UNIVERSITY OF CALIFORNIA SAN DIEGO

Impairment of O-antigen Synthesis in *Synechococcus elongatus* PCC 7942 Alters Outer Membrane Protein Composition

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

in

Marine Biology

by

Michelle Marie Esteban Prieto

Committee in charge:

Bianca Brahamsha, Chair
Douglas Bartlett
Brian Palenik

2019

The Thesis of Michelle Marie Esteban Prieto is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2019

TABLE OF CONTENTS

Signature Page..…..... iii

Table of Contents……..... iv

List of Figures……...……..... v

List of Tables……..... vi

Abstract of the Thesis……... vii

Introduction…… 1

Materials and Methods…….……..... 6

Results……... 14

Discussion……. 19

Figures and Tables……..... 25

References……..... 33

LIST OF FIGURES

Figure 1: Plate grazing assays of *S. elongatus* PCC 7942 and AMC 1908 preyed upon by amoebae HGG1 and LPG1 25

Figure 2: SDS-PAGE analysis of the outer membranes of *S. elongatus* PCC 7942 and AMC 1908 grown in liquid culture 26

Figure 3: SDS-PAGE analysis of the outer membranes of *S. elongatus* PCC 7942 and AMC 1908 grown on plates 27

Figure 4: SDS-PAGE analysis of the outer membranes of *S. elongatus* PCC 7942 and AMC 1908 grown in liquid versus solid media 28

LIST OF TABLES

Table 1: Proteomic analysis of discrete bands cut out of an SDS-PAGE gel ...…….....…… 29

Table 2: Presence and absence of proteins in outer membrane preparations (high speed pellets) of *S. elongatus* PCC 7942 and AMC 1908…….……..... 30

Table 3: Proteins significantly more abundant in the outer membrane preparation (high speed pellet) of *S. elongatus* PCC 7942 compared to AMC 1908…….…..…… 31

Table 4: Proteins significantly more abundant in the outer membrane preparation (high speed pellet) of *S. elongatus* AMC 1908 compared to PCC 7942……...……... 32

ABSTRACT OF THE THESIS

Impairment of O-antigen Synthesis in *Synechococcus elongatus* PCC 7942 Alters Outer Membrane Protein Composition

by

Michelle Marie Esteban Prieto

Master of Science in Marine Biology

University of California San Diego, 2019

Bianca Brahamsha, Chair

Now more than ever the growth of algal farming for biofuels and bioproducts as more sustainable alternatives to fossil fuels and its derivatives has created a need for more efficient cultivation strategies. Unfortunately, however, while open-air ponds are the most cost-effective way to grow the microalgae that are used for biofuels and bioproducts, they are also more susceptible to invasion by predators. Because of this, strategies for crop protection must be developed in order to minimize this pressure from grazers. A mutant of the model

cyanobacterium *Synechococcus elongatus* PCC 7942, called AMC 1908, which lacks the *wzm* (Synpcc7942_1126) gene for O-antigen synthesis, is resistant to grazing by amoebae, which can be algal pool predators. In order to determine the molecular mechanisms that could be behind this resistance, the outer membrane proteins of *S. elongatus* PCC 7942 and AMC 1908 were compared via SDS-PAGE analysis and proteomics. As expected, the protein compositions of *S. elongatus* PCC 7942 and AMC 1908 outer membranes differed quite substantially. Additionally, transport system substrate-binding protein IdiA (Synpcc7942_2175) was found to be more abundant in the mutant AMC 1908 when compared to the wild-type PCC 7942. This protein could be contributing to *S. elongatus* AMC 1908's resistance to grazing by amoebae, but only further testing can confirm this.

INTRODUCTION

The Impact of Grazing on the Cultivation of Microalgae for Biofuels and Bioproducts

In recent years, the microalgal farming industry for the production of biofuels and bioproducts has rapidly grown. This has created a demand for more efficient microalgal cultivation strategies (Georgianna and Mayfield 2012). Currently, open-air ponds are the most cost-effective way to grow the microalgae that are used for biofuels and its co-products (Richardson *et al.* 2014). These ponds are also more susceptible to predation (Shurin *et al.* 2013). Grazing pressure from protozoa, including ciliates, flagellates, and amoebae, is limiting the efficiency of algal growth in open-air ponds (Day *et al.* 2012). This is preventing the production of enough microalgal biomass to meet the growing demands of the biofuel and biotechnology industries (Schenk *et al.* 2008). Strategies for crop protection must be developed in order to mitigate this pressure from grazers (Ducat *et al.* 2011).

Synechococcus elongatus PCC 7942 Characteristics

One of the microalgal candidates for biofuel production belongs to the phylum Cyanophyta, the members of which are commonly referred to as cyanobacteria. Due to their ubiquitous nature and high growth efficiency, cyanobacteria can produce the large amounts biomass necessary for commercial biofuel production (Koller *et al.* 2014). Cyanobacteria are Gram-negative, which means they possess a cytoplasmic membrane, a peptidoglycan layer, and an outer membrane that contains LPS, which is made up of lipid A, core polysaccharides, and O-antigens (Hoiczyk and Hansel 2000). Certain strains can also have an additional surface layer, or S-layer, that lies on top of the outer membrane (Weckesser and Jurgens 1988, McCarren *et al.*

2005). *Synechococcus elongatus* PCC 7942 is a unicellular freshwater cyanobacterium with a small, sequenced genome. Its genetic tractability makes it the perfect model organism for conducting comparative studies in photoautotrophic organisms, like the microalgae used for biofuel production (Cohen and Golden 2015). For the same reason, it is also a favored organism for synthetic biology studies and genetic engineering (Ruffing 2011, Ducat *et al.* 2011). In this study, *Synechococcus elongatus* PCC 7942 is used as a model organism for studying the molecular mechanisms behind grazing resistance.

Known Cyanobacterial Resistance Mechanisms

Cyanobacteria have naturally evolved their own tactics to avoid detection, capture, ingestion, and digestion by predators (Matz and Kjelleberg 2005). With this said, not much is known about the molecular mechanisms that are responsible for these adaptations (Apple *et al.* 2011). Multiple lines of evidence reveal that the cell surface likely plays a major role in defense. In the marine cyanobacterium *Synechococcus* WH8102, the presence of SwmB, a giant cell surface protein, inhibits feeding by the dinoflagellate *Oxyrrhis marina*, whereas the presence of SwmA, an S-layer glycoprotein, makes *Synechococcus* cells more susceptible to grazing by ciliates (Strom *et al.* 2012). *Synechococcus* WH7803 mutants with lipopolysaccharide (LPS) layer modifications have also been shown to be more resistant to grazing by heterotrophic nanoflagellates (Zwirglmaier *et al.* 2009). Further support of this is demonstrated by the resistance of *Synechococcus elongatus* PCC 7942 O-antigen mutants that have a defective LPS to two phylogenetically distinct amoebae (Simkovsky *et al.* 2012). Another way that cyanobacteria could be protecting themselves from protozoan predators is by changing their cell surface features. In some cases, this can prevent cell to cell recognition (Roberts *et al.* 2011). For

intstance, two amoebae, *Acanthamoeba* and *Hartmannella vermiformis*, and the dinoflagellate, *Oxyrrhis marina*, attach to and capture their prey with sugar-binding and mannose-binding lectins. Hindrance of these receptors inhibits prey recognition (Allen and Dawidowicz 1990, Venkataraman *et al.* 1997, Wootton *et al.* 2007). In addition to this, it has been shown that prey morphology, including the formation of stiff filaments in the cyanobacterium *Oscilltoria*, can affect the ingestion ability of *Naegleria*, a genus of amoebae (Xinyao *et al.* 2006). Evidently, the predator-prey interactions between cyanobacteria and their protozoan predators are diverse and complex. Understanding the mechanisms that confer grazing resistance in cyanobacteria at the molecular level is incredibly important. Not only will it uncover the basis behind population mortality and restructuring in natural environments, but it can also lead to the development of more effective strategies for predation limitation in outdoor microalgae pools.

A Model System for Studying Grazing Resistance

Previous studies by Simkovsky *et al.* (2012) and Ma *et al.* (2016) have established the predator and prey interactions between *Synechococcus elongatus* PCC 7942 and two phylogenetically distinct heterolobosean amoebae, HGG1 and LPG1, as a model system for studying the molecular basis for grazing resistance in cyanobacteria. It was found that the impairment of O-antigen synthesis in the LPS of *S. elongatus* PCC 7942 mutants resulted in resistance to grazing by both amoebae. This resistance developed over time and was seen in cells grown between 8 to 10 days. Electron microscopy of an O-antigen mutant indicated the possible presence of an additional layer exterior to the outer membrane that could have been an S-layer (R. Simkovsky, unpublished). In order to investigate this further and identify the molecular mechanism responsible for this resistance, the protein compositions of the outer membranes of

both the wild-type *S. elongatus* strain (PCC 7942) and the O-antigen mutant strain (AMC 1908) were compared.

