Characterization and cellular roles of a bacterial O-GlcNAc transferase in *Synechococcus elongatus* PCC7942

A Dissertation

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Kerry Ann Sokol

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Neil Olszewski, Advisor

May, 2015

DEDICATIONS

I dedicate this work to the family and friends who have patiently supported me, as well as to my advisor, Neil Olszewski, without whose guidance and insight I could not have completed this work. To Mitchell Johnson, thank you for your unending forbearance and support. To Annie Endrizzi, thank you for always having an open and patient ear. To my family and the Fallon family, for understanding the importance of this endeavor to me.

ABSTRACT

The post-translational addition of a single O-linked β-N-acetylglucosamine (O-GlcNAc) to serine or threonine residues is an important element in an ever-increasing array of metazoan cellular and regulatory processes. The enzyme responsible for this modification, O-GlcNAc transferase (OGT), is conserved among a wide variety of organisms and is critical for viability in many eukaryotes, including humans, mice, and *Drosophila*. Although OGTs with domain structure similar to eukaryotic OGTs are predicted for numerous bacteria species, their cellular roles remain unknown. I have identified an OGT in the cyanobacterium *Synechococcus elongatus* PCC 7942 that has active site homology and similar domain structure to eukaryotic OGTs. An OGT deletion mutant was created (Δ*ogt)*, which is viable and has no defect in growth rate, but has several phenotypes of interest. Without agitation, Δ*ogt* cells aggregate, and settle out of media. Compared to wild type cells, the Δ*ogt* cells also have higher free cellular phosphate levels, wider thylakoid lumen, and differential organization of electron dense inclusion bodies. These phenotypes are rescued by re-introduction of the wild type OGT, but are not fully rescued by OGTs with single amino acid substitutions corresponding to mutations that reduce eukaryotic OGT activity. *S. elongatus* OGT purified from *E. coli* hydrolyzes the donor sugar, UDP-GlcNAc, while mutant OGTs that do not fully rescue the deletion mutant have reduced or no activity. These results suggest that the eukaryotic-like OGTs of bacteria affect multiple processes.

Although the substrates for the SeOGT remain elusive, I have uncovered a relationship between the SeOGT and pili through a mutation that suppresses the Δ*ogt* settling phenotype. A single amino acid substitution (alanine 107 to aspartic acid) in the protein PilA suppresses two mutant phenotypes, settling and loss of pili. PilA is a subunit of type four pili (fimbriae), which are frequently found on the surfaces of gram-negative bacteria. Fimbriae are involved in virulence, DNA uptake, twitching motility, and adhesion. Although pilus assembly is a complex process, pili are primarily homopolymers of PilA, which is added or subtracted via ATPases to either extend or retract the pilus. No pili are detectable in Δ*ogt* cells, while pili are restored to Δ*ogt* cells by the $pilA^{(A107D)}$ mutation. When present together, mutant and wild type pilA genes prevent the assembly of pili, suggesting that there is a toxic interaction between mutant and wild type proteins which results in inability to assemble pili. This occurs regardless of the presence of OGT. Glycosylation of extra-cellular appendages have been widely reported. Pili are known substrates of glycosylation and glycosylation states have been reported to be important in the function of pili. The suppression of Δ*ogt* phenotypes by mutant *PilA* suggests that PilA is either a substrate of the OGT or that another protein involved in pilus synthesis is modified by OGT.

Table of Contents

List of Tables

Chapter One

Chapter Two

Appendix I

Appendix III

List of Figures

Chapter One

Chapter two

Chapter three

Chapter four

LIST OF ABBREVIATIONS

ANOVA – Analysis of Variance

BEMAB -- β-elimination Michael Addition of Biotin

CR — Congo Red Dye

ESI-CID-MS/MS -- electrospray ionization-collision induced dissociation tandem mass spectrometry

HsOGT – Human O-GlcNAc transferase

OD – Optical Density

OGA – O-GlcNAc Hydrolase

O-GlcNAc -- O-linked β-N-acetylglucosamine

OGT – O-GlcNAc transferase

PPBs – Polyphosphate bodies

PTM – Post-translational modification

SeOGT – *Synechococcus elongatus* PCC7942 O-GlcNAc transferase

SNP – Single nucleotide polymorphism

T4P – Type four pili

TEM – Transmission electron microscopy

TPRs – Tetratricopeptide Repeats

UDP-GlcNAc -- Uridine diphosphate *N*-acetylglucosamine

WT – Wild type

XcOGT – *Xanthomonas campestris* O-GlcNAc transferase

Chapter One

Introduction

Post-translational modification is the most common mechanism to regulate protein function and quality control. Many of these modifications such as phosphorylation, acetylation, and myristoylation have been studied for decades. A more recently discovered modification type, O-GlcNAcylation, involves the covalent attachment of a single β-N-acetylglucosamine to the hydroxyl group of a serine or threonine residue. The enzyme responsible for this modification is O-linked β-N-acetylglucosamine transferase (EC 2.4.1.255), also known as O-GlcNAc transferase (OGT). This modification has unique properties amongst most other glycosylations, in that it occurs in the nucleus and cytosol (as opposed to the secretory endomembrane system), and, similar to phosphorylation, is labile. The ever-growing number of cellular systems this modification is involved in has led to a quickly expanding body of research to further understand the mechanisms and roles of O-GlcNAcylation.

The O-GlcNAc modification

First discovered in 1984 by Torres and Hart on the surface of lymphocytes (1), it quickly became apparent that this modification is a critical component of many biological processes such as glucose signaling, transcription, and cell cycle progression. Formation of the O-GlcNAc modification consists of the transfer of a

single GlcNAc sugar from the UDP-GlcNAc donor to the hydroxyl of a serine or threonine residue, creating an O-linked GlcNAc (Fig. 1). Over 4000 proteins are currently known to be modified by O-GlcNAc, and new substrate proteins continue to be reported frequently (2). Diverse types of proteins carry the O-GlcNAc modification, including transcription factors, cytoskeletal proteins, hormone receptors, and chromatin-associated proteins. The OGT enzyme is found throughout many metazoan species, including *Drosophila*, *C. elegans*, mice, and humans (3).

O-GlcNAcylation has unique characteristics among post-translational modifications and glycosylation alike. Until recent decades, posttranslational modification (PTM) was thought to occur in the endoplasmic reticulum and Golgi apparatus, with glycosylated and proteins trafficked directly to their final location, such as membranes, vesicles, and the cell exterior. Cytoplasmic and nuclear proteins were not believed be glycosylated. O-GlcNAcylation broke both of these precepts. Not only is this post-translational modification made in the nucleus and cytosol, the newly modified proteins are found to function throughout nucleocytosolic regions (4). Many post-translational modifications such as prenylation or N-linked glycosylation are stable. For O-GlcNAcylation, however, this is not the case. O-GlcNAcylation, like phosphorylation, can occur in a dynamic manner through addition and removal of the glycan (4).

Unfortunately, this modification has several characteristics that make the investigation of its cellular roles difficult. Unlike many other PTMs, there is no

well-defined consensus sequence allowing the prediction of the location of O-GlcNAcylation. Because of the reversible nature of the modification, target substrates can be difficult to identify due to low levels of modification. Among the most vexing problems is the essentiality of the enzyme to most organisms. In mice and *Drosophila*, null mutants result in lethality (5, 6). *C. elegans* and zebrafish can survive without O-GlcNAcylation (7, 8). Currently the simplest model organism in which to study the modification is the multicellular early metazoan *Trichoplax adhaerens* (9), as yeast and *E. coli* do not have OGT homologs.

Cellular roles of O-GlcNAc modification

O-GlcNAcylation was first characterized on the exterior of lymphocyte cells, but soon was discovered to be a very abundant modification on rat liver organelles, particularly the nuclear envelope (1). Human erythrocytes were next to be reported to contain O-GlcNAcylated proteins, including cytoskeletal proteins that interact with actin (10, 11). Just a few years later, the polytene chromosomes of salivary glands in *Drosophila* were reported to be heavily decorated with O-GlcNAcylated proteins (12). O-GlcNAcylation was later studied in *C. elegans*, *Arabidopsis thaliana*, zebrafish, and other model organisms (7, 8, 13). Thousands of metazoan substrate proteins with the O-GlcNAc modification have been reported to date, reflecting the myriad roles of this modification (14, 15). Some of the most well-studied cellular roles affected by O-GlcNAcylation include

nutrient sensing in diabetes (16), cell cycle signaling in cancer (17), and the brain protein tau in neurodegenerative diseases (18).

Phosphorylation crosstalk

O-GlcNAcylation can act in competition with phosphorylation of the same sites or proximal sites, with one modification precluding the other. The crosstalk between O-GlcNAc and phosphorylation states implicates the O-GlcNAc modification's role in more complex cell signaling, via not simply the presence or absence of the O-GlcNAc moiety, but the interference or enhancement of phosphorylation signaling. Notably, unlike phosphorylation signaling, which employs a plethora of site-specific kinases, the cycling of O-GlcNAcylation relies on few highly conserved enzymes. O-GlcNAcylation commonly has only one enzyme for addition, O-GlcNAc transferase (OGT), and one enzyme for removal, O-GlcNAc hydrolase (OGA). In most species, a single, highly conserved gene coding for O-GlcNAc transferase, is present coding for the enzyme responsible for this modification, although some organisms have two OGTs (3). Congruently, a single gene, O-GlcNAc hydrolase (OGA), encodes the enzyme that allows for the dynamic nature of the modification by removing the sugar.

Enzyme structure and function

O-GlcNAcylation is effected by the enzyme O-linked β-N-acetylglucosamine transferase (EC 2.4.1.255), also known as O-GlcNAc transferase (OGT). In

general, O-GlcNAc transferases consist of a variable number of N-terminal tetratricopeptide repeats and two C-terminal catalytic domains with an intervening domain between the catalytic domains. The dual catalytic domains classify the enzyme as a GT41 GT-B glycosyltransferase (CAZY http://www.cazy.org/GT41.html). The TPR domain has been structurally characterized through x-ray crystallography, which shows that it folds into the predicted superhelical fold with a long groove (19). Although numerous known mutations identified critical amino acids and predicted a possible active site location (20), the true structure of an OGT remained elusive until 2008, when two separate groups published the first crystal structure of an OGT (21, 22). These publications both reported on the structure of an OGT homolog from the bacteria *Xanthomonas campestris* (XcOGT). Due to the hydrolytic activity of the enzyme against the sugar donor uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc), it was only possible to crystalize the OGT with either UDP or a UDP-GlcNAc analog to show enzyme-GlcNAc interaction. Finally, the elusive human OGT crystal structure was reported by Lazarus *et al.* (23) and it displayed significant structural homology to the XcOGT. Human OGT (HsOGT) was co-crystallized with UDP and, separately (but for the first time), with a peptide substrate. Most recently, Schimpl *et al.* (24) co-crystallized HsOGT with both a UDP and a UDP-GlcNAc analogue and several peptide substrates (see Table 1 for summary of crystalized OGT and substrates).

All of these crystal structures are essentially identical, having an active site with a deep binding pocket for the sugar donor, and a long crevice for the protein substrate. The two active sites fit tightly together with the TPR domain, with the superhelical groove of the TPRs extending into the active site formed by the two catalytic domains. These initial structures suggest a model wherein the UDP-GlcNAc binds deeply in the active site pocket, followed by a substrate protein that is guided via interactions with the TPRs into proximity with the sugar donor, thereby allowing modification to occur.

The association of the TPRs with protein substrates, suggested by the bacterial crystal structure, are supported by mutational analyses of HsOGT (25) . In humans, three splice isoforms of OGT with different TPR domain lengths are present, each showing differential substrate specificity, which also supports this model (26). A partial TPR may also provide a secure lid to the active site when UDP-GlcNAc is bound, conformationally changing to close the access to the active site and trap the donor (22). Mutations in this active site lid support this role, including the *Arabidopsis* OGT mutations *spy-12* (27) and the SeOGT mutant discussed in this work, G350D (Chapter Two) (28).

The initial structures proposed several key residues, including two histidines near the active site, and a lysine deep in the active site. The authors proposed the lysine K842 interacts with the sugar donor, which is also supported by mutational studies (22). The histidines H498 and H558 were both proposed to be the key catalytic residues for transfer of the sugar to the protein (21). These

residues have been shown via mutational studies to be important for OGT function. In 2012, the HsOGT was co-crystallized with native UDP-GlcNAc (not an analog) and several different peptides (24). These structures demonstrate that the H498 is too far from the acceptor hydroxyl to act as a catalytic base. H558, although critical for enzyme function, is proposed to be structurally inhibited from acting as the catalytic base (24). The authors also report an unusual “back-bent” formation of the sugar donor, and suggest that the alpha-phosphate on the donor substrate itself serves as the catalytic base.

Bacterial glycosylation

Although long believed to be solely a eukaryotic attribute, several glycosylation pathways are now evident in prokaryotes as well. Glycans added to serine or threonine residues via O-linkages have a wide array of structures. Similarly, glycans added to asparagine residues via N-linkages also share numerous synthetic pathways and structures. These pathways are currently being harnessed in the field of bacterial glycoengineering to create novel therapeutics, and show promise in diverse applications, including antibody variants for recruiting immune cells to target cancer and glycosylation of proteins to improve pharmacokinetic properties such as half-life and *in vivo* activity.

Although bacterial glycoproteins are a relatively new discovery, it has become apparent that prokaryotic glycosylation plays an important role. The first glycoproteins found in bacteria were found on the S-layer of the archeon

Halobacterium salinarum, as well as the S-layers of two *Clostridium* species (29). Since then it has been established that both O-linked and N-linked glycoproteins are ubiquitous and play important roles in the prokaryotic world. They have been implicated in functional roles from pathogenicity, antigenicity, enzymatic activity, and cell surface protein adhesion (see (30), and (31) for review). Eukaryotes use five different linkage types to decorate glycoproteins, N-linked glycans on asparagine or arginine, O-linked glycans on serine, threonine, hydroxyproline, and hydroxylysine, phosphate linked glycans on serine, C-linked glycans on tryptophan, and glypiation, which uses a glycosylphosphatidylinositol anchor and a phosphoethanolamine linker to glycosylate C-terminal ends of proteins. Bacteria use mostly N- and O-linkages, but have additional rare modifications, S-linked modification of cysteine and O-linked modification of tyrosine (31). Bacterial glycoproteins can be synthesized sequentially, with each additional sugar added to an existing N- or O-linked chain, or en bloc. The types of sugars included in bacterial glycosylation include glucose, galactose, mannose, fucose, rhamnose, arabinose, GalNAc (N-acetylgalactosamine) and GlcNAc (N-acetylglucosamine) as well as some sugars novel to bacteria, such as 2,4-diacetimido-2,4,6-trideoxyhexose.

N-linked bacterial glycosylation has been well documented in *Campylobacter jejuni* (31). The pathway starts with a lipid-linked oligosaccharide, which is synthesized in the cytoplasm and then translocated to the periplasm via a flippase. Protein glycosylation genes (*pgl)* are responsible for the glycoslylation

of flagella. Flagella are helical filaments on the cell surface that are involved in locomotion and virulence and mutations in the genes responsible for this action inhibit virulence, by affecting adherence, invasion, and colonization (32). Other glycosylation targets are surface proteins and periplasmic proteins.

The study of O-linked glycosylation in the structure and function of flagella and other cell surface proteins has been particularly well studied. In Gram-negative bacteria, a well-documented substrate for glycosylation is the flagellum. Among different species, glycosylation of flagellins varies in number between species from a single glycosylation site in *Listeria monocytogenes* (33), to up to 19 sites in *Burkholderia* spp. (34). Heterogeneity of types of flagellin glycosylation is wide, even among closely related species (see (35) for review), which may be useful in evolving surface recognition and thereby virulence against hosts. In *Listeria monocytogenes,* flagella are proposed to be modified with O-GlcNAc at multiple sites, although the function of this glycosylation remains unclear (33, 36). Bacterial cell surfaces contain other glycoproteins as well, playing roles in adhesion, secretion, signaling, and virulence, among others (see (30) for review).

O-GlcNAc Transferase in the prokaryotic world

Reports of O-GlcNAc transferase-mediated glycosylation in bacteria have been relatively rare, with no known functional roles to this point. In 2004, *Listeria monocytogenes* flagellins were investigated by mass spectrometry and reported to contain several O-linked HexNAc modifications (33). Immunostaining with the

CTD110.1 antibody supported the hypothesis that the flagellins were O-GlcNAcylated, however, the antibody has been reported to react non-specifically (37). The enzyme responsible for this modification however, remains unknown, as do the functional roles for the modification. Another flagellin-associated modification was reported to regulate flagellar motility in *Listeria* (36). This reported enzyme, GmaR interacts with a DNA-binding protein to regulate motility and is reported to modify the flagellin with O-GlcNAc. Intriguingly, GmaR shows no sequence homology to any other known O-GlcNAc transferase. This is striking, as all metazoan OGTs show some domain structure and/or active site sequence homology, suggesting that prokaryotes may have multiple pathways for the creation of the O-GlcNAc modification.

There are several suggested bacterial OGTs that are more similar to eukaryotic OGTs. The most studied is the OGT of *Xanthomonas campestris*, which has been crystallized and has remarkable similarity to the human OGT crystal structure. This enzyme has been reported to be an active OGT, although its cellular roles remain unknown (21). A cell surface protein in the cyanobacteria *Microcystis aeruginosa*, MrpC, which is involved with cell-to-cell contact, has been reported to be O-GlcNAcylated (38). These findings are based on the mass detected by mass spectrometry, however, the precursor ion is reported to be 160Da, whereas the O-GlcNAc moiety is 206Da. Interestingly, the mrpC gene lies in a cluster with two proposed glycosyltransferases, SlpA and SlpB. Both of these proteins share some sequence homology with eukaryotic OGTs, however,

they differ in their domain organization, with TPRs distributed throughout the proteins, not solely at the N-terminal. The O-GlcNAc modification is reported to be important for cell surface retention, however, the effects of blocking this modification are unknown, as is the specific role of the modification.

Due to the difficulty of studying O-GlcNAcylation in eukaryotes, a bacterial model would contribute a novel system for the further research into the cellular roles of O-GlcNAcylation. The photosynthetic cyanobacterium *Synechococcus elongatus* PCC7942 is predicted to have one O-GlcNAc transferase gene, with characteristic eukaryotic-like predicted protein structure. Sokol and Olszewski (28) propose that this is an active OGT similar to eukaryotic OGTs. (see Chapter Two) Phenotypes of null or mutant eukaryote-like OGTs have not yet been reported in prokaryotes and the biological roles of O-GlcNAcylated proteins in bacteria have not been investigated. The varied metabolic and regulatory pathways involving the O-GlcNAc modification in higher organisms suggest that this modification may be a regulatory element in prokaryotic systems as well.

Aims of this study

Because of the important roles that O-GlcNAcylation play in metazoans and the discovery of O-linked glycans in bacteria, this study aims to investigate the O-GlcNAc modification of bacterial proteins. The metazoan O-GlcNAc transferase (OGT) is responsible for the O-GlcNAc modification and bacteria share a homolog of the OGT gene, suggesting that O-GlcNAcylation occurs in

prokaryotes as well. This work seeks to examine the cellular roles of bacterial OGT by examining the phenotypes and activity of mutant OGT strains in the bacterium *Synechococcus elongatus* PCC7942. Additionally, I investigate the relationship between OGT and the pili structural protein PilA and the glycosylation of pili.

Fig 1. The O-GlcNAc reaction.

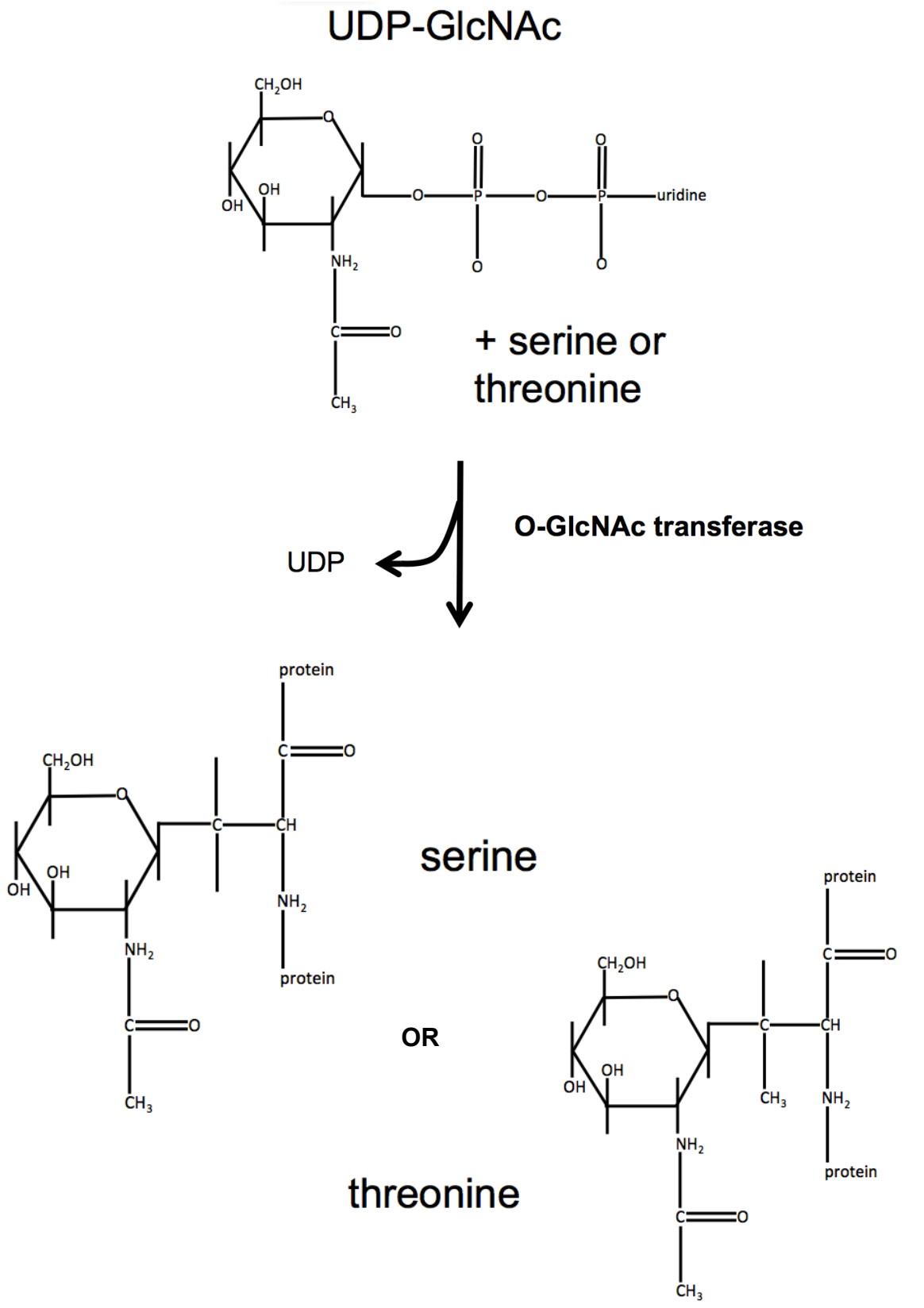

UDP-GlcNAc
+ serine or threonine
UDP
O-GlcNAc transferase
serine
OR
threonine
protein
uridine

Table 1. Table of crystal and co-crystal structures

Year	Group	Species	UDP	UDP-GlcNAc *analogue	Peptide	Reference
2008	van Aalten	Xc		*X		(22)
2008	Davies	Xc	X			(21)
2011	Hanover	Hs	X		X	(23)
2012	Hanover	Hs		*X	X	(39)
2012	van Aalten	Hs	X	*X	X	(24)

Chapter Two has been modified from the previously published work “The putative eukaryotic-like O-GlcNAc transferase of the cyanobacterium *Synechococcus elongatus* PCC7942 hydrolyzes UDP-GlcNAc and is involved in multiple cellular processes.” Sokol, KA, Olszewski, NE. 2015. Journal of Bacteriology, volume 197, issue 2, pages 354-361.

