Manac NT (1988) [9] Cyanobacterial genetic tools: current status. Methods Enzymol 167:208–247. https://doi.org/10.1016/0076-6879(88)67092-3
Huang G, Chen F, Wei D, Zhang X, Chen G (2010) Biodiesel production by microalgal biotechnology. Appl Energy 87:38–46. https://doi.org/10.1016/j.apenergy.2009.06.016
Kaczmarzyk D, Fulda M (2010) Fatty acid activation in cyanobacteria mediated by acyl-acyl carrier protein synthetase enables fatty acid recycling. Plant Physiol 152:1598–1610. https://doi.org/10.1104/pp.109.148007
Kageyama H, Waditee-Sirisattha R, Sirisattha S, Tanaka Y, Mahakhant A, Takabe T (2015) Improved alkane production in nitrogen-fixing and halotolerant cyanobacteria via abiotic stresses and genetic manipulation of alkane synthetic genes. Curr Microbiol 71:115–120. https://doi.org/10.1007/s00284-015-0833-7
Kaiser BK, Carleton M, Hickman JW, Miller C, Lawson D, Budde M, Warrener P, Kageyama H, Waditee-Sirisattha R, Sirisattha S, Tanaka Y, Mahakhant A, Takabe T (2015) Improved alkane production in nitrogen-fixing and halotolerant cyanobacteria via abiotic stresses and genetic manipulation of alkane synthetic genes. Curr Microbiol 71:115–120. https://doi.org/10.1007/s00284-015-0833-7
Kaiser BK, Carleton M, Hickman JW, Miller C, Lawson D, Budde M, Warrener P, Paredes A, Mullapudi S, Navarro P, Cross F, Roberts JM (2013) Fatty aldehyde hydrases in cyanobacteria are a metabolically flexible precursor for a diversity of biofuel products. PLoS ONE 8:e85307. https://doi.org/10.1371/journal.pone.0085307
Klahn S, Baumgartner D, Pfreundt U, Voigt K, Schoen V, Steglich C, Hess WR (2014) Alkane biosynthesis genes in cyanobacteria and their transcriptional organization. Front Bioeng Biotechnol 2:24. https://doi.org/10.3389/fbioe.2014.00024
Liu X, Sheng J, Curtiss R III (2011) Fatty acid production in genetically modified cyanobacteria. Proc Nat Acad Sci USA 108:6899–6904. https://doi.org/10.1073/pnas.1103014108
Liu A, Zhui T, Lu X, Song L (2013) Hydrocarbon profiles and phylogenetic analyses of diversified cyanobacterial species. Appl Energy 111:383–393. https://doi.org/10.1016/j.apenergy.2013.05.008

Lu X (2010) A perspective: photosynthetic production of fatty acid-based biofuels in genetically engineered cyanobacteria. Biotechnol Adv 28:742–746. https://doi.org/10.1016/j.biotechadv.2010.05.021

Mendez-Perez D, Bergemann MB, Pfleger BF (2011) Modular synthase-encoding gene involved in alpha-olefin biosynthesis in Synechococcus sp. strain PCC 7002. Appl Environ Microbiol 77:4264–4267. https://doi.org/10.1128/aem.00467-11

Mitschke J, George J, Scholz I, Sharma CM, Dienst D, Bantscheff J, Voss B, Steglich C, Wilde A, Vogel J, Hess WR (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium Synechocystis sp. PCC6803. Proc Nat Acad Sci USA 108:2124–2129. https://doi.org/10.1073/pnas.1015154108

Ozdemir G, Karabay NU, Dalay MC, Pazarbasi B (2004) Antibacterial activity of volatile component and various extracts of Spirulina platensis. Phytother Res 18:754–757. https://doi.org/10.1002/ptr.1541

Peralta-Yahya PP, Zhang F, del Cardayre SB, Keasling JD (2012) Microbial engineering pathways to biofuels in photosynthetic microorganisms. Biofuels 5:67–78. https://doi.org/10.4155/bfs.13.67

Quintana N, Kooy F, Rhee M, Voshol G, Verpoorte R (2011) Renewable energy from cyanobacteria: energy production optimization by metabolic pathway engineering. Appl Microbiol Biotechnol 91:471–490. https://doi.org/10.1007/s00253-010-3394-0

Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320–W324. https://doi.org/10.1093/nar/gku316

Ruffing AM (2011) Engineered cyanobacteria: teaching an old bug new tricks. Bioeng Bugs 2:136–149

Savakis P, Hellingwerf KJ (2015) Engineering cyanobacteria for direct biofuel production from CO2. Curr Opin Biotechnol 33:8–14. https://doi.org/10.1016/j.copbio.2014.09.007

Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB (2010) Microbial biosynthesis of alkanes. Science 329:559–562. https://doi.org/10.1126/science.1187936

Takatani N, Usuki K, Kato A, Ikeda K, Kojima K, Aichi M, Maeda S, Omata T (2015) Essential role of acyl-ACP synthetase in acclimation of the cyanobacterium Synechococcus elongatus strain PCC 7942 to high-light conditions. Plant Cell Physiol 56:1608–1615. https://doi.org/10.1093/pcp/pcv086

Wang W, Liu X, Lu X (2013) Engineering cyanobacteria to improve photosynthetic production of alkan(e)nes. Biotechnol Biofuels 6:69. https://doi.org/10.1186/1754-6834-6-69

Wolk CF, Fan Q, Zhou R, Huan G, Lechene-Yossief S, Kurtz T, Wojcich E (2007) Paired cloning vectors for complementation of mutations in the cyanobacterium Anabaena sp. strain PCC 7120. Arch Microbiol 188:551–563. https://doi.org/10.1007/s00203-007-0276-z

Xu X, Gu L, He P, Zhou R (2015) Characterization of five putative aspartate amino transferase genes in the N2-fixing heterocystous cyanobacterium Anabaena sp. strain PCC 7120. Microbiology 161:1219–1230. https://doi.org/10.1099/mic.0.000083

Yoon HS, Golden JW (1998) Heterocyst pattern formation controlled by a diffusible peptide. Science 282:935–938

Zehr JP (2011) Nitrogen fixation by marine cyanobacteria. Trends Microbiol 19:162–173. https://doi.org/10.1016/j.tim.2010.12.004

Zhu T, Scalvenzi T, Sassoon N, Lu X, Gugger M (2018) Terminal olefin profiles and phylogenetic analyses of olefin synthases of diverse cyanobacterial species. Appl Environ Microbiol. https://doi.org/10.1128/aem.00425-18