Synechococcus elongatus PCC 7942 Outer Membrane Protein Composition

The first proteomics map of *Synechococcus elongatus* PCC 7942 categorized proteins into 18 functional groups. (Koksharova *et al.* 2006) These groups included: (1) Proteins Involved in Cell Morphogenesis, (2 and 3) Cell Envelope Biogenesis and Peptidoglycan Synthesis Proteins, (4 and 5) Protein Synthesis and Post-Translational Protein Processing, (6) Protein-Protein Interactions, (7) Regulatory Function, (8 to 16) Proteins of General Cyanobacteria Cell Metabolism, (17) Oxidative Stress Defense Proteins, and (18) Unknown and Hypothetical Proteins. Proteins involved in cell envelope biogenesis and peptidoglycan synthesis included those in the outer membrane efflux protein (OEP) family, members of which form channels that allow for the export of different substrates. SomA, an outer membrane porin, was also identified. This protein and a homologous protein that was also identified contain S-layer homology (SLH) domain. Involved in the general (type II) secretion pathway (GSP) in Gram-negative bacteria, HofQ was another common protein hit. OstA, the organic solvent-tolerant protein, also came up in the proteomics data. It is involved in outer membrane permeability and plays an essential role in outer envelope biogenesis. OstA could also be part of a targeting system for outer membrane components. Lastly, UDP-N-acetylmuramyl tripeptide synthase and diaminopimelate epimerase were found to be major players in PCC 7942 peptidoglycan synthesis. (Koksharova *et al.* 2006)

S-Layer Proteins and their Functions

S-layers, which are made up of monomolecular crystalline arrays of proteinaceous subunits, are one of the most common surface structures on bacteria. Most S-layers are made up of a single protein or glycoprotein species with a molecular weight ranging from 40 to 200 kDa. Such proteins typically contain high proportions of acidic and hydrophobic amino acids, lysine being the predominant basic amino acid. In Gram-negative bacteria, the attachment of the S-layer to the outer membrane involves the LPS layer. These Gram-negative S-layers have been found to provide a selective advantage in competitive habitats. In *A. salmonicida*, *C. fetus*, *A. serpens*, and *C. crescentus*, the presence of S-layers conferred resistance to predation by the bacterium *B. bacteriovorus*, but did not shield against protozoan predators, like the ciliate *Tetrahymena* and the flagellate *Paraphysomonas* (Koval 1997). In *Synechococcus* GL-24, the S-layer lattice serves as a template for fine-grain mineralization in high calcium and sulfate ion concentrations. It is shed to prevent clogging of other cell envelope layers (Sara and Sleytr 2000).

Research Hypothesis

Considering what is known about cyanobacterial grazing resistance mechanisms, it is hypothesized that the *S. elongatus* O-antigen mutants compensate for outer membrane defects by altering the protein composition of their outer cell envelope. These changes in the cell surface could then affect recognition or ingestion by the amoebae, protecting the mutant cells from grazing. One possibility is that they are overexpressing S-layer proteins and building a surface layer around their outer membrane. If this is the case, the proteomics data should consistently show an increase in the relative abundance of S-layer proteins in the mutant. The presence or

absence of certain proteins can also lend insight into the driving factors behind the mutant's resistance to grazing by amoebae.

Master's Thesis Goal

Based on this prior research, the aim of this Master's thesis is to determine if there is a difference in the outer membrane protein composition of PCC 7942, the wild-type strain of *S. elongatus*, and AMC 1908, the mutant *S. elongatus* strain that lacks the *wzm* gene involved in O-antigen synthesis, after 10 days of growth, when the mutant cells exhibit resistance to grazing by amoebae. To answer this question, a comparative proteomics analysis was conducted. This analysis could lend insight into whether or not *S. elongatus* O-antigen mutants defend themselves against their amoebal predators by altering the proteins in their outer membranes.

MATERIALS AND METHODS

Growth Conditions

Liquid: *Synechococcus elongatus* PCC 7942 wild-type strains and AMC 1908 mutant strains were routinely grown at 30°C under medium light (20-25 μE m^{-2} s^{-1}) until the cultures reached an optical density (O.D.) of 0.6 at 750 nm. In order to maintain AMC 1908's O-antigen mutation, it was grown in 50mL BG11 medium with a chloramphenicol concentration of 7.5 μg/mL.

Solid: In order to determine the outer membrane compositions of cells grown on a solid surface, *Synechococcus elongatus* PCC 7942 wild-type strains and AMC 1908 mutant strains were also

grown on BG11 agar plates as follows. 1 mL of PCC 7942 and AMC 1908 liquid cultures with an O.D. 750 nm of 0.6 was centrifuged for 2 minutes at 17,000 x g. Supernatants were discarded and the pelleted cells were resuspended in the remaining liquid. 100 µL of each concentrated culture was then used to make lawns on BG11 + sodium thiosulfate plates. AMC 1908 cultures were spread on BG11 + sodium thiosulfate plates with a chloramphenicol concentration of 7.5 µg/mL.

10- to 13-Day Growth Experiments

750 mL of BG 11 liquid media was divided equally into 3 500 mL flasks. Each flask was inoculated with 5 mL of either PCC 7942 or AMC 1908 culture with an O.D. 750 nm of 0.6. The flasks were then grown for 10-13 days at 30°C under medium light (20-25 µE m^{-2} s^{-1}) with no shaking until they reached an O.D. 750 nm of 1. Similarly, 16 BG11 plates inoculated with 100 µL of concentrated PCC 7942 culture (O.D. 750nm = 1) and 8 BG11 plates inoculated with 100 µL of concentrated AMC 1908 culture (O.D. 750nm = 1) were grown for 10 days at 30°C under medium light (20-25 µE m^{-2} s^{-1}).

Outer Membrane Purification

The procedure followed was an adaptation of Brahamsha 1996 and Simkovsky *et al.* 2012 as described below.

Liquid: After 10-13 days of growth, the cells were harvested. Absorbance at 750 nm was measured. The cultures were also streaked on BG11 plates to check for contamination.

Solid: In order to harvest the lawns of cells, 5mL of BG11 was added to each plate, and the cells were gently resuspended with a glass spreader. The liquid then was pipetted off of the plate and into large autoclaved centrifuge tubes. 1mL was reserved to measure absorbance at 750 nm.

Both: The cells were centrifuged at 17,024 x g for 10 mins. at 20°C. Supernatants were discarded. 5mL of BG11 was then added to each of the pellets, and the cells were gently re-suspended. The re-suspended cells were then spun down at 20°C for another 10 mins. at 5,872 x g. Again, supernatants were discarded. After the second spin, 2 mL of ice cold stripping buffer [50mM Tris-HCl, 25mM Na2EDTA, pH 8.0, 15% sucrose wt/vol.] was added to each pellet. The cells were gently re-suspended and incubated on ice for 15 mins. They were then spun at 11,984 x g for 10 mins. at 4°C. After this, 2 mL of each supernatant, which now contains the extracted outer membranes and the periplasmic contents of the wild-type and mutant *S. elongatus* cells, was pipetted into sterile microfuge tubes. These supernatants went through one final centrifuge cycle at 17,000 x g for 3 mins at room temperature to remove any leftover unbroken cells. 980 µL of each supernatant was then transferred into new tubes.

High Speed Pelleting and Sample Concentration

In order to further separate the insoluble outer membranes from periplasmic contents and proteins solubilized by the EDTA treatment, the extracted membrane supernatants were spun at 4°C for 4 hours at 69,028 x g at 10°C. The supernatants, which contain the soluble periplasmic contents as well as cell surface proteins solubilized by EDTA, were then concentrated around 60 fold with Amicon Ultra-15 Centrifugal Filter tubes that had a 10,000 Da molecular weight cutoff

(Millipore) and were spun at 5,872 x g for 30 mins. The filters were then washed once with 4 mL of sterile milliQ and spun for 10 more mins., until the final volume remaining was less than 250 µL. The pellets were resuspended in 80 to 100 µL of sterile milliQ.

Proteomics

The outer membrane pellet and supernatant samples were sent for protein sequencing at UC San Diego's Biomolecular and Proteomics Mass Spectrometry Facility, where they were processed as described below.

