Characterization of TarA, a ToxT-activated small regulatory RNA produced by *Vibrio cholerae*

by

Aimee Lynne Richard

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in the University of Michigan
2012

Doctoral Committee:

Professor Victor DiRita, Chair
Professor David R. Engelke
Associate Professor Philip C. Hanna
Assistant Professor Eric S. Krukonis
Assistant Professor Maria B. Sandkvist
Professor Stephen Lory, Harvard University

To my parents, who have supported me in everything I do.

Acknowledgements

Thank you Vic, first for recruiting me to Michigan, and then for welcoming me to your lab, pushing me to succeed at the highest level, and being an excellent mentor and friend. You've taught me a lot over the past five years, both scientifically and personally. I truly appreciate everything you've done for me. I thank all of my committee members, who have provided guidance, excellent suggestions, and support throughout my graduate school career.

I would also like to thank all of the past and present members of the DiRita lab that I've interacted with over the years. Our lab has been a great place to do science and to collaborate, and I will certainly miss it after I leave. Jeff, thank you for leaving me an awesome project and for being my personal Yoda for all of grad school. Whether it was how to do a Northern blot, where to park for a Tigers game, or anything in between, you've always gone above and beyond to answer my questions, even the dumb ones. You've been a great friend and colleague. Jyl, thank you for making the last five years of my life much better. Having you around to bounce ideas off of, show weird data to, and (occasionally) gossip with has truly been a gift. JJ, thanks for providing me with an automatic enemy in the form of a Yankees fan. All kidding aside, thanks for rocking out with me and keeping me in line. Andy and Mike, thank you for being longstanding members of team penismightier with me. Please keep it alive after I leave!

Becca, Wei Ping, and Jessica, I'm confident the DiRita lab will continue its awesomeness with you three in it. Keep doing great research and hang in there, there is a light at the end of the tunnel of grad school! Finally, Mariah, I wish you had shown up in my life years ago. You've been an amazing friend over the past few months, and I know you'll do great things in grad school, in lab, and in life. Thank you for listening to me complain, making me laugh, and just getting me more than most other people.

To all of the friends I've made in Michigan, whether it was through school, brewing, judging, or some other means, thank you for keeping me balanced. I appreciate and will miss all of you. Phil, you have kept me sane and grounded over the past few months and I've enjoyed every minute of it. Thank you for challenging me and making me a better person. Cherisse, thank you for everything. My life was a lot more exciting with you in it. Dave, through thick and thin, you taught me a lot and have been an excellent friend. Thank you.

Finally, I would like to thank my amazing family. Mom and Dad, thank you for your everlasting love and support. You two are an excellent example of how to work hard, yet still have enough time for your family and friends. You are the best role models anyone could ask for. Adam, we haven't always seen eye to eye, but you're a great brother and friend. Thank you for always finding a way to make me laugh. Allivia, you've grown up so fast and I'm so proud of you. Thank you for always being there for me. To all of my aunts, uncles, cousins, and

grandparents, thank you all for your love and support. It's meant a lot to me, and I wouldn't be the person I am today without you.

Table of Contents

List of Figures

List of Tables

Abstract

Vibrio cholerae produces several virulence factors to establish infection, most of which are regulated by the ToxR regulatory pathway. The direct activator of virulence gene transcription is the protein ToxT. TarA is a ToxT-activated small, regulatory RNA (sRNA) that is conserved throughout toxigenic *V. cholerae* strains. It lacks open reading frames and has a putative binding site for the bacterial regulatory protein Hfq, which is implicated in the function of several sRNAs. Based on these observations, we hypothesized that TarA has a regulatory function. Microarray analysis of a *tarA* mutant and wild type *V. cholerae* revealed increased RNA for *ptsG*, which encodes the major glucose transporter. This result was confirmed, and we further demonstrated decreased *ptsG* transcript levels in response to TarA induction. Additionally, *V. cholerae* overexpressing TarA take up less glucose and grow poorly in glucose-containing media. A working hypothesis is that TarA is stabilized by the protein Hfq and regulates *ptsG* by base-pairing with the 5' untranslated region of the *ptsG* transcript, which targets the transcript for degradation by RNase E. Analysis by electrophoretic mobility shift assays demonstrated that TarA binds the *ptsG* leader sequence *in vitro.* Additionally, a *V. cholerae* mutant lacking *hfq* has lower steady state levels of TarA, presumably due to more rapid turnover in the

absence of its interaction partner. Mutational analysis suggests that TarA possesses at least two redundant Hfq binding sites.

TarA appears to regulate several metabolic genes in addition to *ptsG*, including the sialic acid catabolism genes found at the *nan* locus. In the absence of TarA, *nan* transcripts are diminished, suggesting that TarA activates these genes. TarA may work in conjunction with the repressor protein NanR to precisely control expression of these genes. These data lead to the hypothesis that TarA regulates *V. cholerae* adaptation to the changing nutrient conditions it encounters, from aqueous reservoirs in the environment to the intestine of the host.

Chapter I

Introduction

Cholera: past and present

Cholera is a waterborne gastrointestinal illness that has afflicted humans throughout recorded history. This disease is characterized by massive fluid loss through secretory diarrhea, which contains high levels of mucus and resembles rice water. If untreated, the severe dehydration caused by cholera is often fatal. While cases of cholera have been documented since 500 to 400 B.C.E., these were mainly limited to sporadic outbreaks in isolated areas prior to the early 19th century (Lacey, 1995). Cholera is endemic to the Indian subcontinent and has spread worldwide to cause seven pandemics. The exact dates of the pandemics are somewhat controversial; in many cases, historical records are incomplete and tracking down the first and last case of cholera in a particular pandemic is impossible. For a timeline of the major events in the history of cholera, see Figure 1.1.

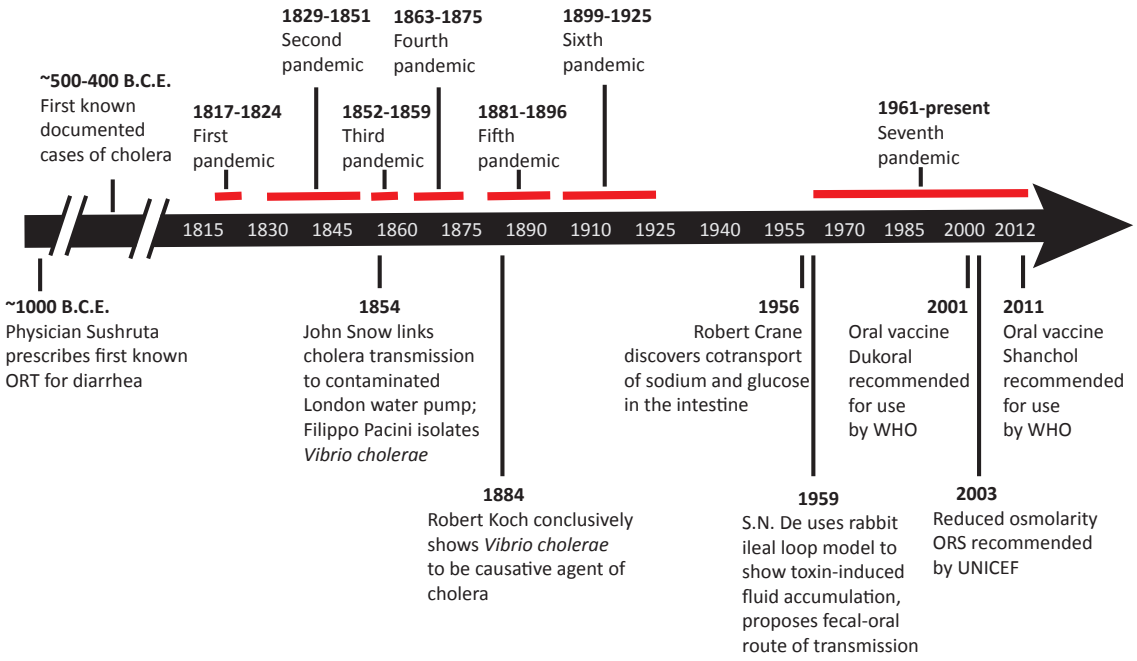


Figure 1.1. Timeline of major events in the history of cholera.

It is widely accepted that the first cholera pandemic started in 1817 with the advent of international trade (Sack *et al.*, 2004). After a season of unusually heavy rainfall and massive flooding in India, cholera sickened hundreds of thousands, with a focus in the province of Bengal, which included the city of Calcutta (Kolkata). British soldiers stationed in this area fell victim to the illness and subsequently transmitted it to Nepalese soldiers against whom they were fighting. From there, the disease reached the countries of China, Malaysia, and Japan via merchant ships. This first pandemic lasted until 1824, when an abnormally cold winter halted spread of the disease (Pollitzer, 1954).

The second pandemic, spanning from 1829 to 1851, also originated in India. In this instance, cholera was spread by Polish and Russian soldiers who carried the disease into their home countries. By 1831, cholera had reached Austria and England, and the following year, Ireland, France, and Belgium all experienced epidemics. In the same year, an outbreak of cholera in Mecca sickened thousands of pilgrims, who brought the illness back to Syria, Palestine, and Egypt. Increased travel and immigration further facilitated explosion of the disease, which reached Scandinavia and crossed the Atlantic Ocean in 1832, entering North America through the cities of Quebec and New York (Pollitzer, 1954). In the course of the second pandemic, cholera reached every continent except Antarctica (Lacey, 1995).

After a brief respite, pandemic cholera returned in 1852, following a severe outbreak of the disease in India. The first years of this pandemic were disastrous, with the United States, Canada, Mexico, the West Indies, England, France, Greece, and Turkey being the most hard-hit countries. The spread of cholera to the latter two was largely blamed on the Crimean War (1853 to 1856), as infected troops traveled from southern France to these countries (Pollitzer, 1954). Meanwhile, in England, an outbreak of disease in a neighborhood of London led John Snow, a physician, to link the transmission of cholera with ingestion of contaminated water (Snow, 1855), a pioneering observation in public health epidemiology. This important discovery had lasting effects; from that point on, England suffered far fewer cholera cases than it had in the past, even as the disease was repeatedly brought into the country. The causative agent of cholera, the Gram-negative bacterium *Vibrio cholerae*, was isolated in the same year by Filippo Pacini, though Robert Koch is widely credited with making this breakthrough in 1884 (Lacey, 1995). Within the next few years, the number of cases worldwide dwindled, and the third pandemic was largely over by 1859 (Pollitzer, 1954).

The fourth pandemic, lasting from 1863 until approximately 1875, marked a shift in the routes of transmission of cholera. New trade routes had been established at this time, providing a means for the disease to be spread even more efficiently. In particular, 1865 was an especially important year, as it was a jubilee year in Mecca, where scores of pilgrims gathered. An outbreak in Mecca

led to massive numbers of cases, and many of these infected people carried cholera to their homelands of Egypt, Syria, Palestine, Mesopotamia, and others. The spread of disease to Egypt was especially fateful, as cholera took hold in Alexandria, a major shipping center on the Mediterranean Sea. This facilitated transmission to other port cities and re-introduced cholera to Europe, where it ravaged several countries, including Austria, Hungary, Germany, the Netherlands, Belgium, and Russia. Later in the decade, the disease gripped northern and eastern Africa, concomitant with the rise of Mozambique and Zanzibar as shipping centers (Pollitzer, 1954).

The fourth pandemic also affected North America, reaching New York City by sea in 1865 or 1866. It spread rapidly once there, traveling by the ever-expanding rail system and reaching as far west as Kansas. In response to the rampant spread of the disease in New York, the nation's first Board of Health was founded in 1866 (Duffy, 1971). While some measure of disease control was taking place in the United States, cholera was spreading deeper into South America at that time, reaching Argentina, Peru, Brazil, and Uruguay. From a global standpoint, the disease was still very much a problem, even as the fourth pandemic waned in the 1870's.

In 1884, as the fifth pandemic raged on, Robert Koch conclusively linked the cause of cholera to infection with the bacterium *V. cholerae* (Koch, 1884). This discovery did not immediately lead to preventative measures, as France, Italy, and Spain saw epidemics throughout the mid-1880's. Of note, cholera was

re-introduced to both England and New York City several times during this period, but there were no new outbreaks due to proper quarantine techniques and improved water sanitation. Other areas of the world were not so lucky; the fifth pandemic claimed hundreds of thousands more lives and persisted until 1896 (Pollitzer, 1954).

As the twentieth century began, severe outbreaks of cholera in Calcutta and Bombay sparked the beginning of the sixth pandemic, which lasted until approximately 1923. This pandemic was marked by increased spread of the disease due to World War I and saw massive numbers of cases in Europe and Russia, with a peak of infection in 1921. In the following years, fewer infections were noted, with Europe being largely free of cholera after 1925 (Pollitzer, 1954).

The seventh pandemic, according to most sources, began in 1961 and continues today (WHO, 2011). This pandemic differed from the first six in several ways. While prior pandemics had originated in severe outbreaks in India, the seventh began in Indonesia and spread from there. Additionally, in contrast to the previous six pandemics, which were caused by *V. cholerae* of the classical biotype, the bacteria responsible for the current pandemic belong to the El Tor biotype, which is generally considered less virulent than its classical relatives. Nonetheless, cholera claimed hundreds of thousands of lives as it spread through Asia, Europe, and Africa, eventually reaching Peru by freighter in 1991. In the late 1980's, studies by the Environmental Protection Agency linking chlorination byproducts with cancer had led Peruvian officials to stop chlorinating

some of the wells in Lima. Thus, when wastewater contaminated with *V. cholerae* entered the main water supply, the disease spread easily and made this epidemic especially devastating (Anderson, 1991). Recent studies have suggested that while adequate chlorination of the water supply would not have prevented the epidemic entirely, it would have saved thousands of lives (Tickner & Gouveia-Vigeant, 2005).

While cholera remains a severe public health problem, it is now limited to developing countries with poor infrastructure, especially those that have recently experienced military conflicts and natural disasters. Cholera remains endemic in India and Bangladesh, and cases generally appear in waves centered around the monsoon season. These waves of cases are often limited by the presence of bacteriophages, or bacterial viruses, that prey on *V. cholerae* (Faruque *et al.*, 2005).

In the last few years alone, Zimbabwe, Iraq, Afghanistan, and Sudan, among others, have seen deadly epidemics of cholera (WHO, 2011). Most recently, Haiti, a country that had not recorded cases of cholera for more than a century, suffered a severe outbreak of disease following a devastating earthquake in January 2010 (Piarroux *et al.*, 2011). Sequence analysis of bacteria isolated from patients indicated that the disease had likely been imported into Haiti by Nepalese troops who arrived as part of post-earthquake humanitarian relief efforts (Chin *et al.*, 2011) and had spread through improperly treated drinking water. While the intensity of the epidemic has waned somewhat, it is expected

that the disease will linger in Haiti until water sanitation can be improved (Butler, 2011).

Symptoms and treatment

Cholera is characterized by the profuse watery diarrhea induced by secretion of cholera toxin into the intestinal lumen (Figure 1.2). This diarrhea, while painless and non-inflammatory, can lead to severe dehydration and death in a matter of hours, with patients losing up to a liter of fluid an hour. Vomiting is less common but does occur, contributing to the observed dehydration. In the absence of treatment, indicators of severe dehydration, such as sunken eyes, loss of skin elasticity, and barely detectable blood pressure manifest within hours. Untreated patients can lose up to ten percent of their body weight within 24 hours of the first symptoms, the period in which most deaths from cholera occur (Sack et al., 2004). In many clinics, cholera cots are utilized, which allow for the measurement of lost fluid for treatment purposes.

Figure 1.2. Photograph of rice water stool collected from a cholera patient. Photo courtesy of Jeffrey H. Withey.

Currently, treatment of cholera is quite effective and focuses on administration of Oral Rehydration Therapy, or ORT, the goal of which is to replace the massive amount of fluid lost by cholera patients. The composition of the fluid given is important, as patients lose both electrolytes and water to the secretory diarrhea of cholera; if only water is given, the patient will suffer severe hyponatremia. As glucose and sodium are co-transported across the intestinal epithelium, sodium absorption by the patient is much more efficient if glucose is present in the solution. While the most commonly used Oral Rehydration Salts (ORS) mixture refers to specific quantities of glucose and salts, the term ORT can be used for most liquids given to a cholera patient.

The idea of replacing fluids lost to diarrhea is not new; in fact, it dates back 3000 years to the Indian physician Sushruta, who prescribed a mixture of water, rock salt and molasses as a treatment for diarrhea (da Cunha Ferreira & Cash, 1990). While ORT in some form, be it rice water, coconut juice, carrot soup, or another fluid, has been used to treat cholera in India for the past three millennia, this development did not catch on in other countries until the mid-1900's (da Cunha Ferreira & Cash, 1990). The formulation most commonly used today (reduced osmolarity ORS) consists of 75 millimolar (mM) sodium chloride, 75 mM glucose, 20 mM potassium chloride, and 10 mM trisodium citrate; this specific combination has been most efficacious at treating severe dehydration (WHO, 2006). In the past, intravenous fluid replacement has also been used to treat cholera, while today it is only used for patients too ill to drink.

In most cases of cholera, antibiotics are ineffective because the disease symptoms are caused by cholera toxin, which is rapidly produced upon infection. Antibiotics target *V. cholerae* itself and by the time they are administered, high levels of toxin production have already occurred. Similarly, antibiotic treatment does not lessen transmission of disease, as the onset of diarrhea has already occurred, triggering shedding of the bacteria into wastewater. As such, antibiotics are now rarely used to treat cholera, and generally only lessen the severity and duration of disease in some cases (Lindenbaum *et al.*, 1967).

Cholera vaccines

While the discovery of *V. cholerae* as the causative agent of cholera prompted research into pathogenesis of the disease, progress was quite slow at the outset. In fact, it was not until 1959 that the transmission of cholera via the fecal-oral route was proposed by S.N. De (De, 1959). Previously, investigators had tried injecting *V. cholerae* into animals via virtually every other route and had not been able to recapitulate disease symptoms. De observed massive fluid accumulation after injecting *V. cholerae* culture filtrates into a ligated rabbit ileal loop. He proposed that this bacteria "multipl(ies) in the small bowel and produce(s) an exotoxin which acts upon the mucosal cells of the small bowel, causing them to secrete large quantities of isotonic fluid… and the result is a watery isotonic diarrhea"(De, 1959). Prior to De's discovery, immunization studies focused on injectable vaccines with limited success. The most widely

used of these preparations was a killed whole-cell vaccine developed in the late 1800's. This fairly expensive vaccine was extensively tested in Bangladeshi children in the 1960's and was found to provide limited protection while causing undesirable side effects (Mosley *et al.*, 1972). Given its poor efficacy and side effects, the vaccine is not recommended by the WHO.

Recently, focus has shifted to oral vaccination strategies, which have been more successful. Vaccines given by this route are especially valuable because they present antigen to the immune system in the same manner that an infection would. While both live attenuated vaccines and killed whole-cell preparations have been developed, the latter have been more fruitful, with greater protection and fewer side effects (for review see (Bishop & Camilli, 2011)). Several trials of oral live attenuated vaccines have been conducted in the US and abroad; one of the largest studies tested the strain CVD 103-HgR on nearly 70,000 volunteers in Indonesia. This trial was largely considered a failure because while only very few subjects developed cholera, the number of cholera cases were similar between the placebo and vaccine groups (Richie *et al.*, 2000). Other live attenuated vaccine strains that have been developed, such as Peru-3, have produced reactogenic diarrhea in patients (Taylor *et al.*, 1994). Recent work with the infant rabbit model of cholera disease has pinpointed the presence of flagellins as the cause of this inflammatory side effect (Rui *et al.*, 2010).

Dukoral, an oral vaccine consisting of recombinant cholera toxin B subunits and killed *V. cholerae* of the Ogawa and Inaba serotypes and classical

and El Tor biotypes, was developed in the 1980's. While relatively expensive and requiring two doses (three doses for young children), this vaccine was shown to be quite effective, providing 85% protection six months after administration in one study (Clemens *et al.*, 1986). Dukoral was licensed in 1991 and has been used in refugee camps with success (Legros *et al.*, 1999). Another oral vaccine, Shanchol, has recently been approved for use by the WHO. Shanchol lacks the recombinant cholera toxin B subunit, making it cheaper than Dukoral and thus more attractive for use in developing countries. While Shanchol also requires two doses, a third dose is not required for children under the age of six (WHO, 2010). It is a killed whole-cell formulation containing *V. cholerae* of both the O1 and O139 serotypes.

Though there are currently two viable vaccines on the market, research in this area continues in hopes of creating a preparation that demonstrates good protection in children, is cost-effective, and is easy to administer to large populations. Studies utilizing outer membrane vesicles from *V. cholerae* have shown particular promise, with these preparations providing immunity to neonatal mice when given to pregnant mice (Bishop & Camilli, 2011). Additionally, the recent re-emergence of an infant rabbit model of disease has provided an important tool for testing of new vaccine candidates (Rui et al., 2010). This model was first used extensively in the 1950's and 60's and has become more favored in recent years, as it recapitulates cholera symptoms well in a small animal (Finkelstein *et al.*, 1964, Dutta & Habbu, 1955). The field of cholera

vaccine research is thus still very active, with many recent and exciting developments towards a safe, effective vaccine.

Vibrio cholerae biotypes and serotypes

To date, over 200 different serotypes of *V. cholerae* have been isolated, but only two of these (O1 and O139) have been found to cause epidemic disease. The serotypes are classified based on the specific O-antigen found on the lipopolysaccharide (LPS), which is a component of the bacterial outer membrane. Within the O1 serotype, there are two subtypes, Ogawa and Inaba, and two biotypes, classical and El Tor. The Ogawa and Inaba isolates are distinguished by genetic differences in one specific part of the O antigen, while the two biotypes differ at both genetic and phenotypic levels. Classical strains of *V. cholerae* caused the first six pandemics of cholera; the current pandemic features an El Tor strain. While these distinctions are helpful in the laboratory, recent research has suggested that there is significant genetic fluidity between the two biotypes of *V. cholerae*. In fact, hybrid, or "El Tor variant" strains have been isolated from cholera patients in Bangladesh, Haiti, India, Thailand, and Malaysia, among other countries (Son *et al.*, 2011, Goel *et al.*, 2011, Na-Ubol *et al.*, 2011, Ang *et al.*, 2010). These strains are constantly evolving through horizontal gene transfer, leading to a diverse population of bacteria in the environment. One mechanism of gene acquisition in the environment may be natural competence, which is induced by chitin, a common component of marine

species with which *V. cholerae* associates in nature (Meibom *et al.*, 2005). The most recently identified genetic elements, VSP-I and VSP-II, or Vibrio seventh pandemic islands I and II, were discovered in the last decade (Dziejman *et al.*, 2002) and contain several yet-to-be-characterized genes that may have roles in the pathogenesis of these strains. These regions, specifically VSP-II, are highly variable and appear to undergo frequent genetic rearrangements (Taviani *et al.*, 2010). In sum, delineation of *V. cholerae* into two distinct biotypes may soon be, if it is not already, an antiquated concept.

Vibrio cholerae **and mechanism of infection/transmission**

Ingestion of water contaminated with *V. cholerae* gives the bacterium access to the gastrointestinal tract of the human host. The infectious dose of *V. cholerae* is quite high because the majority of the organisms are killed by the acidity of the stomach. Bacteria able to survive passage through this acidic environment reach the upper small intestine, where the production of toxin co-regulated pilus (TCP) aids in adherence to the mucus layer (Herrington *et al.*, 1988). While in close association with the intestinal mucosa, *V. cholerae* multiplies rapidly and secretes cholera toxin, which is responsible for the actual disease symptoms. Cholera toxin (CT) is an AB_5 toxin consisting of six subunits: an activity (A) subunit and five binding (B) subunits. Upon binding to GM1 ganglioside on the surface of a host epithelial cell, the toxin is internalized by

receptor-mediated endocytosis and transported to the endoplasmic reticulum (ER) via retrograde transport. In the ER, the catalytic (A1) portion of the A subunit is cleaved from the toxin complex; identified by the cell as a misfolded protein, it is then retrotranslocated to the cytosol for degradation (Tsai *et al.*, 2001). Usually, misfolded proteins in the cytosol are polyubiquitinated on lysine residues, triggering proteasomal degradation. However, the CT A1 subunit contains few lysine residues and re-folds quickly in the cytosol, effectively avoiding this fate (Rodighiero *et al.*, 2002).

Once free in the cytosol, cholera toxin ADP-ribosylates the s subunit of heterotrimeric G proteins, promoting adenylate cyclase activity, which leads to production of cyclic AMP (cAMP). cAMP is a potent second messenger, and its buildup in epithelial cells sparks phosphorylation of the cystic fibrosis transmembrane conductance regulator (CFTR) and activation of chloride channels in the plasma membrane, among other events (De Haan & Hirst, 2004). This triggers high levels of chloride secretion and mucus shedding into the intestinal lumen; recent studies have provided evidence that *V. cholerae* associates with these secreted clumps of mucin (Ritchie *et al.*, 2010). To maintain osmotic balance of the luminal fluid, water is also released, and the overwhelming abundance of fluid produces the secretory diarrhea experienced by cholera patients (De Haan & Hirst, 2004). Shedding of this stool, which is rich in aggregates of mucin and mucin-associated bacteria, thus promotes transmission of *V. cholerae* to environmental reservoirs and new hosts.

In order to efficiently colonize its host and promote spreading, *V. cholerae* tightly controls production of its virulence factors. This regulatory cascade has been extensively reviewed in the literature (Matson *et al.*, 2007). Briefly, the inner membrane-bound proteins ToxR and TcpP coordinately activate transcription of the gene encoding ToxT, an AraC/XylS-like transcriptional regulator (Higgins *et al.*, 1992). ToxT then binds toxboxes (consensus binding sequences) upstream of virulence genes, activating transcription at these sites (Withey & DiRita, 2006). Targets of ToxT include *ctxAB,* the *tcp* locus, and small regulatory RNAs (Richard *et al.*, 2010, Bradley *et al.*, 2011), among others. *V. cholerae* strains lacking *toxR, toxT*, or *tcpA* display severe colonization defects in both suckling mice and human volunteers (Herrington et al., 1988, Taylor *et al.*, 1987, DiRita *et al.*, 1991, Lee *et al.*, 1999), while *ctxA* mutants are able to colonize both human and infant rabbit hosts but do not induce secretory diarrhea (Ryan *et al.*, 2006, Rui et al., 2010).

