August 2015

Characterization of the AtsR/AtsT global regulatory pathway in Burkholderia ceocepacia

Maryam Khodai-Kalaki
mkhodaik

Supervisor
Dr. Miguel Valvano
The University of Western Ontario

Graduate Program in Microbiology and Immunology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Characterization of the AtsR/AtsT global regulatory pathway in *Burkholderia cenocepacia*

(Thesis format: Integrated Article)

by

Maryam Khodai-Kalaki

Graduate Program in Microbiology and Immunology

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

Phosphorylation cascades governed by two-component signal transduction systems provide key signalling mechanisms in bacteria, simple eukaryotes and higher plants, allowing them to translate signals into adaptive responses. These regulatory pathways consist of a transmembrane sensor protein that responds to an environmental cue leading to autophosphorylation, followed by the transfer of the phosphate to a cytoplasmic response regulator. Here, I study AtsR, a membrane-bound hybrid sensor kinase of *Burkholderia cenocepacia*, that negatively regulates quorum sensing related virulence factors such as biofilm, type 6-secretion and protease secretion. *B. cenocepacia* is a Gram-negative opportunistic pathogen which causes severe, chronic respiratory infections in patients with cystic fibrosis and other immunocompromised conditions. This bacterium is also pathogenic in animal, plant, nematode, and insect infection models, and can survive within amoebae and macrophages. Presumably, the ability to survive in various niches requires adaptability to deal with changing environments. I hypothesize that AtsR is part of a multi-protein phosphorelay pathway which plays a critical role in regulation of niche adaptation and survival of *B. cenocepacia* in different hosts. In this thesis, I investigated AtsR function by characterizing the role of critical functional residues within the individual domains of AtsR and identified its cognate response regulator AtsT as a key component of the AtsR phosphorelay pathway. Furthermore, subsets of genes that are directly regulated by the AtsR cognate response regulator were identified by Chromatin immunoprecipitation followed by next generation sequencing (ChIP-Seq) analysis and its corresponding consensus DNA binding site was determined. I also investigated the role of AtsR as a global regulator of *B. cenocepacia* pathogenicity in the *Arabidopsis thaliana*

infection model. Together, these studies identified a new regulatory network that highlights the importance of bacterial virulence and pathogenicity with careful consideration of the host. This work may provide an understanding at the molecular level of bacterial adaptation to ever changing niche environments.

Keywords

Two-component systems, sensor kinase, response regulators, lipopolysaccharide, flagellin glycosylation, innate immunity, next generation sequencing, cystic fibrosis, *Burkholderia cenocepacia*, reactive oxygen species, nitric oxide, *Arabidopsis thaliana, Galleria mellonella*

Co-Authorship Statement

In chapter two of this thesis, Dr. Daniel Aubert provided the deletion mutant strain of AtsT which had not previously been published and made *K56-2atsT*$_{D208A}$ mutant strain. He also contributed to the first figure in that chapter by performing the macrophage infection and biofilm experiments. Maha Alzayer made pMZ24, pMZ25, pMZ33, pMZ34 and pMZ35 plasmids. All other work is that of M.K.

In chapter three of this thesis, Dr. Daniel Aubert performed the microarray experiment and analysis. Microarray hybridization was conducted by Dr. Eshwar Mahenthiralingam and Dr. Andrea Sass. All other work is that of M.K.

In chapter four of this thesis, Dr. Faviola Tavares sub-cloned AtsR mutants carrying point mutation from PET28a vector to PBAD vector. She also performed the PEGylation experiment and analyzed the results. All other work is that of M.K.

In chapter five of this thesis, Dr. Angel Andrade measured the growth rate of mutants and performed the *Galleria* work presented in figure 5.9. He also helped with microscopic pictures. Yasmine Fathy provided the MALDI-TOF experiment and data analysis presented in figure 5.4. All other work is that of M.K.

Dedication

This work is dedicated to my husband Hassan and my two children Ali and Reza who share this journey with me.

Acknowledgments

First and foremost I am so blessed that the greatest privilege was bestowed upon me by God for giving me the ability and opportunity to accomplish this work.

I would like to thank my supervisor, Dr. Miguel Valvano for his guidance and support throughout my PhD research for entrusting me with freedom and responsibility, and for his constant encouragement and belief in my capabilities. It is very hard to in just a few words sum up the years that I had the honor to be his PhD student. Not many people are fortunate to have a supervisor who is both a great scientist and a wonderful leader. His leadership taught me many valuable skills that will be with me for the rest of my life.

I would like to thank my advisory committee members Dr. Carole Creuzenet and Dr. David Litchfield. I am grateful for your suggestions and I benefited from our discussions during the committee meetings. I also had the opportunity to serve as a teaching assistant and would like to thank Drs. Heinrichs, Creuzenet and McGavin for the wonderful experience I gained as a TA in their Bacterial Pathogenesis course.

I am indebted to all my colleagues in the Valvano lab including Katie Bain, Roberto Rosales, Angel Andrade, Sarah Furlong, Mohamed Hammad, Crystal Schmerk, Omar El-Halfawy, Faviola Tavares, Yasmine Fathy, Xiang Ruan, Anna Hanuszkiewicz, Hanna Ostapska, Cristobal Mujica, Stephanie Lamothe and Janet Torres for creating a dynamic and friendly environment that I will greatly miss. Special thanks to Dr. Daniel Aubert who not only has helped me with his advice and discussions during the lab meetings but also has been a wonderful mentor to me for my development as a scientist. Also, thanks to Dr. Lauren Solomon from Dekoter's lab for training me on high throughput genome sequencing data analysis.

Last but not least, thanks to my family overseas for their support and encouragement. Special thanks to my husband Hassan and my kids Ali and Reza who have been supporting me in many ways and cheering me up to keep going. I could not have done this without them.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

1. 1 Overview of two-component regulatory systems

Bacteria like other living organisms, must be able to sense their environment and change their behavior in a timely manner in response to several different cues. They need to sense internal and external stimuli to adequately tune their cellular life. This requires changes in gene expression. One of the common mechanisms for bacterial adaptive responses to their surrounding environment is the two-component regulatory system (TCS). In 1986, Nixon and co-workers introduced the term "two-component regulatory systems" to describe their discovery that nitrogen assimilation proteins NtrB and NtrC of *Bradyrhizobium sp.* share strong similarity with other proteins such as EnvZ and PhoR in *Escherichia coli*, as well as VirA in *Agrobacterium tumefaciens* (1). They suggested that pairs of regulatory proteins with similar amino acid sequences respond to environmental stimuli and control diverse cellular processes. This finding was complemented with another discovery, also in 1986, that the TCSs utilize protein phosphorylation (2). A few years later, the Nobel Prize was awarded jointly to Edmond Fischer and Edwin Krebs for their work on protein phosphorylation as a biological regulatory mechanism. Since then hundreds of two-component signalling proteins have been identified in bacteria [(3) and http://www.ncbi.nlm.nih.gov/Complete_Genomes/RRcensus.html].

TCSs respond to an enormous range of signals (e.g. pH, temperature, nutrient availability, oxygen levels, etc) and control the regulation of a wide variety of cellular functions, virulence factors, chemotaxis, adaptation and survival. This provides a means for bacteria to detect environmental signals and process information and respond appropriately in a constantly changing environment. Therefore, understanding the mechanism of function of these systems is of great importance. Bioinformatic analysis clearly shows a correlation

between the number of two-component regulatory proteins and the genome size (4) and some investigators have taken it to the extreme to use it as a criteria to explain bacterial intelligence (5).

1.2 Signal transduction mechanism

Bacterial TCS convert extra or intracellular signals into modification of gene expression via reversible protein phosphotransfer from histidine to aspartate (3). A canonical signaling pathway consists of a sensor kinase protein and a response regulator protein. The sensor kinase protein detects a stimulus and the histidine kinase (HK) domain autophosphorylates using the ɣ-phosphate of ATP and undergoes a conformational change. This way, the information as a stimulus is converted into a high energy phosphoryl group that is then transferred to a conserved aspartate in the response regulator (RR) protein, which modulates gene expression accordingly (Fig. 1.1A) (6). There are two proteins in a canonical pathway, one with sensor and kinase function and the other as the receiver and effector function. In many cases there is an extended multistep phosphorelay and instead of only two, there are various phosphotransfer reactions (His-Asp-His-Asp). It typically initiates with a hybrid sensor kinase, which has a HK domain and a C-terminal receiver domain (Fig. 1.1B&C). The other essential module is a histidine phosphotransferase (Hpt) domain which may exist as a separate protein or as an addition to the hybrid sensor kinase. In such cases the phosphoryl group is transferred from histidine to aspartate within the hybrid sensor kinase and subsequently, to the histidine residue of the intermediate Hpt protein and finally to an aspartate of a terminal response regulator (Fig. 1.1C) (7). In the *Bacillus subtilis* Kin/Spo sporulation pathway, all three modules (transmitter, receiver, and Hpt) exist as individual

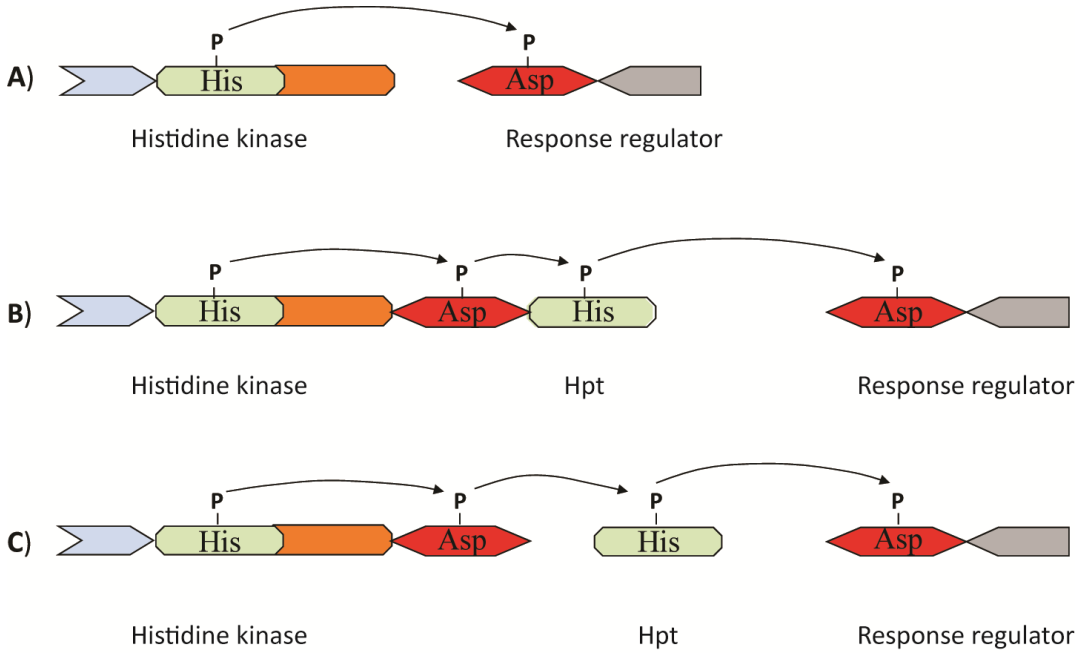


Figure 1.1: Schematic diagram representing two-component signal transduction. A) Schematic of a basic two-component model consisting of a sensor histidine kinase and a response regulator. A conserved histidine autophosphorylates and the phosphoryl group is transferred to an aspartate in the receiver domain of a response regulator. B) Illustration of a phosphorelay system with a hybrid HK containing both receiver domain and Hpt domain. The hybrid HK phosphorylates and transfers its phosphoryl group intramolecularly to its receiver domain and then to the Hpt domain. The phosphate is then transferred to the aspartate of the response regulator. C) A phosphorelay system with a hybrid sensor kinase and a separate intermediate Hpt protein. Blue box: sensory domain, light green box: HK domain, orange box: catalytic domain, red box: receiver domain, gray box: effector domain. His: histidine, Asp: aspartate, p: phosphoryl group, arrows indicate phosphoryl group transfer.

proteins (8), while the *Bordetella* spp. BvgAS regulatory system (9) and the *Burkholderia cenocepacia cblS* and *cblT* genes encode multidomain hybrid sensor kinases (10). Such multistep phosphorelay pathways provide multiple phosphorylation checkpoints which may imply a better fine-tuning mechanism than a typical TCS.

In most cases, sensor kinases can have both kinase or phosphatase activity; thus, it is ultimately the ratio of kinase to phosphatase activity that determines the level of RR phosphorylation (3). For those that do not have phosphatase activity, regulation occurs only at the level of autophosphorylation (11). Either of these activities can be regulated directly or indirectly by stimuli. Phosphotransfer itself can also be regulated. The C-terminal Asp-containing domain of the hybrid sensor kinase VirA in *A. tumefaciens* interacts with the autophosphorylation site and modulates the phosphotransfer ability of the kinase core (12).

1.3 Sensor kinases

A typical sensor kinase protein is an integral membrane protein with at least two hydrophobic transmembrane (TM) helices (13) and a periplasmic sensor domain that is surrounded by the two TM helices. It should be noted that not all sensor histidine kinases possess two transmembrane domains. For example the sensor kinases KdpD and UhpB of *E. coli* have 4 and 8 predicted transmembrane helices, respectively (14). The cytosolic part of the protein is composed of a phosphorylation site and a catalytic ATP-binding domain that together form the kinase core of the protein which is connected to the second TM. Most kinases have a linker domain between the second TM and the kinase domain. The most common linkers include HAMP, GAF and PAS (Fig. 1.2). The length of the

linker domain is variable ranging from 40 to more than 180 amino acids forming α-helical, coiled coil-like motifs (3). Sensor histidine kinases can be divided into two functionally and structurally distinct regions: the extracellular portion that senses stimuli, and the cytoplasmic domain involved in autophosphorylation and phosphotransfer to the cognate response regulator (Fig. 1.2).

HK sensing proteins typically act as homodimers with the kinase domain of one subunit catalysing the phosphorylation of a second subunit (15, 16). With these characteristic properties hundreds of HK proteins can be identified through genome sequences but some aspects such as the type of perceived signals and interactions with auxiliary proteins are still poorly understood.

1.3.1 Kinase domain

The kinase domains contain several signature motifs including the H box with conserved histidine, followed by the N, G1, F, and G2 boxes which make the binding cleft (Fig. 1.2). These motifs may exhibit different spacing across the kinase domain but they are arranged in the same order. The H box contains the site of histidine phosphorylation which is usually located approximately 110 amino acids from the N-box (3). Although the region surrounding the conserved histidine residue might have an influence on the ratio of autophosphorylation and phosphatase activities, substitution of this conserved histidine residue results in lack of enzymatic activities (17). Furthermore, the N box itself has a critical role for kinase activity and ATP binding, as the magnesium ion connects the phosphate groups via hydrogen binding and is associated with the asparagines within the N box (18). Gl and G2 boxes are glycine-rich regions that provide flexibility for the ATP

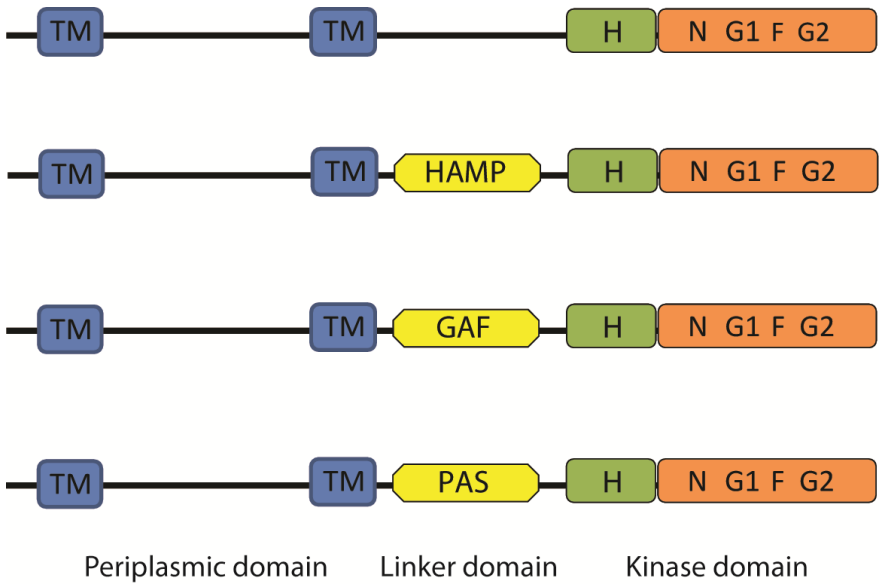


Figure 1.2: Schematic diagrams representing common domain organizations of sensor histidine kinases. Two TM domains are shown but kinases can harbor 1 to 13 TM domains. The cytosolic part consists of linker and kinase domains. Signature motifs H, N, G1, F and G2 are shown with common linker domains PAS, HAMP, and GAF. Abbreviation: HAMP: histidine kinase adenyl cyclase methyl-accepting protein and phosphatase; GAF: cGMP-specific phosphodiesterase adenylyl cyclase and FhlA; PAS: Per Arnt Sim; TM: transmembrane domain.

lid to secure bound ATP. Together these signature motifs form an ATP-binding cavity (19).

1.3.2 Periplasmic sensory domains

Unlike HK and receiver domains, periplasmic sensor domains do not share sequence homology as they may sense various signals. However, most of the known structures of periplasmic sensor domains include distinct structures such as mixed αβ, all-helical, and β-sandwiches. Some well-known examples of periplasmic domains with all-helices are represented by NarX with its sensor domain having an antiparallel four-helix bundle (20) and TorS with two antiparallel helix bundles (21). Examples of mixed αβ fold structure which is the most common form of periplasmic sensor domains exist in PhoQ, DcuS and CitA proteins (22). This form of structure often consists of central antiparallel β-sheets, surrounded by α-helices on both sides (22, 23). Some other HK proteins have β-sandwich folding sensor domains with two or more antiparallel β-sheets stacking back-to-back such as in RetS from *P. aeruginosa* (24).

1.3.3 Cytoplasmic and transmembrane sensory domains

Several HK proteins have been identified that lack a periplasmic sensing domain, such as the chemotaxis kinase CheA and nitrogen-regulating NtrB (3, 25). Although the periplasmic sensory domain is the most common type associated with HKs (22, 26), other classes of sensor domains include cytoplasmic and transmembrane domains. Cytoplasmic-sensing HKs exist both as membrane-embedded as well as fully soluble proteins (22, 26). They can be involved in detection of both chemical and physical stimuli.

Sensing domains that involve multiple membrane-spanning regions are thought to sense changes associated primarily with the membrane; however, some of them have also been implicated in forming protein-protein interactions or binding various natural ligands (22), as well as antibiotics (26, 27) via transmembrane helices.

1.4 Response regulators

Response regulators are cytoplasmic proteins that associate with a histidine kinase sensor protein. This protein family has a receiver domain which contains a conserved aspartic acid residue that accepts the phosphoryl group from the phosphorylated HK. Response regulators have also a DNA-binding effector domain that generates the outputs of the signaling events (28). Like the C-terminus of HKs, the conserved receiver domain of RRs exhibits amino acid sequence homology to other RRs. The transfer of a phosphate group to the receiver domain activates the transcriptional regulatory domain, which consequently results in activation or repression of a given set of genes (29). In contrast to the receiver domain, effector domains are very diverse reflecting the variable functions controlled by TCSs. The majority of RRs are transcription regulators with a winged helix-turn-helix (wHTH) DNA-binding motif, a helix-turn-helix (HTH) motif or β-strands. However, the effector domain in some RRs works as an enzyme or binds RNA or proteins to regulate bacterial cellular processes at different levels (30).

According to Stock and co-authors, phosphorylation of the regulatory domain often changes the effector domain from inactive conformational state to an active form although phosphorylation does not necessarily correspond to activation (3). Activation may result in a release of inhibition (31, 32, 33, 34), can promote dimerization (35),

higher-order oligomerization (36, 37), or interactions with other proteins (38, 39). The half life of the phosphorylated RR may vary from only a few seconds to several hours, depending on the type of response output. For example, processes occurring in a narrow time frame such as chemotaxis require rapid dephosphorylation of RR to fine tune the response. The CheY has a very short half life to allow *E. coli* to respond quickly to changes in chemotactic gradients (38). In contrast, signal transductions that lead to major changes in gene expression have a very long half life (40).

1.5 Specificity in two-component signaling pathways

It is straightforward to identify histidine kinases and response regulators in bacterial genomes based on their sequence homology. Some bacterial genomes encode as many as 250 of these signaling proteins (41). The ability of discriminating a cognate RR from the pool of many others by a phosphorylated HK is extremely important to avoid mistranslation of a signal into an inappropriate outcome. It is believed that interfering crosstalk is rare due to mechanisms such as molecular recognition and substrate competition which play a key role in insulating signalling pathways. In a study by Skerker et al in 2005, a systematic phosphotransfer profiling was performed in *Caulobacter crescentus* genome demonstrating that HKs have a global kinetic preference for their cognate response regulator *in vitro* (42). They proposed that kinetic preference of histidine kinases for their cognate response regulators is fundamental to insulation of signaling pathways and lies in the existence of conserved molecular characteristics that determine protein–protein interaction specificity. By mutating the specificity residues of the histidine kinase EnvZ and matching them to those of RstB, the mutated EnvZ lost its

interaction with its cognate partner, OmpR but phosphorylated RstA, the cognate regulator of RstB (43, 44).

The relative cellular concentration of HK and RR also determines the specificity (45). The higher abundance of the RR with respect to HKs observed *in vivo* provides a competitive environment in which a cognate regulator outcompetes non-cognate regulators. This may be an evolutionary strategy to minimise unwanted interactions.

These mechanisms maintain specificity of TCS pathways at the level of phosphotransfer. However, there are reports of crosstalk between TCSs that normally activate separate pathways (46). It should be noted that phosphotransfer between non-cognate sensors and regulators has been found only *in vitro*, or *in vivo* in the presence of constitutively expressed non-cognate proteins or in the absence of the cognate histidine kinase (47, 48). Therefore, the biological relevance of such crosstalk is unclear because the efficiency of the interaction should be studied under physiological conditions when proteins are present at normal levels.

Spatial or temporal separation of different systems also maintains interaction specificity as some TCS proteins are expressed at certain developmental stages (45). Scaffold proteins can also be involved in forming large protein complexes with hybrid HKs that may reduce the chances for crosstalk (49).

1.6 Two-component systems and signal integration

Signal integration is a combined information processing event that occurs between TCSs in branched phosphorylation mode. There are two generic modes of signal integration in which multiple HKs (multiple independent signals) are able to phosphorylate a single RR

protein (many-to-one), or one single HK phosphorylates multiple RRs (one-to-many) triggering multiple outputs (Figure1.3 A&B). This is an effective way to integrate signals and control multiple outputs at once.

A physiologically meaningful example of "one-to-many" branched regulation is found in bacterial chemotaxis, where the CheA phosphorylates two response regulators CheY and CheB (50). A more complex version of "many-to-one" mode is seen in the signal transduction system controlling sporulation in *B. subtilis* where five distinct sensor kinase proteins, KinA, KinB, KinC, KinD, and KinE phosphorylate the receiver domain of Spo0F, and then the Hpt-containing Spo0B protein. Finally, the signal is transferred from Spo0B to the Spo0A response regulator (7). Each of the five sensor kinases is capable of phosphorylating Spo0F with various efficiencies (51). Another example is the activation of NarL by NarQ and NarX in *E.coli* (52). In *B. subtilis*, signal integration also occurs between the pathways controlling respiration (ResE/ResD) and phosphate utilization (PhoR/PhoP) in which phosphorylation of PhoP is required for expression of ResD and vice versa (53). This is an example of communication between distinct signalling pathways that work in parallel and provide physiological benefits to bacterial cells by coordination and fine-tuning of cellular processes.

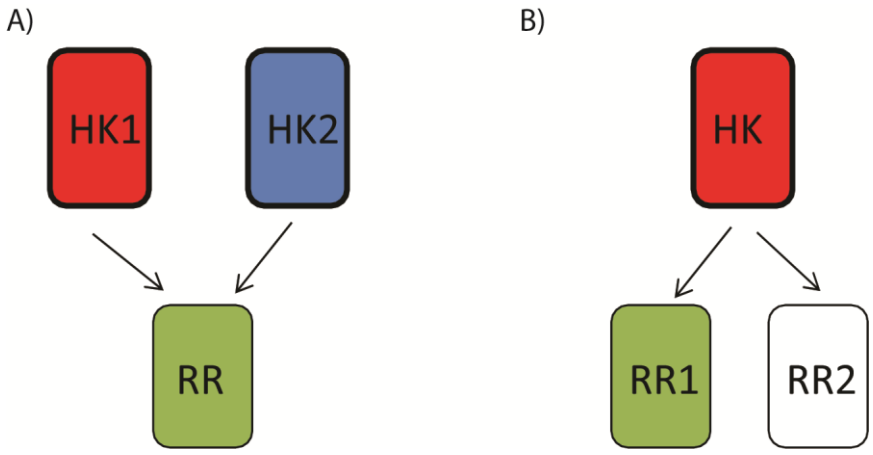


Figure 1.3: Schematic representation of signal integration by two-component regulatory systems. A) One-to-many branched signalling pathway in which one histidine kinase phosphorylates multiple response regulators. B) Many-to-one branched pathway in which multiple histidine kinases perceive similar or different signals and transfer it to one response regulator. HK: histidine kinase, RR: response regulator.

1.7 Genomic organisation of two-component systems

Genes encoding proteins of similar function are usually clustered together in bacterial genomes; this way they are more likely transcribed together and manifest a particular phenotype (54). Such co-localisation in an operon greatly facilitates identification of interacting pairs of TCSs. However, in many cases the signalling partner is not in the vicinity of a TCS protein. It might be due to gene duplication or lateral gene transfer within and between species (55). It has been proposed that when bacteria occupy a wide variety of environmental niches or utilise different metabolic sources then gene duplication occurs with domain shuffling and thus results in duplicated genes with sequence differentiation from the parental gene (55). When such event happens to one of the two genes encoding TCS proteins it would introduce an orphan protein with new domain arrangements and possibly new sensory or regulatory functions which may become involved in cross-regulation. This is a possible explanation for the presence of several orphan Kin histidine kinases in *B. subtilis* which can all phosphorylate the same RR, Spo0F to integrate multiple inputs for fine tuning of the adaptive responses (56).

1.8 Autoregulation

Many TCSs regulate their own expression. In fact, regulation of HK gene expression by its cognate RR and also endogenous regulation of RR itself in the form of positive or negative feedback is not uncommon in bacteria. This autoregulation depends on the amount of input signal such as the one observed for the PhoQ/PhoP system in *E. coli* that under sufficiently high stimulus conditions PhoB levels are transcriptionally auto-tuned to minimize wasteful production of PhoB (57).

Similarly, depending on the intensity of the signal, autoregulation positively modulates the sensitivity of the BvgS/BvgA system in *Bordetella bronchiseptica* (58). Another important factor is the RR functioning as a switch depending on the proportion of phosphorylation (3). As mentioned above, some TCSs have rapid conversion to the inactive state to enable highly sensitive responses. In addition, all TCSs seem to have intrinsic mechanisms regulating RR phosphorylation levels which ensure proper feedback mechanism, thus eliminating the effects of potential crosstalk (59, 60).

Negative autoregulation, in which the response regulator represses its own expression, has been reported for the CovS/CovR system in *Streptococcus pyogenes*, where phosphorylated CovR represses transcription of its operon (61). There are also response regulators such as TorR (62) and LuxO (63) that repress their own expression regardless of phosphorylation status. This might be part of a complex regulatory network that accessory proteins are involved and interact with either histidine kinase or response regulator and thus, autoregulation may affect the entire system by controlling the availability of the interacting proteins (64). In other cases, autoregulation provides a threshold for gene activation and only when a signal persists sufficient levels of HK or RR proteins are produced (65).

1.9 Two-component systems and bacterial virulence

To establish a successful infection, bacteria should coordinate the expression of genes required for niche adaptation and virulence in response to signals from the host environment. Therefore, TCSs play a sophisticated role processing various signals to control expression of virulence factors at different stages during infection. This role was

extensively examined in *Streptococcus pneumoniae* when systemic inactivation of 13 TCSs led to a significant reduction in virulence in an *in vivo* mice model (66). Another well-studied example, AgrA/AgrC TCS in *S. aureus* responds to cell density and controls the transcription of the regulatory RNA III (67). Other examples include the BvgA-BvgS TCS in *Bordetella pertussis* (68), the RscC/B two-component system in enteric *E. coli* that regulates capsule formation (69) and many others that are listed in Table 1.1. The role of TCSs in virulence is not limited to bacteria. Several fungal pathogens such as *Aspergillus fumigatus* (70), *Cryptococcus neoformans* (13), *Blastomyces dermatitidis*, *Histoplasma capsulatum* (71) and *Candida albicans* (72) use phosphorelay signalling to regulate virulence.

Besides the central role of TCS in virulence, some are essential for bacterial growth. In particular, the WalK/WalR TCS responsible for cell wall metabolism is indispensable for growth in several Gram-positive species, including *B. subtilis*, *Staphylococcus aureus*, *Enterococcus faecalis*, *Listeria monocytogenes*, *Staphylococcus epidermidis*, *S. pneumoniae*, *Streptococcus mutans*, and *S. pyogenes* (73). Also, *Helicobacter pylori* has response regulators that were found to be essential for cell growth (74, 75). Essential TCSs are listed in Table 1.2.

Table 1.1: List of selected TCS involved in virulence. Homologues of the described systems are present in other bacteria. The list represented in Table 1.1 is not exclusive.

Organism	TCS	Function	Ref.
S. aureus	AgrC/AgrA	Adhesion, secreted invasive factor	(76)
	ArlS/ArlR	Expression of α-toxin, β-haemolysin, Lipase, Coagulase and protein A	(77)
	SrrB/SrrA	Regulator of virulence factors in response to environmental oxygen levels	(78)
	SaeS/SaeR	Expression of α-toxin, β-haemolysin, Coagulase	(79)
P. aeruginosa	GacS/GacA	Lipase, Elastase, Biofilm	(80, 81)
	AlgR/FlmS	Twitching motility	(82)
	SadR/SadS	Biofilm	(83)
	RetS	T3SS and its effector genes	(83, 84)
	PprB/PrpA	Cell motility, quorum sensing	(85)
	PirS/PirR	Iron aquisition	(86)
	PilS/PilR	Pilus production	(87)
Brucella abortus	BvrR/BvrS	Virulence attenuation, reduced invasiveness in macrophages	(88)
Neisseria meningitidis	MisR/MisS	Composition of LOS inner core	(89)
Clostridium perfringens	VirR/VirS	Toxin and adhesion genes	(90)
Salmonella enterica	PhoQ/PhoP	Cationic antimicrobial peptide resistance, LPS modification	(91)
	OmpR/EnvZ	Osmolarity	(92)
	PmrA/PmrB	Lipid A modification	(93)
	RcsC-YojN-RcsB	Colonic acid capsule synthesis, chemotaxis, motility	(94)
	SsrB/SsrA	T3SS and effector genes	(95)
Shigella flexneri	OmpR/EnvZ	Invasion genes	(96)
E. faecalis	FsrC/FsrA	Protease activity, Quorum sensing	(97)
	VanS/VanR	Vancomycin resistance	(98)
Enterococcus gallinarum	VanS/VanR	Vancomycin resistance	(99)
Vibrio cholerae	ToxS/ToxR	Expression of toxins	(100)
	ArcA/ArcB	Virulence regulator gene *toxT*	(101)
S. pneumoniae	CiaH/CiaR	Competence, resistance to cefotaxim	(102)
	PnpS/PnpR	Competence, phosphate regulation	(66)
	VncS/VncR	Vancomycin-induced cell death	(103)

	MicA/MicB	Virulence attenuation	(104)
S. pyogenes	CovS/CovR	Expression of capsule, pyrogenic exotoxin B, streptolysin, Streptokinase, mitogenic factor	(105)
	FacB/ FacC/ FacA	Expression of streptolysin S, fibronectin/fibrinogen-binding proteins, Streptokinase	(106)
	CsrS/CsrR	Negative regulation of hyaluronic acid and toxins	(107)
Mycobacterium tuberculosis	PhoP/PhoR	Intracellular growth in macrophages	(108)
	PrrA/PrrB	Survival in phagosome environment	(109)
	MprA/MprB	Virulence attenuation	(110)
Listeria monocytogenes	VirR/ VirS	Virulence attenuation	(111)
B. pertussis	BvgS/BvgA	Toxin and adhesion expression, biofilm formation	(112)
H. pylori	FlgR/FlgS	Flagellar motility	(113)
	ArsR-ArsS	Urease and acid resistance genes	(114)
Campylobacter jejuni	DccR-DccS	Colonization defect	(115)
S. typhimurium	PhoQ/PhoP	Polymyxin B	(116)
S. typhimurium LT2	PmrB/PmrA	Polymyxin B	(117)
Legionella pneumophila	CpxR/CpxA	*icmR* and *icm-dot* genes	(118)
Bacteroides fragilis	RprX/RprY	Tetracycline resistance	(119)
Yersinia pseudotuberculosis	PhoP	Virulence attenuation, reduced survival in macrophages	(120)
EHEC	QseC/QseB	Quorum sensing regulation	(121)
	VanS/VanR	Vancomycin resistance	(122)

Table 1.2: List of essential TCSs adapted from Gotoh et al 2010 (123).

TCS	Organism	Function	Ref
WalK/WalR	*S. mutans*	Biofilm formation	(124)
	S. aureus	Cell wall metabolism	(125)
	S. pyogenes	Cell wall metabolism	(126)
	S. pneumoniae	Cell wall metabolism	(127)
	S. epidermis	Cell wall metabolism	(128)
	E. faecalis	Unknown	(129)
	L. monocytogenes	Unknown	(130)
YhcS/YhcR	*S. aureus*	Unknown	(131)
HP165/HP166	*H. pylori*	Unknown	(75)
MtrB/MtrA	*M. tuberculosis*	Unknown	(132)

1.10 Two-component systems and antimicrobial resistance

In addition to being involved directly in virulence, a number of bacterial TCSs have been identified to play a central role in antimicrobial resistance. Considering the number of genes regulated by TCSs, it is not surprising that some are involved in modulating several aspects of adaptive functions related to antibiotic sensitivity. Some of those are listed in Table 1.1 that regulate the resistance to antimicrobial agents. For example, the EvgA response regulator modulates multidrug transporter genes expression and subsequently controls drug efflux in *E. coli* (133). The VanS/R in enterococci provides resistance to vancomycin by controlling the expression of the genes for the synthesis of cell wall precursors that have much lower affinity for vancomycin (98, 134). Same TCS in *S. pneumoniae* triggers multiple cell death pathways in response to penicillin and similar antibiotics, like cephalosporins and glycopeptides (135). Similarly, the LytS/R two-component system contributes to penicillin-induced killing of *S. aureus* (136). These features make TCSs ideal targets for therapy as it permits a drug to target multiple systems in the same bacterium.

1.11 Two-component systems and quorum sensing

To establish a successful infection, bacteria need to reach a critical cell density at the site of infection. Bacterial cell to cell communication via chemical signaling molecules and subsequent coordination of the expression of virulence factors in response to bacterial population density is known as quorum sensing (QS) (137). This allows a population of bacteria to coordinately control the gene expression of the entire community and behave as multicellular organisms. A well-known example is the AgrA/AgrC system in *S.*

aureus. This system is activated in response to a secreted auto-inducing peptide (AIP) in a cell density-dependent manner. At a certain concentration of AIP, the sensor kinase AgrC is autophosphorylated and transfers the phosphate to the AgrA response regulator. Activated AgrA activates the expression of genes encoding secreted virulence factors (76). Vancomycin-induced cell death is also modulated by the VncS/VncR two-component system in *S. pneumonia* in a cell density-dependent manner (135). Another example is the *fsrC* and *fsrA* genes in *E. faecalis*, which encode a sensor kinase and response regulator pair that respond to a biosynthesis-activating pheromone in a cell density manner. The Lux system in *Vibrio harveyi* is a well-studied hybrid quorum sensing circuit. When acyl-homoserine lactone (AHL) inducers are accumulated as a function of cell density they are detected by LuxN, LuxP and LuxQ. LuxN and LuxQ are hybrid sensor kinase and response regulator pairs that transduce information to LuxU and then to the response regulator LuxO that controls the transcription of LuxR which is also required for expression of the luciferase structural operon (*luxCDABE*) (138). These are a few examples that indicate the important role of TCSs in bacterial communication within and between species that allows bacteria to compete with other species inhabiting unique niches.

1.12 Two-component systems as therapeutic targets

Several features of TCSs make them excellent candidates for the development of antimicrobial agents. First of all, the nature of histidine to aspartate phosphotransfer mechanism utilized by bacteria and fungi is distinct from the serine/threonine signaling pathways in eukaryotes. This feature alone makes TCS a good antimicrobial target as for an antimicrobial agent to be effective it should be selective for the bacterium and not the

host. Second, sensor kinases and response regulators have sequence similarity and similar mode of action irrespective of bacterial species (139). This may be used to design broad spectrum inhibitors that target common structural or functional features. Third, TCSs modulate a wide variety of cellular functions and thus targeting TCSs is equal to addressing multiple targets. This feature alone has additional benefits such as minimizing the emergence of resistance in response to antibiotics simply because multiple TCS with different functions exist, and alteration mechanisms utilised by bacteria would no longer be effective. This is unlike available antibiotics that have a very specific mechanism of action against one target only or limited number of targets.

1.13 *Burkholderia cenocepacia*

Burkholderia cenocepacia is an opportunistic pathogen that belongs to the *Burkholderia cepacia* complex (Bcc), a group of Gram-negative non-spore forming bacilli that consists of at least 17 phenotypically similar but genetically distinct species (140). Due to their diverse genetics and metabolic capacity the Bcc strains can be found throughout the environment (141). Also, because of their opportunistic nature they are capable of causing infections in vulnerable individuals including immunocompromised and cystic fibrosis (CF) patients.

Historically, *B. cenocepacia* ET-12 was the dominant Bcc pathogen in the CF community, responsible for the largest CF epidemic across Canada, United Kingdom and Italy prior to 2002 with 67% prevalence followed by *B. multivorans* as the next most dominant Bcc species (17%) (142). However, its prevalence has been reduced and *B. multivorans* is now the most dominant Bcc infection encountered among CF population in

North America and UK (143). *B. cenocepacia* is also associated with poor prognosis and increased risk of developing cepacia syndrome (144). Because of these complications lung transplantation for CF patients infected with *B. cenocepacia* has a poor outcome and patients have a high morbidity and mortality rate (145).

1.13.1 *B. cenocepacia* in the environment

B. cenocepacia has also been isolated from different environmental sources, ranging from urban (146) and agricultural (147) soils, to Mexican radish (148), and raw milk (149). *B. cenocepacia* engages in both beneficial and pathogenic interactions with plants. For example *B. cenocepacia* has been identified to establish an endophytic relationship in the rhizosphere of onion (150) and maize (151). *B. cenocepacia* has also been isolated from the stem and roots of sugar cane promoting the production of the growth hormone IAA and inhibiting hyphal growth of *Fusarium moniliforme* (152).

B. cenocepacia is beneficial to a few plants but there are several examples indicating its role as a phytopathogen. *B. cenocepacia* has been isolated directly from the lesions of field-grown onion bulbs and is the causal agent of banana finger-tip rot (153). Strain K56-2 could cause the plant tissue watersoaking (PTW) phenotype which is an accumulation of fluids in the intracellular spaces of plant tissues caused by a loss of cell membrane integrity (154).

Further evidence that supports the direct environmental acquisition of *B. cenocepacia* in patients was found in a soil isolate that was genetically identical to strains isolated from patients during the 1980s *B. cenocepacia* CF epidemic (147). In a comprehensive genetic study by Baldwin et al. 2007, multiple clinical isolates of *B. cenocepacia* and *B.*

multivorans were not distinguishable from environmental isolates recovered from diverse environments such as river water, onion, radish, maize rhizosphere, pharmaceutical solutions, hospital equipment, shampoo and industrial settings (148). All together, these features make *B. cenocepacia* an ideal model for studying niche adaptation and opportunistic pathogenicity.

1.13.2 Virulence factors in *B. cenocepacia*

The genome of *B. cenocepacia* J2315 is divided into three circular chromosomes and a plasmid encoding 7261 putative coding sequences (CDS) in 8 Mbp of genetic material (155). The genes are designated with a prefix (*Bcal*, *Bcam*, *Bcas*, or *pBca*) and a number. The genome possesses a number of virulence factors including quorum sensing (QS) and secretion systems that are putatively involved in pathogenicity of *B. cenocepacia* [reviewed in (156)].

QS is a form of bacterial communication that detects a signaling molecule relative to cell density to regulate gene expression (157). *B. cenocepacia* has AHL-dependent and independent QS modules: The CepIR and CciIR systems modulate gene expression via two AHL synthases, CepI and CciI, which produce *N*-octanoylhomoserine lactone (C8-AHL) and *N*-hexanoylhomoserine lactone (C6-AHL), respectively, and activate their cognate transcriptional regulator CepR or CciR (158, 159). Another QS module in *B. cenocepacia* that produces *cis*-2-dodecenoic acid (BDSF) in a cell density-dependent manner is AHL-independent (160). These QS regulatory systems together allow *B. cenocepacia* to fine-tune the expression of its virulence factors including biofilm maturation (161), swarming motility (162), and type 6 secretion system (T6SS)

expression (163). Another critical feature utilized by *B. cenocepacia* is the specialized secretion systems to translocate effector molecules into the host cell. The pathogenic potential of types III, IV, and VI secretion systems in *B. cenocepacia* have been examined in several infection models (164, 165), (155, 166) and (167). In particular, the T6SS is the most recently discovered and has been shown to mediate actin rearrangement in macrophages infected by *B. cenocepacia* (168). Mutation in the T6SS of *B. cenocepacia* results in susceptibility to predation by the amoeba *Dictyostelium discoideum* (168). Furthermore, it has been established that the ability of *B. cenocepacia* to cause chronic infection is associated with several virulence factors including the expression of T6SS, which is important for bacterial survival in the *in vivo* rat model of chronic lung infection (167). Our laboratory identified AtsR (Adherence and T6SS Regulator) protein as a global regulator of the expression of several virulence factors including biofilm and the T6SS in *B. cenocepacia* (168). In addition to affecting biofilm and T6SS activity, the deletion of *atsR* in *B. cenocepacia* leads to overproduction of other QS-regulated virulence determinants including proteases and swarming motility (169). Deletion of *atsR* upregulated the expression of the QS genes and increased AHL production, suggesting that AtsR plays a significant role in timing and fine-tuning of virulence gene expression by modulating QS signalling (169). Since the deletion of AtsR affects multiple virulence factors, it is likely that it initiates a signaling cascade with broad activity.

1.14 Hypothesis and project objectives

The work described in this thesis uses *B. cenocepacia* K56-2, isolated from the lungs of a Canadian CF patient. K56-2 belongs to the ET-12 lineage, and is assigned to the same clonal complex as J2315 (170). AtsR is encoded by the gene *Bcam0379* in *B. cenocepacia*

K56-2 strain and is predicted to be a putative hybrid sensor kinase. Therefore, I hypothesize that **AtsR is part of a multi-protein phosphorelay pathway which plays a critical role in regulation of virulence factors in *B. cenocepacia* K56-2 via a cognate cytosolic transcriptional regulator.**

The specific objectives of this project are:

1) To characterize the AtsR phosphorelay pathway and identify its corresponding response regulator.
2) To identify the regulon of AtsR/AtsT signaling pathway by whole genome mapping via microarray and chromatin immunoprecipitation followed by Next Generation Sequencing.
3) To determine the topology of AtsR protein and find the functional residues in the periplasmic loop.
4) To determine the pathogenicity of *B. cenocepacia* in a plant infection model and study the subsequent events of plant-*B. cenocepacia* interaction.

In Chapter Two, I characterized the mechanism of AtsR signal transduction pathway by studying each specific domain and identifying conserved residues required for autophosphorylation within the kinase domain. Furthermore, I identified AtsT as the cognate response regulator that contribute to phosphotransfer and determined its position in the regulatory cascade. The biological significance of AtsR/AtsT phosphorylation was also studied *in vivo.*

In Chapter Three, I studied the AtsR/AtsT regulon using a combination of transcriptome and ChIP-Seq analysis, confirmed AtsT binding to the promoter regions of the target

genes and determined a conserved binding motif for AtsT-DNA binding.

In Chapter Four, the AtsR topology was experimentally determined and functional residues within the periplasmic sensing domain were identified.

In Chapter Five, I took a different approach to study the concept of virulence and pathogenicity of *B. cenocepacia* as an opportunistic pathogen in the context of host and its innate susceptibility to infection. I used *Arabidopsis* as an infection model and hypothesized that modifications in key bacterial molecules such as LPS and flagellin recognized by the innate immune system modulate host responses to *B. cenocepacia.*

1.15 Chapter 1 references

1. **Nixon BT, Ronson CW, Ausubel FM.** 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. Proc Natl Acad Sci U S A **83:**7850-7854.

2. **Ninfa AJ, Magasanik B.** 1986. Covalent modification of the *glnG* product, NRI, by the *glnL* product, NRII, regulates the transcription of the *glnALG* operon in *Escherichia coli*. Proc Natl Acad Sci U S A **83:**5909-5913.

3. **Stock AM, Robinson VL, Goudreau PN.** 2000. Two-component signal transduction. Annu Rev Biochem **69:**183-215.

4. **Ulrich LE, Koonin EV, Zhulin IB.** 2005. One-component systems dominate signal transduction in prokaryotes. Trends Microbiol **13:**52-56.

5. **Marijuan PC, Navarro J, del Moral R.** 2010. On prokaryotic intelligence: strategies for sensing the environment. Biosystems **99:**94-103.

6. **Capra EJ, Laub MT.** 2012. Evolution of two-component signal transduction systems. Annu Rev Microbiol **66:**325-347.

7. **Burbulys D, Trach KA, Hoch JA.** 1991. Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. Cell **64:**545-552.

8. **Fabret C, Feher VA, Hoch JA.** 1999. Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. J Bacteriol **181:**1975-1983.

9. **Cotter PA, Jones AM.** 2003. Phosphorelay control of virulence gene expression in *Bordetella*. Trends Microbiol **11:**367-373.

10. **Tomich M, Mohr CD.** 2004. Genetic characterization of a multicomponent signal transduction system controlling the expression of cable pili in *Burkholderia cenocepacia*. J Bacteriol **186:**3826-3836.

11. **Borkovich KA, Kaplan N, Hess JF, Simon MI.** 1989. Transmembrane signal transduction in bacterial chemotaxis involves ligand-dependent activation of phosphate group transfer. Proc Natl Acad Sci U S A **86:**1208-1212.

12. **Chang CH, Zhu J, Winans SC.** 1996. Pleiotropic phenotypes caused by genetic ablation of the receiver module of the *Agrobacterium tumefaciens* VirA protein. J Bacteriol **178:**4710-4716.

13. **Bahn YS, Kojima K, Cox GM, Heitman J.** 2006. A unique fungal two-component system regulates stress responses, drug sensitivity, sexual development, and virulence of *Cryptococcus neoformans*. Mol Biol Cell **17:**3122-3135.

14. **Zimmann P, Puppe W, Altendorf K.** 1995. Membrane topology analysis of the sensor kinase KdpD of *Escherichia coli*. J Biol Chem **270:**28282-28288.

15. **Surette MG, Levit M, Liu Y, Lukat G, Ninfa EG, Ninfa A, Stock JB.** 1996. Dimerization is required for the activity of the protein histidine kinase CheA that mediates signal transduction in bacterial chemotaxis. J Biol Chem **271:**939-945.

16. **Pan SQ, Charles T, Jin S, Wu ZL, Nester EW.** 1993. Preformed dimeric state of the sensor protein VirA is involved in plant-*Agrobacterium* signal transduction. Proc Natl Acad Sci U S A **90:**9939-9943.

17. **Kanamaru K, Aiba H, Mizuno T.** 1990. Transmembrane signal transduction and osmoregulation in *Escherichia coli*: I. Analysis by site-directed mutagenesis of the amino acid residues involved in phosphotransfer between the two regulatory components, EnvZ and OmpR. J Biochem **108:**483-487.

18. **Hsing W, Russo FD, Bernd KK, Silhavy TJ.** 1998. Mutations that alter the kinase and phosphatase activities of the two-component sensor EnvZ. J Bacteriol **180:**4538-4546.

19. **Foussard M, Cabantous S, Pedelacq J, Guillet V, Tranier S, Mourey L, Birck C, Samama J.** 2001. The molecular puzzle of two-component signaling cascades. Microbes Infect **3:**417-424.

20. **Cheung J, Hendrickson WA.** 2009. Structural analysis of ligand stimulation of the histidine kinase NarX. Structure **17:**190-201.

21. **Moore JO, Hendrickson WA.** 2009. Structural analysis of sensor domains from the TMAO-responsive histidine kinase receptor TorS. Structure **17:**1195-1204.

22. **Cheung J, Hendrickson WA.** 2010. Sensor domains of two-component regulatory systems. Curr Opin Microbiol **13:**116-123.

23. **Sevvana M, Vijayan V, Zweckstetter M, Reinelt S, Madden DR, Herbst-Irmer R, Sheldrick GM, Bott M, Griesinger C, Becker S.** 2008. A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA. J Mol Biol **377:**512-523.

24. **Jing X, Jaw J, Robinson HH, Schubot FD.** 2010. Crystal structure and oligomeric state of the RetS signaling kinase sensory domain. Proteins **78:**1631-1640.

25. **Grebe TW, Stock JB.** 1999. The histidine protein kinase superfamily. Adv Microb Physiol **41:**139-227.

26. **Mascher T, Helmann JD, Unden G.** 2006. Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol Mol Biol Rev **70:**910-938.

27. **Krell T, Lacal J, Busch A, Silva-Jimenez H, Guazzaroni ME, Ramos JL.** 2010. Bacterial sensor kinases: diversity in the recognition of environmental signals. Annu Rev Microbiol **64:**539-559.

28. **Goudreau PN, Stock AM.** 1998. Signal transduction in bacteria: molecular mechanisms of stimulus-response coupling. Curr Opin Microbiol **1:**160-169.

29. **Hakenbeck R, Stock JB.** 1996. Analysis of two-component signal transduction systems involved in transcriptional regulation. Methods Enzymol **273:**281-300.

30. **Galperin MY.** 2010. Diversity of structure and function of response regulator output domains. Curr Opin Microbiol **13:**150-159.

31. **Simms SA, Keane MG, Stock J.** 1985. Multiple forms of the CheB methylesterase in bacterial chemosensing. J Biol Chem **260:**10161-10168.

32. **Kahn D, Ditta G.** 1991. Modular structure of FixJ: homology of the transcriptional activator domain with the -35 binding domain of sigma factors. Mol Microbiol **5:**987-997.

33. **Huala E, Stigter J, Ausubel FM.** 1992. The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription. J Bacteriol **174:**1428-1431.

34. **Grimsley JK, Tjalkens RB, Strauch MA, Bird TH, Spiegelman GB, Hostomsky Z, Whiteley JM, Hoch JA.** 1994. Subunit composition and domain structure of the Spo0A sporulation transcription factor of *Bacillus subtilis*. J Biol Chem **269:**16977-16982.

35. **Fiedler U, Weiss V.** 1995. A common switch in activation of the response regulators NtrC and PhoB: phosphorylation induces dimerization of the receiver modules. EMBO J **14:**3696-3705.

36. **Webber CA, Kadner RJ.** 1997. Involvement of the amino-terminal phosphorylation module of UhpA in activation of uhpT transcription in *Escherichia coli*. Mol Microbiol **24:**1039-1048.

37. **Wyman C, Rombel I, North AK, Bustamante C, Kustu S.** 1997. Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. Science **275:**1658-1661.

38. **Blat Y, Eisenbach M.** 1994. Phosphorylation-dependent binding of the chemotaxis signal molecule CheY to its phosphatase, CheZ. Biochemistry **33:**902-906.

39. **Welch M, Oosawa K, Aizawa S, Eisenbach M.** 1993. Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria. Proc Natl Acad Sci U S A **90:**8787-8791.

40. **Yamamoto K, Hirao K, Oshima T, Aiba H, Utsumi R, Ishihama A.** 2005. Functional characterization *in vitro* of all two-component signal transduction systems from *Escherichia coli*. J Biol Chem **280:**1448-1456.

41. **Davidsen T, Beck E, Ganapathy A, Montgomery R, Zafar N, Yang Q, Madupu R, Goetz P, Galinsky K, White O, Sutton G.** 2010. The comprehensive microbial resource. Nucleic Acids Res **38:**D340-345.

42. **Skerker JM, Prasol MS, Perchuk BS, Biondi EG, Laub MT.** 2005. Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. PLoS Biol **3:**e334.

43. **Whitworth DE, Cock PJ.** 2009. Evolution of prokaryotic two-component systems: insights from comparative genomics. Amino Acids **37:**459-466.

44. **Noriega CE, Lin HY, Chen LL, Williams SB, Stewart V.** 2010. Asymmetric cross-regulation between the nitrate-responsive NarX-NarL and NarQ-NarP two-component regulatory systems from *Escherichia coli* K-12. Mol Microbiol **75:**394-412.

45. **Laub MT, Goulian M.** 2007. Specificity in two-component signal transduction pathways. Annu Rev Genet **41:**121-145.

46. **Wanner BL.** 1992. Is cross regulation by phosphorylation of two-component response regulator proteins important in bacteria? J Bacteriol **174:**2053-2058.

47. **Fisher SL, Kim SK, Wanner BL, Walsh CT.** 1996. Kinetic comparison of the specificity of the vancomycin resistance VanSfor two response regulators, VanR and PhoB. Biochemistry **35:**4732-4740.

48. **McCleary WR.** 1996. The activation of PhoB by acetylphosphate. Mol Microbiol **20:**1155-1163.

49. **Wegener-Feldbrugge S, Sogaard-Andersen L.** 2009. The atypical hybrid histidine protein kinase RodK in *Myxococcus xanthus*: spatial proximity supersedes kinetic preference in phosphotransfer reactions. J Bacteriol **191:**1765-1776.

50. **Kirby JR.** 2009. Chemotaxis-like regulatory systems: unique roles in diverse bacteria. Annu Rev Microbiol **63:**45-59.

51. **Jiang M, Shao W, Perego M, Hoch JA.** 2000. Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. Mol Microbiol **38:**535-542.

52. **Rabin RS, Stewart V.** 1992. Either of two functionally redundant sensor proteins, NarX and NarQ, is sufficient for nitrate regulation in *Escherichia coli* K-12. Proc Natl Acad Sci U S A **89:**8419-8423.

53. **Sun G, Birkey SM, Hulett FM.** 1996. Three two-component signal-transduction systems interact for Pho regulation in *Bacillus subtilis*. Mol Microbiol **19:**941-948.

54. **Lawrence J.** 1999. Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. Curr Opin Genet Dev **9:**642-648.

55. **Alm E, Huang K, Arkin A.** 2006. The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. PLoS Comput Biol **2:**e143.

56. **Raghavan V, Groisman EA.** 2010. Orphan and hybrid two-component system proteins in health and disease. Current Opinion in Microbiology **13:**226-231.

57. **Miyashiro T, Goulian M.** 2008. High stimulus unmasks positive feedback in an autoregulated bacterial signaling circuit. Proc Natl Acad Sci U S A **105:**17457-17462.

58. **Williams CL, Cotter PA.** 2007. Autoregulation is essential for precise temporal and steady-state regulation by the *Bordetella* BvgAS phosphorelay. J Bacteriol **189:**1974-1982.

59. **Perego M, Glaser P, Hoch JA.** 1996. Aspartyl-phosphate phosphatases deactivate the response regulator components of the sporulation signal transduction system in *Bacillus subtilis*. Mol Microbiol **19:**1151-1157.

60. **Tschowri N, Busse S, Hengge R.** 2009. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. Genes Dev **23:**522-534.

61. **Gusa AA, Scott JR.** 2005. The CovR response regulator of group A *streptococcus* (GAS) acts directly to repress its own promoter. Mol Microbiol **56:**1195-1207.

62. **Ansaldi M, Simon G, Lepelletier M, Mejean V.** 2000. The TorR high-affinity binding site plays a key role in both torR autoregulation and *torCAD* operon expression in *Escherichia coli*. J Bacteriol **182:**961-966.

63. **Svenningsen SL, Tu KC, Bassler BL.** 2009. Gene dosage compensation calibrates four regulatory RNAs to control *Vibrio cholerae* quorum sensing. EMBO J **28:**429-439.

64. **Goulian M.** 2010. Two-component signaling circuit structure and properties. Curr Opin Microbiol**13:**184-189.

65. **Fujita M, Gonzalez-Pastor JE, Losick R.** 2005. High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis*. J Bacteriol **187:**1357-1368.

66. **Throup JP, Koretke KK, Bryant AP, Ingraham KA, Chalker AF, Ge Y, Marra A, Wallis NG, Brown JR, Holmes DJ, Rosenberg M, Burnham MK.**

2000. A genomic analysis of two-component signal transduction in *Streptococcus pneumoniae*. Mol Microbiol **35:**566-576.

67. **Bronner S, Monteil H, Prevost G.** 2004. Regulation of virulence determinants in *Staphylococcus aureus*: complexity and applications. FEMS Microbiol Rev **28:**183-200.

68. **Beier D, Gross R.** 2006. Regulation of bacterial virulence by two-component systems. Curr Opin Microbiol **9:**143-152.

69. **Chen MH, Takeda S, Yamada H, Ishii Y, Yamashino T, Mizuno T.** 2001. Characterization of the RcsC->YojN->RcsB phosphorelay signaling pathway involved in capsular synthesis in *Escherichia coli*. Biosci Biotechnol Biochem **65:**2364-2367.

70. **Clemons KV, Miller TK, Selitrennikoff CP, Stevens DA.** 2002. fos-1, a putative histidine kinase as a virulence factor for systemic *aspergillosis*. Med Mycol **40:**259-262.

71. **Nemecek JC, Wuthrich M, Klein BS.** 2006. Global control of dimorphism and virulence in fungi. Science **312:**583-588.

72. **Desai C, Mavrianos J, Chauhan N.** 2011. *Candida albicans* SRR1, a putative two-component response regulator gene, is required for stress adaptation, morphogenesis, and virulence. Eukaryot Cell **10:**1370-1374.

73. **Dubrac S, Msadek T.** 2008. Tearing down the wall: peptidoglycan metabolism and the WalK/WalR (YycG/YycF) essential two-component system. Adv Exp Med Biol **631:**214-228.

74. **Beier D, Frank R.** 2000. Molecular characterization of two-component systems of *Helicobacter pylori*. J Bacteriol **182:**2068-2076.

75. **Dietz P, Gerlach G, Beier D.** 2002. Identification of target genes regulated by the two-component system HP166-HP165 of *Helicobacter pylori*. J Bacteriol **184:**350-362.

76. **George EA, Muir TW.** 2007. Molecular mechanisms of agr quorum sensing in virulent *Staphylococci*. Chembiochem **8:**847-855.

77. **Fournier B, Hooper DC.** 2000. A new two-component regulatory system involved in adhesion, autolysis, and extracellular proteolytic activity of *Staphylococcus aureus*. J Bacteriol **182:**3955-3964.

78. **Yarwood JM, McCormick JK, Schlievert PM.** 2001. Identification of a novel two-component regulatory system that acts in global regulation of virulence factors of *Staphylococcus aureus*. J Bacteriol **183:**1113-1123.

79. **Giraudo AT, Calzolari A, Cataldi AA, Bogni C, Nagel R.** 1999. The *sae* locus of *Staphylococcus aureus* encodes a two-component regulatory system. FEMS Microbiol Lett **177:**15-22.

80. **Gooderham WJ, Hancock RE.** 2009. Regulation of virulence and antibiotic resistance by two-component regulatory systems in *Pseudomonas aeruginosa*. FEMS Microbiol Rev **33:**279-294.

81. **Goodman AL, Merighi M, Hyodo M, Ventre I, Filloux A, Lory S.** 2009. Direct interaction between sensor kinase proteins mediates acute and chronic disease phenotypes in a bacterial pathogen. Genes Dev **23:**249-259.

82. **Whitchurch CB, Erova TE, Emery JA, Sargent JL, Harris JM, Semmler AB, Young MD, Mattick JS, Wozniak DJ.** 2002. Phosphorylation of the *Pseudomonas aeruginosa* response regulator AlgR is essential for type IV fimbria-mediated twitching motility. J Bacteriol **184:**4544-4554.

83. **Kuchma SL, Connolly JP, O'Toole GA.** 2005. A three-component regulatory system regulates biofilm maturation and type III secretion in *Pseudomonas aeruginosa*. J Bacteriol **187:**1441-1454.

84. **Laskowski MA, Osborn E, Kazmierczak BI.** 2004. A novel sensor kinase-response regulator hybrid regulates type III secretion and is required for virulence in *Pseudomonas aeruginosa*. Mol Microbiol **54:**1090-1103.

85. **Dong YH, Zhang XF, Soo HM, Greenberg EP, Zhang LH.** 2005. The two-component response regulator PprB modulates quorum-sensing signal production and global gene expression in *Pseudomonas aeruginosa*. Mol Microbiol **56:**1287-1301.

86. **Rumbaugh KP, Griswold JA, Hamood AN.** 1999. *Pseudomonas aeruginosa* strains obtained from patients with tracheal, urinary tract and wound infection: variations in virulence factors and virulence genes. J Hosp Infect **43:**211-218.

87. **Rodrigue A, Quentin Y, Lazdunski A, Mejean V, Foglino M.** 2000. Two-component systems in *Pseudomonas aeruginosa*: why so many? Trends Microbiol **8:**498-504.

88. **Guzman-Verri C, Manterola L, Sola-Landa A, Parra A, Cloeckaert A, Garin J, Gorvel JP, Moriyon I, Moreno E, Lopez-Goni I.** 2002. The two-component system BvrR/BvrS essential for *Brucella abortus* virulence regulates the expression of outer membrane proteins with counterparts in members of the *Rhizobiaceae*. Proc Natl Acad Sci U S A **99:**12375-12380.

89. **Tzeng YL, Datta A, Ambrose K, Lo M, Davies JK, Carlson RW, Stephens DS, Kahler CM.** 2004. The MisR/MisS two-component regulatory system influences inner core structure and immunotype of lipooligosaccharide in *Neisseria meningitidis*. J Biol Chem **279:**35053-35062.

90. **Ohtani K, Kawsar HI, Okumura K, Hayashi H, Shimizu T.** 2003. The VirR/VirS regulatory cascade affects transcription of plasmid-encoded putative virulence genes in *Clostridium perfringens* strain 13. FEMS Microbiol Lett **222:**137-141.

91. **Bijlsma JJ, Groisman EA.** 2005. The PhoP/PhoQ system controls the intramacrophage type three secretion system of *Salmonella enterica*. Mol Microbiol **57:**85-96.

92. **Feng X, Oropeza R, Kenney LJ.** 2003. Dual regulation by phospho-OmpR of *ssrA/B* gene expression in *Salmonella* pathogenicity island 2. Mol Microbiol **48:**1131-1143.

93. **Wosten MM, Kox LF, Chamnongpol S, Soncini FC, Groisman EA.** 2000. A signal transduction system that responds to extracellular iron. Cell **103:**113-125.

94. **Mouslim C, Groisman EA.** 2003. Control of the *Salmonella ugd* gene by three two-component regulatory systems. Mol Microbiol **47:**335-344.

95. **Cirillo DM, Valdivia RH, Monack DM, Falkow S.** 1998. Macrophage-dependent induction of the *Salmonella* pathogenicity island 2 type III secretion system and its role in intracellular survival. Mol Microbiol **30:**175-188.

96. **Bernardini ML, Fontaine A, Sansonetti PJ.** 1990. The two-component regulatory system ompR-envZ controls the virulence of *Shigella flexneri*. J Bacteriol **172:**6274-6281.

97. **Nakayama J, Cao Y, Horii T, Sakuda S, Akkermans AD, de Vos WM, Nagasawa H.** 2001. Gelatinase biosynthesis-activating pheromone: a peptide lactone that mediates a quorum sensing in *Enterococcus faecalis*. Mol Microbiol **41:**145-154.

98. **Arthur M, Quintiliani R, Jr.** 2001. Regulation of VanA- and VanB-type glycopeptide resistance in *enterococci*. Antimicrob Agents Chemother **45:**375-381.

99. **Arias CA, Courvalin P, Reynolds PE.** 2000. vanC cluster of vancomycin-resistant *Enterococcus gallinarum BM4174*. Antimicrob Agents Chemother **44:**1660-1666.

100. **DiRita VJ.** 1992. Co-ordinate expression of virulence genes by ToxR in *Vibrio cholerae*. Mol Microbiol **6:**451-458.

101. **Sengupta N, Paul K, Chowdhury R.** 2003. The global regulator ArcA modulates expression of virulence factors in *Vibrio cholerae*. Infect Immun **71:**5583-5589.

102. **Giammarinaro P, Sicard M, Gasc AM.** 1999. Genetic and physiological studies of the CiaH-CiaR two-component signal-transducing system involved in cefotaxime resistance and competence of *Streptococcus pneumoniae*. Microbiology **145 (Pt 8):**1859-1869.

103. **Novak R, Henriques B, Charpentier E, Normark S, Tuomanen E.** 1999. Emergence of vancomycin tolerance in *Streptococcus pneumoniae*. Nature **399:**590-593.

104. **Kadioglu A, Echenique J, Manco S, Trombe MC, Andrew PW.** 2003. The MicAB two-component signaling system is involved in virulence of *Streptococcus pneumoniae*. Infect Immun **71:**6676-6679.

105. **Graham MR, Smoot LM, Migliaccio CA, Virtaneva K, Sturdevant DE, Porcella SF, Federle MJ, Adams GJ, Scott JR, Musser JM.** 2002. Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and *in vivo* infection modeling. Proc Natl Acad Sci U S A **99:**13855-13860.

106. **Kreikemeyer B, Boyle MD, Buttaro BA, Heinemann M, Podbielski A.** 2001. Group A *streptococcal* growth phase-associated virulence factor regulation by a novel operon (*Fas*) with homologies to two-component-type regulators requires a small RNA molecule. Mol Microbiol **39:**392-406.

107. **Engleberg NC, Heath A, Miller A, Rivera C, DiRita VJ.** 2001. Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. J Infect Dis **183:**1043-1054.

108. **Perez E, Samper S, Bordas Y, Guilhot C, Gicquel B, Martin C.** 2001. An essential role for phoP in *Mycobacterium tuberculosis* virulence. Mol Microbiol **41:**179-187.

109. **Ewann F, Jackson M, Pethe K, Cooper A, Mielcarek N, Ensergueix D, Gicquel B, Locht C, Supply P.** 2002. Transient requirement of the PrrA-PrrB two-component system for early intracellular multiplication of *Mycobacterium tuberculosis*. Infect Immun **70:**2256-2263.

110. **Zahrt TC, Wozniak C, Jones D, Trevett A.** 2003. Functional analysis of the *Mycobacterium tuberculosis* MprAB two-component signal transduction system. Infect Immun **71:**6962-6970.

111. **Mandin P, Fsihi H, Dussurget O, Vergassola M, Milohanic E, Toledo-Arana A, Lasa I, Johansson J, Cossart P.** 2005. VirR, a response regulator critical for *Listeria monocytogenes* virulence. Mol Microbiol **57:**1367-1380.

112. **Mishra M, Parise G, Jackson KD, Wozniak DJ, Deora R.** 2005. The BvgAS signal transduction system regulates biofilm development in *Bordetella*. J Bacteriol **187:**1474-1484.

113. **Niehus E, Gressmann H, Ye F, Schlapbach R, Dehio M, Dehio C, Stack A, Meyer TF, Suerbaum S, Josenhans C.** 2004. Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori*. Mol Microbiol **52:**947-961.

114. **Pflock M, Dietz P, Schar J, Beier D.** 2004. Genetic evidence for histidine kinase HP165 being an acid sensor of *Helicobacter pylori*. FEMS Microbiol Lett **234:**51-61.

115. **MacKichan JK, Gaynor EC, Chang C, Cawthraw S, Newell DG, Miller JF, Falkow S.** 2004. The *Campylobacter jejuni* dccRS two-component system is required for optimal *in vivo* colonization but is dispensable for *in vitro* growth. Mol Microbiol **54:**1269-1286.

116. **Guo L, Lim KB, Gunn JS, Bainbridge B, Darveau RP, Hackett M, Miller SI.** 1997. Regulation of lipid A modifications by *Salmonella typhimurium* virulence genes *phoP-phoQ*. Science **276:**250-253.

117. **Gunn JS, Lim KB, Krueger J, Kim K, Guo L, Hackett M, Miller SI.** 1998. *PmrA-PmrB*-regulated genes necessary for 4-aminoarabinose lipid A modification and polymyxin resistance. Mol Microbiol **27:**1171-1182.

118. **Gal-Mor O, Segal G.** 2003. Identification of CpxR as a positive regulator of *icm* and *dot* virulence genes of *Legionella pneumophila*. J Bacteriol **185:**4908-4919.

119. **Rasmussen BA, Kovacs E.** 1993. Cloning and identification of a two-component signal-transducing regulatory system from *Bacteroides fragilis*. Mol Microbiol **7:**765-776.

120. **Grabenstein JP, Marceau M, Pujol C, Simonet M, Bliska JB.** 2004. The response regulator PhoP of *Yersinia pseudotuberculosis* is important for replication in macrophages and for virulence. Infect Immun **72:**4973-4984.