Sample preparation

Protein samples were diluted in TNE (50 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA) buffer. RapiGest SF reagent (Waters Corp.) was added to the mix to a final concentration of 0.1% and samples were boiled for 5 min. TCEP (Tris (2- carboxyethyl) phosphine) was added to 1 mM (final concentration) and the samples were incubated at 37°C for 30 min. Subsequently, the samples were carboxymethylated with 0.5 mg/ml of iodoacetamide for 30 min at 37°C followed by neutralization with 2 mM TCEP (final concentration). Proteins samples prepared as above were digested with trypsin (trypsin:protein ratio - 1:50) overnight at 37°C. RapiGest was degraded and removed by treating the samples with 250 mM HCl at 37°C for 1 h followed by centrifugation at 14000 rpm for 30 min at 4°C. The soluble fraction was then added to a new tube and the peptides were extracted and desalted using C18 desalting columns (Thermo Scientific, PI-87782). Peptides were quantified using BCA assay and a total of 1 ug of peptides were injected for LC-MS analysis.

LC-MS-MS

Trypsin-digested peptides were analyzed by ultra high pressure liquid chromatography (UPLC) coupled with tandem mass spectroscopy (LC-MS/MS) using nano-spray ionization. The nanospray ionization experiments were performed using a Orbitrap fusion Lumos hybrid mass spectrometer (Thermo) interfaced with nano-scale reversed-phase UPLC (Thermo Dionex UltiMate™ 3000 RSLC nano System) using a 25 cm, 75-micron ID glass capillary packed with 1.7-µm C18 (130) BEHTM beads (Waters corporation). Peptides were eluted from the C18 column into the mass spectrometer using a linear gradient (5–80%) of ACN (Acetonitrile) at a flow rate of 375 µl/min for 1h. The buffers used to create the ACN gradient were: Buffer A (98% H2O, 2% ACN, 0.1% formic acid) and Buffer B (100% ACN, 0.1% formic acid). Mass spectrometer parameters are as follows; an MS1 survey scan using the orbitrap detector (mass range (m/z): 400- 1500 (using quadrupole isolation), 120000 resolution setting, spray voltage of 2200 V, Ion transfer tube temperature of 275 C, AGC target of 400000, and maximum injection time of 50 ms) was followed by data dependent scans (top speed for most intense ions, with charge state set to only include +2-5 ions, and 5 second exclusion time, while selecting ions with minimal intensities of 50000 at in which the collision event was carried out in the high energy collision cell (HCD Collision Energy of 30%), and the fragment masses where analyzed in the ion trap mass analyzer (With ion trap scan rate of turbo, first mass m/z was 100, AGC Target 5000 and maximum injection time of 35ms). Protein identification and label free quantification was carried out using Peaks Studio 8.5 (Bioinformatics solutions Inc.)

SDS-PAGE

The concentrations of each outer membrane sample were measured using the Pierce BCA Protein Assay Kit according to Thermo Scientific's recommendations and using Bovine Serum Albumin as a standard. Mutant and wild-type pellet and supernatant samples of equal concentrations were loaded onto Novex 12% Tris-Glycine Mini Gels, Wedge Well format, (Invitrogen) with 7 µL of the PageRuler Plus Prestained Protein Ladder (Thermo Scientific). The gels were run at 220V constant voltage for 50 minutes. They were then fixed in a 50% methanol, 7% acetic acid fixative solution and stained with SYPRO Ruby Red (Invitrogen). The gels were then washed with a 10% methanol, 7% acetic acid wash solution and imaged using Bio-Rad's GelDoc System.

Band Sequencing

Seven bands were cut out from a Tris-Glycine 12% gel loaded with mutant outer membrane preparations. The bands were cut into 1mm cubes, suspended in 40 µL of milli-Q water, and sequenced as described above. Hypothetical proteins were identified using NCBI's BLAST database.

Proteomics Data Analysis

Protein band assignments (Figure 2.B. and Table 1) were made by sorting the proteomics data in descending order by PEAKS protein confidence score (-10logP), which was calculated as the weighted sum of the -10logP scores of the protein's supporting peptides, Coverage %, or the percentage of the protein sequence that is covered by supporting peptides, Area Sample, or the total protein area found in each sample, and lastly, the number of high-confidence supporting

peptides (#Peptides). The proteins listed in Table 1 were clearly identified based on their relative abundances (Area Sample > 100,000,000) and confidence scores (-10logP > 260) relative to the other proteins detected in the sample.

Table 2 represents the presence and absence of specific proteins in the wild-type and mutant outer membrane preparation datasets. It was generated by a combination of Python scripts that compared the outer membrane proteomics data files from the *S. elongatus* PCC 7942 and AMC 1908 outer membrane preparations from cultures grown in liquid. The script pulled all of the proteins with the same accession number into one file. The proteins in that file were then sorted in descending order according to Area Sample. Anything with an area less than 180,000,000 was cut off. All proteins that had a -10logP value that was greater than 315, a Coverage % that was greater than 40, a #Peptides value that was greater than 15, and a Spec Sample value that was greater than 80 were included in the table. These cutoffs were based on the proteins that had the highest abundance (Area Sample) and significance (-10logP). Table 2 also lists the peak area ratios for each of the common proteins between *S. elongatus* PCC 7942 and AMC 1908. Using the quantitative proteomics data file that included proteins that had a fold change greater than 2 between the peak areas of each sample, these ratios were calculated by dividing the wild-type peak area by the mutant peak area for the PCC 7942 column and dividing the mutant peak area by the wild-type peak area for the AMC 1908 column.

The resulting data processed as described above was combined with data generated by a separate Python script that compared the outer membrane proteomics data files from the *S. elongatus* PCC 7942 outer membrane preparation from cultures grown in liquid to those from AMC 1908. The script pulled all of the proteins with accession numbers that were present in the mutant sample, but not present in the wild-type sample. The proteins in that file were then sorted

in descending order according to Area Sample. Anything with an area less than 2,200,000 was cut off. All proteins that had a -10logP value that was greater than 25, a Coverage % that was greater than 5, a #Peptides value that was greater than 1, and a Spec Sample value that was greater than 2 were included in Table 2. Again, these cutoffs were based on the proteins in the file that had the highest abundance (Area Sample) and significance (-10logP).

The proteins unique to the wild-type outer membrane preparation in Table 2 were identified in the same way as the previously described data, but for accession numbers present in the wild-type and not in the mutant. The Sample Area cutoff for this file was 12,000,000. This value was chosen according to the relative abundances of the other, less significant proteins in the sample. The rest of the cutoffs were the same as those listed in the paragraph above.

Tables 3 and 4 were generated using the quantitative proteomics data file that included proteins that had a fold change greater than 2 between the peak areas of each sample. Table 3 lists the proteins present in both the wild-type outer membrane preparation and the mutant outer membrane preparation that were significantly more abundant in the wild-type based on ratios calculated by dividing the wild-type peak area by the mutant peak area. All ratios greater than 11 were included.

Table 4 lists the proteins present in both the wild-type outer membrane preparation and the mutant outer membrane preparation that were significantly more abundant in the mutant based on ratios calculated by dividing the mutant peak area by the wild-type peak area. All ratios greater than 5 were included. This cutoff differs from that of Table 3, because the number of proteins that were more abundant in the mutant was much greater than the number of proteins that were more abundant in the wild-type. Cutoffs were adjusted according to the relative abundance of the proteins within each calculated sample set.

RESULTS

Verifying Grazing Resistance of *S. elongatus* AMC 1908

Figure 1 verifies *S. elongatus* AMC 1908's resistance to grazing by two phylogenetically distinct amoebae. Each of the four lawns pictured in Figure 1 were inoculated with two phylogenetically distinct amoebae, HGG1 and LPG1, on the same day. After seven days, it was evident that the amoebae HGG1 and LPG1, which were spotted in the center of plates A and B, respectively, had grazed a significant amount of the *S. elongatus* PCC 7942 lawns. The *S. elongatus* AMC 1908 cultures on plates C and D remained completely intact despite inoculation with the same set of amoebae. This particular set of lawns acted as resistance controls for the solid culture batch that was used for the SDS-PAGE gels in Figure 3 and part of Figure 4. The results pictured in Figure 1 agree with what has been shown for this mutant in the past (Simkovsky *et al.* 2012, Ma *et al.* 2016).