A spatially and temporally regulated pattern of gene expression is crucial for successful *V. cholerae* infection and transmission. Studies using recombination-based in vivo expression technology (RIVET) have identified several genes induced during early and late infection of both human and infant mouse hosts that are required for colonization and disease spread, respectively (Merrell & Camilli, 2000, Lombardo *et al.*, 2007, Schild *et al.*, 2007). As one might expect, gene products promoting colonization, such as toxin co-regulated pilus and the accessory colonization factors, are produced in the initial stages of

infection. Later, once the bacteria have robustly colonized the bowel, the focus shifts to metabolic functions, such as storage of carbon in the form of glycogen (Bourassa & Camilli, 2009), in preparation for exit to the environment. Vibrios freshly shed from the host are hyperinfectious compared to those grown *in vitro*, suggesting that the genes expressed *in vivo* help equip the bacterium not only to survive in nutrient-poor environmental reservoirs, but also to colonize its next host effectively (Merrell *et al.*, 2002).

Regulation of carbon metabolism is also crucial for efficient colonization by *V. cholerae*. Like other bacteria, *V. cholerae* possesses an extensive phosphotransferase system (PTS) that allows it to grow on carbon sources such as glucose, fructose, mannitol, and chitobiose, among others. These PTS systems are specific to certain sugars and are generally comprised of four proteins: enzymes I, IIA, and histidine protein (HPr), which are found in the cytoplasm, and enzyme IIBC, which is the membrane-bound component of the system. The glucose-specific PTS system regulates biofilm formation in *V. cholerae* and has been well-studied (Houot et al., 2010a, Houot et al., 2010b, Houot & Watnick, 2008). Briefly, upon presence of glucose in the media, phosphoenolpyruvate donates a phosphate group to EI. Phosphorylated EI protein then transfers its phosphate group to the HPr, which in turn phosphorylates EIIA. The glucose-specific EIIBC protein, called PtsG in *V. cholerae,* then takes up a glucose molecule and transports it through the membrane, converting it to glucose-6-phosphate with the phosphate donated by

EIIA-P. This results in the production of glucose-6-phosphate, which can be metabolized for energy. Additionally, there is evidence that EI-P and EIIBC (PtsG) repress biofilm growth of *V. cholerae* by repressing transcription of the *vps* genes, thus promoting growth in the planktonic form.

In *E. coli, ptsG* is regulated by a small RNA, SgrS (Vanderpool & Gottesman, 2004). Under conditions of glucose phosphate stress, in which glucose-6-phosphate has accumulated to toxic levels, the regulatory protein SgrR activates transcription of SgrS. SgrS, a 227 nt small regulatory RNA, base-pairs with the leader sequence of the *ptsG* transcript, occluding the Shine-Dalgarno sequence and targeting the transcript for degradation by RNase E (Vanderpool & Gottesman, 2004). This interaction is catalyzed by the small regulatory protein Hfq (Kawamoto *et al.*, 2006). SgrS also encodes a small peptide, SgrT, which relieves glucose phosphate stress as well, presumably by interacting with PtsG at the protein level (Wadler & Vanderpool, 2007).

Small regulatory RNAs in *Vibrio cholerae*

Small regulatory RNAs (sRNAs), while recently discovered, have been increasingly recognized as important regulatory elements of prokaryotes and eukaryotes alike (Toledo-Arana *et al.*, 2007, Liu & Camilli, 2010, Esteller, 2011). As in most types of bacteria, sRNAs are a key component of gene regulation in *V. cholerae.* These RNA molecules are generally between 75 and 400

nucleotides in length and can act in *cis* or *trans* to regulate their targets at the RNA or protein level. Thirteen of these RNAs have been characterized (Lenz *et al.*, 2004, Lenz *et al.*, 2005, Davis *et al.*, 2005, Mey *et al.*, 2005, Davis & Waldor, 2007, Song *et al.*, 2008, Richard et al., 2010, Bradley et al., 2011, Mustachio *et al.*, 2012), while over 500 others have been predicted using bioinformatics and deep-sequencing approaches (Livny *et al.*, 2005, Liu *et al.*, 2009). Many sRNAs require the small hexameric protein Hfq for their function. Hfq binds AU-rich sequences on both sRNAs and their targets and often serves to both stabilize the sRNA and to catalyze base-pairing between the two RNA molecules.

The first small regulatory RNAs found in *V. cholerae* were those involved in quorum sensing, Qrr1-4 and CsrB, CsrC, and CsrD. The three Csr RNAs are among the largest sRNAs discovered and redundantly inhibit the protein CsrA when cell density is low. CsrA indirectly increases the activity of the protein LuxO, which activates transcription of the *qrr* genes. Once transcribed, Qrrs1-4, in conjunction with Hfq, bind the *hapR* transcript, destabilizing it (Lenz et al., 2004, Lenz et al., 2005). As HapR is an inhibitor of virulence factor production and biofilm formation, this results in activation of the virulence regulatory cascade.

Two other sRNAs, MicX and VrrA, affect composition of the *V. cholerae* membrane. MicX is processed into an active molecule of 189 nucleotides and represses VC0972, a putative outer membrane protein, and VC0620, a component of a putative ABC transporter, while positively regulating VCA1041, a

putative phosphomannomutase. While Hfq has been shown to stabilize this sRNA, presence of the protein is not necessary for binding of MicX to its targets (Davis & Waldor, 2007). Similarly, the sRNA VrrA, which binds the *ompA* mRNA to block translation of the protein, does not require Hfq for its activity. Deletion of *vrrA* results in overproduction of the porin OmpA and confers a competitive advantage in an infant mouse model of colonization (Song et al., 2008).

Small regulatory RNAs can also control metabolic functions (Richards & Vanderpool, 2011). The recently described sRNA MtlS acts in *cis* to regulate mannitol uptake through the protein MtlA. MtlS is encoded antisense to the *mtlA* gene and is hypothesized to directly base-pair with the 5' untranslated region (UTR) of the *mtlA* RNA, blocking translation of the MtlA protein. This regulation occurs in the presence of sugars and sugar alcohols other than mannitol; in the presence of mannitol, MtlS is not produced, and the *mtlA* mRNA can be translated (Mustachio et al., 2012). Another sRNA, TarB, is transcribed under ToxT-inducing conditions and controls synthesis of a novel cyclic dinucleotide. Strains lacking TarB, which are therefore deficient in synthesis of this molecule, are attenuated for colonization of the suckling mouse intestine (Davies et al., 2012). TarB also appears to negatively regulate *tcpF*, a gene found on the VPI that encodes a secreted colonization factor (Bradley et al., 2011). It is becoming increasingly clear that sRNAs play important roles in the physiology and virulence of bacteria, particularly in *V. cholerae.*

Purpose of this Study

This work describes identification and characterization of TarA, a 99 nucleotide small regulatory RNA produced by *V. cholerae*. TarA is highly conserved among isolates of *V. cholerae* containing the Vibrio pathogenicity island and is transcribed from a promoter upstream of the gene *tcpI* under conditions in which ToxT is active. There are no open reading frames in the sequence of TarA, which suggests a regulatory role for this molecule. We constructed an isogenic strain of *V. cholerae* lacking the *tarA* gene and used microarray analysis to compare gene expression levels between this mutant and the wildtype strain. Of the roughly twenty potential targets of TarA identified, only one, the glucose import gene *ptsG*, appeared to be repressed by this sRNA.

Chapter II of this dissertation describes the initial characterization of *ptsG* repression by TarA and explores the physiological significance of this regulation. We found that *ptsG* transcript levels drop shortly after induction of expression of *tarA* from the plasmid pHDB3. In cells constitutively overexpressing *tarA*, growth in glucose is severely compromised due to a lack of glucose uptake. Additionally, we determined that TarA is stabilized by the small protein Hfq, which we hypothesized could catalyze base-pairing between TarA and the 5' UTR of the *ptsG* transcript.

Chapter III of this work details further exploration of the mechanism of *ptsG* regulation by TarA. We used biochemical approaches to directly assess binding between TarA and the *ptsG* transcript in the presence and absence of

Hfq. We present evidence that TarA binds both Hfq and the *ptsG* transcript *in vitro*, and that Hfq is not required for base-pairing between the two RNA molecules. We also determined that neither of the predicted Hfq binding sites found in the *tarA* sequence are required for binding to Hfq or the *ptsG* mRNA. We hypothesize that there may be three Hfq binding sites in the TarA molecule that function redundantly in binding of this protein.

Initial characterization of a different set of TarA targets, the *nan* locus, is detailed in Chapter IV of this study. The genes comprising the *nan* locus appeared to be potentially activated by TarA according to our microarray analysis. These genes have been studied in both *E. coli* and *V. cholerae* and encode proteins necessary for catabolism of sialic acid, a 9-carbon sugar. We confirmed some of these findings via qRT-PCR and also began to assess the physiological implications of this regulation by testing for growth defects in media containing sialic acid. While we have not yet been able to identify a phenotype for a *tarA* mutant in sialic acid-containing media, we hypothesize that TarA does regulate this locus in conjunction with the repressor NanR.

Finally, chapter V summarizes the findings reported in this dissertation and fits these discoveries into the greater picture of *V. cholerae* small regulatory RNAs and the pathogenesis of this organism. This work has identified a previously unknown ToxT-activated sRNA produced by *V. cholerae,* TarA, which shares some similarity with the *E. coli* sRNA SgrS but is unique in several ways. Since our initial publication regarding TarA, two other research groups have also

reported finding this molecule, demonstrating that this is a key work in the field of *Vibrio cholerae* and bacterial pathogenesis in general.

References

Anderson, C., (1991) Cholera epidemic traced to risk miscalculation. *Nature* **354**: 255.

Ang, G. Y., C. Y. Yu, K. Balqis, H. T. Elina, H. Azura, M. H. Hani & C. Y. Yean, (2010) Molecular evidence of cholera outbreak caused by a toxigenic Vibrio cholerae O1 El tor variant strain in Kelantan, Malaysia. *J Clin Microbiol* **48**: 3963-3969.

Bishop, A. L. & A. Camilli, (2011) Vibrio cholerae: lessons for mucosal vaccine design. *Expert review of vaccines* **10**: 79-94.

Bourassa, L. & A. Camilli, (2009) Glycogen contributes to the environmental persistence and transmission of Vibrio cholerae. *Molecular microbiology* **72**: 124-138.

Bradley, E. S., K. Bodi, A. M. Ismail & A. Camilli, (2011) A genome-wide approach to discovery of small RNAs involved in regulation of virulence in Vibrio cholerae. *PLoS pathogens* **7**: e1002126.

Butler, D., (2011) No quick fix for Haiti cholera. *Nature* **478**: 295-296.

Chin, C. S., J. Sorenson, J. B. Harris, W. P. Robins, R. C. Charles, R. R. Jean-Charles, J. Bullard, D. R. Webster, A. Kasarskis, P. Peluso, E. E. Paxinos, Y. Yamaichi, S. B. Calderwood, J. J. Mekalanos, E. E. Schadt & M. K. Waldor, (2011) The origin of the Haitian cholera outbreak strain. *The New England journal of medicine* **364**: 33-42.

Clemens, J. D., D. A. Sack, J. R. Harris, J. Chakraborty, M. R. Khan, B. F. Stanton, B. A. Kay, M. U. Khan, M. Yunus, W. Atkinson & et al., (1986) Field trial of oral cholera vaccines in Bangladesh. *Lancet* **2**: 124-127.

da Cunha Ferreira, R. M. & R. A. Cash, (1990) History of the development of oral rehydration therapy. *Clinical therapeutics* **12 Suppl A**: 2-11; discussion 11-13.

Davis, B. M., M. Quinones, J. Pratt, Y. Ding & M. K. Waldor, (2005) Characterization of the small untranslated RNA RyhB and its regulon in Vibrio cholerae. *Journal of bacteriology* **187**: 4005-4014.

Davis, B. M. & M. K. Waldor, (2007) RNase E-dependent processing stabilizes MicX, a Vibrio cholerae sRNA. *Molecular microbiology* **65**: 373-385.

De Haan, L. & T. R. Hirst, (2004) Cholera toxin: a paradigm for multi-functional engagement of cellular mechanisms (Review). *Mol Membr Biol* **21**: 77-92.

De, S. N., (1959) Enterotoxicity of bacteria-free culture-filtrate of Vibrio cholerae. *Nature* **183**: 1533-1534.

DiRita, V. J., C. Parsot, G. Jander & J. J. Mekalanos, (1991) Regulatory cascade controls virulence in Vibrio cholerae. *Proceedings of the National Academy of Sciences of the United States of America* **88**: 5403-5407.

Duffy, J., (1971) The history of Asiatic cholera in the United States. *Bulletin of the New York Academy of Medicine* **47**: 1152-1168.

Dutta, N. K. & M. K. Habbu, (1955) Experimental cholera in infant rabbits: a method for chemotherapeutic investigation. *British journal of pharmacology and chemotherapy* **10**: 153-159.

Dziejman, M., E. Balon, D. Boyd, C. M. Fraser, J. F. Heidelberg & J. J. Mekalanos, (2002) Comparative genomic analysis of Vibrio cholerae: genes that correlate with cholera endemic and pandemic disease. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 1556-1561.

Esteller, M., (2011) Non-coding RNAs in human disease. *Nature reviews. Genetics* **12**: 861-874.

Faruque, S. M., I. B. Naser, M. J. Islam, A. S. Faruque, A. N. Ghosh, G. B. Nair, D. A. Sack & J. J. Mekalanos, (2005) Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 1702-1707.

Finkelstein, R. A., H. T. Norris & N. K. Dutta, (1964) Pathogenesis Experimental Cholera in Infant Rabbits. I. Observations on the Intraintestinal Infection and Experimental Cholera Produced with Cell-Free Products. *The Journal of infectious diseases* **114**: 203-216.

Goel, A. K., M. Jain, P. Kumar, P. Sarguna, M. Bai, N. Ghosh & N. Gopalan, (2011) Molecular characterization reveals involvement of altered El Tor biotype Vibrio cholerae O1 strains in cholera outbreak at Hyderabad, India. *J Microbiol* **49**: 280-284.

Herrington, D. A., R. H. Hall, G. Losonsky, J. J. Mekalanos, R. K. Taylor & M. M. Levine, (1988) Toxin, toxin-coregulated pili, and the toxR regulon are essential for Vibrio cholerae pathogenesis in humans. *The Journal of experimental medicine* **168**: 1487-1492.

Higgins, D. E., E. Nazareno & V. J. DiRita, (1992) The virulence gene activator ToxT from Vibrio cholerae is a member of the AraC family of transcriptional activators. *Journal of bacteriology* **174**: 6974-6980.

Houot, L., S. Chang, C. Absalon & P. I. Watnick, (2010a) Vibrio cholerae phosphoenolpyruvate phosphotransferase system control of carbohydrate transport, biofilm formation, and colonization of the germfree mouse intestine. *Infection and immunity* **78**: 1482-1494.

Houot, L., S. Chang, B. S. Pickering, C. Absalon & P. I. Watnick, (2010b) The phosphoenolpyruvate phosphotransferase system regulates Vibrio cholerae biofilm formation through multiple independent pathways. *Journal of bacteriology* **192**: 3055-3067.

Houot, L. & P. I. Watnick, (2008) A novel role for enzyme I of the Vibrio cholerae phosphoenolpyruvate phosphotransferase system in regulation of growth in a biofilm. *Journal of bacteriology* **190**: 311-320.

Kawamoto, H., Y. Koide, T. Morita & H. Aiba, (2006) Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Molecular microbiology* **61**: 1013-1022.

Koch, R., (1884) An Address on Cholera and its Bacillus. *Br Med J* **2**: 453-459.
Lacey, S. W., (1995) Cholera: calamitous past, ominous future. *Clin Infect Dis* **20**: 1409-1419.
Lee, S. H., D. L. Hava, M. K. Waldor & A. Camilli, (1999) Regulation and temporal expression patterns of Vibrio cholerae virulence genes during infection. *Cell* **99**: 625-634.
Legros, D., C. Paquet, W. Perea, I. Marty, N. K. Mugisha, H. Royer, M. Neira & B. Ivanoff, (1999) Mass vaccination with a two-dose oral cholera vaccine in a refugee camp. *Bull World Health Organ* **77**: 837-842.
Lenz, D. H., M. B. Miller, J. Zhu, R. V. Kulkarni & B. L. Bassler, (2005) CsrA and three redundant small RNAs regulate quorum sensing in Vibrio cholerae. *Molecular microbiology* **58**: 1186-1202.
Lenz, D. H., K. C. Mok, B. N. Lilley, R. V. Kulkarni, N. S. Wingreen & B. L. Bassler, (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae. *Cell* **118**: 69-82.
Lindenbaum, J., W. B. Greenough & M. R. Islam, (1967) Antibiotic therapy of cholera. *Bull World Health Organ* **36**: 871-883.
Liu, J. M. & A. Camilli, (2010) A broadening world of bacterial small RNAs. *Current opinion in microbiology* **13**: 18-23.
Liu, J. M., J. Livny, M. S. Lawrence, M. D. Kimball, M. K. Waldor & A. Camilli, (2009) Experimental discovery of sRNAs in Vibrio cholerae by direct cloning, 5S/tRNA depletion and parallel sequencing. *Nucleic acids research* **37**: e46.
Livny, J., M. A. Fogel, B. M. Davis & M. K. Waldor, (2005) sRNAPredict: an integrative computational approach to identify sRNAs in bacterial genomes. *Nucleic acids research* **33**: 4096-4105.
Lombardo, M. J., J. Michalski, H. Martinez-Wilson, C. Morin, T. Hilton, C. G. Osorio, J. P. Nataro, C. O. Tacket, A. Camilli & J. B. Kaper, (2007) An in vivo expression technology screen for Vibrio cholerae genes expressed in human volunteers. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 18229-18234.
Matson, J. S., J. H. Withey & V. J. DiRita, (2007) Regulatory networks controlling Vibrio cholerae virulence gene expression. *Infection and immunity* **75**: 5542-5549.
Meibom, K. L., M. Blokesch, N. A. Dolganov, C. Y. Wu & G. K. Schoolnik, (2005) Chitin induces natural competence in Vibrio cholerae. *Science (New York, N.Y* **310**: 1824-1827.
Merrell, D. S., S. M. Butler, F. Qadri, N. A. Dolganov, A. Alam, M. B. Cohen, S. B. Calderwood, G. K. Schoolnik & A. Camilli, (2002) Host-induced epidemic spread of the cholera bacterium. *Nature* **417**: 642-645.
Merrell, D. S. & A. Camilli, (2000) Detection and analysis of gene expression during infection by in vivo expression technology. *Philos Trans R Soc Lond B Biol Sci* **355**: 587-599.

Mey, A. R., S. A. Craig & S. M. Payne, (2005) Characterization of Vibrio cholerae RyhB: the RyhB regulon and role of ryhB in biofilm formation. *Infection and immunity* **73**: 5706-5719.

Mosley, W. H., K. J. Bart & A. Sommer, (1972) An epidemiological assessment of cholera control programs in rural East Pakistan. *International journal of epidemiology* **1**: 5-11.

Mustachio, L. M., S. Aksit, R. H. Mistry, R. Scheffler, A. Yamada & J. M. Liu, (2012) The Vibrio cholerae Mannitol Transporter Is Regulated Posttranscriptionally by the MtlS Small Regulatory RNA. *Journal of bacteriology* **194**: 598-606.

Na-Ubol, M., P. Srimanote, M. Chongsa-Nguan, N. Indrawattana, N. Sookrung, P. Tapchaisri, S. Yamazaki, L. Bodhidatta, B. Eampokalap, H. Kurazono, H. Hayashi, G. B. Nair, Y. Takeda & W. Chaicumpa, (2011) Hybrid & El Tor variant biotypes of Vibrio cholerae O1 in Thailand. *The Indian journal of medical research* **133**: 387-394.

Piarroux, R., R. Barrais, B. Faucher, R. Haus, M. Piarroux, J. Gaudart, R. Magloire & D. Raoult, (2011) Understanding the cholera epidemic, Haiti. *Emerg Infect Dis* **17**: 1161-1168.

Pollitzer, R., (1954) Cholera studies. 1. History of the disease. *Bull World Health Organ* **10**: 421-461.

Richard, A. L., J. H. Withey, S. Beyhan, F. Yildiz & V. J. DiRita, (2010) The Vibrio cholerae virulence regulatory cascade controls glucose uptake through activation of TarA, a small regulatory RNA. *Molecular microbiology* **78**: 1171-1181.

Richards, G. R. & C. K. Vanderpool, (2011) Molecular call and response: the physiology of bacterial small RNAs. *Biochimica et biophysica acta* **1809**: 525-531.

Richie, E. E., N. H. Punjabi, Y. Y. Sidharta, K. K. Peetosutan, M. M. Sukandar, S. S. Wasserman, M. M. Lesmana, F. F. Wangsasaputra, S. S. Pandam, M. M. Levine, P. P. O'Hanley, S. J. Cryz & C. H. Simanjuntak, (2000) Efficacy trial of single-dose live oral cholera vaccine CVD 103-HgR in North Jakarta, Indonesia, a cholera-endemic area. *Vaccine* **18**: 2399-2410.

Ritchie, J. M., H. Rui, R. T. Bronson & M. K. Waldor, (2010) Back to the future: studying cholera pathogenesis using infant rabbits. *MBio* **1**.

Rodighiero, C., B. Tsai, T. A. Rapoport & W. I. Lencer, (2002) Role of ubiquitination in retro-translocation of cholera toxin and escape of cytosolic degradation. *EMBO reports* **3**: 1222-1227.

Rui, H., J. M. Ritchie, R. T. Bronson, J. J. Mekalanos, Y. Zhang & M. K. Waldor, (2010) Reactogenicity of live-attenuated Vibrio cholerae vaccines is dependent on flagellins. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 4359-4364.

Ryan, E. T., S. B. Calderwood & F. Qadri, (2006) Live attenuated oral cholera vaccines. *Expert Rev Vaccines* **5**: 483-494.

Sack, D. A., R. B. Sack, G. B. Nair & A. K. Siddique, (2004) Cholera. *Lancet* **363**: 223-233.

Schild, S., R. Tamayo, E. J. Nelson, F. Qadri, S. B. Calderwood & A. Camilli, (2007) Genes induced late in infection increase fitness of Vibrio cholerae after release into the environment. *Cell host & microbe* **2**: 264-277.

Snow, J., (1855) On the mode of communication of cholera. In*.* London: John Churchill, pp.

Son, M. S., C. J. Megli, G. Kovacikova, F. Qadri & R. K. Taylor, (2011) Characterization of Vibrio cholerae O1 El Tor biotype variant clinical isolates from Bangladesh and Haiti, including a molecular genetic analysis of virulence genes. *J Clin Microbiol* **49**: 3739-3749.

Song, T., F. Mika, B. Lindmark, Z. Liu, S. Schild, A. Bishop, J. Zhu, A. Camilli, J. Johansson, J. Vogel & S. N. Wai, (2008) A new Vibrio cholerae sRNA modulates colonization and affects release of outer membrane vesicles. *Molecular microbiology* **70**: 100-111.

Taviani, E., C. J. Grim, J. Choi, J. Chun, B. Haley, N. A. Hasan, A. Huq & R. R. Colwell, (2010) Discovery of novel Vibrio cholerae VSP-II genomic islands using comparative genomic analysis. *FEMS Microbiol Lett* **308**: 130-137.

Taylor, D. N., K. P. Killeen, D. C. Hack, J. R. Kenner, T. S. Coster, D. T. Beattie, J. Ezzell, T. Hyman, A. Trofa, M. H. Sjogren & et al., (1994) Development of a live, oral, attenuated vaccine against El Tor cholera. *J Infect Dis* **170**: 1518-1523.

Taylor, R. K., V. L. Miller, D. B. Furlong & J. J. Mekalanos, (1987) Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 2833-2837.

Tickner, J. & T. Gouveia-Vigeant, (2005) The 1991 cholera epidemic in Peru: not a case of precaution gone awry. *Risk analysis : an official publication of the Society for Risk Analysis* **25**: 495-502.

Toledo-Arana, A., F. Repoila & P. Cossart, (2007) Small noncoding RNAs controlling pathogenesis. *Current opinion in microbiology* **10**: 182-188.

Tsai, B., C. Rodighiero, W. I. Lencer & T. A. Rapoport, (2001) Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. *Cell* **104**: 937-948.

Vanderpool, C. K. & S. Gottesman, (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Molecular microbiology* **54**: 1076-1089.

Wadler, C. S. & C. K. Vanderpool, (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 20454-20459.

WHO, (2006) Implementing the New Recommendations on the Clinical Management of Diarrhoea. In*.* Geneva, Switzerland: WHO Press, pp.

WHO, (2010) Updated WHO position paper on cholera vaccines. In., pp.
WHO, (2011) Global epidemics and impact of cholera. In., pp.
Withey, J. H. & V. J. DiRita, (2006) The toxbox: specific DNA sequence requirements for activation of Vibrio cholerae virulence genes by ToxT. *Molecular microbiology* **59**: 1779-1789.

Chapter II

The *Vibrio cholerae* virulence regulatory cascade controls glucose uptake through activation of TarA, a small regulatory RNA

Note: A modified version of this text was previously published in *Molecular Microbiology* (Richard et al., 2010). Figures 2.1 and 2.3 and Table 2.1 were created by Jeffrey H. Withey (Wayne State University), and the microarray analysis was performed by Sinem Beyhan (UC Santa Cruz).

Summary

Vibrio cholerae causes the severe diarrheal disease cholera. A cascade of regulators controls expression of virulence determinants in *V. cholerae* at both transcriptional and post-transcriptional levels. ToxT is the direct transcriptional activator of the major virulence genes in *V. cholerae*. Here we describe TarA, a highly conserved small regulatory RNA, whose transcription is activated by ToxT from toxboxes present upstream of the ToxT-activated gene *tcpI*. TarA regulates *ptsG,* encoding a major glucose transporter in *V. cholerae*. Cells overexpressing TarA exhibit decreased steady-state levels of *ptsG* mRNA and grow poorly in glucose-minimal media. A mutant lacking the ubiquitous regulatory protein Hfq expresses diminished TarA levels, indicating that TarA likely interacts with Hfq to regulate gene expression. RNAhybrid analysis of TarA and the putative *ptsG* mRNA leader suggests potential productive

base-pairing between these two RNA molecules. A *V. cholerae* mutant lacking TarA is compromised for infant mouse colonization in competition with wild type, suggesting a role in the *in vivo* fitness of *V. cholerae*. Although somewhat functionally analogous to SgrS of *E. coli*, TarA does not encode a regulatory peptide, and its expression is activated by the virulence gene pathway in *V. cholerae* and not by glycolytic intermediates.

Introduction

Asiatic cholera, a severe diarrheal disease, remains a significant cause of morbidity and mortality in many regions of the globe (Sack *et al.*, 2006). Cholera is caused by ingestion of the gram-negative, highly motile curved bacillus *Vibrio cholerae* in contaminated food or water. *V. cholerae* are present in many coastal areas worldwide; however, of the hundreds of known *V. cholerae* serogroups, only the O1 and O139 serogroups can cause cholera (Sack *et al.*, 2004, Reidl & Klose, 2002).