121. **Rasko DA, Moreira CG, Li de R, Reading NC, Ritchie JM, Waldor MK, Williams N, Taussig R, Wei S, Roth M, Hughes DT, Huntley JF, Fina MW, Falck JR, Sperandio V.** 2008. Targeting QseC signaling and virulence for antibiotic development. Science **321:**1078-1080.

122. **Silva JC, Haldimann A, Prahalad MK, Walsh CT, Wanner BL.** 1998. *In vivo* characterization of the type A and B vancomycin-resistant *enterococci* (VRE) VanRS two-component systems in *Escherichia coli*: a nonpathogenic model for studying the VRE signal transduction pathways. Proc Natl Acad Sci U S A **95:**11951-11956.

123. **Gotoh Y, Eguchi Y, Watanabe T, Okamoto S, Doi A, Utsumi R.** 2010. Two-component signal transduction as potential drug targets in pathogenic bacteria. Current Opinion in Microbiology **13:**232-239.

124. **Senadheera MD, Guggenheim B, Spatafora GA, Huang YC, Choi J, Hung DC, Treglown JS, Goodman SD, Ellen RP, Cvitkovitch DG.** 2005. A VicRK signal transduction system in *Streptococcus mutans* affects *gtfBCD*, *gbpB*, and *ftf* expression, biofilm formation, and genetic competence development. J Bacteriol **187:**4064-4076.

125. **Martin PK, Li T, Sun D, Biek DP, Schmid MB.** 1999. Role in cell permeability of an essential two-component system in *Staphylococcus aureus*. J Bacteriol **181:**3666-3673.

126. **Liu M, Hanks TS, Zhang J, McClure MJ, Siemsen DW, Elser JL, Quinn MT, Lei B.** 2006. Defects in *ex vivo* and *in vivo* growth and sensitivity to osmotic stress of group A *Streptococcus* caused by interruption of response regulator gene *vicR*. Microbiology **152:**967-978.

127. **Lange R, Wagner C, de Saizieu A, Flint N, Molnos J, Stieger M, Caspers P, Kamber M, Keck W, Amrein KE.** 1999. Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of *Streptococcus pneumoniae*. Gene **237:**223-234.

128. **Qin Z, Zhang J, Xu B, Chen L, Wu Y, Yang X, Shen X, Molin S, Danchin A, Jiang H, Qu D.** 2006. Structure-based discovery of inhibitors of the YycG histidine kinase: new chemical leads to combat *Staphylococcus epidermidis* infections. BMC Microbiol **6:**96.

129. **Hancock L, Perego M.** 2002. Two-component signal transduction in *Enterococcus faecalis*. J Bacteriol **184:**5819-5825.

130. **Kallipolitis BH, Ingmer H.** 2001. *Listeria monocytogenes* response regulators important for stress tolerance and pathogenesis. FEMS Microbiol Lett **204:**111-115.

131. **Sun J, Zheng L, Landwehr C, Yang J, Ji Y.** 2005. Identification of a novel essential two-component signal transduction system, YhcSR, in *Staphylococcus aureus*. J Bacteriol **187:**7876-7880.

132. **Fol M, Chauhan A, Nair NK, Maloney E, Moomey M, Jagannath C, Madiraju MV, Rajagopalan M.** 2006. Modulation of *Mycobacterium tuberculosis* proliferation by MtrA, an essential two-component response regulator. Mol Microbiol **60:**643-657.

133. **Nishino K, Yamaguchi A.** 2002. EvgA of the two-component signal transduction system modulates production of the yhiUV multidrug transporter in *Escherichia coli*. J Bacteriol **184:**2319-2323.

134. **Arthur M, Reynolds P, Courvalin P.** 1996. Glycopeptide resistance in *enterococci*. Trends Microbiol **4:**401-407.

135. **Novak R, Charpentier E, Braun JS, Tuomanen E.** 2000. Signal transduction by a death signal peptide: uncovering the mechanism of bacterial killing by penicillin. Mol Cell **5:**49-57.

136. **Groicher KH, Firek BA, Fujimoto DF, Bayles KW.** 2000. The *Staphylococcus aureus* lrgAB operon modulates murein hydrolase activity and penicillin tolerance. J Bacteriol **182:**1794-1801.

137. **Miller MB, Bassler BL.** 2001. Quorum sensing in bacteria. Annu Rev Microbiol **55:**165-199.

138. **Schauder S, Shokat K, Surette MG, Bassler BL.** 2001. The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. Mol Microbiol **41:**463-476.

139. **Hoch JA, Varughese KI.** 2001. Keeping signals straight in phosphorelay signal transduction. J Bacteriol **183:**4941-4949.

140. **Vanlaere E, Baldwin A, Gevers D, Henry D, De Brandt E, LiPuma JJ, Mahenthiralingam E, Speert DP, Dowson C, Vandamme P.** 2009. Taxon K, a complex within the *Burkholderia cepacia complex*, comprises at least two novel species, *Burkholderia contaminans* sp. nov. and *Burkholderia lata* sp. nov. Int J Syst Evol Microbiol **59:**102-111.

141. **Vial L, Chapalain A, Groleau MC, Deziel E.** 2011. The various lifestyles of the *Burkholderia cepacia* complex species: a tribute to adaptation. Environ Microbiol **13:**1-12.

142. **Drevinek P, Mahenthiralingam E.** 2010. *Burkholderia cenocepacia* in cystic fibrosis: epidemiology and molecular mechanisms of virulence. Clin Microbiol Infect **16:**821-830.

143. **Govan JR, Brown AR, Jones AM.** 2007. Evolving epidemiology of *Pseudomonas aeruginosa* and the *Burkholderia cepacia* complex in cystic fibrosis lung infection. Future Microbiol **2:**153-164.

144. **Courtney JM, Dunbar KE, McDowell A, Moore JE, Warke TJ, Stevenson M, Elborn JS.** 2004. Clinical outcome of *Burkholderia cepacia* complex infection in cystic fibrosis adults. J Cyst Fibros **3:**93-98.

145. **De Soyza A, Meachery G, Hester KL, Nicholson A, Parry G, Tocewicz K, Pillay T, Clark S, Lordan JL, Schueler S, Fisher AJ, Dark JH, Gould FK, Corris PA.** 2010. Lung transplantation for patients with cystic fibrosis and *Burkholderia cepacia* complex infection: a single-center experience. J Heart Lung Transplant **29:**1395-1404.

146. **Miller SC, LiPuma JJ, Parke JL.** 2002. Culture-based and non-growth-dependent detection of the *Burkholderia cepacia* complex in soil environments. Appl Environ Microbiol **68:**3750-3758.

147. **LiPuma JJ, Spilker T, Coenye T, Gonzalez CF.** 2002. An epidemic *Burkholderia cepacia* complex strain identified in soil. Lancet **359:**2002-2003.

148. **Baldwin A, Mahenthiralingam E, Drevinek P, Vandamme P, Govan JR, Waine DJ, LiPuma JJ, Chiarini L, Dalmastri C, Henry DA, Speert DP, Honeybourne D, Maiden MC, Dowson CG.** 2007. Environmental *Burkholderia cepacia* complex isolates in human infections. Emerg Infect Dis **13:**458-461.

149. **Nornberg Mde F, Mentges ML, Silveira ST, Tondo EC, Brandelli A.** 2011. A psychrotrophic *Burkholderia cepacia* strain isolated from refrigerated raw milk showing proteolytic activity and adhesion to stainless steel. J Dairy Res **78:**257-262.

150. **Jacobs JL, Fasi AC, Ramette A, Smith JJ, Hammerschmidt R, Sundin GW.** 2008. Identification and onion pathogenicity of *Burkholderia cepacia* complex isolates from the onion rhizosphere and onion field soil. Appl Environ Microbiol **74:**3121-3129.

151. **Dalmastri C, Baldwin A, Tabacchioni S, Bevivino A, Mahenthiralingam E, Chiarini L, Dowson C.** 2007. Investigating *Burkholderia cepacia* complex populations recovered from Italian maize rhizosphere by multilocus sequence typing. Environ Microbiol **9:**1632-1639.

152. **Mendes R, Pizzirani-Kleiner AA, Araujo WL, Raaijmakers JM.** 2007. Diversity of cultivated endophytic bacteria from sugarcane: genetic and biochemical characterization of *Burkholderia cepacia* complex isolates. Appl Environ Microbiol **73:**7259-7267.

153. **Lee YA, Chan CW.** 2007. Molecular Typing and Presence of Genetic Markers Among Strains of Banana Finger-Tip Rot Pathogen, *Burkholderia cenocepacia*, in Taiwan. Phytopathology **97:**195-201.

154. **Engledow AS, Medrano EG, Mahenthiralingam E, LiPuma JJ, Gonzalez CF.** 2004. Involvement of a plasmid-encoded type IV secretion system in the plant tissue watersoaking phenotype of *Burkholderia cenocepacia*. J Bacteriol **186:**6015-6024.

155. **Holden MT, Seth-Smith HM, Crossman LC, Sebaihia M, Bentley SD, Cerdeno-Tarraga AM, Thomson NR, Bason N, Quail MA, Sharp S, Cherevach I, Churcher C, Goodhead I, Hauser H, Holroyd N, Mungall K, Scott P, Walker D, White B, Rose H, Iversen P, Mil-Homens D, Rocha EP, Fialho AM, Baldwin A, Dowson C, Barrell BG, Govan JR, Vandamme P, Hart CA, Mahenthiralingam E, Parkhill J.** 2009. The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. J Bacteriol **191:**261-277.

156. **Loutet SA, Valvano MA.** 2010. A decade of *Burkholderia cenocepacia* virulence determinant research. Infect Immun **78:**4088-4100.

157. **Fuqua WC, Winans SC, Greenberg EP.** 1994. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. J Bacteriol **176:**269-275.

158. **Lutter E, Lewenza S, Dennis JJ, Visser MB, Sokol PA.** 2001. Distribution of quorum-sensing genes in the *Burkholderia cepacia* complex. Infect Immun **69:**4661-4666.

159. **Malott RJ, Baldwin A, Mahenthiralingam E, Sokol PA.** 2005. Characterization of the CciIR quorum-sensing system in *Burkholderia cenocepacia*. Infect Immun **73:**4982-4992.

160. **Boon C, Deng Y, Wang LH, He Y, Xu JL, Fan Y, Pan SQ, Zhang LH.** 2008. A novel DSF-like signal from *Burkholderia cenocepacia* interferes with *Candida albicans* morphological transition. ISME J **2:**27-36.

161. **Tomlin KL, Malott RJ, Ramage G, Storey DG, Sokol PA, Ceri H.** 2005. Quorum-sensing mutations affect attachment and stability of *Burkholderia cenocepacia* biofilms. Appl Environ Microbiol **71:**5208-5218.

162. **Huber B, Riedel K, Hentzer M, Heydorn A, Gotschlich A, Givskov M, Molin S, Eberl L.** 2001. The Cep quorum-sensing system of *Burkholderia cepacia* H111 controls biofilm formation and swarming motility. Microbiology **147:**2517-2528.

163. **O'Grady EP, Viteri DF, Malott RJ, Sokol PA.** 2009. Reciprocal regulation by the CepIR and CciIR quorum sensing systems in *Burkholderia cenocepacia*. BMC Genomics **10:**441.

164. **Lamothe J, Huynh KK, Grinstein S, Valvano MA.** 2007. Intracellular survival of *Burkholderia cenocepacia* in macrophages is associated with a delay in the maturation of bacteria-containing vacuoles. Cell Microbiol **9:**40-53.

165. **Tomich M, Griffith A, Herfst CA, Burns JL, Mohr CD.** 2003. Attenuated virulence of a *Burkholderia cepacia* type III secretion mutant in a murine model of infection. Infect Immun **71:**1405-1415.

166. **Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E.** 2005. Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu Rev Microbiol **59:**451-485.

167. **Hunt TA, Kooi C, Sokol PA, Valvano MA.** 2004. Identification of *Burkholderia cenocepacia* genes required for bacterial survival *in vivo*. Infect Immun **72:**4010-4022.

168. **Aubert DF, Flannagan RS, Valvano MA.** 2008. A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in *Burkholderia cenocepacia*. Infect Immun **76:**1979-1991.

169. **Aubert DF, O'Grady EP, Hamad MA, Sokol PA, Valvano MA.** 2013. The *Burkholderia cenocepacia* sensor kinase hybrid AtsR is a global regulator modulating quorum-sensing signalling. Environ Microbiol **15:**372-385.

170. **Mahenthiralingam E, Coenye T, Chung JW, Speert DP, Govan JR, Taylor P, Vandamme P.** 2000. Diagnostically and experimentally useful panel of strains from the *Burkholderia cepacia complex*. J Clin Microbiol **38:**910-913.

Chapter 2

Characterization of the AtsR phosphorelay pathway and identification of its response regulator in *Burkholderia cenocepacia*

Maryam Khodai-Kalaki, Daniel F. Aubert, Miguel A. Valvano

A version of this chapter has been published:

Maryam Khodai-Kalaki, Daniel F. Aubert, Miguel A. Valvano (2013).Characterization of the AtsR phosphorelay pathway and identification of its response regulator in *Burkholderia cenocepacia*. J. Biol. Chem. 2013, 288:30473-30484. doi: 10.1074/jbc.M113.489914

2.1 Introduction

Phosphorylation cascades governed by two-component signal transduction systems provide key signalling mechanisms in bacteria, Archaea, simple eukaryotes and higher plants, allowing them to translate signals into adaptive responses (1). These regulatory pathways consist of a transmembrane protein that responds to an environmental cue leading to autophosphorylation, followed by the transfer of the phosphate to a cytoplasmic response regulator (RR). The sensor protein has a variable region, dedicated to sensing a signal, and a conserved histidine kinase (HK) domain. The RR typically consists of a conserved receiver domain (RD) and a variable effector domain that binds to DNA. Upon perceiving the signal, the HK becomes autophosphorylated on its conserved histidine (His) residue and the phosphate transferred to a conserved aspartic acid (Asp) residue on the RR. More complex phosphorelay systems require multiple phosphoryl transfer reactions involving a hybrid sensor kinase that contains an extra domain serving as an RD, a histidine-phosphotransfer protein, and response regulator proteins (1, 2). Since the mechanism of phosphotransfer is from His to Asp, histidine-phosphotransfer proteins shuttle phosphoryl groups between a hybrid sensor kinase and a RR, providing additional checkpoints of regulation in signaling pathways (3, 4).

Burkholderia cenocepacia is a Gram-negative opportunistic pathogen belonging to the *Burkholderia cepacia* complex, which causes severe, chronic respiratory infections in patients with cystic fibrosis and other immunocompromised conditions (5, 6). Furthermore, *B. cenocepacia* can also be found in maize roots (7), present in sewage (8), and cause banana fingertip rot (9) and onion rot (10). This bacterium is also virulent in zebrafish (11), alfalfa, nematodes, and insect infection models (12), and can survive

within amoebae and macrophages (13, 14). We have previously identified AtsR (Adhesion and Type six secretion system regulator), which negatively regulates the expression of the quorum sensing-regulated virulence factors in *B. cenocepacia* (15, 16). *B. cenocepacia* strains encode two *N*-acylhomoserine lactone-dependent quorum-sensing systems, CepIR and CciIR, which coordinate the expression of ZmpA and ZmpB zinc metalloproteases and other virulence factors during infection (17-21). In the absence of *atsR*, expression of *cepIR* and *cciIR* is upregulated and mediates early and increased *N*-acylhomoserine lactone production, suggesting that AtsR plays a role in controlling virulence gene expression by modulating the timing of quorum sensing signalling (16). AtsR also represses the expression of virulence genes by an *N*-acylhomoserine lactone-independent mechanism (16). Consequently, inactivation of *atsR* in *B. cenocepacia* also leads to increased biofilm formation, adherence to polystyrene and lung epithelial cells, extracellular protease secretion and expression of a type 6 secretion system (T6SS). The latter is exemplified by actin cytoskeletal rearrangement with the formation of characteristic "pearls on a string-like structures" around infected macrophages (15, 16, 22).

AtsR is a predicted membrane protein with two transmembrane domains, and a sensor kinase region (containing the HK and ATPase domains) attached to an RD with conserved Asp residues. AtsR lacks a DNA binding motif found in canonical response regulator proteins, which suggests that AtsR does not bind directly to the promoter regions of target genes and is likely part of a multistep signal transduction pathway. Two genes near *atsR* encode two putative components of the AtsR phosphorelay pathway. One is *Bcam0381* (herein designated *atsT*), a gene co-transcribed with *atsR* (15). AtsT is a putative cytoplasmic transcriptional regulator containing an N-terminal helix-turn-helix

domain related to domains found in repressors (23) and a receiver domain with a conserved Asp at the C-terminus. The other is *Bcam0378*, which locates immediately upstream of *atsR* and encodes a hypothetical protein containing His and Asp residues that form part of a conserved motif, UPF0047 (24). BCAM0378 might be functionally necessary as a histidine-phosphotransfer protein intermediate to transfer the phosphate from AtsR to the putative response regulator AtsT.

In this work, I studied the mechanism of AtsR function by characterizing the role of critical functional residues within the individual domains of AtsR using phosphorylation assays. I identified the conserved residues His-245 and Asp-536 as phosphoacceptor sites in AtsR. Furthermore, I provided *in vivo* and *in vitro* evidence that AtsR is a hybrid sensor kinase that regulates downstream cellular activities through direct phosphorylation of AtsT.

2.2 Materials and methods

2.2.1 Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 2.1. Bacteria were grown in Luria broth (8) (Difco) at 37 °C unless indicated otherwise. *E. coli* cultures were supplemented, as required, with the following antibiotics (final concentrations): tetracycline 30 μg/ml, kanamycin 30 μg/ml and trimethoprim 50 μg/ml. *B. cenocepacia* cultures were supplemented, as required, with trimethoprim 100 μg/ml, tetracycline 150 μg/ml, ampicillin 200 μg/ml and polymyxin B 20 μg/ml.

Table 2.1: Strains and plasmids

Strain or plasmid	Relevant characteristics[a]	Source/reference
Strains		
B. cenocepacia		
K56-2	ET12 clone related to J2315, CF clinical isolate	[b]BCRRC
K56-2Δ*atsR*	Deletion of *atsR* in K56-2	(25)
K56-2Δ*cepI*	Deletion of *cepI* in K56-2	(16)
K56-2Δ*atsR*Δ*cepI*	Deletion of *cepI* in K56-2 Δ*atsR*	(16)
K56-2Δ*atsR atsR*$^{+}$	Chromosomal *atsR* integration in Δ*atsR*	This study
K56-2Δ*atsR*Δ*cepI atsR*$^{+}$	Chromosomal *atsR* integration in Δ*atsR*Δ*cepI*	This study
K56-2Δ*atsR*Δ*cepI atsR*$_{H245A}$$^{+}$	Chromosomal *atsR*$_{H245A}$ integration in Δ*atsR*Δ*cepI*	This study
K56-2Δ*atsR*Δ*cepI atsR*$_{D536A}$$^{+}$	Chromosomal *atsR*$_{D536A}$ integration in Δ*atsR*Δ*cepI*	This study
K56-2Δ*atsR*Δ*cepI atsR*Δ*RD*$^{+}$	Chromosomal *atsR*Δ*RD* integration in Δ*atsR*Δ*cepI*	This study
K56-2Δ*atsR atsR*$_{H245A}$$^{+}$	Chromosomal *atsR*$_{H245A}$ integration in Δ*atsR*	This study
K56-2Δ*atsR atsR*$_{D536A}$$^{+}$	Chromosomal *atsR*$_{D536A}$ integration in Δ*atsR*	This study
K56-2Δ*atsR atsR*Δ*RD*$^{+}$	Chromosomal *atsR*Δ*RD* integration in Δ*atsR*	This study
K56-2Δ*atsR*Δ*cepI* $^{+}$Δ*atsT atsR*$_{D536A}$	Deletion of *atsT* in Δ*atsR*Δ*cepI atsR*$_{D536A}$$^{+}$	This study
K56-2Δ*atsR*Δ*cepI* $^{+}$Δ*atsT atsR*Δ*RD*	Deletion of *atsT* in Δ*atsR*Δ*cepI atsR*Δ*RD*$^{+}$	This study
K56-2Δ*atsR* Δ*atsT atsR*$_{D536A}$$^{+}$	Deletion of *atsT* in Δ*atsR atsR*$_{D536A}$$^{+}$	This study
K56-2Δ*atsR* Δ*atsT atsR*Δ*RD*$^{+}$	Deletion of *atsT* in Δ*atsR atsR*Δ*RD*$^{+}$	This study
K56-2Δ*cepI* Δ*atsT*	Deletion of *atsT* in Δ*cepI*	This study
K56-2Δ*cepI* Δ*atsT atsT* $^{+}$	Chromosomal *atsT* integration in Δ*cepI* Δ*atsT*	This study
K56-2Δ*cepI* Δ*atsT atsT*$_{D208A}$ $^{+}$	Chromosomal *atsT*$_{D208A}$ integration in Δ*cepI* Δ*atsT*	This study
K56-2Δ*atsR*Δ*hcp*	Deletion of *hcp* in Δ*atsR*	(25)
E. coli		
DH5α	F^{-} ϕ80*lacZ* M15 *endA1 recA1 supE44 hsdR*17(r_K^{-} m_K^{+}) *deoR thi-1 nupG supE44 gyrA96 relA1* Δ(*lacZYA-argF*)*U169*, λ^{-}	Laboratory stock
BL21	F^{-} omp (Tgal *dcm ompT hsdS*(Rb $^{-}$ mB $^{-}$) *gal* λ(DE3)	Laboratory stock
GT115	F^{-} *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80Δ*lacZ*ΔM15 Δ*lacX74 recA1 rpsL endA1*Δ*dcm uidA*(ΔMluI)::*pir-116* Δ*sbcC-sbcD*	Laboratory stock
Plasmids		
pAtsRChr	*atsR* cloned in pMH447 for chromosomal complementation	This study
pAtsR$_{H245A}$Chr	*atsR*$_{H245A}$ cloned in pMH447 for chromosomal complementation	This study

pAtsR$_{D536A}$Chr	*atsR*$_{D536A}$ cloned in pMH447 for chromosomal complementation	This study
pAtsRΔRDChr	*atsRΔRD* cloned in pMH447 for chromosomal complementation	This study
pAtsTChr	*atsT* cloned in pMH447 for chromosomal complementation	This study
pAtsT$_{D208A}$Chr	*atsT*$_{D208A}$ cloned in pMH447 for chromosomal complementation	This study
pDAI-SceI-SacB	*ori*$_{pBBR1}$, Tet^R, P_{dhfr}, *mob*$^+$, expressing I-SceI, SacB	(26)
pET28a(+)	Cloning vector, IPTG inducible, N-terminal His, Kan^R	Laboratory stock
pDelatsT	pGPI-SceI with fragments flanking *atsT*	This study
pDelM0378	pGPI-SceI with fragments flanking *bcam0378*	This study
pMK1	pET28a(+) encoding *B. cenocepacia bcam0378*, 6X His: kan^R	This study
pMK2	pET28a(+) encoding *B. cenocepacia atsT*, 6X His: kan^R	This study
pMK4	pET28a(+) encoding *B. cenocepacia atsR-RD*$_{D536A}$, 6X His: kan^R	This study
pMK5	pET28a(+) encoding *B. cenocepacia atsR*$_{D536A}$, 6X His: kan^R	This study
pMZ24	pET28a(+) encoding *B. cenocepacia atsR*$_{205-605}$, 6X His: kan^R	M. AlZayer
pMZ25	pET28a(+) encoding *B. cenocepacia atsR*, 6X His: kan^R	M. AlZayer
pMZ33	pET28a(+) encoding *B. cenocepacia atsR*-HK, 6X His: kan^R	M. AlZayer
pMZ34	pET28a(+) encoding *B. cenocepacia atsR-RD*, 6X His: kan^R	M. AlZayer
pMZ36	pET28a(+) encoding *B. cenocepacia atsR atsR*$_{H245A}$,6X His:kan^R	M. AlZayer
pMH447	pGPI-SceI derivative used for chromosomal complementation	(27)
pRK2013	*ori*$_{colE1,}$ RK2 derivative, Kan^R, *mob*$^+$, *tra*$^+$	(28)

[a] Tp^R, trimethoprim resistance, Kan^R, kanamycin resistance, Tet^R, tetracycline resistance.

[b] BCRRC, *B. cepacia* Research and Referral Repository for Canadian CF Clinics.

2.2.2 General molecular techniques

DNA manipulations were performed as described previously (29). T4 DNA ligase (Roche Diagnostics, Laval, Quebec, Canada) and Antarctic phosphatase (New England Biolabs, Pickering, Ontario, Canada) were used as recommended by the manufacturers. Transformation of *E. coli* DH5α and *E. coli* GT115 was done using the calcium chloride method (30). Mobilization of complementing plasmids and mutagenesis plasmids into *B. cenocepacia* K56-2 was performed by triparental mating using *E. coli* DH5α carrying the helper plasmid pRK2013 (28, 31). DNA amplification by PCR was performed using Bio-Rad C1000^{TM} Thermal Cycler with Taq or HotStar HiFidelity DNA polymerases (Qiagen, Canada). DNA sequences of all primers used in this study are described in Table 2.2. DNA sequencing was performed at the DNA sequencing Facility of York University, Toronto, Canada. BLAST was used to analyse the sequenced genome of *B. cenocepacia* strain J2315.

2.2.3 Plasmid construction and chromosomal complementation

Unmarked and non-polar deletions were performed as described previously (16, 32). To delete *atsT* (*Bcam0381)* and *Bcam0378*, PCR amplifications of regions flanking these genes were performed individually using 2844-2836 and 2840-2839 primer pairs for *atsT* and 4009-4010 and 4011-4012 primer pairs for *Bcam0378*. The amplicons were digested with the restriction enzymes *Xba*I-*Xho*I and *Xho*I-*Eco*RI respectively, and cloned into the mutagenic plasmid pGPI-*SceI* digested with *Xba*I and *Eco*RI giving rise to pDelatsT and pDelM0378. To create His-tagged fusions of BCAM0378, AtsT, AtsR and its truncated versions including AtsR-HK domain (residues 205 to 460), AtsR-RD receiver domain

(residues 488 to 606), AtsRΔTM (residues 205 to 606), and AtsRΔRD, sequences were amplified from *B. cenocepacia* K56-2 genomic DNA using primers 5798-5799; 5108-5109; 4632-4633; 4799-4634; 4800-4736; 4634-4633; 4632-4799, respectively (Table 2.2). Amplicons were digested with the appropriate restriction enzymes and ligated into similarly digested pET28a cloning vector. pET28a (Novagen) was used to engineer C-terminally His-tagged proteins for expression in *E. coli* BL21. Site-directed mutagenesis was performed using the QuikChange Site-Directed Mutagenesis Kit from Stratagene (Santa Clara, California), as recommended by the supplier. Primers were designed with 15-20 nucleotides flanking each side of the targeted mutation. Plasmid pMZ25 was used as a template to create H245A and D536A using 4880-4881 and 5959-5960 respectively. Plasmid pMK2 was used as a template to create D208A using 6997-6998. The resulting PCR products were digested with *Dpn*I and introduced into *E. coli* DH5α. All constructs and replacement mutants were confirmed by sequencing.

Chromosomal complementations of Δ*atsR,* Δ*atsR*Δ*cepI* or Δ*cepI*Δ*atsT* were performed using the pMH447 plasmid (26, 27). Primers 5866-4632 was used to PCR amplify *atsR* and its mutated versions. The amplicon was digested with *Nde*I-*Xba*I and cloned into the similarly digested pMH447, giving rise to pAtsRChr, pAtsR$_{H245A}$Chr, pAtsR$_{D536A}$Chr and pAtsRΔRDChr. Likewise, primers 7020-7021 was used to PCR amplify *atsT* and its mutated version. The amplicons were digested with *Nde*I and cloned into the similarly digested pMH447, giving rise to pAtsTChr and pAtsT$_{D536A}$Chr.

Table 2.2: Oligonucleotide primers

Primer No.	5'-3' Primer sequence	Restriction enzyme*
2836	TTTGCTCGAGTTGTTGGCGGTGATGAGAT	XhoI
2839	TAGGAATTCAGGATCACGCCGTACTTGTC	EcoRI
2840	TTTTCTCGAGCACATCGTGTCGCACTACAA	XhoI
2844	TACGTCTAGAATCTTTAGGGCGATCGGGAAC	XbaI
2869	TTTCATATGTTTATCGTCGTCATCTTTGT	NdeI
4009	TTTTTCTAGACACCGAGCAACGCTACAC	XbaI
4010	TTTTCTCGAGCGTGATGGCCTGTTGCAT	XhoI
4011	TTTTCTCGAGGATATCGTGCTGCATCT	XhoI
4012	TTTTGAATTCACCTCGTCGTGCTCGATCT	EcoRI
4632	AAAACATATGACGCGGCGGCGATGGAAGAA	NdeI
4633	AAAAGCGGCCGCGGCGAGCAGTGTCTCGACGA	NotI
4634	AAAACATATGCGCACGCGCGACGACCT	NdeI
4736	AAAAGCGGCCGCTCAGGCGAGCAGTGTCTCGACGA	NotI
4799	AAAAGCGGCCGCTCATTCGACCGGCAGCGTCAC	NotI
4800	AAAACATATGGCGCTGGTGGTCGACGAC	NdeI
4880	GCGTTCCTCGGGATGGTCAGCGCCGAACTGCGCACG CCGCTG	N/A
4881	CAGCGGCGTGCGCAGTTCGGCGCTGACCATCCCGA GGAACGC	N/A
5108	AAAACATATGTCCACCACCGAGCAGGCCAA	NdeI
5109	AAAAGAATTCTCAGTTCGTCGCGGCCGCTG	EcoRI
5798	TTTTTTTCATATGCAACAGGCCATCACG	NdeI
5799	TTTTGCGGCCGCCTACTCGCCGAGCAGATG	NotI
5866	TTTTTCTAGAGCTTTGTTAGCAGCCGGATC	XbaI
5885	TTGATGGCGAGCGATTCTTC	N/A
5886	CCAGTTCTTCAGCGTGACGA	N/A
5959	CTCGTCGTGCTCGCGCTCGAACTGCCG	N/A
5960	CGGCAGTTCGAGCGCGAGCACGACGA	N/A
6997	GATGGCTACATCCTCGCCTGGATGCTCGGCG	N/A
6998	GTCGCCGAGCATCCAGGCGAGGATGTAGCC	N/A
7020	TTTTTTCATATGTGGAGACCCCATCGAGATG	NdeI
7021	TTTTTTCATATGGCATGCCTGCAGGTCGTCAC	NdeI

*****Restriction endonuclease sites incorporated in the oligonucleotide sequences are underlined. N/A indicates absence of restriction site.

2.2.4 Protein expression and purification

For over-expression and purification of recombinant proteins, a single colony was inoculated in LB broth supplemented with 30 μg/ml of kanamycin and grown at 37 ºC. Optical Density (OD_{600}) was monitored until it reached 0.6. The culture was then shifted to 30 ºC and 0.2 mM isopropyl-beta-D-thiogalactopyranoside was added to induce the expression of proteins. Cultures were incubated for an additional 4 h. Samples were analyzed by SDS-PAGE and stained with Coomassie blue to check the induction of proteins. Cells were harvested by centrifugation at 8,000 x*g* for 10 min. Cell pellets were resuspended in lysis buffer containing [(50 mM Na phosphate pH 8, 0.3 M NaCl, 10 mM imidazole pH 8.0, 10% glycerol, 0.25% Tween 20, and 1X of EDTA-free protease inhibitor cocktail (Sigma))], and disrupted by One Shot cell disruptor (Thermo Scientific, Rockville, MD, USA). After centrifugation at 27,000 x*g* for 30 min at 4 ºC to remove the debris, the clarified cell lysate was loaded onto a Ni^{+2}-binding sepharose beads, washed, and the His-tagged proteins were then eluted using increased gradient concentrations of imidazole (125-500 mM). Fractions were analyzed by SDS-PAGE and stained with Coomassie blue to determine the integrity of the purified protein. Fractions were pooled and buffer-exchanged against dialysis buffer (100 mM Tris-HCl pH 8, 50 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 10% glycerol) using Thermoscientific Slide-A-Lyzer mini dialysis devices. Proteins were concentrated in Amicon ultrafiltration devices (10 kDa molecular weight cut off) and protein concentration was determined by standard Bradford assay (Bio-Rad, Hercules, CA). The purity of proteins was evaluated by SDS-PAGE followed by Coomassie Brilliant Blue staining (33).

2.2.5 *In vitro* phosphorylation assay

For autophosphorylation and phosphotransfer reactions 5 µmol of each protein was added to the phosphorylation buffer containing 50 mM Tris-HCl pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 1 mM DTT and 5 µCi [γ-^{33}P]-ATP (specific activity of 3000 Ci/mmol; 3.3 µM stock solution) (Perkin Elmer). The reactions were carried out at room temperature for the desired time and were terminated by adding 3X sample buffer (32). Experiments were conducted in presence or absence of the response regulator in a final volume of 10 µl. The reaction products were separated by electrophoresis on 14% SDS-PAGE gels. Gels were fixed, exposed to a high-resolution screen (Kodak), and analyzed using a Phosphorimager with ImageQuant software (Molecular Dynamics 5.0; Amersham Biosciences). For pulse-chase experiments, an excess of nonradioactive ATP (20 mM) was added to the reaction mixture following a 10-min pre-incubation with (γ-^{33}p) ATP. Aliquots were taken prior to the addition of unlabeled ATP (time zero) and at various time points after the addition of cold ATP.

2.2.6 Analysis of chemical stability of H245-p and D536-P

Phosphorylated AtsR and $AtsR_{D536A}$ proteins were prepared as described above and treated with 1 M HCl, 1 M NaOH or left untreated for 45 min at RT. The HCl reaction was neutralized with 0.25 vol. of 2 M Tris pH 8 and analyzed by 16% SDS-PAGE followed by Phosphorimager.

2.2.7 Protease, swarming motility and biofilm formation assays

Protease assays were performed as described previously (16). Briefly, 18 h cultures were normalized to an OD_{600} of 1. The bacterial suspension (3 µl) was spotted onto dialysed

brain heart infusion (D-BHI) agar plates containing 1.5% carnation milk. The plates were incubated at 37 °C and examined for zones of clearing around the bacterial spots at 48 h. The protease activity was recorded by measuring the radius of the surrounding halo (from the outside of the spot to the edge of the halo). Swarming motility assays were performed as described previously (16). Three µl of overnight culture, adjusted to an OD_{600} of 1, was spotted on a swarm plate (0.8% Nutrient Broth, 0.5% agar, 0.2% glucose). Plates were incubated at 37 °C for 20 h and the diameters of swarming zones were measured. Biofilm formation assays were performed as described previously (15). Assays were done in triplicate and repeated independently three times.

2.2.8 Macrophage infections and T6SS activity

Infections were performed as previously described (15) using the C57BL/6 murine bone marrow-derived macrophage cell line ANA-1. Bacteria were added to ANA-1 cells grown on glass cover slips at a MOI of 50. Cover slips were analysed by phase contrast microscopy after 4 h of incubation at 37 °C. T6SS activity was recorded as the ability of the bacteria to induce the formation of characteristic ectopic actin nucleation around the macrophages (15, 22).

2.2.9 Western blot analysis

For *E. coli*, overnight bacterial cultures in 5 mL of LB were diluted to an initial OD_{600} of 0.2 and incubated at 37 °C until reaching an OD_{600} of 0.7. At this point, isopropyl-beta-D-thiogalactopyranoside was added to a final concentration of 0.2 mM. Cells were incubated for 4 h at 30 °C and were then harvested by centrifugation at 8,000 x*g* for 10 min at 4 °C. The bacterial pellet was suspended in lysis buffer and the suspension lysed

using a One-shot cell disrupter (Thermo Scientific, Rockville, MD, USA). Cell debris were removed by centrifugation (15,000 x*g* for 15 min at 4 °C), and the clear supernatant was centrifuged at 40,000 x*g* for 30 min at 4 °C. The pellet, containing total membranes, was suspended in lysis buffer. The protein concentration was determined by the Bradford assay (Bio-Rad, Hercules, CA). Staining was performed with Coomassie brilliant blue. SDS–PAGE, protein transfers to nitrocellulose membranes and immunoblots were performed as described (34). For detection of 6xHis-tagged proteins, membranes were incubated with a 1:10,000 dilution of anti-His IgG2a monoclonal antibodies (Amersham, Piscataway, NJ) and Alexa Fluor 680 anti-mouse IgG antibodies (molecular probes). For *B. cenocepacia*, overnight bacterial cultures in 5 ml of LB were diluted to an initial OD_{600} of 0.2 and incubated at 37 °C for 8 h. His-tagged proteins were purified and detected by Western blot using an anti-His antibody as indicated above.

2.3 Results and discussion

2.3.1 Deletion of atsT causes the same phenotypes as ΔatsR mutants

I reasoned that if AtsT (BCAM0381) was involved in the AtsR phosphorelay pathway, Δ*atsT* and Δ*atsR* mutants should have similar phenotypes. Therefore, Δ*atsT* was tested in our macrophage infection model. Phase-contrast microscopy revealed that, in contrast to K56-2, Δ*atsR* and Δ*atsT* noticeably induce the formation of "pearls on a string-like structures" around infected macrophages (Fig. 2.1). These structures depend on T6SS-mediated rearrangements of host actin, and are characteristic of an upregulated T6SS (15, 16, 22). *B. cenocepacia* Δ*atsR*Δ*hcp*, a T6SS defective mutant, was used as a negative control during the infections (25), and as expected it did not mediate the "pearls on a string" phenotype in infected macrophages (Fig. 2.1A). Similarly to Δ*atsR* (15), Δ*atsT*

showed increased biofilm formation as evidenced by a robust ring of biofilm at the air-liquid interface (Fig. 2.1B). Together, these results strongly suggest that AtsT and AtsR are in the same regulatory pathway, which is likely initiated by AtsR.

2.3.2 Identification of phosphoacceptor residues within AtsR

The region of AtsR corresponding to the HK domain (amino acids 233-457) was aligned with the HK domain of two well-characterized hybrid sensor kinase proteins, ArcB from *Escherichia coli* and RetS from *P. aeruginosa* (3, 35-37). The amino acid sequence of AtsR (606 amino acids) is 34% identical and 50% similar to that of ArcB, and 26% identical and 48% similar to that of RetS. The invariant His residue that is the site of autophosphorylation in these proteins corresponds to His-245 in AtsR. Also, the AtsR-RD (amino acids 488-601) showed 52% and 50% similarity at the amino acid level to ArcB and RetS, respectively. This region has an invariant Asp at position 536, which corresponds to the site of phosphorylation in ArcB and RetS.

Truncated versions of AtsR spanning AtsR-HK and AtsR-RD were constructed and purified to analyze their biochemical properties. Residues 1-200, comprising the membrane spanning domains were removed to facilitate protein solubility and purification. Furthermore, to assess the contribution of conserved His and Asp residues to AtsR phosphorylation, His-245 and Asp-536 were individually replaced by alanine (Fig. 2.2A). The truncated proteins along with the mutated polypeptides were assessed for their autophosphorylation activity *in vitro* using [γ-^{33}P]-ATP. Only the polypeptides containing His-245 underwent autophosphorylation (Fig. 2.2B, lanes 1 and 3). Therefore, AtsR HK containing the native His-245 is necessary and sufficient for initiation of the phosphorelay

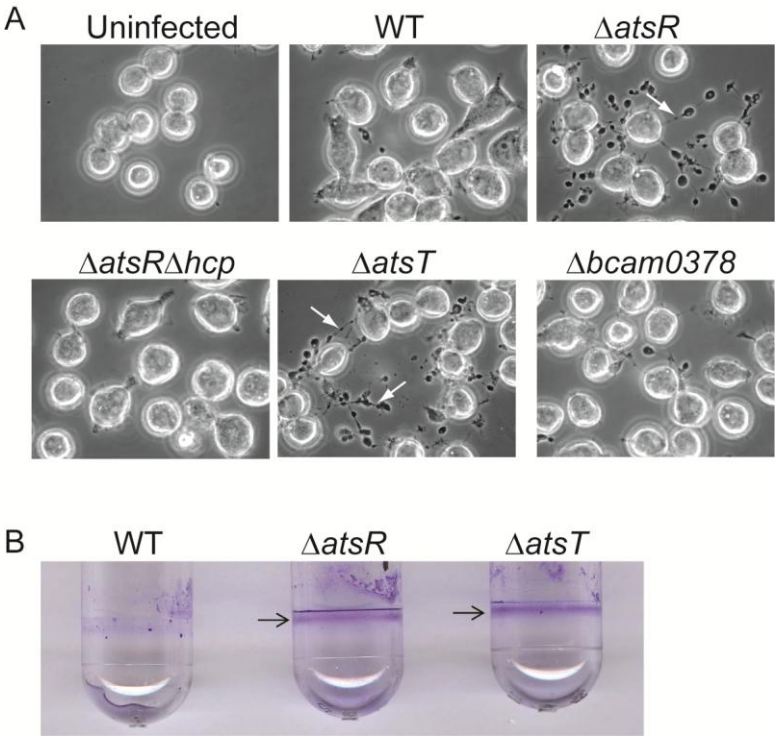


Figure 2.1 T6SS activity and biofilm formation of *B. cenocepacia* K56-2 wild-type and its mutant derivatives. A) Phase-contrast microscopy of infected ANA-1 murine macrophages to assess T6SS activity. Infections were performed at a multiplicity of infection of 50 for 4 h. The arrows indicate the presence of ectopic actin nucleation ("pearls on a string" phenotype; 15; 22) extending from infected macrophages, which denotes T6SS activity. *B. cenocepacia* K56-2 Δ*atsR*Δ*hcp*, a T6SS defective mutant, was used as a negative control during the infections. Experiments consisted of three independent biological repeats where similar results were obtained. B) Biofilm formation by parental strains, Δ*atsR* and Δ*atsT* mutants. *B. cenocepacia* K56-2 wild-type and derivative mutants were tested for biofilm formation by crystal violet staining. Arrows indicate the ring corresponding to the biofilm formation characteristic in Δ*atsR* and Δ*atsT* mutants. The experiment was repeated three times in triplicates and pictures were taken after 24 h of static incubation at 37 °C.

pathway since the replacement of His with alanine at position 245 abolished autophosphorylation (Fig. 2.2B, lanes 2 and 6). Moreover, the AtsR-RD is unable to autophosphorylate in the absence of the HK domain (Fig. 2.2B, lane 4). The ability of AtsR-RD to accept the phosphoryl group from the AtsR-HK was tested in a phosphotransfer assay (Fig. 2.2B, lanes 5-7). Two bands were obtained after incubating AtsR-HK and AtsR-RD together in a reaction with [γ-^{33}P]-ATP (Fig. 2.2B, lane 5). These bands corresponded to the phosphorylated form of these proteins, indicating the transfer of phosphoryl group from the HK domain to the RD domain of AtsR. To test whether Asp-536 within AtsR-RD is the residue accepting the phosphate from His-245, AtsR-RD_{D536A} was added to the reaction with AtsR-HK and [γ-^{33}P]-ATP. No band corresponding to AtsR-RD_{D536A} was detected (Fig. 2.2B, lane 7) clearly indicating that Asp-536 is necessary for phosphorylation.

2.3.3 Chemical stability of phosphorylated proteins

Phosphor amidates, such as histidine or lysine phosphate, are stable in alkali conditions but extremely labile to acid. Acyl phosphates such as aspartate and glutamate are, on the other hand, labile to either pH extremes (38). To compare wild-type AtsR and $AtsR_{D536A}$ chemical stability, phosphorylation reactions were performed for each, and the products were subjected to either no treatment or treatment with acid or alkali. If the wild-type AtsR protein is phosphorylated at both His-245 and Asp-536 and the $AtsR_{D536A}$ is phosphorylated at only His-245, the two proteins should behave distinctively upon treatment with base. As shown in Fig. 2.2C (lanes 5-6), AtsR and $AtsR_{D536A}$ were both labile to acid and relatively stable to base (lanes 3-4), indicating that these two proteins are primarily amidyl phosphates. Considering the quantitative difference between the

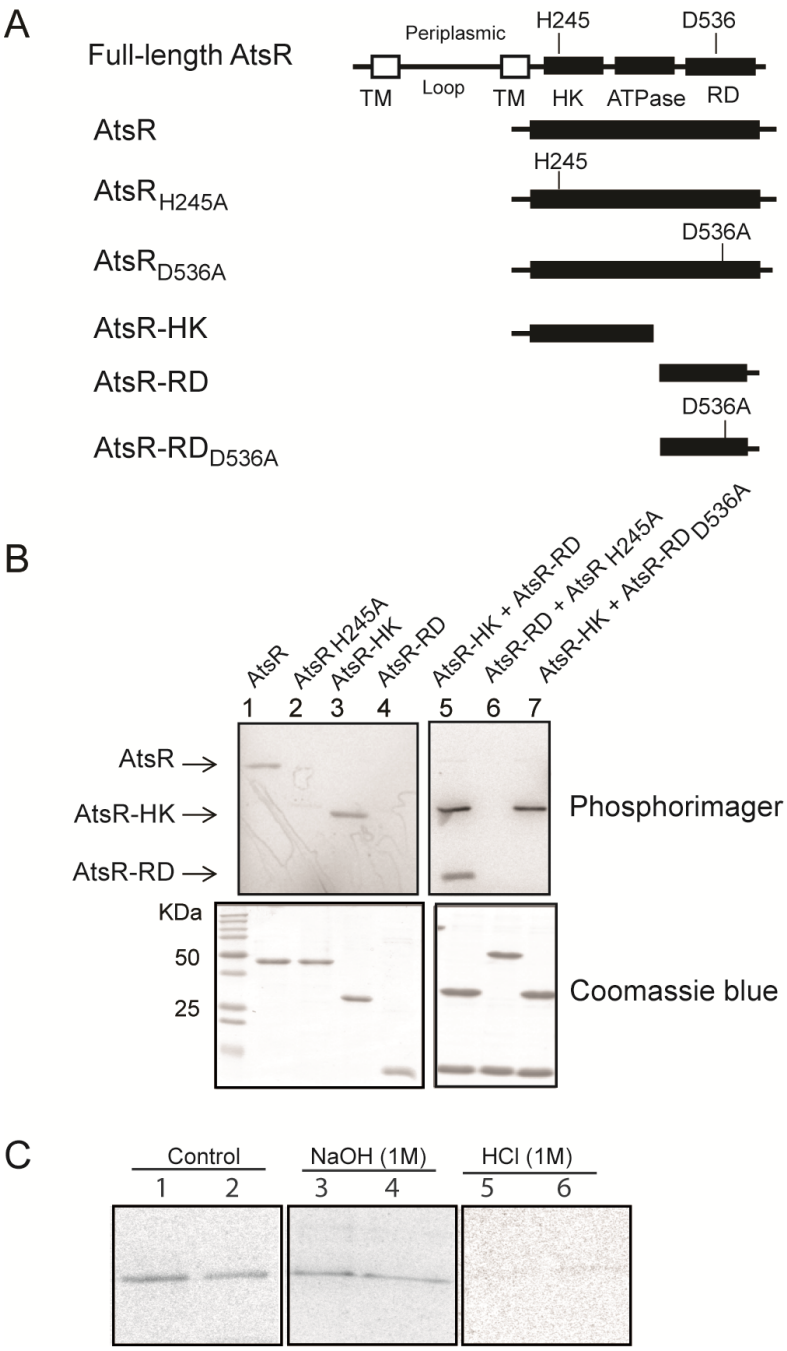


Figure 2.2: Functional analysis of AtsR domains. A) Schematic domain organization of AtsR and its derivatives (domains are not drawn to scale). The predicted sites of phosphorylation are His-245 (H245) and Asp-536 (D536). TM: transmembrane domain, HK: Histidine kinase domain, ATPase: ATPase domain, RD: Receiver domain, A:

Alanine. B) *In vitro* phosphorylation assays. Purified proteins were added in a standard phosphorylation mixture [100 mM Tris-HCl pH 8, 50 mM KCl, 5 mM MgCl2, 1 mM DTT, 5 μCi [γ-^{33}P]-ATP] and incubated for 15 min at 25 °C. Samples were resolved on 16% SDS-PAGE gel. The phosphorylated proteins were visualized using a phosphorimager (top). Phosphorylated and non-phosphorylated proteins were revealed by Coomassie Blue staining (bottom). The location of phosphorylated bands is denoted with arrows. C) Chemical stability of phosphorylated proteins. Phosphorylated AtsR (lanes 1, 3 and 5) and $AtsR_{D536A}$ (lanes 2, 4 and 6) were treated with 1 M NaOH or 1 M HCl or were left untreated for 45 min at RT. The reactions were neutralized with 2 M Tris pH 8 and analyzed by Phosphorimager after SDS-PAGE.

wild-type AtsR protein before and after alkali treatment in Fig. 2.2C, I concluded that at least a portion of the wild-type protein may be phosphorylated at the Asp-536. In quantitative analysis using densitometry, the wild-type AtsR retained 82 ± 6% of its label after alkali treatment compared with the untreated control, whereas the $AtsR_{D536A}$ mutant retained 91 ± 5% of its label. These results therefore suggest that the majority of phosphate present in both AtsR and $AtsR_{D536A}$ is in the form of amidyl phosphate.

2.3.4 AtsR and AtsT form a cognate HK-RR pair

I initially hypothesized that BCAM0378 could be a histidine-phosphotransfer protein component of the AtsR phosphorelay based on the proximity of *Bcam0378* to *atsR* (*Bcam0379*) and bioinformatic data indicating the presence of highly conserved His and Asp residues within this family of proteins. To investigate the relationship between AtsR, BCAM0378 and AtsT we performed phosphotransfer assays employing combinations of these different proteins. The histidine-phosphotransfer candidate, BCAM0378 was not phosphorylated (Fig. 2.3, lanes 1, 2 and 3). To rule out the possibility of a rapid phosphotransfer from AtsR to BCAM0378, we repeated the experiment by reducing the incubation time from 10 min to either 1 or 5 min and increasing the temperature to 30 ˚C. A phosphorylated form of BCAM0378 was never detected in any condition tested (data not shown). However, AtsT was phosphorylated directly by AtsR and independently from BCAM0378 (Fig. 2.3, lane 4). To test whether AtsT autophosphorylates independently from AtsR, we incubated AtsT with $AtsR_{H245A}$ and BCAM0378 (Fig. 2.3, lane 3). No phosphorylated band was detected indicating that AtsT is unable to autophosphorylate, and it requires the HK domain of AtsR with the native His-245 to be phosphorylated. Together, these data suggest that AtsR acts as a conventional TCS that mediates signal

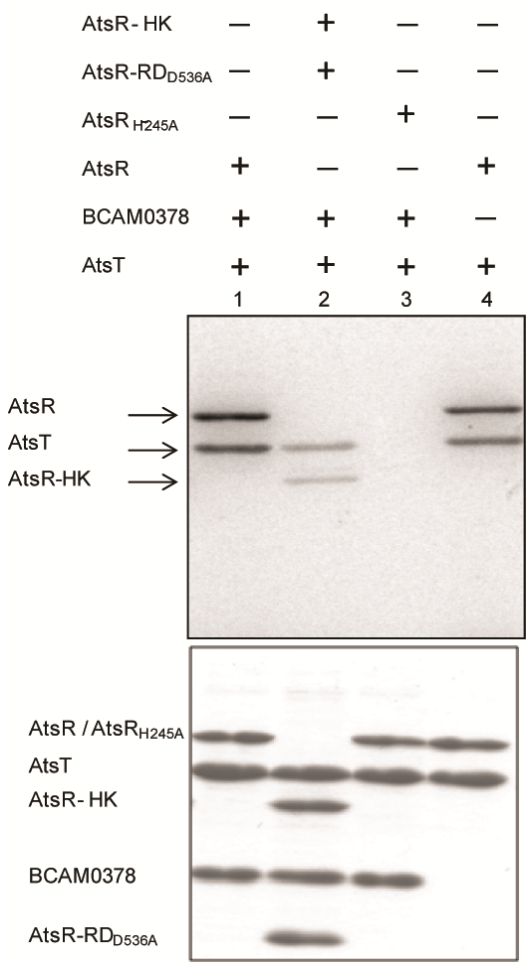


Figure 2.3: Phosphotransfer from AtsR and its derivatives to the AtsT response regulator. Five µmol of AtsR, AtsR$_{H245A}$ and AtsR-HK, was pre-incubated in individual standard phosphorylation mixtures for 15 min at 25 °C, followed by addition of 5 µmol of AtsT and /or BCAM0378. Reactions were terminated by adding 3x SDS-PAGE loading buffer after 15 min. Samples were resolved on 16% SDS-PAGE gel and stained with Coomassie Blue (Bottom). Phosphorylated proteins were visualized by a phosphorimager (Top). Phosphorylated bands corresponding to the expected mass of AtsR, AtsR-HK and AtsT polypeptides are indicated by arrows.

transduction from a sensor kinase to a response regulator, and BCAM0378 is not involved in the transfer of the phosphate from AtsR to the AtsT under the conditions tested. This conclusion was supported by the lack of the "pearls on a string" phenotype in macrophages infected with a Δ*Bcam0378* mutant (Fig. 2.1A), strongly suggesting that BCAM0378 is not part of the AtsR regulatory pathway.

2.3.5 Kinetics of AtsR phosphorylation

A pulse-chase experiment was performed to determine the stability of AtsR phosphorylation in the presence of AtsT and to determine whether phosphorylation of AtsT can take place independently from AtsR in excess of ATP. If AtsT is phosphorylated from the free nucleotide pool, then an excess of unlabeled ATP should compete with hot ATP for AtsT phosphorylation in a pulse-chase reaction. Conversely, if AtsR phosphorylates AtsT, then the label should be chased from AtsR to AtsT. AtsR was incubated with [γ-^{33}P]-ATP for 10 min, and an excess of unlabeled ATP (20 mM) and AtsT were simultaneously added to the reaction. Labelling of AtsT coincided with immediate loss of signal from phosphorylated AtsR (Fig. 2.4A). A decrease of phosphorylation of nearly 50% was observed between time zero (pre-incubation) and 2 min (Fig. 2.4C). Phosphorylated AtsR was highly stable, and the signal remained relatively strong at 60 min. We repeated this experiment and chased phosphorylated proteins for up to 80 min, and phosphorylated bands were still visible (data not shown). To control the quality of unlabeled ATP, AtsR was incubated simultaneously with both labeled and unlabeled ATP for 10 min, followed by the addition of AtsT (Fig. 2.4B). No labelling of AtsR or AtsT was detected in the control reaction indicating that incorporation of γ-^{33}P can be inhibited by an excess of unlabeled ATP, which competes

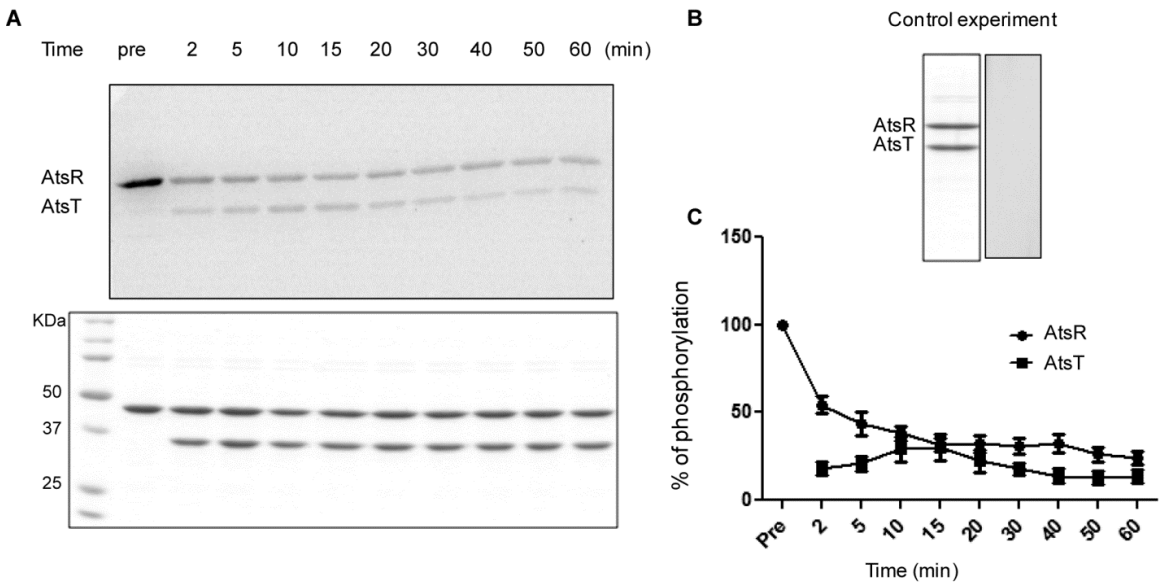


Figure 2.4: Kinetics of phosphotransfer from AtsR to AtsT. A) Five µmol of AtsR was pre-incubated with 5 µCi [γ-^{33}P]-ATP in a standard phosphorylation mixture for 15 min and then 5 µmol of AtsT and 20 mM ATP were simultaneously added to the reaction. The reaction was chased over time at 25 °C. Aliquots were removed before and after the chase at the times indicated. Reactions were terminated by adding SDS-PAGE loading buffer. Samples were run on 16% SDS-PAGE gel and stained with Coomassie Blue (bottom). Phosphorylated proteins were visualized by a phosphorimager (top). The images shown here are the representatives of two independent repeats. B) Five µmol of AtsR was incubated simultaneously with both labeled and unlabeled ATP for 10 min, followed by the addition of 5 µmol of AtsT and incubated for 15 min. Samples were run on 16% SDS-PAGE gels and stained with Coomassie Blue (left) or visualized by a phosphorimager (right). C) The y-axis represents the percentage of normalized absorbance of densitometry readings from bands corresponding to phosphorylated proteins obtained from two independent experiments.

with labeled ATP in the reaction. Thus, AtsR phosphorylation at His-245 is stable, and AtsT phosphorylation is resistant to competition from excess ATP, which confirms that AtsT acquires its phosphate directly from AtsR.

2.3.6 The role of Asp-536 on AtsR-AtsT phosphorylation kinetics

To determine the effect of Asp-536 on phosphotransfer, AtsR and $AtsR_{D536A}$ were incubated with [γ-^{33}P]-ATP for 10 min in individual reactions to make a pool of phosphorylated proteins, followed by the addition of AtsT to each reaction. The rate of phosphate incorporation to the RR was followed as a function of time. The D536A replacement did not abolish the ability of AtsR to autophosphorylate, but the phosphotransfer capabilities of the parental and $AtsR_{D536A}$ proteins were quite different. The transfer of phosphate from AtsR to AtsT took place not only slower but also at a lower rate than from $AtsR_{D536A}$ to AtsT. Indeed, more than 90% of the phosphate was transferred from $AtsR_{D536A}$ to AtsT after 90 sec, and the amount of label gained by AtsT was almost equal to that lost by $AtsR_{D536A}$ (Fig. 2.5A and B). The increase in AtsT phosphorylation correlates with the disappearance of the AtsR and $AtsR_{D536A}$ phosphorylation signals. In contrast to native AtsR, where the phosphorylation signal was maintained from 2 min until the end of the experiment, phosphorylation of AtsT by $AtsR_{D536A}$ was rapidly detected 1 min after addition of RR to the reaction and the signal of $AtsR_{D536A}$ had already disappeared after 2 min under the conditions tested (Fig. 2.5B). Together, these results indicate that replacing Asp-536 with alanine increases and accelerates the phosphotransfer reaction to the RR. It also might be due to the fact that AtsR has two sites of phosphorylation and only His-245 transfers phosphate to AtsT,

therefore, the persistence of the signal on wild-type protein is the result of Asp-536 phosphorylation.

2.3.7 *In vivo* reconstitution of the AtsR signalling pathway

B. cenocepacia secretes two zinc metalloproteases ZmpA and ZmpB (17-21). Deletion of *atsR* upregulates the expression of several quorum sensing regulated virulence factors, including swarming motility and the secretion of zinc metalloproteases, both of which are also positively regulated by the CepI/CepR quorum sensing system (16). Moreover, whereas a Δ*cepI* mutant was protease deficient at 48 h when spotted on D-BHI milk agar plates, further deletion of *atsR* in Δ*cepI* resulted in increased proteolytic activity demonstrating that deletion of *atsR* upregulates protease activity independently of quorum sensing (16). We took advantage of the protease and swarming motility phenotypes to investigate whether AtsR phosphorylation is required for its function *in vivo*. The mutant strains Δ*atsR* and Δ*atsR*Δ*cepI* were complemented at the chromosomal level with full-length *atsR* or *atsR* variants encoding either $AtsR_{H245A}$, $AtsR_{D536A}$ or AtsR lacking the RD (AtsRΔRD) eliminating the putative effect of other conserved Asp residues adjacent to Asp-536. The ability of $AtsR_{H245A}$, $AtsR_{D536A}$ and AtsRΔRD to suppress the phenotypes of Δ*atsR* and Δ*atsR*Δ*cepI* mutant backgrounds was first assessed using protease secretion assay (Fig. 2.6A). As expected, the radius of halo corresponding to casein degradation surrounding the Δ*atsR* spot was bigger compared to that of the wild-type within 48 h of incubation while no halo was present for Δ*cepI* (Fig. 2.6A and B). Proteolytic activity was also detectable for Δ*atsR*Δ*cepI* to a lesser extent compared to WT and Δ*atsR* (Fig. 2.6A, top row, and Fig. 2.6B). Successful complementation was achieved when a chromosomal copy of the *atsR* gene was restored intoΔ*atsR* and in Δ*atsR*Δ*cepI* leading to

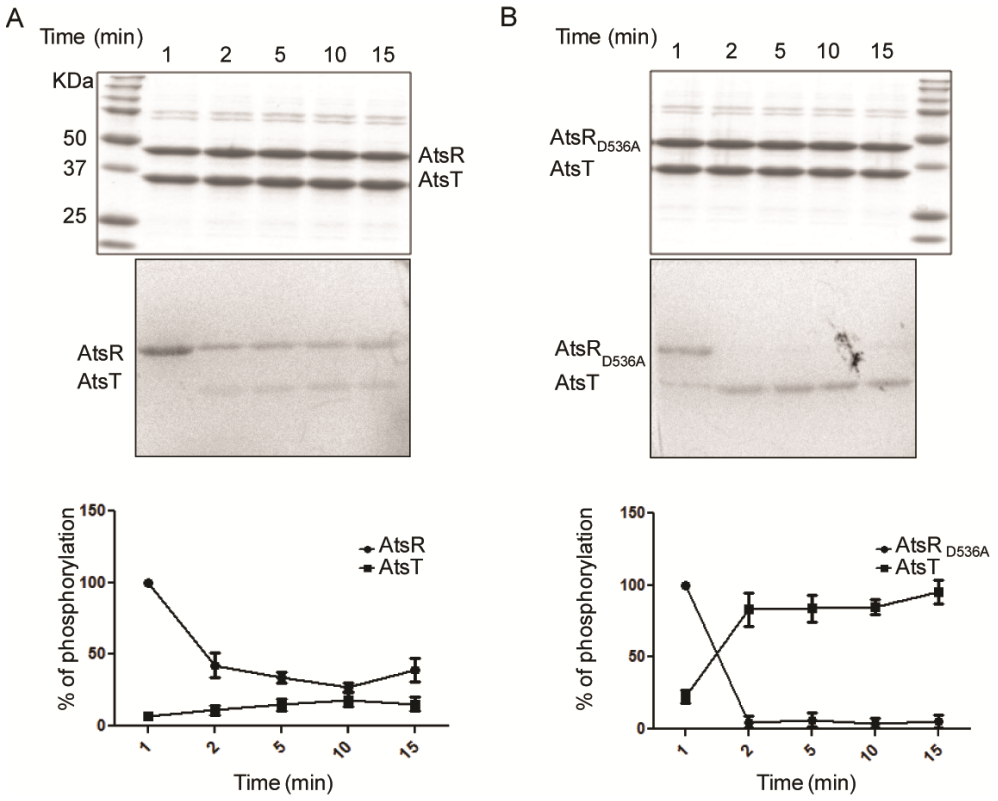


Figure 2.5: Kinetics of phosphotransfer from AtsR and $AtsR_{D536A}$ to AtsT. After 10 min of pre-incubation of 5 µmol AtsR (A) or $AtsR_{D536A}$ (B) in a standard phosphorylation mixture, AtsT was added to the reaction and aliquots were removed at the times indicated. The reaction was performed at 25 °C and terminated by adding SDS-PAGE loading buffer. Samples were run on 16% SDS-PAGE gel and stained with Coomassie Blue (top). Phosphorylated proteins were visualized with a phosphorimager (middle). The images shown here are the representatives of two independent repeats. The y-axis represents the percentage of normalized absorbance of densitometry readings from bands corresponding to phosphorylated proteins obtained from two independent experiments (bottom).

decreased protease activity to WT and Δ*cepI* levels, respectively (Fig. 2.6A, middle and bottom row, and Fig. 2.6B).

Since Asp-536 significantly increases the phosphotransfer from His-245 to AtsT *in vitro*, one could have expected that complementation of Δ*atsR* with $AtsR_{D536A}$ would have a stronger inhibitory effect than *atsR in vivo*. However, similar results were obtained when either *atsR*Δ*RD* or $atsR_{D536A}$ was introduced to Δ*atsR* or Δ*atsR*Δ*cepI*, respectively. This suggests that Asp-D536 does not have a strong modulatory effect on the AtsR-AtsT phosphotransfer *in vivo* under the conditions tested, or alternatively, that additional components not identified in this study may be involved in modulating the AtsR-AtsT phosphorelay pathway.

In contrast, the phosphorylation status of AtsR is critical for its role in expression and/or secretion of proteases, as complementation failed when $atsR_{H245A}$ was introduced. These strains were also tested in a swarming motility assay. In agreement with previous results, only complementation of Δ*atsR*Δ*cepI* strains with *atsR* or *atsR*Δ*RD,* but not $atsR_{H245A}$, could restore swarming motility to Δ*cepI* levels (Fig. 2.7). Protein expression was confirmed by Western blot analysis of bacterial membrane pellets prepared from K56-2 Δ*atsR* and Δ*atsR*Δ*cepI* complemented with $atsR_{H245A}$, $atsR_{D536A}$, *atsR*Δ*RD* and *atsR* and demonstrated that the encoded AtsR variants were similarly expressed (Fig. 2.6C). Thus, a lack of complementation by $atsR_{H245A}$ was not due to a defect in protein expression. Together, these data confirm the *in vitro* results and suggest that His-245 is essential for the function of AtsR *in vivo*.

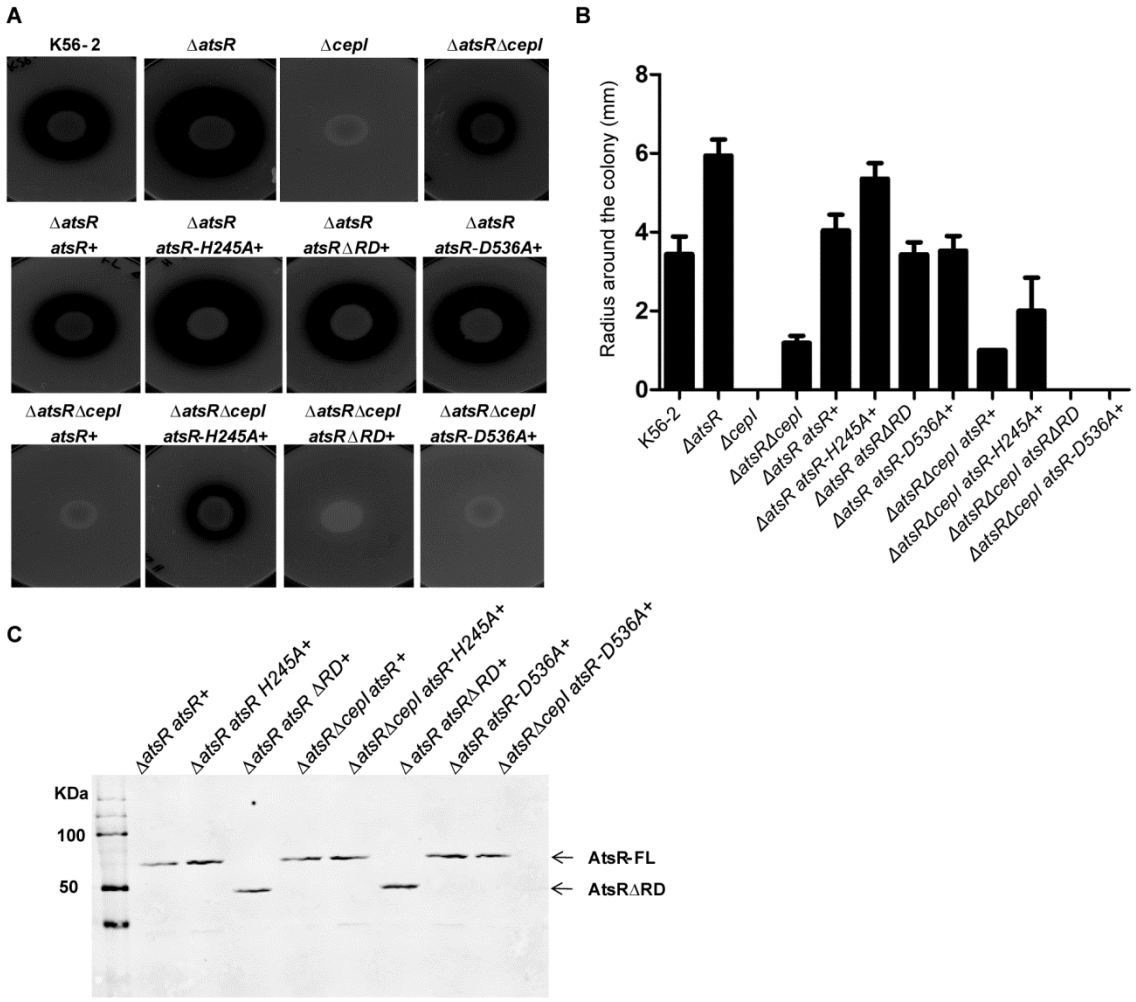


Figure 2.6: Proteolytic activity of *B. cenocepacia* K56-2 wild-type, Δ*atsR*, Δ*atsR*Δ*cepI* and Δ*cepI* mutants and complemented mutants at the chromosomal level in different genetic backgrounds. A) Proteolysis was tested on D-BHI milk agar plates. The plates shown are representatives of three experiments performed in triplicate. Zones of clearing around the colonies were measured at 48 h of incubation at 37 °C. B) Values are average radius ± S.D. in millimeters of three experiments performed in triplicate. C) Anti-His Western blot analysis of His-tag purified AtsR, AtsRΔRD, $AtsR_{D536A}$ and $AtsR_{H245A}$ in *B. cenocepacia*. Arrows indicate the positions of full-length AtsR (AtsR, $AtsR_{D536A}$ and $AtsR_{H245A}$) and AtsRΔRD. AtsR-FL: full-length AtsR.