Chapter Two

The putative eukaryotic-like O-GlcNAc transferase of the cyanobacterium Synechococcus elongatus PCC7942 hydrolyzes UDP-GlcNAc and is involved in multiple cellular processes

INTRODUCTION

Although long believed to be solely a eukaryotic attribute, many glycosylation pathways are now evident in prokaryotes, as bacterial proteins can be modified with a variety of N-linked (40) and N-linked and O-linked glycans (35). Post-translational modification of serine or threonine residues with single O-linked β-N-acetylglucosamine (GlcNAc) by O-GlcNAc transferases (OGTs) is common in eukaryotes. All eukaryotic OGTs share a common domain structure consisting of tetratricopeptide repeats (TPRs) that are involved in protein-protein interactions (41) followed by the glycosyltransferase catalytic region (Fig. 1A). The catalytic region is composed of two highly conserved domains (3), which are linked by a variable length insertion. OGT homologs occur across prokaryotic phyla (3) and a number of these OGTs are “eukaryote-like” (Fig. 1B); this is best illustrated by the similarities in the crystal structures of the *Xanthomonas campestris* (21, 22) and the human OGTs (23, 24). In addition to sharing similar domain structures, key amino acids in and around the active site are conserved in the eukaryotic and bacterial enzymes (Fig. 2).

Unlike other eukaryotic glycosylations, O-GlcNAcylation occurs in the nucleus and cytosol and is dynamically added and removed from proteins. O-GlcNAcylation can regulate cellular processes directly or by acting in competition with phosphorylation (42, 43). Notably and unlike phosphorylation, which employs a plethora of site-specific kinases and phosphatases, the cycling of O-GlcNAcylation relies on two, highly conserved enzymes that add and remove the sugar. An O-GlcNAc hydrolase capable of removing O-GlcNAc modifications from proteins has also been identified in bacteria (44), suggesting that the O-GlcNAc modification may be present and dynamic in prokaryotes. Phenotypes of null or mutant eukaryote-like OGTs have not been reported in prokaryotes and the biological roles of O-GlcNAcylation in bacteria have not been investigated.

Diverse types of eukaryotic proteins, including transcription factors, cytoskeletal proteins, metabolic enzymes, hormone receptors, and chromatin are O-GlcNAc modified. The number and range of physiological systems affected by O-GlcNAcylation states are continuously growing (45). Although OGTs were identified over two decades ago, the regulation of glycan transfer and how this modification regulates cellular processes are not yet fully understood. Perhaps studying the OGT of a less complex organism will facilitate gaining insights into how it regulates cellular processes.

The photosynthetic cyanobacteria *Synechococcus elongatus* PCC7942 has one predicted O-GlcNAc transferase (SeOGT), with a characteristic eukaryotic-like structure. We describe multiple mutant phenotypes of an OGT

deletion strain. We also show homology with eukaryotic OGTs by demonstrating that SeOGT with mutations affecting amino acids important in eukaryotic OGTs do not fully rescue mutant phenotypes. Additionally, we demonstrate a decrease in hydrolysis of UDP-GlcNAc by the mutant SeOGTs.

RESULTS

S. elongatus has a eukaryotic-like OGT.

A Blast search using the catalytic region of the *Arabidopsis* OGT SPY identified a putative *S. elongatus* OGT gene. The SeOGT gene (Synpcc7942_0051) is predicted to encode a 633 amino acid protein (ABB56083) with a eukaryotic-like domain structure of six predicted N-terminal tetratricopeptide repeats followed by the C-terminal catalytic region (Fig. 1B). Protein alignment using standard parameters of Emboss 6.3.1: matcher (http://mobyle.pasteur.fr) shows that SeOGT has 52% similarity to the *Arabidopsis* OGT SPY and 49% similarity to human OGT (HsOGT). The active site domain shows 59% similarity to SPY and 49% similarity to HsOGT (Fig. 2). The full protein shows 45% similarity to the previously identified bacterial eukaryotic-like OGT from *Xanthomonas campestris pv campestris* .

***Phenotypes of an SeOGT deletion mutant*.**

We created a SeOGT deletion strain (∆*ogt*) by replacing the coding region with an antibiotic marker cassette. The ∆*ogt* mutant displayed no obvious growth defect under standard growth conditions. However, without agitation from

shaking or aeration, the mutant cells showed a very distinct phenotype of settling to the bottom of the culture container (Fig. 3A). When the optical density (OD_{730}) at a fixed height above the bottom of cuvettes was monitored, mutant cultures settled but wild type did not (Fig. 3B). The initial density of cultures affected the time to completion of settling, which varied from 24 hours for mid-exponential phase cultures, to a week for stationary cultures. Some settling occurred when wild type cultures were at high cell densities, possibly due to cell death or aggregation, whereas cultures at lower cell densities showed an initial increase in OD_{730}, due to culture growth.

To determine if mutant cells were viable post-settling, cultures were agitated and OD_{730} was measured (Fig. 3B). Both wild type and mutant cultures had similar densities at all time points indicating the mutant cells were viable and had the same growth rate as wild type. Further supporting this, the numbers of colony-forming units were similar after seven days of settling (Table 3). The buoyant densities of mutant and wild type cells were similar when assessed by isopycnic centrifugation on a Percoll gradient (Fig. 4), suggesting that a difference in buoyancy is not the cause of mutant cell settling. To determine if impaired metabolic function was involved in the settling phenotype, cells were treated with sodium azide, a biocide that acts by inhibiting electron transport, depleting the cells of ATP. Treatment accelerated the settling of mutant and wild type cells, and both settled at similar rates, which were greater than untreated

Δ*ogt* cells (Fig. 5). Therefore metabolism contributes to the ability of both mutant and wild type to remain in suspension.

Upon inspection under standard light microscopy, mutant cells had no obvious morphological phenotypes. However, when cells were observed on hemocytometer slides, a higher fraction of the mutant cells were aggregated (Figs. 3C and 3D). To assess possible changes in cell wall or glycocalyx that might lead to increased cell aggregation, cells were examined by TEM. The mutant cells exhibited no obvious defects in cell wall structure or thickness (Fig. 6A) but the thylakoid lumen of mutant cells was swollen (Figs. 6A, 6B). In addition, mutant cells had differential accumulation of electron-dense bodies between the thylakoid membranes (Fig. 7A), with more frequent, larger bodies deposited in mutant cells compared to wild type.

Because polyphosphate bodies (PPBs) are electron-dense, we determined the amounts of inorganic phosphate bound in PPBs by using the exopolyphosphatase ScPPX1 to hydrolyze polyphosphate and measuring the resultant increase in inorganic phosphate (46). For both wild type and mutant cells, relatively small amounts of phosphate were released from polyphosphate, and there was no statistically significant difference in the amounts released (Fig. 7B). Surprisingly, free cellular phosphate was nearly three-fold higher in mutant cells (Fig. 7B).

Transgene rescue of ∆ogt with wild type and mutant SeOGTs.

To determine if the observed phenotypes were attributable to deleting the SeOGT gene, we reintroduced the wild type SeOGT coding region under the control of its predicted promoter into the ∆*ogt* strain via homologous recombination. The resulting strain was fully rescued for the settling (Fig. 8A) and cell aggregation phenotypes (Fig. 3D), while the thylakoid width defect was partially rescued (Fig. 6B).

A previously described G570D mutation in *Arabidopsis*, *spy12*, results in plants with altered hormone responses (47). The G570 residue is conserved in many phyla (Fig. 2), corresponding to G631 and G350 in the human and *S. elongatus* enzymes, respectively. A mutant SeOGT carrying a G350D mutation did not rescue the settling phenotype of the ∆*ogt* strain (Fig. 8A).

Co-crystal structures for the HsOGT and the bacterial *X. campestris* OGT with various UDP-GlcNAc analogs and substrate peptides have recently been determined (22–24, 39). Based on these structures, several catalytically important residues have been proposed. Two histidines, HsOGT H558 and H498, have been predicted to be important in the transfer of GlcNAc to substrates (21–23). Additionally, a lysine residue, HsOGT K842, was shown to interact with the UDP portion of the sugar donor (22). Mutational studies have shown that each of these residues is critical for full enzymatic activity (21, 22, 24).

Although H498 is not conserved, H558 and K842 are conserved and correspond to H280 and K445 in SeOGT. Introduction of mutant *SeOGT* genes carrying the H280A and K445A mutations into Δ*ogt* did not rescue the settling phenotype (Fig. 8A), suggesting that, similar to human OGT, these residues also play an important role in SeOGT activity.

***SeOGT hydrolyzes UDP-GlcNAc*.**

To determine if SeOGT has enzymatic activity, we took advantage of the ability of some sugar transferases to hydrolyze the bond between UDP and the sugar moiety (23). Wild type SeOGT protein, and each of the three mutant SeOGTs were expressed in *E. coli* as MBP fusions and purified. Wild type SeOGT enzyme hydrolyzed more UDP-GlcNAc sugar than either the G350D or K445A enzymes (Fig. 8B). The G350D enzyme was more active than the negative controls (GFP-MBP fusion protein and heat-killed OGT enzyme), suggesting it has some residual activity, but the activity of the K445A enzyme was not different from the negative controls. The H280A enzyme hydrolyzed slightly less UDP-GlcNAc, but was not significantly different from wild type.

Non-informative experiments, over-expression and detection of substrates.

Several methods are currently used to identify O-GlcNAcylated proteins including transfer of labeled galactose to a known terminal O-GlcNAc, antibodies against the O-GlcNAc modification, and chemi-enzymatic methods (48). The GalT assay uses the enzyme β-1,4-galactosyl transferase, which adds a radioactively labeled galactose to proteins with glycosylations terminating in GlcNAc. The BEMAB

method involves the elimination of the β-O-linkage between glycosylated serines or threonines, and substitution at this site with biotin, which labels any O-linkage. Using these methods, several *S. elongatus* proteins were detected but wild type and mutant samples showed the same labeling patterns (data not shown). Several antibodies are currently available to detect the O-GlcNAc modification (48). Western blots using these antibodies detected several proteins but did not reveal differences between the mutant and wild type (data not shown). Since *S. elongatus* PCC7942 has no other predicted OGTs, the proteins detected are either modified by a novel OGT or the antibodies are recognizing other structures as has been previously reported (37). Overexpression of SeOGT did not result in any detected phenotypes. In view of the failure of these approaches to identify O-GlcNAcylation differences between wild type and Δ*ogt,* it is likely that glycoproteomic approaches employing mass spectrometry will be needed to identify the SeOGT substrates. Since prokaryotic substrates are not known, we attempted to modify a known eukaryotic OGT substrates using co-expressed SeOGT and Nup62 but did not detect modified substrates (data not shown), suggesting the enzyme either is not an OGT or that bacterial and eukaryotic OGTs have different substrate specificities.

DISCUSSION

In addition to the eukaryotic-like OGTs, there are bacterial OGTs which have a different domain structure or share little sequence homology with the eukaryotic OGTs (Fig. 1B)(38, 49). The *MrpC* gene of the cyanobacteria *Microcystis aeruginosa*, which encodes a highly abundant extracellular protein that is important in cell-to-cell contact and may be O-GlcNAcylated, lies in a gene cluster with two proposed glycolsyltransferases, *SlpA* and *SlpB* (38). Both glycolsyltransferases share sequence homology with eukaryotic OGTs, but only SlpA contains TPR domains (Fig. 1B). However, SlpA's TPRs have a different arrangement than eukaryotic OGTs. MrpC has two cellular variants: a secreted form and a form associated with the cell surface that is reported to be O-GlcNAc modified. While the O-GlcNAc modification is proposed to be important for cell surface retention, the effects of blocking the modification on protein localization have not been determined. Two bacterial flagellins that may be modified with O-GlcNAc have been reported (36, 49). *Listeria monocytogenes* flagellins were found by mass spectrometry to contain several O-linked HexNAc modifications and immunostaining suggested the modifications were O-GlcNAc (33). However, the enzyme responsible for this modification remains unknown. Shen *et al*. (36) reported a bacterial OGT, GmaR, that interacts with a DNA-binding protein to regulate flagellar motility and can modify the flagellin. However, GmaR has no sequence homology to the catalytic region of eukaryotic OGTs.

While the crystal structure of a eukaryotic-like OGT from *Xanthomonas campestris* has been determined (21, 22), the cellular role(s) of this enzyme are unknown. In *S. elongatus,* deletion of the SeOGT does not affect the growth rate but causes several distinct phenotypes, suggesting the OGT affects multiple processes. Mutant cells settle out of media and aggregate to a greater extent than wild type (Fig. 3). The two phenotypes may be interrelated in that a defect in motility could lead to aggregation or vice versa. Marine strains of *S. elongatus* have motility despite having no known locomotion organelles such as flagella (50, 51). Several membrane-related mechanisms for *S. elongatus* motility have been proposed, including localized expansion and contraction of the cell membrane (52), acoustic streaming (53), mechanochemical filaments near the inner membrane causing surface waves (54), as well as a helical rotor model (55). In a *Synechococcus* species related to *S. elongatus* PCC 7942 several genes involved in motility (56), including ABC transporters, glycosyltransferases, and the cell-surface polypeptide, SwmA, which forms a glycosylated S-layer (57) have been identified. Although the motility genes *SwmA* and *SwmB* are not present in *S. elongatus* PCC 7942, cell surface proteins are likely to have a role in motility. Bacterial glycoproteins residing on the cell surface are often important in both motility-related structures such as flagella, as well as mediating cell-to-cell contact (31) and therefore could be SeOGT substrates.

The Δ*ogt* settling phenotype is not due to a difference in buoyant density, as mutant and wild type cells had identical buoyant densities in a Percoll gradient

(Fig. 4). After treatment with sodium azide, which inhibits cytochrome oxidase and ATP synthesis, the wild type and mutant settling rates were similar (Fig. 5). Interestingly, settling of Δ*ogt* was increased by sodium azide indicating that Δ*ogt* cells are only partially impaired in the ability to remain in suspension.

TEM imaging reveals several macromolecular mutant phenotypes. Mutant cells show a widening of the thylakoid lumen (Fig. 6A, 6B) and a defect in the size and number of electron-dense bodies (Fig. 7A), which we hypothesized were polyphosphate bodies. However, we did not detect a change in the overall amount of phosphate in the form of polyphosphate suggesting the electron dense bodies are not polyphosphate bodies. Surprisingly, free inorganic phosphate was significantly higher in the Δ*ogt* strain (Fig. 7B), suggesting a defect in phosphate uptake or storage by the mutant. Supporting this hypothesis, when cells were transferred to phosphate poor media the settling rate of Δ*ogt* cells was unaffected, while that of wild type increased (Fig. 9). Supplementing the media with additional phosphate did not rescue the Δ*ogt* settling phenotype (data not shown). Since phosphate is a key limiting nutrient in many natural environments, regulation of sensing, uptake, and storage of phosphate in bacteria is tightly controlled by a process that involves phosphorylation of signaling proteins (58). Since O-GlcNAcylation and phosphorylation are reciprocal regulatory modifications in eukaryotes (42, 43), perhaps the interplay of O-GlcNAcylation and phosphorylation regulates phosphate status in bacteria.

Mutational studies, as well as crystal structure models of OGTs with reaction components have led to identification of several critical amino acids (Fig. 2). Although reintroduction of the native SeOGT gene rescued the settling phenotype, introduction of two genes with mutations predicted to impair catalytic activity did not fully rescue the phenotype (Fig. 8A). The H280A mutation, which affects a residue in or near the active site that possibly interacts with substrate proteins, prevented rescue. Likewise, the K445A mutation, which affects a residue predicted to stabilize the sugar donor, prevented rescue. In contrast, the G350D mutation affects the active site lid and cells expressing the mutant protein were weakly rescued. G350 is conserved in the Arabidopsis OGT SPY and is mutated to a D in the *spy-12* mutant. Interestingly, the *spy-12* mutation produces a weak phenotype suggesting that it causes only a partial loss of OGT activity (27).

Since HsOGT has been reported to hydrolyze UDP-GlcNAc (23), the ability of SeOGT to hydrolyze UDP-GlcNAc was assessed. SeOGT hydrolyzed UDP-[^{3}H]GlcNAc (Fig. 8B); furthermore, single amino acid mutations in the active site lid (G350D) and in the catalytic domain (K445A) significantly decreased this activity (Fig. 8B). These results support the homology of SeOGT with eukaryotic OGTs. Although the G350 residue resides in the active site lid, the reduction in activity may be due to instability in the binding pocket, or inhibition of access of the UDP-GlcNAc to the donor binding site. The reduction in activity for the K445 residue is consistent with its predicted role to stabilize the UDP of the donor

sugar moiety in the active site. The loss of hydrolysis in these mutant enzymes and the inability of the mutant strains to fully rescue mutant phenotypes suggest that catalytic activity is required for SeOGT function.

Interestingly, an additional active site mutant, H280A, did not have a significant difference in hydrolytic activity from wild type. In HsOGT this residue corresponds to H558, and is critical for modification of proteins . The residue was initially proposed to serve as a general base for catalysis , but was later found to be too distant from the acceptor hydroxyl to serve this role and therefore may not be directly involved in hydrolysis , which may explain why the mutant protein remains able to hydrolyze UDP-GlcNAc. The H280 strain did not rescue the null mutant settling phenotype (Fig. 8A), suggesting that although the mutant protein is able to hydrolyze UDP-GlcNAc it may be unable to transfer the sugar to the substrate protein.

The availability of a single celled system for the study of O-GlcNAcylation could accelerate research into the function of this protein modification. For this purpose, neither yeast nor *E. coli* are viable model organisms, as they lack OGT. The *S. elongatus* OGT has demonstrated mutant phenotypes suggesting multiple cellular roles, and can hydrolyze UDP-GlcNAc. Conserved amino acids important for OGT function in mammals and plants are also important in the SeOGT. Therefore, studies of the *S. elongatus* PCC 7942 OGT will potentially contribute to the understanding of both bacterial and eukaryotic OGTs.

MATERIALS AND METHODS

Bacterial Growth.

Cultures were grown in BG-11 media (59) with no antibiotics unless specified . Cultures were grown at 22°C under continuous fluorescent light (33.75 μmol photons $m^{-2}s^{-1}$) with aeration. To aerate, an aquarium air pump (Petco) was used to introduce bubbles to the bottom of culture flasks through a Pasteur pipette plugged with sterile cotton.

Generation of OGT deletion and rescue strains.

Constructs used to create deletion and rescue strains were made using MultiSite Gateway Three-Fragment vectors (Invitrogen) according to supplier's protocols (see Table 1 for a list of all vectors and Table 2 for all primer sequences). Briefly, 600bp of sequence directly upstream and 600bp directly downstream of the SeOGT gene (SynPCC7942_0051, GenBank ABB56083.1) were used to make pDONR SeOGT 5' and pDONR SeOGT 3', respectively, and a kanamycin resistance cassette was cloned into pDONR 221 to create pDONR Km^r. These constructs were used in a MultiSite Gateway reaction with pDEST R4-R3 vector to create pENTR Δ*ogt*::Km^r. This construct was transformed into *S. elongatus* PCC 7942 according to a previously established protocol (59). Homologous recombination resulted in replacement of the native gene with the antibiotic marker which was used to select for a deletion strain, Δ*ogt*::Km, noted as Δ*ogt* in figures. The genomic region was sequenced to confirm full deletion of the SeOGT.

The *Synechococcus* genome has two gene-free sites (neutral sites) where insertion of sequences does not cause any confounding downstream effects (59). For rescue experiments, wild type and mutant SeOGT genes under the control of the native promoter were introduced into neutral site two (NSII) of Δ*ogt*. The SoftBerry bacterial promoter prediction software BPROM (www.softberry.com) predicts that the SeOGT promoter is 293bp upstream of the start codon. Wild type SeOGT with 697bp directly upstream of the start codon was first fused to a streptomycin/spectinomycin cassette from pCDF Duet-1 (EMD-Millipore), in opposing orientation, using splice overlap extension PCR. This fusion was used to create pDONR pOGT:SeOGT-Spec^r. MultiSite Gateway cloning was used with the following vectors: pDONR pOGT:OGT-Spec^r, pDONR NSII 5', pDONR NSII 3', and pDEST R4-R3 and the resulting construct, pENTR pOGT:OGT-Spec^r, was transformed into Δ*ogt*. The resulting rescue strain Δ*ogt*::Km^r, NSII::pSeOGT:SeOGT-Spec^r *(*Δ*ogt*:OGT in Figures) was selected using spectinomycin. Transformed colonies were screened and genotyped to confirm the presence of the expected sequences. After confirmation of a stable genotype, cultures were no longer grown under drug selection.