SDS-PAGE Analysis of *S. elongatus* PCC 7942 and AMC 1908 Outer Membrane Preparations

In order to get a first look at outer membrane protein differences between *S. elongatus* PCC 7942 and AMC 1908, purified outer membrane preparations (high speed pellets) along with periplasmic material and any solubilized proteins that were separated from the outer membrane pellets (supernatants) were run alongside wild-type and mutant whole cells grown in liquid culture. Figure 2 shows band differences between wild-type and mutant samples that were extracted from cultures grown in liquid media. Seven bands consistently showed up in the mutant outer membrane preparation samples. (See Figures 3 and 4.) These bands are boxed in

red in Figure 2.A. They are labeled in Figure 2.B., which magnifies the seven bands. Of these bands, Mutant Band 1 (MB1), MB2, and MB7 were not visible in the lanes loaded with wild-type samples. MB3, MB4, MB5, and MB6 were all visible in the wild-type outer membrane preparation sample (Lane 3). These four bands were around the same intensity in both the wild-type and mutant outer membrane preparation samples loaded with equal concentrations in Lanes 3 and 5, respectively. Each of the seven bands was sequenced. Table 1 assigns each band with its identified protein based on LC-MS sequencing. The high-speed supernatant samples, which should consist of periplasmic and solubilized material, differed from the high-speed pellet samples, which should contain mostly outer membrane material. Based on a visual analysis, there are more proteins present in the supernatant samples (Figures 2, 3, and 4). The seven strong sequenced bands that show up in all of the high-speed pellet samples are also either relatively very faint (Figures 3 and 4) or absent (Figure 2) in the high-speed supernatant samples.

SDS-PAGE Analysis of *S. elongatus* PCC 7942 and AMC 1908 Cultures Grown on Plates versus in Liquid Media

To compare the outer membrane profiles of liquid and plate-grown cultures, *S. elongatus* PCC 7942 and AMC 1908 outer membrane preparation samples along with their separated periplasmic and solubilized protein samples were extracted from cultures grown on plates and in liquid media. Figure 2.A. shows a Tris-Glycine 12% gel loaded with samples extracted from *S. elongatus* PCC 7942 and AMC 1908 cultures cultivated in liquid media. Figure 3 is an image of the same type of gel loaded with samples extracted from *S. elongatus* PCC 7942 and AMC 1908 cultures cultivated on plates. Figure 4 shows a side-by-side comparison of the same samples that were loaded into the gels in Figures 2 and 3, minus their whole cell samples. In Figure 4, it is

easier to see that there are similarities in the band patterns between samples extracted from solid versus liquid cultures. The biggest differences amongst the four outer membrane preparation samples (Figure 4, Lanes 1-4) are between the mutants grown on plates (Lane 4) and in liquid culture (Lane 3). The mutant outer membrane preparation extracted from cells grown on plates in Lane 4 contains a lot more visible bands than the preparation extracted from cells grown in flasks (Lane 3). Between 35 and 10 kDa, there are 6 intense bands that are either not visible or are very faint in the flask sample (Lane 3). There are also two bands between 70 and 250 kDa in the plate sample (Lane 4) that are not in the flask sample (Lane 3). There are no notable differences between the wild-type outer membrane preparation samples from cells grown on plates and in flasks, Lanes 2 and 1, respectively.

The supernatant samples in Lanes 6 through 9 of Figure 4 exhibited more variability than the pellet samples in both pattern and abundance between samples extracted from solid and liquid cultures. Among these four lanes, there were clear band pattern differences between the flask (Lanes 6 and 8) and plate (Lanes 7 and 9) samples. At around 20 kDa, there is a notable alternation of intense bands between the flask and plate samples. In the flask samples (Lanes 6 and 8), the first band in the pair of bands at ~20 kDa is more intense. This changes in the plate samples. In Lanes 7 and 9, the lower band in the pair is more intense. At 30 kDa, there is a similar pattern difference. 30 kDa bands in the flask samples (Lanes 6 and 8) are clearly more intense than those in the plate samples (Lanes 7 and 9). Lastly, in the mutant supernatant extracted from cells grown in flasks (Lane 8), there appears to be a band missing just below 35 kDa that is present in the three other lanes. The rest of the differences in Lanes 6 through 9 mostly look like variations in band intensity. For instance, just below the band at 55 kDa, there is an intense band in the mutant plate supernatant sample (Lane 9) that is very faint in both of the

wild-type samples. In addition, above the strong band present in both wild-type supernatant samples (Lanes 6 and 7) at 35 kDa, there is another less intense band at around 45 kDa that is not present in both mutant supernatant samples (Lanes 8 and 9). Other than that, there aren't many notable differences between the wild-type and mutant supernatant samples.

Despite band differences between each of the samples in Figure 4, the seven sequenced bands (Figure 2.B. and Table 1) were present in mutant outer membrane preparations extracted from both liquid and solid cultures. They are boxed in red in Figure 3. Not only are these bands consistent between the mutants grown under both conditions, but they also exhibit similar band intensities (Figure 4). We noticed that cells grown on plates, unlike those grown in liquid, underwent some lysis during the harvesting process and the variability seen in membrane preparations may be due to the contamination from cytoplasmic proteins.

S. elongatus PCC 7942 and AMC 1908 Outer Membrane Protein Composition

Shotgun proteomic analysis of a single sample of *S. elongatus* PCC 7942 and AMC 1908 outer membrane preparations was carried out in order to identify and compare proteins present in the wild-type and in the mutant. In total, 787 proteins were detected in the PCC 7942 outer membrane preparation (high speed pellet) and 697 proteins were detected in the AMC 1908 outer membrane preparation. Additionally, 727 proteins were detected in the periplasmic and solubilized material (supernatant) of PCC 7942 and 800 proteins were detected in the periplasmic and solubilized material of AMC 1908. Table 2 shows 11 proteins that were present at similar abundances in both the wild-type and mutant outer membrane preparations. These proteins are those highlighted in green under both the PCC 7942 and AMC 1908 columns, and

they are consistent with those pointed out in Figure 2. Of the common PCC 7942 and AMC 1908 proteins indicated in Table 2, 6 were present in the seven sequenced bands (Figure 2.B. and Table 1). These 6 included the proteins identified for MB1, MB3, MB4, MB5, MB6, and MB7 (Table 1). Based on peak abundance ratios (Table 2), 5 of these 6 proteins were at the same relative abundance in both the mutant and the wild-type. The only exception was the transport system substrate-binding protein IdiA (MB7), which was significantly more abundant in the mutant compared to the wild-type (Tables 2 and 4) at a ratio of 7.14. This evidence of overexpression is also consistent with MB7's intensity on each of the gels (Figures 2, 3, 4, and 5). Phycocyanin was also among the common proteins in PCC 7942 and AMC 1908. However, because phycocyanin is a photosynthetic pigment that is not associated with the outer membrane (Nagarajan *et al.* 2019), these proteins are likely cytosolic contaminants that occurred due to cell lysis during the outer membrane purification process. Other proteins, like porins and proteins involved in binding and transport, would be expected to be found in the outer membrane (Koksharova *et al.* 2006).

Quantitative Proteomic Differences between *S. elongatus* PCC 7942 and AMC 1908 Outer Membranes

Table 2 also lists proteins that are present in the mutant outer membrane preparation but are missing in the wild-type outer membrane preparation. These proteins are those indicated with a green "P" for "present" under the AMC 1908 column and a red "ND" for "not detected" under the PCC 7942 column. Proteins that are only found in the wild-type outer membrane preparation are indicated with a green "P" under the PCC 7942 column and a red "ND" under the AMC 1908 column. Compared to the other more abundant proteins within the proteomics dataset, these

unique proteins were less significant and had lower coverage percentages (Coverage %) and peptide counts (#Peptides). This may be an indication that they may not represent significant or reproducible differences. Proteomics replicates must be performed to confirm whether or not these proteins consistently show up as only present in the mutant and vice versa.

Table 3 identifies proteins significantly more abundant in the wild-type outer membrane preparation when compared to the mutant outer membrane preparation. Conversely, Table 4 lists the proteins that were significantly more abundant in the mutant outer membrane preparation when compared to the wild-type outer membrane preparation. Among the seven sequenced bands, only MB7 (Figure 2.B. and Table 1), the transport system substrate-binding protein IdiA, is significantly overexpressed in the mutant at a ratio of 7.14 (Tables 2 and 4). This protein was not significantly abundant in the wild-type outer membrane preparation.

DISCUSSION

A First Look at Protein Differences: SDS-PAGE analysis of *S. elongatus* PCC 7942 and AMC 1908 Outer Membranes

According to the results of the gel in Figures 2 and 3, there are band differences between *S. elongatus* PCC 7942 and AMC 1908 outer membrane preparations extracted from cultures grown in both liquid media (Figure 2) and on plates (Figure 3). This supports the hypothesis that the lack of the O-antigen is leading to outer membrane protein composition changes in the mutant. At a qualitative level, this initial SDS-PAGE analysis answers the most basic question: Are there protein composition differences between the outer membranes of *S. elongatus* PCC 7942 and AMC 1908?