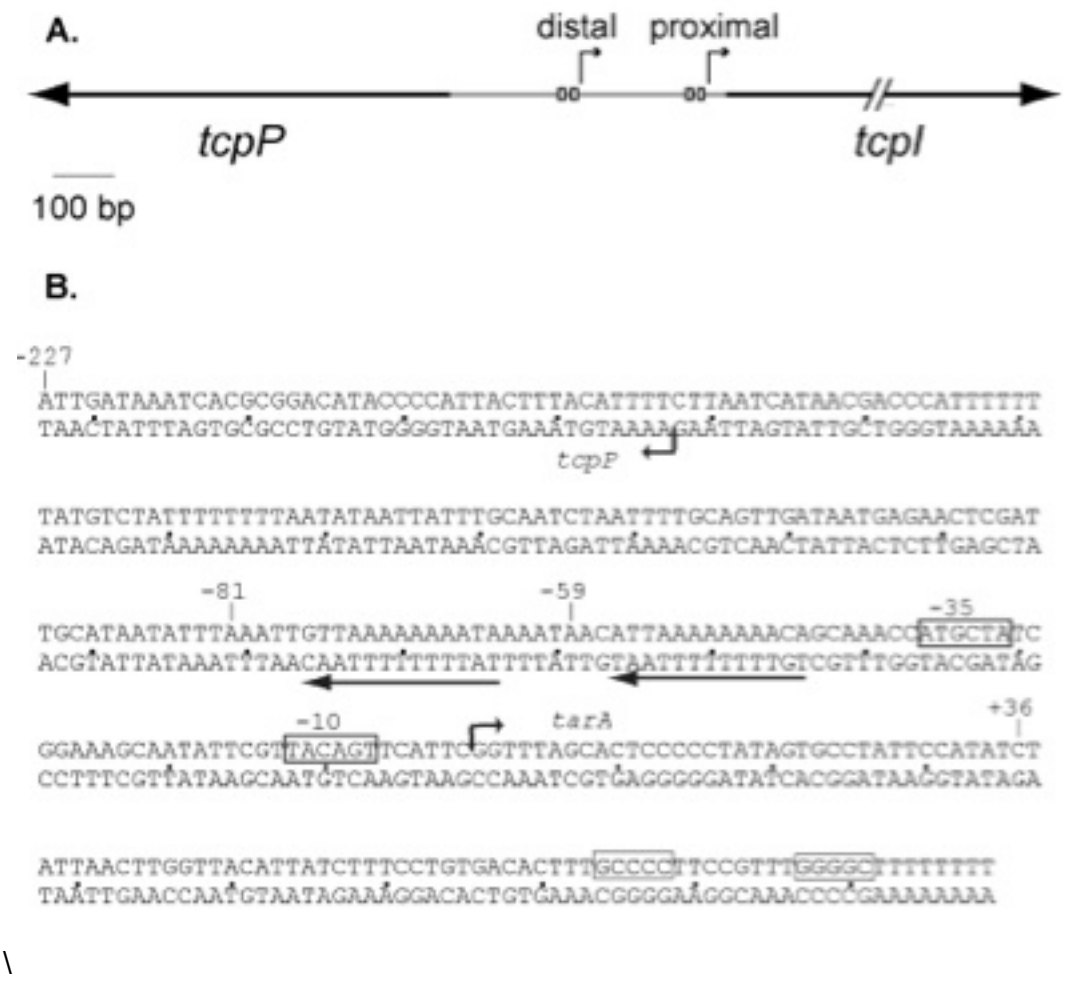
The two major *V. cholerae* virulence factors are the cholera toxin (CT) and the toxin-coregulated pilus (TCP). Secretion of CT into the intestinal lumen by *V. cholerae* directly causes the secretory diarrhea characteristic of cholera (Gill, 1976, Lonnroth & Holmgren, 1973). The genes encoding the two subunits of CT, *ctxAB*, are located within the genome of a filamentous bacteriophage, CTXΦ (Waldor & Mekalanos, 1996). The TCP, a type IV bundle-forming pilus, is required for intestinal colonization by *V. cholerae* and acts to initiate microcolony formation (Attridge *et al.*, 1996, Herrington *et al.*, 1988, Thelin & Taylor, 1996, Taylor *et al.*, 1987). The genes encoding TCP are

located within the Vibrio pathogenicity island (VPI), which may also be a mobile element but is not a bacteriophage (Manning, 1997, Karaolis *et al.*, 1998, Rajanna *et al.*, 2003, Faruque & Mekalanos, 2003). Other putative virulence genes located within the VPI are coordinately regulated with *ctxAB* and *tcp*, but their roles in pathogenesis are unclear.

Expression of genes encoding *V. cholerae* virulence factors is regulated by a cascade involving transcription activators, repressors and proteases (Matson *et al.*, 2007); however, the direct transcription activator of virulence genes is ToxT (Matson et al., 2007). ToxT is a member of the AraC/XylS family of transcription factors (Higgins *et al.*, 1992, Ogierman & Manning, 1992). ToxT binds to a 13 bp degenerate DNA sequence, the toxbox, to activate transcription (Withey & DiRita, 2006), using toxboxes configured as direct repeats, inverted repeats, and single toxboxes to activate transcription of different operons (Withey & DiRita, 2006, Withey & Dirita, 2005b).

Binding sites for ToxT have been characterized at most of the promoters it is known to activate (Withey & DiRita, 2006, Withey & DiRita, 2005a, Withey & Dirita, 2005b). An exception is the *tcpI* promoter, located on the VPI and encoding a putative methyl-accepting chemotaxis protein with an unknown role in pathogenesis (Harkey *et al.*, 1994). Our sequence analysis identified two pairs of putative toxboxes upstream of *tcpI*; one pair, in an inverted configuration, is positioned proximal to the *tcpI* open reading frame (Withey & DiRita, 2006). The other pair, in a direct repeat configuration, is positioned at a site upstream of the proximal toxboxes (Figure 2.1), and thus we refer to them as the distal toxboxes. Primer extension analysis performed previously by others identified two different start sites of *tcpI* transcription that would potentially

correspond to transcription activated by ToxT from each set of toxboxes; though when those studies were carried out, the toxboxes had not yet been identified (Thomas *et al.*, 1995, Murley *et al.*, 2000). We performed experiments to characterize the role of ToxT in transcription from both of these putative sites and have found that there are indeed two separate ToxT-dependent transcripts produced upstream of *tcpI*. However, the *tcpI*-proximal promoter alone is responsible for transcription of *tcpI* itself (JHW and VJD, unpublished).



\

Figure 2.1. Schematic and diagram of *tcpI* promoters and the surrounding region. A.) Schematic. Scale bar: 0.5 centimeters represents 200 bp of DNA. Bent arrows indicate putative transcriptional start sites, and long black arrows represent coding sequences of the known genes *tcpP* and *tcpI*. *tcpI* has been shortened from 1.6 kb for clarity. Toxboxes, or ToxT-recognition sequences, are indicated by paired boxes. B.) Diagram of *tcpI* distal promoter, *tarA*, and surrounding sequences. Bent arrows indicate transcriptional start sites for *tcpP* and *tarA*. Numbers indicate position relative to *tarA* transcriptional start site. ToxT binding sites (toxboxes) for the *tcpI* distal promoter are underlined with horizontal arrows. The -35 and -10 elements for *tarA* are indicated by thick boxes, and thin boxes represent potential base-pairing regions in putative hairpin structure.

Here we describe a small RNA (~99 nt), which we named TarA for *T*oxT-*a*ctivated *R*NA, whose transcription from a gene located between *tcpI* and *tcpP* is activated by ToxT at the promoter distal to *tcpI*. The TarA nucleotide sequence is highly conserved among both epidemic and environmental *V. cholerae* strains that carry the VPI. We demonstrate that TarA influences glucose uptake through its effect on the transcript encoding the glucose transporter PtsG.

Results

ToxT activates transcription of a small non-coding RNA.

To characterize ToxT-dependent activation of *tcpI*, we constructed *lacZ* fusions and assayed β–galactosidase activity in *V. cholerae* having either wild type *toxT* or a *toxT* deletion. Our results indicated that the minimal ToxT-responsive DNA sequence for activating a *tcpI-lacZ* fusion at this *tcpI*-proximal promoter extended only to -94 relative to the proximal *tcpI* promoter. Because a second separate start site for *tcpI* transcription had been previously described at -243 relative to the AUG (Murley et al., 2000, Harkey et al., 1994), we also constructed *tcpI::lacZ* fusions extending to -431 relative to the start site of transcription from the *tcpI*-proximal promoter to include both putative promoters (Figure 2.1A). However, additional upstream DNA, even if it contained the putative *tcpI*-distal promoter, did not affect the level of β–galactosidase produced. This result suggested that the distal promoter, if functional, does not read into *tcpI* under the conditions tested.

To determine whether the *tcpI*-distal promoter is functional, we created a set of *lacZ* fusions whose 3' endpoints were located upstream of the *tcpI*-proximal promoter (at -173 relative to its start site and +36 relative to the distal promoter start site) and with 5' endpoints that extended to -59, -81, or -223 relative to the start of transcription from the distal promoter (Figure 2.1B). Results of β–galactosidase assays using these constructs indicated that the *tcpI*-distal promoter was indeed functional and highly activated by ToxT (Table 2.1). Sequence analysis identified two consensus toxboxes configured as a direct repeat upstream of the *tcpI*-distal promoter -35 box that had the same spacing as the toxboxes we observed at the *tcpA* promoter (Withey & DiRita, 2006) (Figure 2.1B). Deletion of one of the two toxboxes resulted in a complete loss of ToxT-directed transcription from the *tcpI*-distal promoter (Table 2.1, Figure 2.1B (-59 construct)).

5' Endpoint	*3' Endpoint*	*Miller Units (WT)*	*Miller Units (toxT)*	*Fold Difference (WT/toxT)*
-223	+36	11,800 ± 500	804 ± 137	15x
-81	+36	18,000 ± 551	436 ± 44.8	41x
-59	+36	94.6 ± 9.35	93.1 ± 16.5	1x

Table 2.1. Nested *tarA::lacZ* fusion analysis. 5' and 3' endpoints indicate region of *V. cholerae* chromosome ligated into pTL61T plasmid. Fusion plasmids were electroporated into WT O395 and an isogenic *toxT* strain. Cultures were grown under inducing conditions for 3 hours and then tested for β-galactosidase activity, with data reported in Miller units. Fold difference was determined by taking the ratio of wild type activity to *toxT* activity.

Downstream of the start site of transcription is a G/C rich region that could potentially form a stem-loop structure (boxes in Figure 2.1B), followed by a tract of T nucleotides on the non-template strand, suggesting the presence of a transcription terminator. To determine if a small RNA product initiating at the *tcpI*-distal promoter and ending at the putative terminator is produced, we performed Northern blots on whole cell RNA purified from *V. cholerae* having a wild type or a deleted *toxT* gene, probing with an oligonucleotide complementary to the putative sRNA sequence. A single RNA product of the expected size, <100 nt, was detected in the Northern blotting experiments (Figure 2.2). Production of the sRNA required functional ToxT, was detectable beginning at three hours of growth (data not shown), continued through logarithmic phase, and peaked at seven hours of growth under ToxT-inducing conditions; no TarA was detected in non-inducing conditions (Figure 2.2). These findings are consistent with previous observations of ToxT activity (Yu & DiRita, 1999). Based on this evidence, we named this previously unknown sRNA "TarA" for ToxT-Activated RNA. TarA contains a putative binding site for the ubiquitous regulatory protein Hfq and lacks an open reading frame; thus, its function is likely to be regulatory.

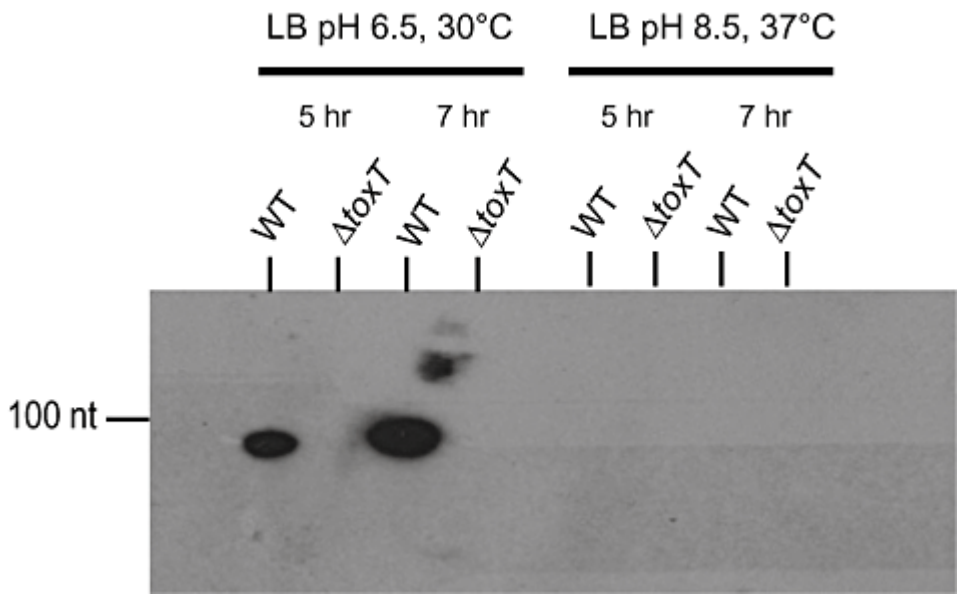


Figure 2.2. TarA is produced under ToxT-inducing conditions. Wild type and *toxT* strains were grown under inducing and non-inducing conditions (LB pH 6.5, 30ºC and LB pH 8.5, 37ºC, respectively) and total RNA was collected at indicated timepoints. Equal amounts of RNA from each sample (as determined by OD_{260}) were subjected to Northern blot analysis, using an oligonucleotide complementary to the putative sRNA sequence as a probe. Methylene blue staining was used to confirm load amounts (data not shown).

TarA is highly conserved among *V. cholerae* strains.

To determine whether TarA has homology to any other sRNA species, we performed searches with its sequence from the classical *V. cholerae* strain O395 using both the RNAfam database and BLAST (Touzet & Perriquet, 2004). The former search indicated that TarA has no homology to any known sRNA family. However, the latter search indicated that TarA has a very high degree of sequence conservation among *V. cholerae* (Table 2.2, Figure 2.3). TarA sequences from eight different *V. cholerae* strains were obtained from the BLAST search; two of these are classical biotype *V. cholerae*, two are El Tor biotype *V. cholerae*, and four are non-O1 *V. cholerae*. The two classical strains have identical TarA sequences and the sequences of the two El Tor strains are also identical (Table 2.2 and Figure 2.3). However, the El Tor sequences have a single nucleotide insertion and two substitutions that differ from the classical sequences. The non-O1 *V. cholerae* strains, commonly known as environmental *V. cholerae* to distinguish them from the O1 epidemic strains, carry the VPI but are not capable of causing cholera. These environmental *V. cholerae* strains have TarA sequences apparently derived from the classical biotype or vice versa. Three of these four non-O1 strains have TarA sequences identical to the classical TarA sequences; the fourth non-O1 strain has a single nucleotide change from the classical TarA sequences at a position that differs from the changes observed between classical and the El Tor biotype strains.

V. cholerae strain	*Biotype*	*O Antigen*	*TarA Conservation*	*Upstream DNA Conservation*
O395*	Classical	O1	99/99 nt	99/99 bp
Z17561*	Classical	O1	99/99 nt	99/99 bp
N16961*	El Tor	O1	96/99 nt	62/99 bp
H1*	El Tor	O1	96/99 nt	75/99 bp
SCE4	Environmental	O8	99/99 nt	75/99 bp
SCE200*	Environmental	O44	99/99 nt	91/99 bp
SCE256	Environmental	O42	99/99 nt	68/99 bp
SCE226	Environmental	O35	98/99 nt	85/99 bp

*Toxigenic strain

Table 2.2. Conservation of TarA and upstream sequences. Asterisks indicate toxigenic strains.

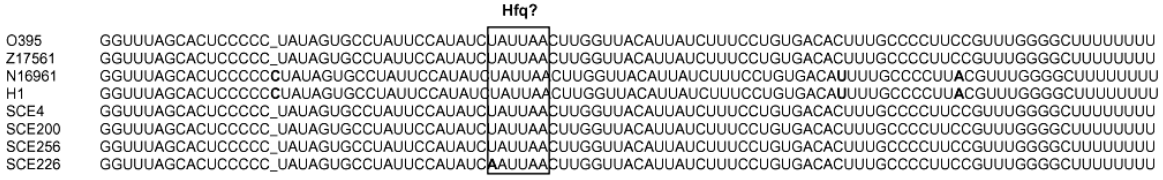


Figure 2.3. Comparison of *tarA* sequences among *V. cholerae* strains. Sequences from strains indicated in Table 2.2 are aligned, with bold letters indicating differences between strains. The box indicates a conserved potential Hfq binding site.

To determine whether the degree of TarA sequence conservation among different *V. cholerae* strains is a general feature of the VPI, we compared the DNA sequences 99 bp directly upstream of the *tarA* start site in the above strains to the sequence in O395. This analysis indicated that the *tarA* sequence was preferentially conserved over the surrounding intergenic sequence (Table 2.2). The degree of conservation of the 99 bp directly upstream of *tarA* varies to a wide degree, even between strains that have identical *tarA* sequences. For example, classical O395 and environmental SCE 256 have 99/99 identical base pairs within the *tarA* gene but only 68/99 identical base-pairs within the upstream sequence. These upstream 99 bp include both the promoter and the two toxboxes that are required for expression of *tarA*; notably, the toxbox sequences are 100% conserved in every strain. We have also observed TarA expression in the El Tor strain N16961 under ToxT-inducing conditions (data not shown). These results strongly suggest that TarA conveys an evolutionary advantage to both environmental and epidemic *V. cholerae* and that its regulation by ToxT is highly conserved.

TarA regulates *ptsG* mRNA levels in *V. cholerae*.

To identify potential regulatory targets of TarA, we probed microarrays of the *V. cholerae* genome with RNA isolated from classical strain O395 and its *tarA* derivative. The *tarA* strain features a markerless deletion of the 99 nucleotide RNA sequence with the promoters left intact. The whole-genome expression data were

analyzed using the Significance Analysis of Microarrays (SAM) program (Tusher *et al.*, 2001) using ≥2.0 fold differences in gene expression and a ≤1% false discovery rate (FDR) as cutoff values. A gene annotated as *ptsG* (VC2013 in the N16961 genome sequence) was upregulated by greater than two-fold in the *tarA* strain compared with wild type. This is consistent with a previous microarray study (Bina *et al.*, 2003), in which VC2013 RNA was shown to be upregulated in *toxT* mutant *V. cholerae* relative to wild type. This pattern of regulation – with VC2013 expression being negatively controlled by ToxT – is in contrast to that of nearly all other known ToxT-regulated genes, such as the toxin co-regulated pilus (*tcp*) and the cholera toxin (*ctxAB*), which are positively controlled by ToxT. That *ptsG* may be the target of sRNA regulation in *V. cholerae* is consistent with the fact that *ptsG* in *E. coli* is also regulated by a small RNA, SgrS (Vanderpool & Gottesman, 2004).

To confirm the microarray result, we used quantitative real time PCR (qRT-PCR) comparing *ptsG* mRNA levels in wild type and *tarA* mutant *V. cholerae* strain O395. Consistent with the microarray experiments, we observed an approximately three-fold increase in *ptsG* mRNA in mutant cells compared with wild type (Figure 2.4), supporting the hypothesis that TarA acts to decrease the levels of *ptsG* mRNA. To be sure that regulation was due to the specific absence of *tarA* and not due to a polar effect on *tcpI*, a *tcpI* strain was also tested. *ptsG* transcript levels in the *tcpI* strain closely resemble those seen in wild type O395, ruling out this possibility.

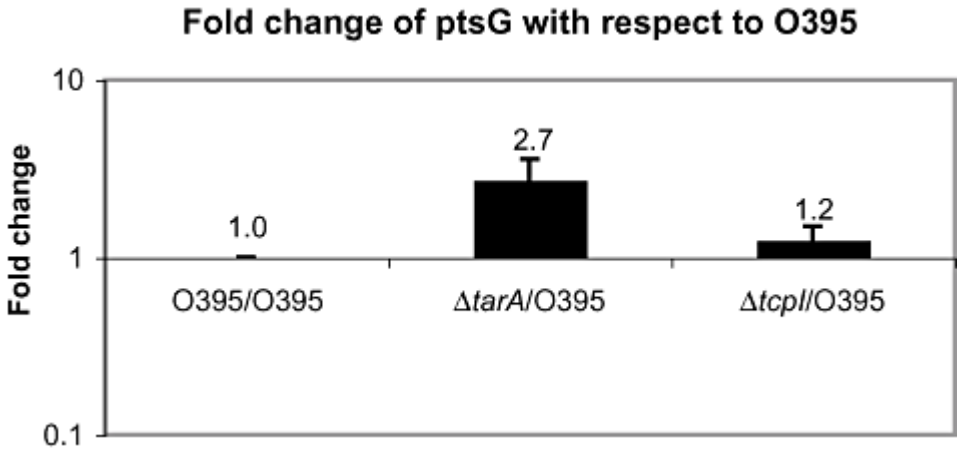


Figure 2.4. Confirmation of *ptsG* regulation by qRT-PCR analysis. Cultures of the listed strains were grown under ToxT-inducing conditions for 7 hours and total RNA was isolated using Trizol. cDNA was prepared from equivalent amounts of RNA for each sample, and cDNA was used for SYBR Green qRT-PCR. Numbers indicate fold change in transcript level between strains, calculated using the ΔΔCT method and *recA* transcript levels as an internal control.

Growth of *V. cholerae* in minimal glucose medium is limited in cells overexpressing TarA.

In *E. coli*, SgrS acts to reduce *ptsG* mRNA levels and to limit glucose uptake under conditions where flux through the glycolytic pathway may be disrupted, thereby allowing the accumulation of potentially lethal phospho-glucoside compounds (Vanderpool & Gottesman, 2004). When *E. coli* is grown with the glucose analogue alpha-methyl-glucoside (αMG), the compound is phosphorylated and taken up by PtsG, but phospho-αMG cannot be catabolized by the normal glucose metabolic pathway and therefore accumulates. Under such conditions, *sgrS* expression is induced and the resulting degradation of *ptsG* mRNA reduces further uptake of αMG. Mutants lacking *sgrS* are highly sensitive to αMG (Vanderpool & Gottesman, 2004). We tested whether *tarA V. cholerae* exhibited enhanced sensitivity to αMG relative to wild type, which would be a physiological consequence of the fact that TarA regulates *ptsG* mRNA in *V. cholerae*. The two strains were grown in LB supplemented with αMG in concentrations ranging from 1-10% weight by volume; we did not see a consistent effect of this glucose analogue on *tarA V. cholerae* at any of the concentrations tested.

Failing to identify an αMG effect, we tested another *ptsG*-related hypothesis. We reasoned that constitutive overexpression of TarA, which would remove the RNA from its normal regulation by ToxT, would affect the growth physiology of *V. cholerae*, leading to diminished growth in glucose-containing minimal media if indeed *ptsG* mRNA levels are regulated by TarA. Overexpression of TarA in LB-grown *V. cholerae* had a slight effect on culture growth rate, although the final optical density achieved by this

culture was close to wild type. Growth of *tarA* and wild type *V. cholerae* carrying the vector alone was unaffected compared to wild type (Figure 2.5A). However, after dilution from LB into minimal MOPS/0.5% glucose, the cells overexpressing TarA were severely deficient for continued growth and resembled the *ptsG* strains, while the other strains grew similarly after a lag phase of approximately three hours (Figure 2.5B). The *ptsG* strain complemented with a constitutively-expressed copy of *ptsG* exhibited partially restored growth in this medium. Analysis of glucose remaining in the supernatants of these cultures revealed similar levels of glucose for approximately four hours, after which the culture overexpressing TarA and that of the *ptsG* mutants nearly stopped removing glucose from the medium while the others continued to do so (Figure 2.5C). Thus, the poor growth of wild type cells overexpressing TarA correlated with decreased levels of glucose uptake, similar to cells lacking *ptsG*.

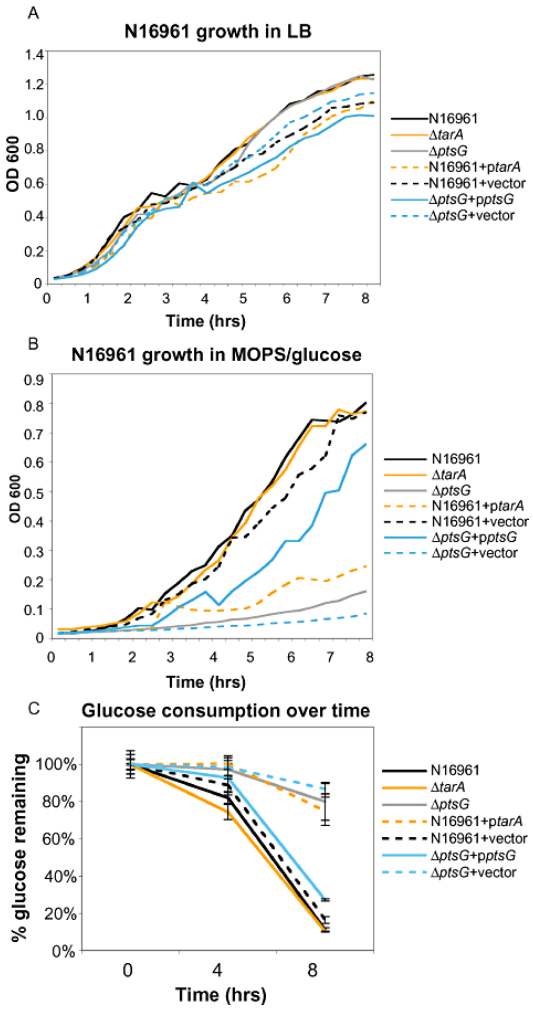


Figure 2.5. Overexpression of TarA results in severely reduced glucose uptake. A.) LB growth. Indicated strains were grown overnight in LB media, subcultured into fresh LB and grown to mid-exponential phase, then washed with PBS and diluted into fresh LB to OD = 0.04. Cultures were grown in a 96-well plate at 30°C with constant aeration, and the OD_{600} was measured every 20 minutes. Growth assays were performed at least three times, and a representative growth curve is shown. (B.) MOPS/glucose growth. Strains were prepared as in (A), but diluted into MOPS/0.5% glucose instead of LB and then grown as in (A). (C.) Strains were grown overnight in LB, subcultured into LB and grown to mid-exponential phase, then washed and diluted into MOPS/0.5% glucose, and grown in flasks. At specified timepoints, supernatants were collected and the glucose concentration was tested using a tetrazolium blue reducing sugar assay. Calculated glucose concentrations were divided by the starting concentration to determine the amount of glucose remaining in the media. This experiment was performed three times, and a representative assay is shown. Error bars indicate the standard deviation among triplicate samples in the representative experiment.