Mutant SeOGTs rescue constructs were created by splice overlap extension PCR using primers introducing mutations that create single amino acid changes (Table 2). The G350D mutant was then fused to a Strep/Spec^r cassette in opposite orientation to the SeOGT gene as described above for the Δ*ogt*:OGT strain. For other constructs, the Strep-Spec cassette was fused to the

downstream recombination region NSII 3' in the pDONR NSII 3' vector, resulting in pDONR NSII 3'-Spec^r vector. MultiSite Gateway was used to create the constructs used to generate these mutant OGT strains. The mutant strains Δ*ogt*::Km^r, NSII::pSeOGT:SeOGT G350D, Δ*ogt*::Km^r, NSII::pSeOGT:SeOGT H280A, and Δ*ogt*::Km^r, NSII::pSeOGT:SeOGT K445A, are noted in the figures as G350D, H280A, K445A, respectively.

Culture Settling.

To determine rates of settling, 1 mL of cultures at similar initial OD_{730} were transferred into 1 mL spectrophotometric cuvettes. The cuvettes were covered with parafilm and allowed to sit under continuous fluorescent light (33.75 μmol photons $m^{-2}s^{-1}$) without agitation or aeration. OD_{730} was measured at 24-hour intervals. Culture growth was determined using paired replicate cultures. Each day one of the replicate cultures was resuspended and OD_{730} was measured. In addition, growth was also monitored by dilution plating of the resuspended cultures on solid BG-11 media to determine colony-forming units.

Cell Aggregation.

Cultures were grown as described to exponential phase with aeration. 50 µL of cultures were diluted into 500 µL BG11 media, 50 µL of this dilution was applied to the well of a hemocytometer with a cover slip, and allowed to sit at 22°C for 20 minutes. Cells were then counted as either single cells or cell clusters with the exception of cells with end-to-end contact, which could represent newly divided cells.

TEM Imaging.

For TEM imaging, exponential cultures fixed with 3% glutaraldehyde were prepared for Transmission Electron Microscopy (TEM) with osmium tetroxide staining according to the methods at Cyanosite (http://www-cyanosite.bio.purdue.edu/protocols/semithin.html) and sectioned. Cells were visualized using a Philips CM12 Transmission Electron Microscope and digitally photographed. Image J software (http://rsbweb.nih.gov/ij/) was used to measure the width of the lumen of cells that presented as round, mid-cell cross-sections.

Production of *Saccharomyces cerevisiae* exopolyphosphatase PPX1.

The exopolyphosphatase PPX1 from *Saccharomyces cerevisiae* was cloned into the expression vector pET11c (Invitrogen), expressed in *E.coli* cells (Rosetta, Novagen) and purified according to Wurst *et al.* (60). Enzymatic activity was confirmed using commercially synthesized polyphosphate (sodium phosphate glass type 45, Aldrich)**.**

Inorganic phosphate measurements.

S.elongatus cultures were grown to exponential phase and cells were pelleted by centrifugation for 15 minutes at 3700xg in a JA-20 rotor (Beckman). The cell pellets were frozen at -80°C, thawed, and re-suspended in French Press Buffer consisting of 20mM Tris-HCl (pH 7.5), 2mM EDTA, 1M sodium perchlorate, 1% SDS w/v (61) and cells were disrupted via two passes through a French Press. Cellular debris was cleared by centrifugation at 4°C for 15 minutes, at 20,000xg in a JA-20 rotor (Beckman) and aliquots of the supernatant were frozen at -20°C.

Free inorganic phosphate and phosphate in polyphosphate bodies were determined using a modified protocol from Werner *et al.* (46). 50 µL of cell lysate was added to 50 µL 1M sulfuric acid and vortexed briefly. The mixture was incubated for five minutes, then neutralized with 50 µL 2M sodium hydroxide and 100 µL Red Tris solution (1M Tris-HCl (pH 7.5), 6% Neutral Red solution (Ricca Chemical Company, Arlington TX), vortexed, then centrifuged to remove precipitated proteins. pH was assessed via color test and adjusted to neutral (acidic solutions are pink, basic solutions are yellow and neutral pH solutions are pinkish-orange). 50 µL aliquots of this neutralized solution were transferred to two separate microcentrifuge tubes. One aliquot was used to determine free inorganic phosphate and the other was used to determine total inorganic phosphate (free plus polyphosphate). Polyphosphate in one sample was hydrolyzed at 37°C following the addition of 142 µL of reaction buffer (20mM Tris-HCl (pH 7.5), 5mM magnesium acetate, 5mM ammonium acetate) and 8 µL PPX1. A pilot experiment indicated the hydrolysis was complete after 60 minutes so samples were incubated for 60 minutes. In some cases, additional polyphosphate was added and the reaction was extended for an additional 60 minutes to confirm that the reaction was not enzyme-limited. The aliquot used to measure free inorganic phosphate was treated similarly to the total phosphate sample except that PPX1 was omitted.

To measure phosphate in these samples, 344 µL of 28mM ammonium heptamolybdate in 2.1M sulfuric acid was added to each tube and tubes were

incubated for 20 minutes at 22°C. To develop color, 256 µL of 0.76mM malachite green oxalate in 0.35% polyvinyl alcohol was added to each sample and incubated for 20 minutes at 22°C. Absorbance at 595 nm was then measured. A standard curve using fresh potassium phosphate solutions was created for each experiment. The amount of inorganic phosphate as polyphosphate was calculated by subtracting the amount of free inorganic phosphate from the amount of total phosphate.

Expression and purification of SeOGT.

Wild type and mutant SeOGT were expressed as MBP-fusion proteins using pRARE-MBP-DEST in *E. coli* BL21(AI) cells. Proteins were purified on amylose columns according to the manufacturers recommendations (NEB), with the exception of removal of salt in the final wash, and flash frozen at -80° C in 20% glycerol. Coomassie-stained SDS PAGE gels indicated that approximately 90% of the protein in each preparation was full length SeOGT (data not shown).

Enzyme catalyzed hydrolysis of UDP-GlcNAc.

Purified SeOGT-MBP enzyme was tested for the ability to hydrolyze radiolabeled UDP-[^{3}H]GlcNAc in an assay modified from Support Protocol 4 in Zachara *et al.* (48). A UDP-[^{3}H]GlcNAc solution was made by drying 2 µCi of 2 mCi/ mL UDP-[^{3}H]GlcNAc (Specific Activity 20-45 Ci, 0.74-1.66 TBq/mmol, Perkin Elmer) under argon and resuspending it in 2 µL of 25mM adenosine mono-phosphate and 48 µL of 50mM Tris (pH 6.8), 5mM $MnCl_2$, 0.5uM UDP-GlcNAc. Each 50 µL reaction included 5 µL of the UDP-[^{3}H]GlcNAc solution (0.2 µCi per reaction), 6 µg

purified enzyme, 50mM Tris-HCl (pH 7.4), 1mM DTT, 12.5mM $MgCl_2$. The reaction incubated at 30°C and was stopped with the addition of 500 µL H_2O. In all experiments, reactions were run for 2 hours because a preliminary time course experiment indicated that reactions with wild type enzyme were 80% complete at this time (determined by time course). After termination, the reaction was immediately loaded onto Dowex AG1 resin (PO_4 form) columns, which retains UDP-GlcNAc. The column was washed with 3 mL H_2O, and radioactivity in the wash was measured using a Packard Tri-Carb 2900TR liquid scintillation counter.

Preparation of protein samples for antibody probing, GalT, and BEMAB

20 mL cultures were collected and gently pelleted. Pellets were stored overnight at -20C. Cells were resuspended in 1 mL cold bead beating buffer (20 mM Tris-HCl (pH 8.0), 5 mM EDTA). 10 µL of PMSF and 10 µL of PUGNAc were added, with a small addition of glass beads. Samples were flash-frozen in liquid nitrogen, thawed in a 37°C water bath, then vortexed for 30 seconds, four times, with a 30 second rest in ice in between each vortexing. These steps were repeated three times. Samples were spun at 4K rpm in Beckman JA18.1 rotor for 1 minute. 750 mL supernatant was collected, centrifuged at 10K for 15 minutes. Resulting supernatant was saved as soluble protein sample. Pellet from 1st spin was resuspended in 1 ml cold bead beating buffer, with 50 µL 25% Triton. Samples were vortexed every 5 minutes for 30 minutes with rest on ice in between vortexing. Samples were centrifuged at 10K for 15 minutes, and 200 µL

supernatant was collected as thylakoid membrane sample. The remaining pellet was used as insoluble sample. All sample fractions were boiled 5 minutes in SDS-PAGE loading buffer and stored at -20°C. Some samples were frozen overnight at -80°C and run through French Pressure Cell two times in either bead beating buffer or French Press Buffer (61). Soluble and insoluble samples were collected.

Antibody probing.

Samples were run on 12% polyacrylamide gels and Coomassie stained, or run on gels and transferred to PVDF membrane for antibody probing.

Probing with anti-GlcNAc CTD110.6 was performed as described in Zachara (48). Probing with additional antibodies was performed as described in Teo *et al*. (62).

Galactosyl transferase assay (GalT).

Samples were run on SDS-PAGE gels and transferred to PVDF membrane. Membrane was blocked overnight at 4°C in 3% bovine serum albumin in 10 mM HEPES (pH 7.9). Membrane was washed in 10 mM HEPES (pH 7.9) for 10 minutes with gentle shaking. Three µCi of UDP-H^3-Galatose were dried under argon and resuspended in 3 µL 25 mM adenosine monophosphate and added to 800 µL labeling buffer (10 mM HEPES (pH 7.9), 1M $MnCl_2$, 1% BSA, protease inhibitor, and 100m units calf intestinal phosphatase. The membrane was sealed in plastic and incubated for 3 hours at 37°C. The membrane was washed 7 times for 10 minutes each in 10 mM HEPES (pH 7.9). The membrane was dried and

sprayed with EN^{3}HANCE (PerkinElmer) and dried for 20 minutes. The membrane was placed on pre-flashed film, then sealed and placed at -80°C for 72 hours.

BEMAB.

Samples were run on 15% SDS-PAGE gels and transferred to PVDF membrane. In a technique adapted from Wells *et al.* (63), the blot was incubated in 5 mL of 1% triethylamine, 0.1% NaOH, 3 mM 50 mM 5-(biotinamido)pentylamine. The blot was then incubated at 37 C for 24 hours with gentle agitation. The blot was washing with ddH20, blocked in PBS with 0.05% Tween and 3% BSA, and probed with Streptavidin-HRP, which was detected by chemiluminescence.

Overexpression of SeOGT.

GeneArt Synechococcus TOPO kit (LifeTechnologies) was used to create an inducible overexpression line. (Primers SeOGT pSyn F caccatgagtactcaggacaacag

SeOGT pSyn R tcaaaatctaacttctttattggcgctc). PCR and transformation of *E. coli* and *Synechococcus* performed as recommended by kit. The overexpression cassette was transformed into the Neutral Site I of the Δ*ogt*. This nickel-induced promoter is weakly constitutive and was found to be active enough to rescue mutant phenotypes without induction. However, to determine overexpression effects, 5 μM $NiCl_2$ was added as per kit instructions to growing cultures, and cultures were examined for phenotypes.

Coexpression of SeOGT with Nup62.

SeOGT-MBP was purified as previously explained. 20 µg SeOGT-MBP and 20 µg purified E.coli-expressed Nup62-His were added to 30 µL reaction buffer (50 mM Tris (pH 7.4), 1 mM DTT, 12.5 mM $MgCl_2$, 0.5 mM Tris (pH 6.8), 5 mM $MnCl_2$, UDP-GlcNAc). The mix was incubated at 30°C for 1 hour, then run on an SDS-PAGE gel and stained to look for band shifts of heavily O-GlcNAcylated Nup62.

Table 1. Plasmids and strains used and created in this study.

Plasmids used in this study	Description of use
pDONR221	DONR vector for Three Way Gateway Cloning
pDONR P4 P1	DONR vector for Three Way Gateway Cloning
pDONR P2 P3	DONR vector for Three Way Gateway Cloning
pDEST R3 R4	DEST vector for Three Way Gateway Cloning
pDONR 221	Source of Km^r gene
pCDF_Duet	Source of Sp^r gene
pRARE-MBP-DEST	Used for protein expression

Plasmids created in this study	Source Plasmid	Description
pDONR Km^r	pDONR221	Kanamycin resistance cassette
pDONR SeOGT 5'	pDONR	Upstream region of SeOGT for deletion
pDONR SeOGT 3'	pDONR	Downstream region of SeOGT for deletion
pENTR Δogt::Km^r	pDEST R3 R4	Final construct for SeOGT deletion transformation
pDONR NSII 5'	pDONR P4 P1	Upstream region of NSII for introduction of genes
pDONR NSII 3'	pDONR P2 P3	Downstream region of NSII for introduction of genes
pDONR NSII 3'-$Spec^r$	pDONR P2 P3	Downstream region of NSII for introduction of genes, fused to Spectinomycin resistance cassette
pDONR pOGT:SeOGT-$Spec^r$	pDONR 221	SeOGT and native promoter, fused to Spectinomycin resistance cassette

pENTR pOGT:SeOGT-Spec^r	pDEST R3 R4	Final construct for Δogt:SeOGT rescue strain
pDONR SeOGT G350D	pDONR 221	Mutant constructs for Δogt rescue
pENTR SeOGT G350D	pDEST R3 R4	"
pDONR SeOGT H280A	pDONR 221	"
pENTR SeOGT H280A	pDEST R3 R4	"
pDONR SeOGT K445A	pDONR 221	"
pENTR SeOGT K445A	pDEST R3 R4	"
pRARE SeOGT	pRARE	Protein expression of SeOGT in E. coli
pRARE SeOGT G350D	pRARE	Protein expression of SeOGT G350D in E. coli
pRARE SeOGT H280A	pRARE	Protein expression of SeOGT H280A in E. coli
pRARE SeOGT K445A	pRARE	Protein expression of SeOGT K445A in E. coli

Table 2. Primers used in this study

Name	Sequence	Purpose of primer pair
Kan-R replacement B1F	GGGGACAAGTTTGTACAAAA AAGCAGGCTTCCGTGTCTCA AAATCTCTGATGT	Replacing SeOGT with KanR cassette
Kan-R replacement B2R	GGGGACCACTTTGTACAAGA AAGCTGGGTAATGAGCTTGC GCCGTCCCGTCA	
SeOGT 5' B4F	GGGGACAACTTTGTATAGAA AAGTTGCCTCATGGGTCGAC ATCATTAAAT	5' upstream region around SeOGT to allow for homologous recombination
SeOGT 5' B1R	GGGGACTGCTTTTTTGTACA AACTTGCCATAGATTCATTTA ACTGCAAC	
SeOGT 3' B2F	GGGGACAGCTTTCTTGTACA AAGTGGCCGATTTTGATATG ATGAGCAATC	3' downstream region around SeOGT to allow for homologous recombination
SeOGT 3' B3R	GGGGACAACTTTGTATAATA AAGTTGCAAACACTTCCAAC AACAGCACC	
Synecho NSII 5’ B4F	GGGGACAACTTTGTATAGAA AAGTTGCCACCCCACGATTC TGAGCTTA	5' upstream region around neutral site II to allow for homologous recombination to create transgenes
Synecho NSII 5’ B1R	GGGGACTGCTTTTTTGTACA	

	AACTTGCGCTTCAAGATTCG GCAGTGA	
Synecho NSII 3’ B2F	GGGGACAGCTTTCTTGTACA AAGTGGCCAACCGTTCCTGC GCGATCGCT	3' downstream region around neutral site II to allow for homologous recombination to create transgenes
Synecho NSII 3’ B3R	GGGGACAACTTTGTATAATA AAGTTGCGCGCCCAGTGGA AACAACTGT	
Synecho pOGT:OGT B1F	GGGGACAAGTTTGTACAAAA AAGCAGGCTTCTCAAAATCT AACTTCTTTAT	SeOGT for creating transgene rescue into neutral site II
Synecho OGT B2R	TCCATTTCATGAATAAGTTGA ATAATTACCAACAGATTAG	
SpecR F	CAACTTATTCATGAAATGGA CTGTCAGATTCATGGTTGAG CC	Spectinomycin resistance cassette to fuse to NSII 3' region to confer resistance to inserted genes
SpecR B2R	GGGGACCACTTTGTACAAGA AAGCTGGGTATCACACTGCT TCCGGTAGTC	
SpecR B2F	GGGGACAGCTTTCTTGTACA AAGTGGCCATTTGTTTATTTT TCTAAATAC	Spectinomycin resistance cassette to fuse to NSII 3' region to confer resistance to inserted genes
SpecR Overlap R	CGCGCAGGAACGGTTTAGT	

	GTGTTATTTGCCGACTACC	
NSII 3’ Overlap F	CACACTAAACCGTTCCTGCG CGATCGCTCTTACTG	NSII 3' region to fuse with Spectinomycin resistance cassette
NSII 3’ B3R	GGGGACAACTTTGTATAATA AAGTTGCCGCGCCCAGTGG AAAC	
SeOGT pRare F	GGGGACAAGTTTGTACAAAA AAGCAGGCTTCATGCAGTTA AATGAATCTATGAGT	Protein expression of SeOGT in E.coli
SeOGT pRare R	GGGGACCACTTTGTACAAGA AAGCTGGGTTCAAAATCTAA CTTCTTTATTGGC	
SeOGT G350D Overlap F	CTTATTGAGCTTTCTGACCAT ACAAATGGTAAC	Overlap PCR to create G350D mutant
SeOGT G350D OverlapR	GTTAACATTTGTATGGTCAG AAAGCTCAATAAG	GGA (Gly) to GAC (Asp)
SeOGT H280A Overlap F	GAAATAACTTACGGAGGCTG TTTTAAAATCACCC	Overlap PCR to create H280A mutant
SeOGT H280A Overlap R	GGGGTGATTTTAAAACAGCC TCCGTAAGTTATTTC	CAT (His) to GCG (Ala)
SeOGT K445A Overlap F	GAATAGTCACTAATTGCTCG ACTTGCAACG	Overlap PCR to create K445A mutant
SeOGT K445A Overlap R	CGTTGCAAGTCGAGCAATTA GTGACTATTC	AAG (Lys) to GCA (Ala)

Table 3. Colony forming units of wild type and null mutant.

	Colonies grown from Dil $1x10^6$	Colonies grown from Dil $1x10^5$	CFUs/mL
WT	27	204	$2.37x10^7$
∆ogt	23	224	$2.27x10^7$

Wild type and ∆*ogt* mutant cultures were grown to mid-exponential phase under aeration, then allowed to sit with no aeration or shaking for 5 days. Cultures were shaken to resuspend settled cells and plated in dilution series. After two weeks, colony forming units (CFUs) were counted. Mutant and wild type cultures share a similar growth rate, despite mutant cell settling.

Figure 1. Domain organization of eukaryotic-like OGTs. (A) Structure of metazoan OGTs, showing N-terminal TPR repeats and two catalytic regions divided by a variable linker region. (B) Structure of putative prokaryotic OGTs: *Synechococcus elongatus* PCC7942, *Xanthomonas campestris campestris*, *Geobacter sp. M21*, *Thiorhodovibrio sp. 970*, *Bradyrhizobium japonicum*, *Microcystis aeruginosa*. Gray shading indicates homology to the eukaryotic catalytic domain. Asterisk indicates enzymes with non-eukaryotic domain structures. The catalytic domain of the *Bradyrhizobium* protein and M. *aeruginosa* SlpA and SlpB show homology to the Arabidopsis OGT SEC, but have an uncharacteristic TPR distribution. TPRs were predicted using TPRpred (http://toolkit.tuebingen.mpg.de/tprpred).

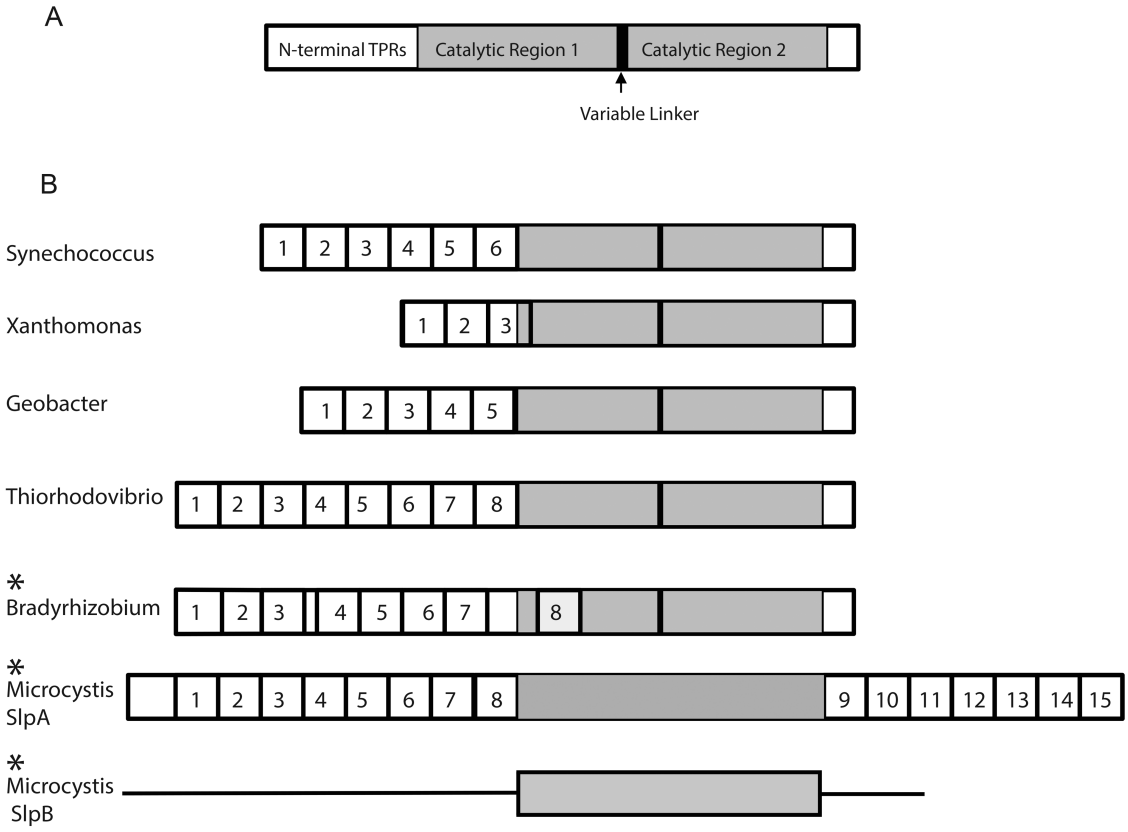

A
N-terminal TPRs
Catalytic Region 1
Catalytic Region 2
Variable Linker
B
Synechococcus
1 2 3 4 5 6
Xanthomonas
1 2 3
Geobacter
1 2 3 4 5
Thiorhodovibrio
1 2 3 4 5 6 7 8
*
Bradyrhizobium
1 2 3 4 5 6 7 8
*
Microcystis
SlpA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
*
Microcystis
SlpB

Figure 2. Alignment of OGT active sites.