Because the outer membrane purification procedure was resulting in cell lysis and cytosolic contamination when outer membranes were extracted from cells grown on plates, the next question was: Are there protein composition differences between the outer membranes extracted from cultures grown in liquid versus solid media? This question was answered in Figure 4, which compares wild-type and mutant liquid culture extracts and plate extracts side by side. Between the outer membrane samples in Lanes 1 to 4, the band patterns are more or less consistent. The flask and plate wild-type samples in Lanes 1 and 2 appear to have the exact same patterns. The mutants share four intense bands with the wild-type samples (Figure 2). Between the mutant outer membrane samples (Lanes 3 and 4), however, there are more bands present in the plate extract (Lane 4) than in the outer membrane preparations extracted from cells grown in liquid culture. This could be an indication that the mutant cultures grown on plates are more sensitive to lysis during the outer membrane purification process. The additional bands could be proteins associated with the cytoplasm. This remains to be determined.

In the sample supernatants (Lanes 6 to 9), which are expected to contain periplasmic proteins and other solubilized material, a similar occurrence is seen (Figure 4). Again, between the wild type samples (Lanes 6 and 7), there aren't striking differences in band pattern. Among the mutant samples extracted from liquid and solid culture (Lanes 8 and 9, respectively), however, there were several band differences. The majority of these differences were differences in abundance, although the plate sample (Lane 9) contained additional visible bands that were not present in the sample extracted from liquid culture. Just like in the mutant outer membrane samples (Lanes 3 and 4), this increase in visible proteins is probably a product of cell lysis during the purification process. There are striking differences in protein abundances between supernatant samples grown in liquid versus solid media at around 20 and 30 kDa. In addition, in

the mutant flask supernatant sample in Lane 8, a band is missing just below 35 kDa. While interesting, these bands were not sequenced. However, if the outer membrane purification protocol can be modified to reduce cell lysis on plates and minimize cytosolic contamination, it would be valuable to do a proteomics comparison between outer membrane pellet and supernatant preparations extracted from cultures grown on plates versus in flasks.

The final question that can be answered via SDS-PAGE analysis is the following: What are some of the notable outer membrane protein differences between the wild-type and the mutant? In Figure 2, MB1, MB2, and MB7 were not visible in the lanes loaded with wild-type samples. These proteins are extracellular solute-binding protein, branched-chain amino-acid ABC transport system periplasmic binding protein, and transport system substrate-binding protein IdiA, respectively (Table 1). Quantitative proteomics data later revealed that these three proteins were not actually absent in the wild-type outer membrane preparation. They just appeared to be absent when run on a gel. This could be due to a number of reasons, including the possibility of multiple proteins overlapping on the gel, solubility issues, and/or post-translational modifications. Nevertheless, according to the proteomics ratios, while extracellular solute-binding protein, branched-chain amino-acid ABC transport system periplasmic binding protein, and transport system substrate-binding protein IdiA were present in both wild-type and mutant outer membrane preparations, only the transport system substrate binding protein IdiA was significantly more abundant in the mutant.

Testing the S-layer Hypothesis

Because S-layer proteins easily solubilize into monomers when cations such as Ca^{2+} are chelated (McCarren *et al.* 2005, Sara and Sleytr 2000), they would be expected to show up in the supernatant of the purified outer membrane samples, which should contain solubilized material. An uncharacterized conserved hypothetical protein containing an S-layer homology domain (Synpcc7942_0443) was identified in both *S. elongatus* PCC 7942 and AMC 1908. Contrary to what was expected, it was present in the high speed pellets of the outer membrane preparations. This may be because it did not solubilize as easily as typical S-layers do. However, it was not more abundant in the mutant sample, and its presence in both *S. elongatus* PCC 7942 and AMC 1908 does not support the idea that it may be responsible for grazing resistance.

Quantitative Proteomics Differences between *S. elongatus* PCC 7942 and AMC 1908 Outer Membranes Identify Proteins that Could be Involved in Resistance

According to the proteomics data, only one of the 3 unique bands, MB1, MB2, and MB7 (Figure 2 and Table 1), identified in the mutant outer membrane samples on gels is more abundant in the mutant outer membrane. This band, MB7, was identified as transport system substrate-binding protein IdiA (Table 1). It was significantly more abundant than the wild-type based on a peak area ratio of 7.14 (Tables 2 and 4). With this said, it was still present in the wild-type at high enough abundances to be included in Table 2, which indicates common outer membrane proteins between the wild-type and the mutant. Table 2 also lists proteins that are present in the mutant outer membranes and not in the wild-type outer membranes. Amongst the proteins in this table, type IV pilus assembly protein PilO (Synpcc7942_2451) stood out. Other pili proteins were present in the wild-type and mutant samples, but not type IV pilus assembly

protein PilO. This was only present in the mutant. Pili are a known bacterial defense mechanism (Harvey *et al.* 2018), and since this particular protein is present in the mutant outer membrane and not in the wild-type outer membrane, it could be involved in resistance, potentially affecting the ingestion or digestion of *S. elongatus* AMC 1908 by amoebae. This type IV pilus assembly protein could even be masking the detection of AMC 1908 by amoebae if they rely on a different surface receptor to recognize their prey (Harvey *et al.* 2018). With all of this said, while this particular protein had a high sample area of 1210000000, its -10logP was 30.82, its peptide count was 2, and it only had 7% coverage. This means it is likely not significant. More proteomics replicates must be carried out in order to determine whether or not PilO consistently shows up in only the mutant and not the wild-type. Electron microscopy and a pili stain could also reveal whether or not the mutants are actually forming pili.

Returning to MB7 (Figure 2.B. and Table 1), which is significantly overexpressed in the mutant pellet (Table 4), in *Synechococcus elongatus* PCC 7942, this particular transport system substrate-binding protein, also known as iron deficiency induced protein, or IdiA (Synpcc7942_2175), allows for growth under iron or manganese stress and accumulates under iron and manganese limited conditions (Nodop *et al.* 2008). IdiA is homologous to iron-binding proteins and is likely involved in transporting iron into the cell (Michel *et al.* 1999, Webb *et al.* 2001). Insertional inactivation of the *idiA* gene that encodes for IdiA in *S. elongatus* PCC 7942 prevented growth under iron and manganese stress (Michel *et al.* 1996). Since the AMC 1908 mutant's resistance develops over time and is seen at 8 to 10 days of growth, it's possible that as the iron and manganese in BG11 are getting used up by the increasing cell density of *S. elongatus* AMC 1908, IdiA begins to accumulate in response (Michel *et al.* 1996). If this is the case, it is unclear why the same phenomenon is not seen in the wild-type. The overexpression of

this transport system substrate-binding protein could then be conferring AMC 1908's resistance to grazing by amoebae by outcompeting the amoebae for iron. This can be directly tested with iron limitation experiments. IdiA is a proposed marker for iron scavenging (Webb *et al.* 2001). If this is true, then AMC 1908 mutants may be using up the iron in BG11, thus leaving none for the amoebae. Ecologically, this could also makes sense, as it could protect mutants in nutrient limited conditions from grazers that would only further lessen their chances of survival.

Future Directions

In order to determine whether or not the protein of interest identified, transport system substrate-binding protein IdiA, plays a role in *S. elongatus* AMC 1908's development of resistance, genetic experiments can be conducted. The *idiA* gene that encodes for said protein could be knocked out in AMC 1908 (Clerico *et al.* 2007, Taton *et al.* 2014, Chen *et al.* 2016), and a grazing assay (Simkovsky *et al.* 2012, Ma *et al.* 2016) could be used to test if the knockout mutant is still resistant to grazing by amoebae or not. Similarly, an overexpression experiment could also be conducted. IdiA could be overexpressed in *S. elongatus* PCC 7942 (Clerico *et al.* 2007, Taton *et al.* 2014, Chen *et al.* 2016). Once again, the grazing assay (Simkovsky *et al.* 2012, Ma *et al.* 2016) could be used to test whether the overexpression of transport system substrate-binding protein IdiA confers resistance to grazing by amoebae in the wild-type.