To more carefully examine association between TarA and *ptsG* mRNA levels, we placed *tarA* expression under control of an IPTG-inducible promoter in pHDB3 and followed steady state levels of *ptsG* mRNA message upon TarA induction. *tarA V. cholerae* carrying this plasmid was grown in LB to mid-logarithmic levels before addition of IPTG to induce *tarA* expression. Thirty minutes after IPTG induction, cells were collected, washed, resuspended in fresh media and allowed to continue growing for another 30 minutes. RNA was harvested from cells at 0, 2, 10 and 30 minutes after addition of IPTG, and then at 10 and 30 minutes after the IPTG was washed out. RNA was analyzed by Northern blotting using a probe specific for TarA or *ptsG* mRNA.

Prior to induction (0 minutes) TarA was undetectable while *ptsG* mRNA was abundant. Within two minutes after induction by IPTG, TarA became detectable and the amount of *ptsG* mRNA diminished. This continued throughout the course of the experiment, with TarA expression remaining at high levels and *ptsG* mRNA detectable at levels considerably lower than those seen prior to induction (Figure 2.6). Cells carrying the cloning vector alone expressed no detectable TarA and very high levels of *ptsG* mRNA. The levels of *ptsG* mRNA were higher in the strain carrying only the cloning vector than in the cells carrying the *tarA* clone without induction, which could be due to undetectable yet functional levels of TarA being expressed even prior to addition of IPTG. There is a slight decrease in *ptsG* mRNA levels in the vector control after thirty minutes of induction and ten minutes of washout (t=40 timepoint), which we attribute to transcript turnover as the cells near stationary phase.

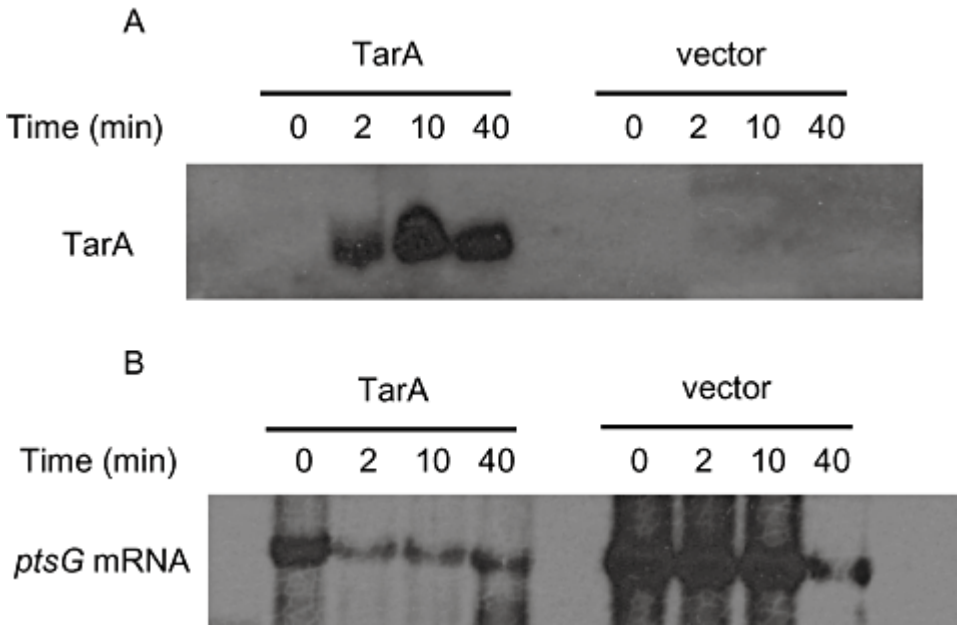


Figure 2.6. Induction of TarA reduces *ptsG* transcript levels. Strains AR059 and AR060, containing an IPTG-inducible copy of TarA and an empty vector control, respectively, were grown overnight in LB. Cultures were diluted 1:200 into fresh LB, grown until OD ~ 0.4, and stimulated with IPTG. Thirty minutes after induction, cultures were washed and resuspended in fresh media lacking IPTG. RNA was collected at indicated timepoints (from the start of induction) using Trizol, and equal amounts of RNA for each sample were subjected to Northern blot analysis. Methylene blue staining was used to confirm load amounts (not shown). Blots were probed with oligonucleotides complementary to TarA (A) and to the *ptsG* transcript (B).

Base complementarity between TarA and the 5' region of the *ptsG* mRNA.

In *E. coli*, SgrS controls *ptsG* transcript levels in part by associating with the 5' end of the *ptsG* mRNA and causing its degradation in a process that depends on Hfq and RNAse E. (Maki *et al.*, 2008, Wadler & Vanderpool, 2007). We used the alignment prediction program RNAhybrid (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/submission.html) to identify potential areas of complementarity between TarA and the leader sequence of the *ptsG* transcript that would be expected for a similar mechanism to be operating in *V. cholerae*. This program produced a structure with a favorable free energy (-40.3 kcal/mol) (Figure 2.7). For purposes of comparison, we tested the potential complementarity of SgrS and the 100 bp directly upstream of *ptsG* in *E. coli* with the same program, which yielded a similarly-composed structure with a minimal free energy of -49.2 kcal/mol (data not shown).

TarA is unstable in the absence of Hfq. Given the putative Hfq binding site in the TarA sequence, we hypothesized that TarA interacts with Hfq. It is well-established that Hfq-interacting small RNAs are markedly less stable in the absence of Hfq (Morita *et al.*, 2004, Kawamoto *et al.,* 2005, (Sledjeski *et al.*, 2001), (Grunze *et al.*, 2002). To explore a potential TarA/Hfq interaction, we assessed TarA levels by Northern blot analysis of RNA from wild type and *hfq* mutant *V. cholerae.* Cells were grown for seven hours in ToxT-inducing conditions prior to RNA isolation and blotting. As predicted, the *hfq* mutant had significantly reduced TarA RNA in comparison to the wild type strain O395 (Figure 2.8).

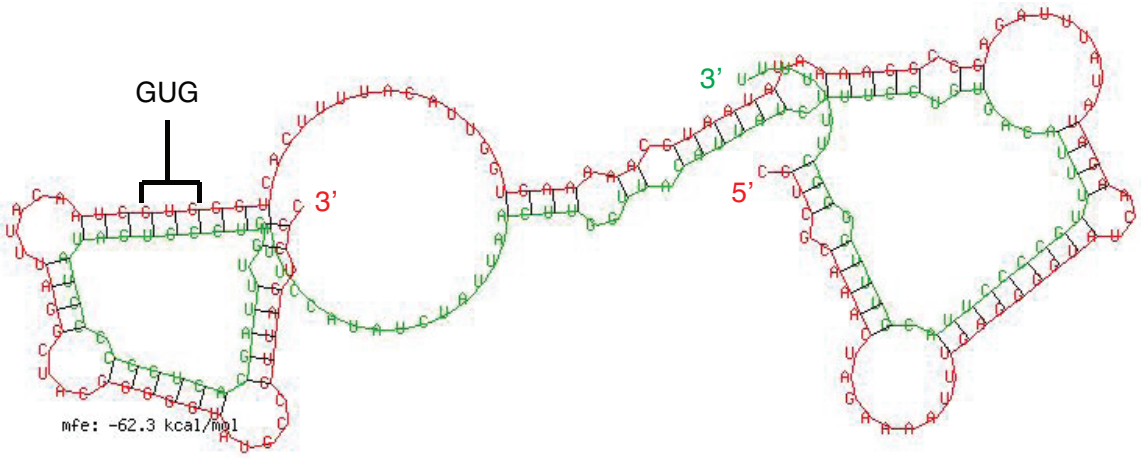
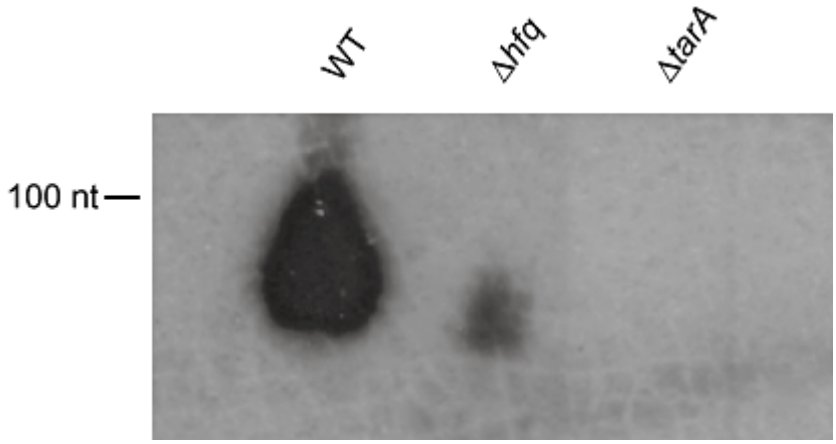


Figure 2.7. Alignment of TarA and *ptsG* upstream sequences. The program RNAhybrid (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/welcome.html) was used to predict potential base-pairing between TarA and the *ptsG* transcript. The TarA sequence is shown in green and the *ptsG* leader sequence in red, with the 5' and 3' ends of each RNA labeled correspondingly. The 5' end of TarA is unlabeled but found to the left of the *ptsG* 3' end. The minimum predicted free energy for this alignment is -62.3 kcal/mol.



Relative amount:	1	0.14	.06

Figure 2.8. TarA is unstable in the absence of Hfq. Wild type, *tarA,* and *hfq* strains were grown under inducing conditions (LB pH 6.5, 30ºC) and total RNA was collected after seven hours of growth. Equal amounts of RNA from each sample (as determined by OD_{260}) were subjected to Northern blot analysis, using an oligonucleotide complementary to TarA as a probe. Methylene blue staining was used to confirm load amounts (data not shown). Protein amounts were quantified using ImageJ software.

Discussion

Here we describe a small, non-coding regulatory RNA, TarA, whose transcription is directly activated by the major *V. cholerae* virulence activator protein ToxT. Activation of a regulatory RNA by ToxT adds another level to the complex *V. cholerae* virulence regulatory cascade and contributes to our understanding of ToxT-dependent gene activation. Based on our previous analyses of other ToxT-activated genes, the promoter architecture of *tarA* most closely resembles that of *tcpA* (Withey & DiRita, 2006). Furthermore, the degree of activation conferred by ToxT to the two genes is very similar – 41-fold for *tarA* and 44-fold for *tcpA*. The two *tarA* toxboxes have 7/7 and 6/7 of the toxbox core determinants conserved (Matson et al., 2007), the highest match to consensus we have observed for any operon having two toxboxes. However, while the degree of ToxT activation of *tarA* is very high, it is not higher than the ToxT activation of *tcpA*, which has a poorer match to consensus at one toxbox. This suggests that the toxbox sequences may have evolved to be maximally efficient at any given locus, such that surrounding sequence that is not conserved in the consensus alignment plays a more important role than previously thought.

TarA joins a growing list of regulatory sRNAs in *Vibrio cholerae.* Four sRNAs, Qrr1,2,3, and 4, are involved in regulating the quorum-sensing response, together with Hfq protein, by negatively affecting the stability of *hapR* mRNA (Lenz *et al.*, 2004). Another three sRNAs, CsrB,C,D, are also involved in quorum sensing; these act to inhibit activity of CsrA protein, which subsequently leads to expression of the Qrr sRNAs (Lenz *et al.*, 2005). Another sRNA is RyhB, which is involved in regulating iron-

containing proteins (Davis *et al.*, 2005). RyhB acts with Hfq to destabilize the mRNAs encoding these proteins. While many other sRNAs have been described in *E. coli*, very few of these are conserved in *V. cholerae*. The exceptions are SsrA, also known as tmRNA, RnpB, Spf, and Ffs (Livny & Waldor, 2007). Recently, large numbers of sRNAs have been identified in *V. cholerae* using bioinformatics and parallel sequencing methods (Livny *et al.*, 2008, Vogel *et al.*, 2003). One of these, MtlS, previously known as IGR7, also contributes to regulation of carbon utilization by negatively regulating a mannitol transport gene, *mtlA* (Liu *et al.*, 2009, Mustachio *et al.*, 2012). A study published after this work appeared in *Molecular Microbiology* confirmed that TarA is produced by *V. cholerae* and also identified a second ToxT-activated small RNA, TarB (Bradley *et al.*, 2011).

Our results suggest that a target of negative regulation by TarA in *V. cholerae* is *ptsG*, encoding a glucose transporter in the phosphotransferase system. In *E. coli*, *ptsG* is negatively regulated by SgrS, an sRNA whose mechanisms of action have been the subject of a significant amount of research. Expression of *sgrS* is activated by a conserved regulatory protein called SgrR, in response to glucose-phosphate stress (Vanderpool & Gottesman, 2004). SgrS is a bifunctional molecule, postulated to inhibit PtsG through two distinct mechanisms. One is by acting as a regulatory RNA directly by participating in a ternary complex with Hfq and the *ptsG* mRNA through specific base-pairing between the two RNA molecules. The site on *ptsG* mRNA that is base-paired to SgrS is in the 5' end of the transcript at the Shine-Dalgarno sequence (Vanderpool & Gottesman, 2004), blocking translation and eventually leading to

degradation of *ptsG* mRNA through a mechanism that requires RNAse E (Maki et al., 2008). We think it likely that TarA also acts in a similar fashion, given the base complementarity between TarA and *ptsG* RNA, overlapping the latter's start codon. Interestingly, the *V. cholerae ptsG* transcript lacks a Shine-Dalgarno sequence, unlike that of *E. coli*. AU-rich sequences within a predicted unpaired region of TarA may also correspond to an Hfq binding domain (Figure 2.7) (Link *et al.*, 2009). Our observation that TarA is less stable in an *hfq* mutant lends support to our hypothesis that Hfq and TarA interact. The second mechanism by which SgrS regulates PtsG is through a small protein, SgrT, encoded at the 5' end of the same RNA that includes SgrS in *E. coli* and in a variety of other species (Wadler & Vanderpool, 2007), (Horler & Vanderpool, 2009). SgrT is hypothesized to act post-translationally on PtsG function (Wadler & Vanderpool, 2007). TarA does not have an open reading frame and is consequently much shorter than SgrS. Another distinction between SgrS and TarA is that, unlike functional SgrS homologues in many other Gram negative bacteria (Horler & Vanderpool, 2009), a homologue of the SgrR regulatory protein is not found encoded next to *tarA* in the VPI. As sequence homology to *sgrR* was used to locate SgrS functional homologues in other enteric bacteria, the fact that an SgrS homologue was not predicted in *Vibrio cholerae* is not surprising.

Unlike *sgrS* in *E. coli*, *tarA* expression is not controlled by the glucose-phosphate response regulator SgrR, but rather by ToxT, the activator responsible for expression of the major virulence factors of *V. cholerae*. The high degree of similarity between the *tarA* and *tcpA* toxboxes suggests that the two are expressed

synchronously during infection. *tarA* makes a slight contribution to the fitness of *V. cholerae* during infection, as competition experiments between wild type and *tarA* mutant *V. cholerae* in the infant mouse model resulted in a small but consistent colonization defect of the *tarA* mutant (avg. competitive index = 0.45). Cells expressing elevated PtsG during infection could be at a competitive disadvantage if glucose is not the preferred in vivo carbon source. This hypothesis is under investigation. An alternative explanation for how the link between *tarA* expression and virulence gene activation may have evolved comes from work demonstrating that *V. cholerae* in stool samples of cholera patients store glucose in the form of glycogen granules, and that glycogen biosynthesis genes are expressed late during infection (Schild *et al.*, 2007). Of note, *toxT* is expressed very early in infection, and the *tarA* promoter most closely resembles that of *tcpA*, which is one of the first genes activated by ToxT. Perhaps *tarA* expression by ToxT during infection, with a subsequent reduction in *ptsG* mRNA, prepares the cells for the apparent metabolic conversion to glycogen production by reducing the amount of glucose transported into the glycolytic pathway.

Materials and Methods

V. cholerae strains and plasmids.

Strains used in this work were classical *V. cholerae* O395, El Tor strain N16961, and derivatives (listed in Table 2.3). The *tarA*, *ptsG,* and *hfq* strains were constructed from O395 and N16961 by mating with *E. coli* SM10λpir using the pKAS32 suicide vector as previously described (Skorupski & Taylor, 1996). The deletion was confirmed by PCR using primers flanking the gene. Strains were grown in LB medium at 37° C (overnight cultures), in LB medium with a starting pH of 6.5 at 30° C (inducing conditions for classical strains), or under AKI conditions (inducing conditions for El Tor strains). Strains were maintained in LB/20% glycerol at -70°C. Antibiotics were used in the following concentrations: ampicillin: 100 μg/ml, streptomycin: 100 μg/ml and chloramphenicol: 30 μg/ml.

Strains used for β-galactosidase assays were either O395 or an O395 *toxT* derivative (VJ740,(Champion *et al.*, 1997)) carrying the respective *tarA::lacZ* fusion in plasmid pTL61T (Linn & St Pierre, 1990). All constructs were cloned between the HindIII and XbaI sites of pTL61T. For constitutive and inducible expression of TarA, plasmids pKK177-3RI (Opdyke *et al.*, 2004) and pHDB3 (Ulbrandt *et al.*, 1997) were used, respectively. In pKK177-3RI, the *tarA* sequence alone was cloned between the EcoRI and HindIII sites, creating plasmid pJM292. The *tarA* sequence preceded by the P_{LlacO} operator sequence was cloned between the BamHI and HindIII sites of pHDB3 to generate pAR6. These plasmids, and the empty vector controls, were electroporated into wild type O395 and N16961 (pKK177-3RI and pJM292) and the corresponding

D*tarA* strains (pHDB3 and pAR6). For *ptsG* complementation, the *ptsG* sequence preceded by a ribosome binding site was cloned between the EcoRI and XbaI sites of pUC19, generating plasmid pAR7. This plasmid and the empty vector control were then electroporated into N16961 *ptsG*. All plasmid sequences were confirmed by the University of Michigan DNA sequencing core.

Strain	Plasmid	Genotype	Parent strain	Source/reference
O395		smR		Lab collection
JW415		O395 *tarA*	O395	This study
N16961		smR		Lab collection
AR43		N16961 *tarA*	N16961	This study
VJ740		O395 *toxT*	O395	Champion *et al.,* 1997
JW335	pJW184	*tcpI8::lacZ*	O395	This study
JW336	pJW185	*tcpI9::lacZ*	O395	This study
JW350	pJW186	*tcpI10::lacZ*	O395	This study
JW339	pJW184	*tcpI8::lacZ toxT*	VJ740	This study
JW340	pJW185	*tcpI9::lacZ toxT*	VJ740	This study
JW353	pJW186	*tcpI10::lacZ toxT*	VJ740	This study
AR15		O395 *hfq*	O395	This study
AR38	pKK177-3RI	smR, apR	N16961	This study
AR39	pJM292	pKK177-3RI with TarA sequence under control of *rrnB* promoter	N16961	This study
AR59	pHDB3, pMMB207	smR, apR, cmR	JW415	This study
AR60	pAR6, pMMB207	pHDB3 with TarA sequence under control of PLlacO promoter	JW415	This study
AR74		N16961 *ptsG*	N16961	This study
AR80	pUC19	smR, apR	AR074	This study
AR81	pAR7	pUC19 with *ptsG* under control of *lac* promoter	AR074	This study

Table 2.3. Strains used in this study.

β-galactosidase assays.

Strains were cultured overnight in LB medium at 37°C, then sub-cultured at a 1/50 dilution into inducing medium and grown for 3 hours at 30°C with vigorous aeration. Bacteria were then placed on ice and chloramphenicol was added to 0.5 mg ml^{-1}. β-galactosidase assays were performed using the basic procedure of Miller (Miller, 1972) as previously described (Withey & Dirita, 2005b).

DNA manipulations.

Plasmids were purified using the Promega Wizard Plus Miniprep kit. Polymerase chain reaction (PCR) was performed using Taq DNA polymerase (Roche) as specified by the manufacturer. Restriction enzymes were purchased from New England Biolabs and used as specified by the manufacturer.

Growth analysis and glucose quantification.

Strains N16961, AR43, AR38, AR39, AR74, AR80, and AR81 were grown overnight in LB, then subcultured 1:100 into fresh LB and grown to mid-log phase (six hours). OD_{600} was measured for each culture, and $1x10^9$ bacteria of each strain were pelleted, washed in PBS, and resuspended to a final OD_{600} of 2 in PBS. These suspensions were then diluted 1:50 (final OD_{600} = 0.04) into LB and MOPS minimal medium supplemented with 0.4% glucose (Neidhardt *et al.*, 1974), and dispensed into a 96-well plate. The plate was incubated at 30°C with constant shaking for 8 hours in an Omega Polarstar plate reader, with OD_{600} measured every 20 minutes. For glucose

quantification, cultures were prepared as described and grown in 125 mL flasks on a ThermoFisher orbital shaker at 37°C. One milliliter samples were taken at 0, 4 and 8 hours of growth. Bacteria were pelleted, and 10-fold diluted supernatants were used in a tetrazolium blue reducing sugar assay as described (Jue & Lipke, 1985). Glucose concentration was determined for each sample and divided by the starting concentration to determine the percent glucose remaining. Assays were run in triplicate three times.

Northern blots.

Whole cell RNA was purified from O395 and VJ740 (*toxT*) *V. cholerae* grown in parallel under inducing and non-inducing conditions (Figure 2), from strains AR59 and AR60 after addition of IPTG (Figure 6), and from strains O395, AR15, and AR24 grown under inducing conditions (Figure 8). At specific points (as detailed in figure legends for each experiment), the OD_{600} of each culture was read and aliquots of each culture equal to 1 ml of a 0.7 OD_{600} culture were removed. RNA purification was performed using the TRIzol reagent as specified by the manufacturer (Invitrogen). For TarA visualization, 7 mg total RNA from each sample were separated by 8% polyacrylamide gel electrophoresis and transferred to an Amersham Hybond nylon membrane (GE Healthcare) (0.45mm) membrane with a Semi-Phor apparatus (Hoefer). An oligonucleotide probe complementary to nt 12-36 of TarA was radiolabeled with γ–^{32}P-ATP with T4 polynucleotide kinase (New England Biolabs). Probing conditions were as previously described (Sambrook, 1989). To visualize *ptsG* transcript, 12.5 mg of each sample were run on a 1.2% agarose/formaldehyde gel and transferred to an Amersham

Hybond nylon membrane (GE Healthcare). Membranes were probed with an oligonucleotide complementary to nt 176-208 of the *ptsG* transcript, labeled as above.

Microarray and qRT-PCR analysis.

Cultures of O395 and *tarA* strains were grown overnight and then diluted 1:100 into LB pH 6.5 and grown at 30°C. Once cultures reached an OD_{600} ~0.35, they were diluted 1:10 and grown until OD_{600} ~0.35, and 2 ml aliquots of the samples were harvested. Cell pellets were resuspended in TRIzol (Invitrogen) and total RNA was isolated according to the manufacturer's protocol as described previously (Yildiz *et al.*, 2004). Whole genome expression analysis was performed using 70-mer oligos representing the open reading frames present in the *V. cholerae* N16961 genome. cDNA synthesis, microarray hybridization, and analysis of expression data were performed as described previously (Beyhan *et al.*, 2006) using two biological and four technical replicates (for each biological replicate).

RNA samples for qRT-PCR were DNase treated, run on an agarose gel to check quality, and quantified by measuring the OD_{260}. 2.5 μg of each sample were treated with Moloney murine leukemia virus (M-MLV) reverse transcriptase (Invitrogen) according to the manufacturer's specifications. For detection of transcripts, primers amplifying a 200-bp region in the middle of the mRNA were used with SYBR Green Master Mix (Stratagene) on a Stratagene MX3000P thermocycler. Primers were designed using the OligoPerfect tool (Invitrogen). Each test was performed in triplicate at least three times, and fold change in expression was calculated using the ΔΔCT

method (Pfaffl, 2001), with *recA* transcript levels used as controls between cDNA samples.

Infant mouse colonization assays.

Four-day-old CD1 mice (Charles River) were inoculated intragastrically with approximately 10^6 bacteria. The inocula were 50:50 mixtures of the strain of interest (either *tarA*$^-$/ *lacZ*$^+$ *V. cholerae* or the *tarA*$^+$/ *lacZ*$^+$ control) and the control strain, which is a *lacZ*$^-$ O395 derivative. Inoculated mice were incubated at 30° C for 16 hours, after which the mice were sacrificed and their intestines were removed and homogenized. Serial dilutions of the individual intestinal homogenates were plated on LB agar plates containing streptomycin and X-gal. Blue and white colonies were counted after overnight growth at 37°C. To determine the ratio of strains in the initial inoculum, the mixtures used to inoculate the mice were diluted into PBS and plated onto LB agar plates containing streptomycin and X-gal, and again the blue and white colonies were counted. The competitive index was calculated by dividing the blue:white ratio of the initial inoculum by the in vivo blue:white ratio.

Acknowledgments

This work was supported in part by AI31645 (to V.J.D.) from NIAID. J.H.W. was supported by a Kirschstein National Research Service Award (1 F32 AI51074) awarded by the NIH. A.L.R. is a trainee of the Genetics Training Program at University of Michigan. We thank Kevin Ginsburg for construction of the N16961 *ptsG* strain.

References

Attridge, S. R., P. A. Manning, J. Holmgren & G. Jonson, (1996) Relative significance of mannose-sensitive hemagglutinin and toxin-coregulated pili in colonization of infant mice by Vibrio cholerae El Tor. *Infection and immunity* **64**: 3369-3373.

Beyhan, S., A. D. Tischler, A. Camilli & F. H. Yildiz, (2006) Differences in gene expression between the classical and El Tor biotypes of Vibrio cholerae O1. *Infection and immunity* **74**: 3633-3642.

Bina, J., J. Zhu, M. Dziejman, S. Faruque, S. Calderwood & J. Mekalanos, (2003) ToxR regulon of Vibrio cholerae and its expression in vibrios shed by cholera patients. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 2801-2806.

Bradley, E. S., K. Bodi, A. M. Ismail & A. Camilli, (2011) A genome-wide approach to discovery of small RNAs involved in regulation of virulence in Vibrio cholerae. *PLoS pathogens* **7**: e1002126.