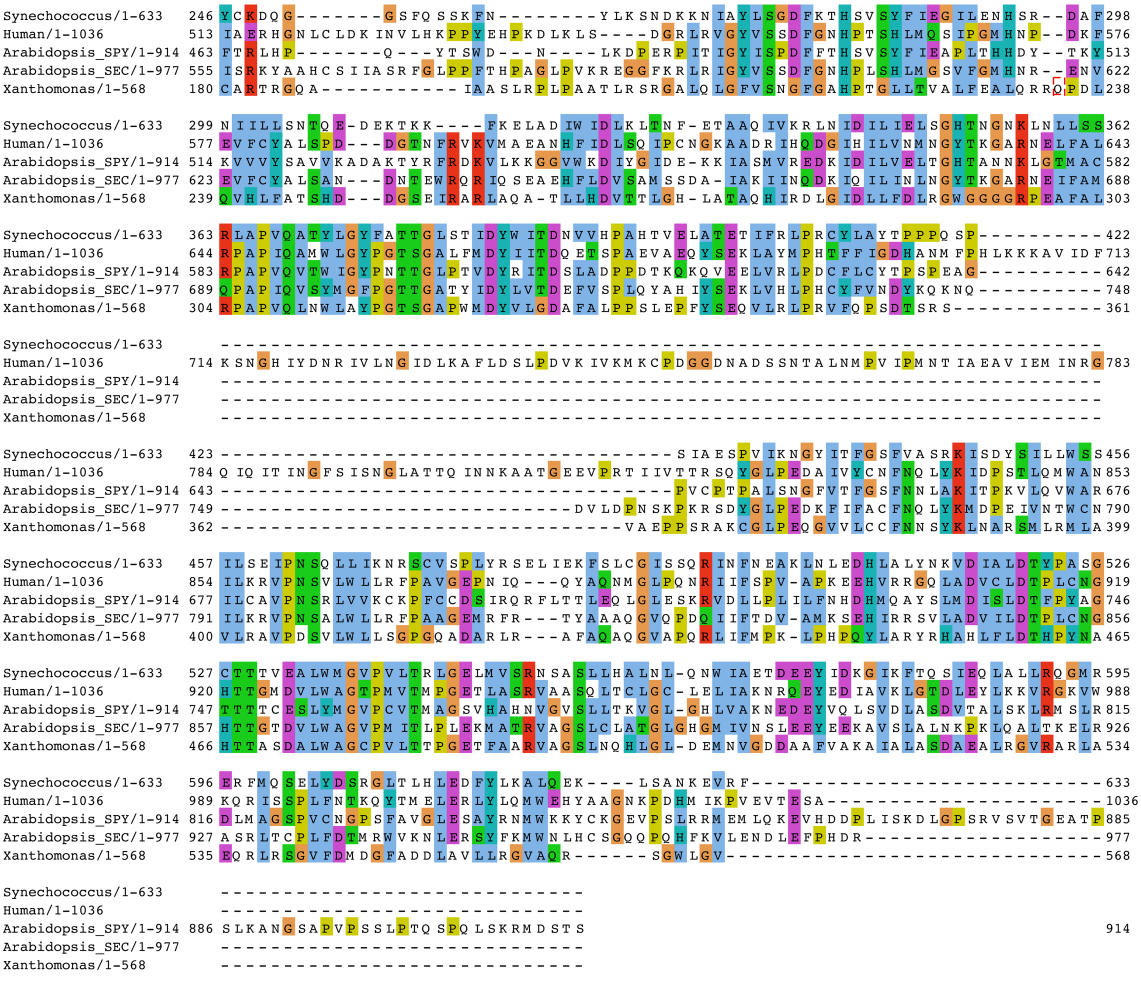

Synechococcus/1-633 246 YCKDQG-------GSFQSSKFN--------YLKSNDKKNIAYLSGDFKTHSVSYFIEGILENHSR--DAF 298
Human/1-1036 513 IAERHGNLCLDKINVLHKPPYEHPKDLKLS----DGRLRVGYVSSDFGNHPTSHLMQSIPGMHNP--DKF 576
Arabidopsis_SPY/1-914 463 FTRLHP-------Q---YTSWD---N-----LKDPERPITIGYISPDFFTHSVSYFIEAPLTHHDY--TKY 513
Arabidopsis_SEC/1-977 555 ISRKYAAHCSIIASRFGLPPFTHPAGLPVKREGGFKRLRIGYVSSDFGNHPLSHLMGSVFGMHNR--ENV 622
Xanthomonas/1-568 180 CARTRGQA----------IAASLRPLPAATLRSRGALQLGFVSNGFGAHPTGLLTVALFEALQRRQPDL 238
Synechococcus/1-633 299 NIILLSNTQE-DEKTKK----FKELADIWIDLKLTNF-ETAAQIVKRLNIDILIELSGHTNGNKLNLLSS 362
Human/1-1036 577 EVFCYALSPD---DGTNFRVKVMAEANHFIDLSQIPCNGKAADRIHQDGIHILVNMNGYTKGARNELFAL 643
Arabidopsis_SPY/1-914 514 KVVVYSAVVKADAKTYRFRDKVLKKGGVWKDIYGIDE-KKIASMVREDKIDILVELTGHTANNKLGTMAC 582
Arabidopsis_SEC/1-977 623 EVFCYALSAN---DNTEWRQRIQSEAEHFLDVSAMSSDA-IAKIINQLKIQILINLNGYTKGARNEIFAM 688
Xanthomonas/1-568 239 QVHLFATSHD---DGREIRARLAQA-TLLHDVTTLGH-LATAQHIRDLGIDLLFDLRGWGGGGRPEAFAL 303
Synechococcus/1-633 363 RLAPVQATYLGYFATTGLSTIDYWITDNVVHPAHTVELATETIFRLPRCYLAYTPPPQSP---------- 422
Human/1-1036 644 RPAPIQAMWLGYPGTSGALFMDYIITDQETSPAEVAEQYSEKLAYMPHTFFIGDHANMFPHLKKKAVIDF 713
Arabidopsis_SPY/1-914 583 RPAPVQVTWIGYPNTTGLPTVDYRITDSLADPPDTKQKQVEELVRLPDCFLCYTPSPEAG---------- 642
Arabidopsis_SEC/1-977 689 QPAPIQVSYMGFPGTTGATYIDYLVTDEFVSPLQYAHIYSEKLVHLPHCYFVNDYKQKNQ---------- 748
Xanthomonas/1-568 304 RPAPVQLNWLAYPGTSGAPWMDYVLGDAFALPPSLEPFYSEQVLRLPRVFQPSDTSRS---------- 361
Synechococcus/1-633 ----------
Human/1-1036 714 KSNGHIYDNRIVLNGIDLKAFLDSLPDVKIVKMKCPDGGDNADSSNTALNMPVIPMNTIAEAVIEMINRG 783
Arabidopsis_SPY/1-914 ----------
Arabidopsis_SEC/1-977 ----------
Xanthomonas/1-568 ----------
Synechococcus/1-633 423 ----------SIAESPVIKNGYITFGSFVASRKISDYSILLWSS 456
Human/1-1036 784 QIQITINGFSISNGLATTQINNKAATGEEVPRTIIVTTRSQYGLPEDAIVYCNFNQLYKIDPSTLQMWAN 853
Arabidopsis_SPY/1-914 643 ----------PVCPTPALSNGFVTFGSFNNLAKITPKVLQVWAR 676
Arabidopsis_SEC/1-977 749 ----------DVLDPNSKPKRSDYGLPEDKFIFACFNQLYKMDPEIVNTWCN 790
Xanthomonas/1-568 362 ----------VAEPPSRAKCGLPEQGVVLCCFNNSYKLNARSMLRMLA 399
Synechococcus/1-633 457 ILSEIPNSQLLIKNRGCVSPLYRSELIEKFSLCGISSQRINFNEAKLNLEDHLALYNKVDIALDTYPASG 526
Human/1-1036 854 ILKRVPNSVLWLLRFPAVGEPNIQ---QYAQNMGLPQNRIIFSPV-APKEEHVRRGQLADVCLDTPLCNG 919
Arabidopsis_SPY/1-914 677 ILCAVPNSRLVVKCKPFCCDSIRQRFLTTLEQLGLESKRVDLLPLILFNHDHMQAYSLMDISLDTFPYAG 746
Arabidopsis_SEC/1-977 791 ILKRVPNSALWLLRFPAAGEMRFR---TYAAAQGVQPDQIIFTDV-AMKSEHIRRSVLADVILDTPLCNG 856
Xanthomonas/1-568 400 VLRAVPDSVLWLLSGPGQADARLR---AFAQAQGVAPQRLIFMPK-LPHPQYLARYRHAHLFLDTHPYNA 465
Synechococcus/1-633 527 CTTTVEALWMGVPVLTRLGELMVSRNSASLLHALNL-QNWIAETDEEYIDKGIKFTQSIEQLALLRQGMR 595
Human/1-1036 920 HTTGMDVLWAGTPMVTMPGETLASRVAASQLTCLGC-LELIAKNRQEYEDIAVKLGTDLEYLKKVRGKVW 988
Arabidopsis_SPY/1-914 747 TTTTCESLYMGVPCVTMAGSVHAHNVGVSLLTKVGL-GHLVAKNEDEYVQLSVDLASDVTALSKLRM SLR 815
Arabidopsis_SEC/1-977 857 HTTGTDVLWAGVPMITLPLEKMATRVAGSLCLATGLGHGMIVNSLEEYEEKAVSLALNKPKLQALTKELR 926
Xanthomonas/1-568 466 HTTASDALWAGCPVLTTPGETFAARVAGSLNQHLGL-DEMNVGDDAAFVAKAIALASDAEALRGVRARLA 534
Synechococcus/1-633 596 ERFMQSELYDSRGLTLHLEDFYLKALQEK----LSANKEVRF---------- 633
Human/1-1036 989 KQRISSPLFNTKQYTMELERLYLQMWEHYAAGNKPDHMIKPVEVTESA---------- 1036
Arabidopsis_SPY/1-914 816 DLMAGSPVCNGPSFAVGLESAYRNMWKKYCKGEVPSLRRMEMLQKEVHDDPLISKDLGPSRVSVTGEATP 885
Arabidopsis_SEC/1-977 927 ASRLTCPLFDTMRWVKNLERSYFKMWNLHCSGQQPQHFKVLENDLEFFPHDR---------- 977
Xanthomonas/1-568 535 EQRLRSGVFDMDGFADDLAVLLRGVAQR------SGWLGV---------- 568
Synechococcus/1-633 ----------
Human/1-1036 ----------
Arabidopsis_SPY/1-914 886 SLKANGSAPVPSSLPTQSPQLSKRMDSTS 914
Arabidopsis_SEC/1-977 ----------
Xanthomonas/1-568 ----------

Figure 3. Mutant cells settle and aggregate. (A) Early exponential phase cultures were transferred to a cuvette and allowed to sit without agitation for 0 or 4 days before photographing. (B) Early exponential phase cultures were transferred to a cuvette and allowed to sit without agitation and optical density (OD_{730}) was measured daily (bars). To monitor growth rate, cells in duplicate cuvettes were resuspended before measuring the OD_{730} (lines). (C) Cells were placed on a hemocytometer slide for 20 minutes and photographed at 100x. The arrows indicate aggregated cells. The scale bar = 2 μM**.** (D) The fraction of cells (WT, n=1934; ∆ogt, n=1964; ∆ogt::OGT, n=1968) in aggregates of two or more was determined. In panels B and D, the data represent the results from three complete biological replicates with the error bars indicating the standard error.

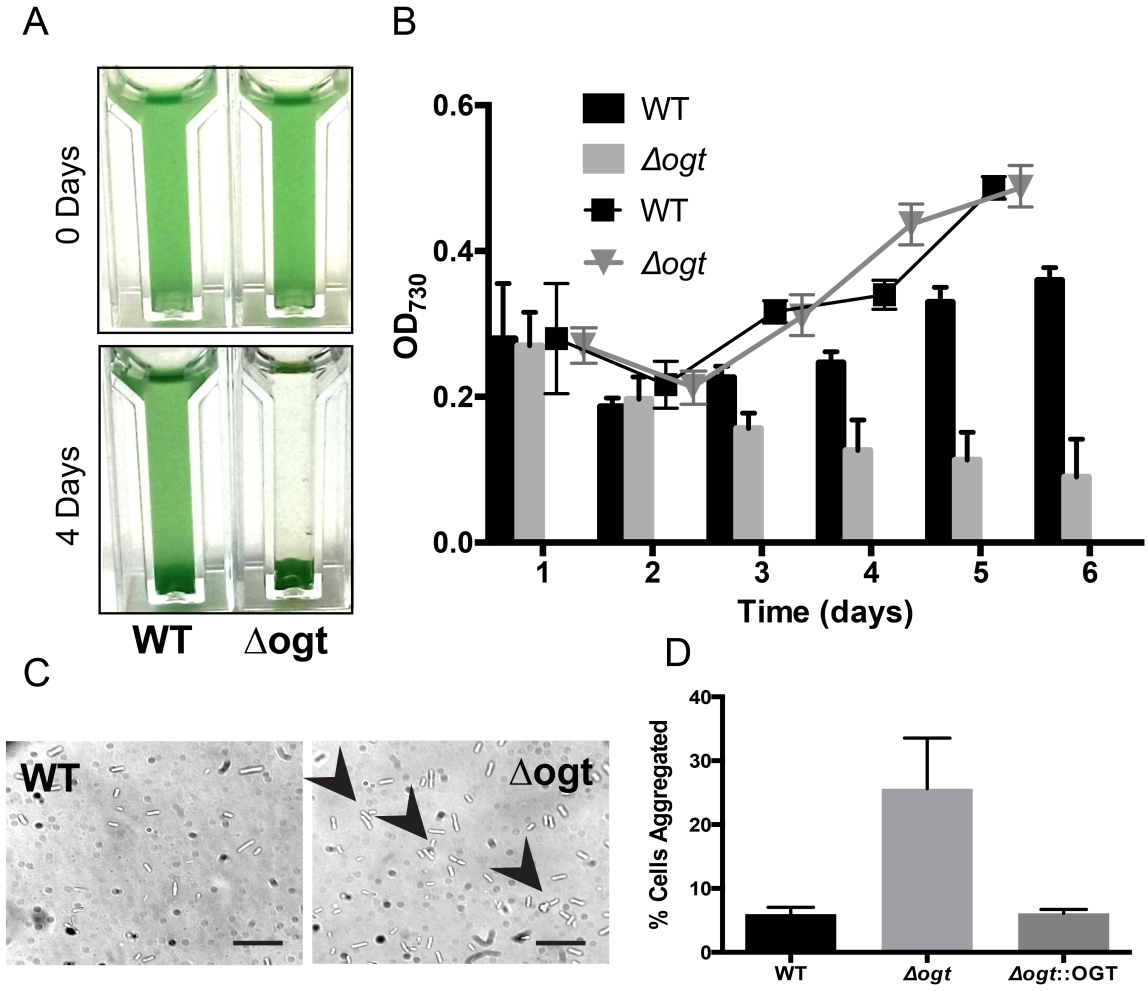

A
0 Days
4 Days
WT
Δogt
B
0.6
0.4
0.2
0.0
OD730
WT
Δogt
WT
Δogt
1
2
3
4
5
6
Time (days)
C
WT
Δogt
D
40
30
20
10
0
% Cells Aggregated
WT
Δogt
Δogt::OGT

Figure 4. Centrifugation in a Percoll gradient shows that wild type and Δ*ogt* cells have identical buoyancies. 1.5 mL PBS. 0.5 mL culture were gently pelleted, then resuspended in 0.5 mL PBS. Samples were overlayed in tubes on 8 mL Percoll (Amersham). The tubes were centrifuged in a Ti75 rotor (Beckman) at 20,000 rpm for 2 hours at 4°C.

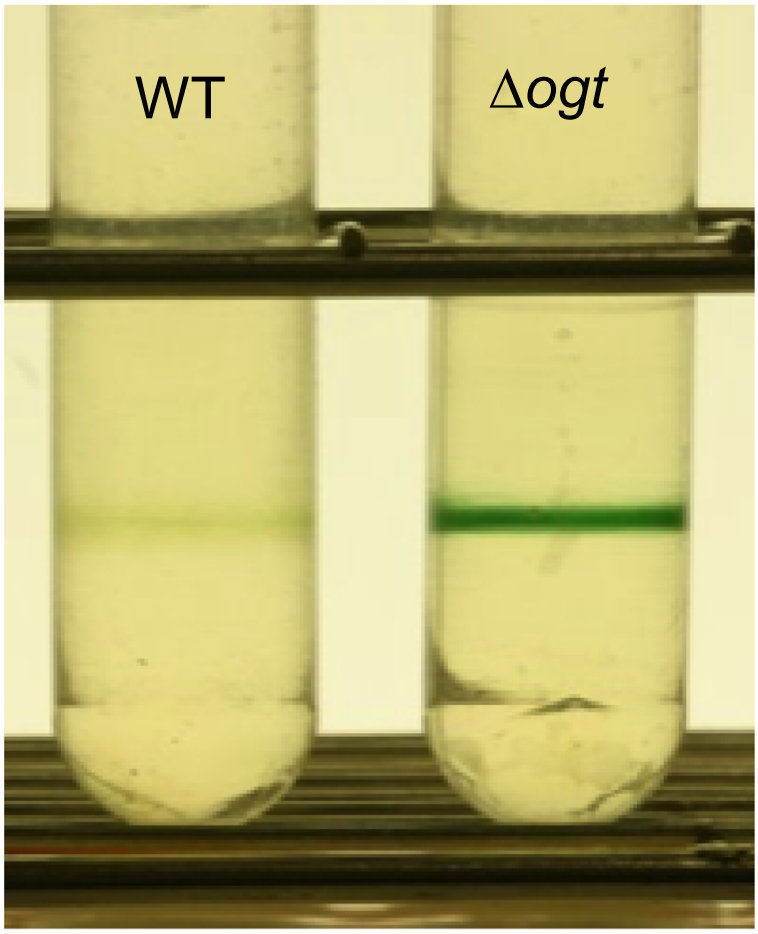

WT
Δogt

Figure 5. Wild type and Δ*ogt* cells settle faster under treatment with 0.05% sodium azide. Experiments were run at a range of initial densities. This figure is representative of other experiments.

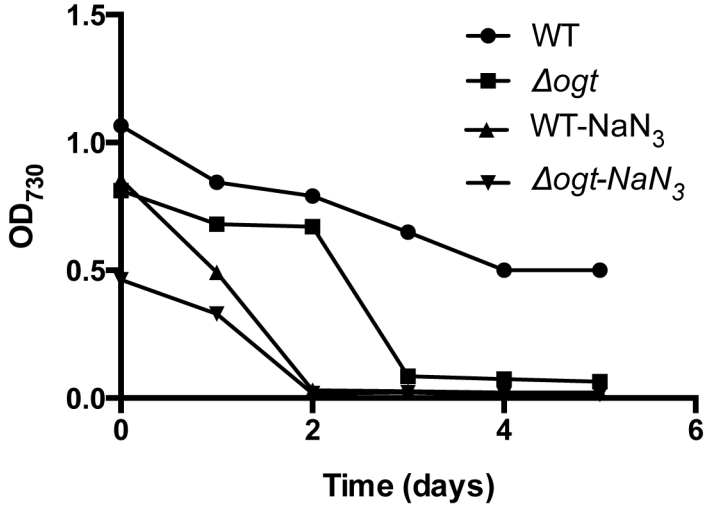

1.5
1.0
0.5
0.0
OD730
0
2
4
6
Time (days)
WT
Δogt
WT-NaN3
Δogt-NaN3

Figure 6. The Δ*ogt* mutant has a wider thylakoid lumen. (A) TEM images show Δ*ogt* mutant cells sharing typical cell structure with concentric thylakoids around a central chamber, but display wider thylakoid lumen (arrows). Scale bar = 200 nm (B) The thylakoid lumen width of wild type, Δ*ogt* mutant, and Δ*ogt* rescued with a SeOGT transgene. One-way ANOVA analysis with standard error of 16 cells per strain reveals that the thylakoid widths are statistically different ($p<0.05$) from each other.

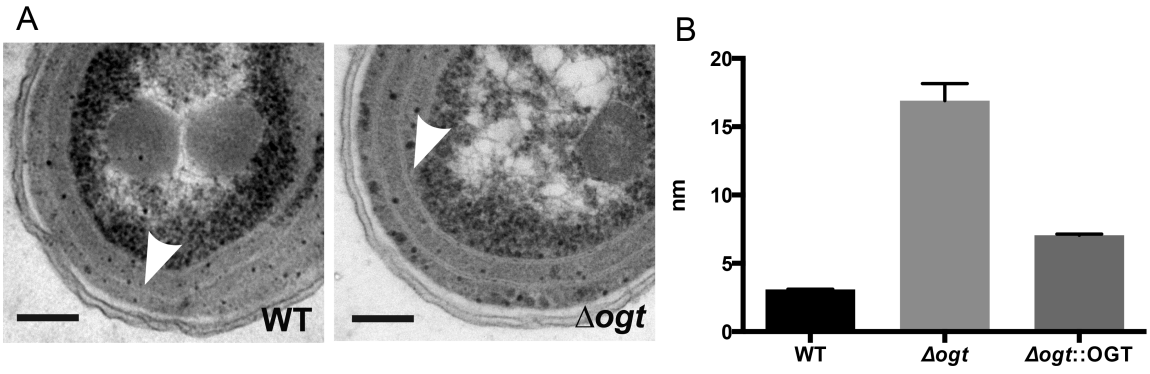

A
WT
Δogt
B
nm
20
15
10
5
0
WT
Δogt
Δogt::OGT

Figure 7. The Δ*ogt* mutant has aberrant electron dense bodies and accumulates inorganic phosphate. (A) Δ*ogt* mutant cells display an increase in electron dense bodies (arrows) distributed between the exterior of concentric thylakoid membrane pairs. The scale bar = 250 nm. (B) Free inorganic phosphate and inorganic phosphate released from polyphosphate bodies (PPBs) by exopolyphosphatase were quantitated. Although there is no difference in the amount of inorganic phosphate released from polyphosphate, mutant cultures contained nearly three-fold more free inorganic phosphate. The data represent the results from three complete biological replicates with the error bars indicating the standard error.