FIGURES AND TABLES

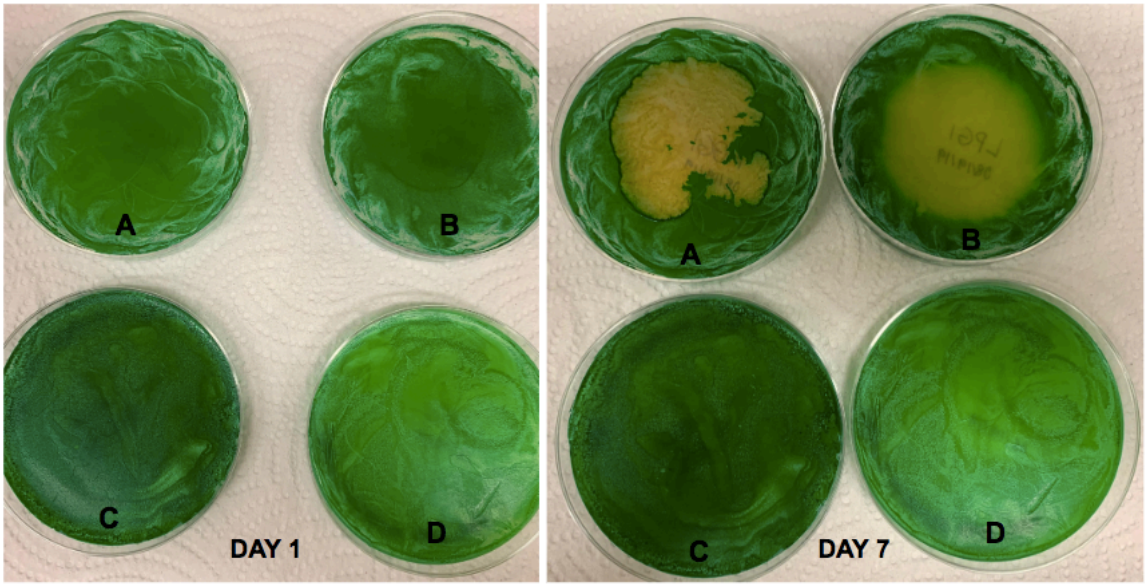

Figure 1: 10-day-old PCC 7942 and AMC 1908 lawns were spotted with the amoebae HGG1 and LPG 1 on Day 1. The plates were left under constant light at room temperature for 7 days. (A) PCC 7942 lawn spotted with HGG1 in the center of the plate. (B) PCC 7942 lawn spotted with LPG1 in the center of the plate. (C) AMC 1908 lawn spotted with HGG1 in the center of the plate. (D) AMC 1908 lawn spotted with LPG1 in the center of the plate.

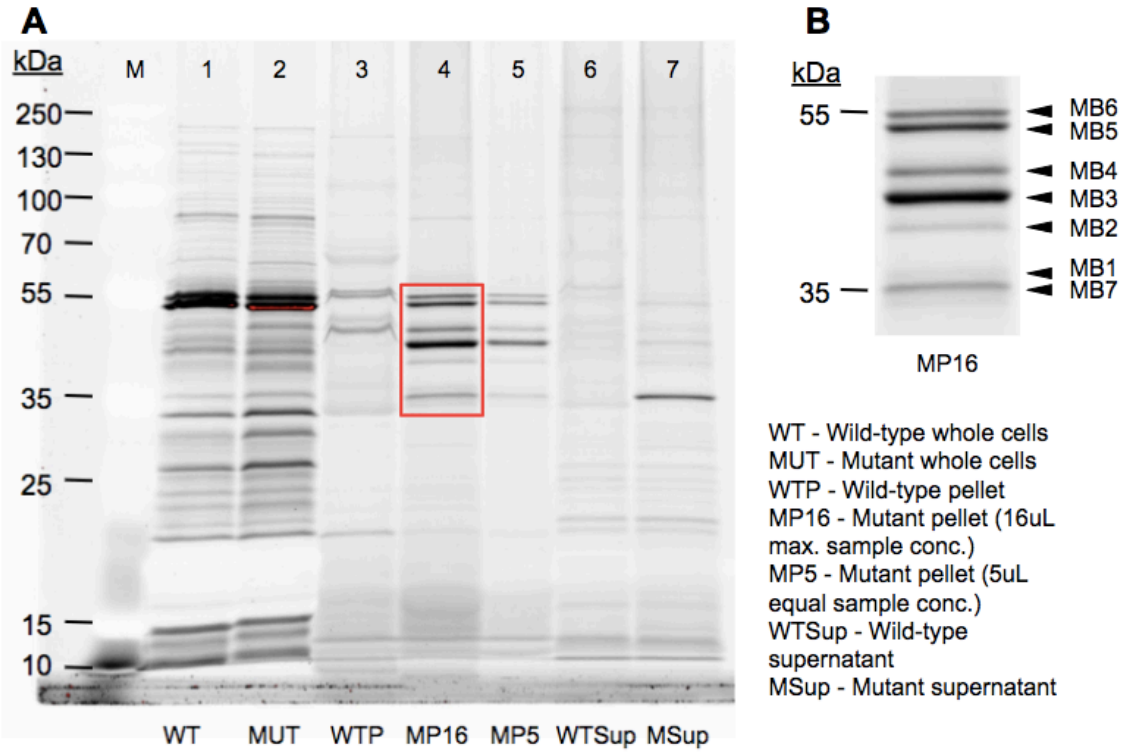

Figure 2: (A) Tris-Glycine 12% gel loaded with PCC 7942 and AMC 1908 whole cells grown in liquid culture (Lanes 1 and 2) and equal concentrations (1.6 µg) of their purified outer membrane samples (Lanes 3, 5, 6, and 7). Lane 4 was loaded with the same sample as Lane 5, but at its maximum protein concentration (5.8 µg). Lane M contains the PageRuler Plus Prestained Protein Ladder. The red box highlights the seven bands that were sequenced. (B) Close-up image of the seven bands that were sequenced. See Table 1 for protein sequence identification results.

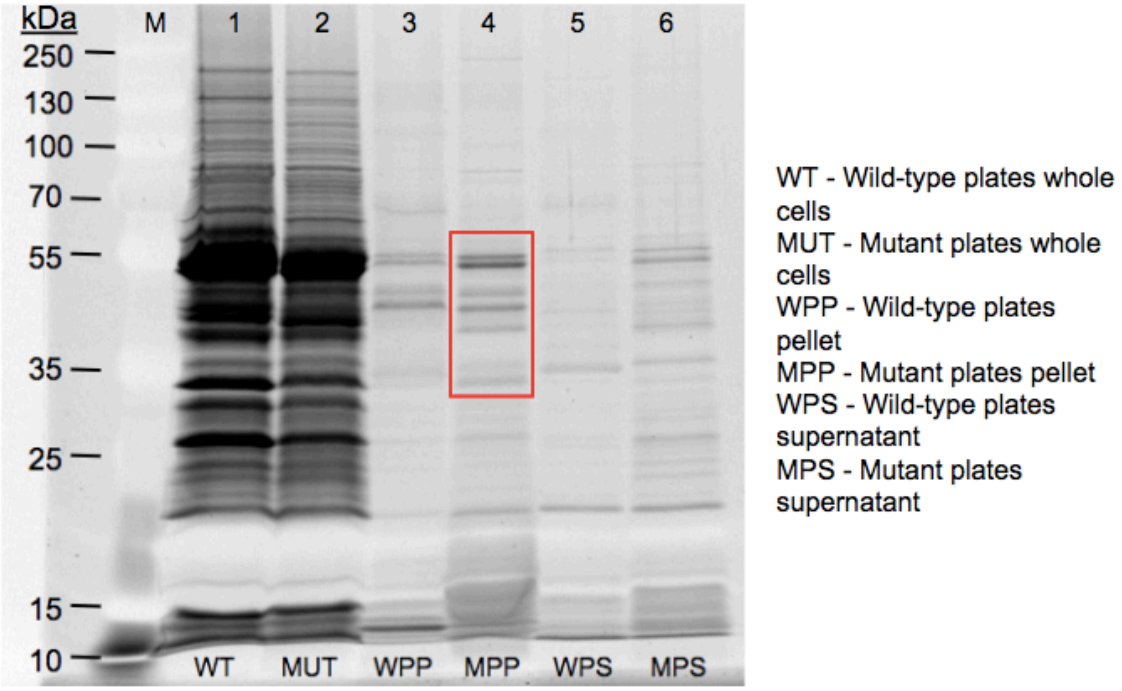

Figure 3: Tris-Glycine 12% gel loaded with PCC 7942 and AMC 1908 whole cells grown on plates (Lanes 1 and 2) and equal concentrations (1.6 µg) of their purified outer membrane samples (Lanes 3-6). Lane M contains the PageRuler Plus Prestained Protein Ladder. The red box highlights the seven bands that were sequenced. (See Figure 2 and Table 1).