Champion, G. A., M. N. Neely, M. A. Brennan & V. J. DiRita, (1997) A branch in the ToxR regulatory cascade of Vibrio cholerae revealed by characterization of toxT mutant strains. *Molecular microbiology* **23**: 323-331.

Davis, B. M., M. Quinones, J. Pratt, Y. Ding & M. K. Waldor, (2005) Characterization of the small untranslated RNA RyhB and its regulon in Vibrio cholerae. *Journal of bacteriology* **187**: 4005-4014.

Faruque, S. M. & J. J. Mekalanos, (2003) Pathogenicity islands and phages in Vibrio cholerae evolution. *Trends in microbiology* **11**: 505-510.

Gill, D. M., (1976) The arrangement of subunits in cholera toxin. *Biochemistry* **15**: 1242-1248.

Grunze, H., S. Kasper, G. Goodwin, C. Bowden, D. Baldwin, R. Licht, E. Vieta & H. J. Moller, (2002) World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of bipolar disorders. Part I: Treatment of bipolar depression. *World J Biol Psychiatry* **3**: 115-124.

Harkey, C. W., K. D. Everiss & K. M. Peterson, (1994) The Vibrio cholerae toxin-coregulated-pilus gene tcpI encodes a homolog of methyl-accepting chemotaxis proteins. *Infection and immunity* **62**: 2669-2678.

Herrington, D. A., R. H. Hall, G. Losonsky, J. J. Mekalanos, R. K. Taylor & M. M. Levine, (1988) Toxin, toxin-coregulated pili, and the toxR regulon are essential for Vibrio cholerae pathogenesis in humans. *The Journal of experimental medicine* **168**: 1487-1492.

Higgins, D. E., E. Nazareno & V. J. DiRita, (1992) The virulence gene activator ToxT from Vibrio cholerae is a member of the AraC family of transcriptional activators. *Journal of bacteriology* **174**: 6974-6980.

Horler, R. S. & C. K. Vanderpool, (2009) Homologs of the small RNA SgrS are broadly distributed in enteric bacteria but have diverged in size and sequence. *Nucleic acids research* **37**: 5465-5476.

Jue, C. K. & P. N. Lipke, (1985) Determination of reducing sugars in the nanomole range with tetrazolium blue. *Journal of biochemical and biophysical methods* **11**: 109-115.

Karaolis, D. K., J. A. Johnson, C. C. Bailey, E. C. Boedeker, J. B. Kaper & P. R. Reeves, (1998) A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 3134-3139.

Lenz, D. H., M. B. Miller, J. Zhu, R. V. Kulkarni & B. L. Bassler, (2005) CsrA and three redundant small RNAs regulate quorum sensing in Vibrio cholerae. *Molecular microbiology* **58**: 1186-1202.

Lenz, D. H., K. C. Mok, B. N. Lilley, R. V. Kulkarni, N. S. Wingreen & B. L. Bassler, (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in Vibrio harveyi and Vibrio cholerae. *Cell* **118**: 69-82.

Link, T. M., P. Valentin-Hansen & R. G. Brennan, (2009) Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 19292-19297.

Linn, T. & R. St Pierre, (1990) Improved vector system for constructing transcriptional fusions that ensures independent translation of lacZ. *Journal of bacteriology* **172**: 1077-1084.

Liu, J. M., J. Livny, M. S. Lawrence, M. D. Kimball, M. K. Waldor & A. Camilli, (2009) Experimental discovery of sRNAs in Vibrio cholerae by direct cloning, 5S/tRNA depletion and parallel sequencing. *Nucleic acids research* **37**: e46.

Livny, J., H. Teonadi, M. Livny & M. K. Waldor, (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. *PloS one* **3**: e3197.

Livny, J. & M. K. Waldor, (2007) Identification of small RNAs in diverse bacterial species. *Current opinion in microbiology* **10**: 96-101.

Lonnroth, I. & J. Holmgren, (1973) Subunit structure of cholera toxin. *Journal of general microbiology* **76**: 417-427.

Maki, K., K. Uno, T. Morita & H. Aiba, (2008) RNA, but not protein partners, is directly responsible for translational silencing by a bacterial Hfq-binding small RNA. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 10332-10337.

Manning, P. A., (1997) The tcp gene cluster of Vibrio cholerae. *Gene* **192**: 63-70.

Matson, J. S., J. H. Withey & V. J. DiRita, (2007) Regulatory networks controlling Vibrio cholerae virulence gene expression. *Infection and immunity* **75**: 5542-5549.

Miller, J. H., (1972). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Murley, Y. M., J. Behari, R. Griffin & S. B. Calderwood, (2000) Classical and El Tor biotypes of Vibrio cholerae differ in timing of transcription of tcpPH during growth in inducing conditions. *Infection and immunity* **68**: 3010-3014.

Mustachio, L. M., S. Aksit, R. H. Mistry, R. Scheffler, A. Yamada & J. M. Liu, (2012) The Vibrio cholerae Mannitol Transporter Is Regulated Posttranscriptionally by the MtlS Small Regulatory RNA. *Journal of bacteriology* **194**: 598-606.

Neidhardt, F. C., P. L. Bloch & D. F. Smith, (1974) Culture medium for enterobacteria. *Journal of bacteriology* **119**: 736-747.

Ogierman, M. A. & P. A. Manning, (1992) Homology of TcpN, a putative regulatory protein of Vibrio cholerae, to the AraC family of transcriptional activators. *Gene* **116**: 93-97.

Opdyke, J. A., J. G. Kang & G. Storz, (2004) GadY, a small-RNA regulator of acid response genes in Escherichia coli. *Journal of bacteriology* **186**: 6698-6705.

Pfaffl, M. W., (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* **29**: e45.

Rajanna, C., J. Wang, D. Zhang, Z. Xu, A. Ali, Y. M. Hou & D. K. Karaolis, (2003) The vibrio pathogenicity island of epidemic Vibrio cholerae forms precise extrachromosomal circular excision products. *Journal of bacteriology* **185**: 6893-6901.

Reidl, J. & K. E. Klose, (2002) Vibrio cholerae and cholera: out of the water and into the host. *FEMS microbiology reviews* **26**: 125-139.

Richard, A. L., J. H. Withey, S. Beyhan, F. Yildiz & V. J. DiRita, (2010) The Vibrio cholerae virulence regulatory cascade controls glucose uptake through activation of TarA, a small regulatory RNA. *Molecular microbiology* **78**: 1171-1181.

Sack, D. A., R. B. Sack & C. L. Chaignat, (2006) Getting serious about cholera. *The New England journal of medicine* **355**: 649-651.

Sack, D. A., R. B. Sack, G. B. Nair & A. K. Siddique, (2004) Cholera. *Lancet* **363**: 223-233.

Sambrook, J., Fritsch, E.F., Maniatis, T., (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schild, S., R. Tamayo, E. J. Nelson, F. Qadri, S. B. Calderwood & A. Camilli, (2007) Genes induced late in infection increase fitness of Vibrio cholerae after release into the environment. *Cell host & microbe* **2**: 264-277.

Skorupski, K. & R. K. Taylor, (1996) Positive selection vectors for allelic exchange. *Gene* **169**: 47-52.

Sledjeski, D. D., C. Whitman & A. Zhang, (2001) Hfq is necessary for regulation by the untranslated RNA DsrA. *Journal of bacteriology* **183**: 1997-2005.

Taylor, R. K., V. L. Miller, D. B. Furlong & J. J. Mekalanos, (1987) Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated

with cholera toxin. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 2833-2837.

Thelin, K. H. & R. K. Taylor, (1996) Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by Vibrio cholerae O1 El Tor biotype and O139 strains. *Infection and immunity* **64**: 2853-2856.

Thomas, S., S. G. Williams & P. A. Manning, (1995) Regulation of tcp genes in classical and El Tor strains of Vibrio cholerae O1. *Gene* **166**: 43-48.

Touzet, H. & O. Perriquet, (2004) CARNAC: folding families of related RNAs. *Nucleic acids research* **32**: W142-145.

Tusher, V. G., R. Tibshirani & G. Chu, (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 5116-5121.

Ulbrandt, N. D., J. A. Newitt & H. D. Bernstein, (1997) The E. coli signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* **88**: 187-196.

Vanderpool, C. K. & S. Gottesman, (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Molecular microbiology* **54**: 1076-1089.

Vogel, J., V. Bartels, T. H. Tang, G. Churakov, J. G. Slagter-Jager, A. Huttenhofer & E. G. Wagner, (2003) RNomics in Escherichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic acids research* **31**: 6435-6443.

Wadler, C. S. & C. K. Vanderpool, (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 20454-20459.

Waldor, M. K. & J. J. Mekalanos, (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science (New York, N.Y* **272**: 1910-1914.

Withey, J. H. & V. J. DiRita, (2005a) Activation of both acfA and acfD transcription by Vibrio cholerae ToxT requires binding to two centrally located DNA sites in an inverted repeat conformation. *Molecular microbiology* **56**: 1062-1077.

Withey, J. H. & V. J. Dirita, (2005b) Vibrio cholerae ToxT independently activates the divergently transcribed aldA and tagA genes. *Journal of bacteriology* **187**: 7890-7900.

Withey, J. H. & V. J. DiRita, (2006) The toxbox: specific DNA sequence requirements for activation of Vibrio cholerae virulence genes by ToxT. *Molecular microbiology* **59**: 1779-1789.

Yildiz, F. H., X. S. Liu, A. Heydorn & G. K. Schoolnik, (2004) Molecular analysis of rugosity in a Vibrio cholerae O1 El Tor phase variant. *Molecular microbiology* **53**: 497-515.

Yu, R. R. & V. J. DiRita, (1999) Analysis of an autoregulatory loop controlling ToxT, cholera toxin, and toxin-coregulated pilus production in Vibrio cholerae. *Journal of bacteriology* **181**: 2584-2592.

Chapter III

Binding interactions between TarA, Hfq, and the *ptsG* leader sequence

Note: The cloning of the TarA internal site deletion construct detailed here was performed by Mariah E. Gavin.

Summary

TarA, a small regulatory RNA produced by the bacterium *Vibrio cholerae*, is a 99 nucleotide molecule that has been previously shown to decrease expression of the the glucose import gene *ptsG*. In a mutant lacking *tarA*, *ptsG* transcript levels are approximately three-fold increased. Transcription of *tarA* is ToxT-induced, and we hypothesize that *V. cholerae* uses this RNA to downregulate glucose uptake in the host intestine. In order to better understand regulation of the *ptsG* gene by TarA, it is necessary to elucidate the mechanism of this repression. We hypothesized that TarA makes specific base-pairing interactions with the 5' untranslated region (UTR) of the *ptsG* transcript, and that this binding may be facilitated by the small protein Hfq. This chapter focuses on efforts to test these hypotheses using electrophoretic mobility shift assays to monitor binding between TarA and other molecules. We demonstrate i) that TarA can bind the *ptsG* transcript *in vitro,* and ii) Hfq is not required for the TarA-*ptsG* interaction, notwithstanding that Hfq protein evidently binds TarA and stabilizes

this molecule *in vivo*. Additionally, we found that TarA contains three potential Hfq binding sites, which likely serve redundant functions.

Introduction

The Gram-negative bacterium *Vibrio cholerae,* like most other bacterial species, produces several small regulatory RNAs that have been well characterized. These RNAs regulate their targets at either the protein level, like the CsrB-D RNAs that are involved in quorum sensing (Lenz *et al.*, 2005) or, more commonly, by interacting with the mRNA encoding their targets (Bardill *et al.*, 2011, Song *et al.*, 2008). The latter class of sRNAs includes the Qrr1-4 sRNAs, which redundantly regulate *hapR* and thus control biofilm formation and virulence factor production. Recently published work has demonstrated that these RNAs directly bind the 5' untranslated region of the *hapR* transcript through base-pairing interactions, and that this interaction is catalyzed by the small protein Hfq (Bardill et al., 2011). As the Qrr sRNAs negatively regulate this gene, this base-pairing is presumed to recruit RNase E to degrade the transcript.

We previously described the ToxT-activated small regulatory RNA TarA, a 99 nucleotide molecule that negatively regulates the gene *ptsG*, which encodes the major glucose transporter of *Vibrio cholerae* (Richard et al., 2010). Our initial studies illustrated that Hfq stabilizes this RNA, and we hypothesized that TarA

directly binds the *ptsG* transcript in a manner similar to the Qrrs. This would also be a similar mechanism to that of the *E. coli* sRNA SgrS, which represses *ptsG* in that bacterium. SgrS has been extensively studied by several groups. Its transcription is induced under conditions of glucose-phosphate stress and upon expression, the sRNA base-pairs with the *ptsG* transcript; this binding is catalyzed by Hfq and results in RNase E-mediated degradation of the *ptsG* transcript (Wadler & Vanderpool, 2007, Maki *et al.*, 2010, Vanderpool & Gottesman, 2004). Mutational analysis of SgrS has defined the minimal base-pairing interactions necessary for this regulation, and recently, the poly U tail of the RNA has been implicated as the required Hfq binding site (Maki et al., 2010, Otaka *et al.*, 2011). For other sRNAs, the Hfq binding site is an internal AU-rich sequence, demonstrating that there is not one specific binding sequence for this protein (Brescia *et al.*, 2003, Zhang *et al.*, 1998).

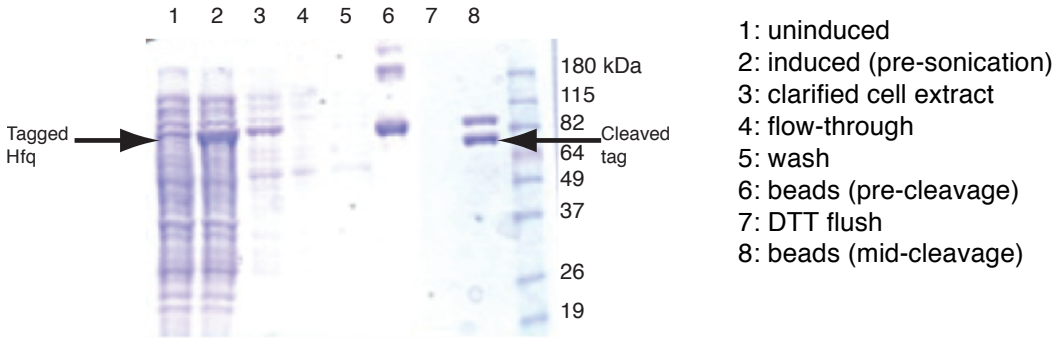
To further understand regulation of *ptsG* by TarA, we sought to test the ability of TarA to bind the *ptsG* transcript *in vitro.* Electrophoretic mobility shift assays have been previously used to query RNA-RNA interactions (Bardill et al., 2011), and we reasoned that this technique would allow us to learn more about several aspects of *ptsG* regulation through the use of various TarA mutants. In this chapter, we use this experimental approach to demonstrate that TarA binds both Hfq and the *ptsG* transcript *in vitro.* We also show evidence that TarA may possess several redundant Hfq binding sites. Finally, while Hfq stabilizes TarA, it appears to be dispensable in the process of TarA-*ptsG* binding.

Results

TarA directly binds Hfq and the *ptsG* leader sequence *in vitro*

To test binding of TarA to its target, *ptsG,* we used an electrophoretic mobility shift assay. The TarA and *ptsG* leader sequence RNAs were generated via *in vitro* transcription, and TarA was end-labeled with ^{32}P. The *ptsG* RNA contained the 204 nucleotides upstream of the GUG and the first 166 codons of the *ptsG* gene. The putative transcript start site was mapped using the SoftBerry BPROM tool (http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb). The *Vibrio cholerae* Hfq protein was expressed in *E. coli* and purified using the New England Biolabs IMPACT system, which utilizes a self-cleaving intein tag with a chitin-binding domain (Figure 3.1). Purified RNAs and protein were mixed together in native buffer, incubated at room temperature, and then run on a non-denaturing polyacrylamide gel. After drying, the gel was exposed to a storage phosphor screen and imaged using a phosphorimager. We first determined that in the absence of *ptsG* leader sequence, TarA interacted with the RNA-binding protein Hfq (Figure 3.2).

A.



B.

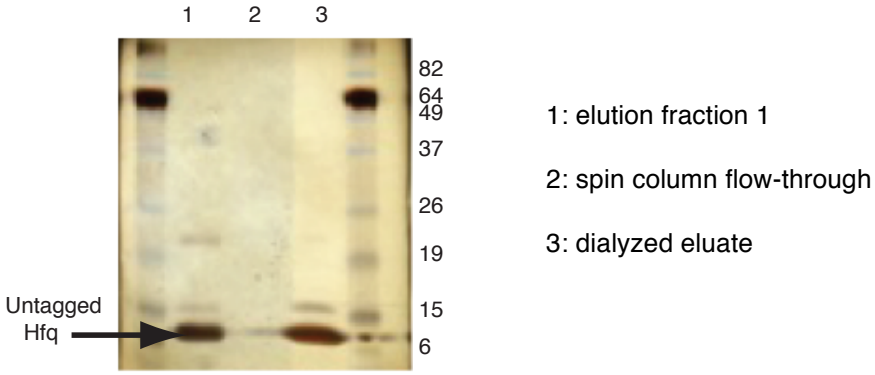


Figure 3.1 Purification of Hfq protein. The *V. cholerae hfq* gene was cloned into the vector pTYB21 and transformed into *E. coli* ER2566. Cells were grown overnight at 16°C in the presence of 0.4 mM IPTG to induce protein expression. Cultures were spun down and cells were lysed by sonication. Cleared cell lysate was then loaded onto a column of chitin beads and allowed to empty by gravity flow. The column was then washed, and cleavage of the chitin-binding tag was induced by addition of buffer containing DTT. Post-cleavage, protein was eluted in column buffer, and elution fractions containing protein were pooled and concentrated. Protein concentration was determined by spectrophotometry and samples were stored at -20°C. A.) Purification fractions were run on a 12.5% polyacrylamide gel and stained with Coomassie Brilliant Blue. B.) Eluted protein was electrophoresed on a 12.5% polyacrylamide gel and silver-stained using the SilverQuest staining kit (Invitrogen).

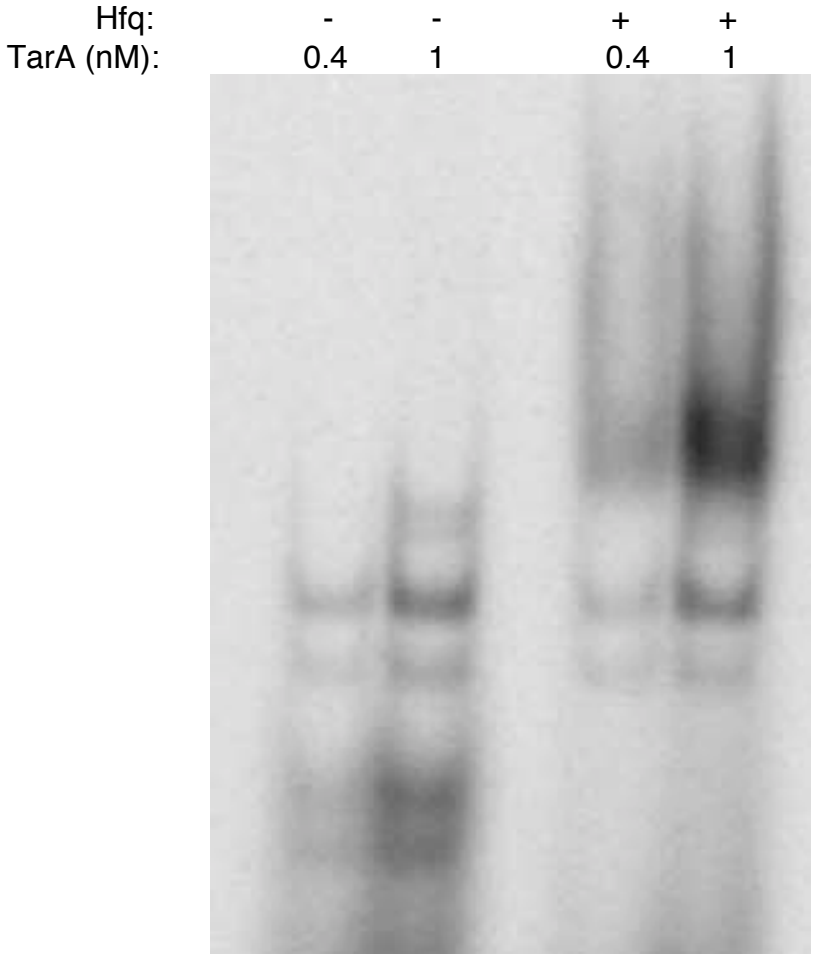


Figure 3.2. TarA binds the protein Hfq *in vitro.* TarA was transcribed *in vitro* gel purified, and end-labeled with ^{32}P. Hfq was expressed in *E. coli* and purified using the IMPACT purification system (NEB). Indicated concentrations of RNA were mixed with 300 nM Hfq in native buffer and incubated at 25°C for 60 minutes, then run on a nondenaturing 5% polyacrylamide gel and exposed to a phosphor storage screen. Image was obtained using a Typhoon phosphorimager (Amersham).

In reaction mixtures containing only labeled TarA and the *ptsG* leader sequence, we observed a shift in mobility of labeled TarA, indicating binding between the species; the intensity of the shifted band increased with the concentration of *ptsG* (Figure 3.3). When Hfq was present as well, at the highest concentration of *ptsG* leader we observed a species with mobility slightly slower than that of the species seen in the absence of Hfq. We interpret this as comprising TarA*, ptsG* leader and Hfq. In contrast, at the lowest concentrations of *ptsG,* we observed less clearly defined species with motility between that of TarA alone and of the putative TarA:*ptsG* species. This is likely a TarA:Hfq complex. Additionally, samples containing Hfq appear to undergo a shift at very low *ptsG* concentrations (2.5 and 5 nM). From this experiment, we concluded that TarA binds both Hfq and the *ptsG* transcript *in vitro,* and that the presence of Hfq may catalyze complex formation between TarA and the *ptsG* transcript.

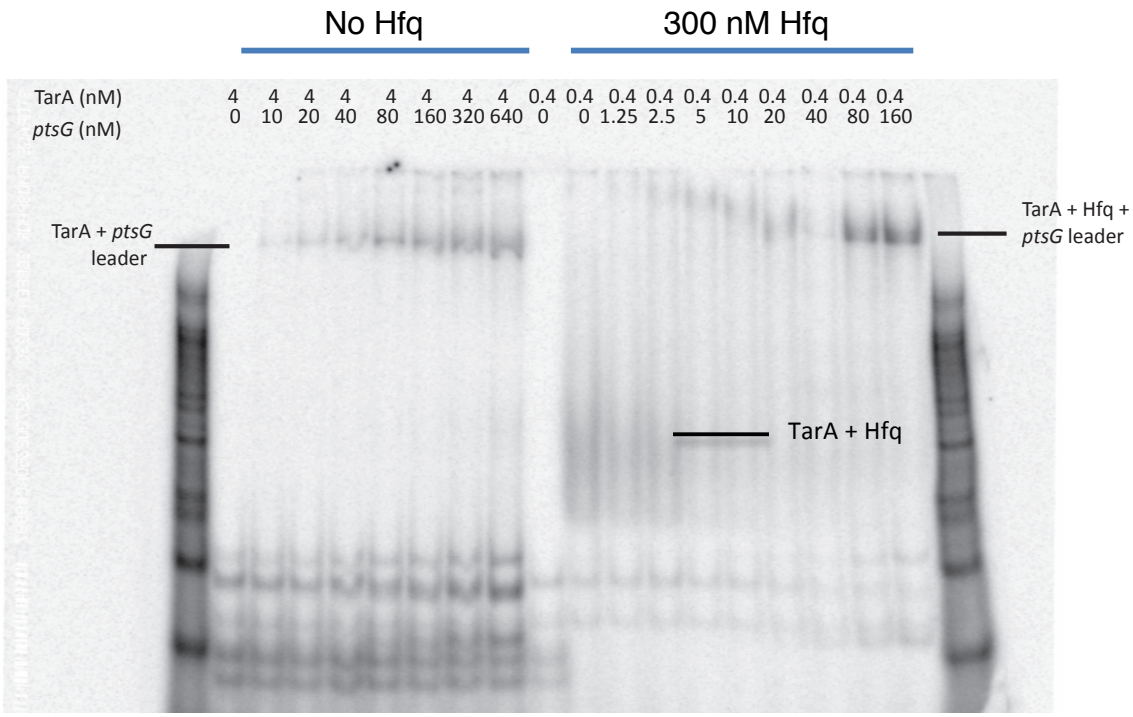


Figure 3.3. TarA binds Hfq and the *ptsG* leader sequence. TarA and a fragment of the *ptsG* transcript were *in vitro* transcribed and gel purified. TarA was end-labeled with ^{32}P. Hfq was expressed in *E. coli* and purified using the IMPACT purification system (NEB). Indicated concentrations of protein and RNA were mixed in native buffer and incubated at 25°C for 60 minutes, then subjected to electrophoresis on a nondenaturing 5% polyacrylamide gel and exposed to a phosphor storage screen. Image was obtained using a Typhoon phosphorimager.

Binding between TarA and the *ptsG* leader sequence is specific

We next sought to determine the specificity of binding between TarA and *ptsG* leader. We reasoned that specific binding would be disrupted by addition of excess unlabeled TarA but not by addition of excess unlabeled, unrelated RNA. We repeated the gel-shift experiment with unlabeled competitor RNA (TarA and yeast tRNA) in 10-fold and 100-fold excess (Figure 3.4). Again we observed an Hfq::TarA complex when those were the only two molecules in the reaction, and this complex disappeared in favor of a larger complex, presumably of Hfq:TarA:*ptsG*, when *ptsG* leader was added. This tri-molecular complex was undetectable when 100-fold excess cold TarA, but not yeast tRNA, was added to the reaction mixture. This result suggests that TarA complex formation is specific.