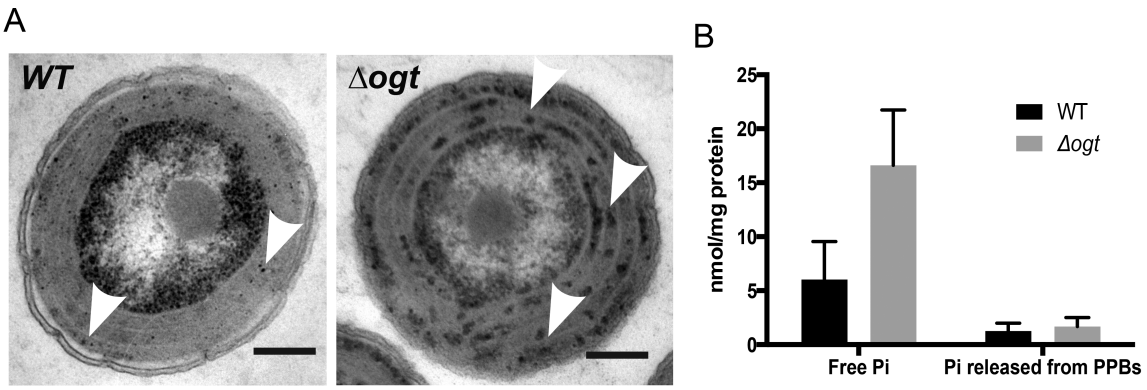

A
WT
Δogt
B
nmol/mg protein
25
20
15
10
5
0
WT
Δogt
Free Pi
Pi released from PPBs

Figure 8. Catalytic region mutations affect SeOGT activity. (A) Wild type and mutant SeOGT genes were introduced into the Δ*ogt* null mutant and the resulting strains were tested for rescue of the settling phenotype. Wild type OGT rescues the settling phenotype, H280A and K445A SeOGT do not and G350D weakly rescues. (B) Catalytic region mutations affect hydrolysis of UDP-GlcNAc. Wild type and mutant MBP-fusion proteins were purified. Enzymatic activity was measured based on the amount (DPM) of [H^3]GlcNAc released from UDP-[H^3]GlcNAc. The values shown are after subtraction [H^3]GlcNAc produced in a reaction with purified MBP-GFP. Reactions were performed in triplicate and analyzed by one-way ANOVA. Different numbers of asterisks indicate that the values are different ($p<0.05$). In both A and B, the error bars indicate the standard error.

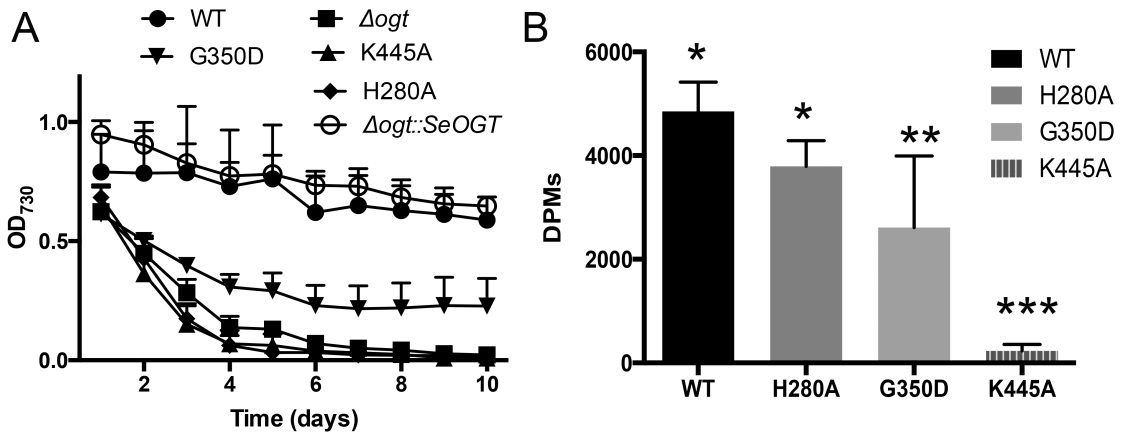
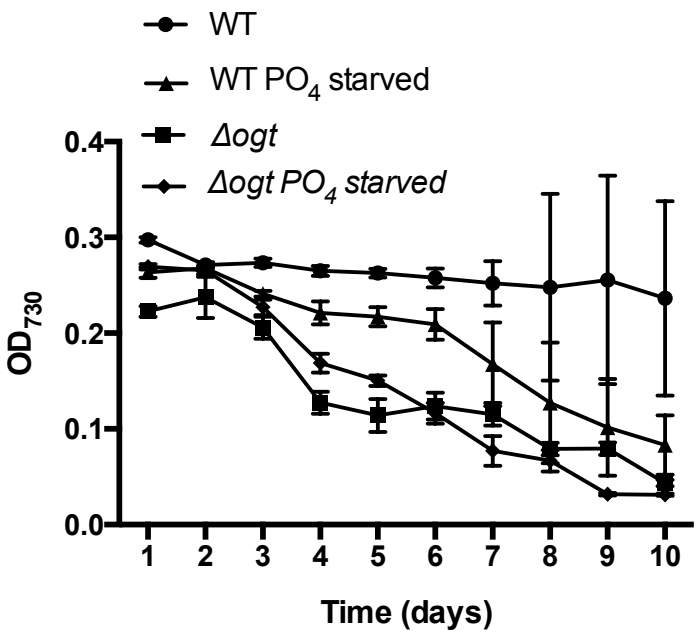

A
WT
G350D
Δogt
K445A
H280A
Δogt::SeOGT
OD730
1.0
0.5
0.0
2
4
6
8
10
Time (days)
B
DPMs
6000
4000
2000
0
*
*
**
***
WT
H280A
G350D
K445A

Figure 9. In phosphate-poor media wild type cells settle. The settling rate of Δ*ogt* cells is not affected. Cultures were grown to OD_{730} 0.5, gently centrifuged, and resuspended in BG-11 media made without K_2PO_4, then allowed to settle in cuvettes.


WT
WT PO4 starved
Δogt
Δogt PO4 starved
OD730
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
6
7
8
9
10
Time (days)

Chapter Three

A missense mutation in PilA suppresses multiple mutant phenotypes in the O-GlcNAc transferase deletion strain of *Synechococcus elongatus* PCC7942

INTRODUCTION

In bacteria, glycans can be added to proteins via O-linkages to serine or threonine residues, or by N-linkage to asparagine residues, and can be constructed sequentially or *en bloc* (64). Bacteria boast a wide variety of glycans due to the diversity of both the monosaccharides used, and resulting glycan configurations. The cell surfaces of diverse bacteria host an array of glycoproteins, which are used in virulence, adhesion, and motility.

In mammals, an important glycosylation is the addition of a single O-linked N-acetylglucosamine (O-GlcNAc) to the hydroxyl group of a serine or threonine residue. In metazoans, this modification has been implicated in a wide variety of cellular roles, including cell signaling, proteasome degradation, and gene transcription. Similar to phosphorylation, O-GlcNAcylation is dynamic. Unlike phosphorylation, which has a large variety of kinases and phosphatases to enable this cycling, O-GlcNAcylation commonly has only one enzyme for addition, O-GlcNAc transferase (OGT), and one enzyme for removal, O-GlcNAc hydrolase (OGA). Interestingly, the O-GlcNAc modification has been reported to compete with phosphorylation on the same residues. Numerous bacterial

genomes contain genes encoding for homologs of the eukaryotic OGT and/or OGA, but the cellular roles of these enzymes and O-GlcNAc modification in bacteria remain unclear.

The OGT of the cyanobacterium *Synechococcus elongatus* PCC 7942 has previously been shown to have hydrolytic activity against the sugar donor, UDP-GlcNAc (Chapter Two) (28). Mutant phenotypes have suggested varied cellular roles for the enzyme. Δ*ogt* mutants exhibit a range of seemingly unrelated phenotypes such as settling, larger electron-dense bodies, wider thylakoid lumen, and high inorganic phosphate levels. The substrates targeted by SeOGT involved in these phenotypes are unknown, as are the roles of the modification.

Type IV pili (T4P) are polymeric filaments found on the surface of many Gram negative bacteria that have multiple functional roles, including pathogenesis, DNA uptake, adhesion, and twitching motility. Pili consist of a homopolymer of the small protein PilA arranged helically, which can be dynamically extruded or retracted through the addition or subtraction of PilA units via assembly ATPases (PilB, PilF), and disassembly ATPases (PilT, PilU). Although the precise mechanisms for polymerization and retraction are not yet known, a core group of highly conserved assembly proteins are critical for pilus assembly, including the PilA subunits, assembly/disassembly ATPases, the core protein PilC, and the secretin channel protein PilQ (see Giltner et al. 2012 (65) for review). Regulation of pili production occurs at the transcriptional level as well as the posttranslational level. Pili are modified by glycosylation, most frequently

via O-glycosylation, which can modulate pili function (65). Defects in different types of glycosylation affect motility, host cell invasion, pilus assembly and others. *Synechococcus elongatus* PCC 7942 has two adjacent annotated pilA-like genes (Synpcc7942_0048 and 0049), however it has been reported that the pilA distal to the predicted promoter region, Synpcc7942_0048, is not commonly active (66).

RESULTS

Suppressor strain.

In the course of growing cultures, a culture of Δ*ogt* was observed to no longer exhibit the settling phenotype. I hypothesized that the culture had a spontaneous suppressor mutation. The strain was isolated and genotyped to confirm the Δ*ogt* mutation was present and that there was no evidence of the wild type OGT gene.

Characterization of suppressor phenotypes.

The suppressor mutant was tested for other phenotypes of the Δ*ogt* mutant. Although Δ*ogt* cultures settle out of media, both wild type and the suppressor strain do not (Fig 1A). Δ*ogt* cells aggregate to a greater extent than wild type, and the suppressor strain was suppressed for this phenotype as well (Fig. 1B). To determine whether other mutant phenotypes are suppressed by this mutation, I analyzed the inorganic phosphate content of cells. Δ*ogt* cells have a nearly two-fold increase in inorganic phosphate, but surprisingly, this phenotype was not suppressed in the suppressor strain (Fig. 2).

Whole genome sequencing reveals a SNP in pilA.

Genomic DNA was purified from wild type, Δ*ogt*, and the suppressor strain and submitted for whole genome sequencing. Upon comparison, a single nucleotide polymorphism (SNP) between Δ*ogt* and the suppressor genomes was identified. This guanine to thymine transversion results in an amino acid change from alanine to aspartic acid (A107D) in the promoter-proximal gene of the tandem pair of pilA-like genes (SynPCC7942_0049, GenBank ABB56081.1) (Fig. 3A, B). The suppressor strain is referred to as Δ*ogt, pilA*$^{(A107D)}$.

The A107D mutation rescues the Δogt settling and pili defect.

PilA is the major structural protein that comprises pili. To study the effects of the mutation on pili formation, we performed a Congo Red (CR) assay that takes advantage of Congo Red's affinity for pili (67). In this assay whole cells are exposed to CR. The residual CR in the supernatant is measured after removal of cells. Surprisingly, the Δ*ogt* mutant displayed a significant decrease in CR binding compared to wild type, suggesting fewer pili. The Δ*ogt, pilA*$^{(A107D)}$ mutation showed reversion to wild type CR binding levels, suggesting restoration of pili formation (Fig. 4A). These results led us to perform electron microscopy imaging to observe pili on whole cells. Consistent with the CR results, TEM imaging revealed pili on wild type and suppressor cells, but showed that Δ*ogt* cells were almost completely devoid of pili (Fig 4B).

Transgene rescues.

To investigate the interactions between OGT and PilA in *pilA*$^{(A107D)}$, several transgene constructs were made. Transgene rescue of Δ*ogt, pilA*$^{(A107D)}$ with the wild type OGT gene has no effect on CR binding (Fig 5A). Interestingly, introduction of a wild type PilA gene into Δ*ogt, pilA*$^{(A107D)}$ reduced CR binding. Introduction of *pilA*$^{(A107D)}$ into wild type cells also reduced CR binding, suggesting a lack of pili.

Introduction of wild type pilA into Δ*ogt, pilA*$^{(A107D)}$ resulted in reversion of the settling phenotype to that of the Δ*ogt* null mutant (Fig 5B). Introduction of *pilA*$^{(A107D)}$ into the wild type strain phenocopied the Δ*ogt* mutant. Congo Red assays reveal that when both wild type PilA and *pilA*$^{(A107D)}$ are expressed together pili production is inhibited. The ability of *pilA*$^{(A107D)}$ to cause settling when introduced to wild type suggests that the mutant and wild type PilA interfere with each other.

Pili are O-glycosylated with GlcNAc terminal modification.

To determine the glycosylation state of wild type, Δ*ogt*, and Δ*ogt, pilA*$^{(A107D)}$ pili were purified. Purified pilus proteins were run on 15% SDS-PAGE gels and Coomassie stained. Very little protein was detected near the predicted size of pili on the stained gel in the Δ*ogt* strain. Proteins on the SDS-PAGE gel were transferred to a nylon membrane for β-elimination Michael Addition with Biotin (BEMAB). This reaction replaces terminal O-linked moieties with a biotin molecule, which is detected using HRP-conjugated streptavidin. A protein of the

predicted pilin size, 15kD, was detected in wild type and Δ*ogt, pilA*$^{(A107D)}$, but was not detected in the Δ*ogt* pilin preparation (Fig. 6). Interestingly, both wild type and the Δ*ogt, pilA*$^{(A107D)}$ showed a strong BEMAB signal, indicating O-linked modifications are present on both pili. Despite there being a small amount of protein on the SDS-PAGE gel detectable by staining, the null mutant did not return any BEMAB signal, suggesting loss of an O-linked moiety on pili in this strain or that the 15kD protein in Δ*ogt* preparations is not PilA.

An assay to detect glycans was also performed in which glycans are oxidized with periodate on a blot, and detected by a reaction with biotin hydrazide followed by immunostaining. This assay labeled a protein of the predicted pilin size (Fig. 7). After β-elimination to remove O-linked glycans, the wild type band faded, but the Δ*ogt, pilA* $^{(A107D)}$ band was much more prominent. This may be a result of N-glycans taking the place of O-glycans in the absence of SeOGT. The suppressor may therefore have a different ratio of N- and O-linked glycans.

To determine whether the O-linked modification is a terminal O-GlcNAc, we performed an on-the-blot GalT reaction, wherein purified pili are run on SDS-PAGE gel, then transferred to a nylon membrane and exposed to galactosyl transferase (GalT) which transfers a ^{3}H-galactose to a terminal O-GlcNAc. Bands are apparent for wild type and Δ*ogt,pilA*$^{(A107D)}$ (Fig. 8). After β-elimination to remove O-linked glycans, no signal was obtained from either of the strains, indicating a terminal O-linked GlcNAc.

DISCUSSION

A spontaneous suppressor of Δ*ogt* mutant phenotypes in *S. elongatus* was identified. Whole genome sequencing uncovered a SNP in the pilA gene, resulting in the mutation $pilA^{(A107D)}$. Electron micrograph imaging revealed a loss of pili in Δ*ogt*, and suppression of this phenotype in Δ*ogt,* $pilA^{(A107D)}$, supporting the role of $pilA^{(A107D)}$ as a suppressor of Δ*ogt* phenotypes (Fig. 4B). It has been well established that many bacterial glycoproteins reside on the cell surface, so it is no surprise that the mutation suppressing the glycosyltransferase mutant phenotypes is a cell surface protein. Similar to flagella, T4Ps have been shown to be posttranslationally modified with a variety of glycans and some of these modifications are critical for several functions including motility, host cell invasion, pilus assembly and others (68, 69). The most common T4P modification, beyond the removal of the N-terminal signal sequence, is O-glycosylation, which can modulate pili function.

Wild type pili require the presence of OGT to produce pili, but the requirement for OGT is alleviated in the $pilA^{(A107D)}$ strain. If PilA subunits are directly modified by OGT, this result may be due to changes in conformation or changes in other local glycan modifications in the Δogt, $pilA^{(A107D)}$ strain that allow subunits to appropriately interact. Although the precise mechanism for pili assembly and production remains unclear, several key parts of the process are known. In one model suggested by Craig and Li (70), PilA subunits are first posttranslationally modified by cleavage of the signal sequence followed by

methylation of the N-terminus. Then PilA units remain as a pool anchored to the inner membrane via their hydrophobic α-helices and may be attracted to the base of the assembly mechanism by chemical complementarity of both the N-terminal region and the globular domains. The filament is extruded subunit by subunit by a cytoplasmic ATPase, PilB, through a secretin channel. Considering this model, defects in either PilA availability, PilA/PilA interaction, or some other aspect of the assembly mechanism could lead to a loss of pili.

The introduction of wild type pilA gene into the Δ*ogt, pilA*$^{(A107D)}$ strain inhibits pili production. Similarly, introduction of *pilA*$^{(A107D)}$ to wild type strain also inhibits pili. Inability to assemble pili with two types of PilA subunit may result from non-compatible subunit interaction, inhibiting homopolymer interaction. Pili defects may also be a result of defects in assembly machinery, as PilA interacts with pili subunits PilC, PilD, PilF and PilG to form a pilus. Pili require a complex set of proteins for assembly, which are generally well-conserved in Gram negative bacteria. These include the PilA subunits, PilC inner membrane core protein, PilQ outer membrane secretin channel, and two ATPases, PilB and PilT, devoted to both assembly and retraction, respectively, as well as several other accessory proteins and chaperones. An *in situ* chemical crosslinking study suggests that several of these components form a large complex linking both the inner and outer membranes (71). Defects in glycosylation of any of these components may prevent assembly.

PilA levels are regulated at the transcriptional level by a two-component system consisting of PilR and PilS. PilR binds the RpoN-RNA polymerase complex at the σ^{54}-dependent promoter of PilA and acts as an enhancer protein, allowing transcription to occur. PilR requires phosphorylation by the membrane-bound histidine kinase PilS, which responds to a yet unknown signal and autophosphorylates before acting on PilR. Studies of pilin gene transcription showed an increase in a pilA transcriptional reporter in pilin mutants, suggesting that PilA may negatively self-regulate expression . If PilA subunits are available at the inner membrane for assembly, but assembly is inhibited, either by defects in subunit interaction or by assembly mechanism defects, the unused PilA pool in the inner membrane may negatively regulate further pilA transcription. A defect in the PilR-PilS system may also negatively regulate pilA transcription, leading to loss of pili.

T4P pilin sequences show low conservation across species, except for several key regions. A variable leader peptide is followed by a distinctive and highly conserved cleavage sequence. At the N-terminus, the hydrophobic α-helix region serves as a transmembrane anchor to retain individual subunits at the ready for processing, and as the core of the final pilus. An amphipathic subdomain packs against a variable-length anti-parallel β sheet region, forming a globular head. Although sequences of these globular regions are divergent, the general structure remains conserved. The final C-terminal proximal loop is the

disulfide-bond loop, or the D-loop, created by cysteine-cysteine interactions, which is critical for pilus assembly and function (72).

Consistent with T4P, PilA of *Synechococcus elongatus* PCC7942 has the conserved cleavage signal sequence, and is predicted to form two N-terminal α-helix subdomains, followed by β sheets. The *pilA*$^{(A107D)}$ mutation in the suppressor strain is located near the predicted D-loop (Fig. 3). Mutations in the D-loop have been reported to affect pilus assembly and motility (72, 73), phenotypes seen in the Δ*ogt* strain. Although the suppressor has an intact D-loop region, the *pilA*$^{(A107D)}$ mutation may result in aberrant local structure, leading to instability of the D-loop. The *pilA*$^{(A107D)}$ mutation results in suppression of the pili assembly and motility phenotypes in Δ*ogt*. I speculate that this may be a result of structural changes or alternative glycosylation patterns allowing the un-modified D-loop to revert back to wild type function.

Notably, other mutant phenotypes are not rescued by the PilA mutation, possibly due to effects of the pilA-like gene distal to the promoter region. However, this tandem arrangement of pilA-like genes has been noted as a common occurrence, and the distal gene has been shown to be inactive (66). The lack of suppression of the inorganic phosphate mutant phenotype may indicate that SeOGT acts in several distinct cellular roles.

Although pili have not been established to have the O-GlcNAc modification specifically, they have been reported to carry a variety of glycosylations including O-linked moieties (74). We have detected both N- and

O-linked modifications on *S. elongatus* pili, both in wild type and the suppressor strain (Fig. 7), but have not determined the nature of the sugar modification. β-elimination of pili, which removes O-linked glycans, reveals that the ratio of N- and O-linkages is altered in the suppressor strain, suggesting a defect in regulation. Radioactive labeling of terminal O-GlcNAc results in visible bands, and β-elimination of O-linked glycans results in loss of bands, indicating terminal O-GlcNAc on pili (Fig.8). Further investigation into the type of glycosylation on *S. elongatus* pili will require structural characterization of the glycans on pili, and determination of the glycan positions on the protein.

MATERIALS AND METHODS

Bacterial Growth.

Cultures were grown in BG-11 media with no antibiotics unless specified . Cultures were grown at 22°C under continuous fluorescent light (33.75 μmol photons $m^{-2}s^{-1}$) with aeration. To aerate, an aquarium air pump (Petco) was used to introduce bubbles to the bottom of culture flasks through a Pasteur pipette plugged with sterile cotton.

Genomic DNA Preparation.