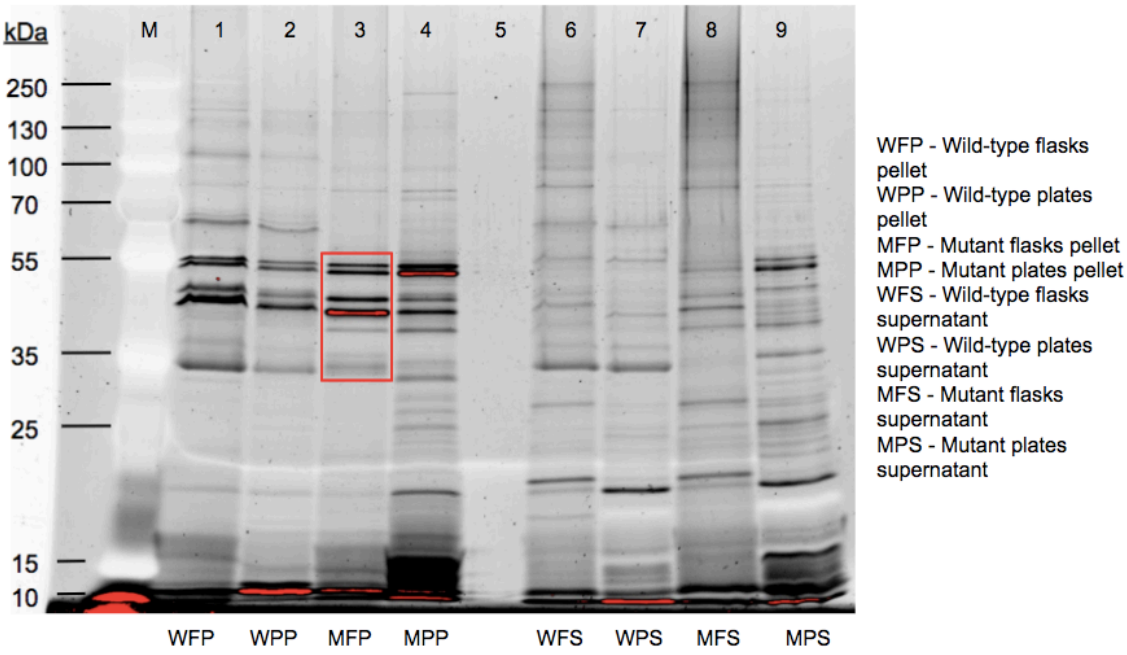

Figure 4: Tris-Glycine 12% gel loaded with equal concentrations (1.6 µg) of PCC 7942 and AMC 1908 outer membrane pellets from cells grown in liquid and on plates (Lanes 1-4) and equal concentrations of their supernatants (Lanes 6-9). Lane M contains the PageRuler Plus Prestained Protein Ladder. The red box highlights the seven bands that were sequenced. (See Figure 2 and Table 1).

Table 1: Proteomic analysis of discrete bands cut out of an SDS-PAGE gel. (See Figure 2).

Band	Mol. Weight on Gel (Da)	Protein Description	Gene Name	Locus Tag	Avg. Predicted Mol. Mass (Da)
MB1	37000	extracellular solute-binding protein	--	Synpcc7942_0246	39326
MB2	43000	branched-chain amino-acid ABC transport system periplasmic binding protein	*livJ*	Synpcc7942_2496	42504
MB3	45000	bicarbonate-binding protein CmpA	*cmpA*	Synpcc7942_1488	49108
MB4	50000	ABC-type nitrate/nitrite transport system substrate-binding protein	*nrtA*	Synpcc7942_1239	48425
MB5	53000	porin; major outer membrane protein	*somA(1)*	Synpcc7942_1464	57139
MB6	55000	porin; major outer membrane protein	*somB(2)*	Synpcc7942_1635	57909
MB7	35000	transport system substrate-binding protein	*idiA*	Synpcc7942_2175	40260

Table 2: Presence and absence of proteins in outer membrane preparations (high speed pellets) of *S. elongatus* PCC 7942 and AMC 1908. P (green) = present, ND (red) = not detected. Ratios under PCC 7942 are WT:MUT peak area. Ratios under AMC 1908 are MUT:WT peak area.

Locus Tag	Protein	PCC 7942	AMC 1908
Synpcc7942_1488	bicarbonate-binding protein CmpA	0.89	1.12
Synpcc7942_1635	porin; major outer membrane protein	1.93	0.52
Synpcc7942_1464	porin	1.11	0.90
Synpcc7942_1239	ABC-type nitrate/nitrite transport system substrate-binding protein	1.03	0.97
Synpcc7942_2444	phosphate binding protein	1.19	0.84
Synpcc7942_0246	extracellular solute-binding protein	0.70	1.43
Synpcc7942_1224	ABC-transporter membrane fusion protein	0.47	2.13
Synpcc7942_1052	phycocyanin subunit beta	0.97	1.03
Synpcc7942_0443	S-layer homology	8.29	0.12
Synpcc7942_1397	hypothetical protein Synpcc7942_1397	1.62	0.62
Synpcc7942_2175	transport system substrate-binding protein IdiA	0.14	7.14
Synpcc7942_2451	type IV pilus assembly protein PilO	ND	P
Synpcc7942_2431	hypothetical protein Synpcc7942_2431	ND	P
Synpcc7942_0997	50S ribosomal protein L32	ND	P
Synpcc7942_0976	hypothetical protein Synpcc7942_0976	ND	P
Synpcc7942_1491	nitrate transport ATP-binding subunits C and D	ND	P
Synpcc7942_1009	XRE family transcriptional regulator	ND	P
Synpcc7942_2248	nucleoid protein Hbs	ND	P
Synpcc7942_2492	ATPase	ND	P
Synpcc7942_2354	peptidylprolyl isomerase	P	ND
Synpcc7942_1613	hypothetical protein Synpcc7942_1613	P	ND
Synpcc7942_2608	hypothetical protein Synpcc7942_2608	P	ND
Synpcc7942_1384	hypothetical protein Synpcc7942_1384	P	ND
Synpcc7942_2257	hypothetical protein Synpcc7942_2257	P	ND
Synpcc7942_2539	hypothetical protein Synpcc7942_2539	P	ND
Synpcc7942_0195	hypothetical protein Synpcc7942_0195	P	ND
Synpcc7942_0905	hypothetical protein Synpcc7942_0905	P	ND
Synpcc7942_1909	hypothetical protein Synpcc7942_1909	P	ND

Table 3: Proteins significantly more abundant in the outer membrane preparation (high speed pellet) of *S. elongatus* PCC 7942 compared to AMC 1908 using a cutoff of 11

Locus Tag	Protein	WP78:MP78
Synpcc7942_1182	NADH dehydrogenase subunit J	100.00
Synpcc7942_0905	hypothetical protein Synpcc7942_0905	50.00
Synpcc7942_1128	hypothetical protein Synpcc7942_1128	50.00
Synpcc7942_0374	hypothetical protein Synpcc7942_0374	50.00
Synpcc7942_1678	hypothetical protein Synpcc7942_1678	33.33
Synpcc7942_0158	hypothetical protein Synpcc7942_0158	25.00
Synpcc7942_0289	preprotein translocase subunit SecA	25.00
Synpcc7942_0318	hypothetical protein Synpcc7942_0318	20.00
Synpcc7942_1515	Ser/Thr phosphatase	20.00
Synpcc7942_0969	carboxymethylenebutenolidase	20.00
Synpcc7942_1632	hypothetical protein Synpcc7942_1632	16.67
Synpcc7942_2152	hypothetical protein Synpcc7942_2152	14.29
Synpcc7942_0128	hypothetical protein Synpcc7942_0128	14.29
Synpcc7942_0424	photosystem q(b) protein	12.50
Synpcc7942_2086	hypothetical protein Synpcc7942_2086	12.50
Synpcc7942_1484	UDP-N-acetylmuramoylalanyl-D-glutamate—2,6 diaminopimelate ligase	12.50
Synpcc7942_1427	ribulose 1,5-biphosphate carboxylase small subunit	11.11

Table 4: Proteins significantly more abundant in the outer membrane preparation (high speed pellet) of *S. elongatus* AMC 1908 compared to PCC 7942 using a cutoff of 5

Locus Tag	Protein	MP78:WP78
Synpcc7942_0632	50S ribosomal protein L10	5.26
Synpcc7942_2491	DNA gyrase subunit B	5.56
Synpcc7942_2006	hypothetical protein Synpcc7942_2006	5.56
Synpcc7942_0896	septum site-determining protein MinD	5.56
Synpcc7942_0226	twin-arginine translocation protein TatA	5.88
Synpcc7942_1944	pyruvate dehydrogenase (lipoamide)	6.25
Synpcc7942_1656	catalase/peroxidase HPI	6.67
Synpcc7942_2175	transport system substrate-binding protein IdiA	7.14
Synpcc7942_0757	hypothetical protein Synpcc7942_0757	7.69
Synpcc7942_0513	ATPase	9.09
Synpcc7942_1381	ATPase	9.09
Synpcc7942_0095	two component transcriptional regulator	9.09
Synpcc7942_0105	NAD synthetase	9.09
Synpcc7942_2118	DUF1816 domain-containing protein	10.00
Synpcc7942_0886	30S ribosomal protein S7	10.00
Synpcc7942_1524	DNA-directed RNA polymerase subunit beta	10.00
Synpcc7942_0313	pentapeptide repeat-containing protein	12.50
Synpcc7942_0628	spermidine synthase	14.29
Synpcc7942_1654	hypothetical protein Synpcc7942_1654	14.29
Synpcc7942_1723	carotene isomerase	20.00
Synpcc7942_2387	c-type cytochrome	25.00
Synpcc7942_2171	starvation induced DNA binding protein	33.33
Synpcc7942_1591	RNA binding S1	33.33

REFERENCES

Georgianna DR, Mayfield SP. 2012. Exploiting diversity and synthetic biology for the production of algal biofuels. Nature **488**:329–335.