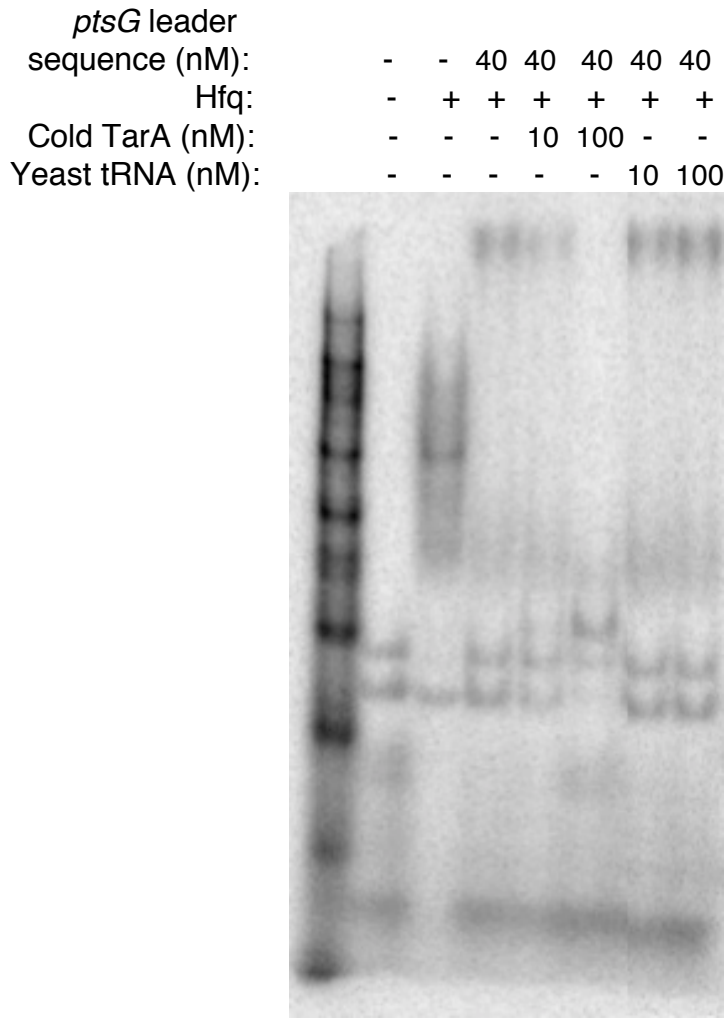


Figure 3.4. Binding between TarA and the *ptsG* leader sequence is specific. Radioactively-labeled *in vitro* transcribed TarA (1 nM) was mixed with the listed RNA and protein amounts in native buffer and incubated at 25°C for 60 minutes. Unlabeled TarA was used as a competitor (cold TarA). Samples were then subjected to electrophoresis on a 5% native gel; the gel was then exposed to a storage phosphor screen and imaged with a Typhoon phosphorimager.

Deletion of either putative Hfq binding site does not affect binding of TarA to Hfq or the *ptsG* transcript

Previously, we had identified two putative Hfq binding sites in the sequence of TarA: an interior UAUUAA sequence and the poly U terminator tail (Richard et al., 2010). We transcribed copies of TarA lacking each of these sites and a double deletion construct missing both sequences. These transcripts were then labeled and used in further gel-shift experiments. All three mutant constructs displayed binding patterns similar to wildtype for both Hfq and the *ptsG* leader sequence (Figures 3.5, 3.6, and 3.7). We observed the same banding patterns as those featuring full-length TarA, suggesting that neither putative binding site was required for Hfq binding or for binding to the *ptsG* transcript. Additionally, they did not appear to serve redundant functions, as the double deletion construct still bound Hfq and the *ptsG* transcript. The labeled double deletion TarA sample was less radioactive than the labeled full-length TarA, explaining the relative faintness of the bands on the right side of the gel in Figure 3.7. Despite the difference in band intensities, it still seems that the shift patterns are similar between the samples.

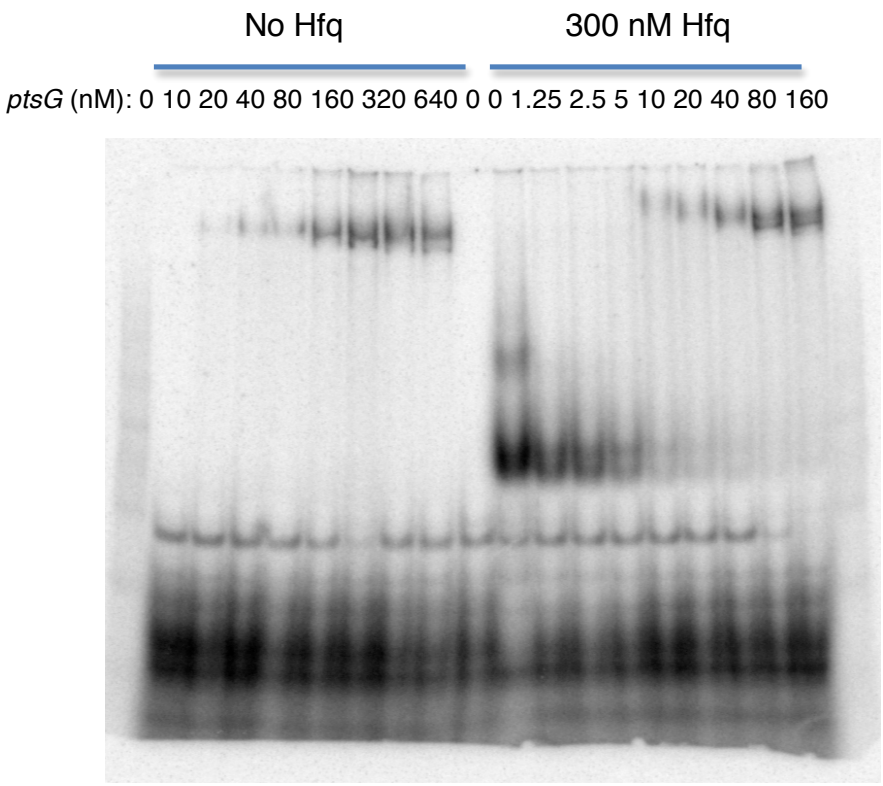


Figure 3.5. Deletion of the internal UAUUAA site has no effect on Hfq or *ptsG* leader binding. RNA and proteins were mixed in the indicated concentrations with 1 nM radioactively-labeled TarA in native buffer and incubated at 25°C for 60 minutes. Products were separated on a native 5% polyacrylamide gel and imaged with a Typhoon phosphorimager.

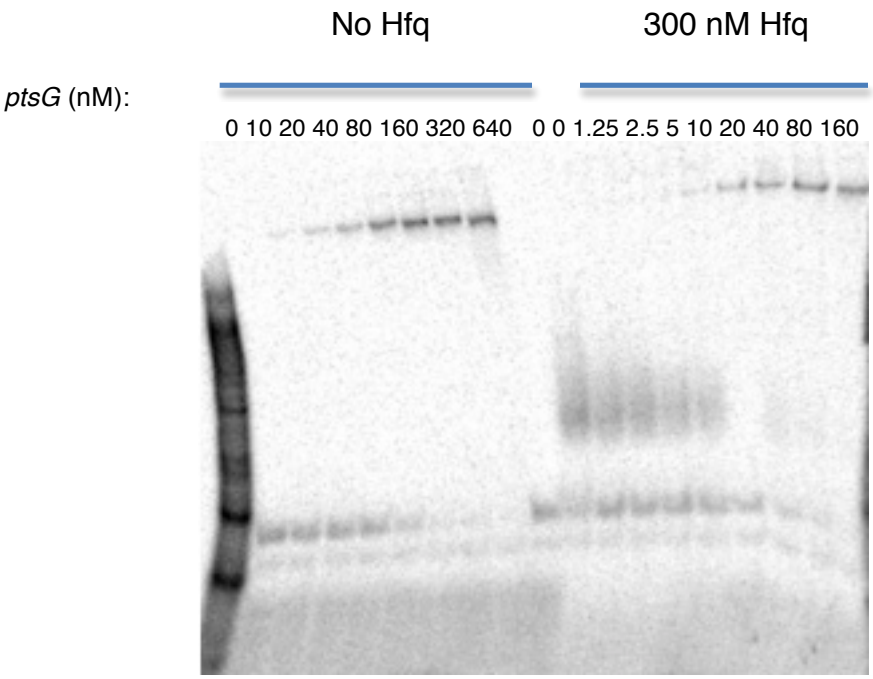


Figure 3.6. Deletion of the poly U tail does not alter Hfq or *ptsG* leader binding. Listed concentrations of *ptsG* leader sequence RNA and Hfq were mixed with 1 nM radioactively-labeled TarA in native buffer and incubated at 25°C for 60 minutes. Samples were then run on a 5% native polyacrylamide gel and imaged with a Typhoon phosphorimager.

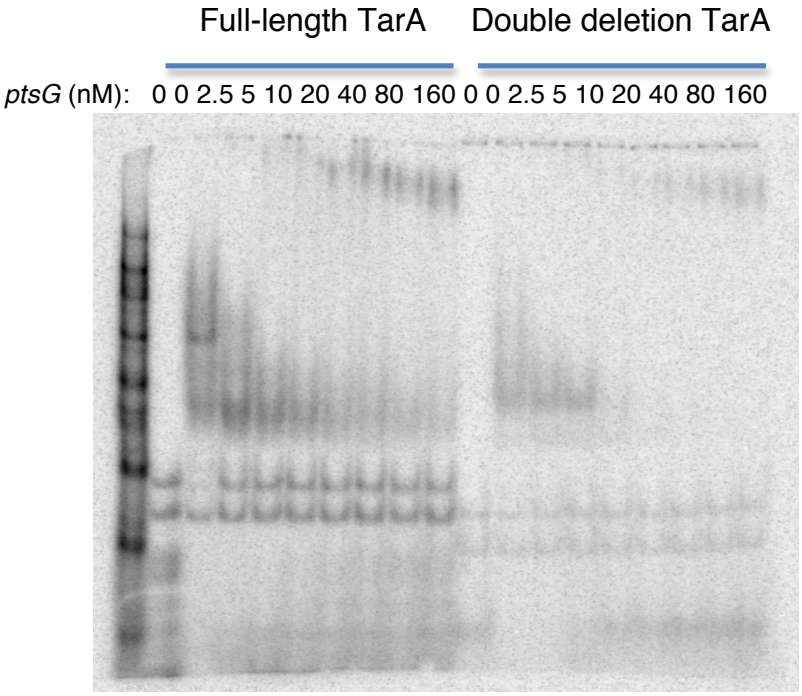


Figure 3.7. Deletion of both putative Hfq binding sites does not alter Hfq or *ptsG* leader binding. Indicated concentrations of *ptsG* leader sequence RNA were mixed with 1 nM radioactively-labeled TarA and 300 nM Hfq in native buffer and incubated at 25°C for 60 minutes. Samples were then run on a 5% native polyacrylamide gel and imaged with a Typhoon phosphorimager.

A. B.

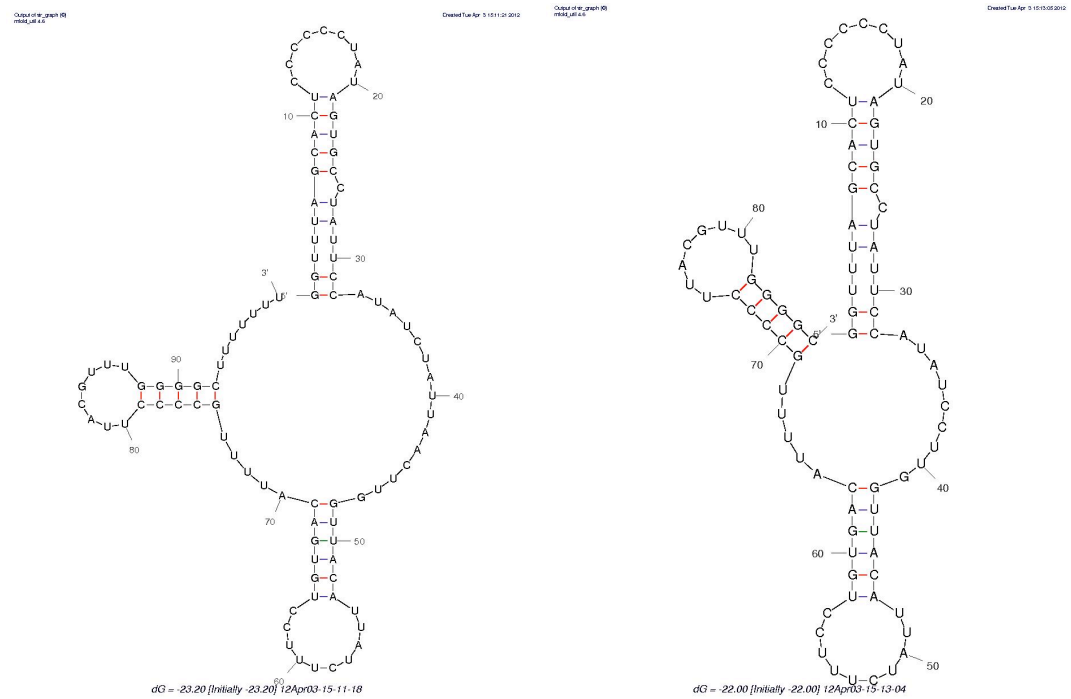


Figure 3.8. Predicted secondary structure of full length and double deletion TarA constructs. The sequences of full length TarA (A) and TarA lacking both the putative internal Hfq binding site and the poly U tail (B) were entered into Mfold (http://mfold.rna.albany.edu/?q=mfold). Pictured are the predicted folding patterns of these molecules. The AUUUUG mentioned as a third potential Hfq binding site is found on the left side of the center loop of both constructs.

Binding of a competitor RNA oligo blocks Hfq binding to the double deletion TarA construct but not to the full-length TarA

A third potential Hfq binding site in TarA, AUUUUG, resembles the binding site for Hfq in *Staphylococcus aureus* (AUUUUUG). This site is accessible in the central loop of the predicted Mfold structure of the double deletion construct (Figure 3.8). We hypothesized that this may be a binding site for *V. cholerae* Hfq as well. To test this, we used an RNA oligo complementary to that sequence as a competitive inhibitor in the gel-shift assay. This oligo features a UAAAAC sequence flanked by two nucleotides on either side. As a control, we utilized an oligo with the same bases but out of order (AGACAGAGUCG).

We repeated the gel-shift assay for both the full-length and double deletion TarA constructs, adding in either the competitor oligo or the scrambled oligo in 100-fold excess. While neither oligo altered Hfq binding by full-length TarA, there was a slight alteration of Hfq binding by the double deletion upon addition of the competitor oligo. When the competitor oligo was present, we observed diminished intensity in the bound band, with less of this effect observed when the scrambled oligo was included (Figure 3.9). However, the major gel-shifted species in the presence of Hfq alone (black arrow) is present in similar amounts at the un-shifted position (blue arrow) when either the competitor or the scrambled oligo are used. As the competitor oligo affects binding of the double deletion TarA but not of the full-length TarA, we conclude that the third site is indeed a binding site for Hfq, and that it functions redundantly with one or both of

the previously identified sites. More experiments will be necessary to determine exactly which sites are required for binding to Hfq.

As neither oligo appears to disrupt binding of either TarA construct to the *ptsG* leader sequence, these data also suggest that Hfq binding is not essential for interactions between the two RNAs. Taking all of our data together, it seems likely that a major role for Hfq is to stabilize the TarA RNA.

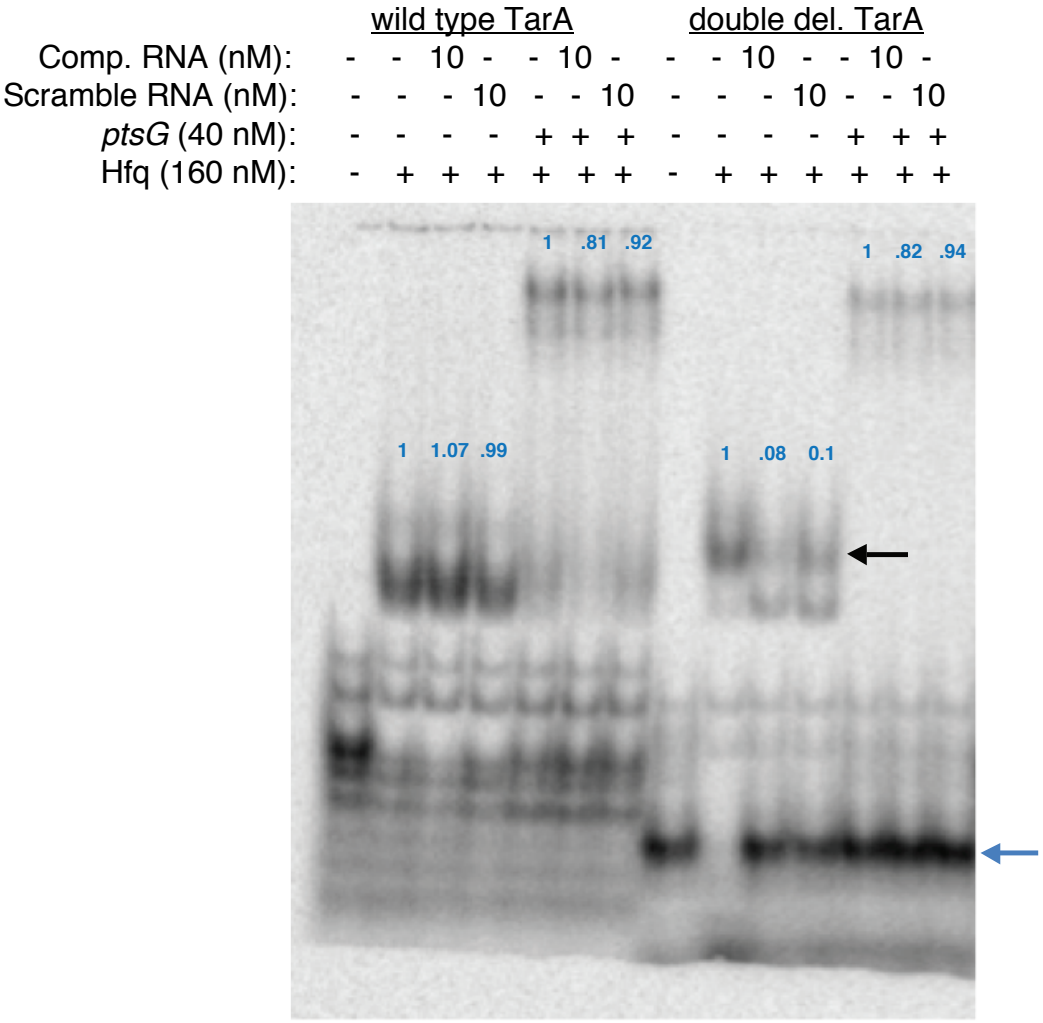


Figure 3.9. Binding of a competitor RNA oligo slightly alters Hfq binding to the double deletion TarA construct but not to the full-length TarA. Indicated concentrations of protein and RNAs were mixed with 1 nM radioactively-labeled TarA in native buffer and incubated at 25°C for 60 minutes. Full-length TarA was used for the first seven samples, while the double deletion TarA was used for the second set. Samples were then separated on a 5% native polyacrylamide gel and imaged with a Typhoon phosphorimager. The density of bands for samples containing competitor oligos were compared to those of the corresponding untreated samples using ImageJ software.

Discussion

In this chapter, we explore the mechanism of regulation of the glucose import gene *ptsG* by the small regulatory RNA TarA. We hypothesized that this sRNA directly binds the 5' untranslated region of the *ptsG* transcript by base-pairing, and that this binding may be catalyzed by the small protein Hfq. To test these hypotheses, we took a biochemical approach and analyzed binding of these species *in vitro* through electrophoretic mobility shift assays. These experiments showed that radioactively-labeled wildtype TarA binds both Hfq and the *ptsG* leader sequence (likely from residue -133 to -2, as predicted by RNAhybrid) *in vitro*. Binding of TarA to *ptsG* occurs in both the presence and absence of Hfq; in Figure 3.3, we observe binding of TarA to the *ptsG* leader at all tested *ptsG* concentrations. This suggests that this protein is not a required catalyst of the binding interaction. Taking these findings together with our previous result that TarA is less stable in an *hfq* mutant strain, we hypothesize that a major role for Hfq in *ptsG* regulation is to stabilize the TarA sRNA (Richard et al., 2010). It is important to note that these experiments only provide evidence that binding of TarA to the *ptsG* transcript can occur in the absence of Hfq; it remains possible that binding would not lead to efficient destabilization of the *ptsG* transcript *in vivo* without Hfq. We have conducted preliminary experiments to determine the role of Hfq in *V. cholerae,* testing *ptsG* transcript levels by qRT-PCR in an *hfq* mutant compared to a wildtype strain. These experiments showed a slight *decrease* in *ptsG* transcript levels compared to the wildtype but also had

a high standard deviation between experiments (data not shown). The fact that *ptsG* transcript levels are not elevated in an *hfq* mutant grown under ToxT-inducing conditions suggests that Hfq is not required for *ptsG* regulation, but further experiments are required to fully test this hypothesis.

If it is confirmed that the primary role for Hfq in *ptsG* regulation is to stabilize TarA, this would be similar to its role in conjunction with the MicX sRNA; Hfq has been shown to promote stability of the sRNA itself, but plays no role in the actual regulation of the target genes VC0972, VC0620 or VCA1041 (Davis & Waldor, 2007). While similar to MicX in this way, it is notable that TarA differs from SgrS, its functional homolog in *E. coli*. Regulation of *ptsG* by SgrS requires Hfq; in an *hfq* mutant strain of *E. coli,* SgrS-mediated destabilization of the *ptsG* transcript is not observed (Morita *et al*., 2004, Vanderpool & Gottesman, 2004). This fundamental difference in the mechanism of regulation adds to the growing list of dissimilarities between these two molecules that serve the same general function.

This study tested the effects of three putative binding sites for Hfq in the sequence of TarA. Two of these sites (UAUUAA and AUUUUG) are internal and consist of AU-rich sequences, while the third is the poly-U tail of the molecule. We used single and double deletion constructs of TarA to assess the roles of the UAUUAA and poly-U sites. We found that all three species were able to bind both Hfq and the *ptsG* transcript in a manner similar to the full-length RNA. These experiments, while not ruling out these sequences as Hfq binding sites,

suggested that if either or both of these sites bind Hfq, they are not the sole binding sites. To test this hypothesis, we analyzed the AUUUUG site, which is one nucleotide shorter than previously characterized Hfq binding sites. We tested requirement for this site by adding a competitor RNA oligo complementary to this sequence to the gel-shift experiments. Addition of the oligo resulted in a slight disruption of binding between the double deletion TarA and Hfq, but not the full-length TarA and Hfq. We interpreted this result to mean that the AUUUUG site is capable of binding Hfq, but that one or both of the previously identified sites can also bind this protein. This suggests that there are either two or three Hfq binding sites present in the TarA molecule. Further experiments with TarA constructs missing the AUUUUG site in combination with the other two potential binding sites would conclusively determine which binding sites are required for interaction with Hfq.

While this study sheds light on the role of Hfq in *ptsG* regulation of TarA, it does not address the requirement for specific base-pairing interactions between the two RNA molecules. The software program RNAhybrid (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/) predicts extensive base-pairing contacts between TarA and the *ptsG* transcript (Figure 2.7). For other sRNAs, such as SgrS in *E. coli* and the Qrrs in *V. cholerae*, a minimal required base-pairing region has been defined (Maki et al., 2010, Bardill et al., 2011). We attempted to mutagenize TarA to study these base-pairing interactions, but were unsuccessful, likely due to the large number of changes introduced. This

question of how TarA and the *ptsG* leader sequence interact is open for future study and would benefit from analysis by alternative methods in the future. Toeprint analysis to specifically determine which base-pairing interactions are required for binding would more clearly define the mechanism of *ptsG* regulation. Additionally, we could use lead probing to pinpoint the specific folding isoforms that are amenable to binding both Hfq and the *ptsG* leader (Lindell *et al.*, 2002). These experiments would likely shed more light on the kinetics of binding between TarA, Hfq and the *ptsG* leader sequence.

Materials and Methods

Bacterial strains and plasmids

Vibrio cholerae N16961 was used as a source of genomic DNA for generation of the *in vitro* transcription template, and *Escherichia coli* strains JM101 and ER2566 were used for cloning and protein expression, respectively. Strains were stored at -70°C in LB broth supplemented with 20% glycerol and cultured in LB broth with shaking at 37°C, except where noted. Ampicillin was used at a concentration of 100 μg/ml when necessary. The plasmid pUC19 was used for cloning of the TarA construct lacking the internal Hfq binding site (UAUUAA). The GENEart Mutagenesis Kit (Invitrogen) was used to clone this 93 nt sequence from a genomic DNA template. *Vibrio cholerae* Hfq was cloned into pTYB21 (New England Biolabs) for protein expression and purification. Plasmids

were purified from overnight cultures using a Qiagen Miniprep Kit, while genomic DNA was isolated by boiling bacterial colonies in ddH_2O.

In vitro transcription of TarA and the *ptsG* leader sequence

Templates for *in vitro* transcription were generated by PCR amplification of either N16961 genomic DNA (full-length and truncated TarA, *ptsG*) or pUC19 carrying the TarA deletion construct. After cleanup using a Qiagen PCR Purification Kit, 2 μg of template DNA was incubated with 4 mM each UTP, ATP, CTP, and GTP and purified T7 polymerase (generously donated by the lab of David Engelke) in N4 buffer containing 20 mM DTT for 4 hours at 37°C. After incubation, RNA was ethanol-precipitated, resuspended in 0.3X TE and electrophoresed on a 6% polyacrylamide gel. Bands were visualized by UV shadowing, excised and placed into 1 milliliter of elution buffer (0.3 M NaAc pH 5.2, 1 mM EDTA, 0.1% SDS), which was then incubated overnight at room temperature. After elution, RNA was again ethanol-precipitated, washed, and resuspended in 20 μl UltraPure water and 10 μl TE pH 8.0. RNA concentration and purity were determined by OD_{260} and OD_{280}, respectively, using a NanoDrop 3000 spectrophotometer.

Labeling of *in vitro*-transcribed TarA.

Purified transcribed TarA was end-labeled with γ-^{32}P-ATP using T4 Polynucleotide Kinase (PNK, New England Biolabs). Briefly, 30 pmol of purified

TarA were mixed with 4 μCi γ-^{32}P-ATP and 1 μl T4 PNK in 1X PNK buffer in a total volume of 60 μl and incubated at 37°C for 60 minutes. Labeled RNA was then ethanol-precipitated, washed and resuspended in 30 μl TE pH 8.0 and stored at -20°C until use.

Expression and purification of Hfq.