Genomic DNA was extracted using a method based on Clerico *et al.* (59). 20 mL of stationary phase culture was pelleted and left at -80°C overnight. The pellet was resuspended in 500 mL 120 mM NaCl and 20 mM EDTA then pelleted and

resuspend in 340 mL 25% sucrose 50mM Tris-HCl (pH 8), 10mM EDTA. 2 mg/mL lysozyme was added and the solution was incubated for 45 minutes at 37°C. 2 μL 10 mg/mL proteinase K and 20 μL 20% sarkosyl were added and the mixture vortexed for 20 seconds and allowed to incubate for 30 minutes at 55°C. 57 μL 5M NaCl and 45 μL of CTAB were added and the mixture was incubated at 65°C for 10 minutes. DNA was extracted using 500uL 24:1 chloroform:isoamyl alcohol by vortexing and centrifuging at 16,000g in a Beckman JA 18.1 rotor at 4°C for 10 minutes. The upper phase was transferred to a new tube and 500 μL phenol was added, then the mixture was vortexed for 20 seconds. This step was repeated once. The upper phase was transferred to a new tube and 2 volumes of 100% ethanol were added. The DNA was spooled onto a glass Pasteur pipette, washed with 70% ethanol, then transferred to a new tube, pelleted, and allowed to dry. The pellet was resuspended in 50 μL water.

Whole Genome Sequencing.

All sequencing was completed at the University of Minnesota Genomics Center using the Illumina Paired-end MiSeq platform (250bp paired-end, insert size = 280bp). Libraries were created for wild type (Adaptor #2 – CGATGT), mutant 1 (Adaptor #7 – CAGATC), mutant 2 (Adaptor #16 – CCGTCC), and the null mutant (Adaptor #4 – TGACCA) using the Illumina TruSeq DNA Sample Preparation Kit. All raw sequencing files have been submitted to NCBI Sequence Read Archive under accession number PRNJA268360 SRPXXXXX. FASTQ files from all sequencing runs were imported into CLC Genomics Workbench Version

6.5 (CLC Bio, www.clcbio.com) and analyzed on the desktop computer described in Dorn et al. (75). Illumina reads were subjected to quality control using the Sequencing QC Report tool. 250 bp paired-end Illumina reads were trimmed and filtered using the following parameters: Ambiguous Trim = Yes, Ambiguous Limit = 2, Quality Limit = 0.05, Use Colorspace = No, Search Reverse Sequence = Yes, Remove 5' terminal nucleotides = Yes, Number of 5' Terminal Nucleotides to Remove = 10, Remove 3' terminal nucleotides = No, Discard Long Reads = No, Save Broken Pairs = No. Trimmed and filtered reads were mapped to the Synechococcus elongatus genome (GenBank accession NC_007604) and plasmid (GenBank accession CP000101) using the 'Map Reads to Reference' tool in CLC Genomics Workbench (Mapping Parameters: Masking mode: No Masking, Mismatch cost = 3, Insertion cost = 3, Deletion cost = 3, Length Fraction = 0.95, Similarity Fraction = 0.95, Global Alignment = No, Auto-detect paired distances = Yes, Non-specific match handling = Ignore). Sequence variants were analyzed using the 'Probablistic Variant Detection' tool in CLC Genomics Workbench (Parameters: Ignore non-specific matches = Yes, Ignore Broken Pairs = Yes, Minimum Coverage = 10, Variant Probability = 0.90, Required Variant Count = 2, Require presence of both forward and reverse reads = Yes, Ignore variants in non-specific regions = No, Maximum Expected Variants = 2). Bioproject information: Bioproject PRJNA268360, Wild type SRR Accession # SRR1685987, deletion mutant (null ogt) SRR Accession # SRR1686015, Suppressor mutant SRR Accession # SRR1686063.

Congo Red assay for pili.

To determine pilin abundance, a Congo Red assay was adapted from Black and Yang (76). Cells were grown to exponential phase, then 1 mL was gently pelleted, and resuspended in 500 µL Congo Red buffer (10 mM Tris (pH 7.5), 1 mM KH_2PO_4, 5 mM $MgSO_4$, and 15 µg/mL Congo Red). Cells were incubated at 25°C for 30 minutes in a dark room. Cells were then pelleted and supernatant was transferred to cuvettes. Absorbance at 490 nm was measured. Cell free samples were used as controls.

TEM.

TEM images were obtained using a modified protocol from Ahlawat et al. (77). Briefly, a 400 mesh copper grid was layered with formvar and carbon film. Grids were put into 10 µl poly-L-lysine, which was then blotted off. Cells were applied to the grid upside down using 6 µl of saturated culture, 2 µl glutaraldehyde and 2 µl sodium phosphate buffer for 20 minutes. Slightly dried grids were stained with uranyl formate. Cells were visualized using a Philips CM12 Transmission Electron Microscope and digitally photographed.

Inorganic phosphate measurements.

S.elongatus cultures were grown to exponential phase and cells were pelleted by centrifugation for 15 minutes at 3700xg in a JA-20 rotor (Beckman). The cell pellets were frozen at -80°C, thawed, and re-suspended in French Press Buffer consisting of 20 mM Tris-HCl (pH 7.5), 2 mM EDTA, 1 M sodium perchlorate, 1% SDS w/v and cells were disrupted via two passes through a French Press.

Cellular debris was cleared by centrifugation at 4°C for 15 minutes, at 20,000xg in a JA-20 rotor (Beckman) and aliquots of the supernatant were frozen at -20°C.

Free inorganic phosphate was determined using a modified protocol from Werner et al. (46). 50 µL of cell lysate was added to 50 µL 1M sulfuric acid and vortexed briefly. The mixture was incubated for five minutes, then neutralized with 50 µL 2M sodium hydroxide and 100 µL Red Tris solution (1M Tris-HCl (pH 7.5), 6% Neutral Red solution (Ricca Chemical Company, Arlington TX)), vortexed, then centrifuged to remove precipitated proteins. pH was assessed via color test and adjusted to neutral (acidic solutions are pink, basic solutions are yellow and neutral pH solutions are pinkish-orange). 50 µL aliquots of this neutralized solution were transferred to microcentrifuge tubes.

To measure phosphate in these samples, 344 µL of 28mM ammonium heptamolybdate in 2.1M sulfuric acid was added to each tube and tubes were incubated for 20 minutes at 22°C. To develop color, 256 µL of 0.76mM malachite green oxalate in 0.35% polyvinyl alcohol was added to each sample and incubated for 20 minutes at 22°C. Absorbance at 595 nm was then measured. A standard curve using fresh potassium phosphate solutions was created for each experiment.

Isolation of pili.

Cultures were grown to saturation and 300 mLs were centrifuged at 4000xg in a JA-10 rotor (Beckman) for 20 minutes. The pellet was resuspended in 5 mL 5 mM Tris (pH 8.0). Samples were sheared for 3 minutes using a polytron

(Kinematica) at level 4, then centrifuged at 12,000xg for 30 minutes in a JA-20 rotor (Beckman). Supernatant was collected, and ammonium sulfate was added to 50%. The solution was stirred overnight at 4°C, then centrifuged at 12,000xg for 60 minutes in a JA-20 rotor. The pellet was resuspended in 300 µL 5mM Tris-HCl (pH 7.5) and dialyzed against 1 L 5 mM Tris-HCl (pH 7.5).

Transgene complementation.

Transgenes were produced using the pSyn vector (Invitrogen) according to manufacturer's protocols. Briefly, genes and putative promoter regions were PCR amplified and cloned into the pSyn_1/D-TOPO vector. Strains were transformed as previously described (Sokol & Olszewski 2015), and transformants selected using spectinomycin resistance. Primer pairs included in the pSyn kit were used to verify strain transformation.

BEMAB.

Samples were run on 15% SDS-PAGE gels and transferred to PVDF membrane. In a technique adapted from Wells *et al* (2002), the blot was incubated in 5 mL of 1% triethylamine, 0.1% NaOH, 3 mM biotin-PEG3-amine (ChemPep). The blot was then incubated at 37°C for 24 hours with gentle agitation. The blot was washing with ddH20, blocked in PBS with 0.05% Tween and 3% BSA, and probed with Streptavidin-HRP, which was detected by chemiluminescence.

GalT on the blot labeling.

Samples were run on 15% SDS-PAGE gels and transferred to PVDF membrane. Membrane was blocked overnight at 4°C in 3% bovine serum albumin in 10 mM

HEPES (pH 7.9). Membrane was washed in 10 mM HEPES (pH 7.9) for 10 minutes with gentle shaking. Three µCi of UDP-H^3-Galactose was dried under argon and resuspended in 3 µL 25 mM AMP and added to 800 µL labeling buffer (10 mM HEPES (pH 7.9), 1M $MNCl_2$, 1% BSA, protease inhibitor, and 100m units calf intestinal phosphatase. The membrane was sealed in plastic and incubated for 3 hours at 37°C. The membrane was washed 7 times for 10 minutes each in 10 mM HEPES pH (7.9). The membrane was dried and sprayed with EN^{3}HANCE (PerkinElmer) and dried for 20 minutes. The membrane was placed on pre-flashed film, then sealed and placed at -80°C for 72 hours.

Glycodetection.

Proteins in pili preparations were resolved by PAGE on a 15% gel and transferred to a PVDF membrane. The membrane was wet with 95% ethanol, rinsed in H2O and washed 3x in 100 mM sodium acetate (pH 5.5) for 5 minutes. Glycoproteins on the membrane were oxidized by incubating the membrane in 50 mL of freshly prepared 100 mM sodium acetate (pH 5.5) containing 20 mM sodium periodate at room temperature in the dark for 30 minutes. The membrane was washed for 5 minutes in 50 mL of 20 mM sodium phosphate (pH7.4) containing 20 mM sodium sulfite and then washed three additional times in 20mM sodium phosphate (pH7.4). After transferring the membrane to a seal-a-meal bag, the glycoproteins were biotinylated by adding 8 mL of 20 mM sodium phosphate (pH 7.4) to which 400µL of DMSO in which 10.4 mg of biotin hydrazide had been dissolved. The biotinylation reaction was performed at room

temperature for 2 hours. The reaction was terminated by washing the membrane 3 times for 5 minutes in PBS. Biotin on the membrane was detected using strepavidin-HRP.

Figure 1. Suppression of Δ*ogt* phenotypes. (A) The suppressor strain is rescued for the *ogt* mutant settling phenotype. (B) The suppressor strain has similar percentage of cells in aggregates as wild type, indicating the suppressor mutation also suppressed the aggregation phenotype. Aggregates of two or more cells were counted, excluding cells with end-to-end aggregates, as these may be dividing cells. Bars in A and B indicate standard error, some bars being too small to be visible. In B, analysis of variance (ANOVA) was performed. Asterisk indicates significant difference ($p < 0.05$).

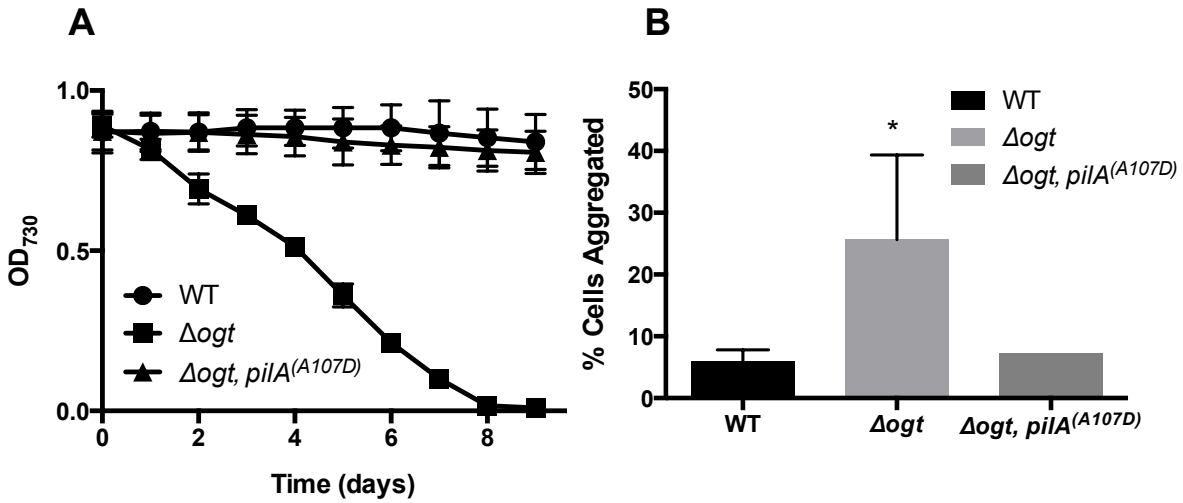

A
B
1.0
0.5
0.0
OD730
0
2
4
6
8
Time (days)
WT
Δogt
Δogt, pilA(A107D)
50
40
30
20
10
0
% Cells Aggregated
*
WT
Δogt
Δogt, pilA(A107D)

Figure 2. Non-suppression of Δ*ogt* phosphate phenotype. Δ*ogt* cells have higher inorganic phosphate than wild type cells, and the suppressor mutation does not rescue this mutant phenotype. Inorganic phosphate is expressed as phosphate per mg of protein in cell lysates. Bars indicate standard error. ANOVA was performed. Asterisks indicate significant difference from wild type ($P < 0.05$).

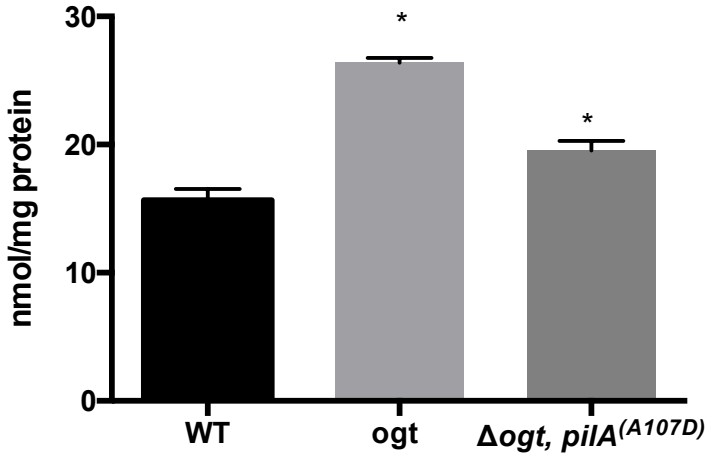

30
20
10
0
nmol/mg protein
*
*
WT
ogt
Δogt, pilA(A107D)

Figure 3. Diagram of missense mutation in PilA. (A) Nucleotide sequence of *PilA* (top line) and the suppressor mutation (bottom line). (B). Amino acid sequence of PilA (top line), consensus sequence (middle line) and the *pilA* suppressor mutation which causes an alanine to aspartic acid change (A107D). The predicted D-loop region adjacent to the mutation is noted.

A

```
1    ATGTCTGACTCCCTCCGTCTTCGTTATCTACAATATCTTGCCCAGCGTAAAGACGAACAA  60
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1    ATGTCTGACTCCCTCCGTCTTCGTTATCTACAATATCTTGCCCAGCGTAAAGACGAACAA  60

61   GGTGAAGAAGAAAAAGGTTTTACCCTTGTCGAGTTGCTGGTCGTTATCATTATCGTTGGC  120
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61   GGTGAAGAAGAAAAAGGTTTTACCCTTGTCGAGTTGCTGGTCGTTATCATTATCGTTGGC  120

121  ATCTTGGCAGCAGTTGCATTACCGAACCTGTTGGATCAAACAGATAAAGCCTATGCCTCT  180
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121  ATCTTGGCAGCAGTTGCATTACCGAACCTGTTGGATCAAACAGATAAAGCCTATGCCTCT  180

181  GAAGGCAAATCAGCAGTCGGTGCTGCTCTTCGTACCCTTAGTGCGGCGACACTAGATCCT  240
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181  GAAGGCAAATCAGCAGTCGGTGCTGCTCTTCGTACCCTTAGTGCGGCGACACTAGATCCT  240

241  AACTACGTCACCAATGCGTCTTGTACACAGCTTGGTATTGGTAGCAGTGCAGGTAACTTT  300
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241  AACTACGTCACCAATGCGTCTTGTACACAGCTTGGTATTGGTAGCAGTGCAGGTAACTTT  300

301  AACATAACTTGCGGCAATGCTAGCCAAGTAACGGCTGCTGGAAGTGGTAAAGCAGCGAAT  360
     ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
301  AACATAACTTGCGGCAATGATAGCCAAGTAACGGCTGCTGGAAGTGGTAAAGCAGCGAAT  360

361  ATTAACGTGACTGGCACAATCGGGACAGACGGTAAATTTACCGTTATTGCAACCAAAGGC  420
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361  ATTAACGTGACTGGCACAATCGGGACAGACGGTAAATTTACCGTTATTGCAACCAAAGGC  420

421  AGCGCAACTCTTTAA  435
     ||||||||||||||
421  AGCGCAACTCTTTAA  435
```

B

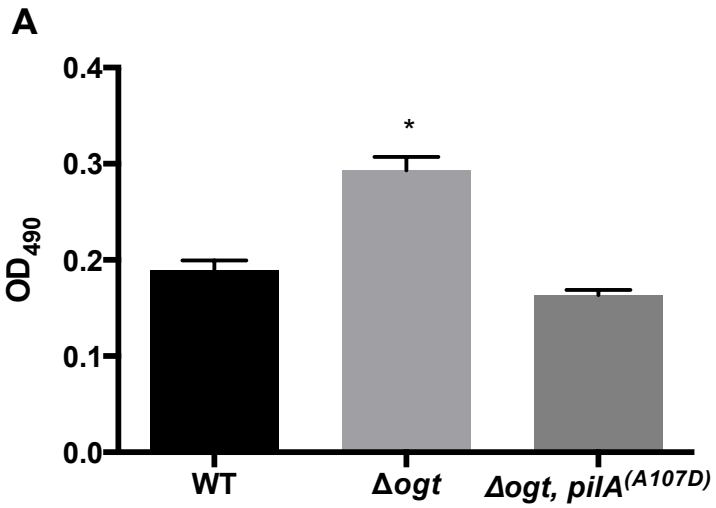
```
     SDSLRLRYLQYLAQRKDEQGEEEKGFTLVELLVVIIIVGILAAVALPNLLDQTDKAYASE  60
     SDSLRLRYLQYLAQRKDEQGEEEKGFTLVELLVVIIIVGILAAVALPNLLDQTDKAYASE
     SDSLRLRYLQYLAQRKDEQGEEEKGFTLVELLVVIIIVGILAAVALPNLLDQTDKAYASE  61

     GKSAVGAALRTLSAATLDPNYVTNASCTQLGIGSSAGNFNITCGNASQVTAAGSGKAANI  120
     GKSAVGAALRTLSAATLDPNYVTNASCTQLGIGSSAGNFNITCGN SQVTAAGSGKAANI
     GKSAVGAALRTLSAATLDPNYVTNASCTQLGIGSSAGNFNITCGNDSQVTAAGSGKAANI  121

     NVTGTIGTDGKFTVIATKGSATL  143
     NVTGTIGTDGKFTVIATKGSATL
     NVTGTIGTDGKFTVIATKGSATL  144
```

D-loop

Figure 4. Suppression of Δ*ogt* loss of pili phenotype. (A) Congo Red preferentially binds to pili. Cells with fewer pili bind less Congo Red dye suggesting a loss of pili in Δ*ogt* cells which is then restored in Δ*ogt,pilA*$^{(A107D)}$. Bars indicate standard error. ANOVA was performed. Asterisks indicate significant difference from wild type ($P < 0.05$). (B) TEM imaging of negatively stained cells reveals a loss of pili in the Δ*ogt* mutant, but reversion to wild type in the suppression mutant Δ*ogt, pilA*$^{(A107D)}$.


A
0.4
0.3
0.2
0.1
0.0
OD490
*
WT
Δogt
Δogt, pilA(A107D)

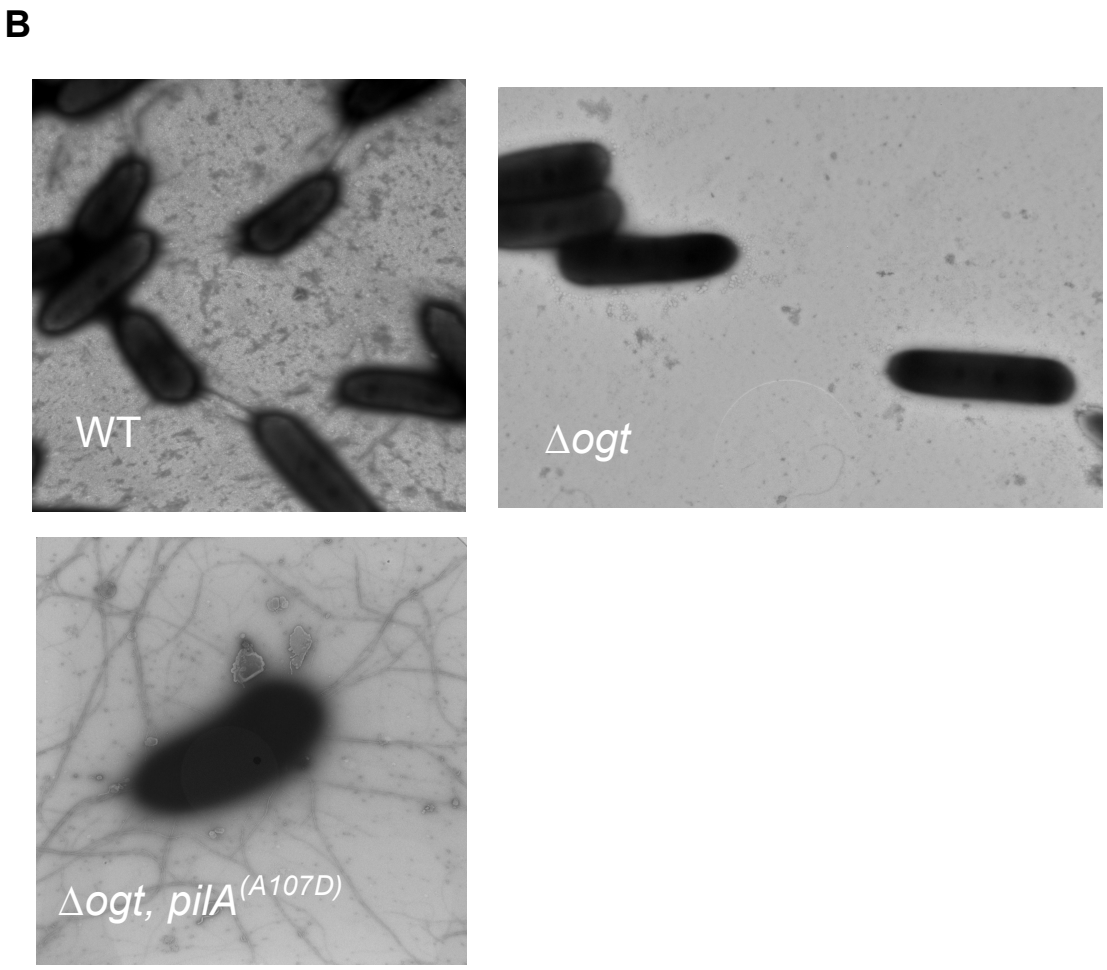
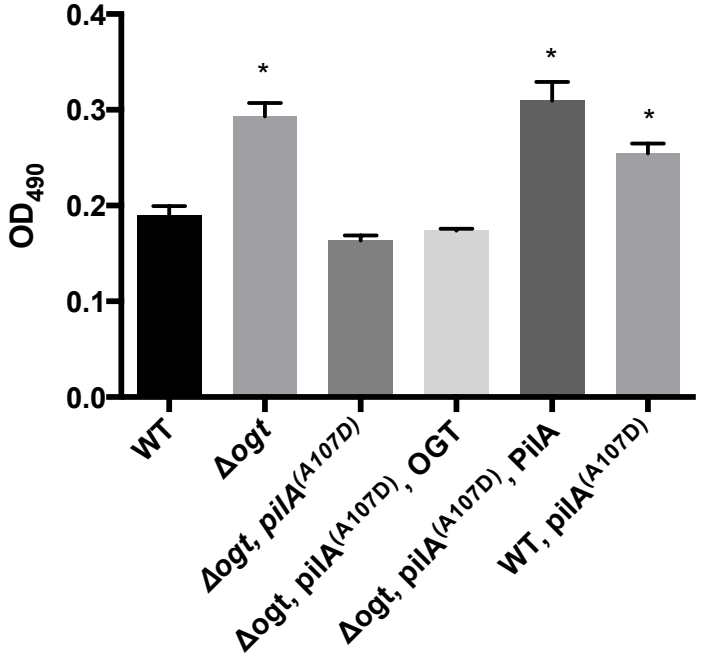

B
WT
Δogt
Δogt, pilA(A107D)

Figure 5. Transgene effects on pili and settling phenotypes. Several transgene experiments were performed. OGT and pilA transgenes were separately introduced to the Neutral Site I of Δ*ogt, pilA*$^{(A107D)}$. Wild type was transformed with *pilA*$^{(A107D.).}$ A) Congo Red was used to determine the presence of pili. Bars indicate standard error. ANOVA was performed. Asterisks indicate significant difference from wild type ($P < 0.05$). B) The settling phenotype was determined for each of the transgene constructs. Bars indicate standard error, with some bars too small to be visible.