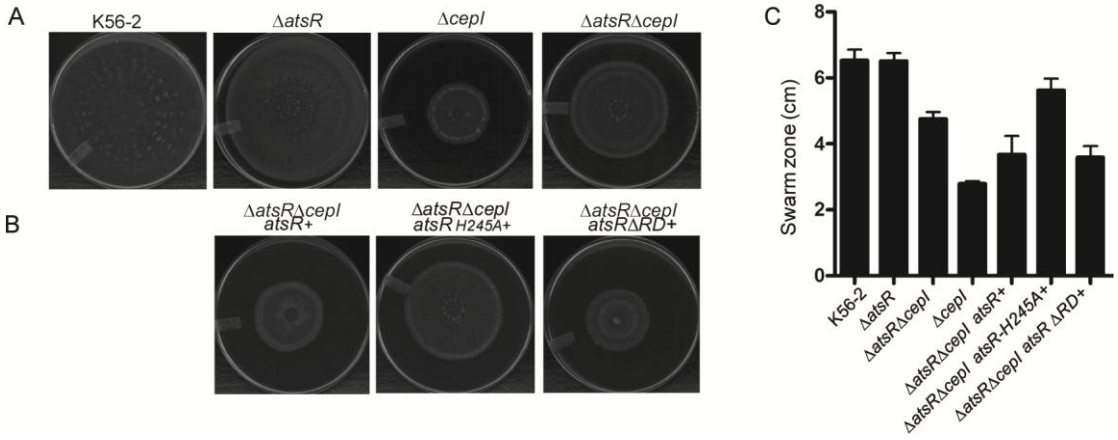


Figure 2.7: Swarming motility. *B. cenocepacia* K56-2 wild type and derivative mutants (A) and Δ*atsR*Δ*cepI* mutant complemented by the integration of *atsR*, *atsR-H245A*, and *atsR*Δ*RD* at the chromosomal level (B) were tested for swarming motility. The plates are representatives of at least three experiments performed in triplicate. The extent of the swarm zone was measured, and error bars represent the S.D. (C).

To confirm that AtsT is the cognate response regulator, which negatively controls the expression of protease activity, *atsT* was deleted in Δ*atsR,* Δ*atsR*Δ*cepI* and Δ*cepI* strains which were then complemented with $atsR_{D536A}$, *atsR*Δ*RD, atsT* or $atsT_{D208A}$. The resulting strains were spotted on D-BHI milk agar plates (Fig. 2.8A & B) and the proteolysis quantified by measuring the radius of clearing around the colonies (Fig. 2.8C). While complementation of Δ*atsR* and Δ*atsR*Δ*cepI* strains with $atsR_{D536A}$ or *atsR*Δ*RD* reduced the proteolytic activity to WT and Δ*cepI* levels (Fig. 2.6), further deletion of *atsT* bypassed the $atsR_{D536A}$ or *atsR*Δ*RD* complementation and resulted in an increase of proteolytic activity. Furthermore, Δ*cepI*Δ*atsT* has the same phenotype as Δ*atsR*Δ*cepI* and as expected, by complementing Δ*cepI*Δ*atsT* with *atsT*, proteolytic activity decreases to Δ*cepI* levels whereas, $atsT_{D208A}$ is unable to complement. These results suggest that (i) AtsT indeed acts as a negative regulator, and (ii) is a direct target of AtsR contributing to the regulatory role of this protein on proteolytic activity and (iii) Asp-208 on AtsT is required for its function.

2.4 Conclusion

The predicted structural features of AtsR suggested this protein could not directly control gene expression because it lacks an effector domain. We demonstrate here three key properties of AtsR: (i) upon autophosphorylation AtsR transfers the phosphate to the response regulator AtsT without the participation of an intermediate histidine-phosphotransfer protein; (ii) AtsR function *in vitro* and *in vivo* depends on autophosphorylation of the His-245 residue, which is absolutely essential for initiation of signal transduction; (iii) the AtsR-RD and more specifically the Asp-536 to some extent plays a role in modulating the stability of phosphorylated AtsR.

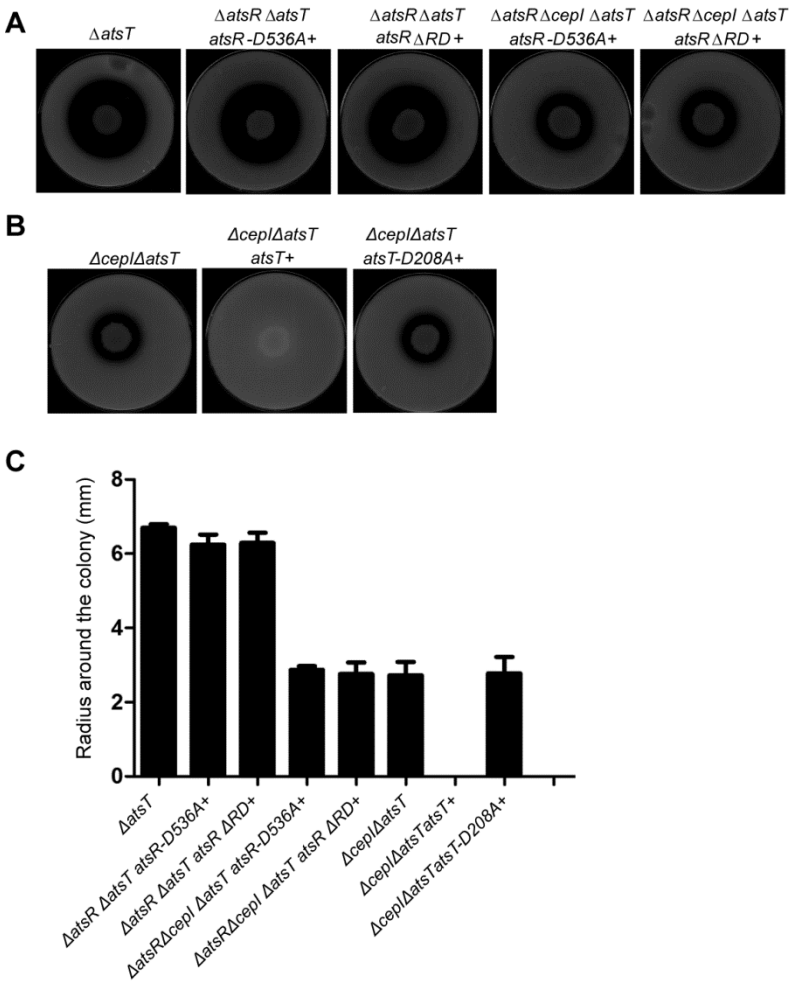


Figure 2.8: Proteolytic activity of *B. cenocepacia* Δ*atsT*, Δ*atsR*Δ*atsT atsR*Δ*RD+,* Δ*atsR*Δ*cepI*Δ*atsT atsR*Δ*RD+,* Δ*atsR*Δ*atsT atsR*$_{D536A}$+ and Δ*atsR*Δ*cepI*Δ*atsT atsR*$_{D536A}$+. A) *atsT* was deleted from *B. cenocepacia* K56-2Δ*atsR atsR*Δ*RD+,* Δ*atsR*Δ*cepI atsR*Δ*RD+,* Δ*atsR atsR*$_{D536A}$+ and Δ*atsR*Δ*cepI atsR*$_{D536A}$+ backgrounds. Mutants were tested on D-BHI milk agar plates. Zones of clearing around the colonies were measured at 48 h of incubation at 37 °C. B) *atsT* was deleted from *B. cenocepacia* K56-2Δ*cepI* background and the resulting strain was complemented with either *atsT* or *atsT*$_{D208A}$ at chromosomal level. Mutants were tested on D-BHI milk agar plates. C) Values are average radius in millimetres of three experiments performed in triplicates. The plates shown are representatives of three experiments performed in triplicate.

Other studies have shown that the sequential phosphotransfer between residues within the same hybrid sensor kinase modulates the phosphotransfer to the cytosolic response regulator and the overall response by determining the specificity for the cognate RR or by regulating the autokinase activity (2, 3, 39-42). The stability of the phosphorylated AtsR and AtsT determined by the pulse-chase kinetic experiments is remarkable as these proteins can maintain the response memory for at least 60 min, which agrees with their physiological function as negative regulators of gene expression. Comparative studies have suggested that the *in vitro* assay of HK autophosphorylation reflects, to a certain extent, the *in vivo* situation (43). For example, ArcB, NarQ, and NtrB with a high rate of phosphotransfer but a low level of autophosphorylation are able to respond quickly to changes in the environment thereby returning to the steady-state levels after transient activation or repression by external stresses, whereas, CheA and BaeS showed high levels of self-phosphorylation even though they had a slow signal transduction rate (43). In the case of AtsR, we find that the phosphate is transferred very quickly within 2 min from AtsR to AtsT, and both proteins are capable of prolonged phosphorylation. This may indicate that unlike some other systems (44, 45), AtsR might not have phosphatase activity although further experiments are needed to validate this notion. Based on the consistency of results obtained by *in vitro* and *in vivo* experimental approaches in this study, we propose a model (Fig. 2.9) in which AtsR autophosphorylates *in vivo* at His-245. The phosphate is then transferred to AtsT and partly to the AtsR receiver domain on Asp-536, as stable phosphorylation of the AtsR receiver domain was detected *in vitro*. Our data strongly support the notion that AtsR phosphorylation has a significant biological relevance as a global virulence regulator modulating the expression of proteases through AtsT.

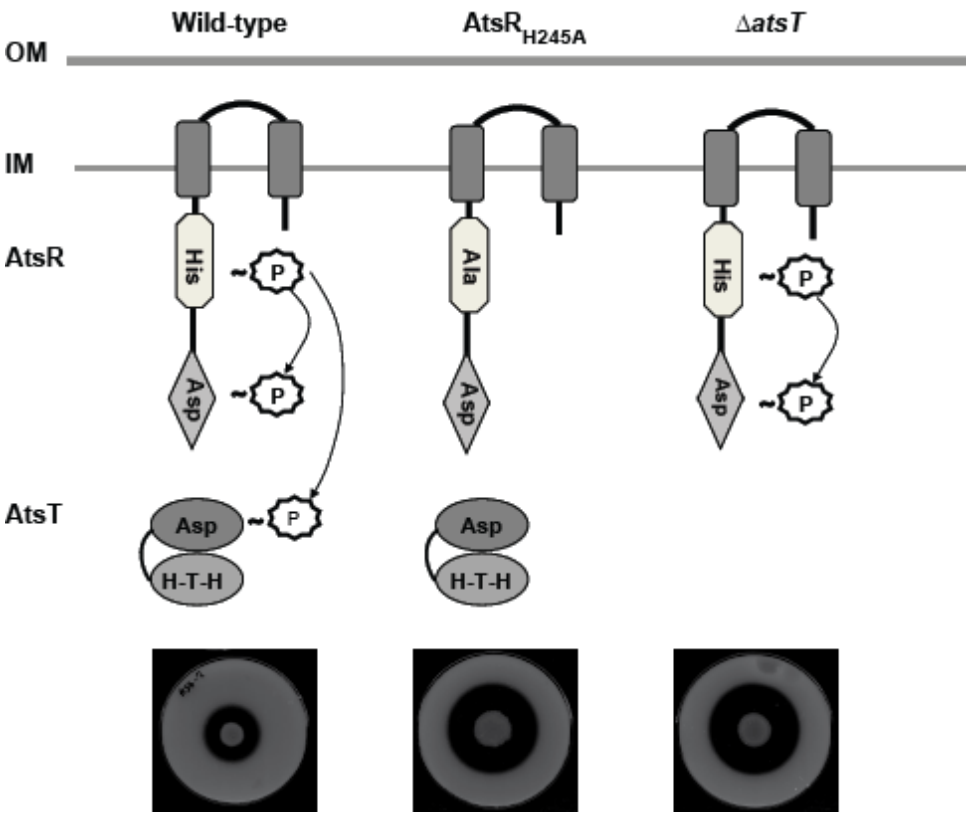


Figure 2.9: Diagram indicating the regulatory relationship between AtsR and corresponding response regulator and its effect on proteolytic activity. OM, outer membrane; IM, inner membrane; His, histidine; Ala, alanine; Asp, aspartic acid; HTH, helix-turn-helix; P, phosphate.

2.5 Chapter 2 references

1. **Wolanin PM, Thomason PA, Stock JB.** 2002. Histidine protein kinases: key signal transducers outside the animal kingdom. Genome Biol **3:**REVIEWS3013.

2. **Krell T, Lacal J, Busch A, Silva-Jimenez H, Guazzaroni ME, Ramos JL.** 2010. Bacterial sensor kinases: diversity in the recognition of environmental signals. Annu Rev Microbiol **64:**539-559.

3. **Georgellis D, Lynch AS, Lin EC.** 1997. In vitro phosphorylation study of the arc two-component signal transduction system of *Escherichia coli*. J Bacteriol **179:**5429-5435.

4. **Uhl MA, Miller JF.** 1994. Autophosphorylation and phosphotransfer in the *Bordetella pertussis* BvgAS signal transduction cascade. Proc Natl Acad Sci U S A **91:**1163-1167.

5. **Mahenthiralingam E, Baldwin A, Dowson CG.** 2008. *Burkholderia cepacia* complex bacteria: opportunistic pathogens with important natural biology. J Appl Microbiol **104:**1539-1551.

6. **Vandamme P, Dawyndt P.** 2011. Classification and identification of the *Burkholderia cepacia* complex: Past, present and future. Syst Appl Microbiol **34:**87-95.

7. **Bevivino A, Costa B, Cantale C, Cesarini S, Chiarini L, Tabacchioni S, Caballero-Mellado J, Dalmastri C.** 2011. Genetic relationships among Italian and Mexican maize-rhizosphere *Burkholderia cepacia* complex (BCC) populations belonging to *Burkholderia cenocepacia* IIIB and BCC6 group. BMC Microbiol **11:**228.

8. **McNeely D, Moore JE, Elborn JS, Millar BC, Rendall J, Dooley JS.** 2009. Isolation of *Burkholderia cenocepacia* and *Burkholderia vietnamiensis* from human sewage. Int J Environ Health Res **19:**157-162.

9. **Lee YA, Chan CW.** 2007. Molecular Typing and Presence of Genetic Markers Among Strains of Banana Finger-Tip Rot Pathogen, *Burkholderia cenocepacia*, in Taiwan. Phytopathology **97:**195-201.

10. **Sotokawa N, and Y. Takikawa.** 2004. Occurrence of bacterial rot of onion bulbs caused by Burkholderia cepacia in Japan. J Gen Plant Pathol **70:**348-352.

11. **Vergunst AC, Meijer AH, Renshaw SA, O'Callaghan D.** 2010. *Burkholderia cenocepacia* creates an intramacrophage replication niche in zebrafish embryos, followed by bacterial dissemination and establishment of systemic infection. Infect Immun **78:**1495-1508.

12. **Uehlinger S, Schwager S, Bernier SP, Riedel K, Nguyen DT, Sokol PA, Eberl L.** 2009. Identification of specific and universal virulence factors in *Burkholderia cenocepacia* strains by using multiple infection hosts. Infect Immun **77:**4102-4110.

13. **Lamothe J, Huynh KK, Grinstein S, Valvano MA.** 2007. Intracellular survival of *Burkholderia cenocepacia* in macrophages is associated with a delay in the maturation of bacteria-containing vacuoles. Cell Microbiol **9:**40-53.

14. **Lamothe J, Thyssen S, Valvano MA.** 2004. *Burkholderia cepacia* complex isolates survive intracellularly without replication within acidic vacuoles of *Acanthamoeba polyphaga*. Cell Microbiol **6:**1127-1138.

15. **Aubert DF, Flannagan RS, Valvano MA.** 2008. A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in *Burkholderia cenocepacia*. Infect Immun **76:**1979-1991.

16. **Aubert DF, O'Grady EP, Hamad MA, Sokol PA, Valvano MA.** 2013. The *Burkholderia cenocepacia* sensor kinase hybrid AtsR is a global regulator modulating quorum-sensing signalling. Environ Microbiol **15:**372-385.

17. **Baldwin A, Sokol PA, Parkhill J, Mahenthiralingam E.** 2004. The *Burkholderia cepacia* epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*. Infect Immun **72:**1537-1547.

18. **Kooi C, Subsin B, Chen R, Pohorelic B, Sokol PA.** 2006. *Burkholderia cenocepacia* ZmpB is a broad-specificity zinc metalloprotease involved in virulence. Infect Immun **74:**4083-4093.

19. **Malott RJ, Baldwin A, Mahenthiralingam E, Sokol PA.** 2005. Characterization of the cciIR quorum-sensing system in *Burkholderia cenocepacia*. Infect Immun **73:**4982-4992.

20. **Malott RJ, O'Grady EP, Toller J, Inhulsen S, Eberl L, Sokol PA.** 2009. A *Burkholderia cenocepacia* orphan LuxR homolog is involved in quorum-sensing regulation. J Bacteriol **191:**2447-2460.

21. **Sokol PA, Sajjan U, Visser MB, Gingues S, Forstner J, Kooi C.** 2003. The CepIR quorum-sensing system contributes to the virulence of *Burkholderia cenocepacia* respiratory infections. Microbiology **149:**3649-3658.

22. **Rosales-Reyes R, Skeldon AM, Aubert DF, Valvano MA.** 2012. The Type VI secretion system of *Burkholderia cenocepacia* affects multiple Rho family GTPases disrupting the actin cytoskeleton and the assembly of NADPH oxidase complex in macrophages. *Cell Microbiol* **14:**255-273.

23. **Galperin MY.** 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. J Bacteriol **188:**4169-4182.

24. **Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A.** 2010. The Pfam protein families database. Nucleic Acids Res **38:**D211-222.

25. **Aubert D, MacDonald DK, Valvano MA.** 2010. BcsK(C) Is an Essential Protein for the Type VI Secretion System Activity in *Burkholderia cenocepacia* That Forms an Outer Membrane Complex with BcsL(B). Journal of Biological Chemistry **285:**35988-35998.

26. **Hamad MA, Skeldon AM, Valvano MA.** 2010. Construction of aminoglycoside-sensitive *Burkholderia cenocepacia* strains for use in studies of intracellular bacteria with the gentamicin protection assay. Appl Environ Microbiol **76:**3170-3176.

27. **Hamad MA, Di Lorenzo F, Molinaro A, Valvano MA.** 2012. Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. Mol Microbiol **85:**962-974.

28. **Figurski DH, Helinski DR.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci U S A **76:**1648-1652.

29. **Sambrook J, Russell DW.** 2006. The condensed protocols from Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

30. **Cohen SN, Chang AC, Hsu L.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc Natl Acad Sci U S A **69:**2110-2114.

31. **Craig FF, Coote JG, Parton R, Freer JH, Gilmour NJ.** 1989. A plasmid which can be transferred between *Escherichia coli* and *Pasteurella haemolytica* by electroporation and conjugation. J Gen Microbiol **135:**2885-2890.

32. **Flannagan RS, Linn T, Valvano MA.** 2008. A system for the construction of targeted unmarked gene deletions in the genus *Burkholderia*. Environ Microbiol **10:**1652-1660.

33. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-685.

34. **Perez JM, McGarry MA, Marolda CL, Valvano MA.** 2008. Functional analysis of the large periplasmic loop of the *Escherichia coli* K-12 WaaL O-antigen ligase. Mol Microbiol **70:**1424-1440.

35. **Goodman AL, Kulasekara B, Rietsch A, Boyd D, Smith RS, Lory S.** 2004. A signaling network reciprocally regulates genes associated with acute infection and chronic persistence in *Pseudomonas aeruginosa*. Dev Cell **7:**745-754.

36. **Laskowski MA, Osborn E, Kazmierczak BI.** 2004. A novel sensor kinase-response regulator hybrid regulates type III secretion and is required for virulence in *Pseudomonas aeruginosa*. Mol Microbiol **54:**1090-1103.

37. **Ventre I, Goodman AL, Vallet-Gely I, Vasseur P, Soscia C, Molin S, Bleves S, Lazdunski A, Lory S, Filloux A.** 2006. Multiple sensors control reciprocal expression of *Pseudomonas aeruginosa* regulatory RNA and virulence genes. Proc Natl Acad Sci U S A **103:**171-176.

38. **Stock JB, Ninfa AJ, Stock AM.** 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol Rev **53:**450-490.

39. **Jourlin C, Ansaldi M, Mejean V.** 1997. Transphosphorylation of the TorR response regulator requires the three phosphorylation sites of the TorS unorthodox sensor in *Escherichia coli*. J Mol Biol **267:**770-777.

40. **Kwon O, Georgellis D, Lin EC.** 2000. Phosphorelay as the sole physiological route of signal transmission by the arc two-component system of *Escherichia coli*. J Bacteriol **182:**3858-3862.

41. **Perraud AL, Kimmel B, Weiss V, Gross R.** 1998. Specificity of the BvgAS and EvgAS phosphorelay is mediated by the C-terminal HPt domains of the sensor proteins. Mol Microbiol **27:**875-887.

42. **Chang CH, Zhu J, Winans SC.** 1996. Pleiotropic phenotypes caused by genetic ablation of the receiver module of the *Agrobacterium tumefaciens* VirA protein. J Bacteriol **178:**4710-4716.

43. **Yamamoto K, Hirao K, Oshima T, Aiba H, Utsumi R, Ishihama A.** 2005. Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. J Biol Chem **280:**1448-1456.

44. **Castelli ME, Garcia Vescovi E, Soncini FC.** 2000. The phosphatase activity is the target for Mg2+ regulation of the sensor protein PhoQ in *Salmonella*. J Biol Chem **275:**22948-22954.

45. **Hsing W, Silhavy TJ.** 1997. Function of conserved histidine-243 in phosphatase activity of EnvZ, the sensor for porin osmoregulation in *Escherichia coli*. J Bacteriol **179:**3729-3735.

Chapter 3

Comprehensive whole genome mapping and characterization of AtsR/AtsT regulatory network in *Burkholderia cenocepacia*

A version of this chapter is being prepared for submission:

Maryam Khodai-Kalaki, Daniel Aubert, Miguel A. Valvano. Comprehensive whole genome mapping and characterization of AtsR/AtsT regulatory network in *Burkholderia cenocepacia*.

3.1 Introduction

Burkholderia cenocepacia is a prevalent member of the *B. cepacia* complex (Bcc), a group of closely related *Burkholderia* species that establish chronic lung infections in cystic fibrosis (CF) patients (1). Infection by *B. cenocepacia* can be transmitted from patient to patient and may lead to "cepacia syndrome", a fatal form of necrotizing pneumonia (2). Experimentally, *B. cenocepacia* was shown to be virulent in several infection models including zebrafish (3), *Arabidopsis* (4), alfalfa, nematodes (5), and insect infection models (6), and can survive within amoebae and macrophages (7, 8). Survival in macrophages was also demonstrated in the lungs of infected patients undergoing transplant (9). This wide range of hosts and survival capabilities underscore the high adaptability of *B. cenocepacia* to establish infection in different niches, which may in part depend on regulatory networks allowing this bacterium to survive in different environments. Two-component system (TCS) regulators play a major role in regulation, as they allow coordinated expression of complex networks *via* changes in the expression of multiple genes in response to environmental cues (10).

We have previously identified AtsR (Adhesion and Type 6 secretion system Regulator) as a global virulence regulator in *B. cenocepacia* that modulates quorum sensing (QS) signalling and consequently the expression of QS-regulated virulence factors including biofilm formation and expression of a Type 6 secretion system (T6SS) (11, 12). I recently showed AtsR is a hybrid sensor kinase that regulates downstream cellular activities through direct phosphorylation of AtsT. Autophosphorylation of a conserved histidine initiates a signal transduction phosphorelay into its cognate response regulator AtsT (13). AtsT is a cytoplasmic transcriptional regulator containing an N-terminal helix-turn-helix

domain related to domains found in transcriptional repressors (14) and a C-terminal receiver domain with a conserved aspartate at position 208 at the C-terminus (13). Deletion of either *atsR or atsT* in *B. cenocepacia* K56-2 or amino acid replacements blocking phosphotransfer from AtsR to AtsT lead to increased extracellular proteolytic activity, attributed to a significant increase in the expression of metalloproteases, and swarming motility, in addition to overexpression of the T6SS (13). This indicated AtsT is regulated directly by AtsR via phosphorylation and consequently, controls the expression of the downstream components in the AtsR/AtsT signalling pathway (13).

In this work, I compare transcriptional responses of *atsR* mutant with AtsT genomic binding-location data to identify direct and indirect regulatory targets of the AtsR/AtsT signalling pathway, which is a key step to elucidating this regulatory network in *B. cenocepacia*. Expression profiles of *B. cenocepacia* lacking functional AtsR have been used to define AtsR-regulated genes. Since transcriptional analysis alone cannot distinguish between direct and indirect regulatory effects, we employed chromatin immunoprecipitation (ChIP) followed by sequence identification of the DNA fragments associated with the precipitated AtsT as the transcriptional regulator. This is a method of choice for mapping protein-DNA interactions on a genome scale in eukaryotes and has been recently utilized for mapping the genomic binding sites of bacterial transcriptional regulators (15, 16).

To determine the binding locations of AtsT in *B. cenocepacia,* bioinformatic analyses of sequences corresponding to AtsT-enriched regions were used to predict individual binding sites and to search for a binding motif. The AtsT binding sites were then validated by DNase I footprinting. Additionally, gene expression profiling was performed

in the Δ*atsR* background. This combination of genome-wide approaches allowed us to identify additional operons under direct AtsT control, thereby providing a more complete understanding of the functional role of AtsR/AtsT signalling pathway in *B. cenocepacia*.

3.2 Materials and methods

3.2.1 Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 3.1. Bacteria were grown in LB at 37 °C unless indicated otherwise. *Escherichia coli* cultures were supplemented, as required, with the following antibiotics (final concentrations): tetracycline 30 μg/ml, kanamycin 30 μg/ml and trimethoprim 50 μg/ml. *B. cenocepacia* cultures were supplemented, as required, with trimethoprim 100 μg/ml, tetracycline 150 μg/ml, ampicillin 200 μg/ml and polymyxin B 20 μg/ml.

3.2.2 General molecular techniques

DNA manipulations were performed as described previously (21). T4 DNA ligase (Roche Diagnostics, Laval, Quebec, Canada) and Antarctic phosphatase (New England Biolabs, Pickering, Ontario, Canada) were used as recommended by the manufacturers. Transformation of *E. coli* DH5α and *E. coli* GT115 was done by the calcium chloride method (22). Mobilization of complementing plasmids into *B. cenocepacia* K56-2 was performed by triparental mating using *E. coli* DH5α carrying the helper plasmid pRK2013 (23). DNA amplification by PCR was performed using Bio-Rad C1000^{TM} Thermal Cycler with Taq or HotStar HiFidelity DNA polymerases (Qiagen, Canada). DNA sequences of all primers used in this study are described in Table 3.2. DNA sequencing was performed at the DNA sequencing Facility of York University, Toronto,

Table 3.1: List of strains and plasmids

Strain or Plasmid	Relevant Properties	Source / Reference
Strains		
E. coli		
DH5α	F⁻ ϕ80*lacZ*ΔM15 *endA recA hsdR*(r_K^- m_K^-) *nupG thi glnV deoR gyrA relA1 Δ*(*lacZYA-argF*)*U169*	Laboratory stock
GT115	F⁻ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80Δ*lacZ*ΔM15 Δ*lacX74 recA1 rpsL endA1*Δ*dcm uidA*(ΔMluI)::*pir-116* Δ*sbcC-sbcD*	Laboratory stock
Plasmids		
PGSVTP-*lux*	Mobilizable suicide vector containing the *lux* operon, OriT; TpR	(17)
PDA154	*pBcas0627*::*luxCDABE* transcriptional fusion in pGSVTp-*lux*	This study
pET28a(+)	Cloning vector, IPTG inducible for N-terminal 6X His fusion, Kan^r	Laboratory stock
pMH447	pGPI-SceI derivative used for chromosomal complementation	(18)
pRK2013	$ori_{colE1,}$ RK2 derivative, Kan^R, mob^+, tra^+	(19)
pDAI-SceI-SacB	ori_{pBBR1}, Tet^R, P_{dhfr}, mob^+, expressing I-SceI, SacB	(20)
pMK3	pET28a(+) expressing AtsT-His, kan^R	(13)
pMZ24	pET28a(+) encoding *B. cenocepacia* $atsR_{205-605}$, 6X His, kan^R	(13)
Strains		
B. cenocepacia		
K56-2	ET12 clone related to J2315, CF clinical isolate	[b]BCRRC
K56-2Δ*atsR*	Deletion of *atsR* in K56-2	(13)
K56-2Δ*atsR::atsR*	Chromosomal *atsR* integration in K56-2 Δ*atsR*	(13)
K56-2 Δ*atsR*Δ*cepI*	Deletion of *cepI* in K56-2 Δ*atsR*	This study
K56-2 Δ*atsR*Δ*cepI::atsR*	Chromosomal *atsR* integration in K56-2 Δ*atsR*Δ*cepI*	This study
K56-2 Δ*atsT*	Deletion of *atsR* in K56-2	This study
K56-2 Δ*atsT::atsT*	Chromosomal *atsT* integration in K56-2 Δ*atsT*	This study

K56-2 Δ*atsT*Δ*cepI*	Deletion of *cepI* in K56-2 Δ*atsT*	This study
K56-2 Δ*atsT*Δ*cepI::atsT*	Chromosomal *atsT* integration in K56-2 Δ*atsR*Δ*cepI*	This study
K56-2::PDA154	K56-2, *pBcas0627*::pGSVTp-*luxCDABE*; TpR	This study
PDA35::PDA154	K56-2Δ*atsR*, *pBcas0627*::pGSVTp-*luxCDABE*	This study
DFA155	K56-2Δ*atsR::atsR*, *pBcas0627*::pGSVTp-*luxCDABE*	This study
DFA156	K56-2Δ*atsR*Δ*cepI*, *pBcas0627*::pGSVTp-*luxCDABE*	This study
DFA157	K56-2Δ*atsR*Δ*cepI::atsR*, *pBcas0627*::pGSVTp-*luxCDABE*	This study
MK111	K56-2Δ*atsT*, *pBcas0627*::pGSVTp-*luxCDABE*	This study
MK112	K56-2Δ*atsT::atsT*, *pBcas0627*::pGSVTp-*luxCDABE*	This study
MK113	K56-2Δ*atsT*Δ*cepI*, *pBcas0627*::pGSVTp-*luxCDABE*	This study
MK114	K56-2Δ*atsT*Δ*cepI::atsT*, *pBcas0627*::pGSVTp-*luxCDABE*	This study

[a] Tp^R, trimethoprim resistance, Kan^R, kanamycin resistance, Tet^R, tetracycline resistance.
[b] BCRRC, *B. cepacia* Research and Referral Repository for Canadian CF Clinics.

Table 3.2: List of primers

Primer No.	Primer Name	Primer Sequence
5195	0627prom-For-Eco	TTTTGAATTCGTCGCTTTGTCTACCACCAG
5196	0627prom-Rev-Eco	TTTTGAATTCGGGTATTTTCCTCACGTTTC
2794	bcam0381-Rev-XbaI	TTTTTCTAGAGTTCGTCGCGGCCGCT
5885	MH447-For	TTGATGGCGAGCGATTCTTC
5886	MH447-Rev	CCAGTTCTTCAGCGTGACGA
7128	Bcal3312 -For	TATAGAATTCCGGAACCCGCGCCAATCTTA
7129	Bcal3312-Rev	TATACCATGGAGCGCGATCAGCCAGTGCAA
7167	Bcas0627- For-IDY682	CGTGCCGTTTTCTGCCGTGTC
7168	Bcam0192- For	CGGAAAGAATATGAGAATGTCCG
7169	Bcam0192-Rev	GGGACTCTCCTAGTAATGTC
7170	Bcam0382- For	CTCGATGGGGTCTCCAGGGG
7171	Bcam0382-Rev-EMSA	GGGTACTTTCCTCAAATTTTTATCGGCG
7204	Bcam0382-For	CTCGATGGGGTCTCCAGGGG
7224	6-FAM-Bcas0627- For	CGTGCCGTTTTCTGCCGTGTC
7225	6-FAM-Bcam0382- For	CTCGATGGGGTCTCCAGGGG
7536	SigE-RT-For	AGGAAACCAACCGTCAGATG
7537	SigE-RT-Rev	GCGACGGTATTCGAACTTGT
7540	MtgA-RT-For	GAAGAGCTACATCCGCAAGG
7541	MtgA-RT-Rev	TCGCGGGAATCTTGTAGTAA
7546	DppA-RT-For	ACTTCACCGACATGGGTCTC
7547	DppA-RT-Rev	GAATTCCATCGCCATGTTCT
7548	FlhD-RT-For	RACCAGCGAAATGCTCAGTG
7549	FlhD-RT-Rev	CATACCCATCGCCTTGTCTT
7550	bcam0382-RT-For	GGAGAAGTGGCTGTCGATC
7551	bcam0382-RT-Rev	CCGAGTACCAGTAGCCGTA
7552	0338-RT-For	ACCAACTACAGCGACGATCC
7553	0338-RT-Rev	ATCAGCTTGACCTGCGTCTT
7554	ArgD-RT-For	GTGATGGTGCCGGTATTTTC
7555	ArgD-RT-Rev	GTCCAGCACCTTCAGCAGTT
7556	Bcal0849-RT-For	CGACATCACCGAGAAGTTCA
7557	Bcal0849-RT-Rev	CCCTTCTGCTTCATCAGGTC
7558	Bcal0511-RT-For	GTTCGATGACGCGAAATTCT
7559	Bcal0511-RT-Rev	GTTGCAGTGGAAGAACACGA
7595	FliC-RT-For	GTTGCACAGCAGAACCTCAA
7596	FliC-RT-Rev	GGTTCAGACCGTTGATCTGG
7597	Bcal0051-RT-For	GAAGCTGTCGTGCAAATGG

7598	Bcal0051-RT-Rev	GCGATAGATGATGGTCGTGA
7599	1293-RT-For	CATCTCCGACTGCAAACCA
7600	1293-RT-Rev	AAGAATCGGTCCATGACAGC
7601	Bcal0340-RT-For	CGACTGCGGAAGTGTTCAA
7602	Bcal0340-RT-Rev	CGTAATGGTTTTGTGCGAACT
7603	Bcas0080-RT-For	ATTCTCGGGTTCACGATGG
7604	Bcas0080-RT-Rev	CGAACAGCAGGTTCAGTCCT
7605	Bcas0034-RT-For	AAGTGGCGTTCTGGGAGAAT
7606	Bcas0034-RT-Rev	CGATGATTTCGCCGTAGAGT
7607	189-RT-For	CGGGTTTTGGTATGACGACT
7608	189-RT-Rev	CTCGATCAGCACGTAATCTCC
7609	Bcas0118-RT-For	GCAGGAACTCGAAGCACAAC
7610	Bcas0118-RT-Rev	GCGAAATTCCCAAAAGCTC
7611	0627-RT-For	GCGACAAGGAATCGAAAAAG
7612	0627-RT-Rev	TACGGATAGCCCGAGTACCA
7613	Bcas0110-RT-For	CGTCACCGATACGAAGATCC
7614	Bcas0110-RT-Rev	TGTCCTTGAAGCTGTCGTTG
7615	Bcal1296-RT-For	GCCGCTCTCAAACATGAAAT
7616	Bcal1296-RT-Rev	CAGGTGTTGAAACCGTGTTG
7617	Bcal1079-RT-For	GGTGTACAGCGACAACCTCA
7618	Bcal1079-RT-Rev	AGTAGTCGGCATTGCTCCAG

Canada. BLAST was used to analyse the sequenced genome of *B. cenocepacia* strain J2315.

3.2.3 Chromatin immunoprecipitation

Despite several attempts for *in vivo* chromatin immunoprecipitation the assay failed (Appendix A) and we developed an *in vitro* assay that captures the AtsT protein bound to its corresponding DNA regions. DNA was isolated from K56-2 WT as described previously (11). Purified AtsT-His or BSA was added to the K56-2 DNA in a reaction buffer containing poly dI-dC, 10 mM Tris (pH 7.5), 120 mM KCl, 20% glycerol, 1 mM $MgCl_2$ and 1 mM $MnCl_2$ and incubated for 30 min at room temperature. Formaldehyde was added to the final concentration of 0.05% for 5 min followed by DNAse I digestion for 1 min. The reaction was incubated at 70 °C for 10 min and the AtsT-DNA complex was pulled down using sepharose beads coated with anti-His antibody. The reaction was incubated with beads for 1 h at 4 °C followed by washing with a buffer containing 100 mM sodium phosphate, 600 mM NaCl and 10 mM imidazole. The AtsT-DNA complex then was eluted using increased gradient concentration of imidazole (120-400 mM). The eluates were incubated at 65 °C overnight to reverse crosslinks. Samples were incubated with 8µL of 10 mg/mL RNaseA for 2 h at 37 °C and then with 4 µL of 20 mg/mL proteinase K at 55 °C for 1h and were purified with Qiagen MinElute Reaction Cleanup Kit and confirmed with PCR using primers to amplify *Bcas0627*. Experiments were performed in triplicate.

3.2.4 Next-Generation Sample Preparation and Sequencing

Ten ng of enriched ChIP DNA from three biological replicates and one input DNA were

submitted to the Sequencing Core Facility in the Robarts Research Institute, University of Western Ontario, Canada, and processed for library construction and Illumina MiSeq 1x50 bp SE sequencing cartridges. Samples were processed by end repairing and the addition of a polyA tail as described by NEBNext ChIP-Seq sample preparation protocols from NewEngland Biolabs. The sequence reads were mapped to the *B. cenocepacia* J2315 genome using CLC genomic workbench software. A read track was created and unmapped reads were deleted. The detection of the enrichment was performed based on the comparison with the control sample to correct for sequencing biases such as genomic regions that are more accessible or repeated regions. Peak shape filter was applied to the experimental data to call peaks and a score was calculated at each genomic position. The score was obtained by extracting the genomic coverage profile of a window centered at the genomic position and then comparing this profile to the peak shape filter. The result of this comparison was defined as peak shape score. Once the peak shape score for the complete genome was calculated, the center of the peak was then identified as the genomic region with the highest peak shape score and the boundaries were determined by the genomic positions where the peak shape score becomes negative. The threshold for reporting peaks was set as Maximum P-value for peak calling 0.001, 0.0005 or 0.0001. Other parameters were set to the most stringent mode possible allowing maximum mismatch of 2, Insertion cost at 3, Deletion cost at 3, Length fraction at 0.5 and Similarity fraction at 0.8. We combined the peaks generated from all three samples and required that a peak be called in at least two of the three experiments to be considered an AtsT binding site. The average fragment length was 126 bp and a window size of 250 bp. The peaks were then annotated with nearby genes including the name of the nearest 5' and 3' genes, and the distance from the nearest upstream or downstream gene.

3.2.5 Motif detection

Sequences associated with peaks were extracted and used for motif analysis using MEME-ChIP (24) with motif occurrence set as ‘‘zero or one per sequence’’, minimum width as 6 and other parameters as default.

3.2.6 RNA isolation and quantitative RT-PCR

K56-2 Δ*atsT* and K56-2 WT cells were grown overnight and then diluted to OD_{600} of 0.05 in 5 ml of LB. Cells were grown at 37 °C to the OD_{600} of 0.7 at 200 rpm then collected by centrifugation at 16000 ×*g* for 1.5 min. RNA was prepared using the Promega SV RNA isolation kit and treated with DNAse I (Qiagen Inc., Mississauga, ON, Canada) following the manufacturer's protocol. cDNA was prepared from 1 µg of RNA using Qiagen cDNA synthesis kit. For ChIP-seq peak validation, relative-abundance quantitative RT-PCR (qRT-PCR) was performed with Wisent Fast SYBR green mix using Applied Biosystems StepOne Plus Real time PCR system. The thermal cycling conditions were 95 °C for 2 min, 40 cycles at 95 °C for 3 sec, 60 °C for 30 sec, and 72 °C for 15 sec. After the PCR a melt curve was obtained by increasing the temperature from 72 °C to 95 °C in 1 °C increments. Each gene was normalized with *RpoE* (*Bcam0918*) as an endogenous control and fold changes were calculated using the $2^{-\Delta\Delta CT}$ Method (25). Data were calculated from 3 independent experiments in triplicate.

3.2.7 Transcriptional fusions to luxCDABE and luminescence assays

The promoter region from *Bcas0627* was PCR amplified using primers 5195-5196. The PCR product was digested with *Eco*RI and cloned into the *Eco*RI digested and dephosphorylated pGSVTp-*lux* plasmid. The resulting plasmid contained the promoter

region of *Bcas0627* fused to the *luxCDABE* reporter system. The plasmid was mobilized into K56-2 and the appropriate mutants by triparental mating. Transconjugants were selected on LB agar plates containing 100 μg/ml of trimethoprim (Tp), 200 μg/ml ampicillin and 20 μg/ml gentamicin.

Luminescence assays were performed as described (12) to measure the expression of *Bcas0627* in different mutant backgrounds including K56-2 WT, Δ*atsR*, Δ*atsT*, Δ*atsR*::*atsR* (Δ*atsR* complemented at the chromosomal level with *atsR*), Δ*atsR*Δ*cepI*, Δ*atsR*Δ*cepI::atsR,* Δ*atsT*::*atsT* and Δ*atsT*Δ*cepI::atsT* (Table 3.1). Overnight cultures were diluted into fresh medium with trimethoprim 100 mg/ml at OD_{600} of 0.05. Three hundred μl of sample was loaded in triplicate in a 100-well honeycomb microtitre plate for Bioscreen C automated microbiology growth curve analysis. Growth rates were followed by measuring the OD_{600} at 37 °C every hour under continuous shaking. At 6 h, 9 h, 12 h, 15 h and 18 h post-inoculation, the Bioscreen was paused and three aliquots from each strain were transferred into a flat bottom 96-well microtiter plate (Microfluor 2White, Thermo Scientific) and luminescence (in RLU) was measured using a Fluoroskan Ascent FL (Thermo Labsystems). Expression levels of each gene of interest in the different strain backgrounds were calculated as RLU/OD_{600} for each of the six time-points. Each experiment was repeated independently two times.

3.2.8 Protein purification and Electromobility Shift Assays

C-terminal His-tagged AtsT and AtsR expressed in *E. coli* BL21 were purified by affinity chromatography as described before (13). Five-prime oligos labeled with IRDYE682 purchased from Operon. The amplified labeled DNA duplexes (150 ng) were

purified with QIAquick PCR purification kit (Qiagen) and incubated with various concentrations of AtsT-His in the presence of poly dI-dC and run on a native polyacrylamide gel (8%) in 1X TBE buffer. Band shifts were visualized by Odyssey imaging system (Licor Biosciences). Experiments were repeated three times.

3.2.9 Non-radioactive DNase I footprinting

To examine AtsT binding sites identified by ChIP-seq, we used fluorescently labeled primers in standard DNase I protection assays using an Applied BioSystems 3730xl DNA analyzer. DNA fragments (approximately 250 bp in size) containing predicted AtsT binding sites were amplified using a primer labeled with the 6-carboxyfluorescein (6-FAM) fluorescent dye. PCR products were purified using the QIAquick PCR purification kit (Qiagen). Binding reaction mixtures consisted of 150 ng of labeled DNA probe and 5-10 µM AtsT in binding buffer (30 mM Tris-HCl pH 7, 1 mM $MnCl_2$, 120 mM KCl, and 1 mM $MgCl_2$). These were incubated at 37 °C for 10 min, transferred to room temperature, and allowed to cool for 2 min. DNase I (0.1 U or 0.2 Unit enzyme; New England BioLabs) was added to the reaction mixtures and digestion was allowed to proceed for 2 min at room temperature. The reaction was stopped by the addition of DNase I stop solution (20 mM EDTA pH 8, 1% SDS, and 200 nM NaCl). The DNA was purified with QIAquick PCR purification kit (Qiagen) and submitted to the Carver Biotechnology center at Illinois University for fragment analysis on an Applied BioSystems 3730xl DNA analyzer. The results were visualized using Gene mapper Software.

3.2.10 Gene expression profiling with microarray

Bacteria were grown at 37 °C for 16 h and then diluted to an OD_{600} of 0.05 in 30 ml LB.

Diluted cultures of *B. cenocepacia* K56-2 and the K56-2 *atsR*::pDA27 mutant, DFA21, were grown at 37 °C until exponential phase, harvested and immediately used for RNA extraction. Total RNA was isolated using a RiboPure-Bacteria kit (Ambion, Austin, TX), according to the manufacturer's instructions. RNA samples were treated with 8 U of DNaseI, which was inactivated using DNase inactivation reagent (Ambion, Austin, TX). RNA was concentrated with LiCl and the concentration of RNA samples was measured using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). RNA integrity was assessed by agarose gel electrophoresis and by measuring the ratio of absorbance at 260 nm to 280 nm (values obtained between 2.0 and 2.2). RNA was prepared in three independent experiments for microarray analysis. Labeled cDNA was synthesized from RNA using CyScribe Post Labeling kit (GE Healthcare) and hybridized to custom *B. cenocepacia* microarrays (Agilent Technologies, Santa Clara, CA, USA) according to the Agilent 60-mer oligonucleotide microarray processing protocol, and scanned by the Mahenthiralingam Laboratory, Cardiff University, Wales as previously described (26). The raw Data was imported into GeneSpring (version 7.3.1) and normalized using the "Agilent FE" procedure. The t-test p-value (< 0.05) was used to filter the genes of confidence with application of two multiple testing correction parameters Benjamini-Hochberg false discovery rate (BH_FDR) and "Bonferroni" correction. An arbitrary cutoff value of 2-fold ratio change (either upregulation or repression) was chosen to identify genes that were differentially expressed between K56-2 wild type and K56-2 *atsR*::pDA27.

3.3 Results and discussion

3.3.1 Microarray analysis of *atsR* mutant and wild-type strains

I previously characterized the global role of AtsR in regulating QS signalling and the expression of QS regulated virulence factors (12). To identify the AtsR regulon, microarray analysis was conducted by comparing the gene expression profiles in wild type and *atsR* mutant strains. ORFs that showed a >2-fold difference were considered to be differentially expressed. Overall, 36 ORFs and 8 intergenic regions were upregulated (Fig. 3.1A&B) and 89 ORFs and 16 intergenic regions were downregulated (Fig. 3.1B) in the absence of a functional AtsR. The upregulated regions encoded 16 conserved hypothetical proteins with unknown functions, 3 putative lipoproteins, 2 T6SS related genes including Hcp (+6-fold) and BCAL1294 (VgrG, +3-fold), and other genes involved in QS, chemotaxis, transport and regulatory functions (Table 3.3). Notably, we observed a significant fold change among genes encoding the putative lipoproteins *Bcam0382* and *Bcas0627* with up to 102 and 99-fold increase, respectively. Also, a 30-fold increase in the expression of CepR2 (*Bcam0188)* was observed. CepR2 is a regulator of QS genes including zinc metaloproteases and of the *Bcam*0192 operon (27), which is also detected in our microarrays. Of the 105 downregulated regions, 12 genes were flagella-related, implying a role of AtsR to stimulate flagellum formation or motility. Six genes encode chemotaxis related proteins including CheY and CheZ and 9 DNA binding regulatory proteins. The remaining genes belong to other functional categories (Table 3.3).

Table 3.3: List of genes differentially upregulated in AtsR microarray analysis

Gene Name	Normalized Data	Description
BCAL0343	6.972	type VI secretion system protein TssD
BCAL0358	3.917	family M1 metalopeptidase /aminopeptidase N
BCAL1294	3.094	type VI secretion system secreted protein VgrG
BCAL1295	2.637	conserved hypothetical protein
BCAL1387	4.638	putative phosphatase
BCAL1952	2.595	LysM domain/BON superfamily protein
BCAL3228	2.688	hypothetical protein
BCAM0067	2.226	putative short chain dehydrogenase
BCAM0068	2.887	Major Facilitator Superfamily protein
BCAM0069	3.061	conserved hypothetical protein
BCAM0070	2.798	putative hydrolase
BCAM0188	2.005	N-acylhomoserine lactone dependent regulatory
BCAM0192	31.04	conserved hypothetical protein
BCAM0194	29.7	conserved hypothetical protein
BCAM0196	15.44	conserved hypothetical protein
BCAM0200	7.265	efflux system transport protein
BCAM0202	2.031	conserved hypothetical protein
BCAM0382	87.05	putative lipoprotein
BCAM0384	44.55	putative lipoprotein
BCAM0633	2.091	conserved hypothetical protein
BCAM0785	2.301	conserved hypothetical protein
BCAM0821	2.575	putative methyl-accepting chemotaxis protein
BCAM0835	3.252	AraC family regulatory protein
BCAM1023	2.082	putative selenocysteine-specific elongation
BCAM1103	3.128	
BCAM1106	2.928	putative oxidoreductase
BCAM1345	2.729	conserved hypothetical protein
BCAM1869	2.173	conserved hypothetical protein
BCAM2153	2.088	putative cytochrome P450 oxidoreductase
BCAM2154	2.715	conserved hypothetical protein
BCAM2425	2.292	conserved hypothetical protein
BCAS0153	2.27	conserved hypothetical protein
BCAS0627	99.35	putative lipoprotein
BCAS0664	2.88	conserved hypothetical protein
BCAS0666	2.446	putative ankyrin-repeat exported protein
IG1_1411086	2.955	upstream of BCAL1297 / interG_chr1_pos_767_1411086:1411195
IG1_2151274	4.529	upstream of BCAL1952 /interG_chr1_pos_1124_2151274:2151732
IG1_2166257	2.995	Spans BCAL1968-1973 /interG_chr1_pos_1130_2166257:2171209

IG1_368927	4.551	upstream of BCAL0343 (Hcp)/ interG_chr1_pos_243_368927:368996
IG2_1491140	3.083	Spans BCAM1343-1348 / interG_chr2_pos_769_1491140:1498982
IG2_2676155	2.241	BCAM2377 + surroundings / interG_chr2_pos_1268_2676155:2677543
IG2_429106	116.1	Spans AtsR-AtsT+ upstream of BCAM0382 (putative lipoprotein)/ interG_chr2_pos_211_429106:432827
IG2_433416	102.1	upstream BCAM0383 (putative lipoprotein) interG_chr2_pos_212_433416:433537
pBCA055	2.267	putative protein

3.3.2 Promoter activity assay

According to our microarray data, *Bcas0627* is one of the most upregulated genes in absence of AtsR. To confirm and validate our microarray results, the promoter activity of *Bcas0627* was monitored in WT and single Δ*atsR* or double mutant *ΔcepIΔatsR* strains using a luciferase reporter system. We showed previously that several virulence factors, including swarming motility and the secretion of zinc metalloproteases are positively regulated by CepI and negatively regulated by AtsR (12). As shown in Fig. 3.1C, *Bcas0627* maximum expression occurs in the absence of AtsR at 9 h. *Bcas0627* expression in *ΔcepIΔatsR* followed the same pattern as in *ΔatsR* except that expression levels were consistently higher. Interestingly, *Bcas0627* promoter activity was almost null in the presence of AtsR either in WT strain or *ΔcepIΔatsR::atsR* and *ΔatsR::atsR* mutants that are complemented with *atsR*. Together, these results confirm our microarrays data regarding *Bcas0627*, which its expression is indeed upregulated in the absence of AtsR.

I previously showed that *ΔatsT* and *ΔatsR* mutants have similar phenotypes including increased biofilm formation, upregulation of T6SS activity and increased extracellular proteolytic activity (13) and showed that AtsR/AtsT pair act as a two-component system both *in vivo* and *in vitro* (13). Since AtsT and AtsR are in the same regulatory pathway, we expect that regulatory role of AtsR on lipoproteins takes place via AtsT. To determine if the putative lipoprotein BCAS0627 is indeed regulated by AtsT, the *Bcas0627*::*luxCDABE* transcriptional fusion was introduced into *ΔatsT, ΔatsT::atsT, ΔcepIΔatsT* and *ΔcepIΔatsT::atsT* mutant strains. Interestingly, the relative light unit (RLU)/OD_{600} values in the Δ*atsT* background are similar to Δ*atsR*, indicating that there is a direct correlation between the *Bcas0627* promoter activity and the presence of AtsT

(Fig.3.1C&D). The highest luminescence was detected in the K56-2Δ*atsT*Δ*cepI* double mutant where *atsT* and *cepI* are both deleted. Similarly, the K56-2Δ*atsT* single mutant has high levels of promoter activity; whereas, mutants complemented with *atsT* have reduced levels of luminescence (Fig.3.1D). Together, these data confirm that the expression of *Bcas0627*is regulated by AtsR/AtsT two component system and that AtsT works downstream of AtsR.

3.3.3 Confirmation of AtsT binding via electromobility gel shift assay

To investigate whether the AtsT binds directly to the DNA sequences upstream of *Bcam0382*, *Bcas0627* and *Bcam0192* identified by microarray, we carried out electrophoretic mobility shift assays (EMSA) with His-tagged AtsT protein. IRDYE682 labeled PCR fragments of about 190-360 bp in size were purified and incubated with various concentrations of AtsT-His (0-40 μM) with a DNA concentration of 150 ng. This varied protein to DNA ratio was necessary to address relative affinity differences between the promoter fragments. We detected higher molecular weight bands for *Bcam0382* and *Bcas0627* promoter regions when incubated with AtsT protein. Only 2 μM and 10 μM of AtsT was sufficient for *Bcam0382* and *Bcas0627* DNA sequences respectively to obtain an almost complete shift, whereas such shift was minimal for *Bcam0192* even with a concentration of 20 μM of AtsT protein in the reaction. Further experiments revealed that the *Bcam0192* promoter sequence has lower affinity for AtsT and binding occurs in the presence of higher concentration of AtsT (Fig. 3.2A). To test if DNA-binding ability of AtsT is dispensable from its phosphorylation status, *Bcam0382* and *Bcas0627* promoter regions where incubated with phosphorylated protein and band shifts were compared to that of the reaction where DNA was incubated with unphosphorylated AtsT. Interestingly, DNA shifts were comparable for both promoter sequences regardless of AtsT phosphorylation status (Fig. 3.2B). Together, these data indicate that *in vitro* unphosphorylated AtsT is capable of binding DNA sites

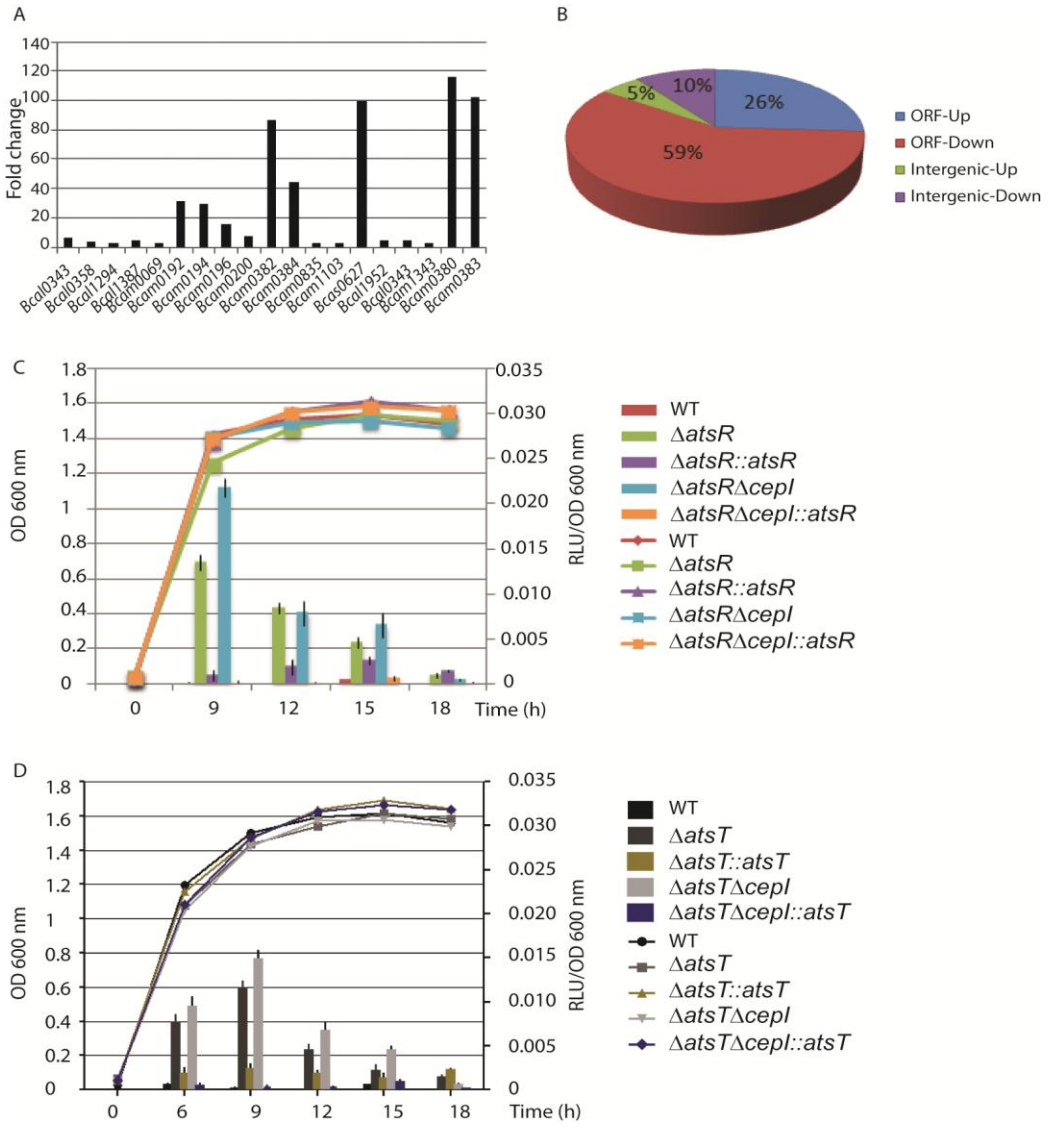


Figure 3.1: Expression of *Bcas0627* is dependent on the presence of functional AtsR and AtsT. A) Expression profiles of representative genes upregulated >3 folds in the absence of a functional AtsR. B) Pie chart representing percentage of up or downregulation of expression profiles regarding their genomic position in *atsR* mutant strain. Promoter activity assay of *Bcas0627* in the presence or absence of AtsR (C) or AtsT (D). Bacterial growth (OD_{600}, lines) and luminescence (RLU/OD_{600}, bars) reflecting *Bcas0627* expression in the different strain backgrounds were measured. Values are average ± SD for triplicates and are representative of three independent experiments.

with varying affinities which may have an impact on the sensitivity of promoters to the phosphorylation state of AtsT. In other words, the different binding affinities to the three promoters we tested here may allow the fine-tuning of expression in response to AtsT-P levels with varying levels of signaling molecule(s).

To test the specificity of AtsT binding to the promoter sequences, DNA binding was carried out in the presence of varying concentrations of unlabeled promoters (as specific competitor). A 100-fold excess of DNA was required to reverse the band shift for *Bcam0382* indicating higher affinity of AtsT for this promoter sequence (Fig. 3.2C, lane 9), whereas, AtsT binding to the labeled *Bcas0627* promoter was reduced by adding a 50-fold excess of unlabeled probe (Fig. 3.2D, lane 7). Further support for the specificity of the interaction with DNA was provided by experiments in which a nonspecific fragment corresponding to the *Bcal3310* upstream sequence was used, and as expected, no shift in DNA mobility was detected with an identical fold excess of nonspecific competitor DNA (Fig. 3.2C right panel & 3.2D lanes 8-10) reflecting that AtsT binding is sequence specific.

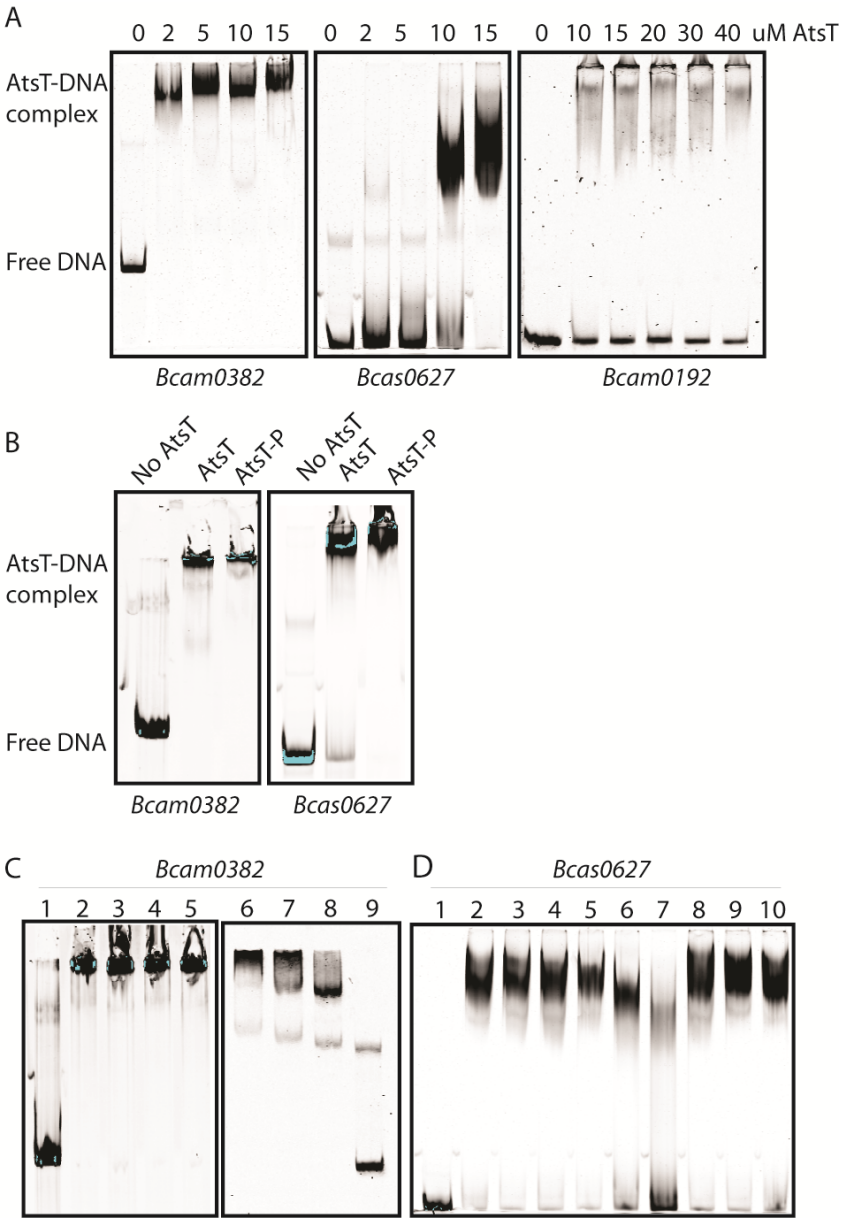


Figure 3.2: Electrophoretic mobility shift *in vitro* with the *Bcam0382*, *Bcas0627* and *Bcam0192* promoter fragments and purified AtsT-His protein. A) The DNA fragments (150 ng) from promoter regions of *Bcam0382* (380 bp), *Bcas0627* (200 bp) and *Bcam0192* (190 bp) were incubated with AtsT-His in the presence of poly dI-dC and run on a native polyacrylamide gel (8%), in 1X TBE buffer. AtsT binding was assessed on the basis of the amount of shifted material for each fragment in the presence of various

concentrations of AtsT-His. B) The DNA fragments (150 ng) from promoter regions of *Bcam0382* and *Bcas0627* were incubated with 5 µM or 10 µM of either unphosphorylated AtsT-His or phosphorylated AtsT-His in a reaction containing 10 µM AtsR and 100 mM ATP and run on a native polyacrylamide gel (8%), in 1X TBE buffer. C) Specificity of AtsT binding to *Bcam0382* promoter sequence in the presence of specific (left panel) and non-specific sequence (right panel) were tested with increasing concentration of DNA in the reaction from 10x (lane 3 and 6), 20x (lane 4 and 7), 50x (lane 5 and 8) to 100x (lane 9). D) Specificity of AtsT binding to *Bcas0627* promoter sequence in the presence of specific (lanes 2-7) and non-specific sequence (lanes 8-10) were tested with increasing concentration of DNA in the reaction from 2x (lane 3), 5x (lane 4), 10x (lane 5 and 8), 20x (lanes 6 and 9) to 50x (lane 7 and 10).

3.3.4 Confirmation of AtsT binding via DNAse I footprinting

To examine and validate the AtsT binding site in the promoter regions of *Bcas0627* and *Bcam0382*, we used fluorescently labeled primers for DNase I protection assays. DNA fragments containing potential AtsT binding sitee were amplified using a primer labeled with the 6-carboxyfluorescein (6-FAM) fluorescent dye. PCR products were purified and digested by DNase I in a binding reaction mixture. The results were visualized using Gene mapper Software. Protein binding sites were indicated by the disappearance of the peaks. As shown in Fig. 3.3, AtsT protected a region from DNase I digestion which is located upstream of the *Bcas0627* and the *Bcam0382* start codons, respectively (Fig. 3.3A&B upper panels). When the same experiment was performed in the absence of Atst, theses regions were not protected (Fig. 3.3A&B lower panels).

3.3.5 Isolation of DNA fragments associated with AtsT

Our aim was to use ChIP to measure the distribution of AtsT across the genome of *B. cenocepacia*. Using the CLC genomics workbench, we analyzed three independent ChIP-seq runs using BSA ChIP as a control. We mapped a total of 17,241,315 reads to the *B. cenocepacia* J2315 genome compared to 238,990 reads for control experiment. After mapping the reads to the reference genome, the ChIP-Seq analysis was done by calling peaks based on detection of signals to identify genomic regions with significantly enriched read coverage and a read distribution with a characteristic shape such as shown in Fig. 3.4A. We combined the peaks generated from all three samples and required that a peak be called in at least two of the three experiments to be considered an AtsT binding site. Representative regions displaying predicted AtsT binding sites for three different p-

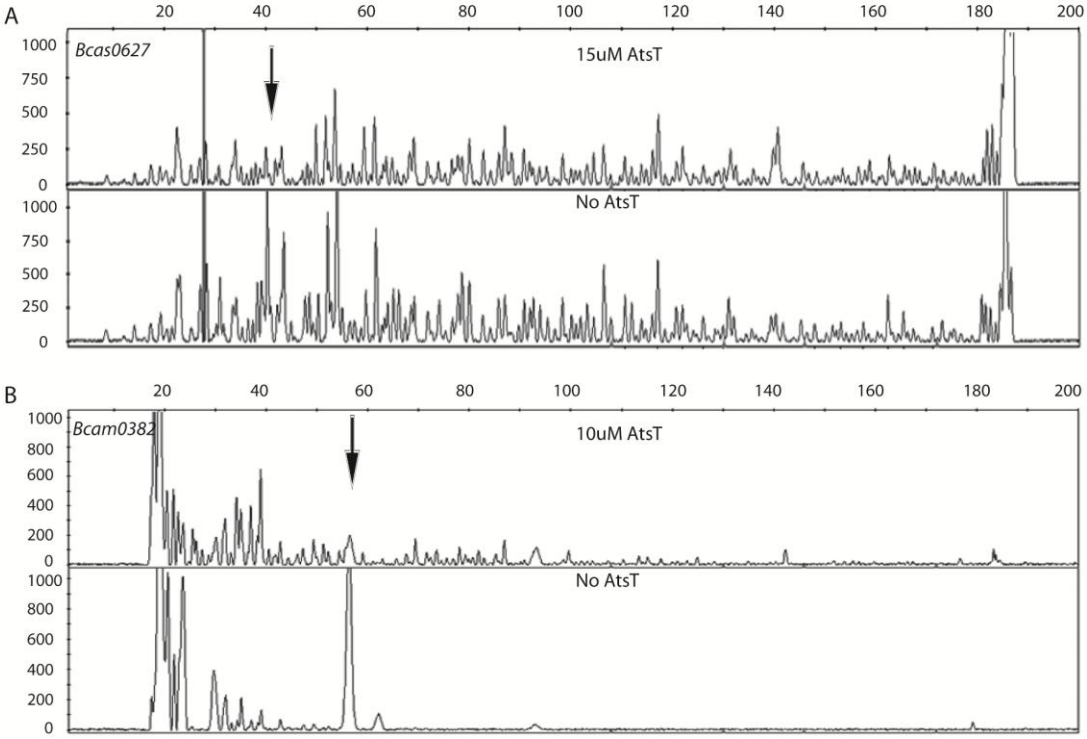


Figure 3.3: Fragment analysis of AtsT-binding site by DNaseI footprinting. Graphs indicate the interaction of 10 or 15 µM AtsT or BSA (no AtsT panel) with promoter regions of *Bcas0627* (A) and *Bcam0382* (B). Arrows point to the missing peaks in both top panels of A and B graphs where AtsT is present in the reaction. This suggests a binding site for AtsT in those areas that is protected from digestion. These pictures are screen shots of the graphs on Gene mapper software.

values annotated with J2315 genome are shown in Fig. 3.4B. A total of 380, 218 and 93 enriched regions were identified for threshold p-values of 0.001, 0.0005 and 0.0001, respectively. We focused on analysing the regions obtained from the p-value setting of 0.0005 and these enriched sequences were extracted. Of these 218 regions, almost 46% were intergenic (Fig. 3.4C) were chosen for further characterization. The rest are inside genes or have overlap with the ends. Complete extracted sequences can be found in Appendix B.