Vibrio cholerae hfq was cloned into the pTYB21 plasmid using the SapI and SacI restriction sites. The plasmid was transformed into the expression strain ER2566. Cells carrying the plasmid were grown at 37°C until the OD_{600} reached 0.1, then isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4 mM and the culture was shifted to 16°C, where it was grown overnight with shaking. Cells were pelleted and resuspended in lysis buffer (20 mM Tris pH 8.5, 500 mM NaCl, 1 mM EDTA, 0.1% Triton X-100) then sonicated on ice. Lysate was cleared by centrifugation at 10,000 RPM for 30 minutes at 4°C. Meanwhile, a 10 ml PolyPrep Chromatograhpy Column (Bio-Rad) was loaded with 4 ml of chitin bead slurry and equilibrated. Cleared lysate was run through the column and allowed to empty by gravity flow. The column was then extensively washed with column buffer (20 mM Tris pH 8.5, 500 mM NaCl, 1 mM EDTA). Cleavage of the chitin-binding intein tag was induced by the addition of column buffer containing 50 mM DTT and incubation at room temperature for 36 hours. Protein was then eluted with fresh column buffer. Eight hundred microliter samples were collected and stored at -20°C. Samples were

tested for presence of protein by electrophoresis on a 12.5% polyacrylamide gel and staining with either Coomassie Brilliant Blue or an Invitrogen SilverQuest Staining Kit. Eluted protein was pooled and dialyzed against fresh column buffer at 4°C to remove the DTT and then concentrated using an Amicon Ultra spin column with a 3000 Da weight cutoff. Protein concentration was assessed by OD_{280} using a NanoDrop 3000 spectrophotometer.

Electrophoretic mobility shift assays

All binding reactions were performed in a final volume of 15 μl. The components of each particular experiment are listed in the figure labels and legends. Assays were performed basically as in (Bardill et al., 2011). Briefly, RNAs were heated to 70°C for 5 minutes then cooled on ice and were mixed (with protein, where indicated) in native buffer (20 mM Tris pH 8.0, 100 mM NH4Cl, 50 mM NaCl, 50 mM KCl, 5% glycerol) and incubated at room temperature for 60 minutes. After incubation, native loading dye was added, and the entire reaction was loaded into a well of a non-denaturing 5% polyacrylamide gel and electropheresed at 80V for 6 hours 30 minutes at 4°C. Following electrophoresis, the gel was dried in a gel-drying frame overnight and exposed to a storage phosphor screen (GE Healthcare). Images were obtained with a Typhoon Trio phosphorimager (Amersham).

References

Bardill, J. P., X. Zhao & B. K. Hammer, (2011) The Vibrio cholerae quorum sensing response is mediated by Hfq-dependent sRNA/mRNA base pairing interactions. *Molecular microbiology* **80**: 1381-1394.

Brescia, C. C., P. J. Mikulecky, A. L. Feig & D. D. Sledjeski, (2003) Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA* **9**: 33-43.

Davis, B. M. & M. K. Waldor, (2007) RNase E-dependent processing stabilizes MicX, a Vibrio cholerae sRNA. *Molecular microbiology* **65**: 373-385.

Lenz, D. H., M. B. Miller, J. Zhu, R. V. Kulkarni & B. L. Bassler, (2005) CsrA and three redundant small RNAs regulate quorum sensing in Vibrio cholerae. *Molecular microbiology* **58**: 1186-1202.

Lindell, M., P. Romby & E. G. Wagner, (2002) Lead(II) as a probe for investigating RNA structure in vivo. *RNA* **8**: 534-541.

Maki, K., T. Morita, H. Otaka & H. Aiba, (2010) A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of ptsG mRNA. *Molecular microbiology* **76**: 782-792.

Morita, T., H. Kawamoto, T. Mizota, T. Inada & H. Aiba, (2004) Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in Escherichia coli. *Molecular microbiology* **54**: 1063-1075.

Otaka, H., H. Ishikawa, T. Morita & H. Aiba, (2011) PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 13059-13064.

Richard, A. L., J. H. Withey, S. Beyhan, F. Yildiz & V. J. DiRita, (2010) The Vibrio cholerae virulence regulatory cascade controls glucose uptake through activation of TarA, a small regulatory RNA. *Molecular microbiology* **78**: 1171-1181.

Song, T., F. Mika, B. Lindmark, Z. Liu, S. Schild, A. Bishop, J. Zhu, A. Camilli, J. Johansson, J. Vogel & S. N. Wai, (2008) A new Vibrio cholerae sRNA modulates colonization and affects release of outer membrane vesicles. *Molecular microbiology* **70**: 100-111.

Vanderpool, C. K. & S. Gottesman, (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Molecular microbiology* **54**: 1076-1089.

Wadler, C. S. & C. K. Vanderpool, (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 20454-20459.

Zhang, A., S. Altuvia, A. Tiwari, L. Argaman, R. Hengge-Aronis & G. Storz, (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *EMBO J* **17**: 6061-6068.

Chapter IV

TarA may play a role in regulation of sialic acid catabolism in *Vibrio cholerae*

Summary

The bacterium *Vibrio cholerae* produces a ToxT-activated small regulatory RNA (sRNA), TarA, which negatively regulates the glucose import gene *ptsG.* In this study, we further describe the regulon of this sRNA by exploring its role in activation of the *nan* locus. The *nan* locus comprises several genes involved in catabolism of sialic acid, a nine-carbon sugar. Transcripts encoding the *nan* proteins were found to be present at lower levels in a *tarA* mutant compared to the wild type strain by microarray. Three of these genes were selected for further study by qRT-PCR; this analysis confirmed the microarray results. However, deletion of *tarA* had no effect on utilization of sialic acid as a carbon source. Further study indicated that *tarA* expression is not induced in sialic acid-containing media. While TarA seems to be involved in regulation of genes for sialic acid metabolism, the mechanism and physiological consequence of this are unclear.

Introduction

Vibrio cholerae, a Gram-negative bacterium that causes the disease cholera, can utilize diverse carbon sources for growth. It is endemic to the Indian subcontinent; in the environment, these bacteria are found in warm stagnant water, where their numbers are enriched during monsoon season. In this setting, the organisms associate with copepods, chironimid egg masses, and other marine species, forming biofilms on their chitinous exoskeletons (for review, see (Butler & Camilli, 2005)). Chitin consists of chains of N-acetylglucosamine, a glucose derivative, and can be broken down and metabolized by several types of bacteria, including *V. cholerae* (Nalin et al., 1979, Veldkamp, 1952).

Transmission of *Vibrio cholerae* to human hosts occurs after ingestion of contaminated water. Bacteria that survive passage through the extremely acidic environment of the stomach progress to the small intestine, where they associate with mucin and rapidly divide. Mucin is secreted by host goblet cells and comprises large glycoprotein complexes that are rich in sialic acid, a nine-carbon sugar found in both prokaryotes and eukaryotes. Presumably, *V. cholerae* use sialic acid liberated from mucin and other sugars present in the small intestine as carbon sources during infection.

Sialic acid catabolism has been extensively characterized in *Escherichia coli* and has also been examined in *Vibrio cholerae* (Vimr & Troy, 1985, Almagro-Moreno & Boyd, 2009). Proteins encoded by the *nan* locus import sialic acid

through the bacterial membrane and break the molecule down into fructose-6-phosphate through a series of enzymatic reactions; fructose-6-phosphate can further be used by the cell for energy. Mutants lacking any of the genes in this metabolic pathway are unable to grow on sialic acid media. Additionally, *Vibrio cholerae* strains missing the *nanA* gene are deficient for colonization of the infant mouse intestine when compared with a wild type strain and are outcompeted by the wildtype strain at a 17:1 ratio (Almagro-Moreno & Boyd, 2009). Thus, it is quite likely that sialic acid cleaved from mucin is a key carbon source for *V. cholerae* in the human intestine.

Multiplication of *V. cholerae* in the human gut induces secretion of cholera toxin (CT) by the bacterium. This toxin stimulates efflux of chloride ions by host epithelial cells into the intestinal lumen. The flow of ions into the lumen promotes fluid accumulation to maintain osmotic balance, and the result of this is profuse secretory diarrhea, which resembles rice water stool and contains up to 10^9 bacteria per milliliter. This results in shedding of the bacterium into the environment and further dissemination of *Vibrio cholerae* to new hosts.

V. cholerae produces several small regulatory RNAs (sRNAs), which control quorum sensing, carbon metabolism, and outer membrane vesicle formation, among other functions. Two of these RNAs, TarA and TarB, are transcriptionally activated by the virulence regulator ToxT, which also controls production of the major virulence factors of *Vibrio cholerae*. In order to discover the regulatory targets of TarA, we performed microarray analysis of a *tarA*

deletion mutant compared to the wild type strain. Our results suggest that the *nan* locus is activated by TarA, as in the *tarA* mutant, *nan* transcript levels were roughly 4 to 20-fold lower than the wild type strain. We thus sought to determine if this RNA governs sialic acid metabolism by *Vibrio cholerae.*

Results

TarA may enhance expression of the *nan* locus.

In order to discover potential regulatory targets of TarA, we analyzed microarrays to compare transcript levels in wild type and *tarA* mutant *Vibrio cholerae* strains. The mutant strain, AR24, features a markerless deletion of the *tarA* sequence and is otherwise isogenic to the wild type classical strain O395. We used the Significance Analysis of Microarrays (SAM) program to identify transcripts with a ≥2.0 fold change with a 1% False Discovery Rate (FDR) (Tusher *et al.*, 2001). In this experiment, fifteen different loci, including six encoding genes with metabolic functions, were differentially regulated between the strains. Of note, expression of several *nan* genes was ≥5-fold increased in the wild type strain compared to the *tarA* mutant (Table 4.1). The genes comprising the *nan* locus have been functionally characterized by other groups and encode enzymes responsible for catabolism of sialic acid, a nine-carbon sugar found in both prokaryotes and eukaryotes (Almagro-Moreno & Boyd, 2009).

Gene name	Fold change	Gene name	Fold change
ptsG	2.72	***siaM***	**0.16**
VC0113	0.48	***nanE***	**0.1**
glmS	0.43	***nanK***	**0.18**
tcpI	0.08	***nagA***	**0.11**
lipB	0.43	*VC2031*	0.53
VC1078	0.54	*glnA*	0.4
VC1659	0.33	*VCA0236*	0.45
nanM1	**0.04**	*gsk-2*	0.51
nanM2	**0.07**	*deaD*	0.47
nanA	**0.12**	*VCA0933*	0.25
siaQ	**0.3**	*nagB*	0.47
siaP	**0.1**		

Table 4.1. Microarray data. Genome expression analysis was carried out to compare transcript levels between wild type *Vibrio cholerae* and an isogenic *tarA* deletion mutant. Whole cell RNA was isolated from cells grown under ToxT-inducing conditions using Trizol, and microarray analysis was carried out as described in Materials and Methods. Shown are the fold changes in transcript level in the *tarA* mutant compared to the wild type strain. Genes found in the *nan* locus are shown in bold.

To confirm the microarray results, we tested transcript levels of several *nan* genes (*nanA, nanK,* and *nanR*) in wild type and *tarA* mutant strains by quantitative real-time polymerase chain reaction (qRT-PCR). While the fold changes in transcript levels observed by qRT-PCR were not as drastic as those seen by microarray, the general trend of activation by TarA appears to be confirmed. We found that *nanA, nanK* and *nanR* transcript levels were 0.35, 0.29 and 0.45 times as abundant in the *tarA* mutant as compared to the wild type strain, respectively (Figure 4.1).

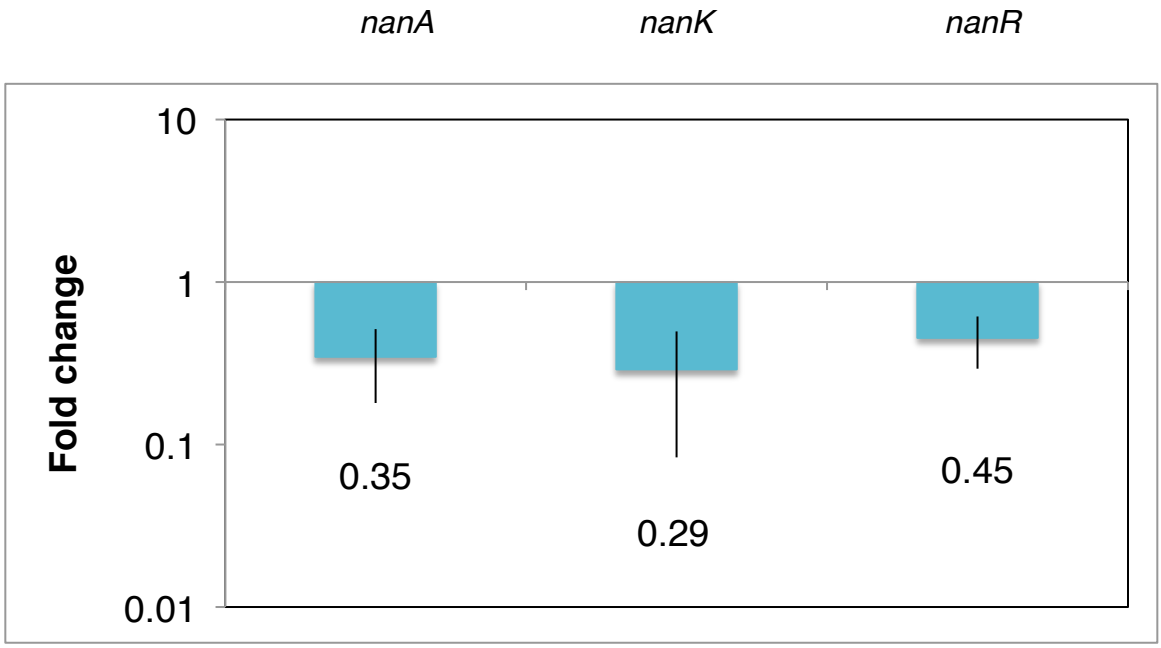


Figure 4.1. Confirmation of *nan* locus regulation by qRT-PCR analysis. Cultures of the listed strains were grown under ToxT-inducing conditions for 7 hours and total RNA was isolated using Trizol. cDNA was prepared from equivalent amounts of RNA for each sample, and cDNA was used for SYBR Green qRT-PCR. Numbers indicate fold change in transcript level relative to the wildtype, calculated using the ΔΔCT method and *recA* transcript levels as an internal control. Experiment performed by Mariah Gavin.

The *nan* locus comprises two transcripts.

Many previously characterized small regulatory RNAs regulate their targets by base-pairing with the 5' untranslated region of their transcripts. In order to more accurately assess regulation of the *nan* locus, it was first necessary to determine which genes were co-transcribed. Analysis of the annotated sequence suggested that the *nan* locus comprises two polycistronic transcripts. To test this, we performed reverse transcriptase polymerase chain reaction (RT-PCR) on cDNA synthesized from RNA isolated from wild type *Vibrio cholerae* and visualized the PCR products via agarose gel electrophoresis. We reasoned that if two neighboring genes were co-transcribed, we would be able to amplify an intergenic fragment by RT-PCR because the transcript would include both of those genes. If the genes were not co-transcribed, the PCR reaction would not produce a band on an agarose gel. We designed oligonucleotide primers that would amplify across each of the predicted intergenic regions and used these for the RT-PCR reactions. As a negative control, we designed primers that would amplify across the intergenic region between the two presumed transcripts. After visualization by ethidium bromide staining of agarose gels, we concluded that the predicted transcript organization was correct (Figure 4.2). The *nan* genes comprise two polycistronic transcripts that are divergently transcribed.

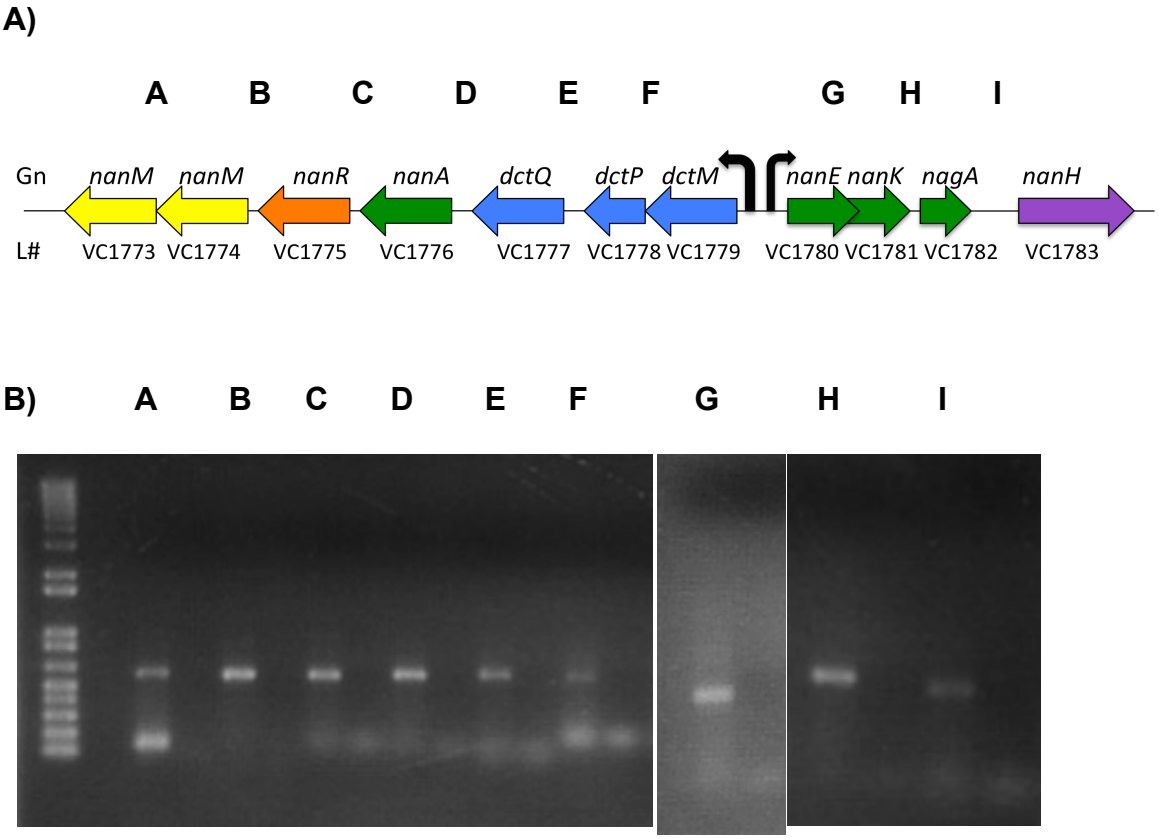


Figure 4.2. The *nan* locus comprises two transcripts. A) Schematic drawing of *nan* locus. Letters indicate junctions queried by RT-PCR. B.) PCR-amplified intergenic regions. Empty lanes in between bands correspond to identical samples run without reverse transcriptase to ensure that RNA samples were free of DNA contamination. Experiments performed by Kaitlin Flynn.

Deletion of *tarA* has no effect on growth in sialic acid-containing media.

Based on the observation that TarA appears to increase expression of the *nan* genes, we hypothesized that the consequence of decreased *nan* transcript levels would be growth deficiency in media with sialic acid as a sole carbon source. We tested growth in both liquid and solid MOPS minimal media containing sialic acid, N-acetylglucosamine (which may be metabolized by *nag* genes found adjacent to the *nan* genes), and mucin as sole carbon sources. In all media tested, the *tarA* mutant grew as well as the wild type, reaching similar final optical densities in the same amount of time (Figures 4.3 and 4.4). On solid media, the mutant strain formed colonies similar in size and number to the wild type (data not shown).

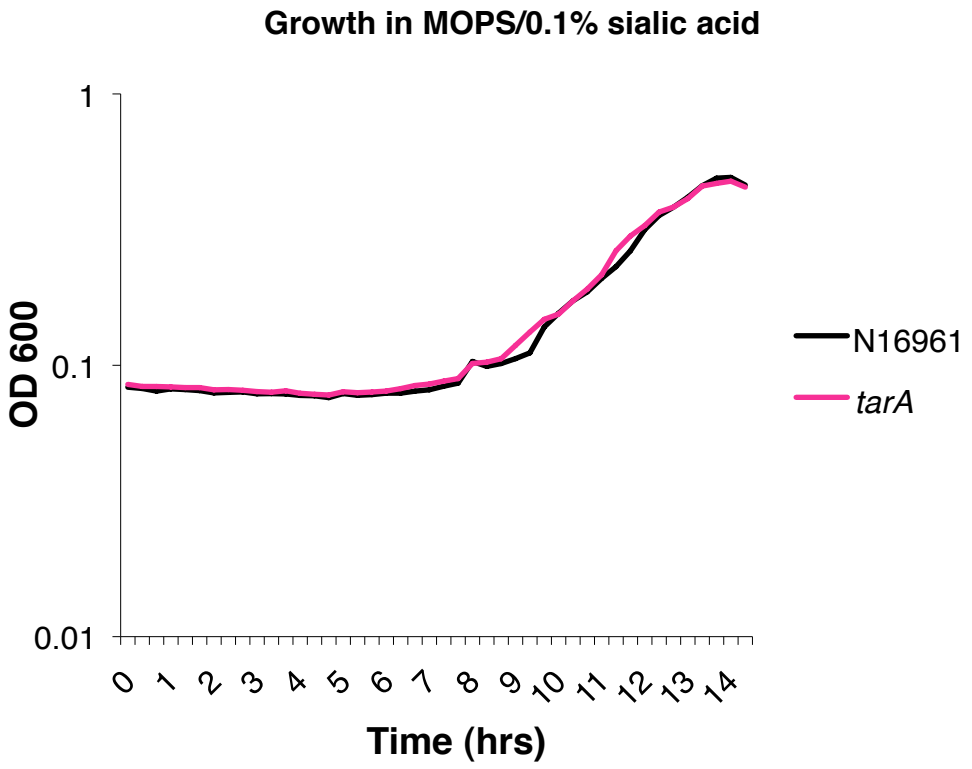


Figure 4.3. Deletion of *tarA* has no effect on growth in media containing sialic acid as a carbon source. A.) Sialic acid growth. Wild type (N16961) and *tarA* strains of *Vibrio cholerae* were grown overnight in LB at 37°C with shaking, then subcultured 1:100 into fresh LB. Once this culture reached mid-log phase, cells were spun down, washed, and resuspended in MOPS minimal media supplemented with 0.1% sialic acid. Cultures were then dispensed into a 96-well plate and grown for 14 hours in a Polarstar Omega plate reader, with OD_{600} readings every 20 minutes.

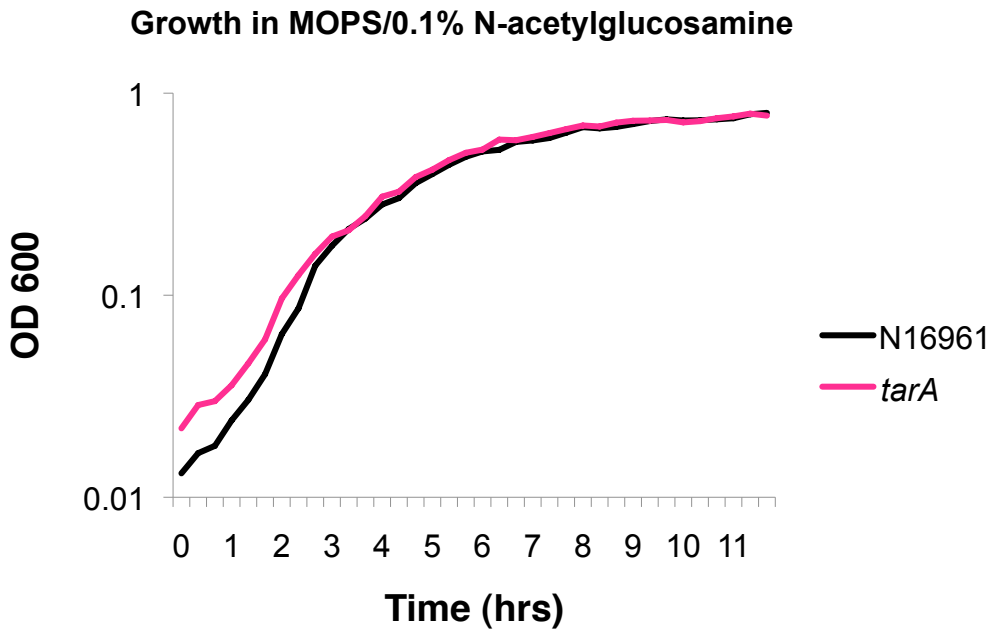


Figure 4.4. Deletion of *tarA* has no effect on growth in media containing N-acetylglucosamine as a carbon source. Cultures were prepared as in Figure 4.3, except they were diluted into MOPS minimal media containing 0.1% N-acetylglucosamine.

***tarA* expression is not induced by the presence of sialic acid.**

tarA expression is induced under conditions in which the virulence activator ToxT is active. For the classical strains of *Vibrio cholerae* analyzed, these conditions consist of growth at 30°C in LB pH 6.5. Under ToxT-inhibiting conditions (LB pH 8.5 at 37°C), no TarA is observed by Northern blot. As TarA appears to either activate transcription of the *nan* locus or stabilize the *nan* transcripts, we reasoned that *tarA* would be induced in the presence of sialic acid. To test this, we grew wild type and *tarA* mutant *Vibrio cholerae* in LB until late log phase, and then washed and resuspended the cells in MOPS minimal media supplemented with 0.1% sialic acid. We isolated total RNA after thirty minutes of incubation in this media, ran the RNA samples on a polyacrylamide gel, transferred to a nylon membrane, and then probed with radioactively labeled probe complementary to the TarA sequence. We did not observe any bands for TarA in the lanes corresponding to the sialic acid-grown samples (Figure 4.5), leading us to conclude that the presence of sialic acid does not induce TarA expression.

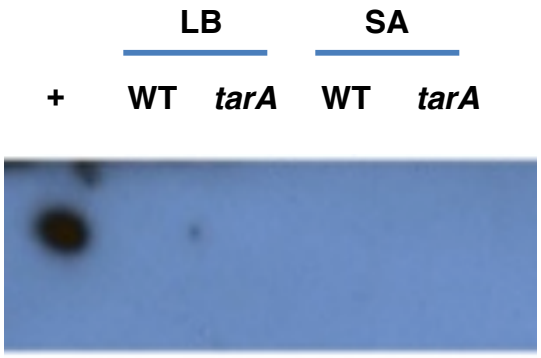


Figure 4.5. TarA is not produced in response to sialic acid. Northern blot analyzing *tarA* expression. Wild type and *tarA* mutant classical *Vibrio cholerae* were grown overnight in LB at 37°C, then subcultured into fresh LB and grown to mid-log phase. Cells were then spun down, washed and resuspended in LB or MOPS minimal media supplemented with 0.1% sialic acid and incubated at 37°C for thirty minutes. Whole cell RNA was isolated using Trizol, and Northern blot analysis was conducted as detailed in Materials and Methods. Wildtype cells grown under ToxT-inducing conditions were used as a positive control.

Discussion

In this study, we sought to further examine the regulon of the small regulatory RNA TarA, which is produced by *Vibrio cholerae.* Based on data from a microarray experiment comparing wild type and *tarA* mutant bacteria, we hypothesized that TarA activates expression of the *nan* locus. We confirmed positive regulation of three of these genes by quantitative real-time PCR, though the microarray indicated a stronger effect of TarA than the qRT-PCR did. Levels of the *nanA* transcript in the wild type were 8.33 fold higher than in the *tarA* strain by microarray, while the qRT-PCR data indicated a more modest 2.8 fold change. Similar results were observed for the *nanK* transcript as well. The results for *nanR*, the putative repressor of this locus, agreed between experiments (2.1 fold by microarray vs. 2.2 fold by qRT-PCR).