A


0.4
0.3
0.2
0.1
0.0
OD490
*
*
*
WT
Δogt
Δogt, pilA(A107D)
Δogt, pilA(A107D), OGT
Δogt, pilA(A107D), PilA
WT, pilA(A107D)

B

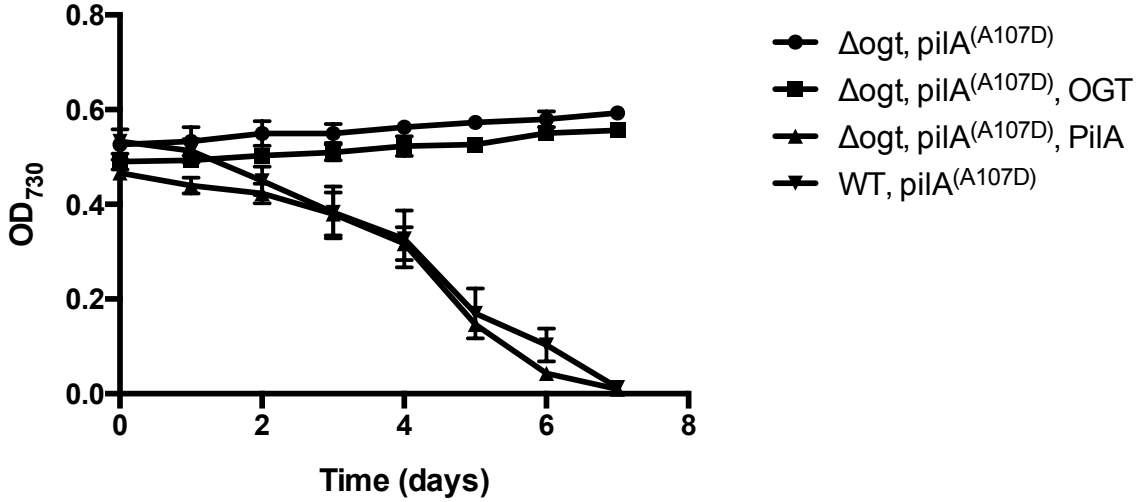

0.8
0.6
0.4
0.2
0.0
OD730
0
2
4
6
8
Time (days)
Δogt, pilA(A107D)
Δogt, pilA(A107D), OGT
Δogt, pilA(A107D), PilA
WT, pilA(A107D)

Figure 6. BEMAB treatment. Twenty μg of isolated pili were run on an SDS PAGE gel and transferred to a PVDF membrane. β-elimination of O-linked glycans was performed, followed by Michael addition of biotin. Biotin was detected using streptavidin-HRP. Both wild type and Δ*ogt,pilA*$^{(A107D)}$ show pili with O-linked glycans. Δ*ogt* shows inhibited pili formation.

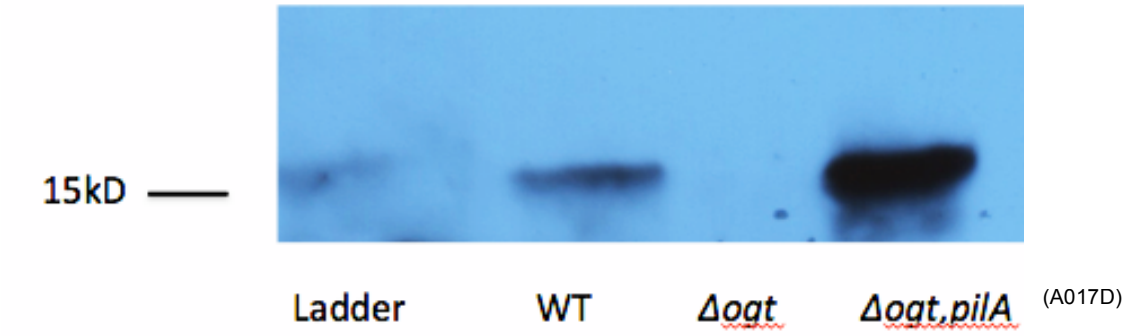
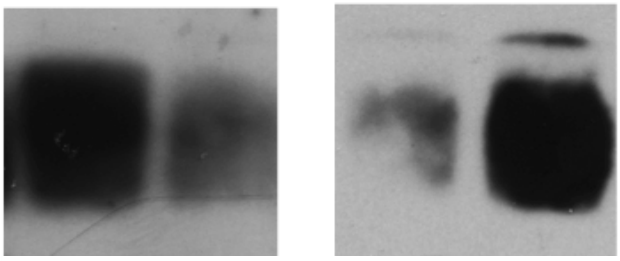

15kD
Ladder
WT
Δoat
Δoat,pilA
(A017D)

Figure 7. Glycosylation of pili. A) Glycodetection on bands running at the predicted size for pili reveal glycosylation on both wild type and Δ*ogt,pilA*$^{(A107)}$. B) After elimination of O-linked glycans, glycosylation ratios change in both strains.

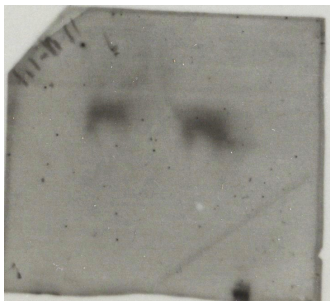
A B



WT Δ*ogt, pilA*$^{(107A)}$ WT Δ*ogt, pilA*$^{(107A)}$

Figure 8. Detection of terminal O-linked GlcNAc. Isolated pili were run on a 15% SDS-PAGE gel and transferred to PVDF membrane. Glycans terminating with O-GlcNAc were labeled with H^3-Galactose A) Both WT and Δ*ogt,pilA*$^{(A107D)}$ show bands of terminal O-GlcNAc near the predicted size of pili. B) β-elimination of O-linkages removes both of these bands, suggesting that both have structures with a terminal GlcNAc.

A



WT Δ*ogt, pilA*$^{(A107D)}$

B

WT Δ*ogt, pilA*$^{(A107D)}$

Chapter Four

Conclusions and Future Directions

The eukaryote-like O-GlcNAc transferase of *Synechococcus elongatus* PCC7942 hydrolyzes the UDP-GlcNAc sugar donor and has a role in settling, aggregation, and inorganic phosphate levels. These phenotypes, of which only settling and aggregation are affected by the suppressor mutation, indicate that some of these cellular processes may be independent. Glycosylation status of pili may affect pili formation, via the PilA subunit. Of particular interest in these phenotypes are the loss of motility, loss of pili, and the increased cellular inorganic phosphate.

Despite the suggestion of multiple cellular roles for SeOGT, the mechanisms by which SeOGT affects these cellular processes remain unclear. To date, no O-GlcNAcylated proteins have been identified in *Synechococcus elongatus* PCC7942. Ultimately, identification of the substrates of the OGT would aid in elucidation of the cellular roles of the enzyme, and allow for *S. elongatus* to serve as a model system for studying the O-GlcNAc modification. To that end, several avenues of study are apparent.

Metazoan O-GlcNAcylation acts in both the nucleus and cytosol on varied systems. The location of the OGT in the bacteria may elucidate which cell fractions to investigate for SeOGT substrate proteins. For instance, loss of SeOGT leads to loss of pili on the exterior of the cell. A significant amount of bacterial glycosylation occurs in the periplasm. If SeOGT is found in the

periplasm or near the inner membrane, we can extrapolate that it may be acting on a protein involved in the pilus assembly mechanism. Similarly, if SeOGT is found near the genomic DNA, it may be acting on transcription factors. To determine the location of SeOGT, immunolabeling could be used in immunogold staining and imaged in sections on a transmission electron microscope. To this end, I have created a strain with an epitope tagged fusion protein of SeOGT tagged with 3xFLAG tag (Fig. 1) in the null OGT background strain, which could be used towards this goal. The FLAG tag can be visualized using a primary antibody and a secondary antibody fused with colloidal gold particles, then examined under transmission electron microscopy.

To further investigate the cellular roles of O-GlcNAcylation in bacteria, mutant phenotypes of the Δ*ogt* strain remain to be investigated. Of particular interest are the high inorganic phosphate found in mutant cells and loss of pili. Because the regulation of phosphate acquisition, storage, and release are well-studied, there are many targets for study in this pathway, including the PhoB-PhoR transcriptional regulators, the *pst* system which transports inorganic phosphate across the cytoplasmic membrane, and ppK and ppX which serve as a phosphate kinase and an exopolyphosphatase respectively and control storage and use of inorganic phosphate. The Pho regulon, which is responsible for phosphate regulation, is conserved in *Synechococcus elongatus*. Investigation into expression of the genes in this regulon could elucidate irregularities in genetic regulation. Co-expression studies in *E. coli* using SeOGT and some of

the key phosphate regulation proteins may uncover an O-GlcNAcylation substrate in the phosphate regulon. Genes involved in pili production are also well characterized. Deletion mutants in both wild type and Δ*ogt* may help elucidate the step in pili production in which O-GlcNAcylation is involved.

Similar to the discovery of the spontaneous suppressor, the settling phenotype of Δ*ogt* could be used in further suppressor screening, analyzing those strains that regain the ability to remain in culture. This approach could uncover additional genes involved in O-GlcNAcylation-dependent motility. Although more time-consuming, other phenotypes could also be used in a suppressor screen, including aggregation and inorganic phosphate levels to determine the role(s) of O-GlcNAcylation in these phenotypes.

The loss of pili in the Δ*ogt* mutant, along with the suppressor mutation in pilA, is compelling evidence that SeOGT is acting on the assembly or stability of pili. Antibodies raised against PilA would serve several purposes in the investigation of this mutant phenotype. The presence of the mutant protein despite the lack of pili would suggest that assembly defects are present, whereas its absence would suggest instability of the translated protein, aberrant localization, or instability of PilA when not incorporated into pili. In the absence of protein, knowledge of transcription rates of pilA could help determine if the gene is being down regulated. Co-localization studies of SeOGT using the previously described SeOGT-FLAG construct and PilA antibodies would be helpful in ascertaining whether SeOGT works directly on the assembly machinery of pili.

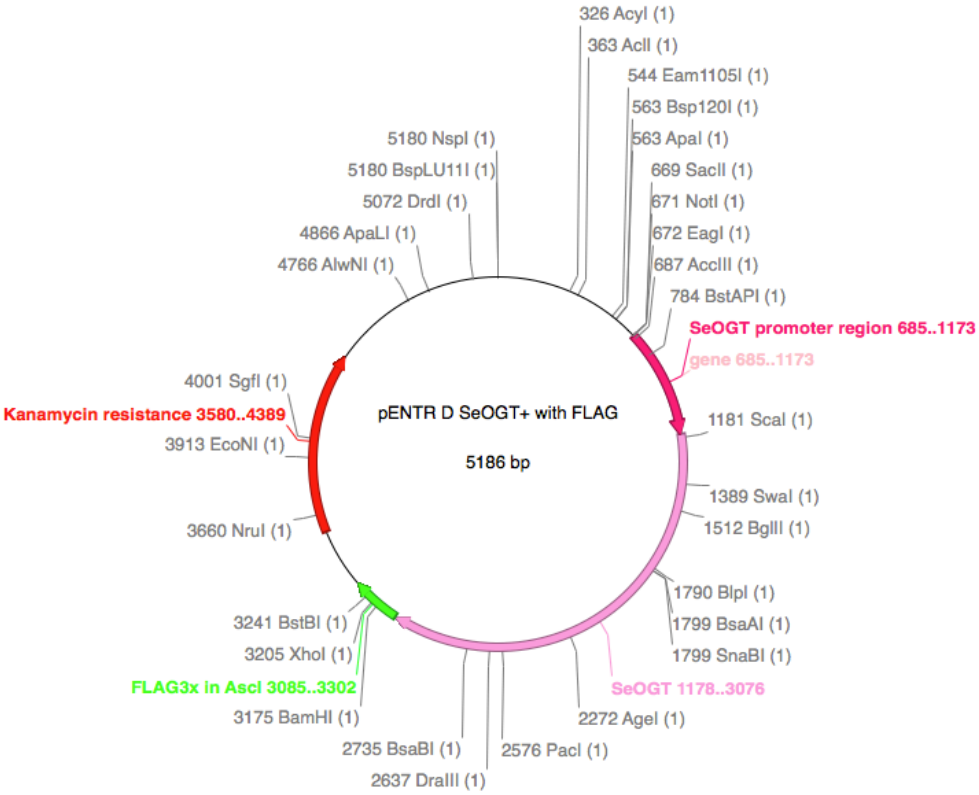
The location of the mutation in the mutant *PilA* raises interesting questions. Proximity to the putative disulfide bond creating the D-loop may suggest that mutant subunits are unable to assemble. Directed mutagenesis of amino acid residues in this region could lead to a better understanding of the relationship and interactions of the PilA subunits.

When mutant pilA is introduced to wild type cells, there is a loss of motility and a decrease in pili, suggesting that wild type PilA and PilA$^{(A107D)}$ may not be able to polymerize into full length pili. This may be due to altered glycosylation patterns in PilA$^{(A107D)}$ affecting polymerization, or conversely, it may be the effect of defects in the assembly machinery when PilA$^{(A107D)}$ is present. Although O-GlcNAcylation of pili has not been demonstrated, differential glycosylation patterns on wild type and suppressor pili are apparent. Further study into the modifications of pili using mass spectrometry may give insight into the functional importance of pili glycosylation, including glycan structure and location of glycans. This information may also give insight into whether glycosylation patterns are involved in PilA-PilA subunit interactions.

As noted in Chapter Two, a more holistic approach may be required to identify SeOGT substrates in *Synechococcus elongatus*. Glycoproteomics, using mass spectrometry, would be useful to uncover a spectrum of O-GlcNAcylated proteins. Additionally, several techniques have been reported to map the sites of O-GlcNAcylation on peptides. Once peptide substrates are identified, the site(s) of modification can be uncovered using techniques such as electrospray

ionization-collision induced dissociation tandem mass spectrometry (ESI-CID-MS/MS).

Figure 1. Map of SeOGT-Flag fusion construct.


326 AcyI (1)
363 AclI (1)
544 Eam1105I (1)
563 Bsp120I (1)
563 ApaI (1)
669 SacII (1)
671 NotI (1)
672 EagI (1)
687 AccIII (1)
784 BstAPI (1)
SeOGT promoter region 685..1173
gene 685..1173
1181 ScaI (1)
1389 SwaI (1)
1512 BglII (1)
1790 BlpI (1)
1799 BsaAI (1)
1799 SnaBI (1)
SeOGT 1178..3076
2272 AgeI (1)
2576 PacI (1)
2637 DraIII (1)
2735 BsaBI (1)
3175 BamHI (1)
FLAG3x in AscI 3085..3302
3205 XhoI (1)
3241 BstBI (1)
3660 NruI (1)
3913 EcoNI (1)
Kanamycin resistance 3580..4389
4001 SgfI (1)
4766 AlwNI (1)
4866 ApaLI (1)
5072 DrdI (1)
5180 BspLU11I (1)
5180 NspI (1)
pENTR D SeOGT+ with FLAG
5186 bp

References

1. **Torres CR, Hart GW**. 1984. Topography and Polypeptide Distribution of Terminal N-Acetylglucosamine Residues on the Surfaces of Intact Lymphocytes. J Biol Chem **259**:3308–3317.

2. **Ma J, Hart GW**. 2014. O-GlcNAc profiling: from proteins to proteomes. Clin. Proteomics **11**:8.

3. **Olszewski NE, West CM, Sassi SO, Hartweck LM**. 2010. O-GlcNAc protein modification in plants: Evolution and function. Biochim. Biophys. Acta **1800**:49–56.

4. **Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O**. 2011. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu. Rev. Biochem. 825–858.

5. **Shafi R, Iyer SP, Ellies LG, O'Donnell N, Marek KW, Chui D, Hart GW, Marth JD**. 2000. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proc. Natl. Acad. Sci. U. S. A. **97**:5735–9.

6. **Gambetta MC, Oktaba K, Müller J**. 2009. Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. Science **325**:93–6.

7. **Webster DM, Teo CF, Sun Y, Wloga D, Gay S, Klonowski KD, Wells L, Dougan ST**. 2009. O-GlcNAc modifications regulate cell survival and epiboly during zebrafish development. BMC Dev. Biol. **9**:28.

8. **Rahman MM, Stuchlick O, El-Karim EG, Stuart R, Kipreos ET, Wells L**. 2010. Intracellular protein glycosylation modulates insulin mediated lifespan in C.elegans. Aging. **2**:678–90.

9. **Selvan N, Mariappa D, van den Toorn HWP, Heck AJR, Ferenbach AT, van Aalten DMF**. 2015. The early metazoan Trichoplax adhaerens possesses a functional O-GlcNAc system. J. Biol. Chem. M114.628750–.

10. **Holt GD, Haltiwanger RS, Torres R, Hart GW**. 1987. Erythrocytes Contain Cytoplasmic Glycoproteins. J Biol Chem **262**:14847–14850.

11. **Hart GW, Haltiwanger RS, Holt GD, Kelly WG**. 1989. Glycosylation in the nucleus and cytoplasm. Annu. Rev. Biochem. **58**:841–74.

12. **Kelly W, Hart G**. 1989. Glycosylation of chromosomal proteins: Localization of O-linked N-acetylglucosamine in drosophila chromatin. Cell. **57**:243-251.

13. **Jacobsen SE, Binkowski K, Olszewski NE**. 1996. SPINDLY, a tetratricopeptide repeat protein involved in gibberellin signal transduction in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. **93**:9292–6.

14. **Vaidyanathan K, Durning S, Wells L**. 2014. Functional O-GlcNAc modifications: Implications in molecular regulation and pathophysiology. Crit. Rev. Biochem. Mol. Biol. **49**:140–163.

15. **Love DC, Hanover JA**. 2005. The hexosamine signaling pathway: deciphering the “O-GlcNAc code”. Sci. STKE **2005**:re13.

16. **Hart GW**. 2014. Nutrient Regulation of Cellular Metabolism & Physiology by O-GlcNAcylation. J. Biol. Chem. **289**:34422–23.

17. **Ma Z, Vosseller K**. 2014. Cancer Metabolism and Elevated O-GlcNAc in Oncogenic Signaling. J. Biol. Chem. **289**:34457–65.

18. **Yuzwa S a, Cheung AH, Okon M, McIntosh LP, Vocadlo DJ**. 2014. O-GlcNAc modification of tau directly inhibits its aggregation without perturbing the conformational properties of tau monomers. J. Mol. Biol. **426**:1736–52.

19. **Jínek M, Rehwinkel J, Lazarus BD, Izaurralde E, Hanover JA, Conti E**. 2004. The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. Nat. Struct. Mol. Biol. **11**:1001–7.

20. **Lazarus BD, Roos MD, Hanover JA**. 2005. Mutational analysis of the catalytic domain of O-linked N-acetylglucosaminyl transferase. J. Biol. Chem. **280**:35537–44.

21. **Martinez-Fleites C, Macauley MS, He Y, Shen DL, Vocadlo DJ, Davies GJ**. 2008. Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation. Nat. Struct. Mol. Biol. **15**:764–5.

22. **Clarke AJ, Hurtado-Guerrero R, Pathak S, Schüttelkopf AW, Borodkin V, Shepherd SM, Ibrahim AFM, van Aalten DMF**. 2008. Structural insights into mechanism and specificity of O-GlcNAc transferase. EMBO J. **27**:2780–8.

23. **Lazarus MB, Nam Y, Jiang J, Sliz P, Walker S**. 2011. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. Nature **469**:564–7.

24. **Schimpl M, Zheng X, Borodkin VS, Blair DE, Ferenbach AT, Schüttelkopf AW, Navratilova I, Aristotelous T, Albarbarawi O, Robinson D, Macnaughtan M, van Aalten DMF**. 2012. O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis. Nat. Chem. Biol. **8**:969–74.