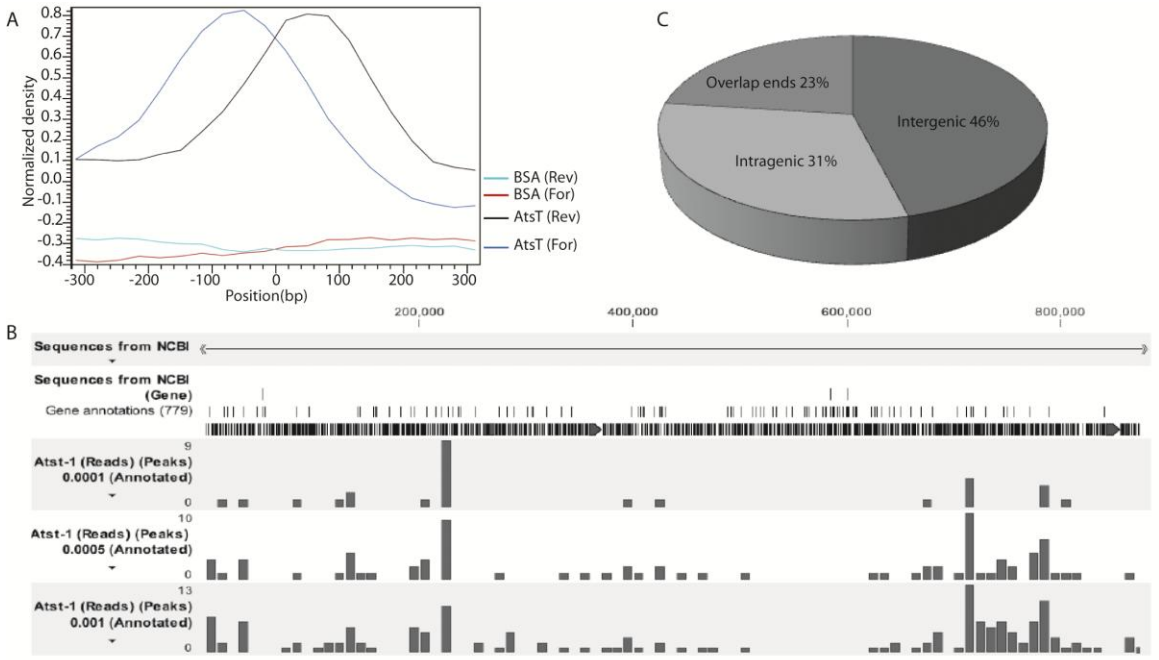


Figure 3.4: Pattern of AtsT ChIP-Seq peaks identified upstream of the AtsT-regulated genes (A). Representative image of genome-wide binding sites of AtsT with 3 different threshold settings (B). Pie chart of AtsT-binding peaks annotated relatively to the nearest gene translation start in *B. cenocepacia* geneome (C).

Table 3.4: List of genes annotated with the enriched AtsT loci obtained from ChIP-Seq analysis

Chromosome	Region	Length	5' gene	3' gene	Function
NC_011000	3253014..3253166	153	*BCAL2969*	*BCAL2969a*	hypothetical protein
NC_011000	3247579..3247713	135	*BCAL2965C*	*BCAL2965D*	hypothetical protein
NC_011000	3247833..3248013	181	*BCAL2965D*	*BCAL2966*	hypothetical protein; K06919 putative DNA primase/helicase
NC_011000	3244246..3244373	128	*BCAL2962a*	*BCAL2962b*	hypothetical protein
NC_011000	3245237..3245374	138	*BCAL2962b*	*BCAL2963*	putative phage-related DNA-binding protein
NC_011000	3252472..3252600	129	*BCAL2969*	*BCAL2969a*	hypothetical protein
NC_011000	3244006..3244151	146	*BCAL2962*	*BCAL2962a*	hypothetical protein
NC_011000	3254478..3254652	175	*BCAL2971*	*BCAL2971a*	hypothetical protein
NC_011002	36502..36690	189	*BCAS0033*	*BCAS0034*	metallo-beta-lactamase superfamily protein
NC_011000	3246856..3246992	137	*BCAL2965*	*BCAL2965a*	hypothetical protein
NC_011002	204538..204677	140	*BCAS0188A*	*BCAS0189*	hypothetical protein; K11312 cupin 2 domain-containing protein
NC_011000	3241596..3241765	170	*BCALr2960*	*BCAL2961*	putative integrase
NC_011000	3247147..3247344	198	*BCAL2965a*	*BCAL2965b*	hypothetical protein
NC_011001	432522..432657	136	*BCAM0381*	*BCAM0382*	hypothetical lipoprotein
NC_011000	3253510..3253673	164	*BCAL2970*	*BCAL2971*	hypothetical protein
NC_011002	716403..716546	144	*BCAS0654*	*BCAS0656*	putative transposase
NC_011002	130355..130492	138	*BCAS0117*	*BCAS0118*	putative H-NS family DNA-binding protein; K03746 DNA-binding protein H-NS
NC_011000	253951..254114	164	*BCAL0218c*	*tufA1*	elongation factor Tu
NC_011002	422485..422621	137	*BCAS0360*	*BCAS0361*	LysR family regulatory protein
NC_011000	3737727..3737897	171	*BCAL3411*	*mtgA*	monofunctional biosynthetic peptidoglycan transglycosylase
NC_011002	122093..122278	186	*BCAS0109*	*BCAS0110*	periplasmic solute-binding protein; K10018 octopine/nopaline transport system substrate-binding protein
NC_011000	812763..812898	136	*BCAL0747*	*BCAL0748*	hypothetical protein
NC_011002	710870..711047	178	*BCAS0649*	*BCAS0650*	putative transposase
NC_011000	365924..366070	147	*BCAL0339*	*BCAL0340*	putative lipoprotein
NC_011002	88283..88413	131	*cpoF*	*BCAS0080*	putative pyridine nucleotide-disulphide oxidoreductase family protein
NC_011001	3025995..3026135	141	*BCAM2678*	*BCAM2679*	hypothetical protein
NC_011002	138427..138569	143	*BCAS0124*	*BCAS0125*	hypothetical protein
NC_011001	2790921..2791083	163	*BCAM2466*	*BCAM2467*	putative transposase
NC_011000	61713..61898	186	*BCAL0050*	*BCAL0051*	periplasmic solute-binding

					protein; K10013 lysine/arginine/ornithine transport system substrate-binding protein
NC_011002	390461..390594	134	*BCAS0334*	*BCAS0335*	putative haemagglutinin-related autotransporter protein
NC_011000	596436..596589	154	*BCAL0543*	*dppA*	putative periplasmic dipeptide transport protein; K12368 dipeptide transport system substrate-binding protein
NC_011002	713043..713195	153	*BCAS0650b*	*BCAS0652*	putative transposase
NC_011000	34503..34658	156	*atpB*	*atpE*	F0F1 ATP synthase subunit C (EC:3.6.3.14); K02110 F-type H+-transporting ATPase subunit c
NC_011001	7278..7422	145	*BCAM0005A*	*BCAM0006*	hypothetical protein
NC_011002	11987..12129	143	*BCAS0010*	*BCAS0011*	hypothetical protein
NC_011001	2720007..2720185	179	*BCAM2415*	*BCAM2416*	hypothetical protein
NC_011002	788081..788216	136	*BCAS0717*	*BCAS0718*	transposase; K07497 putative transposase
NC_011002	748460..748598	139	*BCAS0680*	*BCAS0683*	
NC_011000	1246576..1246713	138	*BCAL1146*	*BCAL1147*	glycosyltransferase
NC_011001	2223385..2223519	135	*BCAM2006b*	*BCAM2007*	TonB-dependent siderophore receptor; K16090 catecholate siderophore receptor
NC_011002	629782..629939	158	*BCAS0577*	*ansP*	L-asparagine permease; K11738 L-asparagine permease
NC_011000	3504662..3504804	143	*BCAL3208*	*BCAL3208a*	hypothetical protein
NC_011001	2739339..2739504	166	*BCAM2426*	*BCAM2427*	hypothetical protein
NC_011001	3070082..3070253	172	*BCAM2713*	*BCAM2714*	putative transposase
NC_011000	657042..657190	149	*BCAL0601*	*BCAL0602*	MerR family regulatory protein
NC_011002	754523..754670	148	*BCAS0688*	*BCAS0689*	metallo-beta-lactamase superfamily protein
NC_011000	3474261..3474410	150	*BCAL3177*	*BCAL3178*	LysR family regulatory protein
NC_011002	370809..370953	145	*BCAS0321*	*BCAS0321a*	hypothetical protein
NC_011000	203026..203161	136	*BCAL0181*	*BCAL0182*	putative plasmid recombinase
NC_011001	2912664..2912801	138	*aer*	*BCAM2565*	putative methyltransferase
NC_011002	333449..333583	135	*BCAS0297*	*BCAS0298*	flp type pilus subunit; K02651 pilus assembly protein Flp/PilA
NC_011002	810233..810398	166	*BCAS0735*	*BCAS0736*	TetR family regulatory protein; K09017 TetR/AcrR family transcriptional regulator
NC_011002	390719..390843	125	*BCAS0334*	*BCAS0335*	putative haemagglutinin-related autotransporter protein
NC_011000	3820929..3821084	156	*BCAL3483*	*BCAL3484*	putative transposase

NC_011000	919933..920067	135	*BCAL0848*	*BCAL0849*	subfamily M48B metalopeptidase; K07387 putative metalloprotease [EC:3.4.24.-]
NC_011001	24329..24498	170	*adc*	*BCAM0024*	putative undecaprenyl-diphosphatase
NC_011002	687343..687498	156	*BCAS0626*	*BCAS0627*	putative lipoprotein
NC_011002	712628..712787	160	*BCAS0650a*	*BCAS0650b*	pseudogene
NC_011000	400302..400447	146	*cspD*	*BCAL0369*	putative amino acid permease
NC_011001	506891..507029	139	*BCAM0451*	*BCAM0452*	family S10 serine peptidase
NC_011002	400289..400485	197	*BCAS0342*	*BCAS0343*	Major Facilitator Superfamily protein
NC_011000	567233..567372	140	*BCAL0515*	*BCAL0516*	hypothetical protein
NC_011000	136340..136504	165	*fliC*	*rpsU*	rpsU; 30S ribosomal protein S21; K02970 small subunit ribosomal protein S21
NC_011001	3069020..3069154	135	*BCAM2711*	*BCAM2712*	hypothetical protein
NC_011000	1704707..1704879	173	*BCAL1538*	*BCAL1539*	hypothetical protein
NC_011000	452189..452342	154	*BCAL0412*	*BCAL0413*	hypothetical protein
NC_011001	90660..90811	152	*BCAM0080*	*BCAM0081*	putative outer membrane protein-OmpW family
NC_011002	500495..500632	138	*BCAS0432*	*BCAS0433*	putative DNA-binding protein
NC_011002	751453..751614	162	*BCAS0682*	*BCAS0684*	pseudogene
NC_011000	146554..146750	197	*BCAL0123*	*flhD*	transcriptional activator FlhD; K02403 flagellar transcriptional activator FlhD
NC_011000	1405446..1405606	161	*BCAL1292*	*BCAL1293* (here *bcal1294* is in the microarray VgrG protein)	
NC_011002	748152..748291	140	*BCAS0679A*	*BCAS0680*	putative TniB-like transposition protein
NC_011000	159157..159313	157	*cheZ*	*BCAL0137*	hypothetical protein
NC_011002	4443..4616	174	*BCAS0003*	*BCAS0004*	putative transposase; K07486 transposase
NC_011000	3456900..3457078	179	*BCAL3164*	*BCAL3165*	putative exported hydrolase; K07518 hydroxybutyrate-dimer hydrolase
NC_011000	32943..33127	185	*BCAL0028*	*BCAL0029*	putative ATP synthase protein I AtpI; K02116 ATP synthase protein I
NC_011002	869202..869370	169	*BCAS0769*	*BCAS0770*	LysR family regulatory protein
NC_011000	3718958..3719122	165	*BCAL3394*	*maeB*	maeB; malic enzyme (EC:1.1.1.40); K00029 malate dehydrogenase (oxaloacetate-decarboxylating)(NADP+)
NC_011000	3375805..3375977	173	*BCAL3083*	*BCAL3084*	putative guanine/cytidine deaminase
NC_011000	3496591..3496745	155	*BCAL3199*	*tolQ*	putative TolQ transport transmembrane protein; K03562 biopolymer transport protein TolQ

NC_011000	3297325..3297457	133	*cspA*	*BCAL3007*	hypothetical protein
NC_011001	3136490..3136633	144	*BCAM2772*	*BCAM2773*	putative DNA-binding protein
NC_011002	708654..708802	149	*BCAS0646*	*BCAS0646A*	putative DNA-binding protein
NC_011002	723986..724116	131	*BCAS0661C*	*BCAS0662*	hypothetical protein
NC_011000	1147412..1147552	141	*BCAL1058*	*argD*	bifunctional N-succinyldiaminopimelate-aminotransferase/acetylornithine transaminase protein
NC_011001	1018568..1018719	152	*BCAM0922*	*BCAM0923*	putative lipoprotein
NC_011002	784816..784973	158	*BCAS0716*	*BCAS0717*	hypothetical protein
NC_011000	562335..562477	143	*BCAL0510*	*BCAL0511*	putative deoxygenases; K00477 phytanoyl-CoA hydroxylase
NC_011001	1173919..1174053	135	*BCAM1078*	*BCAM1079*	hypothetical protein
NC_011002	209614..209811	198	*BCAS0194*	*BCAS0195*	putative transcriptional regulator
NC_011000	3671381..3671556	176	*BCAL3352*	*BCAL3353*	putative outer membrane autotransporter
NC_011000	3529532..3529665	134	*BCAL3227*	*BCAL3228*	hypothetical protein
NC_011000	3532227..3532381	155	*BCAL3228*	*BCAL3229*	hypothetical protein
NC_011002	149887..150045	159	*BCAS0135*	*BCAS0136*	putative amino acid permease
NC_011002	135581..135723	143	*BCAS0122*	*BCAS0123*	putative transcriptional regulator
NC_011002	466699..466882	184	*BCAS0398*	*citH*	citrate-proton symporter; K03288 MFS transporter, MHS family, citrate/tricarballylate:H+ symporter
NC_011002	682721..682913	193	*BCAS0622*	*BCAS0623*	putative NAD(P)H dehydrogenase (quinone) (EC:1.6.5.2); K00355 NAD(P)H dehydrogenase (quinone)
NC_011001	2521772..2521943	172	*BCAM2251*	*BCAM2251A*	hypothetical protein
NC_011002	357264..357459	196	*iorA*	*BCAS0321*	hypothetical protein
NC_011002	714169..714365	197	*BCAS0652*	*BCAS0653*	putative transposase
NC_011000	139291..139424	134	*rpsU*	*BCAL0117*	hypothetical protein
NC_011000	179296..179450	155	*BCAL0153*	*BCAL0154*	histone-like nucleoid-structuring (H-NS) protein; K03746 DNA-binding protein H-NS
NC_011000	257756..257891	136	*rplA*	*rplJ*	rplJ; 50S ribosomal protein L10; K02864 large subunit ribosomal protein L10
NC_011002	205769..205901	133	BCAS0190	BCAS0191	putative endoribonuclease

3.3.6 Motif identification from ChIP-seq data

The extracted sequences of all 101 intergenic AtsT binding regions for the threshold p-value of 0.0005 were searched for a common motif using MEME-ChIP. The software identified 3 predicted motifs that were enriched in ChIP-Seq peak regions and corresponding sequence logos were built. Inspection of the alignment reveals a match for the first predicted motif to the binding area obtained from DNAse I footprinting for *Bcam0382* and *Bcas0627* upstream sequences. This motif has a position p-value less than 0.0001 which has 21 nucleotides that the first 4 bases (TTGC) and G at position 12 are the most conserved bases (Fig. 3.5A). Alignment with upstream regions of annotated genes indicated that this motif is found in all 101 enriched ChIP-Seq peak sequences. A screenshot of this alignment is shown in Figure 3.5B.

3.3.7 Comparison with genes differentially transcribed in microarray data

A detailed transcriptome analysis comparing wild-type and AtsR mutant described AtsR dependent gene regulation during exponential growth in LB media. To identify direct relevant target genes, we analyzed the AtsT ChIP-seq dataset against transcriptional profiling data of control and *atsR* mutant strains. Of the 101 upstream sequences that were enriched, only 6% were associated with genes differentially transcribed in the microarray experiment. Two of these genes that encode putative lipoproteins (*Bcas0627* and *Bcam0382*) have been shown by footprinting assay to bind AtsT directly (Fig. 3.3). Also, several of these genes are the first in apparent operons containing additional genes that are derepressed in the *atsR* mutant strain. For example, the upstream region of *Bcal0340* may serve dually as a binding site for T6SS gene cluster located both downstream of *Bcal0340* including genes encoding TssA to H proteins and upstream of

Figure 3.5: Sequence logos generated by MEME-ChIP. A) The most highly significant motif identified in all 101 ChIP-Seq peaks with a p-value of less than 0.0001. The height of each letter represents the relative frequency of each base at each position in the consensus sequence. B) Representative alignments of the predicted motif against enriched ChIP-Seq peak regions annotated to *B. cenocepacia* genome.

Bcal0340 including genes encoding TssJ to L proteins. *Bcal0343* encodes Hcp protein and is a recognized AtsR-regulated gene (11). The remaining sites include sequence regions of a hypothetical protein (BCAM2425), a putative phage integrase (BCAM1024), BCAL1293 which is in the vicinity of the putative type six secreted VgrG protein BCAL1294 and BCAS0667 (type six secreted VgrG protein ortholog).

3.3.8 Validation of ChIP-Seq results by qRT-PCR

To validate binding of AtsT to these loci, we chose a subset of 15 genes and used quantitative RT-PCR (qRT-PCR) to determine the fold enrichment of their transcripts in WT versus Δ*atsT* strains. Quantification of fold enrichment of 14 genes identified by ChIP-seq showed greater enrichment in the Δ*atsT* mutant than the WT parental strain except for the *FliC* and *FlhD* genes that were expressed slightly more in WT strain (Fig. 3.6). These results support our ChIP peaks as authentic AtsT binding sites and are in line with our previous finding on the negative regulatory role of AtsR/AtsT on swarming motility (12, 13).

3.4 Conclusion

We have defined the direct and indirect AtsR/AtsT regulon by comparing the expression profiling of an AtsR mutant with a modified chromatin-IP procedure, to experimentally identify sites bound by AtsT throughout the *B. cenocepcia* genome. ChIP-seq is a method of choice for identifying regulator-binding sites in eukaryotic systems, and we show here its utility for mapping the genomic binding sites of a bacterial transcriptional regulator. Since the expression level of a transcriptional regulator may vary depending on experimental conditions, affecting its binding potential, we have used a modified

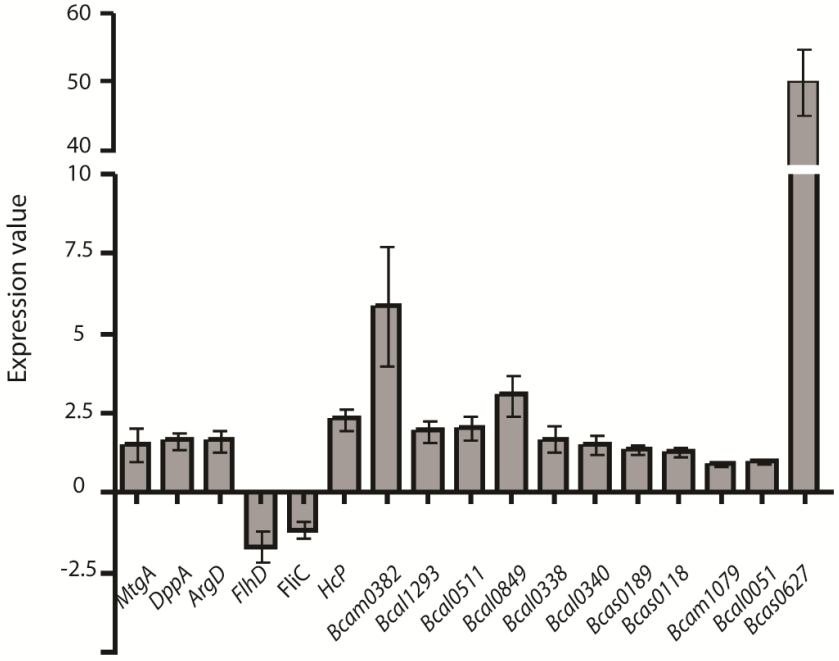


Figure 3.6: Transcript expression analysis by qRT-PCR. The expression levels of 16 genes identified by ChIP-Seq analysis were studied in K56-2 WT and Δ*atsT* and their relative expression value was normalized to *RpoE* gene. Bars indicate means ± SD.

approach of ChIP method to overcome these difficulties in genome-wide binding analysis. In this modified ChIP method we have the benefit of pulling down DNA with the purified regulator protein in which the transcriptional regulator has a greater potential to bind low affinity sites and outcompete interfering proteins under *in vivo* conditions. On the other hand, using a purified transcriptional regulator renders the possibility of producing false-positive ChIP peaks by binding to nonspecific sites. To address this issue we used BSA as a control and had a stringent setting for analysing our high throughput sequencing results. We identified 101 binding sites that are the AtsT target regulon (Table 3.4). The function of nearly 30% of these genes is unknown as they are listed as hypothetical proteins. The remainder have known or predicted functions in a wide range of cellular processes as indicated in Table 3.4. Although we aimed to identify genes directly regulated by AtsT, transcriptome analyses of the *atsR* mutant strain identified several genes regulated by AtsR, including genes encoding regulators that might function as intermediates in regulatory cascades (Table 3.3). A comparative analysis of these genes with ChIP-Seq data showed that there are only a few genes in common. This result reflects the possibility that AtsT might be involved in other signalling pathways and act as a regulatory arm for several receptor proteins besides AtsR. Indeed, as AtsT plays a key role in the regulation of greater number of genes compared to AtsR, its activity levels may vary with the nature of the stimulus that might be perceived by multiple sensor proteins. Furthermore, we studied in depth the *Bcas0627* and *Bcam0382* genes which were shown to be upregulated in transcriptomic studies in the absence of a functional AtsR (Table 3.3). We confirmed AtsT peaks by EMSA and identified the binding site by foot printing. The AtsT binding loci were analyzed and this resulted in prediction of a

consensus sequence. This prediction was consistent with the results obtained from DNAse I footprinting and the EMSA of the same locus.

Overall, combining this AtsT binding location analysis with expression analysis of *atsR* mutant strain, our results indicate that AtsT directly regulates the expression of nearly 100 genes that are spread across all three chromosomes of *B. cenocepacia* and include metabolic, regulatory, transport, cell-surface associated and environmental information processing genes as well as conserved genes of unknown function for which differences in expression have been observed in *atsT* mutant versus wild-type strains. Furthermore, our bioinformatic and DNase I footprinting analyses reveal the AtsT binding site architecture that likely has important implications for global gene regulation in *B. cenocepacia*.

This study is the first to characterize the AtsR/AtsT regulon in *B. cenocepacia*, an important regulatory pathway for QS and T6SS related genes, and is one of only a few studies to apply ChIP-seq to bacterial regulators. This significantly expands our understanding of the AtsR/AtsT regulon and highlights the complexity of this regulatory pathway.

3.5 Chapter 3 references

1. **Mahenthiralingam E, Baldwin A, Dowson CG.** 2008. *Burkholderia cepacia* complex bacteria: opportunistic pathogens with important natural biology. J Appl Microbiol **104:**1539-1551.

2. **Isles A, Maclusky I, Corey M, Gold R, Prober C, Fleming P, Levison H.** 1984. *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. J Pediatr **104:**206-210.

3. **Vergunst AC, Meijer AH, Renshaw SA, O'Callaghan D.** 2010. *Burkholderia cenocepacia* creates an intramacrophage replication niche in zebrafish embryos, followed by bacterial dissemination and establishment of systemic infection. Infect Immun **78:**1495-1508.

4. **Khodai-Kalaki M, Andrade A, Fathy Mohamed Y, Valvano MA.** 2015. *Burkholderia cenocepacia* lipopolysaccharide modification and flagellin glycosylation affect virulence but not innate immune recognition in plants. MBio **6**.

5. **Cardona ST, Wopperer J, Eberl L, Valvano MA.** 2005. Diverse pathogenicity of *Burkholderia cepacia* complex strains in the *Caenorhabditis elegans* host model. FEMS Microbiol Lett **250:**97-104.

6. **Uehlinger S, Schwager S, Bernier SP, Riedel K, Nguyen DT, Sokol PA, Eberl L.** 2009. Identification of specific and universal virulence factors in *Burkholderia cenocepacia* strains by using multiple infection hosts. Infect Immun **77:**4102-4110.

7. **Lamothe J, Huynh KK, Grinstein S, Valvano MA.** 2007. Intracellular survival of *Burkholderia cenocepacia* in macrophages is associated with a delay in the maturation of bacteria-containing vacuoles. Cell Microbiol **9:**40-53.

8. **Lamothe J, Thyssen S, Valvano MA.** 2004. *Burkholderia cepacia* complex isolates survive intracellularly without replication within acidic vacuoles of *Acanthamoeba polyphaga*. Cell Microbiol **6:**1127-1138.

9. **Schwab U, Abdullah LH, Perlmutt OS, Albert D, Davis CW, Arnold RR, Yankaskas JR, Gilligan P, Neubauer H, Randell SH, Boucher RC.** 2014. Localization of *Burkholderia cepacia* complex bacteria in cystic fibrosis lungs and interactions with *Pseudomonas aeruginosa* in hypoxic mucus. Infect Immun **82:**4729-4745.

10. **Stock AM, Robinson VL, Goudreau PN.** 2000. Two-component signal transduction. Annu Rev Biochem **69:**183-215.

11. **Aubert DF, Flannagan RS, Valvano MA.** 2008. A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in *Burkholderia cenocepacia*. Infect Immun **76:**1979-1991.

12. **Aubert DF, O'Grady EP, Hamad MA, Sokol PA, Valvano MA.** 2013. The *Burkholderia cenocepacia* sensor kinase hybrid AtsR is a global regulator modulating quorum-sensing signalling. Environ Microbiol **15:**372-385.

13. **Khodai-Kalaki M, Aubert DF, Valvano MA.** 2013. Characterization of the AtsR hybrid sensor kinase phosphorelay pathway and identification of its response regulator in *Burkholderia cenocepacia*. J Biol Chem **288:**30473-30484.

14. **Galperin MY.** 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. J Bacteriol **188:**4169-4182.

15. **Kahramanoglou C, Seshasayee AS, Prieto AI, Ibberson D, Schmidt S, Zimmermann J, Benes V, Fraser GM, Luscombe NM.** 2011. Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. Nucleic Acids Res **39:**2073-2091.

16. **Davies BW, Bogard RW, Mekalanos JJ.** 2011. Mapping the regulon of *Vibrio cholerae* ferric uptake regulator expands its known network of gene regulation. Proc Natl Acad Sci U S A **108:**12467-12472.

17. **Bernier SP, Nguyen DT, Sokol PA.** 2008. A LysR-type transcriptional regulator in *Burkholderia cenocepacia* influences colony morphology and virulence. Infect Immun **76:**38-47.

18. **Hamad MA, Di Lorenzo F, Molinaro A, Valvano MA.** 2012. Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. Mol Microbiol **85:**962-974.

19. **Figurski DH, Helinski DR.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci U S A **76:**1648-1652.

20. **Hamad MA, Skeldon AM, Valvano MA.** 2010. Construction of aminoglycoside-sensitive *Burkholderia cenocepacia* strains for use in studies of intracellular bacteria with the gentamicin protection assay. Appl Environ Microbiol **76:**3170-3176.

21. **Sambrook J, Russell DW.** 2006. The condensed protocols from Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

22. **Cohen SN, Chang AC, Hsu L.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc Natl Acad Sci U S A **69:**2110-2114.

23. **Craig FF, Coote JG, Parton R, Freer JH, Gilmour NJ.** 1989. A plasmid which can be transferred between *Escherichia coli* and *Pasteurella haemolytica* by electroporation and conjugation. J Gen Microbiol **135:**2885-2890.

24. **Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS.** 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res **37:**W202-208.

25. **Livak KJ, Schmittgen TD.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods **25:**402-408.

26. **Sass A, Marchbank A, Tullis E, Lipuma JJ, Mahenthiralingam E.** 2011. Spontaneous and evolutionary changes in the antibiotic resistance of *Burkholderia cenocepacia* observed by global gene expression analysis. BMC Genomics **12:**373.

27. **Malott RJ, O'Grady EP, Toller J, Inhulsen S, Eberl L, Sokol PA.** 2009. A *Burkholderia cenocepacia* orphan LuxR homolog is involved in quorum-sensing regulation. J Bacteriol **191:**2447-2460.

Chapter 4

Functional and topological analysis of AtsR, a global regulator of quorum sensing in *Burkholderia cenocepacia*

A version of this chapter has been prepared for submission:

Maryam Khodai-Kalaki, Faviola Tavares, Miguel A. Valvano. Functional and topological analysis of AtsR, a global regulator of quorum sensing in *Burkholderia cenocepacia*

4.1 Introduction

In bacteria, two-component systems (TCS) play critical roles in transducing environmental stimuli into the cell to regulate cellular functions (1). A classical TCS is composed of a transmembrane sensor histidine kinase that autophosphorylates in response to specific signals and consequently regulates the activity of a cytoplasmic response regulator which can control transcription through DNA binding (2). We previously identified AtsR, a membrane protein of *Burkholderia cenocepacia*, that negatively regulates quorum sensing and virulence factors such as biofilm, type 6-secretion and protease secretion (3, 4). The *in silico* structural analysis of AtsR reveals that this protein is a hybrid sensor kinase, which falls in the category of TCSs. Based on secondary structure prediction, AtsR consists of two transmembrane domains, a periplasmic "sensing" domain and a cytoplasmic C-terminal domain carrying a transmitter domain with a conserved histidine (5) residue and a catalytic domain. In our previous work we studied the individual domains of the cytoplasmic C-terminus end of this protein and identified the conserved residues within its kinase domain that are absolutely required for its biological function. Furthermore, we identified its cognate response regulator AtsT and characterized the phosphotransfer mechanism in AtsR/AtsT signalling pathway (6).

Topology is a fundamental aspect of the structure of membrane proteins. Since AtsR is a membrane-bound receptor and regulates virulence factors in *B. cenocepacia*, studying the structure of this protein is crucial. Knowledge of AtsR topology and its functional residues will provide insight into the potential signals perceived by this protein and its role in pathogenicity and host adaptation. Here, we report a topological analysis of AtsR using LacZ and PhoA reporter fusions and sulfhydryl labeling of novel cysteines that are

introduced into a cysteine-less AtsR by PEGylation. We further identify the functional residues within the periplasmic loop that are required for AtsR function *in vivo*. Our results confirm that the experimental topological model of AtsR is in agreement with the predicted model.

4.2 Materials and methods

4.2.1 Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 4.1. Bacteria were grown in Luria broth at 37 °C unless indicated otherwise. *E. coli* cultures were supplemented, as required, with the following antibiotics final concentrations: tetracycline 30 μg/ml, kanamycin 30 μg/ml and trimethoprim 50 μg/ml. *B. cenocepacia* cultures were supplemented, as required, with trimethoprim 100 μg/ml, tetracycline 150 μg/ml, ampicillin 200 μg/ml and polymyxin B 20 μg/ml. The β-galactosidase (LacZ) or alkaline phosphatase (PhoA) activities were assessed by examining the blue-colony phenotypes on LB plates containing a final concentration of 40 mg/ml of 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-Gal) and 5-bromo-4-chloro-3-indolylphosphate (XP), respectively.

4.2.2 General molecular techniques

DNA manipulations were performed with T4 DNA ligase (Roche Diagnostics, Laval, Quebec, Canada) and Antarctic phosphatase (New England Biolabs, Pickering, Ontario, Canada) were used as recommended by the manufacturers. Transformation of *Escherichia coli* DH5α and GT115 was done using the calcium chloride method (10). Mobilization of complementing plasmids and mutagenesis plasmids into *B. cenocepacia*

Table 4.1. List of strains and plasmids

Strain or Plasmid	Relevant Properties	Source or Reference
Strains		
E. coli		
CC118	Δ(*ara leu*) Δ*lac phoA galE galK thi rpsL rpsB argE recA*	Laboratory stock
DH5α	F^- ϕ80*lacZ*ΔM15 *endA recA hsdR*(r_K^- m_K^-) *nupG thi glnV deoR gyrA relA1 Δ*(*lacZYA-argF*)*U169*	Laboratory stock
BL21	F^- omp (Tgal *dcm ompT hsdS*(Rb^- mB^-) *gal* λ(DE3)	Laboratory stock
GT115	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80Δ*lacZ*ΔM15 Δ*lacX74 recA1 rpsL endA1*Δ*dcm uidA*(ΔMluI)::*pir-116* Δ*sbcC-sbcD*	Laboratory stock
Plasmids		
pAH01	pBAD vector expressing Flag-Wzx-K367-PhoA	(7)
pAH18	pBAD vector expressing Flag-PhoA	(7)
pAH1809	pBAD vector expressing Flag-Wzx-T242-PhoA	(7)
pBADNTF	pBAD24 vector for N-terminal Flag fusions, AmpR	(7)
pMH447	pGPI-SceI derivative used for chromosomal complementation	(8)
pRK2013	$ori_{colE1,}$ RK2 derivative, Kan^R, mob^+, tra^+	(9)
pDAI-SceI-SacB	ori_{pBBR1}, Tet^R, P_{dhfr}, mob^+, expressing I-SceI, SacB	(8)
pMK23	pBAD expressing Flag-AtsR-S100-PhoA	This study
pMK24	pBAD expressing Flag-AtsR-Y160-PhoA	This study
pMK25	pBAD expressing Flag-AtsR-Q215-PhoA	This study
pMK26	pBAD expressing Flag-AtsR-T300-PhoA	This study
pMK27	pBAD expressing Flag-AtsR-L402-PhoA	This study
pMK28	pBAD expressing Flag-AtsR-S100-LacZ	This study
pMK29	pBAD expressing Flag-AtsR-Y160-LacZ	This study
pMK30	pBAD expressing Flag-AtsR-Q215-LacZ	This study
pMK31	pBAD expressing Flag-AtsR-T300-LacZ	This study
pMK32	pBAD expressing AtsR-L402-LacZ	This study
pMLTraP-BCAS0627	promoter of *bcas0627* cloned in front of full *lacZ* Cloning vector, IPTG inducible for N-terminal 6X His fusion, Kan^r	Valvano Lab (Unpublished)
pET28a(+)		Laboratory stock
pMZ25	pET28a(+) expressing AtsR-His	(6)
pMK37	pET28a(+) expressing AtsRcycless-His	This study
pMK40	pET28a(+) expressing AtsRcycless-V14C-His	This study
pMK41	pET28a(+) expressing AtsRcycless-M21C-His	This study
pMK42	pET28a(+) expressing AtsRcycless-V38C-His	This study
pMK43	pET28a(+) expressing AtsRcycless-G41C-His	This study
pMK44	pET28a(+) expressing AtsRcycless-G49C-His	This study
pMK45	pET28a(+) expressing AtsRcycless-G120C-His	This study
pMK46	pET28a(+) expressing AtsRcycless-Q181C-His	This study
pMK47	pET28a(+) expressing AtsRcycless-A182C-His	This study
pMK48	pET28a(+) expressing AtsRcycless-G187C-His	This study
pMK50	pET28a(+) expressing AtsRcycless-M205C-His	This study
pMK51	pET28a(+) expressing AtsRcycless-A328C-His	This study
pMK52	pET28a(+) expressing AtsRcycless-G484C-His	This study

pFT159	pBAD expressing AtsR-His	This study
pFT160	pBAD expressing AtsRcycless-His	This study
pFT205	pBAD expressing AtsRcycless-V14C-His	This study
pFT183	pBAD expressing AtsRcycless-M21C-His	This study
pFT184	pBAD expressing AtsRcycless-V38C-His	This study
pFT185	pBAD expressing AtsRcycless-G41C-His	This study
pFT186	pBAD expressing AtsRcycless-G49C-His	This study
pFT187	pBAD expressing AtsRcycless-G120C-His	This study
pFT188	pBAD expressing AtsRcycless-Q181C-His	This study
pFT206	pBAD expressing AtsRcycless-A182C-His	This study
pFT189	pBAD expressing AtsRcycless-G187C-His	This study
pFT190	pBAD expressing AtsRcycless-M205C-His	This study
pFT191	pBAD expressing AtsRcycless-A328C-His	This study
pFT192	pBAD expressing AtsRcycless-G484C-His	This study
pMK76	pBAD expressing AtsR-Q56A-His	This study
pMK77	pBAD expressing AtsR-Y61A-His	This study
pMK78	pBAD expressing AtsR-D65A-His	This study
pMK79	pBAD expressing AtsR-F79A-His	This study
pMK80	pBAD expressing AtsR-D80A-His	This study
pMK81	pBAD expressing AtsR-S103A-His	This study
pMK82	pBAD expressing AtsR-W105A-His	This study
pMK83	pBAD expressing AtsR-K127A-His	This study
pMK84	pBAD expressing AtsR-N155A-His	This study
pMK85	pBAD expressing AtsR-R162A-His	This study
pMK86	pBAD expressing AtsR-Q56N-His	This study
pMK87	pBAD expressing AtsR-Q56R-His	This study
pMK88	pBAD expressing AtsR-D65E-His	This study
pMK89	pBAD expressing AtsR-D65K-His	This study
pMK90	pBAD expressing AtsR-S103T-His	This study
pMK91	pBAD expressing AtsR-S103V-His	This study
pMK92	pBAD expressing AtsR-R162K-His	This study
Strains		
B. cenocepacia		
K56-2	ET12 clone related to J2315, CF clinical isolate	[b]BCRRC
K56-2Δ*atsR*	Deletion of *atsR* in K56-2	(6)
K56-2Δ*atsR::atsR*	Chromosomal *atsR* integration in K56-2 Δ*atsR*	(6)
MK93	Chromosomal *atsRcycless* integration in K56-2 Δ*atsR*	This study
MK94	Chromosomal *atsR* integration in K56-2 Δ$atsR_{Q56A}$	This study
MK95	Chromosomal *atsR* integration in K56-2 Δ$atsR_{Q56N}$	This study
MK96	Chromosomal *atsR* integration in K56-2 Δ$atsR_{Q56R}$	This study
MK97	Chromosomal *atsR* integration in K56-2 Δ$atsR_{D65A}$	This study
MK98	Chromosomal *atsR* integration in K56-2 Δ$atsR_{D65K}$	This study
MK99	Chromosomal *atsR* integration in K56-2 Δ$atsR_{D65E}$	This study
MK100	Chromosomal *atsR* integration in K56-2 Δ$atsR_{S103A}$	This study
MK101	Chromosomal *atsR* integration in K56-2 Δ$atsR_{S103T}$	This study
MK102	Chromosomal *atsR* integration in K56-2 Δ$atsR_{S103V}$	This study
MK103	Chromosomal *atsR* integration in K56-2 Δ$atsR_{162A}$	This study
MK104	Chromosomal *atsR* integration in K56-2 Δ$atsR_{162K}$	This study

MK105	Chromosomal *atsR* integration in K56-2 Δ$atsR_{F79A}$	This study
MK106	Chromosomal *atsR* integration in K56-2 Δ$atsR_{D80A}$	This study
MK107	Chromosomal *atsR* integration in K56-2 Δ$atsR_{K127A}$	This study
MK108	Chromosomal *atsR* integration in K56-2 Δ$atsR_{N155A}$	This study
MK109	Chromosomal *atsR* integration in K56-2 Δ$atsR_{W105A}$	This study
MK110	Chromosomal *atsR* integration in K56-2 Δ$atsR_{Y61A}$	This study

[a] Tp^R, trimethoprim resistance, Kan^R, kanamycin resistance, Tet^R, tetracycline resistance.

[b] BCRRC, *B. cepacia* Research and Referral Repository for Canadian CF Clinics.

K56-2 was performed by triparental mating using *E. coli* DH5α carrying the helper plasmid pRK2013. DNA amplification by PCR was performed using Bio-Rad C1000^{TM} Thermal Cycler with Taq or HotStar HiFidelity DNA polymerases (Qiagen, Canada). DNA sequences of all primers used in this study are described in Table 4.2. DNA sequencing was performed at the DNA sequencing Facility of York University, Toronto, Canada. BLAST was used to analyse the sequenced genome of *B. cenocepacia* strain J2315.

4.2.3 Plasmid construction and chromosomal complementation

To create His-tagged fusions of AtsR and its derivatives, sequences were amplified from *B. cenocepacia* K56-2 genomic DNA using primers mentioned in Table 4.2. Amplicons were digested with the appropriate restriction enzymes and ligated into similarly digested pET28a cloning vector (Novagen) to engineer C-terminally His-tagged proteins. Site-directed mutagenesis was performed using the QuikChange Site-Directed Mutagenesis Kit from Stratagene (Santa Clara, California), as recommended by the supplier. Primers were designed with 15-20 nucleotides flanking each side of the targeted mutation. Plasmids expressing AtsR derivatives in pET28a were used as a template to clone His-tagged AtsR versions into pBAD24 using 4632-7266 primer set. The resulting PCR products were digested with *Dpn*I and introduced into *E. coli* DH5α. All constructs and replacement mutants were confirmed by sequencing. Chromosomal complementations of Δ*atsR* were performed using the pMH447 plasmid as described previously (8). Primers 7351-4632 was used to PCR amplify *atsR* and its mutated versions. PhoA fusions to residues S100, Y160, P215, N300, and K402 were constructed by amplifying the AtsR gene using primers mentioned in Table 4.2. Each PCR product and pAH18 were digested

Table 4.2. Oligonucleotide primers

Primer No.	Primer Name	5'-3' Primer sequence
252	pBAD Forward	GATTTAGCGGATCCTACCTGA
258	pBAD Reverse	GACCGCTTCTGCGTTCTGAT
4632	AtsR-For-NdeI	AAAACATATGACGCGGCGGCGATGGAAGAA
5033	PhoA-NotI-For	TATGCGGCCGCACCTGTTCTGGAAAACCGGGCTG
5034	PhoA-NotI-Rev	CTATGCGGCCGCTTTCAGCCCCAGAGCGGCTT
5640	LacZ-For-XbaI	CGACTCTAGAATTACGGATTCACTGGCCG
5866	AtsR-REV-XbaI	TTTTTCTAGAGCTTTGTTAGCAGCCGGATC
5885	MH447-For	TTGATGGCGAGCGATTCTTC
5886	MH447-Rev	CCAGTTCTTCAGCGTGACGA
6423	LacZ-Rev-HindIII	AAAAAGCTTTCAGCGCCATTCGCCATTC
6693	AtsR-100-Rev- XbaI	TTTTTCTAGACGACGGCCGCTGCAGGAA
6694	AtsR-160-Rev- XbaI	TTTTTCTAGAGTAGTTCGCGAGGCCGTTAACT
6696	AtsR-215-Rev- XbaI	TTTTTCTAGAGTTTTGCCGTTCGAGCAGGTCG
6743	AtsR-300-Rev-XbaI	TTTTTCTAGACGTCGAGCGCAGCTTGATGTA
6697	AtsR-402-Rev-XbaI	TTTTTCTAGAGCAGATGTTGTTTCGCGATGCCG
6719	AtsR-Fl-For-BamHI	AAAAGGATCCTATGACGCGGCGGCGATG
6728	AtsR-For-C255A	CCGCTGCAGGCGATCGCGGGATCGATCGAGATC
6729	AtsR-Rev-C255A	GATCTCGATCGATCCCGCGATCGCCTGCAGCGG
6730	AtsR-For-C515A	GACGCCGATCTGGCGGGCACCGGGCAG
6731	AtsR-Rev-C515A	CTGCCCGGTGCCCGCCAGATCGGCGTC
6763	AtsR-For-C195A	GGTGTGCTCCTGGCAGCGCTGTTCCTG
6764	AtsR-Rev-C195A	CAGGAACAGCGCTGCCAGGAGCACACC
6771	AtsR-For-M21C	GGCTCGCTGTGGATCTGTGGGTTCGCCGCGTGG
6772	AtsR-Rev-M21C	CCACGCGGCGAACCCACAGATCCACAGCGAGCC
6773	AtsR-For-G49C	GAGGGGCCGCGCGAATGTGTGTTCTGGACCGCC
6774	AtsR-Rev-G49C	GGCGGTCCAGAACACACATTCGCGCGGCCCCTC
6775	AtsR-For-A328C	CGCGAGAAGATCGTCTGCAAACGGATCGCGGTC
6776	AtsR-Rev-A328C	GACCGCGATCCGTTTGCAGACGATCTTCTCGCG
6777	AtsR-For-G484C	GCGCTGGCGTACCACTGTCGGCGCGCGCTGGTG
6778	AtsR-Rev-G484C	CACCAGCGCGCGCCGACAGTGGTACGCCAGCGC
6835	AtsR-For-V14C	AGAAAAATCATCCTGTGTCTCGGCTCGCTGTGG
6836	AtsR-Rev-V14C	CCACAGCGAGCCGAGACACAGGATGATTTTTCT
6837	AtsR-For-V38C	CTCCTCGCCACCTCGTGTAACGAGGGCGTGCTC
6838	AtsR-Rev-V38C	GAGCACGCCCTCGTTACACGAGGTGGCGAGGAG
6839	AtsR-For-G41C	ACCTCGGTCAACGAGTGTGTGCTCGAGGGGCCG
6840	AtsR-Rev-G41C	CGGCCCCTCGAGCACACACTCGTTGACCGAGGT
6841	AtsR-For-A182C	AAGGACAAACAGCGCTGTATCCTGATCCTCGGC
6842	AtsR-Rev-A182C	GCCGAGGATCAGGATACAGCGCTGTTTGTCCTT
6843	AtsR-For-Q181C	CAACTGAAGGACAAATGTCGCGCGATCCTGATC
6844	AtsR-Rev-Q181C	GATCAGGATCGCGCGACATTTGTCCTTCAGTTG

6845	AtsR-For-G187C	GCGATCCTGATCCTCTGTATCGTGCTCGGTGTG
6846	AtsR-Rev-G187C	CACACCGAGCACGATACAGAGGATCAGGATCGC
6847	AtsR-For-L201C	GCGCTGTTCCTGCTGTGTCTCTATACGATGCGC
6848	AtsR-Rev-L201C	GCGCATCGTATAGAGACACAGCAGGAACAGCGC
6849	AtsR-For-M205C	CTGCTGCTCTATACGTGTCGCACGCGCGACGAC
6850	AtsR-Rev-M205C	GTCGTCGCGCGTGCGACACGTATAGAGCAGCAG
6873	AtsR-For-G120C	GAGATCGACGCGCTGTGTGAATTCATGACGCGC
6874	AtsR-Rev-G120C	GCGCGTCATGAATTCACACAGCGCGTCGATCTC
7266	AtsR-Rev-HindIII	AAAAGCTTTCAGTGGTGGTGGTGGTGGTGCTCG
7322	AtsR-Gln56A-For	TTCTGGACCGCCGCGGCATACCGCAACGTGTAT
7323	AtsR-Gln56A-Rev	ATACACGTTGCGGTATGCCGCGGCGGTCCAGAA
7324	AtsR-Tyr61A-For	CAATACCGCAACGTGGCAACGCGCTTCGATCGT
7325	AtsR-Tyr61A-Rev	ACGATCGAAGCGCGTTGCCACGTTGCGGTATTG
7326	AtsR-Asp65A-For	GTGTATACGCGCTTCGCACGTCAGCTGATTCTT
7327	AtsR-Asp65A-Rev	AAGAATCAGCTGACGTGCGAAGCGCGTATACAC
7328	AtsR-Phe79A-For	CGCGAGGACGAAGACGCAGATCACCTGCAGATG
7329	AtsR-Phe79A-Rev	CATCTGCAGGTGATCTGCGTCTTCGTCCTCGCG
7330	AtsR-Asp80A-For	GAGGACGAAGACTTCGCACACCTGCAGATGCAG
7331	AtsR-Asp80A-Rev	CTGCATCTGCAGGTGTGCGAAGTCTTCGTCCTC
7332	AtsR-Ser103A-For	CGGCCGTCGGAAGTGGCACAATACTGGCTGCGC
7333	AtsR-Ser103A-Rev	GCGCAGCCAGTATTGTGCCACTTCCGACGGCCG
7334	AtsR-Tyr106A-For	GAAGTGTCCCAATACGCACTGCGCATCCCGCGC
7335	AtsR-Tyr106A-Rev	GCGCGGGATGCGCAGTGCGTATTGGGACACTTC
7336	AtsR-Lys127A-For	ATGACGCGCCTGGCACGCGAGGTGCCG
7337	AtsR-Lys127A-Rev	CGGCACCTCGCGTGCCAGGCGCGTCAT
7338	AtsR-Asn155A-For	TATTGGCCGAAAGTTGCCGGCCTCGCGAACTAC
7339	AtsR-Asn155A-Rev	GTAGTTCGCGAGGCCGGCAACTTTCGGCCAATA
7340	AtsR-Arg162A-For	CTCGCGAACTACTTCGCCGCGATCGAAATGGCT
7341	AtsR-Arg162A-Rev	AGCCATTTCGATCGCGGCGAAGTAGTTCGCGAG
7448	AtsR-Arg162D-For	CTCGCGAACTACTTCGATGCGATCGAAATGGCT
7449	AtsR-Arg162D-Rev	AGCCATTTCGATCGCATCGAAGTAGTTCGCGAG
7450	AtsR-Arg162K-For	CTCGCGAACTACTTCAAAGCGATCGAAATGGCT
7451	AtsR-Arg162K-Rev	AGCCATTTCGATCGCTTTGAAGTAGTTCGCGAG
7452	AtsR-Gln56N-For	TTCTGGACCGCCGCGAATTACCGCAACGTGTAT
7453	AtsR-Gln56N-Rev	ATACACGTTGCGGTAATTCGCGGCGGTCCAGAA
7351	AtsR-FL-XbaI-Rev	AAAATCTAGATCAGTGGTGGTGGTGGTGGT

with SmaI and XbaI, purified and ligated together using T4 ligase. Similarly, AtsR-LacZ fusions were constructed and inserted in pLA23 after removing the inserted WcaJ protein with primer 5640 and reverse primers mentioned in Table 4.2 for each truncated version of the AtsR protein.

4.2.4 Protein expression, membrane preparation and Western blot analysis

For expression of recombinant proteins, a single colony was inoculated in LB broth supplemented with appropriate antibiotics and grown at 37 °C. Optical Density (OD_{600}) was monitored until it reached 0.6. The culture was then shifted to 30 °C and arabinose was added to a final concentration of 0.2% (w/v) to induce the expression of proteins. Cultures were incubated for an additional 4 h. Cells were harvested by centrifugation at 8,000 x*g* for 10 min. Cell pellets were resuspended in lysis buffer containing [50 mM Tris-HCl, pH 8 and 1X of EDTA-free protease inhibitor cocktail (Sigma)], and disrupted by One Shot cell disruptor (Thermo Scientific, Rockville, MD, USA). Total membranes were isolated by centrifugation at 40,000 x*g* and resuspended in 50 mM Tris-HCl, pH 8. Protein concentration was determined by the Bradford protein assay (Bio-Rad). Total membrane preparations were used for immunoblotting and PEGylation. Membrane proteins were separated by 14% SDS-PAGE. Proteins were transferred to nitrocellulose membrane. The primary anti-FLAG monoclonal antibody (Sigma, Saint Louis, MO, USA), was diluted to 1: 10,000 and applied, then the secondary antibody goat anti-mouse Alexa fluor 680 IgG antibodies (Invitrogen Molecular Probes, Eugene, OR, USA) was diluted to 1: 20,000 and applied. For detection of 6xHis-tagged proteins, membranes were incubated with a 1:10,000 dilution of anti-His IgG2a monoclonal antibodies (Amersham, Piscataway, NJ) and Alexa Fluor 680 anti-mouse IgG antibodies (molecular probes).

Immunoblots were developed using LI-COR Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE, USA).

4.2.5 Protease assay

Protease assays were performed according to Aubert *et al*., (2013). Briefly, 18 h cultures were normalized to an OD_{600} of 1. The bacterial suspension (3 µl) was spotted onto dialysed brain heart infusion (D-BHI) agar plates containing 1.5% carnation milk. The plates were incubated at 37 °C and examined for zones of clearing around the bacterial spots at 48 h.

4.2.6 Alkaline phosphatase (PhoA) fusions and alkaline phosphatase assay

β-galactosidase and alkaline phosphatase activities were quantitated as previously described (11).

4.2.7 Thiol-specific chemical blocking and modification

Thiol-specific chemical blocking with *N*-ethylmaleimide (NEM) followed by modification with methoxypolyethylene glycol maleimide (PEG-mal) under nondenaturing and denaturing conditions was performed as described previously (12). For modification under nondenaturing conditions, membrane protein was split in three equally parts for three treatments: 1) 1-h incubation at room temperature with 1 mM PEG-mal, followed by addition of dithiothreitol (DTT) to 45 mM and further incubation for 10 min to quench the reaction; 2) 30-m incubation with 5 mM NEM to block accessible sulfhydryl groups, followed by immediate ultracentrifugation (16,000 × *g*, 1 min). The pellet was washed twice with 0.5 ml of ice-cold 50 mM HEPES pH 6.8 with 5

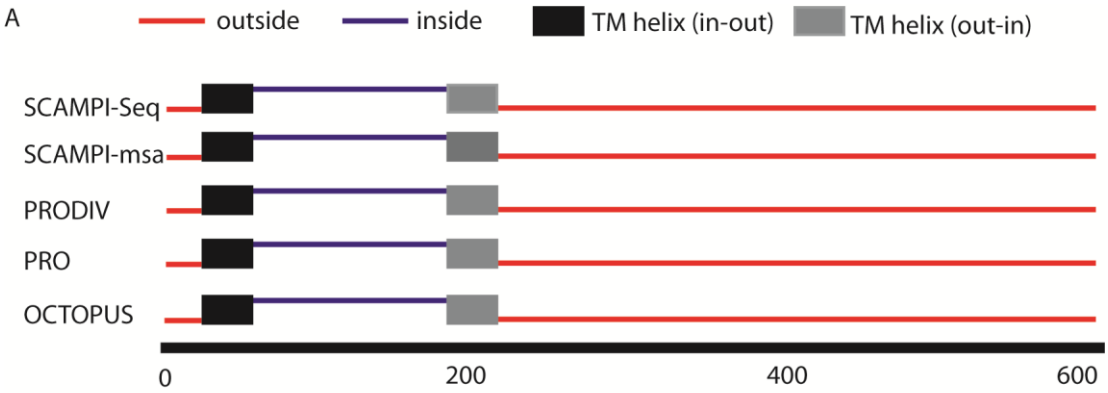
mM $MgCl_2$ plus protease inhibitors buffer to remove excess NEM. After another ultracentrifugation, the pellet was resuspended in 40 µl of resuspension buffer followed by reacting with 1 mM PEG-mal and quenching with DTT as above; 3) untreated control.

For modification under denaturing conditions, membranes were divided into 2 aliquots. One was directly treated with 1 mM PEG-mal in the presence of 2% SDS followed by addition of DTT and incubation for 10 min as described above. The other half was first incubated with 5 mM NEM for 30 min in the absence of SDS and washed as above. The sample was then incubated with 1 mM PEG-mal in the presence of 2% SDS at room temperature for 1 h. The reaction was finally quenched by incubation with 45 mM DTT for 10 min. All aliquots were mixed with sample buffer and subjected to SDS-PAGE.

4.3 Results and discussion

4.3.1 *In silico* analysis of AtsR topology

We used available TOPCONS (v2.0) web server for consensus prediction of membrane protein topology that implements 5 different algorithms including, OCTAPUS, PRO, PROVID, SCAMPI-seq and SCAMPI-msa to predict the membrane topology of AtsR (13). Topology prediction is mainly based on three attributes including number of TMs, TM-position, and periplasmic or cytoplasmic location of N- and C-terminal domains. According to all these algorithms AtsR has two predicted transmembrane domains from residues 10–30 (in-to-out) and residues 184–204 (out-to-in) separated by a large periplasmic loop (Fig. 4.1A). The prediction also suggests a short cytoplasmic N-terminus and a long C-terminal cytosolic tail region (residues 204-606) (Fig. 4.1B).



B

	TM1	TM2
SCAMPI-Seq	11-31	184-204
SCAMPI-msa	11-31	184-204
PRODIV	12-32	182-202
PRO	10-30	184-204
OCTOPUS	11-31	184-204

Figure 4.1: Graphical representation of the topological predictions of AtsR based on TOPCON web server. The numbers indicate the position of the amino acid residues (606 amino acids for AtsR). The location of cytosolic and periplasmic regions and the positions of predicted TMHs are indicated (A). The range of TM borders have also been predicted (B).

4.3.2 Topological analysis of AtsR using LacZ and PhoA reporter fusions

To validate the *in silico* predicted membrane topology of AtsR we employed a dual reporter fusion strategy based on LacZ and PhoA reporters, which were C-terminally fused to Flag-tagged AtsR derivatives that are truncated at various residues and expressed in *E. coli* strains CC118 lacking *phoA* and DH5α (*lacZ*ΔM15), respectively (Fig. 4.2A). Membrane expression of fusion proteins were tested by Western blot to make sure that fusion proteins are properly expressed (Fig. 4.2B). LB-agar plates were supplemented with chromogenic substrates X-P or X-Gal and colonies were examined for blue or white phenotypes. Also, the alkaline phosphatase and β-galactosidase activities were measured. If the PhoA reporter locates in the periplasm, the enzyme will be active as it requires a reducing environment to form disulphide bonds and will cleave the X-P substrate and therefore, bacterial colonies will turn blue (14). If PhoA locates in the cytoplasm, the enzyme is inactive, and bacterial colonies will be colourless. In contrast to PhoA, LacZ functions only in the cytosol generating blue colonies indicative of locations that are in the cytoplasmic space. Flag-tagged AtsR-LacZ fusions at residues S100 and Y160, which are located in the predicted periplasmic domain, did not exhibit β-galactosidase activity *in vivo* or *in vitro* (Fig. 4.3B). On the contrary, the corresponding Flag-AtsR-PhoA fusions at the same positions had a positive alkaline phosphatase enzymatic activity *in vitro* (1313 ± 50, 599 ± 42 Miller units) and strong blue colonies in the presence of X-P substrate (Fig. 4.3A), suggesting that S100 and Y160 reside in the periplasmic space. Despite relatively low *in vitro* activity (29 ± 90, 18 ± 43 Miller units), the LacZ fusion to T300 and L402, residues predicted to be in the cytosol, yielded blue colonies on X-gal plates, and the corresponding PhoA fusion gave no color on X-P plates, suggesting these

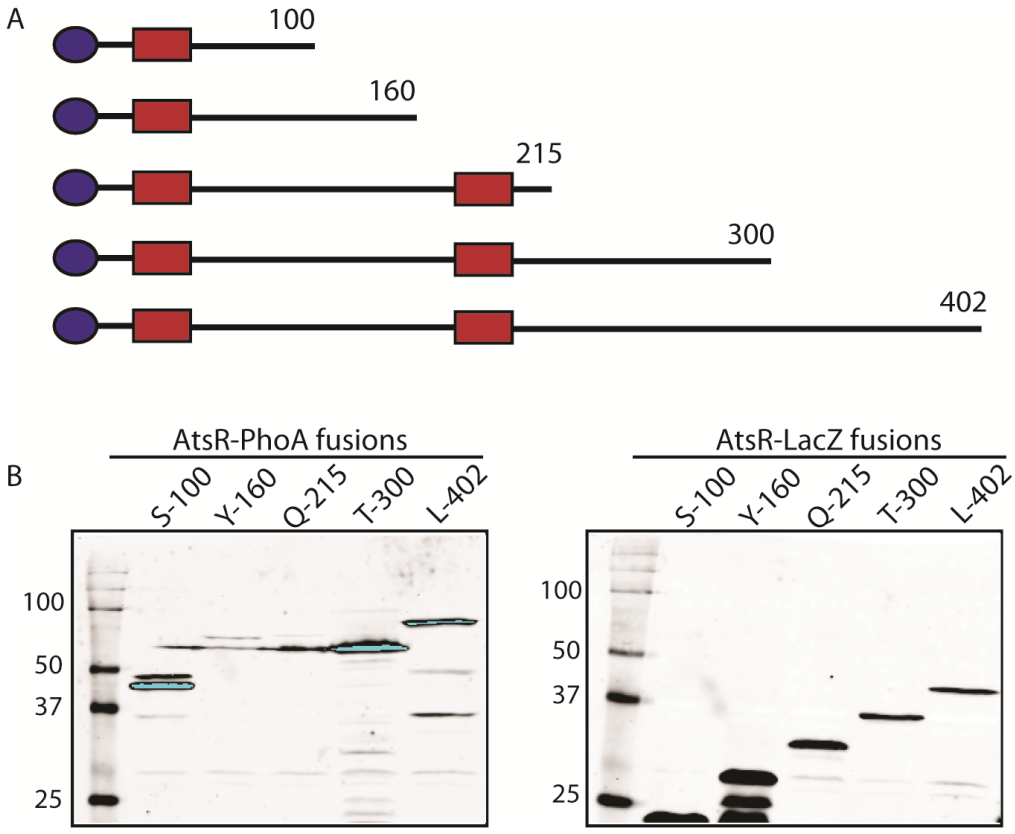


Figure 4.2: A) Schematic drawing of Flag-tagged AtsR truncated proteins fused to either PhoA or LacZ from C-terminus. Blue circle represents Flag tag, red boxes represent TMs and numbers correspond to the amino acids where fusion occurs. B) Membrane expression of Flag-AtsR-PhoA or -LacZ fusion proteins in CC118 (left) or DH5α cells (right). Flag-AtsR-PhoA or -LacZ proteins were expressed from the arabinose-inducible vector pBADNTF. Total membranes were isolated from cells expressing constructs and immunoblotting was performed using anti-Flag antibodies.

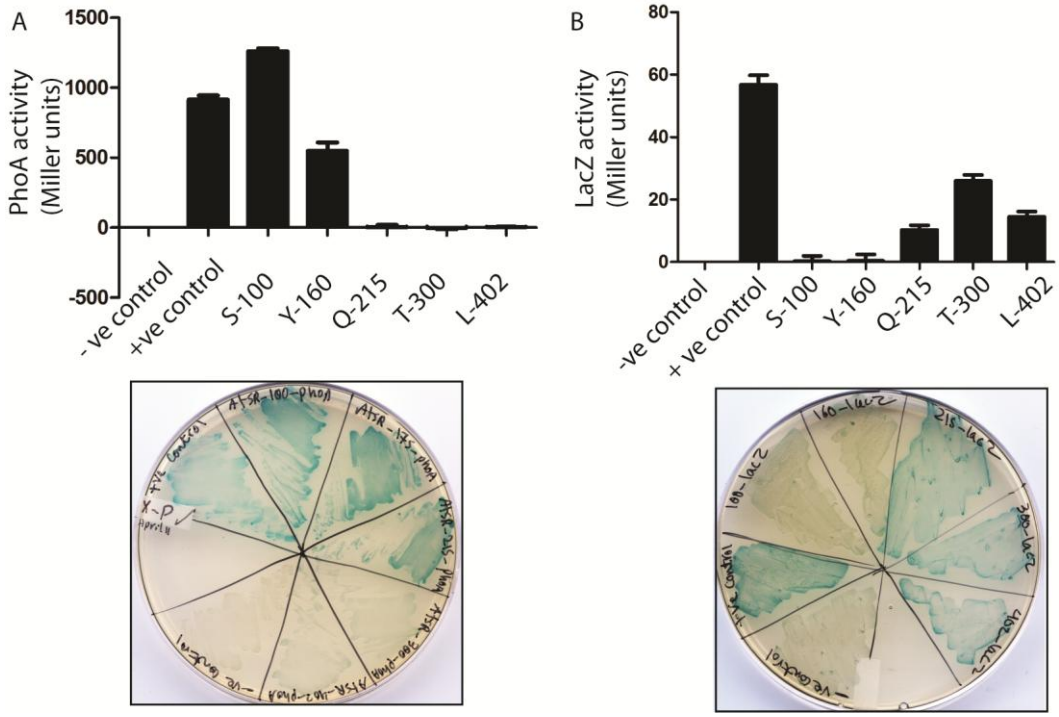


Figure 4.3: Alkaline phosphatase (A) and β-galactosidase (B) results of Flag-AtsR-PhoA/LacZ fusion proteins expressed in CC118 or DH5α cells. *In vitro* alkaline phosphatase and β-galactosidase assay were measured by the hydrolysis of PNPP and ONPG, respectively. Plate assays are shown in the bottom panel indicating the cleavage of X-gal or X-P substrate on plates.

residues are exposed to the cytosolic space (Fig. 4.3B). These results are in agreement with the predicted topology of AtsR. Based on our plate assay we obtained inconclusive results for LacZ and PhoA fusions at Q215 as both produce colonies with blue phenotype (Fig. 4.3 lower panel). This observation is not uncommon in fusions to amino acids located at the borders of TMs (15). However, Q215 yields very low *in vitro* β-galactosidase enzymatic activity (14 ± 21 Miller units) and no alkaline phosphatase activity (Fig. 4.3A&B). Here, the enzymatic data provide topological information for this residue which is predicted to be in the cytoplasmic face of the second TM. Taken together, these results confirm the *in silico* predicted topology of AtsR protein suggesting that AtsR is an inner membrane protein with a sensor domain facing the periplasm and a C-terminal cytoplasmic domain.

4.3.3 Topological analysis of AtsR using sulfhydryl reactive chemistry

To examine the borders of TMs in a more detailed analysis, we used a substituted cysteine labeling approach using sulfhydryl reactive reagents. In this method, the cysteine (Cys) residues are modified by the covalent attachment of PEG-Mal in the presence of a water molecule and the labeled proteins can be readily detected by Western blot by an increase of approximately 5 kDa to their apparent molecular mass (16, 17). Using this strategy the native structure of the protein is maintained with minimal disruption compared to reporter fusion method. We replaced the three native cysteines in AtsR at positions 195, 255 and 515 with alanine (Ala) residues to create a cysteine-less protein, $AtsR_{cysless}$. To make sure that the cysless version of AtsR is functionally active, the K56-2Δ*atsR* mutant was chromosomally complemented with $AtsR_{Cysless}$ and its corresponding proteolytic activity was quantified. This strain retained the protease activity at comparable

levels to the WT strain (Fig. 4.4A&B) indicating that it has a similar structure to the native AtsR protein and the three cysteine residues are dispensable for the function. Following this, several residues were replaced and novel cysteines were introduced at various positions around and within TMs by site-directed mutagenesis (Fig. 4.5). Labeling was performed in total membrane preparations with and without treatment with NEM (permeable to both inner and outer membrane) under denaturing and nondenaturing conditions to distinguish between exposed residues and those hidden in the lipid bilayer. NEM can permeate membranes and blocks the cysteine residues prior to PEG-mal modification in either side of the membrane to prevent them from becoming accessible to PEG-mal and causing a band shift. This property is useful to detect buried cysteine residues within TMs because NEM needs water molecules to be reactive and therefore, cysteine residues in the TM can be characterized by a band shift only after PEG-mal treatment in the presence of SDS. If a cysteine resides in the periplasm, PEG-mal can attach to the residue causing a clear band shift in both SDS denaturing and nondenaturing conditions. This band shift will not appear if NEM blocks the cysteine before PEG-mal treatment. On the contrary, if a cysteine resides in the cytosol, it can only be exposed to PEG-mal if membranes are disrupted by SDS. As shown in Fig. 4.6A, with PEG-Mal treatment, there is a distinct band shift observed in the WT AtsR protein that is not present in the control condition. This band shift is more pronounced under denaturing condition as C195 that is supposedly hidden in the membrane is exposed and therefore two shifts occur. As expected the $AtsR_{cysless}$ does not generate any band shift in any treatment. Interestingly, residues M21C and G187C were not labeled by PEG-mal in the absence of SDS suggesting that these residues were blocked by NEM and therefore prevented from becoming accessible to PEG-mal (Fig. 4.6B). This indicates that M21C

and G187C are buried in the membrane as predicted by the *in silico* topological model. Labeling of membranes with PEG-Mal under non-denaturing or SDS-denaturing conditions resulted in a band shift for AtsR protein carrying novel cysteines at V38C, G41C, G49C and G120C indicating their accessibility to PEG-mal suggesting a periplasmic location for these residues (Fig. 4.6C).