We reasoned that if TarA is a positive regulator of the *nan* locus, it is likely that deletion of *tarA* from the genome and the resulting lower levels of *nan* transcripts would render the bacteria less well-equipped for sialic acid metabolism. To test this, we grew both mutant and wild type strains in minimal media containing sialic acid as the sole carbon source. We were surprised to discover that the *tarA* mutant grew equally well in this media, reaching the same final optical density at the same rate as the wild type strain. We also tested growth of these strains on solid media containing mucin, which contains sialic acid, and again saw no effect for the *tarA* deletion mutant. Finally, we tested media supplemented with N-acetylglucosamine, as the *nagA* gene found at the

nan locus encodes N-acetylglucosamine-6-phosphate deacetylase, a protein involved in catabolism of this sugar (Yamano *et al*., 1997). We saw no growth defect for the *tarA* mutant in this media.

It is important to note that prior to dilution in minimal media for these growth experiments, cells were grown in plain LB at 37°C until they reached mid-log phase. For the El Tor strains used in this experiment, these conditions are not virulence-inducing, and TarA is not expressed. Thus, if the presence of sialic acid induced TarA production, we would expect to see a defect for the *tarA* mutant strain with regards to sialic acid growth. On the other hand, if sialic acid does not induce expression of TarA, the *tarA* mutant would behave like the wild type when grown in sialic acid media. Based on the phenotypes (or lack thereof) seen in sialic acid and mucin-containing media, we predicted that TarA expression would not be induced by the presence of sialic acid. We already established that TarA gene expression is under ToxT control (Richard *et al*., 2010) and, if it were also subject to control by sialic acid we would be interested in exploring the mechanisms of that. To examine this, we grew cells in LB to mid-log phase, then washed and resuspended them in MOPS media containing sialic acid, collecting whole cell RNA samples after thirty minutes of incubation in this media. We tested for TarA expression by Northern blot analysis and found that no detectable TarA was produced after sialic acid exposure.

Taking the growth phenotypes and the Northern blot data together, the fact that the presence of sialic acid does not induce TarA expression seems to

explain the lack of growth phenotypes observed in the different media tested. Additionally, during colonization of the human intestine, TarA would already be expressed because ToxT is active under those conditions, so there would be no need for control by sialic acid in this environment. Repeating these experiments with pre-growth in AKI – the conditions that stimulate *toxT* expression in El Tor *V. cholerae* – could be instructive. This would maximize TarA levels in the wild type strain prior to dilution into sialic acid media, allowing for the biggest difference in growth phenotypes to be observed. We predict that growth of cells in AKI conditions prior to exposure to sialic acid will cause the wild type strain to grow more quickly in MOPS/sialic acid than those grown in LB, as they will be pre-loaded with TarA and consequently will have enhanced expression of the *nan* genes. Additionally, these experiments could be performed with the plasmid-borne inducible *tarA* used in the *ptsG* regulation experiments.

The experiments detailed here suggest that further investigation into the exact regulatory role of TarA on the *nan* locus is certainly warranted. Among the proteins encoded by the *nan* locus is NanR, a putative repressor of this set of genes. While the role of NanR in regulation of sialic acid metabolism has been well studied in *E. coli*, its role in *V. cholerae* is less established. In *E. coli,* NanR binds to the *nanA* promoter in the absence of sialic acid, blocking transcription of *nanA* and the downstream genes *nanT, nanE,* and *nanK*. When sialic acid is present, binding of this molecule to NanR relieves repression at this locus (Kalivoda *et al.*, 2003). This type of regulation is presumed for *Vibrio cholerae*

but has not been directly shown. Additionally, the *nan* genes of *V. cholerae* are organized differently than those of *E. coli:* they make up two divergent transcripts separated by an intergenic region of only 31 bases. Future experiments should explore whether NanR binds this region in the presence and absence of sialic acid.

Based on our gene expression analysis, *nanR* transcripts are approximately 2-fold decreased in a *tarA* deletion mutant, while transcript levels for the other genes at that locus are more strongly down-regulated. This suggests that TarA induces or de-represses the putative repressor of this locus, which would ultimately lead to down-regulation of the *nan* genes. As this scenario does not agree with our observations, we favor a model in which TarA regulates the *nan* locus directly, and that there are simply two different mechanisms by which these genes are controlled (Figure 4.6). In the environment, where there is presumably little sialic acid present and TarA is not expressed, NanR may bind the intergenic region between the two *nan* transcripts, preventing transcription of the entire set of genes. Once introduced to the host, virulence-inducing conditions prompt activation of ToxT and expression of TarA. TarA accumulates to high levels in time for passage to the small intestine, where free sialic acid binds NanR, removing it from the DNA it is bound to and allowing for transcription of the *nan* genes. TarA could then potentially bind the *nan* transcripts and either stabilize them or conformationally change the upstream untranslated region to allow for ribosome binding. As *nanR* is co-transcribed with the other *nan* genes,

this scenario would explain the apparently contradictory microarray results mentioned previously.

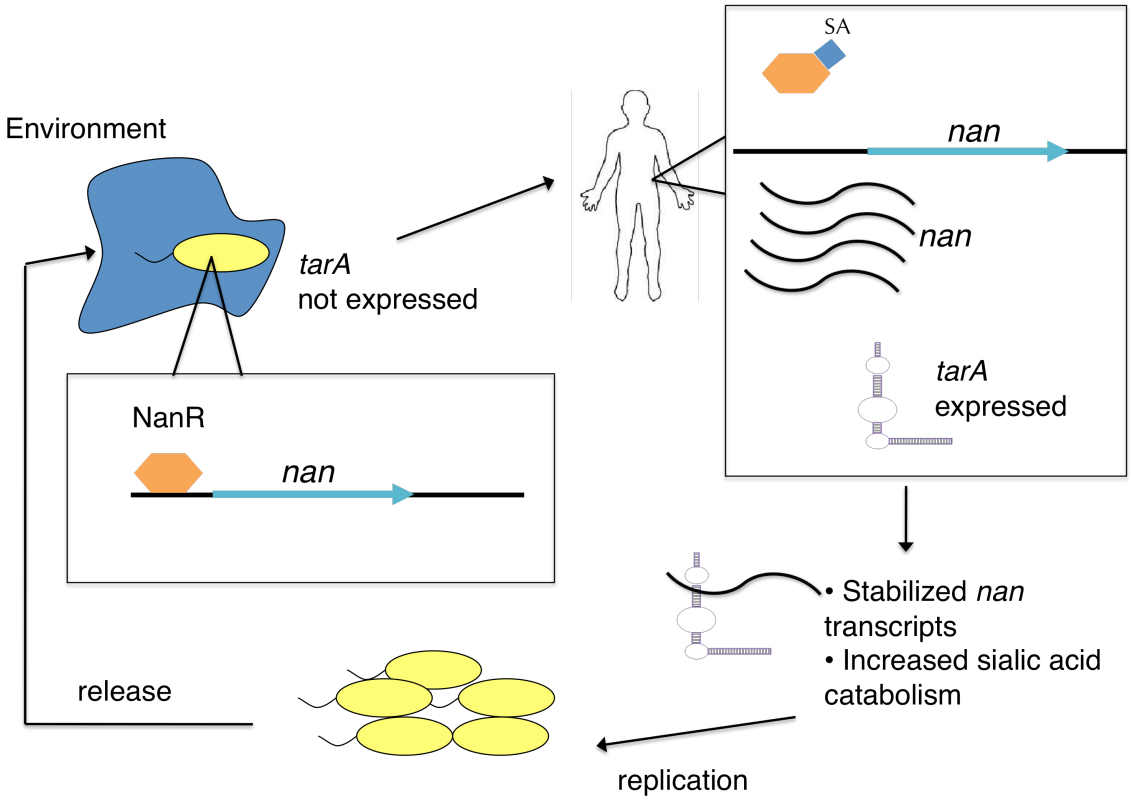


Figure 4.6. Model of *nan* regulation by TarA and NanR.

This study has shown that the small RNA TarA likely plays a role in regulation of sialic acid metabolism. To our knowledge, control of this process by a small RNA has not been described in any other bacterial species. Further characterization of this locus and its regulation by TarA has the potential to define a new paradigm in the study of bacterial physiology.

Materials and Methods

Bacterial strains and growth conditions.

The experiments detailed in this study utilized the classical strain O395, the El Tor strain N16961, and their derivatives AR24 and AR43, which are markerless *tarA* deletion mutants in those backgrounds, respectively. Creation of these strains is detailed in the Materials and Methods section of Chapter II. Strains were stored at -70°C in Luria-Bertani (LB) broth supplemented with 20% glycerol, and were streaked on solid LB agar supplemented with 100 μg/ml streptomycin. For routine growth, bacteria were cultured in LB broth at 37°C with aeration at 200 rpm. For experiments requiring ToxT-inducing conditions (classical biotype), cells were grown in LB pH 6.5 at 30°C with shaking. Growth under ToxT-inhibiting conditions utilized LB pH 8.5 at 37°C.

RNA isolation for microarray analysis and qRT-PCR.

Cultures of O395 and *tarA* strains were grown overnight and then diluted 1:100 into LB pH 6.5 and grown at 30°C. Once cultures reached an OD_{600} ~0.35, they were diluted 1:10 and grown until OD_{600} ~0.35, and 2 ml aliquots of the samples were harvested. Cell pellets were resuspended in TRIzol (Invitrogen) and total RNA was isolated according to the manufacturer's protocol as described previously (Yildiz *et al.*, 2004). Whole genome expression analysis was performed using 70-mer oligos representing the open reading frames present in the *V. cholerae* N16961 genome. cDNA synthesis, microarray hybridization, and analysis of expression data were performed as described previously (Beyhan *et al.*, 2006) using two biological and four technical replicates (for each biological replicate).

RNA samples for qRT-PCR were treated with DNase I (Roche) according to manufacturer's instructions, run on an agarose gel to check quality, and quantified by measuring the OD_{260}. 2.5 μg of each sample were treated with Moloney murine leukemia virus (M-MLV) reverse transcriptase (Invitrogen) according to the manufacturer's specifications.

Quantitative Real-Time PCR (qRT-PCR) and Reverse Transcriptase PCR (RT-PCR).

For detection of transcripts by qRT-PCR, primers amplifying a 200-bp region in the middle of the mRNA were used with SYBR Green Master Mix (Stratagene) on an ABI 7500FAST thermocycler. Primers were designed using the OligoPerfect tool (Invitrogen). Each test was performed in triplicate at least three times, and fold change

in expression was calculated using the ΔΔCT method (Pfaffl, 2001), with *recA* transcript levels used as controls between cDNA samples. For amplification across intergenic regions by RT-PCR, prepared cDNA was amplified by standard PCR using Taq polymerase (Invitrogen) and primers that would amplify only the intergenic region of interest. Five microliters of each 50 microliter reaction were then run on a 1% agarose gel, and bands were visualized after ethidium bromide staining.

Analysis of growth in sialic acid-containing media.

Strains N16961 and AR43 were grown overnight in LB, then diluted 1:100 into fresh LB and grown to mid-log phase (~six hours). OD_{600} was measured for each culture, and $1x10^9$ bacteria of each strain were pelleted, washed in PBS, and resuspended to a final OD_{600} of 2 in PBS. These suspensions were then diluted 1:50 (final OD_{600} = 0.04) into LB and MOPS minimal medium supplemented with 0.1% sialic acid or N-acetylglucosamine (Neidhardt *et al.*, 1974), and dispensed into a 96-well plate. The plate was incubated at 30°C with constant shaking for 14 hours in an Omega Polarstar plate reader, with OD_{600} measured every 20 minutes. Each experiment was run in triplicate three times.

Northern blots.

O395 and AR24 (*tarA*) *V. cholerae* were grown overnight in LB with shaking at 37°C and then subcultured into fresh LB broth and grown at 37°C with shaking. Once the cultures reached mid-log phase (OD_{600} ~0.5), cells were enumerated, spun down,

washed with PBS, and resuspended in either MOPS minimal media supplemented with 0.1% sialic acid or fresh LB. These cultures were then incubated at 37°C with shaking for 30 minutes. Whole cell RNA was then purified from roughly 1.5 x 10^9 cells from each culture. RNA purification was performed using the TRIzol reagent as specified by the manufacturer (Invitrogen). For TarA visualization, 7 μg total RNA from each sample were separated by 8% polyacrylamide gel electrophoresis and transferred to an Amersham Hybond nylon membrane (GE Healthcare) (0.45 mm) membrane with a Semi-Phor apparatus (Hoefer). An oligonucleotide probe complementary to nt 12-36 of TarA was radiolabeled with γ–^{32}P-ATP (Perkin-Elmer) with T4 polynucleotide kinase (New England Biolabs). Probing conditions were as previously described (Sambrook, 1989).

References

Almagro-Moreno, S. & E. F. Boyd, (2009) Sialic acid catabolism confers a competitive advantage to pathogenic vibrio cholerae in the mouse intestine. *Infection and immunity* **77**: 3807-3816.

Beyhan, S., A. D. Tischler, A. Camilli & F. H. Yildiz, (2006) Differences in gene expression between the classical and El Tor biotypes of Vibrio cholerae O1. *Infection and immunity* **74**: 3633-3642.

Butler, S. M. & A. Camilli, (2005) Going against the grain: chemotaxis and infection in Vibrio cholerae. *Nature reviews. Microbiology* **3**: 611-620.

Kalivoda, K. A., S. M. Steenbergen, E. R. Vimr & J. Plumbridge, (2003) Regulation of sialic acid catabolism by the DNA binding protein NanR in Escherichia coli. *Journal of bacteriology* **185**: 4806-4815.

Nalin, D. R., V. Daya, A. Reid, M. M. Levine & L. Cisneros, (1979) Adsorption and growth of Vibrio cholerae on chitin. *Infection and immunity* **25**: 768-770.

Neidhardt, F. C., P. L. Bloch & D. F. Smith, (1974) Culture medium for enterobacteria. *Journal of bacteriology* **119**: 736-747.

Pfaffl, M. W., (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* **29**: e45.

Richard, A. L., J. H. Withey, S. Beyhan, F. Yildiz & V. J. DiRita, (2010) The Vibrio cholerae virulence regulatory cascade controls glucose uptake through activation of TarA, a small regulatory RNA. *Molecular microbiology* **78**: 1171-1181.

Sambrook, J., Fritsch, E.F., Maniatis, T., (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Tusher, V. G., R. Tibshirani & G. Chu, (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 5116-5121.

Veldkamp, H., (1952) Aerobic decomposition of chitin by micro-organisms. *Nature* **169**: 500.

Vimr, E. R. & F. A. Troy, (1985) Identification of an inducible catabolic system for sialic acids (nan) in Escherichia coli. *Journal of bacteriology* **164**: 845-853.

Yamano, N., N. Oura, J. Wang & S. Fujishima, (1997) Cloning and sequencing of the genes for N-acetylglucosamine use that construct divergent operons (nagE-nagAC) from Vibrio cholerae non-O1. *Bioscience, biotechnology, and biochemistry* **61**: 1349-1353.

Yildiz, F. H., X. S. Liu, A. Heydorn & G. K. Schoolnik, (2004) Molecular analysis of rugosity in a Vibrio cholerae O1 El Tor phase variant. *Molecular microbiology* **53**: 497-515.

Chapter V

Discussion

This dissertation details the discovery and characterization of TarA, a small regulatory RNA produced by the bacterium *Vibrio cholerae.* The goal of the study was to determine the function and regulatory targets of this small RNA after its initial discovery by a former postdoctoral fellow in the lab, Dr. Jeffrey Withey. Our findings are discussed in detail below.

TarA is a conserved, ToxT-regulated small RNA

The small RNA TarA is transcribed from a promoter upstream of the gene *tcpI*, which is located on the Vibrio pathogenicity island (VPI) and is divergently transcribed from the gene *tcpP.* Identified during a study of toxboxes, or ToxT recognition sequences, this RNA is 99 nucleotides in length and is transcribed under conditions in which the protein ToxT is active (Figure 2.2). Its sequence is highly conserved among *Vibrio cholerae* isolates that possess the VPI (Table 2.2). TarA is readily detectable by northern blot after three hours of growth under virulence-inducing conditions but is not seen in cells lacking *toxT* or those grown under inhibiting conditions. We hypothesized that this RNA plays a regulatory role in *Vibrio cholerae* because it does not contain any open reading frames, is

highly conserved, and possesses putative binding sites for the small RNA-binding protein Hfq.

TarA represses the glucose import gene *ptsG*

To identify potential regulatory targets of TarA, we performed microarray analysis comparing transcript levels in a *tarA* deletion mutant to a wild type strain. Of the 23 genes identified, only one, *ptsG*, was present at levels higher than our two-fold cutoff in the *tarA* mutant than the wild type strain, suggesting that it is negatively regulated by TarA. Quantitative real-time PCR (qRT-PCR) confirmed this, demonstrating a 2.7-fold increase in transcript levels for the *tarA* mutant relative to the wild type. We reasoned that in a strain overexpressing TarA, *ptsG* would be expressed at lower levels, and growth in glucose would be hampered by inefficient uptake of the sugar from the media. Upon testing this hypothesis, we observed that growth in glucose is severely attenuated in this strain while growth in LB is unaffected (Figure 2.5).

Regulation of *ptsG* by TarA is likely due to base-pairing between TarA and the *ptsG* leader sequence

Many bacterial sRNAs function in *trans* by binding the mRNA encoding their target gene. This interaction is in many cases catalyzed by the protein Hfq, which can bind both sRNAs and their targets. Formation of this ternary complex can recruit either RNase E or the ribosome to degrade the transcript or promote

translation, respectively. As TarA is stabilized by Hfq (Figure 2.8) and negatively regulates *ptsG*, we hypothesized that it base-pairs upstream of the GUG of the *ptsG* transcript with the help of Hfq. Using electrophoretic mobility shift assays (EMSA), we found that TarA binds the *ptsG* leader sequence *in vitro* independently of Hfq (Figure 3.3). TarA does bind purified Hfq *in vitro* (Figure 3.2), but presence of this protein is not required for TarA-*ptsG* binding. We hypothesize that a major role of Hfq in *ptsG* regulation is to stabilize TarA and protect it from degradation by RNases in the cell. To further elucidate this, we wanted to determine the exact binding site for the Hfq protein on TarA. We cloned deletion constructs of TarA that lacked one or both of two putative Hfq binding sites, an internal UAUUAA sequence and the poly U terminator tail. When analyzed by EMSA, all three of these constructs bound Hfq and the *ptsG* leader sequence as efficiently as the full-length molecule (Figures 3.5, 3.6 and 3.7). Preliminary work using a competitor RNA oligo complementary to a third potential binding site, AUUUUG, indicated that this site appears to be important for Hfq, and that it likely functions redundantly with one or both of the other sites (Figure 3.8).

Future perspectives on *ptsG* regulation

While this study provides many details about the mechanism and physiological consequences of *ptsG* repression by TarA, it also prompts several new questions. As *tarA* expression is induced by ToxT, which is activated early

in infection upon access to the gastrointestinal tract of the host, it follows that *ptsG* expression is repressed in the human small intestine, the site of colonization. This seems counterintuitive at the outset, as glucose is one of the easiest sugars to metabolize. However, glucose is actively absorbed through the intestinal epithelium, and thus free glucose may be limited in this environment. Several studies have found that expression of glucose transporters in the small intestine is increased in cholera patients and infected animals (Flach *et al.*, 2004, Schiller *et al.*, 1997), and that the ability of these transporters to take up glucose is unaffected by cholera infection (Rohde & Cash, 1973). As glucose and sodium are co-transported in the intestine, this is likely an attempt by the host to reclaim ions from the intestinal lumen, and would result in lower local glucose concentrations. Repression of *ptsG* by TarA may simply shift the bacterium from uptake of glucose to usage of a more abundant carbon source. As *ptsG* has been shown to indirectly negatively regulate biofilm production, this could result in increased planktonic growth in the intestine.

Additionally, studies from the Camilli laboratory have shown that vibrios actively store sugars in the form of glycogen prior to expulsion from the host; this provides nutrient stores once the bacteria are released into the environment, which is nutrient-poor (Bourassa & Camilli, 2009). In *V. cholerae*, glycogen is continuously synthesized by the enzyme GlgC1 and degraded by GlgX. Under nitrogen-limiting conditions, GlgX is repressed, leading to accumulation of glycogen. Preliminary evidence has shown that another glycogen synthesis

gene, *glgC2*, is induced by RpoS, which is active during late stages of infection, suggesting that *V. cholerae* both increases glycogen production and decreases glycogen catabolism late in infection. Temporally, repression of *ptsG* by TarA may coincide with the final stages of glycogen storage. In this scenario, the bacteria would down-regulate glucose uptake because they have stockpiled enough nutrients for the next stage of the infectious cycle. Future experiments exploring the temporal expression of *tarA* and *ptsG* in an animal model of infection would likely shed light on this intriguing question.

This work also provides some insight into the mechanism of *ptsG* regulation by TarA and how this sRNA is different from its functional homolog in *E. coli,* SgrS. Like SgrS, TarA binds the 5' untranslated region of its target mRNA, and induced expression of the sRNA leads to destabilization of the *ptsG* transcript. However, *tarA* expression is induced by the virulence regulator ToxT, while *sgrS* is regulated by the protein SgrR and is expressed under conditions of stress induced by glucose-phosphate accumulation. Repression of *ptsG* by SgrS requires Hfq, which binds to a single site on the sRNA, the poly U tail. Conversely, TarA can bind the *ptsG* leader sequence in the absence of Hfq, and *ptsG* transcript levels in an *hfq* mutant are not increased compared to the wild type strain. Hfq does appear to be important for stabilization of TarA, but unlike SgrS, TarA possesses multiple redundant binding sites for Hfq.

An especially interesting area for future study is the kinetics of binding between TarA, Hfq and the *ptsG* transcript. TarA is predicted to fold into one

particular secondary structure by the prediction software Mfold (Figure 3.8). Our EMSA analysis suggests that there are several folding isoforms present under non-denaturing conditions, based on the multiple radioactive bands observed with unbound TarA. Furthermore, it appears that different folding isoforms are more favorable for binding Hfq and the *ptsG* transcript. In lanes where a shift due to Hfq binding is observed, we see depletion of one unbound TarA band, whereas in lanes where TarA is binding *ptsG*, a different TarA band diminishes in signal as binding increases. This phenomenon is well illustrated on the right side of Figure 3.5. Future experiments to determine the exact folding conformations adapted by TarA and the ability of each of these isoforms to bind both Hfq and the *ptsG* transcript will further elucidate the mechanism of binding between these species.

TarA regulates the *nan* locus in *Vibrio cholerae*.

Chapter IV of this study examines another set of potential TarA targets, the *nan* locus. The genes of this locus comprise two divergent transcripts and encode the proteins required for sialic acid catabolism in *V. cholerae.* In a *tarA* mutant strain, transcript levels for these genes are 3 to 20-fold lower than the wild type strain, suggesting that TarA positively regulates their expression. We hypothesized that a *tarA* mutant would be deficient for growth in minimal media containing sialic acid as a sole carbon source due to lowered levels of *nan* gene expression. However, the mutant strain grew just as well as wild type in this

media. Upon further examination, we found that the conditions under which we grew these cells did not induce *tarA* expression, which explains why we did not observe a difference between the wild type and *tarA* mutant strains.

Additionally, the protein NanR is encoded at this locus and is a putative repressor of these genes that has been well characterized in *E. coli* but is less well studied in *V. cholerae*. In *E. coli,* this protein binds the *nanA* promoter when sialic acid is not present, blocking transcription of *nanA* and its downstream neighbors. In media containing sialic acid, NanR binds this carbohydrate and no longer blocks the *nan* promoter (Kalivoda *et al.*, 2003). We propose that both TarA and NanR serve to regulate the *nan* genes in *Vibrio cholerae,* allowing the bacterium to most efficiently adapt to changing nutrient conditions in the environment and the host. When little to no sialic acid is present, such as during growth in the environment outside of human hosts, NanR likely binds the intergenic region between the two *nan* transcripts on the chromosome of *V. cholerae*, blocking transcription. Once the bacterium is ingested by the host, ToxT activates expression of *tarA* as *V. cholerae* passes through the gastrointestinal tract. The intestinal lumen is rich in mucin, which contains sialic acid that can be cleaved off by the *V. cholerae* protein NanH. In this environment, sialic acid could then bind NanR and relieve repression of the *nan* genes. TarA would be present at high levels at this point, and could serve to stabilize *nan* transcripts and/or promote translation of these transcripts. In this way, TarA and NanR could coordinately regulate expression of these genes to

best respond to environmental conditions. Further study is required to test these hypotheses and to flesh out the exact role of these regulators in sialic acid catabolism.

Summary

This work describes the ToxT-activated RNA TarA and adds a new member to the existing repertoire of small regulatory RNAs in *Vibrio cholerae*. We show that this molecule regulates carbon metabolism in this organism by repressing the glucose transporter PtsG and activating the *nan* locus, which encodes genes involved in sialic acid metabolism. TarA serves as a link between the virulence regulatory cascade and metabolism, allowing the bacterium to adapt to the changing nutrient conditions it encounters as it cycles between the environment and its niche in the host, the distal small intestine. These discoveries enrich our knowledge of *Vibrio cholerae* and small regulatory RNAs in general, and contribute to a deeper overall understanding of bacterial pathogenesis.

References

Bourassa, L. & A. Camilli, (2009) Glycogen contributes to the environmental persistence and transmission of Vibrio cholerae. *Molecular microbiology* **72**: 124-138.

Flach, C. F., S. Lange, E. Jennische & I. Lonnroth, (2004) Cholera toxin induces expression of ion channels and carriers in rat small intestinal mucosa. *FEBS letters* **561**: 122-126.

Kalivoda, K. A., S. M. Steenbergen, E. R. Vimr & J. Plumbridge, (2003) Regulation of sialic acid catabolism by the DNA binding protein NanR in Escherichia coli. *Journal of bacteriology* **185**: 4806-4815.

Rohde, J. E. & R. A. Cash, (1973) Transport of glucose and amino acids in human jejunum during Asiatic cholera. *The Journal of infectious diseases* **127**: 190-192.

Schiller, L. R., C. A. Santa Ana, J. Porter & J. S. Fordtran, (1997) Glucose-stimulated sodium transport by the human intestine during experimental cholera. *Gastroenterology* **112**: 1529-1535.