25. **Lubas W, Hanover J**. 2000. Functional expression of O-linked GlcNAc transferase. Domain structure and substrate specificity. J. Biol. Chem. **275**:10983–8.

26. **Lazarus BD, Love DC, Hanover JA**. 2006. Recombinant O-GlcNAc transferase isoforms: identification of O-GlcNAcase, yes tyrosine kinase, and tau as isoform-specific substrates. Glycobiology **16**:415–21.

27. **Silverstone AL, Tseng T-S, Swain SM, Dill A, Jeong SY, Olszewski NE, Sun T-P**. 2007. Functional analysis of SPINDLY in gibberellin signaling in Arabidopsis. Plant Physiol. **143**:987–1000.

28. **Sokol KA, Olszewski NE**. 2015. The putative eukaryotic-like O-GlcNAc transferase of the cyanobacterium Synechococcus elongatus PCC7942 hydrolyzes UDP-GlcNAc and is involved in multiple cellular processes. J. Bacteriol. **197**:354–361.

29. **Sleytr UB, Thorne KJ**. 1976. Chemical characterization of the regularly arranged surface layers of Clostridium thermosaccharolyticum. J. Bacteriol. **126**:377-383.

30. **Upreti RK, Kumar M, Shankar V**. 2003. Bacterial glycoproteins: functions, biosynthesis and applications. Proteomics **3**:363–79.

31. **Nothaft H, Szymanski CM**. 2010. Protein glycosylation in bacteria: sweeter than ever. Nat. Rev. Microbiol. **8**:765–78.

32. **Schmidt MA, Riley LW, Benz I**. 2003. Sweet new world: glycoproteins in bacterial pathogens. Trends Microbiol. **11**:554–561.

33. **Schirm M, Kalmokoff M, Aubry A, Thibault P, Sandoz M, Logan SM**. 2004. Flagellin from Listeria monocytogenes Is Glycosylated with β -O-Linked N-Acetylglucosamine. J. Bacteriol. **186**:6721–27.

34. **Scott AE, Twine SM, Fulton KM, Titball RW, Essex-Lopresti AE, Atkins TP, Prior JL**. 2011. Flagellar glycosylation in Burkholderia pseudomallei and Burkholderia thailandensis. J. Bacteriol. **193**:3577–87.

35. **Iwashkiw JA, Vozza NF, Kinsella RL, Feldman MF**. 2013. Pour some sugar on it: the expanding world of bacterial protein O-linked glycosylation. Mol. Microbiol. **89**:14–28.

36. **Shen A, Kamp HD, Gründling A, Higgins DE**. 2006. A bifunctional O-GlcNAc transferase governs flagellar motility through anti-repression. Genes Dev. **20**:3283–3295.

37. **Tashima Y, Stanley P**. 2014. Antibodies that detect O-GlcNAc on the extracellular domain of cell surface glycoproteins. J. Biol. Chem. **289**:11132–11142.

38. **Zilliges Y, Kehr J-C, Mikkat S, Bouchier C, de Marsac NT, Börner T, Dittmann E**. 2008. An extracellular glycoprotein is implicated in cell-cell contacts in the toxic cyanobacterium Microcystis aeruginosa PCC 7806. J. Bacteriol. **190**:2871–9.

39. **Lazarus MB, Jiang J, Gloster TM, Zandberg WF, Whitworth GE, Vocadlo DJ, Walker S**. 2012. Structural snapshots of the reaction coordinate for O-GlcNAc transferase. Nat. Chem. Biol. **8**:966–8.

40. **Baker JL, Çelik E, DeLisa MP**. 2013. Expanding the glycoengineering toolbox: the rise of bacterial N-linked protein glycosylation. Trends Biotechnol. **31**:313–23.

41. **Iyer SPN, Hart GW**. 2003. Roles of the tetratricopeptide repeat domain in O-GlcNAc transferase targeting and protein substrate specificity. J. Biol. Chem. **278**:24608–16.

42. **Wang Z, Udeshi ND, Slawson C, Compton PD, Sakabe K, Cheung WD, Shabanowitz J, Hunt DF, Hart GW**. 2010. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. Sci. Signal. **3**:ra2.

43. **Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O**. 2011. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu. Rev. Biochem. **80**:825–58.

44. **Dennis RJ, Taylor EJ, Macauley MS, Stubbs KA, Turkenburg JP, Hart SJ, Black GN, Vocadlo DJ, Davies GJ**. 2006. Structure and mechanism of a bacterial beta-glucosaminidase having O-GlcNAcase activity. Nat. Struct. Mol. Biol. **13**:365–71.

45. **Copeland RJ, Han G, Hart GW**. 2013. O-GlcNAcomics-Revealing roles of O-GlcNAcylation in disease mechanisms and development of potential diagnostics. Proteomics. Clin. Appl. 1–10.

46. **Werner TP, Amrhein N, Freimoser FM**. 2005. Novel method for the quantification of inorganic polyphosphate (iPoP) in Saccharomyces cerevisiae shows dependence of iPoP content on the growth phase. Arch. Microbiol. **184**:129–36.

47. **Silverstone AL, Tseng T-S, Swain SM, Dill A, Jeong SY, Olszewski NE, Sun T-P**. 2007. Functional analysis of SPINDLY in gibberellin signaling in Arabidopsis. Plant Physiol. **143**:987–1000.

48. **Zachara NE, Vosseller K, Hart GW**. 2011. Detection and analysis of proteins modified by o-linked N-acetylglucosamine.Current protocols in molecular biology.ed. Frederick M. Ausubel. 12.8:12.8.1–12.8.33

49. **Schirm M, Kalmokoff M, Aubry A, Thibault P, Sandoz M, Logan SM**. 2004. Flagellin from Listeria monocytogenes Is Glycosylated with β -O-linked N-acetylglucosamine. J. Bacteriol. **186**:6721-6727.

50. **Waterbury JB, Willey JM, Franks DG, Valois FW, Watson SW**. 1985. A cyanobacterium capable of swimming motility. Science **230**:74–6.

51. **Palenik B, Brahamsha B, Larimer FW, Land M, Hauser L, Chain P, Lamerdin J, Regala W, Allen EE, McCarren J, Paulsen I, Dufresne A, Partensky F, Webb EA, Waterbury J**. 2003. The genome of a motile marine Synechococcus. Nature **424**:1037–42.

52. **Ehlers KM, Samuel AD, Berg HC, Montgomery R**. 1996. Do cyanobacteria swim using traveling surface waves? Proc. Natl. Acad. Sci. **93**:8340–8343.

53. **Ehlers KM, Koiller J**. 2011. Could cell membranes produce acoustic streaming? Making the case for Synechococcus self-propulsion. Math. Comput. Model. **53**:1489–1504.

54. **Wolgemuth CW, Igoshin O, Oster G**. 2003. The Motility of Mollicutes. Biophys. J. **85**:828–842.

55. **Ehlers K, Oster G**. 2012. On the mysterious propulsion of Synechococcus. PLoS One **7**:e36081.

56. **Mccarren J, Brahamsha B**. 2005. Transposon Mutagenesis in a Marine Synechococcus Strain : Isolation of Swimming Motility Mutants. Society **187**:4457–4462.

57. **Brahamsha B**. 1996. An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium Synechococcus. Proc. Natl. Acad. Sci. **93**:6504–6509.

58. **Hsieh Y-J, Wanner BL**. 2010. Global regulation by the seven-component Pi signaling system. Curr. Opin. Microbiol. **13**:198–203.

59. **Clerico EM, Ditty JL, Golden SS**. 2007. Specialized techniques for site-directed mutagenesis in cyanobacteria. Methods Mol. Biol. **362**:155–71.

60. **Wurst H, Shiba T, Kornberg A**. 1995. The gene for a major exopolyphosphatase of Saccharomyces cerevisiae. J. Bacteriol. **177**:898–906.

61. **Stam W, Stulp B**. 1988. New taxonomic methods: DNA/DNA hybridization. Methods Enzym. **167**:125–32.

62. **Teo CF, Ingale S, Wolfert MA, Elsayed GA, Nöt LG, Chatham JC, Wells L, Boons G-J**. 2010. Glycopeptide-specific monoclonal antibodies suggest new roles for O-GlcNAc. Nat. Chem. Biol. **6**:338–43.

63. **Wells L**. 2002. Mapping Sites of O-GlcNAc Modification Using Affinity Tags for Serine and Threonine Post-translational Modifications. Mol. Cell. Proteomics **1**:791–804.

64. **Tytgat HLP, Lebeer S**. 2014. The sweet tooth of bacteria: common themes in bacterial glycoconjugates. Microbiol. Mol. Biol. Rev. **78**:372–417.

65. **Giltner CL, Nguyen Y, Burrows LL**. 2012. Type IV pilin proteins: versatile molecular modules. Microbiol. Mol. Biol. Rev. **76**:740–72.

66. **Bhaya D, Bianco NR, Bryant D, Grossman A**. 2000. Type IV pilus biogenesis and motility in the cyanobacterium Synechocystis sp. PCC6803. Mol. Microbiol. **37**:941–951.

67. **Rollefson JB, Stephen CS, Tien M, Bond DR**. 2011. Identification of an extracellular polysaccharide network essential for cytochrome anchoring and biofilm formation in Geobacter sulfurreducens. J. Bacteriol. **193**:1023–33.

68. **Tan RM, Kuang Z, Hao Y, Lee F, Lee T, Lee RJ, Lau GW**. 2015. Type IV Pilus Glycosylation Mediates Resistance of Pseudomonas aeruginosa to Opsonic Activities of the Pulmonary Surfactant Protein-A. Infect. Immun. **6**:227–239.

69. **Nguyen LC, Taguchi F, Tran QM, Naito K, Yamamoto M, Ohnishi-Kameyama M, Ono H, Yoshida M, Chiku K, Ishii T, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y**. 2012. Type IV pilin is glycosylated in Pseudomonas syringae pv. tabaci 6605 and is required for surface motility and virulence. Mol. Plant Pathol. **13**:764–74.

70. **Craig L, Li J**. 2008. Type IV pili: paradoxes in form and function. Curr. Opin. Struct. Biol. **18**:267–77.

71. **Hwang J, Bieber D, Ramer SW, Wu C-Y, Schoolnik GK**. 2003. Structural and Topographical Studies of the Type IV Bundle-Forming Pilus Assembly Complex of Enteropathogenic Escherichia coli. J. Bacteriol. **185**:6695–6701.

72. **Harvey H, Habash M, Aidoo F, Burrows LL**. 2009. Single-residue changes in the C-terminal disulfide-bonded loop of the Pseudomonas aeruginosa type IV pilin influence pilus assembly and twitching motility. J. Bacteriol. **191**:6513–24.

73. **Horzempa J, Comer JE, Davis SA, Castric P**. 2006. Glycosylation substrate specificity of Pseudomonas aeruginosa 1244 pilin. J. Biol. Chem. **281**:1128–36.

74. **Egge-Jacobsen W, Salomonsson EN, Aas FE, Forslund A-L, Winther-Larsen HC, Maier J, Macellaro A, Kuoppa K, Oyston PCF, Titball RW, Thomas RM, Forsberg Å, Prior JL, Koomey M**. 2011. O-linked glycosylation of the PilA pilin protein of Francisella tularensis: identification of the endogenous protein-targeting oligosaccharyltransferase and characterization of the native oligosaccharide. J. Bacteriol. **193**:5487–97.

75. **Dorn KM, Fankhauser JD, Wyse DL, Marks MD**. 2015. A draft genome of field pennycress (Thlaspi arvense) provides tools for the domestication of a new winter biofuel crop. DNA Res. **2**:121-131.

76. **Black WP, Yang Z**. 2004. Myxococcus xanthus Chemotaxis Homologs DifD and DifG Negatively Regulate Fibril Polysaccharide Production. J. Bacteriol. **186**:1001–1008.

77. **Ahlawat YS, Pant RP, Lockhart BEL, Srivastava M, Chakraborty NK, Varma A**. 1996. Association of a Badnavirus with Citrus Mosaic Disease in India. Plant Dis. **80**:590–592.

Appendix I - Toxicity of SeOGT in E.coli

There is anecdotal evidence of toxicity of SPY-like OGTs in *E.coli* cloning strains, but no documented evidence (Olszewski, personal comm.). Upon multiple cloning attempts of SeOGT, very few transformants were observed in several *E.coli* cell lines. Additionally, numerous colonies contained plasmids with mutated or shuffled SeOGT constructs, which presumably resulted in inactive SeOGT. *E.coli* does not have any SPY-like OGT homologs. To investigate the possibility that SeOGT is toxic to *E.coli*, Rebeccca Choi, an undergraduate assistant from Macalaster College (St.Paul, MN) undertook several experiments to determine whether introduction of mutant SeOGT into *E.coli* results in more transformants than wild type SeOGT.

The Gateway plasmid pDONR SeOGT described in Chapter Two, including the SeOGT plus 500bp upstream to include native promoter, was used for wild type transformations. The previously described “active site lid” mutant corresponding to the *spy12* mutation SeOGT, G350D, was used for mutant transformations. Electro-competent *E. coli* strains DH5α and MC1061 were electroporated using 400ng DNA, and dilutions were plated (10^{-1} to 10^{-4}) on selective media. Following an overnight incubation at 37°C, colonies were counted on appropriate dilution plates to determine Colony Forming Units per milliliter of electroporation (CFUs/mL). Restriction digests were performed using PvuII enzyme (NEB) on

two wild type and two mutant transformants to determine whether the resulting plasmids had altered sequences.

	DH5α	MC1061
WT SeOGT	40	$9.9x10^5$
G350D mutant	60	$>1x10^7$

Table A1. Colony Forming Units per milliliter of electroporated culture are over 10,000 fold lower for wild type SeOGT compared to mutant SeOGT in MC1061, and slightly lower in DH5α, which shows significantly fewer overall transformants.

Wild type SeOGT with native promoter results in over 10,000 fold fewer transformants after electroporation. Additionally, genetic shuffling was observed in both of the wild type transformants (data not shown), resulting in mutant SeOGT genes within the vector. These results support previous observations that a wild type SPY-like OGT decreases transformation efficiency and results in shuffled gene sequences. Although cultures with intact SeOGT sequence can be recovered and maintained for cloning purposes, it appears that SeOGT is toxic to *E. coli*. The mechanism for this toxicity remains unknown, as *E. coli* does not have a native OGT, and therefore, most likely does not have native substrates. Because SeOGT has been shown to have hydrolytic activity against the UDP-GlcNAc sugar donor, and the G350D mutation decreases this activity (28), wild

type enzyme may be hydrolyzing the native UDP-GlcNAc pool. Because bacterial cell surfaces are sites of high concentration of glycolipids containing GlcNAc moieties, the depletion of the UDP-GlcNAc sugar donor pool may be affecting cell walls and decreasing viability.

Appendix II - Other experiments

Glycostaining of cellular proteins

100 μL of high density cultures was centrifuged and resuspended in 1 mL PBS. Cultures were sonicated on ice (1 second on, 4 seconds off for 25 pulses). Proteins from soluble, insoluble, and thylakoid fractions of wild type, Δ*ogt*, and Δ*ogt*, SeOGT were run on 12% SDS-PAGE gels, one of which was Coomassie stained and the other subject to glycostaining. Glycostaining was performed according to manufacturer's protocol for Pro-Q Emerald 488 Glycoprotein Gel and Blot Stain Kit (Molecular Probes). No differences in protein glycosylation were detected between any of the genotypes.

Xanthomonas campestris OGT does not rescue Δogt

The OGT from *Xanthomonas campestris* has been crystallized and shows homology to the human OGT as well as the SeOGT(see Chapter One for review). XcOGT was cloned and introduced into the Δ*ogt* strain using previously described methods (Chapter Two). Transformants were genotyped and tested for rescue of the settling phenotype. No rescue was observed.

Appendix III - List of Strains Created

KS#	Description	Date	Strain	Selection	Notes
1	pENTR Synecho NSII 5'	9/6/09			
2	pENTR Synecho NSII 3'	9/6/09			
3	pENTR 3' SeSPY	11/14/08			
4	pDONR221 SeSPY+AadI	7/9/09l	a-sel	Sp	
5	"	7/31/09	top10		
6	pDONR221 SePY+Promoter+AadI				
7	pMON772/NEO201	8/19/09			
8	pENTR SeSPY+/aadA1 for knockout (pDONR 221)	9/1/09			
9	pDEST SeSPY+promoter+aadI	9/13/09		Sp	
10	pRARE SeSPY in BL21AI for expression	6/2/10	BL21AI	Cm	
11	pDONR221 SeSPY12A (G350A)	6/24/10			
12	pDONR221 SeSPY12D (G350D)	6/24/10			
13	pDONR2-3 Cm-NSII 3' for complementation	10/21/10	MC1061	Cm Km	
14	pRARE SeSPY MBP in BL21AI used for expression 2010-12	12/1/10	BL21AI	Cm Km	
15	pRARE GUS-MBP from Katie (same as above)	12/1/10	BL21AI	Cm	
16	pACYC neg contol - no MBP (from -80 freezer stock of comp cells - AtSEC?)			Cm	
17	pRARE MAL HsOGT from JO342 stock	12/10/10	BL21AI	Cm	NEO 1360
18	pDONR NSII 3 + aadaI	12/30/10		Sp	
19	pENTR SeSPY 5"	11/8/10			
20	pCOLA AtSPY	2/2/11	BL21AI	Km	
21	pCOLA AtSEC	2/2/11	BL21AI	Km	
22	pCOLA AtSEC spy12 mutation	2/2/11	BL21AI	Km	
23	pDEST XcOGT with spec marker from NEO	2/25/11	MC1061		
24	pRARE HsPLBD* not expressed?	3/1/11	BL21AI	Cm	
25	pRARE HsOGT	3/5/11	(?)	Cm	
26	pCOLA SEC9TPRs	3/5/11	MC1061	Km	
27	pCOLA SEC11TPRs	3/5/11	MC1062	Km	
28	pCOLA AtSEC (KS21) for expression	4/5/11	Rosetta	Km	
	pDONR SeOGT G350D mutant for expression				no promoter, using original primers from Neil,
29	BAD - TRUNCATED	6/15/11	MC1061	Km	8/2008 - post BP rxn
30	pRARE-MBP-SPY	6/18/11	Bl21AI	Chl	Use Chl25 for expression - NEO1303
31	pRARE-MBP-SEC	6/18/11	Bl21AI	Chl	NEO1316

KS#	Description	Date	Strain	Selection	Notes
32	pRARE-MBP-GUS	6/18/11	Bl21AI	Chl	NEP 1304
33	pRARE-MBP-GFP	6/20/11	DH5-a	Chl	fromJO150 negative control for hydrolysis assay (in BL21AIs!)
34	pRARE-MBP-SeOGT G350D GOOD	6/24/11	DH5-a	Chl	
35	pRARE-MBP-GFP	7/22/11	BL21AI	Chl	for expression
36	pRARE-MBP-SeOGT G350D **GOOD**	8/18/11	BL21AI	Chl	for expression
37	pCR 2.1 PPX1	11/4/11			
38	pet11C	11/4/11	DH5a	Amp 75	
39	pDONR221 PPX1	11/18/11	DH5-a	Km	for entry into Gateway expression vectors
40	pET32a PPX1	12/4/11	TOP10	Amp	
41	HiMAR Minimariner	1/27/12	WM3064	Km	In mating strain from Jeff Gralnick - needs 10uL/mL 60mM DAP!!!
42	pDONR221 aadA1	2/5/12		Km	
43	pDEST aadA1 SeSPY 5' 3'	2/10/12	Top10	Amp/Sp	For making new SeOGT knockout with a spec marker in place of OGT
44	pDONR SeOGT H280A	2/20/12	Top10	Km	
45	pDEST H280A	2/26/12	Top10	Amp	
46	pDONR SeOGT K445A	4/16/12	Top10	Km	
47	pSyn SeOGT	5/27/12	Top10	Sp 75	for inducible expression in Synechococcus - see kit
48	Synecho knockout - spec	5/2/12			
49	pENTR SeOGT	6/8/12	Top10	Km	for flag tagged OGT
50	pENTR SeOGT-Flag BAD	7/3/12	Top10	Km	for Gateway cloning - BACKWARDS FLAG (Jackie's construct)
51	pDEST SeOGT-flag BAD	12-Aug	Top10	Amp	
52	pDONR SeOGT H280A - no stop	8/12/12	Top10		for pRARE expression - no stop
53	pENTR SeOGT H280A - no stop	12/18/12	Top10		sequenced
54	pENTR SeOGT K445A - no stop	12/18/12	Top10		sequenced
55	pRARE SeOGT H280A	12/20/12	Top10	Cm	
56	pRARE SeOGT K445A	12/20/12	Top10	Cm	
57	KS 55 for expression	1/2/13	BL21AI	Cm	
58	KS 56 for expression	1/2/13	BL21AI	Cm	
59	pRARE SeOGT for expression	1/2/13	BL21AI	Cm	
60	pET32 Nup62	3/20/13	BL21AI	AMP	for co-expression with SeOGT
61	XXXXXXXXXX				

KS#	Description	Date	Strain	Selection	Notes
62	pDONR 2-3 SpecR-NSII 3'	5/23/13	MC1061	Km/Sp	for easier Gateway cloning
63	pDONR pilA	10/24/13	Top10	Km	Not sequenced
64	pDEST SeOGT-Flag GOOD	12/13/13	Top10	Amp	good construct - made by NEO from KS51
65	pSyn PilA*	12/21/13	Top10	Sp	*Sequenced 8/18/14 - may be bad
66	pDONR PilA 3'	6/16/14	Top10	Km	for Gateway cloning
67	pDONR PilA 5'	6/18/14	Top10	Km	
69	pDEST pilA			Km	

KS-C	Cyanobacteria
1,4	PCC 7942 WT
5,8	Δogt::Km
9,12	Δogt::Km, SeOGT Spec
13, 16	Δogt SeOGT G250A
17,20	Δogt SeOGT G350D
21	Δogt, OGT-FLAG
	pSyn (Sp)
22	Δogt, H280A
23	Δogt, K445A
24	Δogt, pilA(A107D)
25	Δogt, pilA(A107D), OGT
26	Δogt, pilA(A107D), PilA
27	WT, pilA(A107D)