Furthermore, we investigated the residues around the membrane hinges by introducing novel cysteines at positions V14C, A182C, G180C and M205C. None of the introduced cysteines were labeled when treated with NEM; however, they all have a band shift when treated with PEG-mal under denaturing condition in the presence of SDS (Fig.4.7A). This might be due to close proximity of these residues to the membrane bilayer borders where are accessible to NEM but not to PEG-mal in non-denaturing condition. Finally, after treatment of membranes novel cysteines at residues A328C and G484C were modified by PEG-mal in the presence of SDS but did not show a band shift when pretreated with NEM suggesting they are on the cytosolic face of the membrane (Fig. 4.7B). This observation corresponds to the PhoA/LacZ reporter fusions. The Western blotting analysis suggested all single cysteine AtsR proteins were expressed at levels comparable with that of the WT and localized in the membrane fraction with the expected molecular mass of 68 KDa (Fig. 4.6&7). Therefore, we concluded that the effect of cysteine residue replacement in each mutant is tolerated and does not cause structural changes in the protein. Collectively, our data from LacZ/PhoA reporter fusions combined with PEG-mal experiments support the predicted topology of AtsR. However, the predicted TM borders shown in Fig. 4.1B need to be refined based on our PEGylation data as follows: the first TM starts from V14 and the second TM is stretched from Q180 to M205.

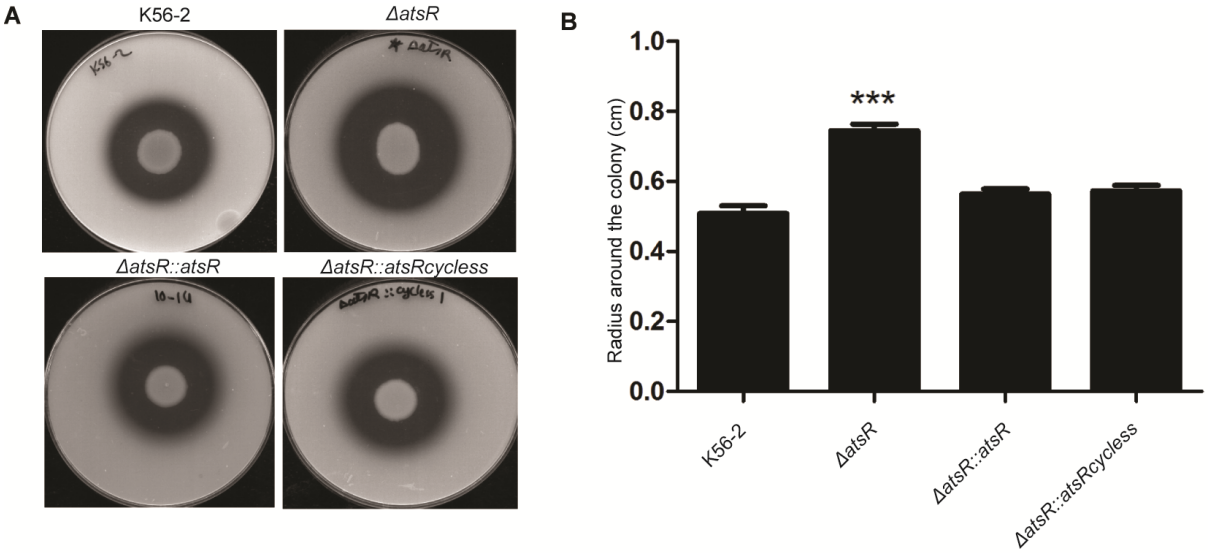


Figure 4.4: Proteolytic activity of K56-2, K56-2Δ*atsR*, K56-2Δ*atsR::atsR*, K56-2Δ*atsR::atsR-cycless* mutant strains on milk agar plates (A) and quantification of protease activity of the same strains (B). Proteolysis was tested on D-BHI milk agar plates. The plates shown are representatives of three experiments performed in triplicate. Zones of clearing around the colonies were measured at 48 h of incubation at 37 °C. Values are average radius ± standard deviation in centimeters of three experiments performed in triplicate. Statistical analysis was performed by paired *t*-test using two-tailed P-values. Significant differences in comparison with *B. cenocepacia* parental strain (K56-2) as control are indicated by *** ($P < 0.005$).

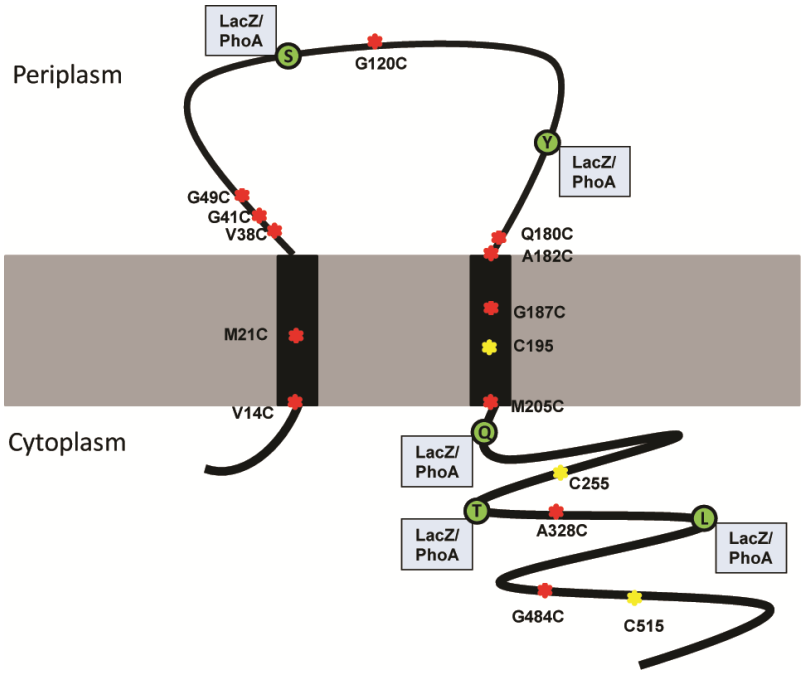


Figure 4.5: Schematic topology of AtsR predicted by *in silico* programs and validated by reporter fusion studies. Yellow stars represent native cysteine residues in AtsR protein and red stars indicate point mutations to introduce novel cysteines. Green circles correspond to the location of LacZ or PhoA fusion. Numbers indicate corresponding amino acid residues. Abbreviations: V: valine, M: metionine, G: glycine, S: serine, Q: glutamine, A: alanine, Y: tyrosine, C:cysteine.

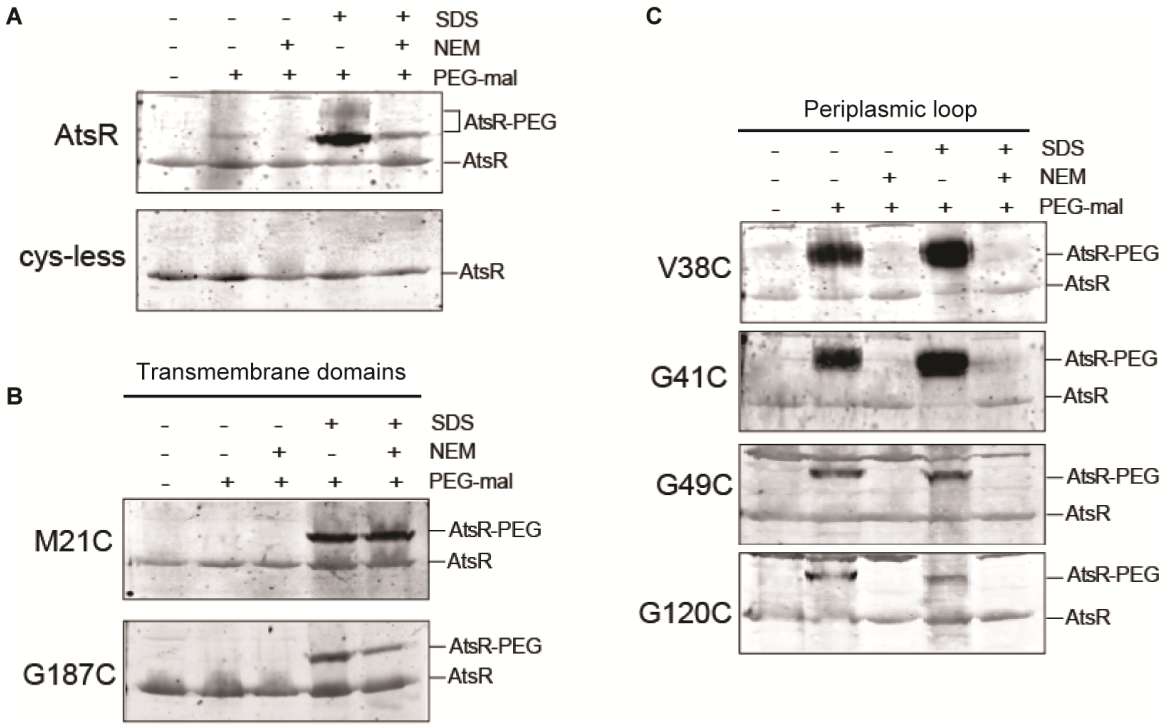


Figure 4.6: PEG-mal labeling of $AtsR_{cysless}$ proteins carrying cysteine replacements with amino acids in periplasmic and TM domains. A, B, and C) Cells were harvested and total membrane fractions were isolated by centrifugation and resuspended in HEPES/MgCl2 buffer. Membrane was incubated with buffer alone or 1 mM PEG-mal with or without 1% SDS for 1 h at room temperature or pretreated with NEM for 30 min at room temperature. Proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes, followed by immunoblot with anti-His antibodies.

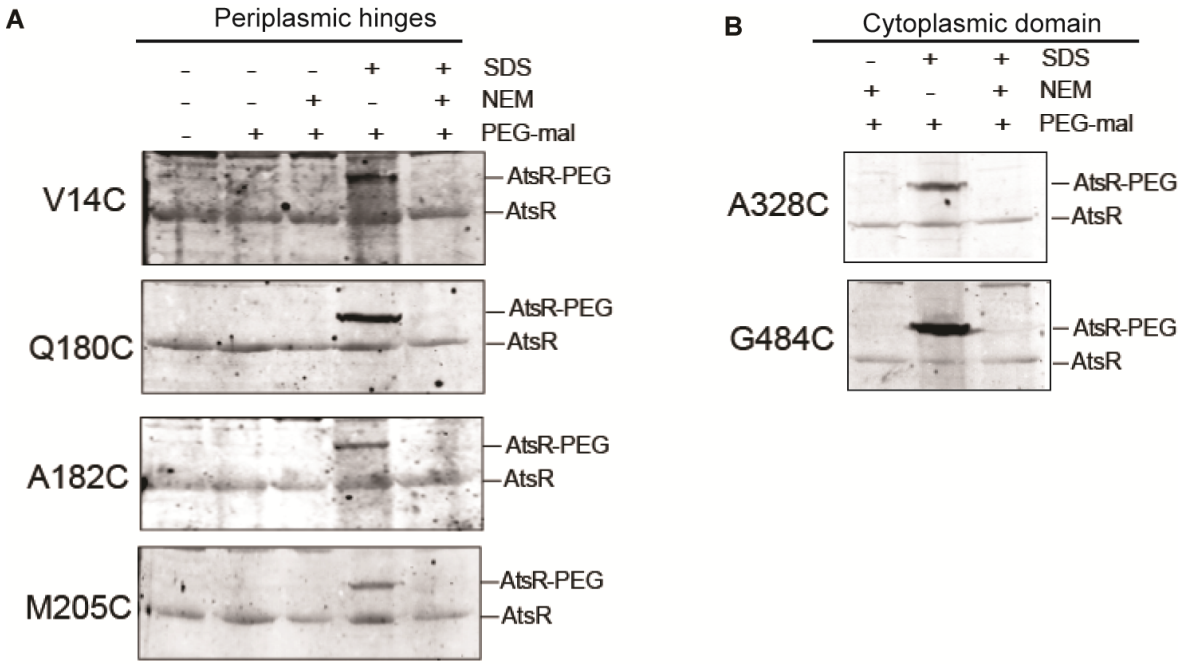


Figure 4.7: PEG-mal labeling in total membrane preparations expressing $AtsR_{cysless}$ proteins with cysteine replacement with amino acids in periplasmic hinges and cytoplasmic domain. A and B) Cells were harvested and crude membrane fractions were isolated by ultracentrifugation and resuspended in HEPES/MgCl2 buffer. Membrane was incubated with buffer alone or 1 mM PEG-mal with or without 1% SDS for 1 h at room temperature or pretreated with NEM for 30 min at room temperature. Proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes, followed by immunoblot with anti-His antibodies.

4.3.4 Analysis of functional residues within periplasmic loop

While the histidine kinase and ATPase domains of AtsR have typical residues of TCS proteins, the predicted periplasmic region is unique to *Burkholderia* species. Examination of the predicted secondary structure of the periplasmic loop of AtsR using the PSI-PRED program (18) revealed the presence of five α-helices (Fig. 4.8), supporting the idea that the periplasmic loop of AtsR could form a pocket to interact with a signalling molecule itself. Therefore, we investigated whether replacing conserved amino acids within predicted periplasmic loop could affect the function of AtsR.

To identify functional residues within the periplasmic loop, the sequence was aligned with corresponding loop of AtsR homologues proteins within *Burkholderia cepacia* complex and 10 conserved residues including Q56, Y61 and D65 residing in the first α-helix, F79 and D80 from the second α-helix, S103, W105 and K127 from the third α-helix, and N155 and R162 from the fourth α-helix were replaced by alanine using site-directed mutagenesis and then introduced into *K56-2*Δ*atsR* mutant strain of *B. cenocepacia* at chromosomal level. These mutant derivatives of AtsR were tested and their protease activity was measured. As shown in Fig. 4.9, results are statistically significant only for mutants carrying Q56A, D65A, S103A and R162A substitution compared to the WT indicating the importance of these residues. To make sure that such phenotype is robust, proteolytic activity of three different colonies for each mutant strain were quantified all exhibiting reproducible data. R162 and D65 are charged amino acids, whereas S103 and Q56 are non-charged polar amino acids. To investigate whether the charge or the nature of the amino acid is important at these positions we replaced D65 with glutamic acid (E) and K, S103 with threonine (T) and valine (V), and Q56 with

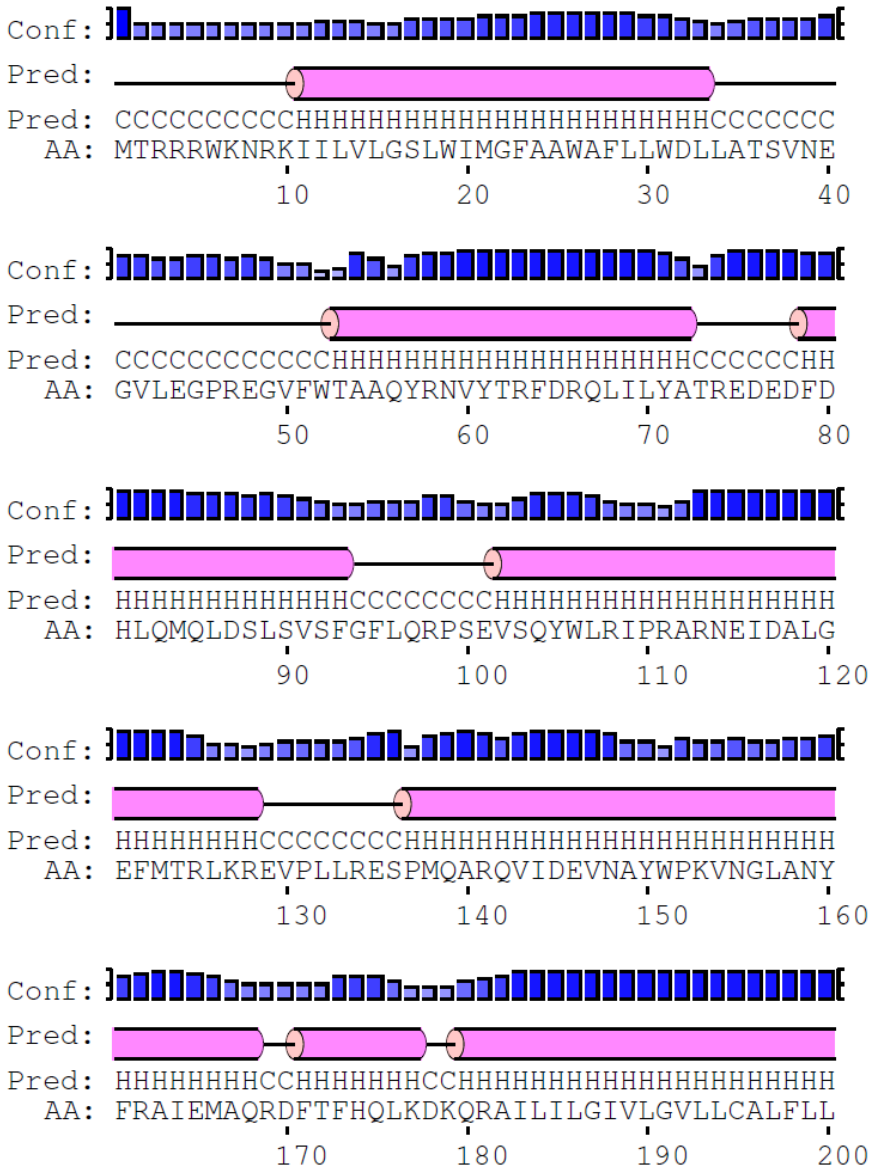


Figure 4.8: PSIRED model representing AtsR periplasmic loop (starting from amino acid 33 to 180) with 5 predicted α-helices. Pred represents graphical representation of the structural prediction that includes *α*-helical regions (pink cylinders) and coils (black lines). Blue boxes indicate the confidence of prediction for each residue as indicated by the height of each rectangle. The amino acid numbers correspond to the position in the wild-type AtsR protein.

asparagine (N) and R. Interestingly, neither S103 replaced with V nor Q56 replaced with R could restore the protease activity of *ΔatsR* phenotype indicating that the nature of amino acid is important at these positions since replacing the hydroxylic S with aliphatic V at position 103 and Q with positively charged R at position 56 might have changed the nature and structure of the protein resulting in a non-functional AtsR. In contrast, replacement of Q56 with N and S103 with T restored the protease activity of mutant strain to WT levels (Fig. 4.9B). Furthermore, D65 substitution with opposite or same charge amino acids does not complement the *ΔatsR* phenotype. This result together with the increased protease activity observed in D65 to Ala substitution indicates that the nature and size of the amino acid is important other than charge. Nevertheless, together these data suggest that S103, D65 and Q59 are indispensable for AtsR function by maintaining the tertiary structure of the protein making contacts with the signal or interacting with auxiliary periplasmic proteins to sense the environmental cues and initiate the AtsR signaling pathway.

4.4 Conclusion

This study provides evidence of the topology of AtsR, demonstrating the agreement between predicted topology with our experimental analysis. A combination of fusion reporters to the LacZ and PhoA proteins and labeling of cysteine replacement mutants with sulfhydryl reactive reagents allowed us to experimentally validate topological model for AtsR that is composed of a short N-terminal cytoplasmic region, a periplasmic loop flanked by 2 TM helices, and a C-terminal cytoplasmic domain. A cys-less version of AtsR localized to the membrane and remained functional indicating that the native cysteines are dispensable for AtsR function. Furthermore, we identified 4 residues in the

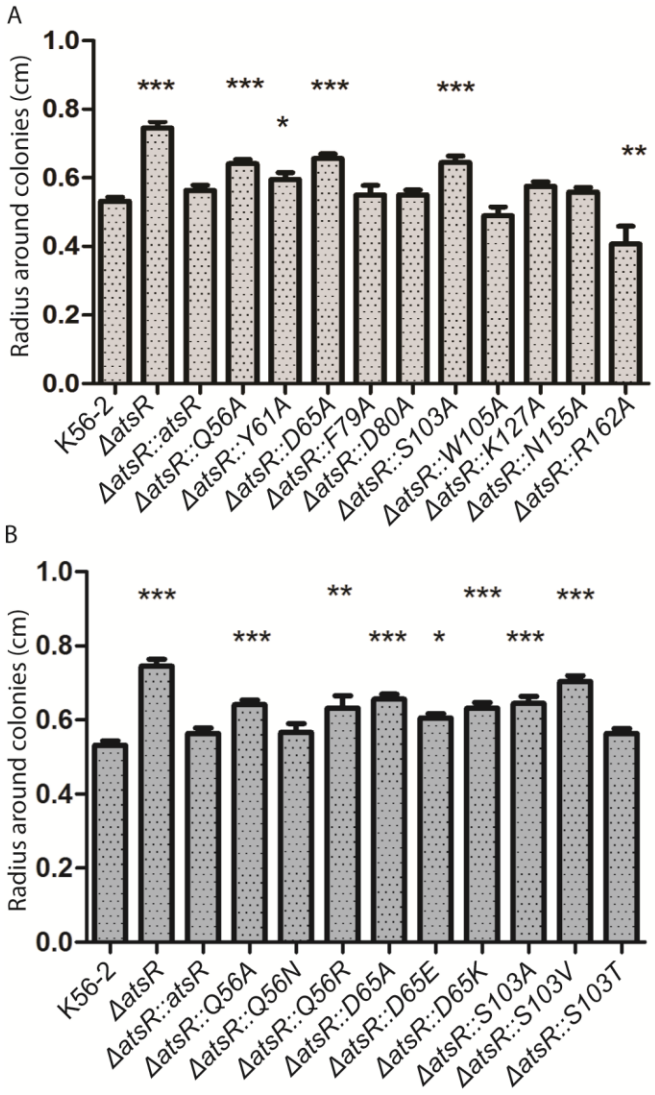


Figure 4.9: Protease activity of K56-2Δ*atsR* strains complemented with AtsR mutants at chromosomal level. Quantifications of proteolytic activity of mutants with alanine substitution (A) or with the same or opposite charged or uncharged amino acids (B). Results obtained based on 3 individual experiments in triplicate for each mutant. Statistical analysis was performed by paired *t*-test using two-tailed P-values for mutants in comparison with *B. cenocepacia* parental strain (K56-2) (*** corresponds to $p<0.0005$, ** corresponds to $p<0.005$ and * corresponds to $p<0.05$). Similar stats were obtained with comparing mutants with K56-2Δ*atsR::atsR* strain.

periplasmic region of AtsR that are absolutely required for the functionality of the protein. While the histidine kinase domain and the ATPase domain at the cytoplasmic C-terminus are highly conserved features in all histidine kinases, the predicted periplasmic domains are variable elements suggesting their potential role in signal perception. Based on PSIPRED modeling system the periplasmic region of AtsR has five α-helices (Fig. 4.8) and is likely involved in sensing environmental cues. Further studies will provide additional insights into the structural or functional significance of these residues within the periplasmic loop regarding facilitating an interaction with a signalling molecule or maintaining the binding pocket.

4.5 Chapter 4 references

1. **Parkinson JS, Kofoid EC.** 1992. Communication modules in bacterial signaling proteins. Annu Rev Genet **26:**71-112.

2. **Galperin MY.** 2010. Diversity of structure and function of response regulator output domains. Current Opinion in Microbiology **13:**150-159.

3. **Aubert DF, Flannagan RS, Valvano MA.** 2008. A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in *Burkholderia cenocepacia*. Infect Immun **76:**1979-1991.

4. **Aubert DF, O'Grady EP, Hamad MA, Sokol PA, Valvano MA.** 2013. The *Burkholderia cenocepacia* sensor kinase hybrid AtsR is a global regulator modulating quorum-sensing signalling. Environ Microbiol **15:**372-385.

5. **Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama A, Murakami H, Hosoyama A, Mizutani-Ui Y, Takahashi NK, Sawano T, Inoue R, Kaito C, Sekimizu K, Hirakawa H, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita A, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K.** 2001. Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. Lancet **357:**1225-1240.

6. **Khodai-Kalaki M, Aubert DF, Valvano MA.** 2013. Characterization of the AtsR hybrid sensor kinase phosphorelay pathway and identification of its response regulator in *Burkholderia cenocepacia*. J Biol Chem **288:**30473-30484.

7. **Marolda CL, Li B, Lung M, Yang M, Hanuszkiewicz A, Rosales AR, Valvano MA.** 2010. Membrane topology and identification of critical amino acid residues in the Wzx O-antigen translocase from *Escherichia coli* O157:H4. J Bacteriol **192:**6160-6171.

8. **Hamad MA, Skeldon AM, Valvano MA.** 2010. Construction of aminoglycoside-sensitive *Burkholderia cenocepacia* strains for use in studies of intracellular bacteria with the gentamicin protection assay. Appl Environ Microbiol **76:**3170-3176.

9. **Figurski DH, Helinski DR.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci U S A **76:**1648-1652.

10. **Craig FF, Coote JG, Parton R, Freer JH, Gilmour NJ.** 1989. A plasmid which can be transferred between *Escherichia coli* and *Pasteurella haemolytica* by electroporation and conjugation. J Gen Microbiol **135:**2885-2890.

11. **Manoil C.** 1991. Analysis of membrane protein topology using *alkaline phosphatase* and *beta-galactosidase* gene fusions. Methods Cell Biol **34:**61-75.

12. **Tavares-Carreon F, Patel KB, Valvano MA.** 2015. *Burkholderia cenocepacia* and *Salmonella enterica* ArnT proteins that transfer 4-amino-4-deoxy-l-arabinose to lipopolysaccharide share membrane topology and functional amino acids. Sci Rep **5:**10773.

13. **Tsirigos KD, Peters C, Shu N, Kall L, Elofsson A.** 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res doi:10.1093/nar/gkv485.

14. **Derman AI, Beckwith J.** 1991. *Escherichia coli* alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm. J Bacteriol **173:**7719-7722.

15. **Furlong SE, Ford A, Albarnez-Rodriguez L, Valvano MA.** 2015. Topological analysis of the *Escherichia coli* WcaJ protein reveals a new conserved configuration for the polyisoprenyl-phosphate hexose-1-phosphate transferase family. Sci Rep **5:**9178.

16. **Koch S, Fritsch MJ, Buchanan G, Palmer T.** 2012. *Escherichia coli* TatA and TatB proteins have N-out, C-in topology in intact cells. J Biol Chem **287:**14420-14431.

17. **Nasie I, Steiner-Mordoch S, Schuldiner S.** 2013. Topology determination of untagged membrane proteins. Methods Mol Biol **1033:**121-130.

18. **McGuffin LJ, Bryson K, Jones DT.** 2000. The PSIPRED protein structure prediction server. Bioinformatics **16:**404-405.

Chapter 5

Burkholderia cenocepacia lipopolysaccharide modification and flagellin glycosylation affect virulence but not innate immune recognition in plants

Maryam Khodai-Kalaki, Angel Andrade, Yasmine Fathy Mohamed, Miguel A. Valvano

Aversion of this chapter has been published:

Maryam Khodai-Kalaki, Angel Andrade, Yasmine Fathy Mohamed, Miguel A. Valvano (2015). mBio 6(3):e00679-15. Doi:10.1128/mbio.00679-15

5.1 Introduction

Burkholderia cepacia complex (Bcc) denotes a group of ubiquitous Gram-negative bacterial species isolated from water, soil, plants, insects, industrial settings and hospital environments (1). Some Bcc strains have beneficial traits, as they utilize complex carbon sources to degrade toxic compounds in pesticides and herbicides, while others serve as bioremediation agents (2). However, Bcc strains are opportunistic pathogens causing serious infections in immunocompromised humans, including chronic lung infection in cystic fibrosis (CF) patients (3). The most common Bcc clinical isolates are *B. cenocepacia* and *B. multivorans*. In particular, *B. cenocepacia* has gained notoriety for its ability to cause lethal necrotizing pneumonia and its transmissibility among CF patients (3).

The opportunistic nature of infections by *B. cenocepacia* makes it difficult to study the pathogenic mechanisms of this bacterium, since "virulence" depends on the host's context and its innate susceptibility to infection. Several infection models for *B. cenocepacia* have been established, which include mice and rats (4), zebrafish embryo (5), the nematode *Caenorhabditis elegans* (6), *Galleria mellonella* moth larvae (7, 8), and plants like alfalfa (4) and duckweed (9). Plants have evolved sophisticated mechanisms to perceive and respond to nearly constant attacks from potential pathogens (10). These mechanisms involve passive protection, such as that provided by the waxy cuticular skin layer, and innate immune responses. As with the innate immune system in mammals and insects, the plant immune system also perceives and responds to various elicitor molecules, which are conserved microbial structures referred to as microbe- or pathogen-associated molecular patterns (PAMPs). These include bacterial flagellin, lipopolysaccharide (LPS) and

lipoteichoic acid, and fungal chitin and ergosterol (11). LPS consists of lipid A, embedded in the outer leaflet of the outer membrane, which is linked to core oligosaccharide and polymeric O antigen (12), while flagellin is the building block of the flagellar filament essential for bacterial motility. LPS and flagellin are typical PAMPs that stimulate plant cells and tissues to generate reactive oxygen species (ROS), alkalinisation of the extracellular medium, callose deposition, and nitric oxide (NO) burst (13-15). These responses lead to programmed cell death localized to the tissue surrounding the infection site, which is known as the hypersensitive response (11, 16).

Arabidopsis thaliana is an established model organism for plant biology, but also a tool to understand molecular mechanisms of human diseases (17, 18), particularly concerning the initial stages of infection (10, 19). Previous work by others reported a rapid burst of NO after treatment of *A. thaliana* cultured cells and leaves with *B. cepacia* lipid A preparations, which was associated with the transcriptional activation of defense-related genes (20). Further, treatment with *B. cepacia* purified lipid A and core oligosaccharide LPS components resulted in different plant transcriptional responses (21), suggesting that plant cells perceive LPS components differently. Lipid A structural modifications alter recognition and responses by the innate immune system in mammals (22) and in plants (23). LPS contributes to reduce the permeability of the outer membrane of Gram-negative bacteria, acting as a barrier for plant-derived antimicrobial substances (24). *B. cenocepacia* exhibits extraordinary intrinsic resistance to antimicrobial peptides and other antibiotics (25). We have shown that the *B. cenocepacia* lipid A and one of the sugars of the inner core oligosaccharide are modified by the incorporation of L-Ara4N, and this modification is not only essential for bacterial survival (26) but is also the major determinant of resistance to antimicrobial peptides (27). The essentiality of the LPS

modification by L-Ara4N suggests the possibility that this amino sugar could also contribute to modulate innate immune responses.

Flagellin is a major activator of innate immunity in animals, where is recognized by the Toll-like receptor 5 (TLR5) (28), and in plants (29, 30). Recognition of flagellin by plants activates disease resistance mechanisms through an oxidative burst during early stages of infection (13, 15). The *A. thaliana* Flagellin Sensing2 (FLS2) protein is a PAMP receptor kinase with an extracellular domain containing 28 leucine-rich repeats, a transmembrane domain, and an intracellular serine/threonine kinase domain (31, 32). These interactions activate production of ROS , triggering a complex defense response including induction of pathogenesis-related genes (15). We have recently shown that the *B. cenocepacia* flagellin is glycosylated and its glycosylation status modulates innate immune responses in mammalian cells (33), but the role of flagellin glycosylation in plant immune responses to *B. cenocepacia* infection has not been examined.

In this work, we investigated the pathogenicity of *B. cenocepacia* in *A. thaliana* and evaluated the significance of LPS modification with L-Ara4N and flagellin glycosylation in virulence and in *Arabidopsis-B. cenocepacia* interactions by studying innate immune responses by wild type and mutant plants. Further, we compared bacterial virulence *in planta* and in the *Galleria mellonella* infection model to clarify the global role of flagellin glycosylation and LPS modification on pathogenesis. We demonstrate that flagellin glycosylation and LPS modification with L-Ara4N play a significant role on bacterial survival upon infection, but do not alter the perception of these molecules by the plant innate immune receptors, indicating these modifications are only critical to establish infection.

5.2 Materials and methods

5.2.1 Plants and growth conditions

Seeds of *Arabidopsis thaliana*, ecotype Columbia (Col-0), *Atnoa1 (*SALK number: CS6511), *fls2* (SALK number: 121477), and *dnd1-1* (SALK number: 066908C) mutants were surface-sterilized by sequential soaking in 75% ethanol for 1 min, rinsing three times with sterile water, soaking in 20% bleach for 15 min, and rinsing five times with sterile water. Sterilized seeds were kept at 4 °C in the dark for 3 days for vernalization and were planted in magenta boxes containing Murashige & Skoog (MS) basal salt mixture (Sigma-Aldrich) supplemented with 1% sucrose (W/V) and 0.8% agar, and placed in a growth chamber at 22 °C with a light intensity of 80 µE $m^{-2}s^{-1}$. Seedlings were grown under 16 h light and 8 h darkness for 21 days. Mutants were obtained from the Arabidopsis Biological Resource Center (ABRC). Tobacco plants (*Nicotiana benthamiana*) were grown at 26 °C with a 16-h photoperiod.

5.2.2 Bacterial strains and growth conditions

B. cenocepacia K56-2 and its isogenic mutants (Table 5.1) were cultured at 37 °C in Luria Bertani broth overnight. *B. cenocepacia* K56-2 is a clinical isolate that belongs to the ET-12 epidemic strain lineage (57). Strain MSS25 carries a polar insertion in the *fliC* gene, which prevents the synthesis of the flagellin filament and also likely inactivates the downstream genes *fliD1* and *fliT*, encoding a flagellar hook and an accessory protein, respectively (35). To analyse growth rate of mutants, overnight cultures were diluted into fresh medium at a starting optical density (OD) at 600 nm of 0.05. Samples were then aliquoted in a 100-well honeycomb microtitre plate and growth rates were followed by

Table 5.1: Bacterial strains used in this study

Strain	Characteristics	Source and/or reference
K56-2	*B. cenocepacia* ET12 clone related to J2315, CF clinical isolate	BCRRC[a]
K56-2 Δ*atsR*	K56-2 carrying an unmarked deletion of the *atsR*gene	(37)
MH55	K56-2, Δ*arnT-arnBC*$^{+}$ *lptG*$_{D31H}$ (*lptG*^{S})	(27)
ΔBCAL0111	K56-2, unmarked deletion of *flmQ* (BCAL0111)	(33)
MSS25	K56-2, *fliC*::pSM62, insertional inactivation of *fliC*	(35)

[a] BCRRC, *B. cepacia* Research and Referral Repository for Canadian CF Clinics.

measuring the OD_{600} every hour under continuous shaking for 24 h in a Bioscreen C automated growth curve analyser (MTX Lab Systems, Vienna, VA, USA). All bacteria grew at similar rates in Luria Bertani medium (Fig. 5.9B).

5.2.3 Plant inoculation methods

Five to 6 week-old *A. thaliana* or *Nicotiana benthamiana* plants were used for initial inoculation experiments. For soil drainage infection, plants were watered with bacterial suspensions in MES (4-morpholineethanesulfonic acid sodium salt) buffer. For surface spraying method bacteria from overnight cultures were suspended in 10 mM MES buffer at different CFU. For inoculation by the syringe-infiltration method, bacteria were infiltrated into the leaves using a needleless syringe. For wound inoculation seedlings were wounded at the leaf surface by scratching. Wounded seedlings were immediately inoculated with bacterial suspensions. MS-agar plates containing inoculated seedlings were sealed with parafilm and placed in growth chambers for 7 days. Plants were monitored for disease symptoms daily. At day 7, three leaves from each plantlet were excised, washed with 5% bleach for 1 min and rinsed with sterile water. The leaf was blotted dry on sterile filter paper and imprinted on LB agar plates to determine if there were any bacteria on the surface of the leaves. The imprinted plates were incubated at 37 °C for 24 h. The leaves were then weighed and macerated in 500 µl PBS (phosphate-buffered saline) with a micro-pestle, serially diluted and plated on LB agar plates in triplicates. Only leaf samples that did not show any bacterial growth on the imprinted plates were counted to avoid counting contaminating bacteria from leaf surfaces. Other mutants were similarly infiltrated into leaves and HR symptoms were visualized by bleaching the leaves in ethanol (Fig. 5.2C). Ethanol bleaching was used to provide

enhanced visualization of HR development in leaves and facilitate comparison with the parental strain.

5.2.4 RNA isolation, cDNA synthesis and RT-PCR

Total RNA was extracted from frozen leaves of *A. thaliana* using the SV Total RNA Isolation System (Promega, USA) and 1 µg of RNA was reverse transcribed using Quantitect Reverse Transcription kit (Qiagen, USA). The coding sequence of *PR-1* and *ACTIN-2* genes, available through GenBank, was used to design the following primers: Forward 5'-GATGTGCCAAAGTGAGGTG- 3' and reverse 5'-CTGATACATATACACGTCC-3' to amplify *PR-1* and forward 5'-TGCTCTTCCTCATGCTAT-3' and reverse 5'-ATCCTCCGATCCAGACACTG-3́ to amplify *ACTIN*. PCR was carried out with an initial denaturation step at 94 ^{o}C for 4 min, followed by 30 cycles of denaturation (30 sec at 94 ^{o}C) for *PR-1* and 21 cycles for *ACTIN*, annealing (30 sec at 53 ^{o}C) and extension (1 min at 72 ^{o}C). After the last cycle, a final extension was carried out for 5 min at 72 ^{o}C. PCR products were visualized on 1% agarose gels using UV light in a Gel-Doc system (Alpha Innotech, USA). Relative mRNA levels of *PR-1* were normalized relative to untreated control and internal control (*ACTIN-2*) by densitometry using ImagJ.

5.2.5 LPS extraction, mass spectrometry, and flagellin purification

LPS was extracted from equal biomass using proteinase K method as previously described (58). Extracted LPS was dissolved at 1 mg/ml in water containing 2.5 mM $MgCl_2$ plus 1 mM $CaCl_2$ and shaken for 3 h on a rotary shaker. If not mentioned otherwise, experiments were performed with K56-2 LPS or with buffer A containing 1.0

mM $CaCl_2$ and 2.5 mM $MgCl_2$ (pH 7.6), as a control. Flagellin purification was performed as described earlier (33). Lipid A of strains MH55 and K56-2 (Table 5.1) was examined by mass spectrometry (MS) using matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF). Cells were grown overnight in Luria-Bertani broth, centrifuged (10,000 ×*g*), washed twice with phosphate buffer (10 mM Na_2HPO_4, 1.7 mM KH_2PO_4), and freeze-dried. Lipid A was extracted and desalted, as previously described (59). For MS analysis, dihydroxybenzoic acid (DHB) (Sigma Chemical Col, St. Louis, MO) was used as a matrix and prepared to saturation in acetonitrile: 0.1% triflouroacetic acid (1:2 v/v). A 2-µl sample was loaded on the target, dried, covered by 1 µl of DHB, and inserted in a Bruker Autoflex MALDI-TOF spectrometer. Data were acquired and analyzed using the Flex Analysis software.

5.2.6 *In vivo* NO and ROS analysis

NO was measured in guard cells of leaf epidermal peels prepared from leaves of 3-week-old plants (grown as described above) using diaminofluorescein diacetate (DAF-2DA) fluorescence as described (60) with the following modifications. Leaves were submerged in buffer (pH 5.7) containing 5 mM MES-KOH, 0.25 mM KCl, and 1 mM $CaCl_2$, with 5 µM DAF-2DA and transferred either to reaction buffer containing 100 µg/mL LPS or pre-incubated for 10 min in reaction buffer containing 50 µM sodium nitroprusside (SNP) or 200 µM N^G-nitro-L-Arg-methyl ester (L-NAME) and then for 20 min with LPS. The epidermal peels were placed underneath a cover slip on a microscope slide with several drops of reaction mix. In a similar way ROS was detected by 2 µM 2, 7-dichlorofluorescein diacetate (DCF) and flagellin treatment following the incubation of leaves in MES/KOH buffer for 2 h.

5.2.7 Microscopy

Fluorescence and phase-contrast images were acquired using a QImaging Retiga-SRV camera on an Axioscope 2 (Carl Zeiss).

5.2.8 Fluorometric Quantification of NO and ROS

To monitor NO accumulation in *Arabidopsis* leaf peels, the DAF-2DA fluorescence was measured at excitation of 480 nm and emission of 521 nm, using a Cary Eclipse fluorescence spectrophotometer (Varian, Inc., Mississauga, Ontario, Canada). The plate was rocked for 20 sec before measuring. Similarly, for determining accumulation of ROS, DCF was added to a final concentration of 2 µM. Fluorescence was measured at excitation of 480 nm and emission of 521 nm in 96-well white plates. Background fluorescence of each probe in buffer control was subtracted. Autofluorescence of the buffer, without adding the probes, was measured and corrected by subtraction from the fluorescence signals. The suspensions were protected from light throughout the assays to avoid photo-oxidation.

5.2.9 Bacterial Swarming and Swimming Motility Assays

Swarming motility assays were performed as described previously (37). Assays were done in triplicate and repeated independently three times. Bacterial swimming motility was analyzed on soft agar plates (1% Bacto tryptone in 0.3% agar). The OD_{600} of overnight cultures was adjusted to 1 and 2 µl of culture were inoculated in the centre of agar plate. Diameter of growth zone was measured after 24 h of incubation at 37 °C.

5.2.10 *Galleria mellonella* infection model

G. mellonella larvae were obtained from Recorp and stored in the dark at 4 °C. Injections of 10 µl containing 1×10^4 CFU diluted in 10 mM $MgSO_4 \cdot 7H_2O$ supplemented with 100 µg ampicillin μl^{-1} were given to the larvae into the haemocoel through the hindmost right proleg. Infected larvae were incubated at 30 °C and larval survival was monitored up to 75 h as judged based on lack of movement in response to stimuli. Control larvae were injected with 10 µl of the same buffer without bacteria. Ten larvae were used for each condition and the experiment was repeated on three independent occasions.

5.3 Results and discussion

5.3.1 Pathogenicity of *B. cenocepacia* in *A. thaliana* and tobacco seedlings

We investigated the ability of *B. cenocepacia* K56-2, a CF clinical isolate, to induce HR in *A. thaliana* and tobacco. Seedlings from both species were inoculated by different methods with 10^1 to 10^8 bacteria and examined for symptoms over one week. A simple infection model was developed in which leaves were wounded, and then inoculated with dilute suspensions of *B. cenocepacia*. Disease was scored for every seedling on an index of 1 to 6 based on the extent of HR symptoms and tissue necrosis using a standard HR scoring system (34). Depending on the bacterial dose, infection of seedlings via wound inoculation or syringe infiltration led to necrosis (brown areas on leaves at higher doses) or chlorosis (yellowing of the leaf tissue), which were clearly visible after day 4 post infection. In contrast, infections by soil drainage or surface spray gave few symptoms even at day 7-post infection (Fig. 5.1A). There was a direct correlation between inoculum size and HR symptoms (Fig. 5.1B). Similar results were observed in tobacco plants, where infection *via* roots or leaves spraying did not cause distinguishable symptoms in plantlets (Fig. 5.1C). Since several mutants in innate immunity pathways and pathogen

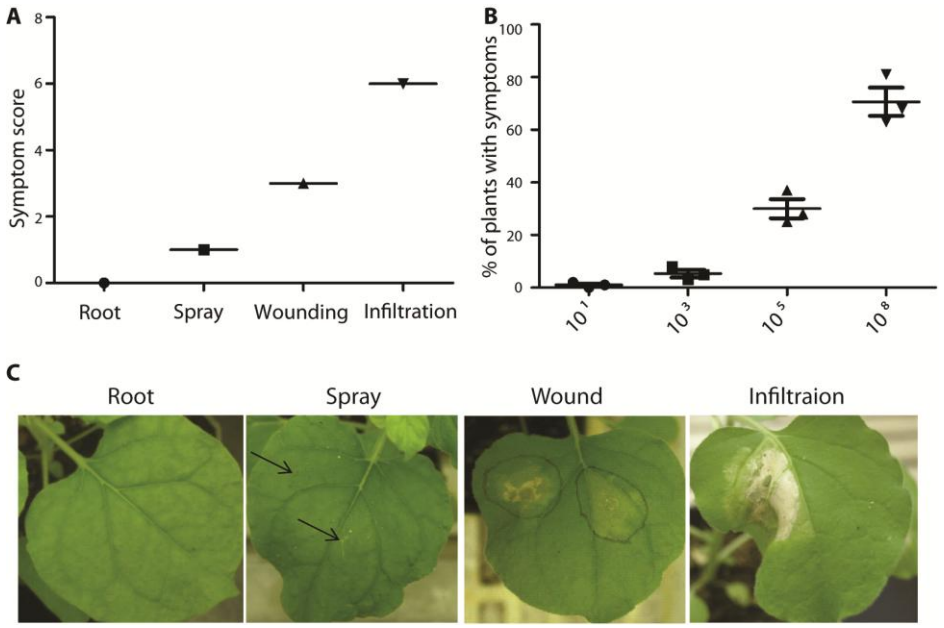


Figure 5.1: HR in plant leaves using various inoculation methods. A) Quantitative scoring of HR in leaves of wild-type *Arabidopsis* inoculated with K56-2 by different inoculation methods. Symptom scores assigned using an index from 1-6 based on the extent of tissue collapse and color change where: 0 is no symptom; 1, random yellow or dark spots; 2, chlorosis over the inoculation site; 3, chlorosis and mild tissue collapse at and around the inoculation site; 4, more than 50 % collapse at and around the inoculation site; 5, 100% collapse at the inoculation site; 6, 100% collapse beyond the inoculation site. The average score was calculated based on a minimum of 45 leaves per treatment cumulative from 3 individual repeats. B) Effect of inoculum size on the incidence of infection. Values indicate percent ± SE of seedlings showing disease symptoms at 7 days post infection inoculated with doses ranging from 10^1 to 10^8 CFU; 15 plants per infection dose were tested. Data represent the average of three experiments. C) Representative images of 5-week-old tobacco plants inoculated with a dose of 10^8 CFU of *B. cenocepacia* by different inoculation methods. Arrows indicate mild symptoms that are barely visible resulted from spraying bacteria on leaves.

receptors are available for *A. thaliana* ecotype Colombia Col-0, we used only seedlings from this species for the remainder of this study, which were inoculated with 10^8 CFU via wound inoculation or syringe infiltration.

5.3.2 L-Ara4N modification of LPS and flagellin glycosylation are required to cause pathology in *A. thaliana*

B. cenocepacia requires L-Ara4N modification of LPS for viability and resistance to antimicrobial peptides (27). However, a suppressor mutation in the *lptG* gene, which encodes an essential protein for the export of LPS to the outer membrane, restores viability in the absence of L-Ara4N LPS modification (27). We therefore investigated if lack of L-Ara4N in the LPS of this mutant (MH55, Δ*arnT lptG*$_{D31H}$) could affect *B. cenocepacia* survival in *A. thaliana*. Infection of *A. thaliana* seedlings with Δ*arnT lptG*$_{D31H}$ caused mild symptoms of chlorosis in 27% of inoculated plants (Fig. 5.2A&C); very few bacteria were recovered from leaves, as compared to plants infected with the K56-2 parental strain (Fig. 5.2B). These data demonstrate that L-Ara4N modification of LPS is critical for *B. cenocepacia* pathogenicity. The poor infectivity of the suppressor mutant is most likely due to the lack of L-Ara4N in the LPS of Δ*arnT lptG*$_{D31H}$, as revealed by mass spectrometry of purified lipid A (Fig. 5.4), which makes it extremely sensitive to naturally produced antimicrobial peptides in plants, as we previously demonstrated with polymyxin B (27).

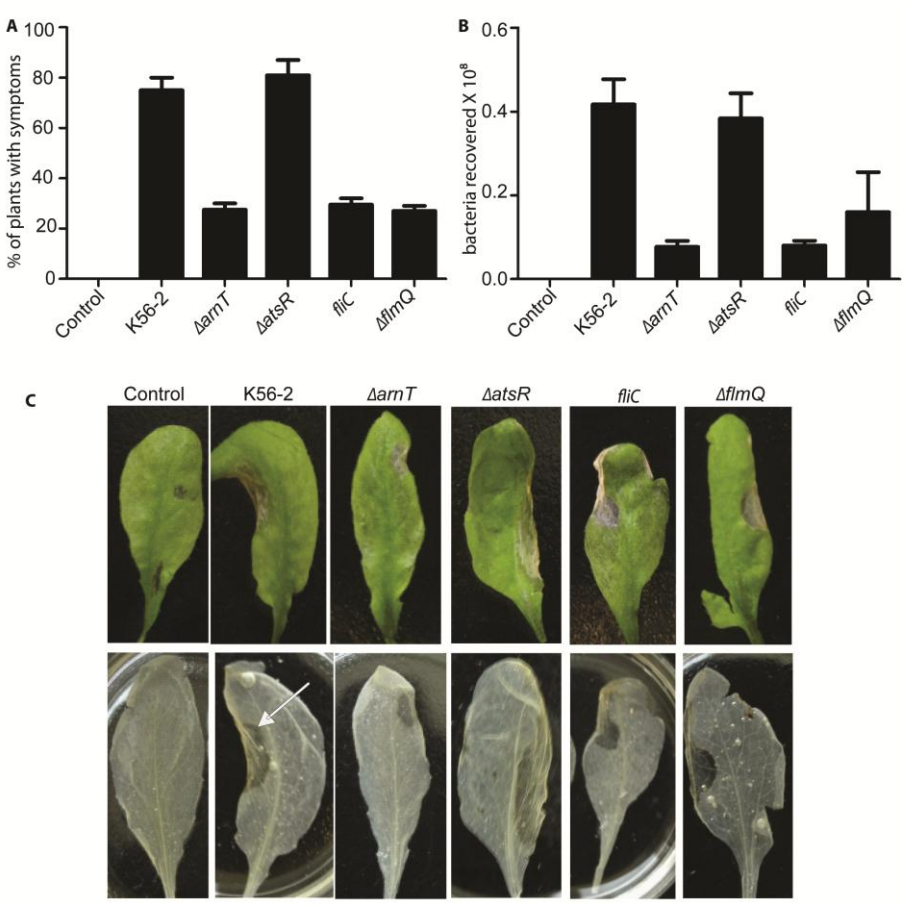


Figure 5.2: Infection and replication of *B. cenocepacia* strains in *Arabidopsis* seedlings. A) Values show the percentages ± SE of the seedlings showing disease symptoms visible on day 7 post infection after wound inoculation with 10^8 CFU of *B. cenocepacia* strains; 15 plants per infection dose were tested. Data are the average of three repeats. B) Graph represents bacterial counts on day 7 post infection from *Arabidopsis* leaves infected with *B. cenocepacia* by wound inoculation. C) Photographs of representative leaves (from three replicate treatments) from *Arabidopsis* plants inoculated with *B. cenocepacia* strains by infiltration (Upper row). Leaves were removed from plants and bleached in ethanol (Lower row). An arrow indicates the inoculated region, which is more transparent and flattened undergoing tissue collapse compared with the rest of the leaf and wild-type control treated with buffer.

Next, we evaluated the significance of flagellin in *A. thaliana* pathogenicity. For these experiments we used a *fliC* mutant that is nonmotile and unable to make flagellin (35), and the hypermotile, flagellated Δ*atsR* mutant (Table 5.1). The *atsR* gene encodes a global regulatory repressor protein that controls motility and the expression of various virulence factors (36-38). There were no statistically significant differences in disease symptoms between plants infected with Δ*atsR* and K56-2 at day 7 postinfection (Fig. 5.2). However, symptoms appeared two days earlier in Δ*atsR*-infected plants, and these results were reproducible in each experimental replica. These observations agree with the current model for the regulatory role of AtsR (37, 38) indicating that removal of inhibitory regulation on motility as well as other AtsR-controlled factors increases bacterial pathogenicity. Therefore, it is not surprising that Δ*atsR* provoked similar lesions to those found with the parental strain K56-2 (Fig. 5.2C). In contrast, loss of flagellin in the *fliC* mutant resulted in a strain that induced only mild HR symptoms in less than 30% of plants infected by wound inoculation (Fig. 5.2A), revealing the critical role of flagella in virulence and also suggesting that flagellar driven motility is important for infection. The wound inoculation method mimics entry of microbes into host plants through natural wounds, which results in bacterial spreading to intercellular spaces. In contrast, infiltration of bacteria with a syringe into the apoplast could bypass the first steps of the natural infection process and therefore, the requirement for bacterial motility. To probe this notion, we infected *Arabidopsis* seedlings by syringe infiltration into leaf abaxial surfaces. Under these conditions, the *fliC* mutant elicited similar HR symptoms as those induced by parental K56-2 (Fig. 5.2C). HR also correlated with higher numbers of bacteria recovered from leaves, a difference that was particularly pronounced when compared with wound inoculation (Fig. 5.3A&B).

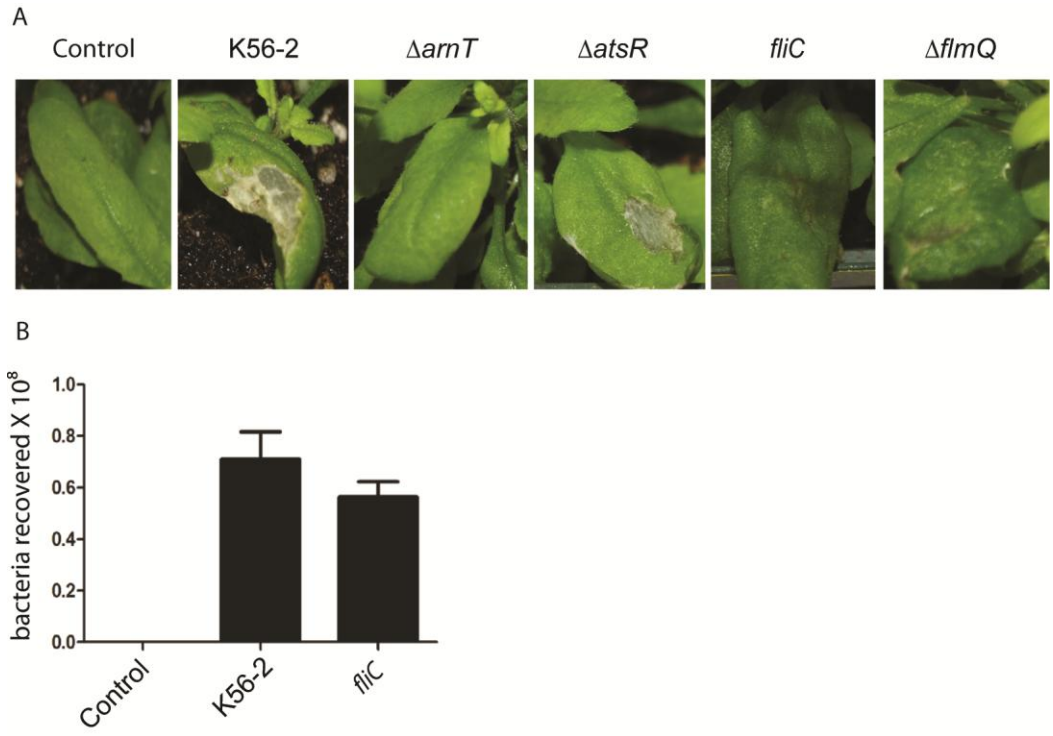


Figure 5.3: Infection of *Arabidopsis* seedlings by the wound inoculation method. A) Representative images of leaves inoculated with a dose of 10^8 CFU of the various strains are shown. B) Bacterial counts on day 7 postinfection recovered from *Arabidopsis* leaves inoculated with *B. cenocepacia* by infiltration.

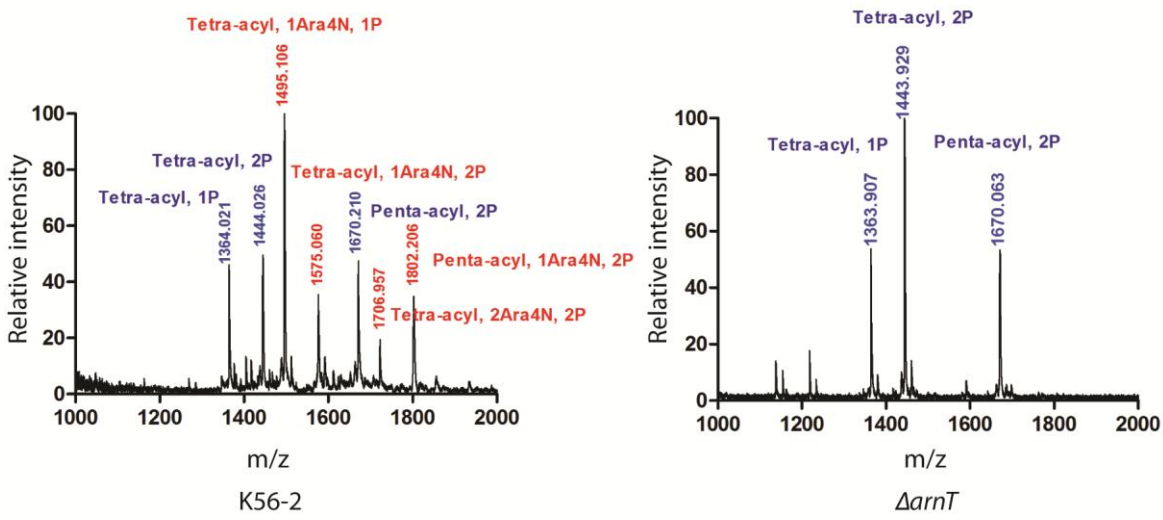


Figure 5.4: MALDI-TOF spectra of purified lipid A produced by the parental *B. cenocepacia* strain K56-2 and the Δ*arnT lptG*^{s} mutant strain MH55 (Table 1). Lipid A samples were prepared and processed as indicated in materials and methods. The profiles represented were obtained using the negative-ion mode. Ion peaks are color coded to indicate ions corresponding to species without L-Ara4N are found in the MH55 lipid A spectrum.

Since post-translational modification of flagellin by glycosylation affects the virulence of *Pseudomonas syringae* in tobacco plants (39, 40), we compared the ability of parental and Δ*flmQ* (Δ*bcal0111*) *B. cenocepacia* strains (Fig. 5.2A) to cause disease in *A. thaliana*. The *flmQ* gene encodes a flagellin glycosyltransferase in *B. cenocepacia* (33). In comparison to the parental strain, Δ*flmQ* was less pathogenic for *A. thaliana* (Fig. 5.2A & 5.3A) and exhibited reduced bacterial survival (Fig. 5.2B). Further, Δ*flmQ* caused significantly less pathology on the leaves with less than 30% of plants having disease symptoms, although slight chlorosis was observable (Fig. 5.2C), suggesting that non-glycosylated flagellin retains its elicitor activity. The reduced virulence of Δ*flmQ* on *A. thaliana* leaves also correlated with bacterial defects in swarming and swimming (Fig. 5.5B&C), consistent with our previous report indicating that glycosylation affects bacterial motility (33). Together, these experiments demonstrate that L-Ara4N modification of LPS and flagellin glycosylation are important for infectivity in *A. thaliana.*

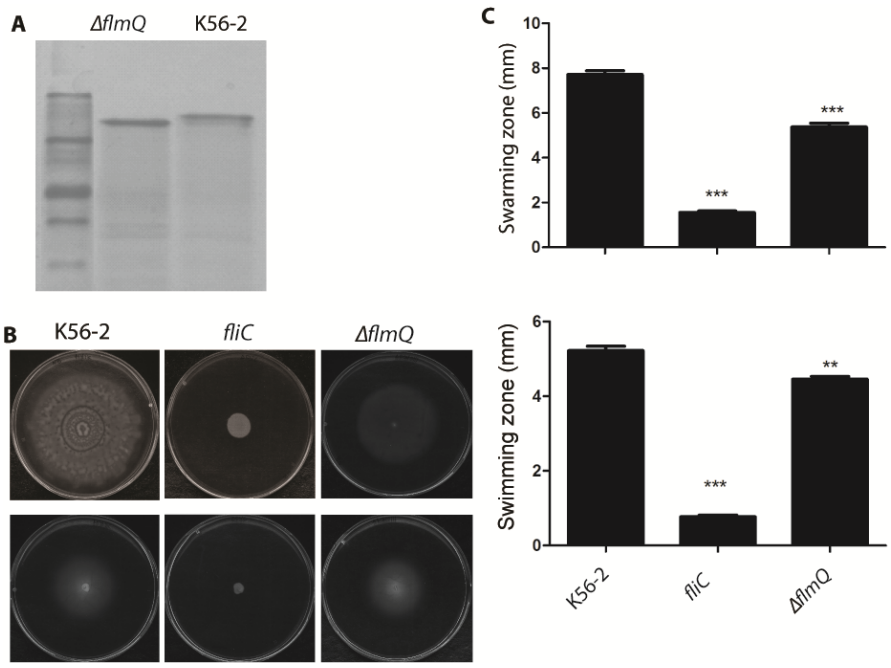


Figure 5.5: Characterization of flagellin mutants. A) SDS-PAGE analysis of purified flagellin from WT and K56-2 Δ*flmQ* mutant strain of *B. cenocepacia.* B) Swarming (Upper row) and swimming (Lower row) motility on soft agar plates of *B. cenocepacia* strains. Data represent three independent experiments. C) Quantification of motility zone. Statistical analysis was performed by paired *t*-test using two-tailed P-values. Significant differences in comparison with *B. cenocepacia* parental strain (WT) as control are indicated by *** ($P < 0.005$) and ** ($P < 0.027$).

5.3.4 LPS induces robust and rapid nitric oxide burst in *A. thaliana*

The reduced pathogenicity of Δ*arnT lptG*$_{D31H}$ could be due to reduced bacterial survival during the initial stages of infection or alternatively, poor recognition of LPS devoid of L-Ara4N by the plant innate immune system. We therefore investigated the levels of NO production in *A. thaliana* guard cells upon exposure to *B. cenocepacia* LPS from mutant and parental sources. Guard cells are specialized epidermal cells in the leaf epidermis surrounding stomatal pores that come into close contact with bacterial pathogens. LPS typically induces NO generation in *A. thaliana* (20, 34). NO burst was detected *in vivo* using the NO-specific fluorescent dye DAF-2DA. Treatment of guard cells with ultrapure *E. coli* LPS control resulted in a rapid burst of green fluorescence, indicative of NO production (Fig. 5.6A). Robust LPS induction of NO in intact guard cells could be reproduced using phenol-extracted LPS from both WT K56-2 and mutant strain lacking L-Ara4N (Fig. 5.6B). Quantification of NO burst using a spectrofluorometric assay that detects NO accumulation indicated that LPS devoid of L-Ara4N has similar levels of elicitor activity suggesting that LPS modification does not affect the innate immune response in *A. thaliana in vivo*.

Further, we compared LPS-induced NO production in guard cells isolated from parental *A. thaliana* and the *Atnoa1* and *dnd1* mutants. *Atnoa1* is a loss-of-function mutant in Nitric Oxide Associated protein 1 (NOA1), which is required for arginine-dependent NO generation and involved in signal cascades responding to PAMPs (20, 41). The defense no death1 (*dnd1-1*) mutant has a null mutation in the *CNGC2/DND1* gene and displays reduced NO generation (34, 42). Minimal responses were detected in *Atnoa1* and *dnd1* guard cells compared to wild type (Fig. 5.7A, left panel). LPS-induced NO production

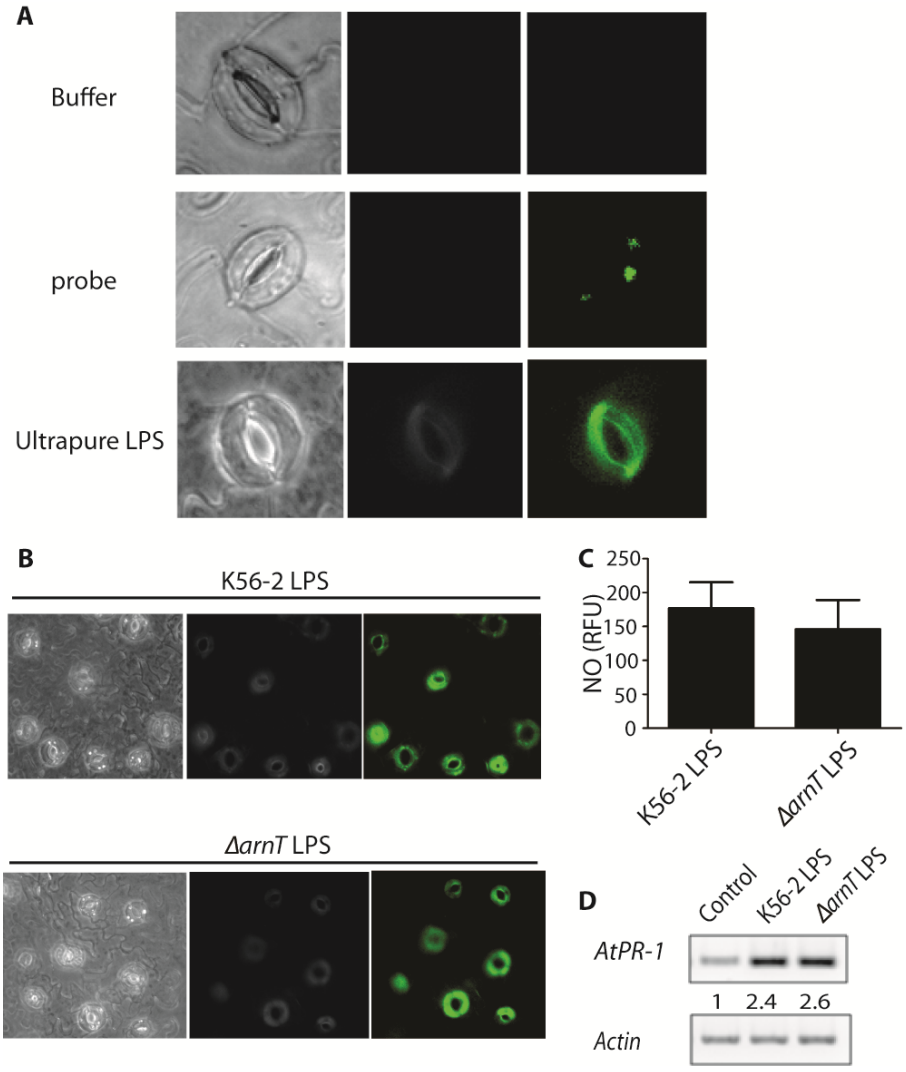


Figure 5.6: Role of L-Ara4N in NO generation in *Arabidopsis* guard cells. A) Leaf epidermal peels prepared from parental (WT) leaves were loaded with buffer control (top panel) or the NO-sensitive dye DAF-2DA (middle panel) prior to incubation with 100 µg/mL ultrapure LPS from *E. coli* (bottom panel). B) Leaf epidermal peels prepared from WT leaves were treated with LPS isolated from K56-2 or K56-2 Δ*arnT* strains. This experiment was repeated twice and a minimum of three epidermal peels was used as treatment replicates. C) Fluorometric quantification of NO generated in WT *Arabidopsis* leaves. D) Induction of *PR-1* gene expression in *Arabidopsis* leaves by LPS. Leaves were treated with buffer (lane 1), K56-2 LPS (lane 2) or LPS isolated from K56-2 Δ*arnT* strain

(lane 3) for 24 h followed by RNA preparation. *ACTIN* was used as an internal control for RT-PCR. Numbers represent relative levels of *PR-1* being normalized relative to control by densitometry.

was reduced dramatically in cells treated with the NOS inhibitor N^G-nitro-L-Arg-methyl ester (L-NAME) (Fig. 5.7A, right panel). Similarly, LPS-dependent NO generation was not observed in *dnd1-1* or *Atnoa1* mutants after treatment with L-NAME, while incubation with the NO donor sodium nitroprusside (SNP) resulted in green fluorescence in parental *A. thaliana* and both mutants. Therefore, reduced NO burst in *Atnoa1* and *dnd1-1* cells was due to a defect in NO production (Fig. 5.7A, middle panel). NO was quantified in LPS-treated leaves of parental and mutant plants using a fluorometric assay. LPS treatment yielded 180 relative fluorescence units (RFUs) in parental leaves, compared with 78% and 66% reduction in RFUs observed for *Atnoa1* and *dnd1-1* mutants, respectively (Fig. 5.7B). LPS-induced NO production in parental leaves was repressed by 6-fold upon treatment with the NO synthase inhibitor L-NAME (Fig. 5.7B). In contrast, parental and mutant leaves showed similar RFUs in the presence of SNP. Together, these results demonstrate that *B. cenocepacia* LPS triggers a robust NO burst in *Arabidopsis* leaves and both intact LPS and LPS lacking L-Ara4N can be equally recognized by the plant innate immune system.