Richardson JW, Johnson MD, Zhang X, Zemke P, Chen W, Hu Q. 2014. A financial assessment of two alternative cultivation systems and their contributions to algae biofuel economic viability. Algal Res **4**:96–104.

Shurin JB, Abbott RL, Deal MS, Kwan GT, Litchman E, McBride RC, Mandal S, Smith VH. 2013. Industrial-strength ecology: trade-offs and opportunities in algal biofuel production. Ecol Lett **16**:1393–1404.

Day JG, Thomas NJ, Achilles-Day UE, Leakey RJ. 2012. Early detection of protozoan grazers in algal biofuel cultures. Bioresour Technol **114**:715–719.

Schenk PM, Thomas-Hall SR, Stephens E, Marx UC, Mussgnug JH, Posten C, Kruse O, Hankamer B. 2008. Second generation biofuels: High-efficiency microalgae for biodiesel production. Bioenergy Res **1**:20–43.

Ducat DC, Way JC, Silver PA. 2011. Engineering cyanobacteria to generate high-value products. Trends Biotechnol **29**:95–103.

Koller M, Muhr A, Braunegg G. 2014. Microalgae as versatile cellular factories for valued products. Algal Res **6**:52–63.

Hoiczyk E, Hansel A. 2000. Cyanobacterial cell walls: news from an unusual prokaryotic envelope. J Bacteriol **182**:1191-1199.

Weckesser J, Jurgens UJ. 1988. Cell walls and external layers. Methods in Enzymology **167**:173-188.

McCarren J, Heuser J, Roth R, Yamada N, Martone M, Brahamsha B. 2005. Inactivation of *swmA* results in the loss of an outer cell layer in a swimming *Synechococcus* strain. J Bacteriol **187**:224-230.

Cohen SE, Golden SS. 2015. Circadian rhythms in cyanobacteria. Microbiol Mol Biol Rev **79**:373-385.

Ruffing AM. 2011. Engineered cyanobacteria to generate high-value products. Trends Biotechnol **29**:95-103.

Matz C, Kjelleberg S. 2005. Off the hook—how bacteria survive protozoan grazing. Trends Microbiol **13**:302–307.

Apple JK, Strom SL, Palenik B, Brahamsha B. 2011. Variability in protist grazing and growth on different marine *Synechococcus* isolates. Appl Environ Microbiol **77**:3074–3084.

Strom SL, Brahamsha B, Fredrickson KA, Apple JK, Rodríguez AG. 2012. A giant cell surface protein in *Synechococcus* WH8102 inhibits feeding by a dinoflagellate predator. Environ Microbiol **14**:807–816.

Zwirglmaier K, Spence E, Zubkov MV, Scanlan DJ, Mann NH. 2009. Differential grazing of two heterotrophic nanoflagellates on marine *Synechococcus* strains. Environ Microbiol **11**:1767–1776.

Simkovsky R, Daniels EF, Tang K, Huynh SC, Golden SS, Brahamsha B. 2012. Impairment of O-antigen production confers resistance to grazing in a model amoeba-cyanobacterium predator–prey system. Proc. Natl. Acad. Sci. U. S. A. **109**:16678–16683.

Roberts EC, Legrand C, Steinke M, Wootton EC. 2011. Mechanisms underlying chemical interactions between predatory planktonic protists and their prey. J Plankton Res **33**:833-841.

Allen PG, Dawidowicz EA. 1990. Phagocytosis in *Acanthamoeba*: I. A mannose receptor is responsible for the binding and phagocytosis of yeast. J Cell Physiol **145**:508-513.

Venkataraman C, Haack BJ, Boudada S, AbuKwaik Y. 1997. Identification of Gal/GalNac lectin in the protozoan *Hartmanella vermiformis* as a potential receptor for attachment and invasion by the Legionnaire's disease bacterium. J Exp Med **186**:537-547.

Wootton EC, Zubkov MV, Jones DH, Jones RH, Martel CM, Thornton CA, Roberts EC. 2007. Biochemical prey recognition by planktonic protozoa. Environ Microbiol **9**:216-222.

Xinyao L, Miao S, Yonghong L, Yin G, Zhongkai Z, Donghui W, Weizhong W, Chencai A. 2006. Feeding characteristics of an amoeba (Lobosea: *Naegleria*) grazing upon cyanobacteria: Food selection, ingestion and digestion progress. Microb Ecol **51**:315–325.

Ma AT, Daniels EF, Gulizia N, Brahamsha B. 2016. Isolation of diverse amoebal grazers of freshwater cyanobacteria for the development of model systems to study predator-prey interactions. Algal Research-Biomass Biofuels and Bioproducts **13**:85-93.

Koksharova OA, Klint J, Rasmussen U. 2006. The first protein map of Synechococcus sp. strain PCC 7942. Microbiology **75**(6):664-672.

Koval SF. 1997. The effect of S-layers and cell surface hydrophobicity on prey selection by bacterivorous protozoa. FEMS Microbiol Rev **20**:138–142.

Sara M, Sleytr UB. 2000. S-Layer Proteins. J Bacteriol **184**(4):859-868.

Guttman M, Betts GN, Barnes H, Ghassemian M, van der Geer P, Komives EA. 2009. Interactions of the NPXY microdomains of the low density lipoprotein receptor-related protein 1. Proteomics **9**(22):5016-28.

McCormack AL, Schieltz DM, Goode B, Yang S, Barnes G, Drubin D, Yates JR 3rd. 1997. Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. Anal Chem **69**(4):767-76.

Brahamsha B. 1996. An abundance cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium *Synechococcus*. PNAS **93**:6504-6509.

Nagarajan A, Zhou MW, Nguyen AY, Liberton M, Kedia K, Shi TJ, Piehowski P, Shukla A, Fillmore TL, Nicora C, Smith RD, Koppenaal DW, Jacobs JM, Pakrasi HB. 2019. Proteomic Insights into Phycobilisome Degradation, A Selective and Tightly Controlled Process in The Fast-Growing Cyanobacterium *Synechococcus elongatus* UTEX 2973. Biomolecules **9**(8):374.

Harvey H, Bondy-Denomy J, Marquis H, Sztanko KM, Davidson AR, Burrows LL. 2018. *Pseudomonas aeruginosa* defends against phages through type IV pilus glycosylation. Nature Microbiology **3**(1):47-52.

Nodop A, Pietsch D, Hocker R, Becker A, Pistorius EK, Forchhammer K, Michel KP. 2008. Transcript profiling reveals new insights into the acclimation of the mesophilic fresh-water cyanobacterium Synechococcus elongatus PCC 7942 to iron starvation. Plant Physiol. **147**(2):747-763.

Michel K, Thole HH, Pistorius EK. 1996. IdiA, a 34 kDa protein in the cyanobacteria *Synechococcus* sp. strains PCC 6301 and PCC 7942, is required for growth under iron and manganese limitations. Microbiology **142**:2635-2645.

Michel K, Krüger F, Pühler A, Pistorius EK. 1999. Molecular characterization of *idiA* and adjacent genes in the cyanobacteria *Synechococcus* sp. strains PCC 6301 and PCC 7942. Microbiology **145**:1473-1484.

Webb EA, Moffett JW, Waterbury JB. 2001. Iron Stress in Open-Ocean Cyanobacteria (*Synechococcus*, *Trichodesmium*, and *Crocosphaera*): Identification of the IdiA Protein. Applied and Environmental Microbiology **67**(12):5444-5452.

Clerico EM, Ditty JL, Golden SS. 2007. Specialized Techniques for Site-Directed Mutagenesis in Cyanobacteria. Circadian Rhythms **362**:155-171.

Taton A, Unglaub F, Wright NE, Zeng WY, Paz-Yepes J, Brahamsha B, Palenik B, Peterson TC, Haerizadeh F, Golden SS, Golden JW. 2014. Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria. Nucleic Acids Research **42**(17):e136.

Chen Y, Taton A, Go M, London RE, Pieper LM, Golden SS, Golden JW. 2016. Self-replicating shuttle vectors based on pANS, a small endogenous plasmid of the unicellular cyanobacterium *Synechococcus elongatus* PCC 7942. Microbiol. **162**:2029-2041.