5.3.5 Flagellin perception in *A. thaliana* occurs irrespective of the flagellin glycosylation status

We also investigated if the glycosylation status of the *B. cenocepacia* flagellin alters the immunity response of *A. thaliana*, which involves a rapid and transient burst of ROS that is subsequently associated with programmed cell death (43). Treatment of plants with *B. cenocepacia* flagellin triggered ROS generation in parental *A. thaliana* leaves, but not in leaves from the *fls2* mutant, which lacks the flagellin receptor (Fig. 5.8A). This result agrees with previous reports in *Arabidopsis* and tomato plants demonstrating the

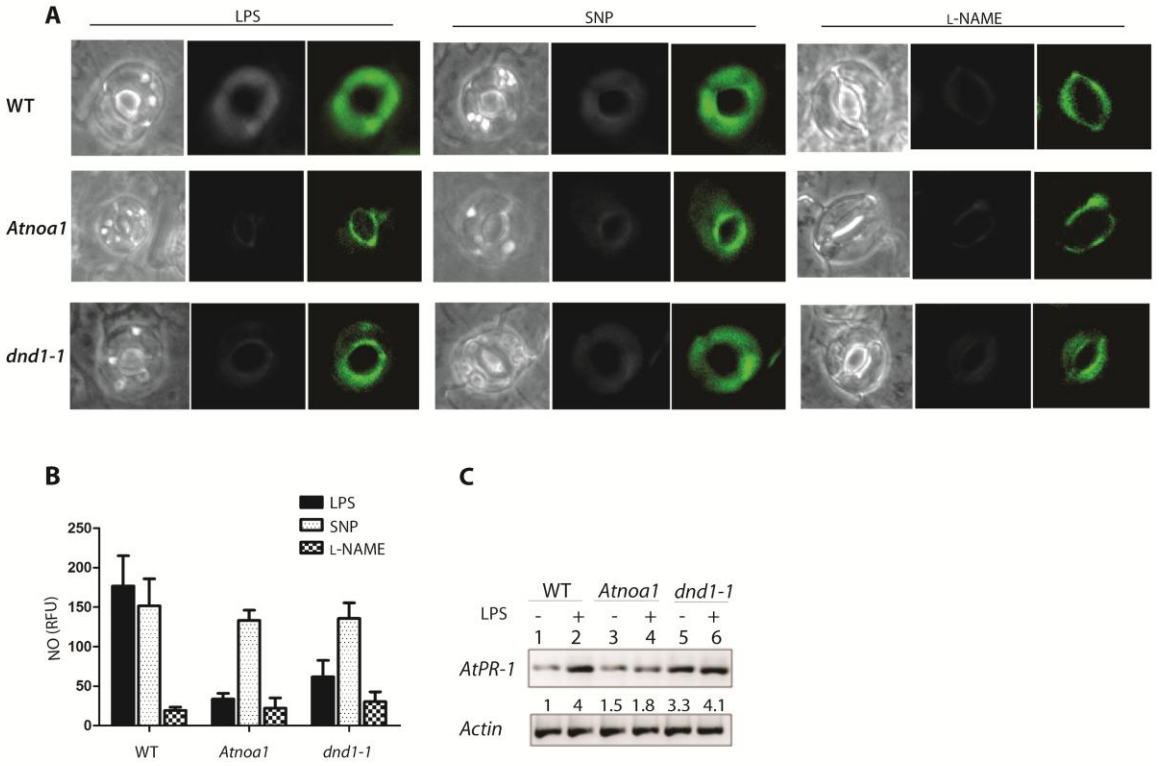


Figure 5.7: LPS activates NO production in WT, *Atnoa1* and *dnd1-1* guard cells. A) Leaf epidermal peels prepared from WT (top panels), *Atnoa1* (middle panels) or *dnd1-1* plants (bottom panels) were loaded with 5 µM DAF-2DA prior to incubation in reaction buffer alone (buffer control), 100 µg/mL LPS (left panel), 50 µM SNP (middle panel) or 200 µM L-NAME (right panel). In each case, corresponding fluorescence and bright-field images are shown; the area of the peel subjected to analysis was greater than that shown in each case. This experiment was repeated three times using at least three epidermal peels per experiment. B) Fluorometric quantification of NO generated in leaves of WT, *Atnoa1* and *dnd1-1 Arabidopsis* seedlings. C) Induction of *PR-1* gene expression in *Arabidopsis* leaves by LPS. *Arabidopsis* leaves were treated with buffer (lanes 1, 3 & 5) or LPS (lanes 2, 4 & 6) for 24 h followed by RNA preparation. RT-PCR was performed with cDNAs for *PR-1* and *ACTIN*. Numbers represent relative levels of *PR-1* being normalized relative to control by densitometry.

importance of flagellin perception for ROS production and disease resistance (13, 15). Fluorometric quantification of ROS in parental and *fls2* mutant leaves showed a 3-fold reduction indicating that flagellin recognition was necessary for eliciting ROS production (Fig. 5.8B). However, no significant differences were detected in ROS production by leaves treated with non-glycosylated flagellin, despite that infection with bacteria lacking glycosylated flagellin produce less symptoms of infection (Fig. 5.8C). This result suggests that reduced pathogenicity of *A. thaliana* exposed with the *B. cenocepacia* Δ*flmQ* is due to reduced bacterial motility rather than differential perception of glycosylated vs. non-glycosylated flagellin by the FLS2 receptor.

5.3.6. Induction of Pathogenesis Related gene-1 (*PR-1*) by LPS and flagellin

Host defense responses in *A. thaliana* depend on ROS and NO production and induction of host defense or stress-associated genes, such as genes encoding Pathogenesis Related proteins (20). We investigated *PR-1* gene expression by RT-PCR, since the expression of this gene is considered a hallmark of defense response in plants (20, 21). *PR-1* expression was significantly increased in parental *A. thaliana* leaves 24 h after treatment with either intact or L-Ara4N devoid LPS (Fig. 5.6D). In contrast, *PR-1* mRNA accumulation was reduced in *Atnoa1* mutant plants (Fig. 5.7C, lanes 3&4) suggesting a functional link between LPS-induced NO production and *PR-1* expression. Similarly, *PR-1* expression is activated by flagellin irrespective of its glycosylation status in parental *A. thaliana* but not in the flagellin-insensitive *fls2* mutant (Fig. 5.8C). As expected, *dnd1-1* mutant plants expressed high levels of *PR-1* mRNA, which is due to the constitutively active salicylic acid signalling pathway in this mutant (44, 45). Together, our findings indicate that modifiedLPS and flagellin induce *PR-1* transcription at levels similar to that of intact LPS

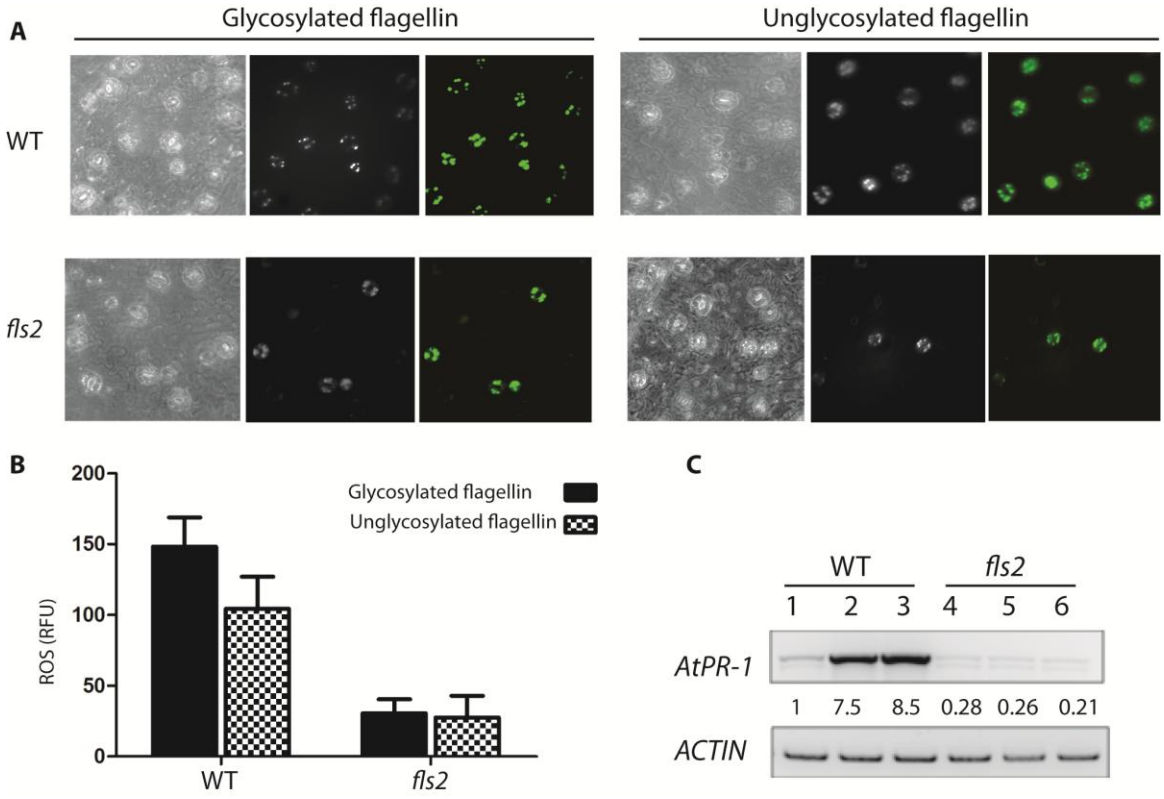


Figure 5.8: Flagellin activates ROS generation in guard cells of WT and *fls2* leaves. A) Leaf epidermal peels prepared from WT (top panels) or *fls2* plants (bottom panels) were loaded with the ROS-sensitive dye DCF prior to incubation with glycosylated (left panel) or unglycosylated flagellin (right panel) from K56-2. In each case, corresponding fluorescence and bright-field images are shown; Images are the representative of three individual experiments. In each experiment, at least three epidermal peels were used as treatment replicates. B) Fluorometric quantification of ROS generated in leaves of WT and *fls2 Arabidopsis* seedlings. C) Induction of *PR-1* gene expression in *Arabidopsis* leaves by flagellin. *Arabidopsis* leaves were treated with buffer (lanes 1 & 4), WT (lanes 2 & 5) or unglycosylated flagellin (lanes 3 & 6) for 24 h followed by RNA preparation. RT-PCR was performed with cDNAs for *PR-1* and *ACTIN-2*. Numbers represent relative levels of *PR-1* being normalized relative to control by densitometry.

and flagellin. This agrees with our results regarding induction of oxidative burst indicating that L-Ara4N modification of LPS and flagellin glycosylation are critical for infectivity in *A. thaliana* but do not alter the perception of these molecules by the plant innate immune receptors.

5.3.7 Comparison of *Arabidopsis* and *Galleria* infection models

The previous results indicate that flagellin and LPS modifications alter the infectivity of *B. cenocepacia* in *A. thaliana*, but the plant can perceive these molecules irrespectively. Therefore, we tested the ability of *B. cenocepacia* strains utilized in the plant infection model for survival and pathogenicity in G. *mellonella*. *G. mellonella* moth larvae have been increasingly used as an infection model for many bacterial pathogens including several *Burkholderia* species (8, 46, 47). *B. cenocepacia* parental and mutant strains were injected into *G. mellonella* and survival was monitored for over 72 h. As with *A. thaliana*, K56-*2* Δ*atsR* exhibited high virulence denoted by no larvae surviving at 48 h post injection with 10^4 CFU (Fig. 5.9A), compared to 5 ± 3% survival of larvae infected with parental K56-2. In contrast, Δ*arnT lptG*$_{D31H}$ was greatly attenuated, as shown by larvae survival rates of more than 80% at 48 h post infection (Fig. 5.9A). At infection doses of 10^4 CFU, similar attenuation levels were observed in the *fliC* mutant, while Δ*flmQ* showed an intermediate phenotype (50 ± 13.5% survival). At 72 h post infection the larvae survival rates with Δ*arnT lptG*$_{D31H}$, *fliC*, and Δ*flmQ* were 77 ± 14, 43 ± 12, and 28 ± 7%, respectively. Control larvae injected with buffer remained viable until the end of the experiment. These results are comparable to the data obtained from the *A. thaliana* infections, confirming the significance of flagellin glycosylation and L-Ara4N modification of LPS in the infectivity of *B. cenocepacia*.

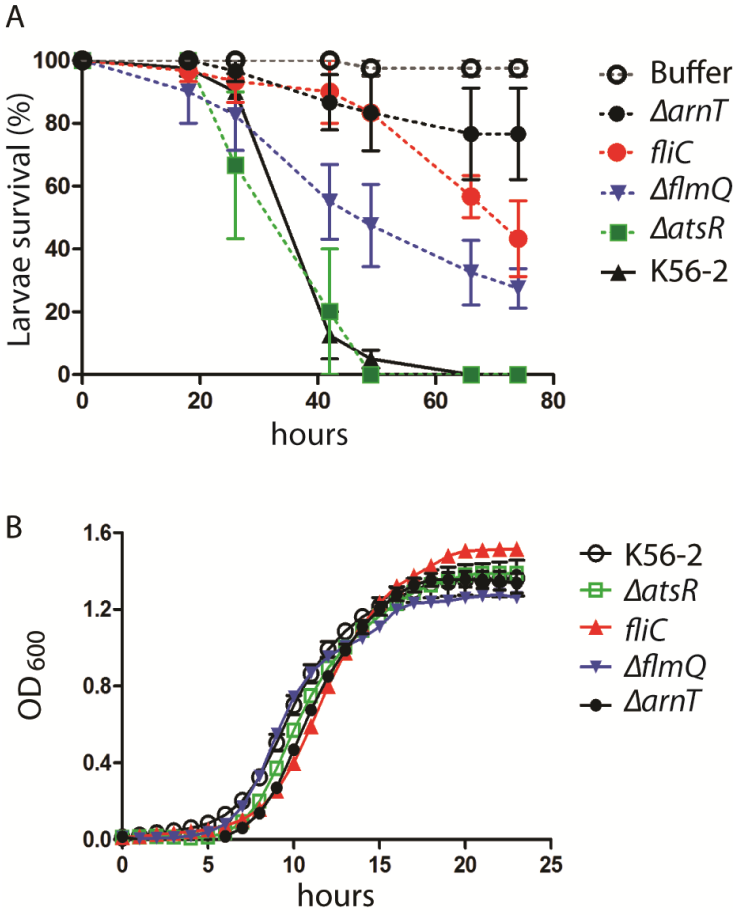


Figure 5.9: Virulence of *B. cenocepacia* strains in *G. mellonella*. A) Larvae were infected with 1 × 10^4 CFU of the indicated *Burkholderia* strains. Larval survival was monitored during 75 h post infection. Data are means ± SEM. B) Growth rate of parental *B. cenocepaica* strain K56-2 and its isogenic mutants in LB medium. Data points represent the mean of 6 replicas.

5.4. Conclusion

Virulence and pathogenicity requires considering both the host and the microbe; the use of these terms to define a microbe in isolation of its host has recently been debated (48, 49). We show here that modifications in the LPS molecule and flagellin glycosylation do not affect recognition by plant innate immune receptors, but are still required for bacteria to establish an infection. Our experiments using the *B. cenocepacia* K56-2-*A. thaliana* infection model reveal that plant seedlings sense LPS and flagellin as PAMPs and induce a burst of NO and ROS that contributes to basal resistance and eventually leads to a hypersensitive response against *B. cenocepacia*. Redox signalling has emerged as a main regulator of cellular function in plant pathophysiology (50). Pathogen recognition turns on a signalling cascade where free radicals, such as NO and ROS, act as infochemicals for activation of various defense genes that mediate disease resistance in plants (14). NO and ROS production upon K56-2 LPS or flagellin treatment potentiated the expression of *PR-1*, which is a disease resistance marker. In contrast, *PR-1* gene expression was reduced when *Atnoa1* and *fls2* mutant plants were treated with LPS or flagellin, indicating that LPS-responsive genes are NO inducible (20). In our work, we did not observe a distinguishable difference in the expression of the *PR-1* gene after treatment of leaves with intact or modified versions of LPS and flagellin.

Although we did not use purified lipid A in our study, these observations contrast with a previous report indicating that isolated lipid A of a non pathogenic mutant of *Xanthomonas campestris*, consisting mainly of penta-acylated lipid A with phosphoethanolamine substitutions of its phosphate groups, was unable to elicit *PR-1* expression (23). The authors concluded that the substitution of both lipid A phosphate

groups with phosphoethanolamine would neutralize the net negative charge of the lipid A, which could affect binding to putative plant receptors (23). In contrast, the *B. cenocepacia* lipid A phosphate groups are substituted by L-Ara4N molecules and disappearance of these residues in the Δ*arnT lptG*$_{D31H}$ would expose negative charges, thus maintaining the ability of lipid A to activate *PR-1* expression. Importantly, the lipid A of *B. cenocepacia* is tetra- and penta-acylated, but the acyl chains are longer than those of the *X. campestris* lipid A (51). Therefore, the fact that parental and L-Ara4N devoid LPS from *B. cenocepacia* induce similar levels of *PR-1* gene expression suggests that the acyl chain composition may also play a role in LPS detection by plant innate immune receptors.

Irrespective of K56-2 LPS elicitor activity, *B. cenocepacia* mutants unable to produce L-Ara4N-modified LPS have extremely high sensitivity to Polymyxin B (27), suggesting that they would also be highly sensitive to plant antimicrobial peptides. Further, we show that post-translational modification of flagellin by glycosylation affects bacterial motility and virulence in *Arabidopsis*, but does not significantly alter the elicitor activity of the protein. A correlation between defects in flagellin glycosylation in *Pseudomonas syringae* and reduced ability to induce HR in plants has been reported (39, 40), but the molecular mechanism is unknown. Our experiments indicate that reduced virulence but increased activation of plant defenses by the Δ*flmQ* mutant producing non glycosylated flagellin is due to bacterial motility defects that prevent bacteria to establish infection. This could be clearly demonstrated by showing that the attenuation of the K56-2 *fliC* mutant strain is bypassed if bacteria are directly infiltrated into the plant deeper tissues. A role for flagellar motility was also reported for several plant pathogens such as *Erwinia*

carotovora (52), *Ralstonia solanacearum* (53), and several pathovars of *P. syringae* and *Xanthomonas campestris* (54-56).

In conclusion, we demonstrate that flagellin glycosylation and LPS modification with L-Ara4N play a significant role on bacterial survival during the early stages upon infection, but do not alter the perception of these molecules by the plant innate immune receptors, indicating these modifications are only critical to establish infection. Our experiments therefore illustrate the notion that microbe's perception by the host and establishment of infection are interrelated but independent events.

5.5 Chapter 5 references

1. **Mahenthiralingam E, Baldwin A, Dowson CG.** 2008. *Burkholderia cepacia* complex bacteria: opportunistic pathogens with important natural biology. J Appl Microbiol **104:**1539-1551.

2. **Parke JL, Gurian-Sherman D.** 2001. Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. Annu Rev Phytopathol **39:**225-258.

3. **Drevinek P, Mahenthiralingam E.** 2010. *Burkholderia cenocepacia* in cystic fibrosis: epidemiology and molecular mechanisms of virulence. Clin Microbiol Infect **16:**821-830.

4. **Bernier SP, Silo-Suh L, Woods DE, Ohman DE, Sokol PA.** 2003. Comparative analysis of plant and animal models for characterization of *Burkholderia cepacia* virulence. Infect Immun **71:**5306-5313.

5. **Vergunst AC, Meijer AH, Renshaw SA, O'Callaghan D.** 2010. *Burkholderia cenocepacia* creates an intramacrophage replication niche in zebrafish embryos, followed by bacterial dissemination and establishment of systemic infection. Infect Immun **78:**1495-1508.

6. **Huber B, Feldmann F, Kothe M, Vandamme P, Wopperer J, Riedel K, Eberl L.** 2004. Identification of a novel virulence factor in *Burkholderia cenocepacia* H111 required for efficient slow killing of *Caenorhabditis elegans*. Infect Immun **72:**7220-7230.

7. **Uehlinger S, Schwager S, Bernier SP, Riedel K, Nguyen DT, Sokol PA, Eberl L.** 2009. Identification of specific and universal virulence factors in *Burkholderia cenocepacia* strains by using multiple infection hosts. Infect Immun **77:**4102-4110.

8. **Seed KD, Dennis JJ.** 2008. Development of *Galleria mellonella* as an alternative infection model for the *Burkholderia cepacia* complex. Infect Immun **76:**1267-1275.

9. **Thomson EL, Dennis JJ.** 2013. Common duckweed (*Lemna minor*) is a versatile high-throughput infection model for the *Burkholderia cepacia* complex and other pathogenic bacteria. PLoS One **8:**e80102.

10. **Nishimura MT, Dangl JL.** 2010. *Arabidopsis* and the plant immune system. Plant J **61:**1053-1066.

11. **Newman MA, Sundelin T, Nielsen JT, Erbs G.** 2013. MAMP (microbe-associated molecular pattern) triggered immunity in plants. Front Plant Sci **4:**139.

12. **Whitfield C, Trent MS.** 2014. Biosynthesis and export of bacterial lipopolysaccharides. Annu Rev Biochem **83:**99-128.

13. **Felix G, Duran JD, Volko S, Boller T.** 1999. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J **18:**265-276.

14. **Torres MA, Jones JD, Dangl JL.** 2006. Reactive oxygen species signaling in response to pathogens. Plant Physiol **141:**373-378.

15. **Gomez-Gomez L, Felix G, Boller T.** 1999. A single locus determines sensitivity to bacterial flagellin in *Arabidopsis thaliana*. Plant J **18:**277-284.

16. **Jones JD, Dangl JL.** 2006. The plant immune system. Nature **444:**323-329.

17. **Xu XM, Moller SG.** 2011. The value of *Arabidopsis* research in understanding human disease states. Curr Opin Biotechnol **22:**300-307.

18. **Koornneef M, Meinke D.** 2010. The development of *Arabidopsis* as a model plant. Plant J **61:**909-921.

19. **Ausubel FM, Katagiri F, Mindrinos M, Glazebrook J.** 1995. Use of *Arabidopsis thaliana* defense-related mutants to dissect the plant response to pathogens. Proc Natl Acad Sci U S A **92:**4189-4196.

20. **Zeidler D, Zahringer U, Gerber I, Dubery I, Hartung T, Bors W, Hutzler P, Durner J.** 2004. Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. Proc Natl Acad Sci U S A **101:**15811-15816.

21. **Madala NE, Molinaro A, Dubery IA.** 2012. Distinct carbohydrate and lipid-based molecular patterns within lipopolysaccharides from *Burkholderia cepacia* contribute to defense-associated differential gene expression in *Arabidopsis thaliana*. Innate Immun **18:**140-154.

22. **Needham BD, Carroll SM, Giles DK, Georgiou G, Whiteley M, Trent MS.** 2013. Modulating the innate immune response by combinatorial engineering of endotoxin. Proc Natl Acad Sci U S A **110:**1464-1469.

23. **Silipo A, Sturiale L, Garozzo D, Erbs G, Jensen TT, Lanzetta R, Dow JM, Parrilli M, Newman MA, Molinaro A.** 2008. The acylation and phosphorylation pattern of lipid A from *Xanthomonas campestris* strongly influence its ability to trigger the innate immune response in *Arabidopsis*. Chembiochem **9:**896-904.

24. **Newman MA, Dow JM, Molinaro A, Parrilli M.** 2007. Priming, induction and modulation of plant defence responses by bacterial lipopolysaccharides. J Endotoxin Res **13:**69-84.

25. **Loutet SA, Valvano MA.** 2011. Extreme antimicrobial Peptide and polymyxin B resistance in the genus *Burkholderia*. Front Microbiol **2:**159.

26. **Ortega XP, Cardona ST, Brown AR, Loutet SA, Flannagan RS, Campopiano DJ, Govan JR, Valvano MA.** 2007. A putative gene cluster for aminoarabinose biosynthesis is essential for *Burkholderia cenocepacia* viability. J Bacteriol **189:**3639-3644.

27. **Hamad MA, Di Lorenzo F, Molinaro A, Valvano MA.** 2012. Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. Mol Microbiol **85:**962-974.

28. **Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR, Eng JK, Akira S, Underhill DM, Aderem A.** 2001. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. Nature **410:**1099-1103.

29. **Taguchi F, Shimizu R, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y.** 2003. Post-translational modification of flagellin determines the specificity of HR induction. Plant Cell Physiol **44:**342-349.

30. **Sun W, Dunning FM, Pfund C, Weingarten R, Bent AF.** 2006. Within-species flagellin polymorphism in *Xanthomonas campestris pv campestris* and its impact on elicitation of *Arabidopsis* FLAGELLIN SENSING2-dependent defenses. Plant Cell **18:**764-779.

31. **Chinchilla D, Bauer Z, Regenass M, Boller T, Felix G.** 2006. The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. Plant Cell **18:**465-476.

32. **Gomez-Gomez L, Boller T.** 2000. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Mol Cell **5:**1003-1011.

33. **Hanuszkiewicz A, Pittock P, Humphries F, Moll H, Rosales AR, Molinaro A, Moynagh PN, Lajoie GA, Valvano MA.** 2014. Identification of the flagellin glycosylation system in *Burkholderia cenocepacia* and the contribution of glycosylated flagellin to evasion of human innate immune responses. J Biol Chem **289:**19231-19244.

34. **Ali R, Ma W, Lemtiri-Chlieh F, Tsaltas D, Leng Q, von Bodman S, Berkowitz GA.** 2007. Death don't have no mercy and neither does calcium: *Arabidopsis* CYCLIC NUCLEOTIDE GATED CHANNEL2 and innate immunity. Plant Cell **19:**1081-1095.

35. **Saldias MS, Lamothe J, Wu R, Valvano MA.** 2008. *Burkholderia cenocepacia* requires the RpoN sigma factor for biofilm formation and intracellular trafficking within macrophages. Infect Immun **76:**1059-1067.

36. **Aubert DF, Flannagan RS, Valvano MA.** 2008. A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in *Burkholderia cenocepacia*. Infect Immun **76:**1979-1991.

37. **Aubert DF, O'Grady EP, Hamad MA, Sokol PA, Valvano MA.** 2013. The *Burkholderia cenocepacia* sensor kinase hybrid AtsR is a global regulator modulating quorum-sensing signalling. Environ Microbiol **15:**372-385.

38. **Khodai-Kalaki M, Aubert DF, Valvano MA.** 2013. Characterization of the AtsR hybrid sensor kinase phosphorelay pathway and identification of its response regulator in *Burkholderia cenocepacia*. J Biol Chem **288:**30473-30484.

39. **Taguchi F, Takeuchi K, Katoh E, Murata K, Suzuki T, Marutani M, Kawasaki T, Eguchi M, Katoh S, Kaku H, Yasuda C, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y.** 2006. Identification of glycosylation genes and glycosylated amino acids of flagellin in *Pseudomonas syringae pv. tabaci*. Cell Microbiol **8:**923-938.

40. **Taguchi F, Yamamoto M, Ohnishi-Kameyama M, Iwaki M, Yoshida M, Ishii T, Konishi T, Ichinose Y.** 2010. Defects in flagellin glycosylation affect the virulence of *Pseudomonas syringae* pv. *tabaci* 6605. Microbiology **156:**72-80.

41. **Crawford NM, Guo FQ.** 2005. New insights into nitric oxide metabolism and regulatory functions. Trends Plant Sci **10:**195-200.

42. **Clough SJ, Fengler KA, Yu IC, Lippok B, Smith RK, Jr., Bent AF.** 2000. The *Arabidopsis dnd1* "defense, no death" gene encodes a mutated cyclic nucleotide-gated ion channel. Proc Natl Acad Sci U S A **97:**9323-9328.

43. **Torres MA, Jones JD, Dangl JL.** 2005. Pathogen-induced, NADPH oxidase-derived reactive oxygen intermediates suppress spread of cell death in *Arabidopsis thaliana*. Nat Genet **37:**1130-1134.

44. **Yu IC, Parker J, Bent AF.** 1998. Gene-for-gene disease resistance without the hypersensitive response in *Arabidopsis dnd1* mutant. Proc Natl Acad Sci U S A **95:**7819-7824.

45. **Ahn IP.** 2007. Disturbance of the Ca(2+)/calmodulin-dependent signalling pathway is responsible for the resistance of *Arabidopsis dnd1* against *Pectobacterium carotovorum* infection. Mol Plant Pathol **8:**747-759.

46. **Wand ME, Muller CM, Titball RW, Michell SL.** 2011. Macrophage and *Galleria mellonella* infection models reflect the virulence of naturally occurring isolates of *B. pseudomallei*, *B. thailandensis* and *B. oklahomensis*. BMC Microbiol **11:**11.

47. **Schmerk CL, Valvano MA.** 2013. *Burkholderia multivorans* survival and trafficking within macrophages. J Med Microbiol **62:**173-184.

48. **Casadevall A, Pirofski LA.** 2014. Microbiology: Ditch the term pathogen. Nature **516:**165-166.

49. **Casadevall A, Pirofski LA.** 2015. What Is a Host? Incorporating the Microbiota into the Damage-Response Framework. Infect Immun **83:**2-7.

50. **Lindermayr C, Sell S, Muller B, Leister D, Durner J.** 2010. Redox regulation of the NPR1-TGA1 system of *Arabidopsis thaliana* by nitric oxide. Plant Cell **22:**2894-2907.

51. **Silipo A, Molinaro A, Ierano T, De Soyza A, Sturiale L, Garozzo D, Aldridge C, Corris PA, Khan CM, Lanzetta R, Parrilli M.** 2007. The complete structure and pro-inflammatory activity of the lipooligosaccharide of the highly epidemic and virulent Gram-negative bacterium *Burkholderia cenocepacia* ET-12 (strain J2315). Chemistry **13:**3501-3511.

52. **Hossain MM, Shibata, S., Aizawa, S., and Tsuyumu, S.** 2005. Motility is an important determinant for pathogenesis of *Erwinia carotovora* subsp. carotovora Physiol Mol Plant Pathol **66:**134-143.

53. **Tans-Kersten J, Brown D, Allen C.** 2004. Swimming motility, a virulence trait of *Ralstonia solanacearum*, is regulated by FlhDC and the plant host environment. Mol Plant Microbe Interact **17:**686-695.

54. **Haefele DM, Lindow SE.** 1987. Flagellar motility confers epiphytic fitness advantages upon *Pseudomonas syringae*. Appl Environ Microbiol **53:**2528-2533.

55. **Hattermann DR.** 1989. Motility of *Pseudomonas syringae* pv. glycinea and its role in infection. Phytopathology **79:**284-289.

56. **Beattie GA, Lindow SE.** 1995. The secret life of foliar bacterial pathogens on leaves. Annu Rev Phytopathol **33:**145-172.

57. **Johnson WM, Tyler SD, Rozee KR.** 1994. Linkage analysis of geographic and clinical clusters in *Pseudomonas cepacia* infections by multilocus enzyme electrophoresis and ribotyping. J Clin Microbiol **32:**924-930.

58. **Loutet SA, Bartholdson SJ, Govan JR, Campopiano DJ, Valvano MA.** 2009. Contributions of two UDP-glucose dehydrogenases to viability and polymyxin B resistance of *Burkholderia cenocepacia*. Microbiology **155:**2029-2039.

59. **El Hamidi A, Tirsoaga A, Novikov A, Hussein A, Caroff M.** 2005. Microextraction of bacterial lipid A: easy and rapid method for mass spectrometric characterization. J Lipid Res **46:**1773-1778.

60. **Guo FQ, Okamoto M, Crawford NM.** 2003. Identification of a plant nitric oxide synthase gene involved in hormonal signaling. Science **302:**100-103.

Chapter 6

General discussion

6.1 Overview

The opportunistic nature of *Burkholderia cenocepacia* is manifested by its amazing ability to adapt to and survive in a wide range of environmental and anatomical niches, which explains its pathogenesis in human (1), plants (2), nematodes (3) and animals (4). This requires an ability to sense and respond to the surrounding environment, which depends on regulatory networks that wire signal perception to coordinated gene expression and adaptive responses. Our data indicate that the AtsR protein has a major role in this process due to its nature as a sensor kinase and its high conservation in *Burkholderia.*

6.2 AtsR mechanism of action

It has been established that quorum sensing and T6SS are required for *in vivo* infection and therefore AtsR is involved in regulating virulence factors. However, the predicted structural features of AtsR suggested this protein could not directly control gene expression because it lacks an effector domain. In chapter two, we identified AtsR's cognate response regulator and demonstrated key properties of the AtsR to AtsT phosphotransfer mechanism. First, upon autophosphorylation AtsR transfers the phosphate directly to the response regulator AtsT without involving other intermediate protein. Second, the His-245 residue of AtsR is absolutely required for function both *in vitro* and *in vivo*. Third, the C-terminal receiver domain of AtsR plays a role in modulating the stability of phosphorylated AtsR. It is very common in hybrid sensor kinases that the aspartate residue on the receiver domain modulates phosphotransfer and thus determines the specificity or regulation of the autokinase activity (5-8). Fourth, the

long half-life of phosphorylated AtsR and AtsT maintains the response memory for longer periods allowing the regulation of gene expression. Finally, the rapid phosphotransfer from AtsR to AtsT within 2 minutes represents the specificity of signal transduction based on molecular recognition indicating that they are true cognate pairs. Non-specific phosphotransfer between non-cognate pairs occur with prolonged incubation times at a much slower rate *in vitro* (9). For more in depth analysis to determine the critical residues for molecular recognition a combination of traditional single amino acid mutagenesis and more computational and systematic approaches as well as structural studies is required.

6.3 AtsR topology

To begin understanding the signal(s) or conditions that lead to activation or deactivation of the AtsR/AtsT phosphorelay we wanted to determine the critical amino acids residing in the sensory domain of AtsR. We first experimentally validated the predicted topology of the protein and then studied its sensory domain. We further identified 4 residues in the periplasmic region that are absolutely required for the function of the protein. Based on PSIPRED modeling system the periplasmic region of AtsR has four major α-helices. A well known example of periplasmic domain with antiparallel four-helix bundle is NarX. NarX responds to nitrate and nitrite and regulates anaerobic respiration. A high-resolution X-ray crystal structure of the NarX periplasmic sensor domain from *E. coli* revealed that nitrate-binding induces conformation changes that result in a piston-type displacement between the N- and C-terminal helices of the periplasmic domain (10). This might represent a conserved mechanism of ligand binding in histidine kinases with antiparallel helix bundles. This might also be the case for the AtsR sensor domain with those

identified 4 residues within the periplasmic loop facilitating an interaction with a signalling molecule or maintaining the binding pocket.

The middle gene of the *atsR* operon is *Bcam0380*, which encodes a protein with a molybdopterin domain that is found in a variety of oxidoreductases. We tested if any of the cysteine residues in the AtsR protein is required for its function to link it functionally to the neighboring gene *Bcam0380*. A cys-less version of AtsR localized to the membrane and remained functional indicated that the native cysteines are dispensable for AtsR function. Currently, experiments are undergoing to identify potential signalling molecules.

6.4 Cellular functions under the regulation of AtsR/AtsT

Our data strongly support the notion that AtsR phosphorylation has a significant biological relevance as a virulence regulator modulating the expression of genes through AtsT. This notion was further confirmed by our comparative genetic and biochemical analysis in chapter 3.

We compared the expression profiling of *atsR* mutant with a high throughput sequencing procedure. This combination defines the direct and indirect AtsR/AtsT regulon in the *B. cenocepacia* genome. Some of the identified genes encode proteins with unknown function and the rest have known or predicted functions in a wide range of cellular processes including T6SS, metallopeptidases, lipoproteins, quorum sensing related genes and a few genes encoding regulatory proteins that might function as intermediates in regulatory cascades (Table 3.3 and Table 3.4). Our results also have an interesting feature which is the existence of a negative feedback loop for regulating the transcription levels

of *atsR* operon. Such negative autoregulation is in line with the prolonged nature of phosphorylated AtsR observed in our work. Assuming that the optimum goal of a two component circuit is to reach a steady-state protein concentration to detect environmental signals and to avoid overloading of the receptors, a negative regulatory circuit with a relatively stable phosphorylation status such as that of AtsR works far better to maintain steady-state levels by *do-novo* synthesis rather than protein turn-over. This way, unoccupied receptors are always available to perceive new signals and adjust responses accordingly.

Another interesting observation is the variable affinity of AtsT DNA-binding to different promoters regardless of its phosphorylation status *in vitro*. This indicates that different binding affinities to the three promoters we tested here may represent a "fine-tuning" mode of action. In other words, this highly flexible mode of occupancy at promoters would likely render the expression of these genes relative to changes in AtsT-P levels; thus, it does not resemble an on-off switch. However, it should be noted that AtsT behaviour under *in vivo* conditions might be different than *in vitro*. For the majority of response regulators the phosphorylated form typically has a higher affinity for their target promoters (11). Also, certain response regulators bind DNA *in vivo* only when they are phosphorylated and produced at physiological levels (12, 13). This might be due to the accessibility of promoters or the presence of accessory proteins that promote or prevent phosphorylation.

Although we aimed to identify genes directly regulated by AtsT, comparison of gene profiles from *atsR* mutant with ChIP-Seq data showed that there are a few genes present in both profiles. This result reflects the possibility of cross talk between two component

systems and that AtsT might be involved in other signalling pathways and act as a regulatory arm for several receptor proteins besides AtsR. Overall, AtsT directly regulates the expression of nearly 100 genes and is predominantly a repressor of genes encoding proteins associated with T6SS as well as lipoproteins, transposases, transport system binding proteins, glycosyltransferases, to name a few for which differences in expression have been observed in Δ*atsT* mutant and wild-type strains. By *in silico* analysis of ChiP-seq data the AtsT-regulated genes can be grouped into different functional categories such as DNA repair and recombination, glycan biosynthesis and metabolism, mitochondrial biogenesis, environmental information processing and membrane transport. The genes that are required for biofilm maturation and quorum sensing as a hallmark of AtsR/AtsT function are three main classes that are involved in encoding surface proteins, the biogenesis and maintenance of membrane integrity, and those that encode regulatory factors that affect *N*-acyl homoserine lactone production or perception. Examples of these classes have been identified in our combined approaches (Tables 3.3 and 3.4). Furthermore, our bioinformatic and DNase I foot printing analyses reveal the AtsT binding site architecture that likely has important implications for gene regulation in *B. cenocepacia* in many ways whether in the CF lung environment or its interaction with other environments.

6.5 *Burkholderia cenocepacia* as an opportunistic model bacterium

B. cenocepacia cause opportunistic infections in plants, insects, animals and humans, suggesting that "virulence" depends on the host and its innate susceptibility to infection. In chapter 5, I used *B. cenocepacia* as an opportunistic model bacterium and the host plant *Arabidopsis thaliana* to investigate the role of bacterial surface molecules, namely

lipopolysaccharide and flagellin, in contributing to infection and also in eliciting a host response. I hypothesized that modifications in key bacterial molecules recognized by the innate immune system modulate host responses to *B. cenocepacia*. Indeed, modification of lipopolysaccharide (LPS) with 4-amino-4-deoxy-L-arabinose and flagellin glycosylation attenuates *B. cenocepacia* infection in *Arabidopsis thaliana* and *Galleria mellonella* insect larvae. However, *B. cenocepacia* LPS and flagellin triggers rapid bursts of nitric oxide and reactive oxygen species in *A. thaliana* leading to activation of the *PR-1* gene. These responses were drastically reduced in plants with *fls2* (flagellin FLS2 host receptor kinase), *Atnoa1* (Nitric Oxide Associated protein 1), and *dnd1-1* (reduced production of nitric oxide) null mutations. Together, our results indicate that LPS modification and flagellin glycosylation do not affect recognition by plant receptors, but are required for bacteria to establish overt infection.

6.6 Final remarks

The results from this work provide insight into the diversity of AtsR/AtsT regulon and facilitate the potential of gene manipulation for redirecting information flow within bacteria to eliminate virulence in *B. cenocepacia*. We provided a comprehensive description of AtsR/AtsT phosphorelay mechanism that would be helpful for future studies investigating the nature of the AtsR activating signals and whether AtsR function can be manipulated pharmacologically to control *B. cenocepacia* infection. We further showed that the progression of *B. cenocepacia* virulence is both host- and strain-dependent. In fact, virulence and pathogenicity are properties ascribed to microbes, which require careful consideration of the host since the microbe's capacity to establish a niche in a given host is a critical feature associated to infection. *B. cenocepacia* is a perfect

example of microbes whose ability to cause disease is intimately related to the host's ability to recognize and respond to the infection.

6.7 Chapter 6 references

1. **Mahenthiralingam E, Baldwin A, Dowson CG.** 2008. *Burkholderia cepacia* complex bacteria: opportunistic pathogens with important natural biology. J Appl Microbiol **104:**1539-1551.

2. **Lee YA, Chan CW.** 2007. Molecular Typing and Presence of Genetic Markers Among Strains of Banana Finger-Tip Rot Pathogen, *Burkholderia cenocepacia*, in Taiwan. Phytopathology **97:**195-201.

3. **Uehlinger S, Schwager S, Bernier SP, Riedel K, Nguyen DT, Sokol PA, Eberl L.** 2009. Identification of specific and universal virulence factors in *Burkholderia cenocepacia* strains by using multiple infection hosts. Infect Immun **77:**4102-4110.

4. **Vergunst AC, Meijer AH, Renshaw SA, O'Callaghan D.** 2010. *Burkholderia cenocepacia* creates an intramacrophage replication niche in zebrafish embryos, followed by bacterial dissemination and establishment of systemic infection. Infect Immun **78:**1495-1508.

5. **Georgellis D, Lynch AS, Lin EC.** 1997. In vitro phosphorylation study of the arc two-component signal transduction system of *Escherichia coli*. J Bacteriol **179:**5429-5435.

6. **Jourlin C, Ansaldi M, Mejean V.** 1997. Transphosphorylation of the TorR response regulator requires the three phosphorylation sites of the TorS unorthodox sensor in *Escherichia coli*. J Mol Biol **267:**770-777.

7. **Kwon O, Georgellis D, Lin EC.** 2000. Phosphorelay as the sole physiological route of signal transmission by the arc two-component system of *Escherichia coli*. J Bacteriol **182:**3858-3862.

8. **Perraud AL, Kimmel B, Weiss V, Gross R.** 1998. Specificity of the BvgAS and EvgAS phosphorelay is mediated by the C-terminal HPt domains of the sensor proteins. Mol Microbiol **27:**875-887.

9. **Podgornaia AI, Laub MT.** 2013. Determinants of specificity in two-component signal transduction. Current Opinion in Microbiology **16:**156-162.

10. **Cheung J, Hendrickson WA.** 2009. Structural analysis of ligand stimulation of the histidine kinase NarX. Structure **17:**190-201.

11. **Hoch JA.** 2000. Two-component and phosphorelay signal transduction. Current Opinion in Microbiology **3:**165-170.

12. **Piggot PJ, Hilbert DW.** 2004. Sporulation of *Bacillus subtilis*. Current Opinion in Microbiology **7:**579-586.

13. **Shin D, Groisman EA.** 2005. Signal-dependent binding of the response regulators PhoP and PmrA to their target promoters *in vivo*. J Biol Chem **280:**4089-4094.

Appendix A

AtsT *in vivo* Chromatin Immunoprecipitation

After several attempts to purify DNA segments co-purified with AtsT, very low amounts of DNA were obtained (ranging from 1- 2 ng/µl). The minimum DNA concentration for a successful ChIP-Seq is 5 ng/µl. Here in this section I show the results of *in vivo* chromatin immunoprecipitation experiments and troubleshooting steps.

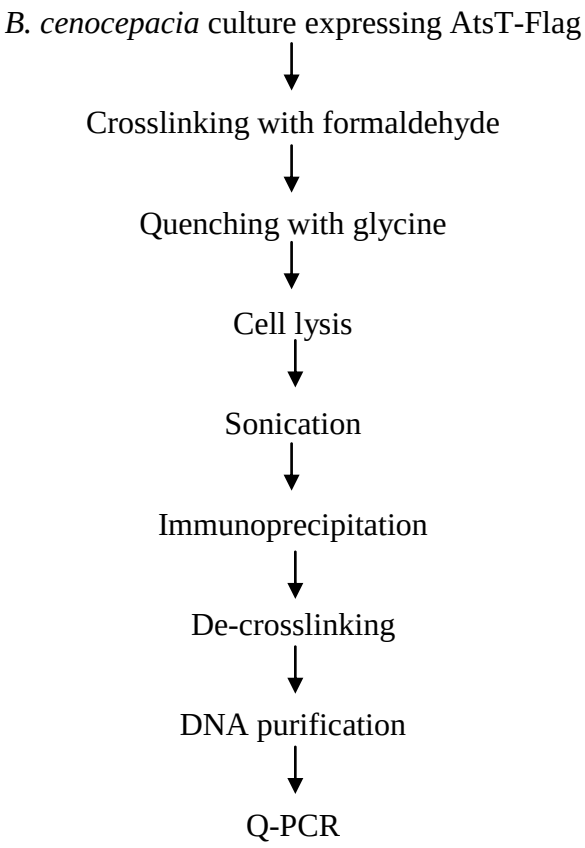


Figure A.1: A step-by-step preview of the ChIP procedure.

To create C-terminal Flag-tagged AtsT, *atsT* gene was amplified using 5108-2794 primers and cloned into pDA17. The resulting plasmid was introduced into K56-2 Δ*atsT* and tested on DBH-I milk plates to check if the flag-tagged AtsT is functional (Fig. A.2). Exponentially growing culture in LB was incubated with formaldehyde to a final concentration of 1%. Crosslinking was quenched by adding glycine to 0.5 M final concentration. Cell pellets were washed in PBS and resuspended in lysis buffer [50 mM Tris-HCl (pH8.0), 10 mM EDTA, 1% SDS] plus protease inhibitor mixture (Sigma). The cells were sonicated for 25 cycles of 30 sec on/ 30 sec off at max power with a Bioruptor, and debris was pelleted by centrifugation. A 1 ml of sample was taken as input control. Following centrifugation,1/10 volume of 10% TritonX-100 in lysis buffer was added to each sample followed by 100 μL of Dynal-Protein G beads coated with anti-flag antibody (Sigma), and samples were incubated overnight with rotation. The beads were washed 1 × with low salt buffer [20 mM Tris (pH7.5), 150 mM NaCl, 2mM EDTA, 1% Triton, 0.1% SDS], 1× in high salt buffer [20 mM Tris (pH7.5), 500 mM NaCl, 2mM EDTA, 1% Triton, 0.1% SDS] and then with LiCl buffer (0.25 M LiCl, 1% Nonidet P-40, 1% Na-deoxycholate, 1 mM EDTA, 10 mM Tris-HCl), and 1× in Tris-EDTA pH 8.0. Immunocomplexes were eluted with elution buffer (1% SDS, 0.1 M NaHCO3). Samples were incubated at 65° C for 4 h to reverse crosslinks, followed by incubation with DNAse free RNase A for 1 h at 37°C and then with proteinase K at 55°C for 2h and were purified with Qiagen PCR Reaction Cleanup Kit and quantitated with Sybr-green kit (Invitrogen). There was no difference between input and output DNA Q-PCR result which suggests that chromatin was not pulled down. All attempts for optimizing ChIP conditions failed and did not result in any improvement of flag-beads binding affinity which indicates that the flag tag is not accessible likely because of the crosslinking. To solve the problem, I

made N-terminal flag-tagged AtsT. No improvement was observed despite every possible trouble shooting.

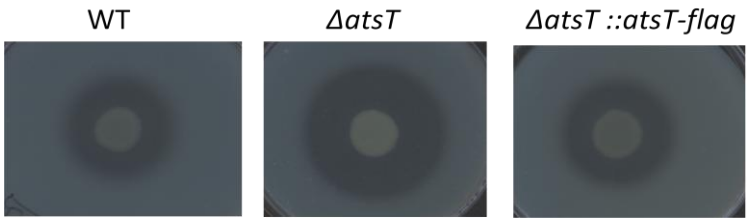


Figure A.2: *B. cenocepacia* expressing AtsT-Flag is functional. Complementation assay using AtsT-Flag on DBH-I milk agar plates to check the functionality of the tagged protein in parental and Δ*atsT* mutant strains.

Aim 1: To determine if AtsT-Flag is detected by Western blot

Aim 2: To determine the optimum time of AtsT expression

Which OD is the best for AtsT to be crosslinked?	No significant difference between OD_{600} of 1 or OD_{600} of 0.5.

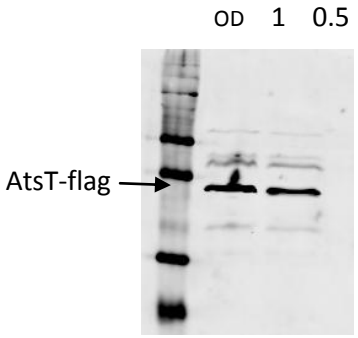


Figure A.3: Western blot result of *B. cenocepacia* expressing AtsT-flag. Numbers refer to OD_{600} of each culture.

Aim 3: To test the effeciency of lysis conditions

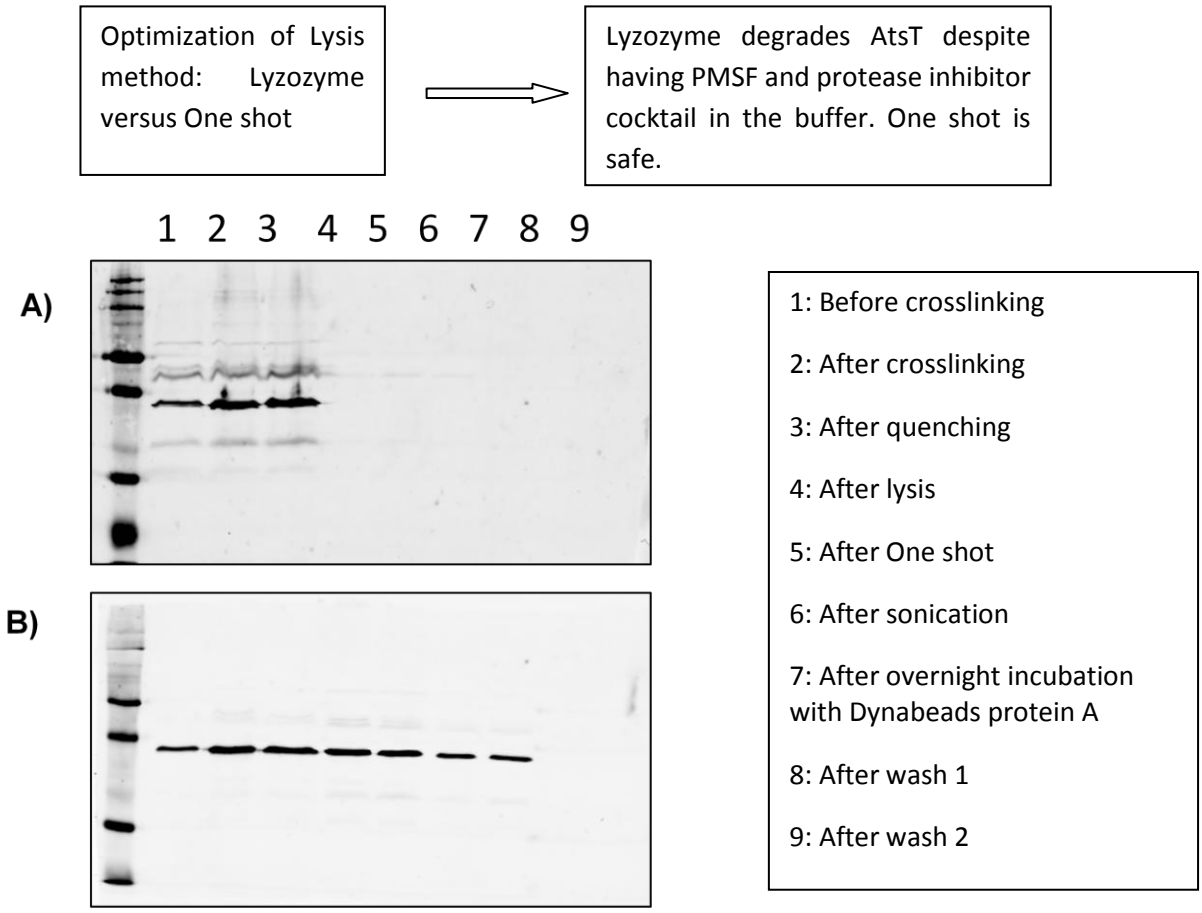


Figure A.4: Western blot results using Anti-flag to compare the efficiency of lysis methods. A) Lyzozyme was used in the lysis buffer along with Tris, NaCl, EDTA, sucrose, PMSF and protease inhibitor cocktail. Cells were incubated on ice for 40 min. B) Cells were incubated with lysis buffer without lysozyme and instead were disrupted by One shot system.

Aim 4: To determine AtsT-Flag binding affinity to different magnetic beads

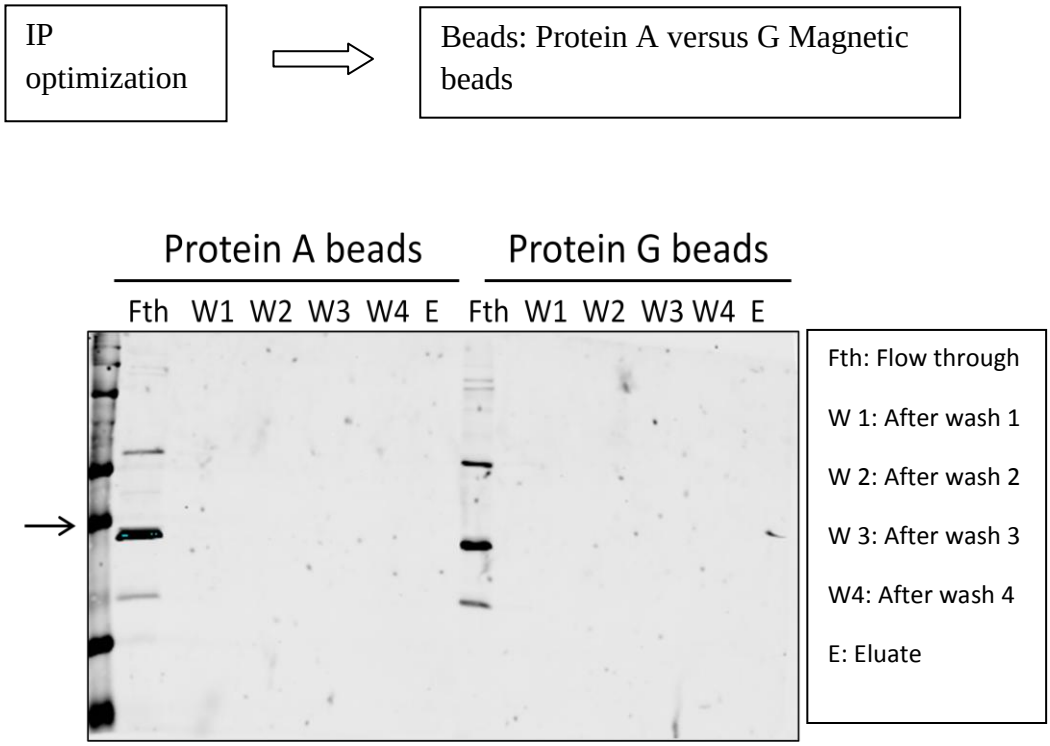


Figure A.5: Optimization of Chromatin immunoprecipitation by comparing binding affinity of protein A versus G. Western blot result using anti-flag for *B. cenocepacia* expressing AtsT-flag. Using protein A Dynabeads (left) or protein G (right) did not improve binding affinity.

Aim 5: To improve binding affinity

1- Block the protein G Dynabeads with BSA before adding anti-flag+sample:

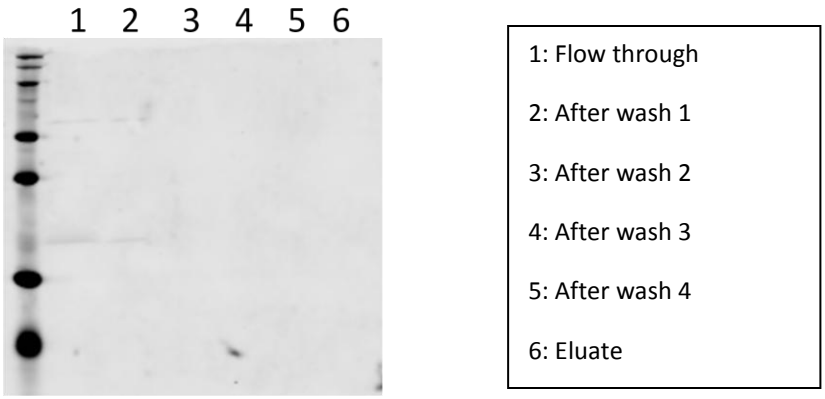


Figure A.6: Optimization of Chromatin immunoprecipitation by blocking beads with BSA before loading the sample. Western blot analysis of immunoprecipitated AtsT-Flag troubleshooting. Blocking beads with BSA before addition of the sample seems to reduce binding affinity.

2- Increase incubation time of anti-flag with beads from 30 min to 2 h.

3- Incubate anti-flag with sample before adding to the beads instead of adding anti-flag to the beads:

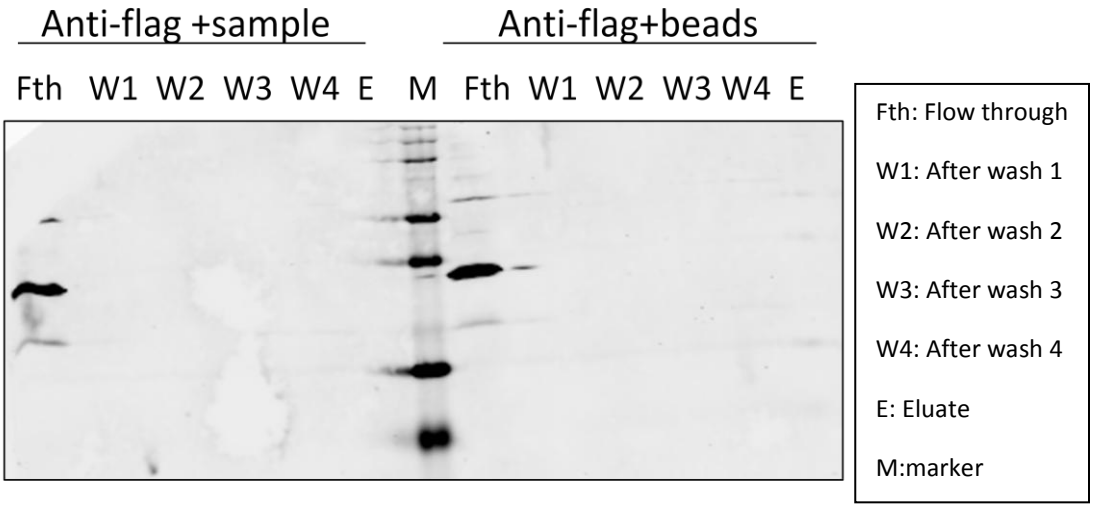


Figure A.7: Optimization of Chromatin immunoprecipitation by applying antibody to the sample versus to the beads. Incubating antibody with sample before adding to the beads has no advantage over incubating antibody with beads (2 h at 4˚C) followed by addition of the sample.

4- Reduce incubation time lysis from 40 min to 15 min

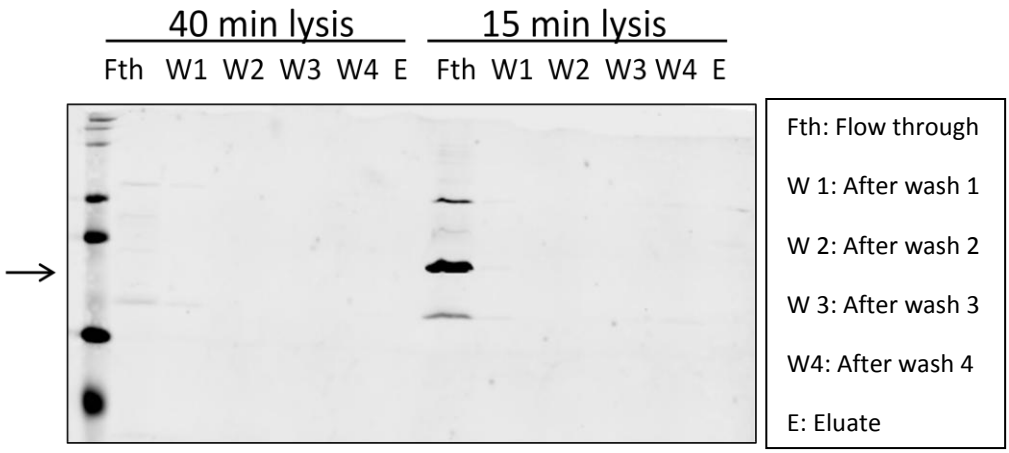


Figure A.8: Optimization of Chromatin immunoprecipitation by changing the incubation time. Western blot representing Co-Immunoprecipitation steps of AtsT-Flag using Protein G Magnetic Dynabeads. No difference observed with different incubation time in lysis buffer.

5- Boil beads with 1× SDS loading dye for 10 min to test if there is any protein (AtsT) bound to the beads.

6- Check AtsT-Flag binding affinity to magnetic beads versus sepharose beads

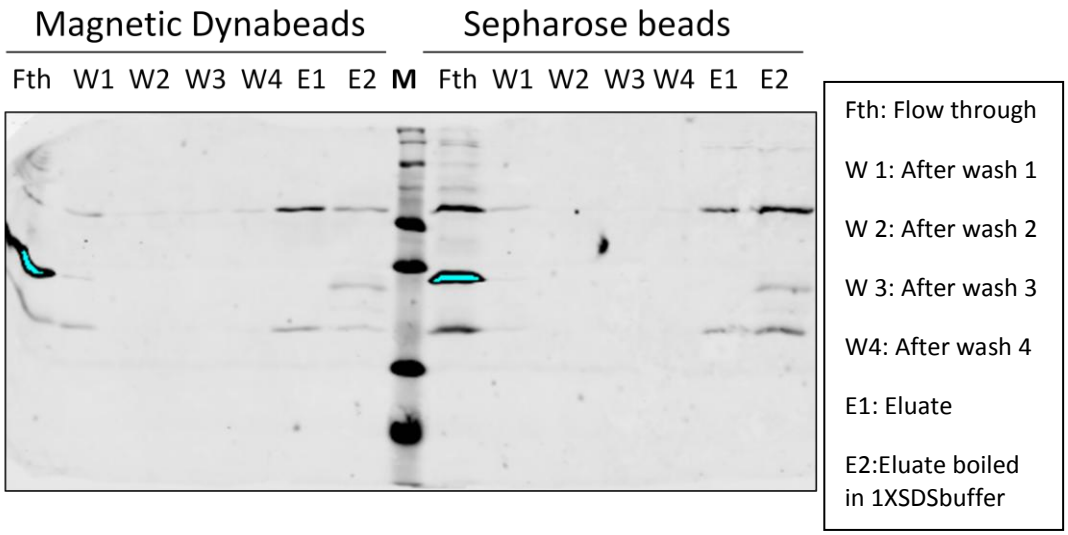


Figure A.9: Optimization of Chromatin immunoprecipitation by changing the beads. Western blot representing a control for crosslinking step of AtsT-flag using Protein G Magnetic Dynabeads and sepharose beads.

7- Changing elution stringency with either low pH or 2% SDS also didn't help.

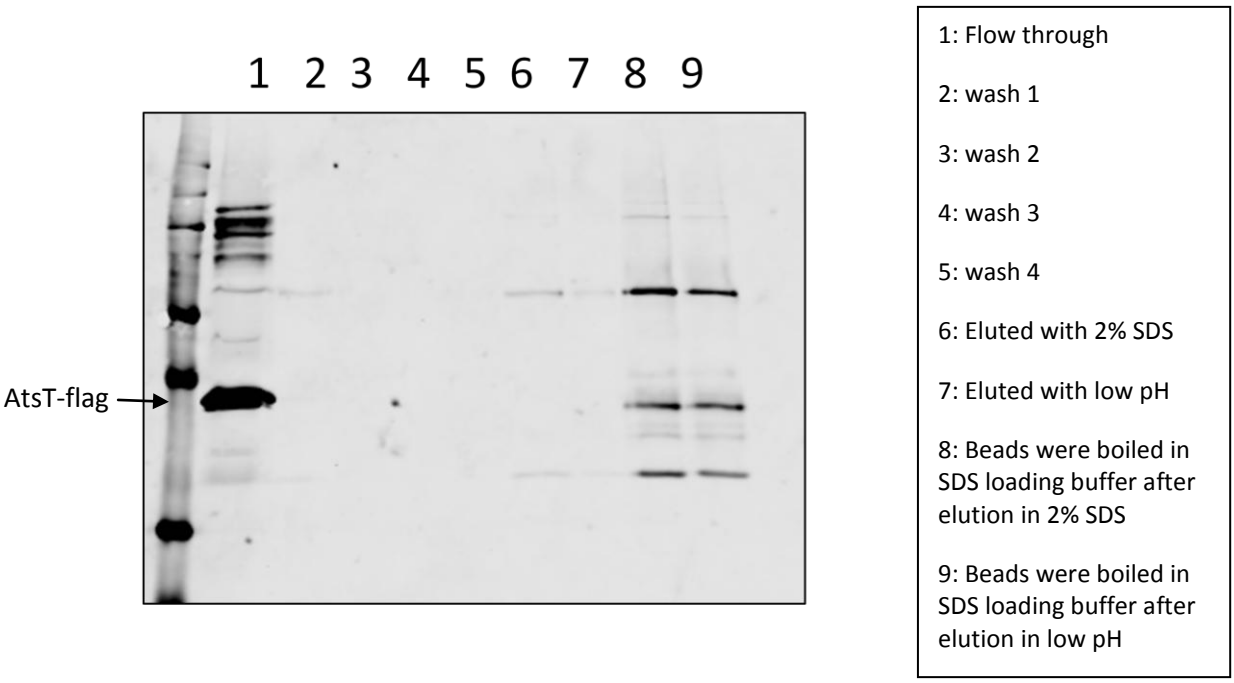


Figure A.10: Optimization of Chromatin immunoprecipitation by changing the stringency of elution condition. Western blot representing ChIP steps of AtsT-flag using Protein G Magnetic Dynabeads with modified elution conditions.

8- Test the accessibility of C-terminal versus N-terminal flag tagged AtsT.

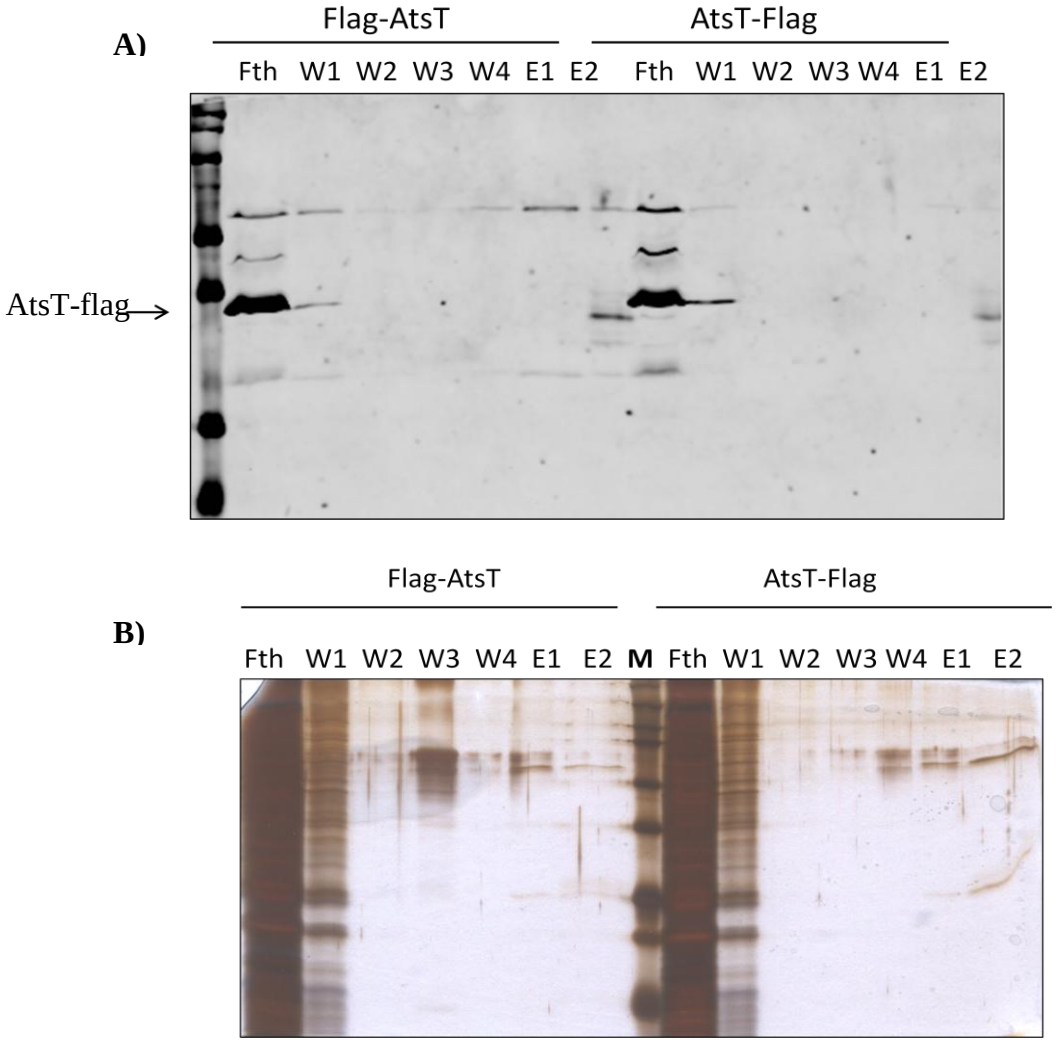
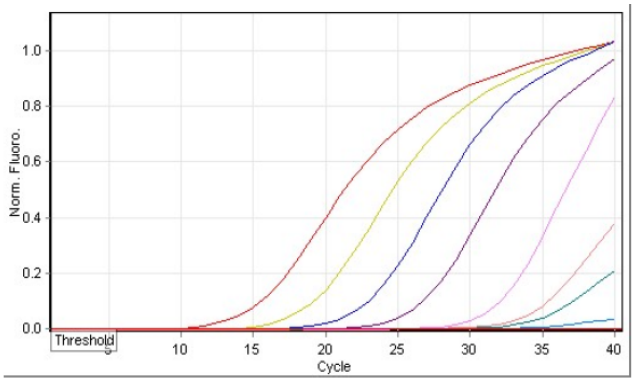
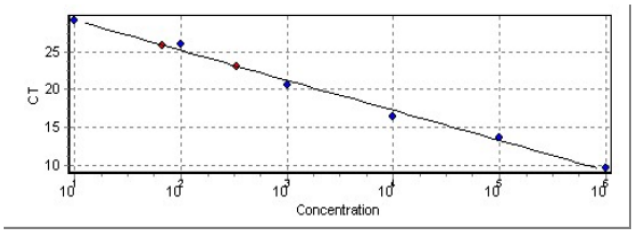


Figure A.11: Optimization of Chromatin immunoprecipitation of AtsT with different position of Flag tag. A) Western blot representing Co-Immunoprecipitation steps of AtsT fused to Flag tag (N-terminally) or (C-terminally) using Protein G Magnetic Dynabeads. Similar to previous results, AtsT is mostly in the flow through and barely in the elution fraction (E2). E1 represents elution fraction eluted with 1% SDS+ 0.1 M NaHCO3 and E2 represents eluate boiled with 1× SDS loading dye. B) Same samples as in A were run for silver staining instead of western blot.



Standard Curve



No.	Colour	Name	Type	Ct	Given Conc (Copies)	Calc Conc (Copies)	% Var
1		0627 10e6	Standard	9.63	1,000,000	830,733	16.9%
2		0627 10e5	Standard	13.63	100,000	81,570	18.4%
3		0627 10e4	Standard	16.38	10,000	16,471	64.7%
4		0627 10e3	Standard	20.56	1,000	1,450	45.0%
5		0627 10e2	Standard	26.01	100	61	39.2%
6		0627 10e1	Standard	29.09	10	10	1.7%
7		output	Unknown	23.09		333	
8		input	Unknown	25.83		68	

Legend:
NEG (NTC) - Sample cancelled due to NTC Threshold.
NEG (R. Eff) - Sample cancelled as efficiency less than reaction efficiency threshold.

Figure A.12: Relative quantitative expression of *Bcas0627* using input and output DNA. Based on the graphs, output DNA amplified 4.8 times more *Bcas0627* than input DNA. Comparing this with Microarray, amplification is minor (*Bcas0627* was upregulated almost 100 times in *atsR* mutant). Besides, this low amount of DNA (1.2 ng) is not sufficient for High throughput sequencing.

Any of the above mentioned conditions did not improve elution of the AtsT from the beads. I believe it is very likely due to the crosslinking that the binding site of Flag tag is altered and therefore tag is not accessible. I have also used straptavidin column and biotinylated antibody but neither were helpful. Attempts to create AtsT tagged with HA and His was not successful as none of the plasmids in a complementation assay failed to functionally complement atsT phenotype in protease assay.

Appendix B

Extracted sequences of AtsT using CLC genomics workbench

A. Extracted sequences from ChIP-Seq peaks with P value > 0.001

>9276..9410_NC_011000
TTCGATCCACGACCTAGAAGCCGCGATCAATTTCTGGCGCGCCCGCTCGCCGTCGAGCGG
CGACGAACTCAAACTCTGCGAAGAGGCCAGCGCGCTTTCCAAGCCGTATGCGCTGCTGAT
TGTACAGCGCGAAAG
>21292..21426_NC_011000
CACCCCCGCTACGCATCGTGACGGTCGAACGATCGTTTGGGCCGGCATGCGGCGCGCGCC
CGGCGTGTCACGGGAAAACGAGCATTGGCGGGGTAGGACGAGATAGTTAATTTATTAACT
TGTTTTTGGGGACCG
>26734..26884_NC_011000
ACGAGTTCCACACATGCTGACGCGACTATCAACCAGAGCTTTTGCACAGGCACGACGATG
TGCGTTGAATCGCTCGTAGGCATCGCCGAAGTACGCCTACCAAGGCATTTCACCGGTCCT
CAGCCAAAAAATCGCCGAAATGCACAGACCC
>32943..33127_NC_011000
TTCACATCATGAGGCGTCGTTTTATGCAGGTTTTTGTCGCGTCTAAAGCCTTGAATCACT
TGCAACTATCGCTTACAATCGCCCGGATTTGTTAGCTAGGCACCCCTGAAAGCTTTCAGA
AAGCTTGGGGTCGGGTTCAAGGATTTCAGTGGAAGATTGCGATGGCGGGTCAGGCGCCGG
ACGAC
>34503..34658_NC_011000
TTCCACATCCTGATTGTTCTGTTGCAGGCATTCATTTTCATGATGCTGACGCTGGTGTAT
CTCGGCCAGGCGCACGACAAGCACTAAGCGCGGCGTGCAAGAAAGAGTTTCCGTTTTAGT
TTTTTAAATCTCAGTTCCAAGTCTTTTCACAAAGGA
>42969..43118_NC_011000
CAGATAAGGCTTTCCGGGATTTTTTGCTGTTTTCGGGGTCCTGAAAAAGGTCATTTTTAC
CCATATTTTCCATCGCGTTTTCAGGGTGTGCGCGCGATGCGTTTTTCACGCGTATCGGGC
GCGATGAATTGCCGATTTTCACGCGAGAAA
>54766..54919_NC_011000
CATCGCGTACCAGACGCATACTCGGCCAGCCTTCCAGATGGTTACTGACATGTTGCGCAA
GGCTTTGCGTCGCCTGCAACCCGACGAAAGACCCATGTTGCACTCTGATCAAGGCTGGCA
TTATCAAATGCCCGGCTGGCGGCTCATGCTCGAG
>60818..60942_NC_011000
TACGTTGATAGGGAAACCGGCGGGATCGTGAATAAGGCAATCAAGCGGGCGAAATTCCTT
GAGTTCTTCGCGAACCGCGCGGCCTGTCTTATCGGGATGGAAGCGTGTGGCGGAGCCCAC
CATTG
>61221..61367_NC_011000
GGGCGAAGCTGGACAAGGAAATACCGGTCGTGCTTGGCCGAATCGCGGAGCGCTTGCCAG
CGGCGCTAATCGATACGTTGCGCGAGCAATGGAACGGGTTGGCAAAGCTCGACGAGCAGA
TCGCGGAAATCGAACGCCGGATGCGCG
>61713..61898_NC_011000
GGACGATCTGGGCTGTACTGGCCCATGACCGGCCGTACCAAAAGGGTTACGTGAGCGTGA
AGCCGGCCTGATCGGTGAATGCCACGTTGAAGATCAACGTTTAACACTGGGTGAACGTCG
AAAGGTTGCGCAGGAATCGAGTGTGATGACAAGCCAGGTAGGACCGGGACTCGCTAAACC
TGAATC
>114436..114625_NC_011000
ACTTCGACGACCTGGCCGACGAAGTTCTCGCCGTTCTGGAACAGGTTGAATCCCTTGAGT
TTGCGAGGCATACCCATCGTGTTTGACTCCTGGTGAGGCCGGTCGTTACGCGCTCACGCG
CGCGGCGAAATCGGCGAGATAACGGTCGGTGATGCGCTGGCGCAGCATCAGGTTTTCGAG
CGGCGGAACC
>136340..136504_NC_011000
TCGTTCATTGCTGGCCGCCTCAGAGCAGGATTTCGGCGCACCAAAACAAAACTTGAGGGG
CACGTGGAGATTCGGCCCGATTTTTTATCCTTCTTTGTCATTTCGGACCGGCCCGGCACG
GCGCCGGATTAAACGGGGGCAGAAACCGGCCCGATGCGGCTCGGG
>139291..139424_NC_011000

GCGGGCGAGGCAGACGGTCATGGCGAGCGAGTAAAAATGCGAGTAAGAATCATTCGCATT
GCGTTAAAATGTCGCGCATCGTATCAAAACCCGCCAACCGCGCAACAATTCCATGCCGCG
TACAAGGGCATCCG
>146554..146750_NC_011000
AATCCGCCTTGTCGGAAAAATTCCTACACGGTTTACAGACAAATCCTGCACGGCATACGC
CACTTCGCTGATACCGGCATAAATGCCGTTTGGCCAAAATTCGCCCTGTAAGACTATAAA
CGAGCCGGCTGAGCGTCCAGTGCGCACTGACCGAGCAAACGTTTTCATAACATGAATGGC
CAAAGCAAAGCTCTTTA
>159157..159313_NC_011000
CCCCGTTCCCGCCGCCGTTTTCCCGTGTCCGCCGCCGATGCGCCGCGCGGCGTGATCCTG
CGTCGAATTGACACGTGGCCACACTCCGCAGGCATACTACGCGTCAGCAGCAACGTAGCG
TCATTCGTAAGATTCATAAGTTGGCAGGCGGCGGGGG
>179296..179450_NC_011000
ATGTGACGTGACGTCGCGGGTCAGTGATGAGTCCGTTTAGCAAACCGCGTCATGTGGTTT
TCGCAAAACTGTCGTCCGATGAAAGCTTTGATTTTTTCGAATGCAGGGTAATTGCGTATG
TTGATAGACCTAATACTCAATCGCTAATGTCCTAA
>180018..180206_NC_011000
TTTGTTTTCCGGGCGCCGGCGGCCCGCTCGGCGCCGTCATAAAAATGTAAGAAACGTCGT
GTGAATGAAAATGCGACACGTTCGTATAAGCGGTTGATTTAAATGAGAAATTTGTGGGTA
TTGGAAGGGCTTGCGGGGCGCTTTTGATAGAATTGCGTATCGCACACCGATGCGCACACC
GATATCTCA
>188094..188290_NC_011000
GTGCCAAATTTGTGCCAAGGTGTGCAGCGCATCCCCCACGCCGCATGCATGACATTGCAG
CAAATTGCGGGTACGATTCTTGGCAATCGCGCTGAAGACGTACAGAACGGCTACGAGACA
AAAGCAAAGGGCAAGGGGTCTAGCAAACCGCACCTTGCCCACAAAGCGTGGAGCCACGGA
CTGTTCAGCAGCCCCTG
>189808..189937_NC_011000
GGGTCTGCCTCACACCTCGGTCAAGGCTCGCGAGAAGGCGCATGCTCGTAAGCGCGGTAG
ACCAAAATGATGCTGACATCGCGTGCCTATCGTTGTGGATAACCGAGGCTCGACTGTGAA
TCACCACGTG
>189969..190110_NC_011000
TGCTTTCACGGATGAGCTACTTCGTGCTTTTAGGGATGGGAACATCCCGAGAAGCCTTGC
GTAGTAAGGGTTTCAAGAACCGTAACGCGCGCGCGCGATTTAACTCTATGTAAACCTTAA
CGTCTGAGTTGTGTACCACGGT
>190543..190674_NC_011000
CCGCCAGCCGGGCATTTGATAATGCCAGCCTTGATCAGAGTGCAACATGGGTCTTTCGTC
GGGTTGCAGGCGACGCAAAGCCTTGCGCAACATGTCAGTAACCATCTGGAAGGCTGGCCG
AGTATGCGTCTG
>194354..194488_NC_011000
GGAGATCACTTGTGAAGATGAAAATTTTAGCGGCTAAAACGGTCGTACTCGCGGTGGCGA
TCGGCGGTGCGGTATCGCATTCCGCGCTGGCTGGTGGTATTCCAGTGATCGACGTTACGA
ACGTAATGCAGACAA
>195334..195490_NC_011000
AAAAATCAGTGAGGGCCACGCTATGAGATCCATGTTTTTCAAACGAGGCAATCTTTTTGC
GCTCGGTTTGATGATGCTCCTTCACGTTGCGACTGCCTTTGCCGCACCGATGGACGTCAA
TGTCGTTGATTCGGTGATGGGCAAGTACCAATCTGCG
>198277..198429_NC_011000
GCGCGACTCTAGTCGAGGTCGATCCGGCAATTTGTGCAGTCTGGCGCACAGTTTTTTCCA
GTGACGCCAGAAAATTGTGCTCTCGAATCCGAGCCTTTCGCTTCACGGAAGAGAATGTGG
AAGAATTGCTCGACAAAAAATTCAGAACGGATG
>198973..199102_NC_011000
GACACTTCCAAGGTCATTGATGATCCCTCTCTCCGACCAGTAATACGTGCTCGAATTGCT
GAGTCGCCGCGCTCGCCAACGGAGATTTCTCTTGAATCAGTTGTCCAATTGTCTTGTGTG
AGACCGACGG

>199448..199619_NC_011000
CCTCGATCGCTTCTTCAACAAGCTTCTCTCGGAGAGCGTGCAAGAAATCATCGCCGGAAA
GGCGAACAATTTTTGCTGCCTCACCGCCGGAGGTTATTTTACGCGGTACACCATCGCGCA
CCAGTTTGTGAAATTCCTGCGTTTCGTCTAATTGATCGAATGGTTGGGCGAT
>201152..201308_NC_011000
GACATATTTCAAGCGCGCGAGCGAGGTCGGTGACAGTGGTTTGCTCGGAATGAAATCGCC
CACGTTGCGCGATTGTTTCGGTGACGCCAGATGACCGAACCCAGACAGCGGTATCGCCAG
ATATTCCAATGCCGCTTAATGCATCAGCCACCCGAGA
>202436..202586_NC_011000
TGTCCGGGAAGAATGCCAAGTATTCGCTACTGCCGACGACTTGGGCAATGTCAGCGGTGT
CCATCGGATGGAACATGACGCCAGACGAATCCAAATCCCGAACATCGATGCCATCAAACG
GTACGATGTGCAGATGAGCATGGTCAGTACC
>203026..203161_NC_011000
ACATCGCATGTGTTTTCGGAATCCCGTGCGATGTCGCATAAACTATTGATTTTAAAGGAA
AAGTAGGTTTTCGTCAAATTTCGCCAGATGATCTGCGCAGATGCGATGCGTCATAGTGCG
CCGTTCACACTAGCGA
>235491..235673_NC_011000
TTGCGTCGGCATCTTTTGGACGCGTCTGCAGTGGCCGGCATGTTCGCCGGTTTTACGGGA
TGGTAGCACTGGGGCAAACCCGCCGCCCATCGGTTGCAACGCGGGGAATGCGCGGCGGCG
CCCCGCGAAAACGGGGCGGATGATGCGTTACATCATGTAAAAGCATGTAAAGCTGAGTTG
ATC
>245973..246126_NC_011000
GTACGGAAGTTGAGCCCCGCAGTCGCAACGAAGTCGTTCAGAAAGTTGTTGACGAGCTGC
GAAACACGGTTCATAATCTCGCTTCTCTGCTGCTGAAAACGCAGCGCTGCTGAGGGGTTC
AAGATCCCTCGCAGACATGCTCTTTAAAAATTAA
>253951..254114_NC_011000
CAGTTCGATCCTGCCCATGGGCACCAGCAGTAAAAAGTCAGTGTCTTGTTTGCGCGCGGC
GCAACCTGTGAAATTCCTTTTGGGAGTTGAAAATGGCCAAGGAAAAGTTTGAGCGGACCA
AGCCGCACGTGAACGTTGGTACGATTGGTCACGTTGACCACGGC
>254849..255044_NC_011000
CCAAGGTCAGGCAGGCGACAACGTGGGTATCCTGCTGCGCGGCACGAAGCGTGAAGACGT
GGAGCGTGGCCAGGTTCTGGCGAAGCCGGGTTCGATCACGCCGCACACGCACTTCACGGC
TGAAGTGTACGTGCTGAGCAAGGACGAAGGCGGCCGTCACACGCCGTTCTTCAACAACTA
CCGTCCGCAGTTCTAC
>257551..257712_NC_011000
CGCAAGATCGCACTGTCGAGCACGATGGGCGTCGGCGTGCGTGTCGACCAGGCTACGCTG
GCAGCACAGTAAGCAAGTATTCGGGCCGTTTCGTTCGCGAAGCGGCCCACATGGGCTTTG
GGCGGTTGTTCGTGCAGCAGGGCGGACAACCGGTTATCAAAG
>257756..257891_NC_011000
GCCAACGCAGATGGCGAACCCGAAAAAGTTTTGCAGTGGTGAAGCCGCAGGCCGCGAAGC
GATTCGCGGCGTGAGGTGGAAATACTCCTAACGAGGTCGGACGCCGTTGTTGAACGAGGT
ACGTGAGGTTTCGGCC
>259316..259470_NC_011000
AAGTTCCGTTCTTGCTGGCTACTCAGCTTGAATCATTCAGCACGTTTCTGCAAGCCGATG
TGCCCGCGACGCAACGTAAGCCCGAGGGTTTGCAGGCCGCGTTCACGTCGGTATTTCCCA
TTGTTTCGCACAACGGTTTCGCGCGCCTCGAGTTC
>274108..274264_NC_011000
GCGGCACGAAGCGTGAAGACGTGGAGCGTGGCCAGGTTCTGGCGAAGCCGGGTTCGATCA
CGCCGCACACGCACTTCACGGCTGAAGTGTACGTGCTGAGCAAGGACGAAGGCGGCCGTC
ACACGCCGTTCTTCAACAACTACCGTCCGCAGTTCTA
>293238..293377_NC_011000
AGCTTGCCCTTCACGCCCTGGTCGCGATACGCGTCGAGCCCCTTCTGCGTGTTGCGCAGC
GCATTGATGCTTTCTTGGCGCGTCAATTTGTCGCACTTCGTCAGCAGCGTGTGGATCGGC
TTGCCGGTCGGCGCGAACCA

>359830..359969_NC_011000
TGCTGCGCGCGTTGTACGAGTGGTGCACCGATAACGGTTACACGCCGCATATCGCGGTGA
GGGTCGACAACTCGACGCGCGTGCCGCGTCAGTTCGTGCGTGACGGCGAGATCGTGCTCA
ACATCAGCTTCGAGGCGACG
>365148..365321_NC_011000
CCGGTTAATGTGGCGTTTGCGATGTAATCGATTTGGTCGTGCCCGAATTAAATCGCCAAG
TGTGTTAACTCTTGTTAATTGACATTCGTGCGTCATGTTTAACCGCAGGAGACCGAAGAC
AGCATATTCAGCGGTTGCGAAGGCGCACCCTGTTGCGGTGCGATGACCGTGCCG
>365924..366070_NC_011000
ATCCCCGACATTGCACCGTTCAACTATTAAGCACGCCCGCGTCAGGCGAGCATCCGGAAA
TTTTTTCCGCCATCCGGGCGGCGCGTCGTGCAACCAGTAAAACTTGCGCATTCCGGTCGA
AAGAAAAACCGCTGCGAGCCCGCAGCG
>368811..368977_NC_011000
CGGGCAAGCCGGGCTCGTATCGTTCGGTCGCGTTCCTGCGTCCGCACTTTCAGCTCGACG
AACTCTCGATTTCTCTGCGACTTGTCGCTGATCTGCCCAAACCGGCAAATTCATAAGCGA
GTAAATCGCCCGGGCGGGCCGCCTGGGTAGGACGTTAAAACCACCTC
>383808..383943_NC_011000
CCAGTAAATTTTGCACATCGTTTAATGCGCCGCGCGTAACGCTTGCCGATTTAACAATCT
GACTCCGCGTGCGCGCGGCGCGATCGTTCGCACGGCGGGTAAGACGGCCGTGAAACTGGG
CAACGGATTCGAAAAC
>400055..400202_NC_011000
CGAGAAGTGCGCGAACAGATCGTCGCCGCCGCCTTCCGGGGTGATGAAGCCGAAGCCCTT
TGCGTCATTGAACCACTTGACGGTACCGGTTGCCATGTTACTTGTTCCTAAAAAGATAAA
ACGGGGCCGAAGCCCATGGGGTGCGGAA
>400302..400447_NC_011000
TACGGCTAATCGGTCTGGTGGTCAACCGTATTCCCATGCTCTCAGGGTTATTGCGGAGAT
CGGGTTTACAAAGCTTGCAAGCGATGCGAATTTTTTGTAGGGGCCGATCGATTTTGGCGC
CACGTTCGAGGTGCAACCGTTAAACT
>427479..427638_NC_011000
CATCGCGTACCAGACGCATACTCGGCCAGCCTTCCAGATGGTTACTGACATGTTGCGCAA
GGCTTTGCGTCGCCTGCAACCCGACGAAAGACCCATGTTGCACTCTGATCAAGGCTGGCA
TTATCAAATGCCCGGCTGGCGGCTCATGCTCGAGCAACGC
>452189..452342_NC_011000
ATATGATTTTCTGGAGGATTAAAATTAATTTTGAGCGCTATCTTTGTCATAATTTTCTGG
AAATAATTGATGTTGGCGTGTGGTGAACGTTTGGGATACTATTTTAGGTTCGGCATATGA
CGTCAAAAAACACGGGGCTGCTTATTTATTTCGA
>454300..454422_NC_011000
AGTGTTTGATTTGGTCGTTACGCGTGTTCAGCAATGGATTTGCTTTCAAAATCAGACGCG
CAGCACCGTCACCTTTTGGGCGAAAAACGACACGCCAAGCGCTCGGCGGGTTATGCTCGC
CGC
>461607..461739_NC_011000
ATGGAGGGAAAGGCGATGCTTGTGACCATGAGCCGTGACATCGCAGCGCGGCTATATGAC
GAGATTTGCAAGCTACGGCCGGCATGGCATGACGAGGACGACGCACAGGGCGCAATTAAG
GTTGTCATGACCG
>497535..497683_NC_011000
TGTTCTCTGCATGTTGTTGACAATCTACAAATCATCAATTTTAATGTCAAGCATGGAAAA
CACCTCCCTTCGGCGCGGCGCGGGCGCCGACGGCTCACACCCTTCGACCTGAAAGGAATG
AACGTCATGACTGAAACGCCTGTATCCGT
>562335..562477_NC_011000
GGTATGAACGGCCTCGGATGGCACAACGGATCTTACCGAAACCCGTTGCGCGCCGCGCCC
GCCAGAATTGACAGGCCGCGTCAGGCTACAGGGGATTCATGACGGGCAAAAGTCGCGATA
AACCCTTGTGCCGCAAGGCGTGT
>567233..567372_NC_011000
CTTTTCGCGCTTCGATAAACCGTGAACAATGGGATCACGCGCAACGCTCCGCCGAAAGCG

GCGCGGATCGGGCATTTGCCGGCTCGCCAGCGGGCCGGGTTTCAAGATTCAAGAGTGGGG
AACCATCATGCATATCGGGT
>582886..583032_NC_011000
TTCGCGCGCGCCCCGGCCTTGTCGCTTCGTATGCGTATGGATTGAGCCGGCGGTCGATTT
ATGTAATGATATCCATTCTCATTTTCATGCTGCCCGGGCGGCGCGCACTCCCTGCGCGCC
GCCTGTCGTTTTGAACGAATGGATACT
>596436..596589_NC_011000
AACGCGGGGAAACGGTGTGTCGCCCCTTCGCGTTCGCACCCGAATTGTTTAGTCGTCGAA
CGAAACATGCCGCGAAATGGACGGAATTGTCGCTCGCGCGCGTGGTTGCATTTTTTACCC
ACACTAAAAATCGGCAAAAACCCGGCGGGCGTAT
>596665..596796_NC_011000
ACCGGTTCCGATCTGGAGACCTGGAGGAAACATGGAATACAACCGTCTGTTGCACACCCT
GCGTGTTACCGCCATGGCAGGCGTGGCAGCGGCATCGCTCGGCGTCGCGGGTTCTGCATT
CGCACAGATCCC
>642911..643059_NC_011000
GCCCTCTACATTGAATTCGGTCACGTCGGTCACCCACTTCTGATTCGGGCGTTCAGCATC
AAACTGACGCCGCAGTAGATTCGGGGCAACGCGGCCGACTTCGCCTCGCCAGGAACGATA
TCTCTTTGGGCGTACCAGAGATTTCAATT
>657042..657190_NC_011000
CGAGCCAATAACGGCGTCAAATGTTTTTTAATGGTCCTGCATCAGGGCTTTCCCGAGCAG
TGCGAACAAAAGATTCGCAAAAGGAATAATCGGCATGCCGTTCATTATTTCGAACGGCTT
TCAGGGTTTTCCCCGCTGATGTGCGGCAT
>745992..746129_NC_011000
TCGTTCGGCAAGTTCGCCGTCGATGAAGATTTCGCGCTGCGCGTTCTTGTCGACGATTTC
GTATCCGCCGGCATTCATCAGGTGATGGCCGGCCTGCGGCGCAAATTCGACGACGCAGTA
GTTGGGGCTGTTGTAGAT
>795820..795990_NC_011000
CTGCAGATCGTTGAAAACGTCGCTGGGTTGCTTCATGTGATTTCCCTTCCTTGCACAAAA
AATGTGCATCTTCAGTTGCGTACGTGCCCATGATGCACGCATGACCGGAAATGGTGCGCG
CGGGGCGCGAATTCCGGGCATGGCGCCCGCATGGGTGAATGGCCCGGCGAT
>812763..812898_NC_011000
TAATGCGTCAGCTTGTCGTCCCGGGGTCCCCGCAATACGGGGCGACGAGTCCGATGCAGG
CATGCATCGCATGCCATTCGCGTGAGGCGGCGATGGAAGCAGCGAGGGTTTTGGCGGAGA
CGGCGGACAGCGAACC
>837668..837830_NC_011000
AGCGAGAACATCGTGCTCGATGATGGGGATGTCATTACGGCCGGCGCCATCCTTGCATGG
AATGATCTTGGTTTCTTGATCGTCGATCGCGTCATGGGCGCAAATGTCATGCTGGCCACA
TCGCGTTTCATGCTCAGCGAAGGGGTGCGTGACGACCAGCGGC
>838191..838352_NC_011000
GAGTCGATAACGCCTTTCGGATGAACGATGTCCATCGTGTTTCGCACGTTCGATGTGCTC
TCGGCGTGCCGGCCGTGAAGCCGCTCGATCGTCGCGATGCCGTGAACGTGATGCGCGTCA
CGCGTCACGGCATGACGGGCGGATGGTAGGTGAGTGTGCTCG
>919933..920067_NC_011000
GTGCCGGATCGCCGCGTCATGTCAATCCACCGAGTAATGAACGTGTCATCCGATGTAAAC
ATCTGGTAATTTTCTGCCGGATTATGCAAAATCGCCGCTCCCAAATACAACTATCGAGGG
ATTCCGGGATGCAAC
>1036324..1036462_NC_011000
TTCAGGTCTAGAGCATGAATTCTAGCCCGATGAAAGTCTTTCGTTCCAAGCGACCAGCGG
TACTTTTTCTTATTGCGAATGATTCTCAGGTAGTACACGCCGTGACGTGACTGGTACAGA
CGGGGTAACTTGTTCATGG
>1121819..1121963_NC_011000
GACGCATACTCGGCCAGCCTTCCAGATGGTTACTGACATGTTGCGCAAGGCTTTGCGTCG
CCTGCAACCCGACGAAAGACCCATGTTGCACTCTGATCAAGGCTGGCATTATCAAATGCC
CGGCTGGCGGCTCATGCTCGAGCAA

>1147001..1147192_NC_011000
GGCTGCAAGGTACTTCAGGGGAGGTCATGCGATTGCGTTCTGCACAAACCGCAGACTGCC
AGGCACTTTTTCTGCTCTCGCATTTGTGTAATTCGATCGTTTGACGAGCGGTGTTTTCGC
GACAGCGATGTTGTTTGCCCCCCTATTTTTCTTGCGAGTGGCGAATTTCCTGCGTATCGA
GAGTTGGTTCCC
>1147412..1147552_NC_011000
TTGGTTACTTTTCTTTGGCAAGACAAAGAAAAGTGACCGCTGCCCCGCGCAGGGGCGACG
CTAATAGACCGACACGATCACAGGTGCTTCACGCGACCGGAAGCCGACCACAGAAAAGCC
GGAAAAAGCCCCAAAATACCG
>1156549..1156736_NC_011000
TCTCCCCGTTGTACAAGTCCATCACCGGCGATAGATAGAGCTTTTGGCCCTCTACATTGA
ATTCGGTCACGTCGGTCACCCACTTCTGATTCGGGCGTTCAGCATCAAACTGACGCCGCA
GTAGATTCGGGGCAACGCGGCCGACTTCGCCTCGCCAGGAACGATATCTCTTTGGGCGTA
CCAGAGAT
>1225455..1225610_NC_011000
CTCTAGCGCGATCTGCTGTCGTCCATGTAAACGAAAGGTGTTTTATGACGTTGGATGAAA
TTCCACCGGTTCTGAGTGTCAAGCAATTTTGCTCAATCATCGCCAAATCCAAGTCGTGGA
TCTACGAAGAGTGGAAACGCGAAGGAAGCGATCTTC
>1228831..1228988_NC_011000
CGTTGGTTGGTCCTGGCGTTGACGCTGCGCTACAAGCCTCAGTTTCGCCGTTGGATTACG
CCAGAAGTGATCCAGCAGCATCGTGACCGCTTCTTCGCCGCTCGTCGGTTTAACACGTTG
ATGTCAGGTATTGCGAACTACGTATGGGCTATCGAACA
>1246576..1246713_NC_011000
ATACGAGGTATACAACCGGCGGCACGCGTATCGACAACGAAACTGACGCGGGGCGAAACA
ACTCGGCCCGACAGTGCAAAAAACGCTCCAATCGACTTTCCTAGAATCGATGATTTTAGA
GAAATTCCGTCAATATTT
>1278932..1279064_NC_011000
AACGGATGACGACGCTTTATGTCACCCTGTGCGCTGTTCGCGATCATTTGAAAAAACTCG
GGCAAAATCCACCCGGCGCCCTTGAGCCGCAAATTGCTGCAACCTCCACGGATGCCAGTG
GGCAGGATCATTT
>1284699..1284832_NC_011000
TCTTTCTTTGCTTGGTGGGCATGCGTCCACCGCTGCCTCTATCCCGGTGATTGCCTGACG
CGTCAGATCACGCGCTACTTCAAGACGTGAGGCGGTTGGTTTTTCCGTGAGTGCGGCGAT
TGACGATCTCGCGC
>1285245..1285400_NC_011000
CAGGCTGAAGAAAATCAGATAGTCGCAAGCTTGGCGAAAGTACTTAGGCTTCGTCGCGAA
CCACGCTCCGATCGCGAGAAATGGGGCGAGGGAGATCCACAATATCAGGATTCCGTCAAC
GCGAGGCTCCACATAGGCAATGAAAACAAATCCGAA
>1312803..1312959_NC_011000
TTGCTTCTGCAATCGGTTGCGCGGTTCCCCAATCCCCCAGTGTCACGTTTCCATTCGGAA
GTCGAATACTTGGAATACCTGCGCGGCATGTGGCGTCAGCGCAAGCTTGGCTACAAGGAC
AAGGAATTCGCTGACGGCATTCTCACTATGGGGTGTA
>1313626..1313770_NC_011000
CGGATTTGTTGCCGTATCGGTTTATTCTGGCTGACGAGTTTCAGGATACTGCGCCGGGGC
GTGGTGACATGATTCGCAAGATGCTACATGCGCGCGAGGACAGTTTGTTCTTTGCAGTGG
GGGATGACTGGCAGGCCATCAACCG
>1314717..1314878_NC_011000
CCTGGTTCCAATCGTTTCGAGCTGAATACGCATGCATTGAGGCGCCCAAGTGGACCCAGG
TCGTCAATTTGTCTTTCACAATCCCGAAGGGTTTCTTGCAATCTCCGTACAGGCGACGCT
GAGGGATACATCTGCGGCGGACGAACCGAACTATGCTCGGTT
>1405446..1405606_NC_011000
GGATCTGAGCGGTGCCGGAATCATGGGCGGTGCCTTGCACGGCGACGTGAAGCGTCAGTC
GTCGAGCTGCTTCGGCGCGCGCGGCGCGACGCATTCGACCCAGGTCGATGTCGACGAAAT
CAAACTTCTCGGGGAAGGCAAAATACTCGATCAGATGCCTG

>1428987..1429140_NC_011000
AAGTGCCGGTTGATCAACTTTCCGGACGAGCAGCGCCGTTTCATCGCGACGCTCCCCGAT
TTCGTGCATGCCGCGTCCGATGCTACACGCTACACGAGACGCTTTCCTGACATCCCGCCG
ACGTGGAACTATCGGTCGGCGGACCCCCTGCCTC
>1436790..1436941_NC_011000
GTTGCGCGCAGTCCGAGGCCGGTTGCACCAGAATCCCATACGAACGCTTGAGATTTGCCG
ACGGGGCAAAAATGTCCGTTGACCCGCGCCGCCGTGAAATTGATCTTTGCCACTCAGACT
GCCCTCTTGAGAAAGTCGGGTTGCGTGCATTC
>1656753..1656896_NC_011000
ACGCCAAAATAACCCGGCAGCAAATCGGGACAATCACCAATCTTGTGCAATTCGCCCGTG
TCACGTGCGTTTGCGAACGTATTGCACCGCAGCGAAACATGTGCGGCGGGTGCCGGATTC
GTTCGCCGAAGCCGGGTTCGCACA
>1704707..1704879_NC_011000
GCAGTCGTGATTGTGTCGAACCGTAACCCACTATAGACGTTCCGGCGCATTTTCGCGATC
CGGCCGAATGGCCAAAGCCGTTTCGTTACGTCGCGTTACGACTCTCGCAGCCCCGTATCA
CCTCGGCGCGGGCGGCGCGCTTACATTGGTTGCGTGAGCCGGGCGGGCCGGCT
>1758173..1758326_NC_011000
CATTGGAATTCAGTCGCACCAATTTCAATCGTGAGATGCTGCGAAATATGACAGCTTGGG
AGTTACGCGCTACGCTCGCATTCGCGATGACAAAATGGCAATCGTATTGGGAAGCGCGAA
ATGCTTAAACTTGCCCGGACGATATCGAGCTGCC
>1762565..1762702_NC_011000
AGCGCAAGGCTGAAATGGACGATGGCGCTGTCGTATCGTGGCCGGCCATGCGCCCGCTGT
TGCGCCTGATGAAGATTCGCGCGAACGATCACCACGCGTTCGATTGCGAATATCAGAACG
ATCCGACCAATAGCGAGA
>1818999..1819137_NC_011000
AGAGCATCGCCAATCGCGAACGCTTTTTGCGTCGCGTCAAGAACTACATTCGTCGTGCCG
TGTCTGACGCGGTGCGCGACCGTAGCATCAAGGATATCCAGAGCACGCAGAGCATCACGA
TCCCGCGCAAGGACATCGC
>2687201..2687357_NC_011000
TCATCGCGTACCAGACGCATACTCGGCCAGCCTTCCAGATGGTTACTGACATGTTGCGCA
AGGCTTTGCGTCGCCTGCAACCCGACGAAAGACCCATGTTGCACTCTGATCAAGGCTGGC
ATTATCAAATGCCCGGCTGGCGGCTCATGCTCGAGCA
>2736764..2736908_NC_011000
AACGTTTCGACCCTGCTCCGATCATCTATCCGACAATAACGCCGCCCGCGCGCACGCGGC
CTGCCGTTCGGCGATTTGTCAAAGCTTTGCAGTTTTCCTATGACAATTGCGAAACAATGA
CGCGGCATCAGGGCGTGCATGCGTC
>2971388..2971550_NC_011000
AGACGCATACTCGGCCAGCCTTCCAGATGGTTACTGACATGTTGCGCAAGGCTTTGCGTC
GCCTGCAACCCGACGAAAGACCCATGTTGCACTCTGATCAAGGCTGGCATTATCAAATGC
CCGGCTGGCGGCTCATGCTCGAGCAACGCTCGTTGGCGCAAAG
>3000190..3000361_NC_011000
CGCGCTTGGCCGCATGCTGCAAAGCAAAGCCGGAAAAACCCTGAAAATGGGTTCTTTACA
ACGACCCTCTAAACGTCTAAAAATTCCACTTGACACTCCAATAAAGAGCGCCAACAATCA
AGCTGGCACTTTTTTCAATTTGCCGGTTGGCGAAACGAAAGAGGCCGAGGTG
>3052746..3052882_NC_011000
GTGCACGCACGATTCTCGCGCGTTCGCTATACTGGCCCCAGATTGAAGCGTTTGCATAAC
ATGACGACTCTCGCCGATCTCCGCATCAACTATTCACGTGCTTCGCTCGACGAAGCCGAT
GCCGCCCCCGACCCCTT
>3241596..3241765_NC_011000
CCAAAGCTTTACCAGTGACAGCAAGAAACCCGTGACAGCTCAAGGCTTCGCGGGTTTTTT
ACTTCCAGCAACGCCAAAGTTTCACCATCGACAGCCGGGGGTATTTGGGGGTATTTTCGG
GGGTATCAGACATACCCCCGTGCACAGATACCCCCATGCCACTGACCGAC
>3241865..3242019_NC_011000
CGACTCAAGTACCGCGCTGCCGGTAAAGAGAAACTTCTGGCGCTCGGCGTCTATCCTGAC

GTGACGCTAGCGGCGGCGCGCCGTAAGCGCGACGAGGCCCGGGAAAAGCTGGCCGCTGGC
ATCGATCCAGGCGAGGCAAAAAAGGCCGAAAAGCG
>3242459..3242611_NC_011000
CCACTGCTCTTCCAGCGTCCCGGTGAGCTGCGCGCCGCCGAGTGGTCCGAGTTCGACCTA
GACGCAAGCACGTGGGAAATTCCGAGCGAGCGGATGAAGCGCACCAAGCAAGGCAAAGCG
TCCGGCGGCGCGCACATCGTCCCACTGTCGGCA
>3242630..3242822_NC_011000
CGCGATCTGCACGCCTTGACGGGCAACAGCCGGTTCCTGTTTCCGAGTGTGCGGACGAAG
GACCGCCCCATGTCCGACAACACCATCAACGGTGCGCTGCGCCGGCTCGGCTACGATGGC
GACACGATGACCGGTCACGGCTTCAGGGCGATGGCTCGAACGATTCTCGATGAAGTGCTC
GGCGTGCCGGTTG
>3242903..3243030_NC_011000
CGCCGCGAGATGATGCAGCGCTGGGCCGACTACCTCGACCAGCTCAAGGCAGGTGCCGAA
ATCATCCCCATCACGGCCGCGATGGGCAAATGACGTCGGACAGCCTGATATGGCTGCATG
ACTCTTCT
>3243346..3243472_NC_011000
GATAGTCGGCAACACTTCCGGCCGGAAAAATTCGACAAACTCGCGTGTGTATTTCATCCG
CGTATTCAGGCCCAGCGCACGTTTCTCCCCGACGGAAAGCGGCTCGTCGCCTTGCGTAGC
GCCGGCG
>3243739..3243870_NC_011000
CAATCCGGGATCGCTCAGTATCGCAAAAGCAGGCCGCGCGCGCACGGCTAACACAATCGA
TGTGCAACAAGTGCAACGCGTGCAACACGACGATCAGCAAAGGAGCGCAGGTGTTGCACG
TCGGCTCGATCG
>3244006..3244151_NC_011000
TCCCTTGCCGGGTTGGCGCGGTTCCCACACAGGGGGCTTTGCGTGAAGGGACGCAATGGG
TATCAAAGAGGACCTTGCCGAGCGTAAGGCCGGCTTCGTCAGTTTGCGCGCGCTGCTGGA
AGACATATGCGCCCATGAGAACGTGA
>3244246..3244373_NC_011000
CGTTCACGGTATTACCGGATCGCGCTCGACGACCGCGTGGCGCGCGCTTCAACACATCGT
CCGCGAAGGAAAATTCGCAACATCATGGGAGCTCAATCCTGATGACATTCCGTTCTAGCG
AGTACGAC
>3244565..3244728_NC_011000
AGACCCGCACGCTTCCGGATACAAATCCGACGACGAAGAGGCGGAAATTCATAGATGGCG
CACCGTGCTTTCCCGCGCGCTCTACGCCACCGATTCCGGCGATCGCCTTGAGGCTGAATC
GGAAGGGGGAGACACGCCTGATTGGTACATCAAGCCGGCAAATC
>3244854..3244992_NC_011000
GCAGAGAACGATCGCCTTCGTACGGAGATTGCACGCCTCTCCGAACGCATCGTCGGAAAA
AATCGTGAAATCGCCAGCCTCGAACGCACAATCGTACGAAATTCTCTGTTGGCGTCAGCG
CGTCAAGAGGATCTGGAAC
>3245237..3245374_NC_011000
TACCTTGAGCGAAGCGCAGCGAAAGGCGGTGGAACTGGTCGCCTGCCCCGTTGACCGTAA
TCCTGCCGCCAAAAGACAACCCTGATGGGTGGGTGTGCGTACCCTGAGGGTCACAGTCAT
CAATGCAAGAACCATTAG
>3245405..3245582_NC_011000
ACAACCCTGAGGGTTCGGCGATCCGCCGACCCTGAACCTTCAGGGTATTGCCCTGAGGGT
AGCGTCGTCGCTTTCAATGCAGCCATTGCAAACAACGCCCAACCGGGCAAAGGAACTGCA
ATGGCAAGCAATGCCGCAACCGCGCAGCCAGCCGGCGGCGCACCTCTTCCCGTTCCCT
>3245898..3246077_NC_011000
GCGAGCCGGTGCAGTTCACAGCCGACGACTGGATCAACGGCACTCAGGTCGCTGAACGCT
TCGGCAAGCGGTTGCGAAACTGGCTGCGCCTCGCTGAAACGAAGGCATACATGCGCGCCT
TGGCCGCTGCTATAAATAGCTCAGATGTGAGCGGTTTAATCCGCCGTCAGCCAGGCCGTG
>3246606..3246724_NC_011000
CACGACCTTCGAAGCCATGCGATTCCTTGACGTGTTCGACCCGCGGCCGCGCATCCACGA
ACTGCGCCATCGGCACGGCCATCACATCACGACCGCCATGCGCGCCGAACAGACCGAAT

>3246856..3246992_NC_011000
GGCGTGGAAACCCGTGGACCATAGGCGGACGAACCGCCGCACACGCGGCTATTTTTTCGT
TTGCGCGCTTCTGCACGCCAGTTTTCTATGGCGGGCCGGGCGAGGGCCAGCTTCGGCTGG
GCCGGTTTCCTATGGGC
>3247147..3247344_NC_011000
CTGTCCTCGAGATCGTGTCGGCAACTCTGCGCGCCGCCGCGCTCGCGCCGACCGTCGATG
CTGCGCTCGACACCTGCGGCGACGCCCTCGCCCGCGTGGCTGACCTTGCTCGTGCGGAGG
TGCGCCATGGCTGACAAGATCACGACGCTCGATGACGCCCTGAGCGACAGTATCGATAGC
GCCAGCTGTTCCAAGGAA
>3247579..3247713_NC_011000
AATCACCATCGTTGAGAGGCTTCGCTCGGACACTCGGTGGAACGTCACGTTGCTCGGCGA
TCCCGCCTTTGGTTTCGCGCGCAACGGAGCGCCCGTCATCACCAACAATTTCGCGTTCCG
CGACTCGTCGCTGGT
>3247833..3248013_NC_011000
GATTCGCTGCTGGCCGACCCGCTACCGGTCGCCTCCAAGTTGGCAGCCGAGACGGGCGCA
TTGTTTCCCTTTCGTTGCGGGACACGCTGGCGTGCGCCGACCGCGTCGACACATTGGGAT
TGGTCCGTATCGTGTCAGCGCTGGAACACAAAAGCGCTAACGGCTCCCTTCTCGTCGACG
A
>3248162..3248299_NC_011000
AAGCCGCACTCGCCTGTATCCCGCCGGATGTGTCGCGCGAGACGTGGTATCGCATCGCCG
CCGCTCTCAAACATGAAATGGGCGATGCTGGTTTCGATCTGTTCGACGGATGGAGCCGCG
GCCATGACGGCTACGCAG
>3248406..3248585_NC_011000
CCGACCGACGCCGGCCGATCCTGCCGAGCTGGAGCGGCGGCGCGCAGCACGTCAGGCCCG
CGATGACGCCGAAACGCAAAAACGCGAGCTGGCCCGGAACAACGCGGCCGCGCTCTCTCT
GACGCTCTGGATGAAATCGGAAGCCGCACGCGACGACCATCCCTATCTGATGCGTAAAGG
>3248596..3248793_NC_011000
CCGCCCACCCTGCACGAAATCGACGCCGACAAGCTGGCGCGGTTGCTGAACTACACGCCC
AAGGCGAACGGCGAGCCGCTTAGCGGCCGGGTGCTGCTGGCTCGCATCCACATCGGCTCG
ATGCTCACCAGTCTCGAGCTAATCGATGAGGCAGGGCGGAAATCGGCATTGGCTGGCGGC
GCAAAATCGGGCGGTTGG
>3249281..3249477_NC_011000
GGATCAAGCCGGGCGTCTACTGGCACGGTGTGAAGCCGGGCCGGGGTGACGCGCCGCCGG
AGCTGATCGACAAGTGGATCTGTTCGCCGATGAAGGTCCACGCCGTGACACGCAACCGGG
AGGATGCCGAGTATGGCCGCCTGCTGGAAATCCTGTCCCCTGCCGGCCACTGGAAGAAAT
GGGCGATGCCCATGTCG
>3250030..3250169_NC_011000
TACGAATCCGCGTATGCGCTCGTCAACGGCATGGGAAAAACGCGGGCCAACCGGCACGGC
GAGGCACGGCAGCCGGCCCGCTGGCGCGTGTTCCTGCTGTCCACGGGCGAACTGACCATC
GCGGCCCGCATGAGTGCCGG
>3250684..3250875_NC_011000
GACCGGGCTGGTTACTGGAAGCAGGACGGCGAGCGGCGTCTCTATCTGTTCACGTCAGGT
GGATTACGCGAGGCAACCAAGGGCCACGACTTCAGCCGGGCGCTGGGCGCATTGGACCAA
GCAAAGGCGATCGCAGCACGAGACGCGGGAGACAAGGGGAAGCGCTCCAAAAAAACGCGC
ATCCCGGGCGGC
>3251157..3251310_NC_011000
AGACATGCGTGCTGATGAGAATCCGCGCCGCCCGGTCGGCTGCGAGTACGGGGCCGAGCT
GATGCTTGCATGGGGGCGGCGCGTATCCGCCGCCGAGGTTCGCAACATGCGCGGCGAGCT
GTTCGACCTGATCCATGAATTGGCCGAGGTCGAG
>3251389..3251585_NC_011000
CATCACTTCCGCCAGCGCGTGGCCGACACTCGAGCGGCAACCGCGGCGCGCAATGTGGCG
GAAATCGTTGAATCGGGGCTGCTGCTCACGAAAGCGCCCCGGCACCCCTGAAAACTGGAG
GAATCATGGCGAAAGAGAAAGACGATGACTTTTCTTGCGAAATGGCGACGAACGGAAAGA
ACTCGTTCTGGCGTGTC

>3251695..3251846_NC_011000
TCGTTCCCGCAGTCGGCGCAACACATTGAGAAATCGGAGATTGAAATGTCAGTGGAACAG
TACCTGCGGAAAGCGTCGCTTGTGATCGGCGGCGACGACGGGAACGGCCTGGACTTCTCA
GATATGCGATTCACGTTCGAGATCCGGCGGGG
>3252472..3252600_NC_011000
TACGAGTTTGGCCTAGGCATCAAGGATCAGCCACAGAATCGATTCGATGCGCTGCATAAC
AAACTAAACAATCCGACCGGAGAATACGGCTTCTATCGCGTCATGATTGCAGAACACCGG
GGTGACACC
>3253014..3253166_NC_011000
GTTTTGAATATCGATCACGTCACCGGCTTTCGCTTGACCAGTCGGCGCCACAAACACGTC
ATTGGACGATTGCTTTGGCGGTGCGGCAATAGCCAAGGTCGTTGCAAACAGTGCGGCGAC
GGCAACAATTTTTCTCATCATTTAGGTGCTCAA
>3253231..3253381_NC_011000
GCGTGAAACTTGGTCTGCGCTGCCTGGGGTAATCGAATCCTTCAACCCGGATGCGATGAC
GTGTACGGTTCAGCCGGCAATCAAATTGCGGTCGCGTGACGAGCAAGGTCAACTCAGCAG
CCTTATGCTTCCGCTGCTGGTCGATTGCCCC
>3253510..3253673_NC_011000
CTCGACGCGCGTTCACGATCTGTCTGACGGCATGGTGTTCGTCGGCTTCCGATCCCGGCC
GCGAGTGCCGGGTGCAATCAGCACGACGACCACTCAGCTCCGCACCGACGACGGTTCGAT
GCACGTCGAGCTCGACCCGGCCGGCGGCGTGGTGTCGCTGAAGG
>3253802..3253937_NC_011000
GCGAACGGCAAGCGCGTCGACGACACGCACACGCACCCCGACGCCCAGGGCGGCAACACC
GGCCCGGTCAACTAATGGAGGCATGATGATCGACGGACTGATTAGCGGGAAACTCTACGG
CACGGCAGCCGAGCGC
>3254478..3254652_NC_011000
TCGCCCGCAAGGGGGTAGACATGGCCTGCATTGAACCCGAGGCAGTTGAGCCGCAGCGCC
TCGTCCCGCTCGACACGTGGATTTACGAGAACCGCTACATGCAACGGAAGCGGGAGCCGT
GACGAGCAAGGGCAATGCATGGCGACTGATCGAACCGGGTTGCACGGAAGCATTG
>3255162..3255352_NC_011000
CGCAAGCCTCCGGCCAGGCGTTTCTTGAGCCCACAAAAAGGCGGCACAACCGCCAGATGT
TTGAAAAAGAATCAACGGAAATCAATGCCGGCGGACGGTCATGCGCATACGCAAATCGAC
ATCGCCACCGCCATGCTGACCGGAGCGTCTTGCTCGGATTTCAAATCTGATTTTCCCGAA
GCGCTCGAAGC
>3255380..3255520_NC_011000
TGACACTGGCGTGCTGCTGCTCGACACCTCGCGCGATGATTTCTGACGAAGCGCGTGCCT
CTCCTTCGATCGGCACGGCCAGAGGAAACGGAAAATCTTTCGTGGGCCTCTCCCTGCGAA
GCGCGGCCCTCGAGTTGCGAC
>3255748..3255886_NC_011000
CGACGCATCATTCTTCGAGCGGCTTACCGCGCCCATGCTTCGTCACCTCCGGTATATCGC
CAAGCTGACCCAAGGCGAACACACTGTCGACGACCTGAAAAGCGAAGCCTGGATCATCGC
TGAGGAAATCAAGGCTGAA
>3256110..3256294_NC_011000
AAATCAAGGAAACCGCGCAAATCGTCGCTGCGCAGTTCTCTGAAGCGGTCGCGTACTTCC
ACATGTTCGACCAGTTCGACGGCGATAGCGCCGCGCTGGCGAACCACTTTGCGGTGAAGG
GACGCACGCTACGCACACGTCTCCGGCGGGCAGAAAGGACCGCGCGCAATCAGCCGTCAG
TTTTC
>3256328..3256523_NC_011000
GCGCCCCGCGGTACATGGGGCCGACAGCCTGCCCTCAACCAGTTCCGCCGTATCTGCGGA
TCCGCTCGCCCTGCGCAGTTGCATCTCTTTCTGCGATTCGGGACGGTGTTTGGGCCGACG
ACACGCCGCTATGCGGGCCGCGCATCAACCGTGGCGATCGAGGGGTAGCATCGAGAGTCT
GGGGGTATTTCTGGGG
>3289490..3289685_NC_011000
TTCATTTCGCCATGCGTCGCATGCGCACGAATCGTGCCGTACGCAGACAAATCACCTGGA
AAATCTGCTTAAGGGTGCATAACATTACTTAACGATTCTGACCCGTCTATCGCGTTACAT

TGCGTGGGCGGCGCGGCGGTCGCAGCCGCTTTCAACCAGCTTGCGGCGCGCGGCGGCCGG
TACGCGGTCATTTCGA
>3297325..3297457_NC_011000
ACTCGATAAGATCAAAAAAGCTATTTTTATAGCCGTTACCCGAAATCCCGAAACATTGCG
CCAGCACCCGCACCAGCACGCCGCCCGAGGACGGAGTGCTGACTTTATAAATTGGACAAT
CCGGGTGCGTCAA
>3309297..3309433_NC_011000
TCGCCCGGCCGGACGAACGCTGCACGCCATCGCTTTCAATCTTCCAACAGGCGCGTCCAC
CGCGCACGCTGTCACCTGGCGCGTCACATCGCGATGTGCAAGGTACGACGCGGCACCCGG
CCGTCATCATCAGCAAT
>3338874..3339031_NC_011000
GCCAACGAGCGTTGCTCGAGCATGAGCCGCCAGCCGGGCATTTGATAATGCCAGCCTTGA
TCAGAGTGCAACATGGGTCTTTCGTCGGGTTGCAGGCGACGCAAAGCCTTGCGCAACATG
TCAGTAACCATCTGGAAGGCTGGCCGAGTATGCGTCTG
>3368069..3368207_NC_011000
GACTTCCTGCAAGCAGCGCGTCGTATGCCGCTCGTTTGTCGAGACCTCGGCGTAATGCGT
CAATGCCCCGCTCCATTTGCGCGAGTGCCGCCAGGAAATCGTGCGTATACGCGGCCTCCA
GCTCTGCCTGATTTGGGCT
>3368914..3369048_NC_011000
CGAAGACTGAGGCGCTTAGAATCCAGCTTCACTGGCGCAAACACTTTCAGGCGATGGACG
CGGAAAAACTCAGGAACGAAAGCCCGCTGATTGCTGGCGACGGCATGATTTCAATCACAG
CGGCGGCGGAAGCTA
>3375805..3375977_NC_011000
GCACTCACGGTCCGCACCAAGAAAAAAGGACACGCCCTCCGTCACCGGGAAAGCGTGTCC
TAAAACCAGTGCCAGGGAGTGCGCCGTCAGCGCCGGCACCGGTCAACCTTGCGTCGCGCG
AGCCGATCGCTCAGCTACCCTGCGCCACTTGCCAGGCCTCGTACAGATCGGGC
>3410899..3411041_NC_011000
AACGCTGCCGTTGAAACGGACGAGATAGGCGAATAACCATGCGGCGACAACCGCCGTCAG
GTCGAACAGGAAAGCACTCAGGGACAGCCAAGATGCTCTGGATTGCAACATCGGCGGGAA
ACCTCAAGAATTGTTTTCGGCAG
>3421412..3421570_NC_011000
TGCGCCGAGCCTATCCGCCAGCGCGGCCTCGAGCGCCGCATGCTTCTGCCCACCATTCGT
CAACCACGCGCTACCCCATATGTCCGCCAACTGCGCCGAATATTCGTCGAGCGGCGGCAA
CATTGGCTTGGTCACATAGATCGGCTTGTCGAATACGTC
>3422012..3422155_NC_011000
TCCGCAGCGTCACTGTGTCGCCGTAAGTTACCTGTCGCGTCACCCGATCCAGTTCGTCGA
TCAGTTGCACGTTGACGAACGTCGCGCGGCCATTCCGGATTCGCTGAAATCCGGCACGCT
TCTGGAATTCCTCTGCACCAGTAC
>3423025..3423160_NC_011000
CAAAAGCATCTTCAACACGCCGAGCCAGTTCGGCGCAGTGCCCCAGATCACGACATTGCG
GCTTTGCTCGATGACGAACGCGACGGGATTCAACGAAACGATATGCTGATAGTGGGCTGG
CAGCGAACTGACCGGG
>3438425..3438560_NC_011000
CCGGCTCATCGCCCGGCAGAGAATTGTTAGCACTCTCGTGCAGCGAGTGCTAATTATATG
GACGGGTTTTGACAAATTCAAGAAGGGGAACCGGAGGAATTTTCGGGGCGGCGGCGCAGC
GGGGAAAACAGACAAA
>3445314..3445455_NC_011000
CCAATTCCAAAAACGGGATTGATTCGTCTGCCTAATGACCGGTAATAACGTTTTTTAATC
GATTTCAGGCGAAATTCGTGAATTGCGGAGCCCGAATCGTCTGATGAAAAAGGTCGCCGG
AACGCGGATGTAACGTAACGTT
>3456900..3457078_NC_011000
TACCCCACCCATGTTTTTGCAACCGCACAATGGAATCAATTCCATTGCAAACAACCATTC
AACGGTTCGCGGCAATATGAATTCATTTCACGAATTCGCACGCGCACGAAATCGATGACG
AAGAATCACCCCGTGAAGCCCGCCCACCGGGGCGCCTTCACTTCGAACGGACGAACGAT

>3468838..3469006_NC_011000
CGCTCGACCGACGTTTTTCTGACAGTTTGCACGCTATCACCGTCAAATTCCGCAACCCAT
AGCCCATCACGCATGCGGCCCATATAACTGCCGCGATATGCGCACCGCGCCGGCCGACAG
CCTTGTCGCATGGCGCGCCCGGGCCCGCCCTGGTCGCCCGCGACCGTTC
>3474261..3474410_NC_011000
GTACCCGCGATTTCGATTTTACGCAAGCCGCCCGTTTTCTCAGGTTGACGGATATCGGTT
CCGCATTTAACAATTGTCTGCGGAATCGAAAAATCTCATTTTCGATACGAGAGCGAACTG
GGGCCGCGCGATGCGTTTTCGACGTTGCGT
>3492918..3493072_NC_011000
GGTCGAGCGCCCCGTTCCATCGCAATCGCCCTCCGCCACGGTGATGCGACGACTGTAGCG
GCGCGTGCAAACGTGCGCCATTTGCCTAACGGCCGAGGCGACATGCGATCGCGACCGACA
CAATCGGCGCGAATCGAACCAATGGAGCAGACGAT
>3496591..3496745_NC_011000
GCCTATGAACTTGCCGTTCTGATTCGATACTAAGCAAGGTTCGGCGTCAGGCATGGCCGG
CGCGCCGCGCCGCACGCACGCCGGACACCCTTGCCGGGCACCCCCGAAGGGACGTTACTC
AAACCTCTATGAACACTTCTCAAGACCTGTCGATC
>3499091..3499251_NC_011000
TGCTGATCACCGGTGTCGGGTCGACCCAGTTCCCCATTGCCACCGCCAATTTCACGAACG
AGGCGAACCTGCCGCAGCAGGTCACGTCGATCGTCCGCGCCGACCTCGCCCGCAGCGGTA
AGTTCACCAACATCGACGCAGGCAGCACGCCCGTGCCCGAG
>3502593..3502720_NC_011000
TCGCGACCGATGCCGTCACGCGATCTCTCGTGATTCGCATTCGTGGCGCACTGACAGACG
ACGAAATGAAATCGGTGGACGGTGCATTGGAACAATTCTCCCAGAAATGGGCGCAGACCG
GCGCAATC
>3503722..3503868_NC_011000
ACGTTGATAGGGAAACCGGCGGGATCGTGAATAAGGCAATCAAGCGGGCGAAATTCCTTG
AGTTCTTCGCGAACCGCGCGGCCTGTCTTATCGGGATGGAAGCGTGTGGCGGAGCCCACC
ATTGGGCGCGGCAATTGACGCAGATGG
>3504118..3504294_NC_011000
AGGGGCGGGCGAAGCTGGACAAGGAAATACCGGTCGTGCTTGGCCGAATCGCGGAGCGCT
TGCCAGCGGCGCTAATCGATACGTTGCGCGAGCAATGGAACGGGTTGGCAAAGCTCGACG
AGCAGATCGCGGAAATCGAACGCCGGATGCGCGAATGGAAGAAGGAAGACAAGGCGG
>3504662..3504804_NC_011000
TTACGTGAGCGTGAAGCCGGCCTGATCGGTGAATGCCACGTTGAAGATCAACGTTTAACA
CTGGGTGAACGTCGAAAGGTTGCGCAGGAATCGAGTGTGATGACAAGCCAGGTAGGACCG
GGACTCGCTAAACCTGAATCGTG
>3505791..3505948_NC_011000
CCGCGCACCACACCGCTTCCAACATGCGCCGAACATGTGCAAACAAGCCCGAGCCAGACT
AAGCGTTGAACTCGGCCCGCTAAAGCGGGAGTGGGACGCATATTGTGCACAGCGTGGCCT
GACACCCAGTGAGGGTTTGCGGCAGTTCGTTGCGAAGG
>3506915..3507058_NC_011000
AATCCACGCAAACGAGATTTGCATCAATGGCGAGAAGCCTTCGCCCAACAACTTCGCGCA
CGCGGCGTCGACGCAAATGCAACGCCCAGAATGGCACGCGGCCAAACGCGGCGTCTTTCG
AAGCAGGCAGCTATTTGGGGGATG
>3522066..3522224_NC_011000
CCCTAACTCCTCCGCCGGCAAATTTCGGCATTCGGGAGCAACTGCGCCGTAATTCCGCCA
GATGAAATAATCGCTGTGTTGCGGATTTACGTCCAGTTTATGAAAAACATGAGGCAGCGC
GGGAACGTGATCGCCGTAGAAACAAACGACCGTTTCACG
>3524627..3524822_NC_011000
TAAAAATCATGCGCTGATCGATGCCGAGTTCCCGAGCCAGTGTCGCTGCGAATTGCTTGT
AGGCGATCAGGCCCGTGTCCAGCGGATGGTTCTTGATGACAAGCCGTGCACGAGACGGGG
CGTGGCGAGCAAATGAACGCAATACGGCCGAAATGGCAGCGCAAATGCCGTCGAATGGAG
AATGCGCGACGATCTG
>3526596..3526790_NC_011000

AATAATTCCAAGCCGCGCCCCCACCGGCGTAGGCCAGGTCGCCGCCACAAAAATTGACGC
GGGTCACATGGTGGCCGCGGCTGCCAAGAACATGCGCGAGCTGGCTGAAAAAAGGGGACG
CCGTGCCTTGCAAGGCAAGGAAAGAACGACTCATAGATCAACTATGCAGTAATGGAAAAA
CCGGCAAACATCGCC
>3528519..3528675_NC_011000
GGCGGCGGCTTGCTCTTTCCCTATATATATTGAATGTGACTTCCGAGATCAGGTCATCGC
TAAACCTCATCAGCCCCATTAATTTGTAAGCCGCGTAATCGTAAAACTCTTCGATCGGGT
GAAATTTCGCCTCACCATACGGGGCATCAAAAGATGT
>3529532..3529665_NC_011000
TATTGACACATCAGATCGAGATTTCTCGACAAATAAGAAAGAAAACAAAACAAACTCCGG
CGCCGACGTTGCAATTAAGCCACGTCGATACGAAATGCACGCCGCACTCCTTGACAGACG
CAAATTGACCGGAT
>3529702..3529858_NC_011000
GTATAGGCATGAATCGTCAGACTCGATTCAATCTCCGACGACCTGGGATGCTAATACCGA
ACGCGGGATTCGACCGTCAAGACCGGCCGAAAACGGATCCTGGTCAGGCCTCGGGCGCAA
ATTTCCGGACAGCGGGCACGACCGTTGCCATGTCCGA
>3530177..3530365_NC_011000
ATACTGCGCCAATCTTGCCGGTAGTTCGCGAAACTGAATCGAATCAAATCGATCATGACG
CGTCGTGAGAATTGGTGATCCGTCGAAATCCCACCGAATCCCGCAATATACAAGGTCGCA
ACTTGTCGCCGCTCACGCAGCTCAGAATACGCAAACAAGGGAGGGAAATCAGGAGTAGAC
CATCAAACC
>3531599..3531767_NC_011000
CCACCCGCAGCGAACAGCGCCTCGCACTCCTGCTCGGTCCAACGCTTTCCGGTCGTGTTG
AATTGATTCGTCTTATTTAACAATTCGAATGATCTGACGTATTTTGCATCATCGCGGCGA
TTAATTACTCCTGGCTTTACCCGGATCTCCAGGGATTGCAAATAGGCCT
>3532227..3532381_NC_011000
TGACAATCTCATCCTCCTCGGCCAAAACCCCTTTCCACAAGGTATTATCAAGATCGGTGA
CAATCAGTTTTATCGGATTTTTTGATTTCAAAATCCTTAGCGCCTGATCAACGCGCGAAA
TCAATGCCAGATAAAACTCCGAGCTATCGTTGGCA
>3533711..3533851_NC_011000
CGCCATCGGATTCTAACCTTGAAGAGAAATTACGATAAAGATTGGCATTTTTCCCAATTC
GAAGCCGAACTGGAATTCGCGATATTACGATGTGACACCCAGCGACATGCTCTCGAAACC
ACCGATCTTGACTAAGACGTT
>3533895..3534068_NC_011000
CTCGAGGAGATAGCCACGCGCACGGTTACGGCTCACCGAACCCTGACGAGTATGGCGCTA
GCTGTCGGCCGCAACGGGCGCTCATCTCCCGGATCAGAGATTGCACGTCGCGATACTCAC
GGTCTGGTCGACGGTACAACCGCTTCAGCGACATCGTTCTTACGATGCTCGCCT
>3539187..3539319_NC_011000
CGATGCAGATCCTCAAAACAGAAGCAACATCGACAGATATTTATTTGTAGTCGGCGCACC
CGCTGATTTTTCGCTTGACGAGTCATTCTGCGACAAGATCATTCGTGCAGAAAGCGTAAC
CCCACCTCAATTT
>3545074..3545222_NC_011000
CCCAATAGACGGCGTGGATAAAATCAGCGTCGTGAAGCTTCAACAGATTTCGCAAATCAC
TTTCGTCGTATGCCGACTTCACCACCGGGATAAGCTCCCGCCTACCGGACAGGTACTCGA
TTTCCGTGAAGAAATGTGGATCCAGGTTA
>3545491..3545687_NC_011000
TGGCGCTCGCGTAAAAACTAAGAAGCGCCTCACCCTTCAATCGATTCGAATGAAAGACGA
AATCGAAGAATTGGTGAATGTAGTAGAAATTCGAAAATTTCTCGAACAACCCCATCCGCA
GGAGTTCATCCTTGATTTTTCCGTAATTTACAAAATGCGCTTTTTCCTTTCCGTTAGCCG
CGCTGGACGATCCAGCA
>3554768..3554956_NC_011000
CCGAAATGACAAATCACTTGGGCCAGCATGAAAGCAAGCGGCCCGCACATGGAAAATATT
GTTTCGTTGAAAGCTTGTGTATTTGCTTCAACGTCACACTGCGGTCACGAGAATGAAGGA
TTTGAAGGAAAAATTACGTAAAAATACGATGCCCACAAAGCGTTCGCACAGCCCGGCGTG

GCAGAATGC
>3560578..3560767_NC_011000
GAAATTTCAGGGGAAGCTCAGTCCGATCATGCTCGCTTCACTATCGAGTCCGGTGTACAG
GTATTTTTCGCCGACCCACATAGTCCATGGCAACGTGGCACGAATGAGAATACGAACGGC
CTATTGCGACAATACTTTCCGAAGGGTACAGACCTATCCCGATGGAGTGCACGAGAGATC
CAAGCCGTTG
>3561824..3561963_NC_011000
AGTCCCGCACGGCCATGCGCGCGATACGCCGCAACCCAACGCTGCAACATTCCGGAATGG
ATGCCGTGACGGGCTGCCAGTGCGTCAAAACCACTATCGCCACGCAAATATGCGCTGACC
ACCTGTTGCTTGAATTGCTC
>3600393..3600538_NC_011000
TCTCCGTTAAACATGCACAAACAATACATGTTTAAAAACGGACGCCCCCCGGGCAAATTG
TCGCAGTGCGCAAAAACGCGACGGCGGGCTAGCGGACGCGCCCGCGTTCCCACAGCACGT
CGGCGCCGCCGCCGGCGCGGTTCAGC
>3639090..3639283_NC_011000
GGCGTCGGCCAGCGCCTTGAAAACGAGGTCCATGCCGCTCCCGGGTCGATGACCTCATTA
TGCAACCAAATGGTTGCATGTCAAGGGTGACAACTCTACTGACAATACGGTGTCAGCAGC
CCGCGATCACACTGCAATCCCGGCGCAATGCAGGCTGGTTTCCACACCGAATCCGCATAA
CAACGCAAACAACG
>3666829..3666990_NC_011000
CCGGCAGGGAAATCATCGGCCGGCGCCCCGCTCGCGCGGCGCCGGCATGGCGCGTCGTGC
AGCGCATCAGCCGATCGTTGCACCGCGGGCTACAATGCCGGATCGAATCCGAATCAGGCA
ACGCTGGAGTGCAAGCATGGAATGCAAAGTAAGCTGGATGGG
>3671381..3671556_NC_011000
TTTATCGATAATCGACGCCCATTCGAATCAGTAGACTCCGCCCATCCTCGGTGTAACTAA
ACACGATTGGCGCGACACGCAGCCATAGCCGCCGCACGCCCCTTCGGATTATTTCCCCAC
CGCGGAATGCGAATTCGCCGAATGGCGGCGAATGCCACGTCAAAAGAAGAAAGAGG
>3718958..3719122_NC_011000
CTGTTAAATCAAGTAGTTGACGACCAACAAAGGCGTTAGCGTAGCGTTATTGGCCGCGCG
AATCAATGTCCGTTTGGGAATGGAAAGCCTCGCGATGCCGCGAATTCCGCATTTTTTACC
GAAATCGCACTACATTTATCTTACATAAGATAGAATGCGTGCGGC
>3737727..3737897_NC_011000
CGCGAGCCCGTCAGCTTTGAAAAATTTTGAAATATTTGGCGTGCAAGCGGTGCAGCGCAC
ACCTTTCAGGAATGCTTCGATTCGCGAAACGTTGCATTTCACGGCGCTTTCAGCGACTTT
CGCGTTTACCCGCGGCGGCTAAAAAAGCACGGATTTCGTGCGACGCCGGTC
>3769617..3769782_NC_011000
GGTCGAGCGTGATTTCAGCGTCAATGTCAGCCGGTATCTGTTCTACTTTGAGTTTTTCGC
CAACGGCAACCTTGTACTGCTTGCCGCCGGTTTTTATGACCGCGTACATTGAGAACCTCA
CTCTGGATCCAATTTTTCCGCACACCACGCGCGGAAAACACGTGAT
>3809894..3810058_NC_011000
TGACCTTGATTGCCTTGCCGTTCTCGACGGTCACGCGCATTGCACACGTGTCCGGACAAT
CATGCGGGCAAACGGCCCGGGCTACCTGGGTGACGGCGTTCATCGTTTTGAATGCTGAGA
TGGGGGTACGCAAATTTTACGTGGACCACGCCGAATAACGCGTTC
>3819773..3819925_NC_011000
GGCCCAGCAGCGTCGCTGTGTTGAGCGTGACCGAACACTGGTCGACCGTGCCGAAATTCA
ACGTGCTCATCGTGTTCTGCGAATATCACCGGGCGAAAACTGCCCGCGCTGCCGGATGCG
CGCGGCAATGCCCGCGGATATCGTGATCCCGTT
>3820929..3821084_NC_011000
GTGAGCCCGTTCGCCGAGCCTTTGCATCCTGGATGCTGGCGATTCTACAGACATGACGGC
TCACGTATCGCTCGAGCACTTCACAGCGTTCCGCTTCGTTCATCCAGCCGCTGCACGCAG
TGCATGTCCTGAGCGGAGCACAGTCGCCGGACTGTG
>3868390..3868532_NC_011000
CGCGCGCCGATGCATTTTTTCTCGATGGCGCCGCGGCGTCGAGCGACGGAATTCGCGCGC
TCGCAAAAATGGCCGGCGAGCACGCAACGTTTGCGACGCATGCAACGTCCGATGCCGTGA

AACACGCGCTCGGCGAGACAGGT
>7278..7422_NC_011001
CCCCAACCCCTGAACACCGACCGGCAACCGGATCCGGTGTCATGTGAAACAAAGGCCGAA
ATTCGCGGTTCGCTGGAAAATGCTTGCAAGACAAGGACTTGGCTGACTTTCTTCGCTTTG
TTTCACGGAATTCCGAGCTTCCTTC
>7465..7617_NC_011001
TAAATTCACAACGGACCCTCGAAAAAGTATCAATAAATACGTATTTCGTTATGATTACAA
CATTTCATAGTCTGCGAGTGTGGTCATGACCCTGGACGAAATGCGTCAAGTCATTCGGGA
AGAACTGGAATCGCTGCGCGCAAGCGGTGCACG
>24329..24498_NC_011001
GTTGCCGTTGTGCGATGCCGCATCTGGAATTAATTGTTGCCGATTTCGAATAATTTGCGG
CGTAAATATTGAGATGAATGAGACAAACGGCCTGGACGGCGTGCCGGCCGGGCCATGGTC
GCACGCATCGATGACGCCATCGTGTGAAAACGGTGAGTCTTTACGGGAGA
>27545..27716_NC_011001
TTCATGAATAAGTTAAACAGGGAAGATTATAAAAACGCCGAATAAAACTATTTTAAAAAT
CGAAACGGTTTTATTTTATTTTCCATCGTGCAAACAAGAATTCCGCGTTGGCAAAGGTCA
TTAATCCGACTCCGTCATTCAATGTCATCGTGTCGGCCATGGCGCATGGATC
>27780..27909_NC_011001
AAGGTAAAGCATTGCAATATAAAATCGTTATCCAGGTATGCCGGTTGCGGAGCACGATTT
TATTTGGCAGGATTTTCATCGGGATGCCAGACTTTAAAACAACAAAGCGATCGACCATTC
AATTTGCCGA
>71907..72071_NC_011001
CGATCTGGTCTACATTCCGAGCGGCAAGGAGCCGAGAATTCATGCCGGATCGCGTAGCGG
AGTTCCGTTTCCGGCTCGCGCCACGAATGAAAAAGCCGCTGACTATCAATACACGCCACG
GTTGCTCACCGAGTCGAGAATGCAGCAGATTCGATTGGCTCGAGT
>72240..72397_NC_011001
GAAGGACATCATCAGAAAGACAGCCGAATCGTTCGGTGCTCACGGAATCGAAGGAGTAGA
CCTGGAAAAATGGTCACGGCCTTTTGCTGGTAAATTATTTCGTAATGCGAAAGAGTATTC
GGGCGAACTGGAGAAGTGGATTCGAAACGATATCGAGA
>82658..82804_NC_011001
GCAGCATTACTTCAAGCTGCTGGGCATCATGGCAATGACCGAATTGCTTGACCCGTCACA
GTACGAAAAGCTGGCGCGCGGGTGCCGGCGTGTCGGGTTAGGCCGTGACGCTGAGCTTGA
CTACTACGACGAGCACGTGACGATTGA
>90660..90811_NC_011001
CGTTCGGCATTGAATTGAGCAAATAGATTCGGAATCCGAATCTATGGAGCGGGAATTTCG
CTTGACGAAGCAGCACCCGACGTTTTTGGCGTCGCGCGTCACACGCGCGGGAACAGCCCG
CAAGTCATGCGTGGCGCACGGCGATGAAGACA
>158028..158164_NC_011001
CCTCCTGGATTTATTCGACACGGTGTCGGCTGTCGCAATTCGCGTGCCGCGACGCGGCGC
TGTTGAAAGCCGGCGCAAGCGCGTTGTCTGAATCGCCTGAATTCATGCGATGAATACCGG
CTGATTCAGAAAATAAA
>174315..174491_NC_011001
CACGAAAACTCGGTCGTGACAAAGTCGTTGAGGCAAAGGCTGTGGAGCATGCATGCCGGG
GACAAAGGCGCACAGGAAGACTTTTTCAAATCTTTTATGGCGTGGCATGACATCGTCGAG
CAGAACAAGGTGAACAAAACAAAGAACCTTCCACCGGAAAAGTCACTCGTAAAATTT
>176960..177093_NC_011001
TTGCCGAGCAGCTGGCGCCACACGATCGAGCTCCCGAGATTGCACGTCAAATGAGGCAGT
GTGCGACCAATCAAGGTTACGGAGATGCTGCCAATTATTTGGGAATTAATTTTAAAAACG
ACTCAAAATATTCG
>301854..301994_NC_011001
CGGCGTGAAAATCTGGCACCGTAAGCGCGATCGCGCACGTGCCGTTAAAGTAGTCACCGA
TGGGATGAGGCGCGAGCGCGTCAAGCATGTCGTATGCGCGCAGCGTCATGAGTGCCGGAA
TGCTCTGAGTGAGTTGTACTG
>309189..309383_NC_011001

ATCGAGCGCCGTTGCAAACGAGACGTCCCGAAGCTGCCCGATCGCGAATGGCGGGCGACG
CGATGCGAGCTCGTGATGCAAACGTCCCGCACAGGACAATGCCAGATCATGTTCGAAGCT
GGTCGAGCAAACGATGATCTGGTCTGGCCGGGGAAGTTGTGAGGCACCGTCATACAACAG
GCGTGCCGCGCATAG
>315865..315999_NC_011001
ACTGTACCGTGCATTCGGTGGCATTAACGCGATCGCGGGCGTTGACGTAGAAATTGGCGC
GAGAGACTTGCTTTGCGTCATCGGCCCGAACGGCGCTGGAAAAAGCACGCTGGTTGGGCT
GATGTCGGGCGCGAT
>432522..432657_NC_011001
CGCAAACTTCAAACACGCTAAGGGCCTTGACATGCCGTTTTTCTCCATTATTATTGCTCA
CCACGTCACTATTGTTTGTTGCGTCATGTTTGGAGATCGATTCCGGCCCTGGTCGGACAG
CGAAAAGCGGCAGCAA
>506891..507029_NC_011001
TGGATGACGCAGCGCACGGCGATGCTGCGATGGAAGCGTGATTGAAGCGAATCGAATCAT
CGAATGACGCGCGGCGCACGCCGCGGCATTGCATTAATCGATAGGTGATAATAAAAATGA
AAGAATGATCCTGTCGATT
>690393..690538_NC_011001
GCGCGGGCTGTGGACGCGATCGAACAAGGTGTCGCCATGCTCCGCATGTTTCCGTTCACA
TGCCGCAAGGTCGATCACGGAAATCCATTTTTGCGCGAGCTGATCGTGTCCTTCGGCACG
TCGGGCTATGTCCTGCTGTTCGAGAT
>796805..796970_NC_011001
GGTGTGTCCCTGAAGTTCAAGATGATCGTTTCGTGCATCGTTCGATGCCGGCGCCGAACT
TGTGTCGCGACGGGCCGGAATATCGCGATTGCGGTGGCGATTCTGGAAAGCGTTCGGATC
GTGACGGTGAGGGACAAATGAGAAAAATGTCATCGCAATGTCATGC
>853272..853410_NC_011001
CTGCGAGGCAGCCTCACGCCCATCTATGCGATTTGTGACGCCGAAGACGGAGTCGCAGCA
GATTTTGTCAGCTTTGCACCGCGTGCGTGATTCTCTGGTGCGGGATCGAGTGAAGACGGC
CAACCAATTACACGGGTTC
>865164..865315_NC_011001
CCAAAGTCTGCTTGTCTTCTCGAAACGGCGTGGCTGTGAACTGAATGCATGGTTTTCCTT
TAAAAGCTGTTTTCACGCTACGCCAAGTTTTTGCTGCAACATGATGTGCTTCATCGACGA
TTAAGTGGCTTACGAGGTTTACTAGCTGAGTC
>904358..904493_NC_011001
GGAGTATCTGAATCCGGCCCTCAATATGACCGACCCCGAGCCTTGCAGCGCATGCAAACC
TCATGGCGCCGGTCTCGCAAACGTTATCGTCTCTCGAATCCGGAAATGATCTGATTTAAC
TGTTTCCGTCTCTTTC
>1018568..1018719_NC_011001
CGTGCGAACGAATTATTCGACGGCGCGTACTTCGGTGCAGCGGAAGCGAAAAAACGTGGA
GCGGAATCGTTTGCGTGGCAGGCAATCCGTCGGATTGCTTGCATCGGACGGTATCAAGAA
CGAAAGCGCGATGTCAAACTTTTTTTGATTAT
>1141060..1141230_NC_011001
ATCGTGACGATCTGCTTTCGTGCGCCCGTGGCAATTGCATATACCTCGGCATCCCATTGG
CGGCCTAACTCGTCAGCCTCGCGCGCAGAAAGTCCCGCAGGGATGAGCTTTGTTTTTCGG
ATGCGGTCACCCTCGATGACGCGCTCGAACGTCCACCGGTAGCGACGTTTG
>1142128..1142243_NC_011001
TGATACGTTGCCGACGATGAAGCGCCGGCCGCCATTGACTTTCGTGAAGCCGATCCATTT
CCGCACCGTTGCGTCGTCCAGCGTGGCATCTGCCGTAGATGGGTTGATGCCGAAAT
>1152472..1152608_NC_011001
CAGGGCTGTGCGTCACTCCTGAGTCATCATCCACAGGGCGAATATGTATGATGCGAAAAC
ATCATGTTCGATTTCGCCCGATGCAACCGCTACGACAACTTCCTCAAGTTGCGGAGTTTT
AGGAATCGGGTAGCCTT
>1155880..1156030_NC_011001
CGCGAGGAACAAGCTCAGTTTTGTTTCGCCTGCGGGACAACTCTTGGCTACATGCGAAAG
GCGCTCAGCGTTGGCGCGATCTTTGGACCAGCTCTTTGCGTTTCCATCGAGCGGGAATCA

AAGGGGCTCGTTACGCGTCGCGATCTCCATC
>1173919..1174053_NC_011001
ATCATGGGTGACATGCTGCAGGGTTCGCAGTGGACGCCGCTTTCGCCCGGTGATCGTGAT
CGCCCGACGCACATCGAGGTGAAGCAGGACGTAAGGATCGAGGTGCGGGGCAGCGCCGAC
CCTGATGCGACGGCC
>1175452..1175603_NC_011001
GGCTGAACGGGCTGCGGCTCGCCGCCGCAGCCAGCCCGCTCGTATTTATCGCTGCGGCTT
TTGTGGCGTGGAGGGTTGCGGCGTAGGCGTCCCCTTGTTGAACCCATCGCGCGTTGGCGC
TCTATAAACAGGCTGTGGGTCCGGCACGACAG
>1180640..1180780_NC_011001
AGTCCGGCGACCTGATCTATTTCTTCGGGACGAACGCCGGCGTCGTGACGCTTCCCCAGG
GAGCATCGATTCATGCGCCGAGTGCGCGCTTCGTGATCTACAACGGTAGCCTCGCGAATC
TTACCGTGAGTGCTTTTTCAG
>1182271..1182421_NC_011001
ATAGCGTTACGATCAACTGCCCCGCGGTGAAATTGATTGGCGCGGGGGGCGGCGGATATA
ACGACAATGCTCCCACCATCAACGCTACGACCATCAATCGTTGGGTCGGCGCGCAAGGTG
GAACGGCTCTCAAGATCATGCCGGCGGCTGA
>1184741..1184878_NC_011001
GCCCGACCGCCTGACGCTCTGGGATTCCTGCAAGAAGGCGCTCGGCGTGACCTGACGCGC
CAGTTTTCGAATTTTCGTAATCCCGCCCGGCCGCGCGCCGGGCGTTTTCATTTCCGGACC
AGACATGACCCGATGCAG
>1869681..1869816_NC_011001
TGGTGTCTCTCGCATCAAAACGCATAAAGAAACCGTGCGTCCTACGAACGCAAGGTGCGC
ATCATACCGGGAACCGGGGCGGTTTTTTCATCGGATGAATGACTTGCGATGCATTTGGGA
GGAATCTGCCGGCACG
>1981923..1982067_NC_011001
TGCTGGTGCCAGTGCATCTCGCGAATGCCGCCGCGCGTCAGCCGCATGTTGACGCCGGAA
ATGGTTTCGGAAATCGCAAAATCGTCCTGCGTCACTTCGCGCGCCCATCCGCCATTTTGA
ATGCGTTTATGAGCATTATTGAAAG
>1992655..1992787_NC_011001
CTGGTACGCCCAAAGAGATATCGTTCCTGGCGAGGCGAAGTCGGCCGCGTTGCCCCGAAT
CTACTGCGGCGTCAGTTTGATGCTGAACGCCCGAATCAGAAGTGGGTGACCGACGTGACC
GAATTCAATGTAG
>2029783..2029917_NC_011001
GCAACGCCTTATTTTCGCTTTCCGACTGCGTCTGATTCGGCTCGCAATGGCTTTGAATCG
GCGTAACCGTGCGTATTTCGCACGGCACGCGCGCACTCGAACATCATTAAAATTAAAATG
GACAACCAACGAATT
>2223385..2223519_NC_011001
CCTGGGCCGCCTCGCGACCACCCCGTCACATTCGTCAACAATTCCCCCACAAACGCGAAC
AAGAACGCTTCTCATTTAACAGGAAATTACATAGAATGCCGCCTGAAAACAAATCGTAAT
AATTCCCGCCGGCAT
>2521772..2521943_NC_011001
AGGGAAAACACGTACGTTTTTGAGATGGGTGGGTGGCGATGGGGGCGCAACCGCGACGCG
GCGACAACGTTTGCTTCGGGTGCAATCGGGTTGCTATTTTGGCAAGAAGGTGGGAGCGGG
TTCCTTTCTGGGCTTGGCGTTCGGCGACGGGGATTTCCCGGTGATGCATTTT
>2720007..2720185_NC_011001
GTACCACCCCGATAACGATTCAGCCCGCTGCGTGATTTGCATTGGCCGAAAGATTGGTTC
ACGCCGCCGGCAGGTTTTCGTACGCTGAAAGTCCGAACTTTTTCCGCGCACGCCATGAAC
CGTACGATCCGCTACAAGGGCTATGAAGTCGCTCCCGCTGCCGCCCGGTTGCCGAACGG
>2722155..2722293_NC_011001
CGGTTCATTTATCACTGTTTGTCAATCTCTAAAATCCTCCTCCAATAAACTCCGCCCATC
AATTTTGAGATTCGCGCGAGCGCACGGAGTAACCCTCTAACAAAAGAGACCGGTGTGTGG
CGTGCACATTTGGAGGTAG
>2739339..2739504_NC_011001

CCGACAGTGCGACAGGTTACACAACGCGAATTACAAAATCCCAAAATCCGGATCGTTTAT
CTTTTTTGCCTAAACAGTTTTATCCGAATCGGCTTTACCCGAATTTGCTCGTCGCGCAAA
TTGGGGCGCCGCCAATTGATCAGATAAATACCGTGCCTCGAGGTGC
>2790921..2791083_NC_011001
CGCCCGGCACGGCATCGATGTGCCAAAGTGGAAGTGTCAGCCACCACTTAACCAGAGGGA
ACGCTCATCATGAAGGTTACGACCGTTGGTATCGATCTGGCAAAGAACGTCTTTCAAATC
CACGGCGTCGACGCGCACGGCAAAGCCGTCCTCAAGAAGCAGA
>2791288..2791448_NC_011001
TCGCTCGGCCCAACATGCGCTTCGTCCCCATCAAGACCGAATCGCAGCAGGCCGTCCTCG
CCATACATCGCGCACGTCAGAGTTTCGTCAAGGCGCGCACCGCGCAAGCTAATCAGATTC
GTGGGCTCCTCGCCGAATTCGGCCTGGTGATTCCCCAAGGT
>2912664..2912801_NC_011001
AATAATCCGGCGATAATCCCGACACGCGTTTATCAACTTGTCACAACGCGCGGCCCGATG
CGCCAATGCCAGCATCCGTTGCGCGCGGCGAAACGATTTGACTTATCCTGCTCCATTCCA
CCTGAATTCGAATTCCAT
>2986244..2986412_NC_011001
CGCCAACGAGCGTTGCTCGAGCATGAGCCGCCAGCCGGGCATTTGATAATGCCAGCCTTG
ATCAGAGTGCAACATGGGTCTTTCGTCGGGTTGCAGGCGACGCAAAGCCTTGCGCAACAT
GTCAGTAACCATCTGGAAGGCTGGCCGAGTATGCGTCTGGTACGCGATG
>3025995..3026135_NC_011001
TTTCTGTTCTCTGTCAAACGTAACCAATCGCAAATGACATCACGCGGTCGCAATCCGCGG
CCCTCGCGTGCGCCGGCCCGTGCGGCGGGATTTTTTTTCGGGTTGGCCCGCGACTATTGG
CGTGCGATAAAGGGTTTTCCC
>3032042..3032183_NC_011001
TAGTCGAACTATCAACGAATCTGAAATATTGCACAGGATTCGACAAAAATCTGAAACGGC
GCCATCGCGCCGGAATGATGGCTGGACGAGGGCCGTCGACGATGATCGACCAATGGTGTC
ACTATTGACACCACAGGCTACG
>3069020..3069154_NC_011001
AAGCGCTTCGCCCGATCGTGGTCTCCTCAGCCGGTCATTTCGGGATCAATAACCGCTCTT
CTTCTTGCACGCGAAGAGAAGATTTCGCACGGCGTCATCCGTCGATTGCTGGATCACGTC
GAAATCTCCGTTGCA
>3070082..3070253_NC_011001
TCGAAGCTCTTACCACGATTCAGGTTTAGCGAGTCCCGGTCCTACCTGGCTTGTCATCAC
ACTCGATTCCTGCGCAACCTTTCGACGTTCACCCAGTGTTAAACGTTGATCTTCAACGTG
GCATTCACCGATCAGGCCGGCTTCACGCTCACGTAACCCTTTTGGTACGGCC
>3090010..3090149_NC_011001
CTTGAGACAGAAACGGCTTCGAAATCCCCGCGGCGGTAGCGGTGTCGTCGAGCGTGCGCT
TCAGACGTTGGCGCAACGCGCGAATCTTGCTGCCGAGTGCGACAGCGTCTGCCGATCGCG
AGTTTTCAGTGGGGGGAACC
>3136490..3136633_NC_011001
TGTTCGCCTCCTGAATGGACCGGGCCTGCCACGCCGTCCGCGGCCGTTTCCCCGGTCACA
CGTGACACACGATACACCGAATGACACTTTGTGTCATGCGTTACAATCGGTTTTTCGAAG
GACACCTGGCGCCCCATGACGATT
>3183824..3183999_NC_011001
CTTGTTTTGCGCCACCAGATGCATGAACAGGCCCGGCTTGTTCCGCTTGTCGACCGTCAC
CACCCACTGCTTGATCTTGCTGGCCGTGGCGTTGGGCGCGCTGACGCTGATATCGACCGG
GTTGCGCAGAATGCCCGCCGCCATCGTGCGGATCTCGTCGGAGAACGTGGCGGAGA
>481..630_NC_011002
TTCCGGCGATCGCATTCGCTCGCGCTGTGGGAAAACACGGTGCGCTACGAAGGAATCGGC
ATCACCGCGAAGATTCCGTTGCCGAAATTCCGTGACCTCATCCTCGGCCAGGACAAGGCG
TCGCAGTCGTACAAGGAATACAAGCTGTTC
>713..844_NC_011002
GCCGCAGCGTGGAAGCCGTGCAGTTCAAGGTGACGCGCAAGCAGAGCACCGATACGGTCG
AGGACGGCGACGTCAAGAACGAAGCGCTGGTCGAGGAAGTCGCGAAGTTCGGCATTCCGC

GCTCGGAAGCGC
>2614..2749_NC_011002
GTCGAGCAGATGCTCCCCTATTTCTTCAGCCCTTAGAATTTTTAAAATTTCTTCACGCAT
TTTAGTTTTTCCATTGACGCTATTCAGCATCAATGCGGAAGAAAGTTGAGGATTTCGAAC
GGCGCGCATCGACTCG
>3108..3271_NC_011002
TACTGACGGCGATCTTGAAAGCCATTAAATTCCTCCACGGAGTTTTTCGTAATATATGCA
GTTATTAACGGAATGTCCAGAAAAGCGTTCGAAGCAGAGGATATTTTGTTGATCGTTGTG
GTCGTTACGTGACAAGTGGGCCGAATGGCTAAGGCGTTACTCGT
>3892..4087_NC_011002
TTGGCACGATCGGTTGCGCGGCCAAATTCCTGCCGGCGTGCGGACCCTCAAGGGCGGACG
CAATGCTGCCGAGCTCTGTCGCAAAGGCTGCATGGGGCTGTCATGATATGTCCCCATTTA
TAGGAGCGCGACCCTGCGTTTCCACGCGTGCCGGCCGATAGGGTAGGTTGCTTCGTGGCC
GCCAGCCGGTGCCTGC
>4196..4392_NC_011002
TTTCCCGAAAATAAGCACTCGCCCGGAAATCCTTGTCGGGCAATGGTTTGCGCCATATTC
TTCATGCTGTTTCACGCAAAACGGATGCTGCACCCTCGGGAAATAGGGGCGTATGAGTTT
TCAACGCCCGGGCATTGAAGATTGATACGGCTGCGGCTAGATTTGTTTTGATTGATATGT
GATTTGTTGTTTTTGAT
>4443..4616_NC_011002
TGCTCTACCATTGAATTCCCTTGAGCCTCCAATCGAGTACGGGGCGCACCGATATTTAGC
GCCAGTTGTCACGTGACAACGCACACGTCGGTCGCGCGATTCATGGAATGGTGGGATTCG
GAGTAGATAGAAGTTGACGCCGGGCATTCGTGGGCCGGTAATCGCAGCGCGACC
>10402..10561_NC_011002
ATGTCGATCTGCCGATGGTGAATCTCCTCGCGCGTGATGCCGTCGTGTGGGGTGTCGTCT
TGGGAAACCATGCATCCGCTCCATCTCGTGTAACCGGGGCTTTTCGCAGGCAACGGGATC
GCGCCCGGCTGACCGGGCGCGTGGGGAAAAGCAGGCGAGA
>11987..12129_NC_011002
TTTCGACCGTGCTGATGGAACGTAAATAATACGCATCATCGATTAATCCGGTGCGTCAGT
CAGATAATTCGTTTTACGAAATGATGGCGTGAAAAACGTCGATCGACGCGCGCGCGGTCC
CGTGTGCCCATCCCGACAAAACG
>35441..35613_NC_011002
TGATGCGCGCCTCGGCGCGCGAGCTGACGCCGCCGCATGCAATGAACTGAACCCCAAAAG
TTGGACGCCAACTTTTGGGTTTTTTTCATGGCGAGATATACCGAGCAGTTCAAGCTATCG
GTAGTGAAGCGGTGCATCGGGGGGGAAGCCATACGCGCGGTGGCGCGAACGCA
>35699..35834_NC_011002
TACGATGTGGCGTTCAGGCTCAAAGTTGTTCAGCACATGCGCGAGCATGGCGTGTCGAGC
AAGGAAGCGGCGGCGCGGTTCAATATTCGCAATCCGAGCGCGGTGCTCGAATGGGCTCGC
CGCTATGATGATGGTG
>35977..36148_NC_011002
CATGGAGAACGCGTATCTAAAAAAGCTGAGGGCCTTGGTTCAGGCGCAGGCTGTACCGCG
CAAAAAGCGCAAATAGTGCTTGAGCTAAGGCAACAATTCCCGCTGGCTGGCTTGCTCCGG
GTGGCCGGACTGGCTCGCAGCACGTTTTATTACCAGTGCAAGGCGCTGGCCG
>36502..36690_NC_011002
GCGTACCAGACGCATACTCGGCCAGCCTTCCAGATGGTTACTGACATGTTGCGCAAGGCT
TTGCGTCGCCTGCAACCCGACGAAAGACCCATGTTGCACTCTGATCAAGGCTGGCATTAT
CAAATGCCCGGCTGGCGGCTCATGCTCGAGCAACGCTCGTTGGCGCAAAGCATGTCGCGT
AAAGGCAAT
>36866..36999_NC_011002
GGACTCAGCCTCATCCGGCCTAGCCGGTAACCGTCCAACTTCGTGGGTTCAGTTCAGCAA
ATGCGGCGGCGTTTTTTCTTGCCGGTACGGATACGGTTCGAGCAAACCGTCAGTGCATCG
CCTGCGGTGCGCGT
>37598..37783_NC_011002
CGGCCGATTCAAGGCGCATCACCGAGCGCCCGGCGCGCGTGAAGTATTCGAAACCGTCGC

CCACGCCCGAGTCGATCAGGATCAGGCGATCGGCGCTGCGCACGAGCGCGATGTTGAGCG
CCCAGTCGAACATGTCGCGTTGCAGGAAACGGCCGTCGAACCACGCATTGCGGTCTGTTT
CGCTCA
>76922..77058_NC_011002
ACGTGGCGAGGATGAGCAAGCGGCAGTTTACGCGCCGCTTCAAGGATGAAACCGGCCTGA
CGCCTGCAAAGGCAGTCGAGCGCATCCGGCTCGAAACGGCCCGGGATTTCGTGACGAGTT
CGTCGTGGACCATCGAC
>80960..81094_NC_011002
AAAAAAACCGTTCCTCTAGACTCGATTCCAACGTGAGTGAAGCGCGCGGCGCGGCCGATG
GCGAAGCGGCAATGTTCGCGATGCGTCGGTGCGACGGATTGATCGCGCGAGCCGCGCATG
CGCATCCAATGCGAG
>88283..88413_NC_011002
CGTGCATCGCATGGCATGTGCGTCGTCGATGCCGATCGCGTCAATGCGGATCTGCTCGCC
TTTTTGCAATCGTGACGCGGGAACTTGCGCGACTCACACACTTCGAGCGGGACGAACGGA
ACCTGGAGAGC
>103526..103672_NC_011002
CGATGCTGCATGGCCGCCCCGTCAAGCAATACGCCGTATCCTGACGCAAAGGAGAAACGC
GCCCCATGCACGCATTCGCAACTTCGGCCGCGCCCGTGACGATCTTCACGGCGCGCCGCA
TCCTGACGATGAACCCGGCGCAACCGT
>115278..115421_NC_011002
CGAGGTATCCGTCGTCTTTGAATCATATGGATTTTTTATCAAAAATATTTCATCGTGGTT
CGACCCACTGGGATACTTCTTCTTCGCTGGCGTCGATTGACCGAGTGTGACCCTTAGGAC
AAGGAGCGTAAAGAATGCGATACG
>115436..115587_NC_011002
GGTCCGACCAAAACCGCGGGTGAAACGGTGGTGGCTCTCCAACGCAAGGTCAACGAAGCG
GTGAGCATCGGAGCGGTATTAGTGGGTGGGATTAGCGTTGTTTTTGGGCGTGCACCGGGA
TTCATCGCCAAGGACGCACATGAAGCATTTCA
>122093..122278_NC_011002
CGGCGCACGCCGGATGCCTGTAATTAGGGGATTTCCCGCGTTTGCTTCGTTTTTATTGAT
CGCTGCTTGTCGCGCTGTTTTTCTTCTTGCATGCTTGTATAGACAAATTCCTGAGGCGGC
TTGCTGCACCGAAGGGATGGCGAACTTCAACGCAAGGAGAAATCTCGTGAATGGCAAGCG
CATCAA
>122589..122748_NC_011002
CTGTTCTCGCGCCCGTATGCGGCGACGCCGGCGACGTTCGCCGTCACCGATACGAAGATC
CTGCCGAAGGCCGCGCCCGGCGCGCCGATCGTCAAGCTGACGGGTGACGCCAACGCGGAC
AAGCCGACCGTCGACACGCTGCGCAAGCAGCTGAAGGGCA
>130355..130492_NC_011002
TGCATGCGTACGCGGCGTGCCGCCCGTTGCCGGCTTGGTTCAAGCTGCCAGTGACGCACG
ACGTCGTTTAGTGCGCATCGCGCAAATCACGCCGGCTCAGGTGGACGATTCACCTATTTG
ACAGCCCCGATCGGCGGA
>134226..134393_NC_011002
ACGCCGGAATCCACCTGGCCCCACATGGTGACCGAACTGCTTTGTGCGTGTGCGAGCGAA
ACCAGCGAGGCGCCGCATCCGGCGATCAGAAAGGTGCGAGGCAGGATTTTCATGTGTCAG
GGACCGGGAAGCGGGCGGACAGACCAACCGAGCGATATATACAAATGA
>135581..135723_NC_011002
TCGTGATCGCGATCATGTCGTTCGCGTTCTTTGCGCAGGGCGTCGCGTCGGCATCGTGGT
CACTGATCGGCGACATCGCGCCGCGCAGCATGCTCGGGGTCACGGGCGGCGCGGTCAACT
TCGTCGGCAACCTGTCTGGTATC
>138427..138569_NC_011002
TTCGCCAACGGCTGCCGTTTCAGGCTGCGGTCATCGCGATTGCCGCGGGGCGGCATCGGC
TGCATGGCGCGGACTGAAAGCTTGGCATTCAGCGAAATCGCCGAAACATGACGCTCTCGT
GAATCATCGGATAAATCGCTGTA
>139663..139840_NC_011002
GCCGAACGACTCGTACAGATGGCGCGCCGCATCGAGCCCCTTGAAGGTCCACAGGTAGGT

TTCGCGAAACCGGTGTTCGTCGACGAAACGCATCGCCAGCGTCATCAATTGACGCCCGAC
GCCAGAGCCGCGCAACGCGTCGTTGACGATGAACCAGCGCAGGTGCGCGACGCCCGTC
>147580..147742_NC_011002
CCTGACCTGCCCGTTGCGGCCGCTTTTGTGTTCAAAGCCGTGATTTCATCGAATTCCGGC
CTGTTCCGCCAATACATTGTCGGACCGCGACCTATAACCGTGGGTTATGCTCTCGCGATC
GTGCGGCGTCGAAAGGAGGGGAACATGGCAAGGCGAACGCTGA
>149887..150045_NC_011002
GTCCGGAAATATCCCTCGATTACGTCCGCTTTTCTCATTCGTCGTCGATTTTTTGGCGCA
CCACACTTTCTTCCAGTCGTGCCCCGCCGGCAGCGCCGGACGCACGCGTGTTGAAGTCAT
GGACGCCACCCACGTGAGGACAGCATGGACAACGACAGG
>191751..191910_NC_011002
AACCGGTCGAACCCGGAATCGAATCGTGTGGATCATCGCTCGGTTTTTTGTTGCGGAAAC
GCGTGGCGCGCGCACCGATCGTCTGGCGATGCCTGAACGAACGCCGCAATGGCTTCCACC
ACACCGCTTGCCAAAGGGAGATCCGGGTTCGAATACGCGG
>191910..192095_NC_011002
GCAAGATTGTCGTCGAGCGTATCCGTGTGAACGGCCTTGAGCACGTGATTGGCATTCGCG
ACGATCGTCAGCCGGGCGCCTGGACTCGCGTCCTTCAACCGTTGCGCGTCCTGCACGCCC
ACCTGAATATCGCGCGCGCCCTGCACGATCAGCACCGGTTTCGCATACCCGGCCAGAAGG
GCCGCC
>192775..192946_NC_011002
CGGCATCAACGTACCGAATCCCGCGCCTCGCCAAGCACCCATTCCGCAAACGCCTGTACC
GGCGCCGCGCTCAATTCGATCGGCTCGGGCAGCACGAGGCAGAAGTTTTTTGTCACCGCT
TTGCCGTCCGGCCAGGGCGCGACCAGACGCCCTTGCTCAAGCTCGGCCCCGA
>196141..196273_NC_011002
CCGCGCCGTTAACCCGTTCGCAGCCTTCGCGCCTCGTCTTGCGTCACGGCGCGCGTCGTT
CGTGTTCCGGAATTCACATCCGCCGCTGGCCGTGCCCGGCCGGCTCAAGGGAACGTCATG
CGAAAACAACTCA
>198089..198219_NC_011002
AACGCGAGCATTTCGCCTTGGCTGACGGCTTCCAGATCGCCGTGAAAATGCTTGCCCAGC
GACATGCGGCCGAGCCCGCTGGGTTCGGCGACGCGACTCAGGGATTCCGGATTCAGCTTC
ACTTCGAATGG
>204538..204677_NC_011002
GACTGGTCTTTGCCGAATTCAATTGCACGCGACGCGGGGTTTCCCGGGAACAAGCCGTCG
CGATCGACATTGCAGGCGACGAACAGAGTTGGCGCAAAGATCGACATCGTCACCATTGCG
TGCAAACCTGAGTGTTGCAA
>205769..205901_NC_011002
CTTCAGGTTGTGTCCATACATGATAGCGGTTTCAACCGGAAATGATAAAGAAAATTGCGC
CGACTACGAATTCCATTGCCGATATGGCAATGCACGACTGCTTCGAACACCTGAGAAGTG
GTTTGACTTGAAT
>205901..206057_NC_011002
TCGTTTGCGTCCGATCTTCTCCGGCGAGACACGCTGCGGAGACTCAATTATTCTTTATCT
TTTTTCAGAGGGATTGGGGAATTCACTGATGAACAGGGATGACACGCTGAAACAAGTTGC
AGCCGAATTCGGATTCGATCCGGACGAAACGATCAGG
>209614..209811_NC_011002
GCCGCCACTTGCGACAACCTTTCGTTCCGCGCCACCCGCGCCACCCGCGCCACTCGCGCA
GGCCGCCATCCGCCGTCTGCGTCAACGCCGATTCCACGGCCGCCGTGTCAAAAACGACCA
TTACGTCAGACGAGAAACGGTAACAAACGTGGTATTTCGATCAAACCCGCATCGGCAATT
TGAACAATGAGGCCGATT
>223035..223181_NC_011002
GTCATCCCGCGCGCCAGCTGGCGGATGCGGCTGATGAACAGCAGGTCCCACGGTTCGAGC
GTTTCGGTGGCGAAGCGATCGTAGTCGTTCGCCGGGCGCAGGTTCTTGAGCAGCATGACG
TTCCTTGTGCGAAAAGGGTGAGGGTTC
>223470..223640_NC_011002
CGCGCATGTGCGGCGGACGCGGATCGGCGACGAACAGCGGATGGCCGAAGCCGTACAGCG

TGCGTTTCGCGTCGAGCGCCGCATCGATCGCCGCGCGGCCGGCCGCGTCGAGCGCGGCGG
CCGGCGCGCCGGCGCCGCCCGGTATGGATCGCGCGAGTGTCGTGAGCCACT
>224470..224614_NC_011002
CGTGCGCCATCGCGGCGATCACGGCCGCGTGCTGGTCGACGACGGGCAGCACATTCGCGA
AGTAGTAGCGCGCGCTTTTCAGCTTGCCGCGATAGAACGCGAGTTCGGCCGGTGCATCGG
CCGTCTCCGGCGCATCGAGCCGGCG
>224894..225038_NC_011002
GAGCTTGTCGCGCACGAGATCCTGCGCCTGGATGTAGTTCGTGCCTTCCCAGATCGACAG
GATTTTCACGTCGCGCGCGTTCTGCTCGACGGGGCTCGCATCGGTATAGCCGAGCCCGCC
GTGCACCTGGATCGCCGTTTCGCAG
>225441..225624_NC_011002
CCGCCGAGCAGCCAGCCTTTCGTCGTGCCGTTCGCGCCGAATACCAGATGCGTGTTCGCG
CAGCCCTTCAGCCCCATCTTGCGCGGCAGGCCGATGCAGTCGACATGGTTCGGCTGCAGC
TCGCCGGTTTCGTCGTCCGGCCAGAAGCGCGGCACGACGAGGCACGACAGCGAGAACGAA
TCCG
>225734..225835_NC_011002
TTCGAGCGGTGTCGCGCGCAGCGCCACGGCGGTGAGGTCGGTGCCGGCCTGCGGTTCGGT
CGCGCAGAAGCACGCGTCCCAGCGGTACGCTTCGAGCGGCGC
>225857..226051_NC_011002
CGTGCCGTGCATCTGCAGCAGCTTGATCGCGGGGCGCGTGAAGCCGCCGTAGGTCATGAA
AGACGGATTCGCGCCCATGAACATCTCGTAGATCATCTGCGTGACGATCGGCGGCAGCCC
GTGCGCGGTGCCGAACAGCGTGTTCGACCATTCGTCGCGAAAGCGCGCCCACAGCGCGTG
GAAATGCGACGGAAT
>226170..226304_NC_011002
TCGTGCGTGCCGTACAGGCCGTGCTCCGCCAGGAACCGCTTGTCGGCCTCGAACAGTTCC
CACAGAAAGAAGCGCAGCTCGCGCAGGCTCACCTTGTATGCGCTCATCATGCGTCTCCTT
CGATGTCGAGCGCGG
>226846..226984_NC_011002
GCCCTCGTCGAGCAGCCCGTCGAGCAGCATCAGCGCCGAATACCAGCCGAGCGTGCCGTT
CTGGCCGAGCGCGAACGGCGTCACGCGCTGCAGCCCGAGTTCGTACTGCATGCGGCCGAC
GACCGAGTCGCCGATCCCT
>257037..257206_NC_011002
ACTTGTCGGGCGGGCGCGGGGACGGCCCGGCGAGGTGGTTCGCTGTCGACGAAATGATTC
GGAAAAATGTGCCGAAGCGATTTTGATAATTCCGGTCTTTATTTCCGGTCAAATGGACTG
GTGGAATCGAAATGGCCCGAATGATGAATTTGTGAGCGTGCTGTCGAAAT
>257402..257537_NC_011002
CCAGACGCGGCTGGAATCGGTTCTTTCCCCGACTTCGGTTGCCACGAAAATTGCTGTTGC
CGTCGCCTCCATGGCAGCCGATATTCAATGTCACATTGAATGATGGCATTCTCAGATGAT
GAGAGGGCGATGACAA
>275097..275279_NC_011002
GATGATCATTTCGACCTCCAAAATACCGAAAAAGTGTTTGGCATTGACGAGCCGCGTCAG
CTCGCGTTGGCGAAGAATCCCTGTTTGCCGCGATTGTCGTCGATGATGTTCTCGGCATAG
CGGACCGACGCACGACGCGCGTCACCGGCGTTGTCGAAGGGGGACGTGTCGCGCAATCTG
AAG
>287086..287241_NC_011002
GCCCGCGTACCGACGGATGGACGGCCATGCCGCGCAGGCGCCAGGCCCCGGCGTTCGAAA
CGCCGGATTGTGCTTCTTTACACACCGATGCCACGGCAACGATCGTATCGTTATCGCGAA
CGGCAAAATGCATTGTCGTGGGGGCGAGATCGCCGT
>287455..287595_NC_011002
CCTATCCGCCACTTCTTTTCCAATCATACGCTAACTGCGATCACAGATGACGTCGAAAGT
CTCGACAGATGCAGAGCTTGCCTACGACGAGATCCGATCGAGGATCTTTGACGGCCGCCT
TGCGCCTGGACAGAAAGTATC
>289327..289477_NC_011002
GGCTTTCCGGTTCAGACGACATGGCCGCGGAAACCGTGACGCGTCAGGCAAGCGTGTAGG

TGCTTCGGAAGCGACGGGCGGCGGGCGTTTTTCAATGTCGCGCCCGGCGAATGTCACGCA
TTGGGGGGCGCCCGCTCGGCCATCCAGGCGC
>289553..289691_NC_011002
AGAAAGATGATGGTTGGTTCAGCGTCCGCAGATTTTTTCATGCTGCGCTTAGGCACGCCG
CATCCGTCGTGGACATTCACCGATGACAGCAACATGCTCGGCATCGGCGATTGCGACGGG
GTGATGAGTGCCATGTTCG
>315627..315763_NC_011002
CGATGCTGCTGCCGGCGCGGCTGGTCGAGCGGTAACGACAATCCGGTTCGCCGCGGCGCA
GCGCCGCGGCGCCTTTTTTGACGGGAATGACATGAAACAGCAGATTGGCCTGATCGGCGT
GGGTTTGATGGGGCATG
>316780..316952_NC_011002
GGGTATTACTCGCATGACAAGGCCATTGCCGACGATGCCGAAAAGATCGCGATTCAGCGC
TTCAACAGCAAGCAGTTCGATGCGATCTACGACAGCTTCGCCGAGAACGAGAAGCAGGCT
ATTCCGAGGAAACAGGCGCTGGAAGCCATGCACGCCACCTTTGATCGTTTCGG
>333449..333583_NC_011002
CGGCCGTTTGATGTTGTGGTGCCTGCGGATCGAATCGCTCTGAAATGGCGTTGTCCGGTG
AATTCATTGTTGAAGAAAAAATCGCGTGATGCAAACCACGCACGGCCGGCCGGCATTTAA
TGTGTTCAGCGATTC
>357264..357459_NC_011002
GCAGATCCGGTAAAAAATCGGCATCCGGGAATTTGAGTTTTTTGCCTTTCTTAATAATTT
GCGTGATTAATTTAATAAGAATCGCTTTCATCCGGAATGAAAAGTAGTTAATGAAGATGA
AAGAGTGTTGAGTTTTGCCAAAAAAATGTCGGCCCCATGGAAAACGTTTGCGAGTCGGGC
CTATATTTGACTCGCG
>370809..370953_NC_011002
TCAACGCTTTGCCATATCGGGCAATCCGCGAACGATTCGCGCCCCCTTCGCGGCGGCAGG
GTCCACGCGTCATCGCTCGAGCTGTTCCGCGGTTGTCCAACCGAAATGAGGTTTCTCATG
AAAACGAAAGCGAAAGCGAACGTGA
>385598..385735_NC_011002
AGGTAGATGAGCAGGGTCATCGTTCTTCTCCTTGCAGTGTTGGCCCGTGACGCCGATTGC
CGGTCGTGCAAAATTTACCGCTCGCGCGGCGAGGCGCGATGAAAGGCGCGATCCGCGCAT
GCGGGAGCATGTTGCGAG
>390461..390594_NC_011002
TTTGGCAAAGTATTTATCCGCTCGATAAAAGATTTTGAATGACGGCGCCAGCGCGAAATC
GTTTGCGGAATGTGTGTATTTTGCGTAAGCCTTTGAATAATATAAATTTTACAATCTATA
AAGAATGCCGGCTT
>390719..390843_NC_011002
CGGTCCAGGCCGGGGTTCGCTGCACCGAAGTGAAACTCCGGTTGCCGGAGATGAAATAAA
AGAAGCGTGGTGGGTCGCGGTCAGCGCGGCGAAAAATCGGGGAAGTACATGACGACGGGT
TGAAT
>395117..395314_NC_011002
GCTGGCCACCGCCCGATTCTTCGAGCGATACCTCGGCTTCACGATCGCGGGAAGCTGGGA
AAAATCCTGGATCCTCAAACGCCCGGGCTTCGACGTCGTGATCGATCACGTCGGCGACGG
CGTGCCCGCATGGCCCGCCAATTTCCACGTCGGCTTCGAACTGCCGAGTCTCGATGCCGT
GCACGCGTTGTTCGCGCG
>400289..400485_NC_011002
AAATCCGCGATTGCACATGCACGCTGCTGCTAAGCGCCTCTGCCGCGCAACTTTATGACG
CACGCCGATTAATTTATCGGTTTATATCGCTTGGGATTGCCGGATCGAGGGAACAGAATT
TCCGTTGCCGGAACCCGGCTGCAGCGGGGCCGCATGCGCTCGACGCCGCGACGGTCGCGA
AGGCGCCTGCCCGGTTC
>422485..422621_NC_011002
TCGGGCACGGCGGGCGGCCTCGCGCCGACATCTGTGTTCCAAACAGTACATTTGCGCGGC
AATTTAATGCAAATGAGAGGCATTTGCAATCGTAATTTTGAGATTGGGAGGGCGAAGCGA
AATGGCGGGCCGCACGG
>422954..423085_NC_011002

CGCAGCCACCGGTTGCCCGCCTCGTGGTGATATCGCGCATGCCAATGCTGCCGTACCGCG
AATCCGTCGACCGGGATCGGGCAGGCATGAACCGCCAGGTCGTTCGCGTGGGCCAGCGTC
TCGCCGATATGG
>441167..441320_NC_011002
GATCGCTCGAACGACGAGCATCGAAACATCGTTCGCGCGATACGCGATGGTGACGTGGAC
GCAGCGTGCGCGGCTTTGTCCGCGCACGTTCTGGCGGCGGGGGATGCGCTGGCTCAGGCG
TTGCGTAGGCGGTAAGGACACGGGTCAAACCGGC
>466699..466882_NC_011002
CGTTTGGGACGCCGGCTGCAATGGCGCCGGCGCGGGTCAAATTCTCGTTAACGGCACGGT
TAACAAAAAATCGAGGCCACGCGTGGTGGCCCGCCGACGATGGCGGCATCGCACGGCGAC
GCAAACGTATGCGGCCTCGCAGCGAGATGCCCGATCGCCATCCGGACCTGCGCAGCGTTT
GCCT
>473509..473680_NC_011002
CCAGGCCGGGATGGTCCTGAACTGGCACTTCTGCGCGGAGGACCAGATCCAGCCGGCGCA
GTATCTGAAGAATTACGGTATCTGGAACAGCGGCCCGTATTCGATCGTGATGACGGCCGG
CCCTGACGACGTGACGATCGGCGATGGTCTCGGCATCAACCTGACGAACAAC
>500495..500632_NC_011002
TACGAAGCACGGAGAATGTCATGTCGGAGTCCCAATCCGTCGACACGATCGGCGCGCTCG
CCCATTTCATTCGCGCGGCGCGAATCCAGCAGGGGTTCACCCGGGACGAACTCGCGAACG
CGACCGGGCTGTCGCCGA
>629782..629939_NC_011002
ACCACTCCAGCAAAAACCGGGCCATTGCCGTCAAACGCTTGCCACGCCGGCACGAGACGG
AAATTTCAGCGAACCGCTCGGCGTGCCGAATTCCTCATTTCGGTGCACGAACCGGCGCCA
TCGGTAACGGTTGCCCTCGCAGCGCGCGCCGGCGCACC
>632267..632432_NC_011002
GACCTTTGAGGCGGCCACGTCGAAGGTTGCTGTGTACTCGCACCCGACGACGTGTGCCGA
CCGCTCGATCAGGCAATCCGCAAAATCCGCCTTCGGTGATGCCGCGAAAACGCGAAGTGC
CTTTCGCACCGTGTCAGCGCCTTCAACGGCCAGTTCCTTCGTAGCG
>642577..642718_NC_011002
TCGCGAGTTCCATCTACATCTTTTTCCTCGCGCTGAAGATGTGGACGCGTAAAGCAAGCT
GGGCGGCGCGCGGCGCTTCATTGCGCGACGCCGCGCTTCGGAAACCGCATCGCTATGACG
AGCCCGCCGTCGGCCGCGTTGG
>644708..644867_NC_011002
GGATACGATCTGACTGGCTTTCGAAATATTACTGTGATTTGCGATTCCTAACGGAACCGG
CACAAGGGTCGGGCCGCGAGGGTATTACGTATGCCGAAATCCCGATTCTCTTCCCCCTCT
CCAGGCGCGTCACTGGACAGTAAGTTTCCGTAAGCACAAA
>669830..669976_NC_011002
AAAATGACCACCTCACCACCACCGAAATCCCCCGACGATACCGTGCAGAAAGAGCCGAGC
CGCGAGGAACTGCTCAAGGAAATCGTGTATTTGCGCGCGGAGGTGGCATACCTAAAAAAG
CTCGATGCGCTGCTGCAATCCAAGAAG
>670266..670409_NC_011002
TCTGGTACGCCCAAAGAGATATCGTTCCTGGCGAGGCGAAGTCGGCCGCGTTGCCCCGAA
TCTACTGCGGCGTCAGTTTGATGCTGAACGCCCGAATCAGAAGTGGGTGACCGACGTGAC
CGAATTCAATGTAGAGGGCCAAAA
>670615..670758_NC_011002
CTCACCCAGAGCATGTCGCGCAGAGGCAACTGCCTTGATAACGCAGCCATGGAGAGCTTC
TTCGGCACGCTGAAATCCGAGTGCTATCGGCTGACGCGCTTTGCGAGCGTGGAGCAGTTG
CGAGATGCGTTGAGCCGCTACATC
>682721..682913_NC_011002
CATGACAGCCGCTGTTCTTTCATGCAAACCTCGTTTGCGAAACTGGCCAATTCCCGTTCA
TTCTCGCAACGCGTACCGTATCAGGCAAATCCCGGCGCAAGCCGCCTTCACCGGCACACC
ATGAGGAACGCCGCCCATGCACACGCTGATCGTTCATTGCCATCCCGAACCCCGCTCGTT
CAACGCCGCGCTA
>684076..684220_NC_011002

TCGATATGCCGAAAGCGAAGTTGATGCAATCGCCGAACCAGACGCCCGAGACAGGCAGAT
CGTGACCGAAGAAGGCGCTGTCCTGCAACGCGGTGCTGAATACGCCACGAATGGCGCCCT
CCTTGTCGTCGGTGATGGAAAGTGT
>684279..684414_NC_011002
CGATCACGCGCCAATGGGCCGAGCGTTCCAGGCGCCGGCGCCTTCTCGCGGCAATCGGGC
TCGCAAAGCCGGCGCCGCGCATGGGCGAGCACCGTCGTGACAGAAAATCCTGGCGCGAAT
CCAGCAATAGTGATGC
>687343..687498_NC_011002
GAGTAAACGGCGATACATGGGTATTTTCCTCACGTTTCATTGGCGCCGTATTGCTTCGCG
CGCTCGCCCGCGAGCATGCCCCCGGCCGGTGCGGGGCGCGGCGAACATCGTCCGCGGGCC
CGCGTGCCAAGGCTCTCTCAGGTTCTCTTTTCGGCT
>708654..708802_NC_011002
CGCGGCACAATGGGCTACGGTTGCTCCACATGTCGCTTCCCGATTTCGCGCACGCGAAAA
TCACCGTCCGCCCTTCTATCGCACCGTGCCGTTCTTGCAGTGTCCACGTCACATGCCTTT
GCTCGCATTGCACAATGCGGGGCGGCCGC
>710870..711047_NC_011002
CGACGTTTATCTCGCGCGTGACGATGCGGCGAGCACAGTCTCGATCCGCCTATCTGGCTC
CATTTCGCGTGACGATGCAAAAGTGGTCGACAGACGGATCCAGAAGTTTGCACACGATCA
TTGCAGCGCCGGCGCCATCCTGTGGCGCGGATGGAACAATTCGTGGAAATGGTTCGTG
>711601..711766_NC_011002
GACACGACATCACACTAAGGTTGCTGAATCCATCGAGCGCCGGAACTCGCGAGGCGAGCG
GTATTTCAGCGCGCTATGGGGATGCTTCTCGTTGTAATGCTCGAACGCGATGGCCAAGTT
GCGTGCAGCAGTCGCTGCGTCCGGTTTCGGCATGAAGGCGACGTAG
>712628..712787_NC_011002
CATGCAGGTCGAAGCCCGCTGGCCCCAATGGAATCCGCTGACTCATGCCTCATTTTTGCC
ACGGACCGTGATGTCCAGAAGACGATGCAGGTTTGATGAGTCCCGGTCTTACCTGTTTGC
CATTACAACACTACCTTGCGCAACCTTGGACTACAGATCG
>712824..712973_NC_011002
CTGCAGTAACTGGAGGCGGATAGCTCTCGCAGTATGAGCGGCCATGAGCTAGCAGAGCCC
AGATCGTCCGCGCGATCTTGTTGGCCAGTGCAACGACTACGACATTGAGCGGACGCCGCA
TGCTCAACTGCACTGCCTGGAGCAGCCCCT
>713043..713195_NC_011002
GGCAGGAAGTGATGGAAGTGTTGACGGGCCCTGAGCGTCGGCGGCGCTGGACGGCGGAGC
AGAAGCTGGCGATGGTTCGCGAGAGTTTCGAGCCGGGGAAATCGGTTTCGATGGTCGCGC
GCCAGCACGGCGTGAATCCGAACCAACTGTTCC
>713580..713747_NC_011002
GGCCATCAGGTCAAATTGATGCCAGGCAAACTGGTCAAAGCGTTCGTGACAGGAAACAAG
AACGATGTCGCCGATGCGCGCGCAATCTGGGCAGCTGCACGACACCCGGGCGTGAAAGCC
GTCCCGGTGAAGACCGAGGCCCAGCAAGCGGTTCTGGCACTGCATCGT
>714169..714365_NC_011002
TCGGCATCAGCAAGAGAGGCGATCGATACCTGCGCACGCTGATGATTCATGGCGCACGGT
CCGTCATGAAGCACGCAAAAGACCCGGGTTCCTGGGCAGTGCAGTTGAGCATGCGGCGTC
CGCTCAATGTCGTAGTCGTTGCACTGGCCAACAAGATCGCGCGGACGATCTGGGCTCTGC
TAGCTCATGGCCGCTCA
>715002..715138_NC_011002
GTCTTCCGCCGAGGCACCGACCCGCTTTACGCCTTCACCGCCACGCTCGGGTTTAGTCGG
TGGCGCTGGGCACGCTTCACCACCGATGAACGCGCCGAGACGCTGGTCGCCTGTCATCAT
GCGCTGTTCGAAGCCTT
>715452..715591_NC_011002
GTGCATCCGGTTACAGGGCTGGCACCCGCGATTCTGTTGGAGCAACGCGAACGGGCTTGC
TTGCGCGATATGCCGGGTTACGCGGTGCCCCGATTGCCGGCTCGTGCCGTCGCTCGGCCA
CGCGTTGATCCAGCAATGTC
>715660..715832_NC_011002
ACGCATCGACGGGCACTGCCAGAGTTTGAAGCTCGAGGGGCTCATGCACCGATACGTCGC

TCTGGCCAGCGACGCGGCGGCCAAACAATGGAGTTTCCTCGATTTCCTCGAGAACGCGCT
CGCGCACGAACGAGAGACGCGGCAAGTTCGTTCGCGGCAAACACTGGTGCGCA
>715923..716069_NC_011002
CCCTGCGCTTCATCGAACGGGGCGAGAACGCCGTGTTGCTTGGCCCGTCAGGCGTGGGCA
AGACGCATCTCGCGATAGCGATCGGATACGCTGCAACGCAGGCCGGCATCAAGACGAAGT
TCATTACGGCGGCGGATCTAATGTTGC
>716189..716321_NC_011002
AGCCACTTCTTCCAGATCGTCGCCAAACGCTACGAGCGCGGTTCAATGATCCTGACCAGC
AACCTGCCGTTTGCGCAGTGGGACGAGACCTTCGGTGGCAACACCACACTGACTGCAGCG
ATGCTCGACCGCA
>716403..716546_NC_011002
CATCGAAGAAGTAACGACTGGCTCAGTTTTACTTTGCGCGATCAGGCGCGAAGTGGCTCA
GATTTCAAATGCGTTTGACAACATCACGGTCCGTGGCAAAAATGAGGCATGAGTCAGCGG
ATTCCATTGGGGCCAGCGGGCTTC
>721034..721218_NC_011002
CATCGTTCGTCTCTGATTCCATGCGATCACACCAGAGCATAACAAAACAGTACGCCACGA
GCGATTGAGCAACGGCTCATATCCTATCGCGTCATTGTCCGCCAACTACAATCTCACTGA
GCTGAAACGCGCCGGTGCGTCGCGCATCCGGTTGAAAAGCTCACCTCCCGACCGCTTCAT
GTCGC
>722434..722589_NC_011002
AGAAGTCGACCACCCCGATGTATGATTCTCACACGTTACTCTCGTGAAATCGAGGCCACC
CGGGCCATTCTCTGACCAGTCACCGGGAGTTGCGAAATCCTACTTCTCAGATTGCGAAAT
GACCTCTCCGAAATCGAACCCTTTTAATTCATTTTG
>722621..722756_NC_011002
AATACTGTAGCAAGACTTTCTTTTAGGCGAAATATATCAGCCCCTTCAATGCGACTCTTC
AGAAGAAATATTTTATTAAAAACCGTCGCCTGAGGGTAGACTTTGAATTTTGATGCATCC
ATATCGACGGCATTCT
>723279..723428_NC_011002
CCGTCCGGCTTGACAACGCCCAATCAGAAAAACGGGGCGAACGACGTTGTTAGACAACTG
AGCGCTGCCATTTGTCGTTAAAATTAAGAAACTCTTTTTTCTTCACCATCCATTTACCCG
CCTTTTCCTTCATCGTGAATCGGGATATCG
>723747..723893_NC_011002
CATTCATCGCCTTCAACCCTTAATCATTGAAAATACCGCCCTAGTGCGAGATCTCAGCGA
CGATGACATCCATATTTTAGCCAGGCCAGTCGTCGACCGCGACGAACTGCATTTGCTGTG
TTCCGGCGACGATCATATTCAGAAATG
>723986..724116_NC_011002
CTATCCCAATATAGGACCTTTTCAACGTTATCGAATCGCGCCAGCGCGTAGTAATTACCC
CCTGGATCATCTCCAATAACAAAGAAATCGTTTAGAAAATCGACCTCACCTCTGACACGT
TCGTTTTGCTC
>733584..733733_NC_011002
ACGTTTGCAAGCGAAAGATACCTGAAATTCCTTTGCAATTTCCCTGCTGGCACCCTCTTC
CGTCACCAGTCAAATAAGCGGTTGGGAACCGGCGCAGGGTTCGCCGCGTCGCAGTACCAC
ACCGGTCTCTGCCGAGATCAACACGATTTG
>733861..733989_NC_011002
GCTGAAGCGGGCTTGCCCGGCATCCAGCGCCGCATCCGTTCAGTCGAGAGGGACACGAGC
GCGTCGAATCGCCGGCCGGCATCGGGCGCAAACATGCCGAAGCGGGCGCACAGCCGCTTG
AGCTCATCG
>734531..734673_NC_011002
CTTGCCCAGCCTTCGTTTCGATTGGCTCCGCATTGCTCGAGAACAAGTTGACTACTGGTG
TGCAAAACAACCGGAGATTGTCGGCCGACGCTGCAATCAGTCGTTGGGCGCAACGTGAGT
CAGGGTGGACGCCGGCAAGGGCC
>735231..735392_NC_011002
ATCCACGCGGAATTGGGCCATCGCACAAGATGGAAACGGCCGCAGGAACGCTCCGTGGAT
TGTGTTAATCAACGCGTGCGTAAACTCCGGCACCTCATCTTCGAGTCGGATGTCGTGCCG

AGCGCACATCAACGCGAACGATTGCAGCAGCCGTTCGACATG
>739281..739413_NC_011002
GCGTGACCGGATGTCTATGACAGAAATCGGCTATTGCATGACGTACTTCGTCGTTGGCGC
TCGAAATGGCATGCAGCCAACGAAACCATTGGTCCCGCCGGAATACCTCCGAAAACGGAC
TCATCAATTCGAT
>742668..742840_NC_011002
AGATCCCACACTCGAAGCGCCCGCGGAAACGTCAACGTCGCAATCAAGCCGTTACGTTCG
GCCTCACCGCTCAAGTAAAACGTCCCCGGCCGGCAATCGTTGCGCGCGAAGCGCGCCTCC
CAGACCGTCGTTTCGGGCTCGAGAGCCAAGTACAGCCAGCCGAACGGCGAAGT
>745619..745768_NC_011002
GCCGAAGCCGCAGCGTTCGCCCAGGCGGGATGGACGGTCCTTAAGGACCCGACACTGCTC
GACGACGACAGCCGCAAGAAACTCGCCAAGGCCTTGCAGGAAGATCGCGACGATTGGATC
GAGACATTCGAGACTTTCGATGCTATTCGT
>748152..748291_NC_011002
ACGTTCGACCCCGCTTTTCTCGTCAAACTCGCTTGGATGTGCGCGGCGAAATTTGGCCGT
TATTTTCGTCTTGCCAATATTTGAATCACCGTGCAACAGCAGGCACGGCATACGTTCGCG
GCGGGGCGTCTCAAAGAGGC
>748460..748598_NC_011002
GAAGGCCTGAACTGACGTCGCGTAGTCGACCGCCTCAGCCTCCTGCTCCTTCACGTCATG
TCTGTACGGGACAATCCTGTCGCGCAACTCTTCCAATGGGACGATCTTTTTGCGTGACCG
TCGACGTGCACGGGCCGTT
>749516..749650_NC_011002
CTTCCACCTGCATGGGCAGGATGCGTCAGGCAGGATGGTGTTCCGTAAGAAGCTCACGCG
CAGCCAGATGTTTACGCTGCTGGGCAATTTCCCGAGTTGTACGGTGGTCATGGAGGCCTG
CGCGGGTGCTCACTG
>749700..749884_NC_011002
CCTCAGTTCGTCAAACCGTTCCGGCAAGGCAACAAGAACGATTTCGCGGACGCCCAGGCG
ATCTGCGAAGCAGCTTCGCGTCCGAGCATGCGTTTCGCGAGTCCGCACAACGAAGCTCAG
CAGATCGTTTCAGCCTTGCACCGCGTGCGCGAGGGCCTTGTGCGTCAGCGCACCTGCACG
ATCAA
>751081..751215_NC_011002
TGCAGTCCGGCAAGCTCACATCTGCGTCGAACCTCCAGCACGAGATCGGTTGCAGGATGC
GCCAAATGTCGTTGAGTGGGCAGCCAGACGTGCACTACTTCGTCAATCACTGCATCAACG
GCCCCTCCCAATCGA
>751453..751614_NC_011002
TCGTCAGACGCGTTAGCGGAATTAAAGGACACGTATAACGCGTTAGCAAAATAAACGGAC
ATCCGATCGTGCGCGCGTTGCCAAAACGCAACGCATGAAAAGACGCGCGAGACGGCAGCA
ATGCCTTATCTCGCGGGGCCGTGGCCGTGTCTTACATTAGCG
>754026..754195_NC_011002
ACGCAAATGCCGACACGAACTGAACCGTTCGGTTTAAAATGTCCGCATCTCATCGCGCGG
TGTTCCGAAATGGGCAGACCCAAAAAATTCAGCCGAGACGGCGTGTTGGAAAAGACGCTG
CCCGTCTTTTGGAAATACGGCTTCGCCGGGACATCGCTCCAGCAGCTGGA
>754523..754670_NC_011002
CGCCACCATCGCCACGATCGTCACGGTGTTCTTTGCCGGTTATTGCATCGAACGTAATCT
GCCCGAACTGGCGGAAAGCGACGCGGTCGGCTCTTTTATGCAAGCGCTGCGCGTTCTCTG
AAGCACCATCGCTCGGTTGCTAGTCGAT
>760020..760212_NC_011002
ATACGATTTTGCGCGACAGCGCCCGATCGGCGGAGCGTTCCGCCTTCCTCTCGGCCTTCT
TCACTTCCTTCCGGCTGACGACGGCGTCGGAGGACGGCCCTGAGTCCTGCGCTGCTGCAT
GGGCGGAAACGCCGAAAGCCAAATACGTTACGTACAGAACCGACGCAATAATTTTCTTCA
TCACAGGCTCTTT
>769985..770117_NC_011002
AAAAACATGCGTTCGTCGATACCGGCGCCCCGCCGCGGGCGTGTGGGCATCGACCGTGGG
TAAAGTCGCCGTTCTTCCACTTGCATCGAGCATCCCGCTGCGCTTCGTTAGCGAGCGAAC

GACGTGCACGCGT
>771571..771715_NC_011002
CAGCGCACGCGCCGCTGCGGCGCCGACACCGTCATTCCGATGCGGGATCCAGCCCCAGTG
CTTGCGCGCCAAGCACGAATTCGGGCAGCGTTCGCGCGCCCATTTTGCGCATCGCATGGC
CGCGATGGATCTTCACGGTGATTTC
>774535..774732_NC_011002
ATACCTCGGTTCGGCATACCAGCGGGTCCTTGCCCTTCAATTCAAAAAAATGGCGAGATC
GCGCCTTCGAGGTCCGCGGGCTCGACCGCATCCTATGCCGCCGGCATCGGCTCGCAAACT
ATACTTCGGTATAACGCCGCGCCCGATATGCATGCTGCCACCCAGTCTGTGCGTCTCACG
CCGTAGCCGACAGCTGCT
>778118..778272_NC_011002
GGCGGCAAGTGACAACGCGCCGATCGATTCTGGTTCGTCCATTTCGCCGTCGCCAAAGAA
TCCCCAGACCCTGCGTTCCTCGGTAGCCGACAGGCCGCGGTTTTGAAGATAGCGCATGAA
ACGCGCCTGATAGATGGCATTGATCGGACCGATGC
>778375..778507_NC_011002
GAGTGTCGTCGAGAAATCCTTCGAGATACGCCCGCGCATAAATGCCCGGCGACGAATGCG
GCTGGAAGAACACCAGGTCGCCTCCGCCTGCGGCGTGCGCACCGCGAAAGAAATGGTTGA
AGCCGACTTCGAA
>779069..779201_NC_011002
ATCCGGCGCCCGCCGGCTCGATCGCATCGACATCGCGATCTTGCAGCAATTGCAGAACGA
CGCCCGCATCACTAACGCCGAACTGGCGCGCGCCGTCAATCTGTCATCGACACCATGTTT
CAACCGCGTACGA
>780426..780578_NC_011002
AAACCTTCGGCCGGCCCGAACAGCCCATAGCCTGCATTATTCACCACGACATCGACCCTG
CCGTATTCTGCAAACGCTCGCTGCACCACCTTGGGGATGGCTGCGACGTCGGTGACATCC
AATGGGAGTCGCCGAAACGATTCCGGGTACCTC
>780578..780702_NC_011002
CTCGTGCAGGTCTTCGACGGATTGCAGTTCACGTACCGTGCCAACGACACGGTCACCACG
CGCGAGCAGTTGTTCGCTCATTTCGCGGCCAAAGCCGCTGTTCACAGCCGTAATGAACCA
GGTAT
>782686..782820_NC_011002
ATACCTCGCCATCGACGCTTTTTTCCTTGACGGTCATGACGCGCGCACAAAATCCGCGTC
CACATGTGCGCTGGATGCGATCCGGCACGGACGCCGTCGTGGTGGCTCTTCGCCAGAAAG
GCCGTTTCCTCATGA
>782876..783018_NC_011002
GGGCATGAGTAGTACCACGCCCTCAGTCCGCGCTATGGACGCTCAGTGGAAATTCCCAAC
GCCCGGAATTCCACGTTCCAACGTTCAGGAACGGTTCGATCCGTTTTGTCTCGATAGCGT
CGCGAGCAGAGACGAGAGGGTCG
>783293..783486_NC_011002
TATTGGTCGCTGGCTTGCCGAAATGGCTGGACCGGAACCGAGTGCGCAAGCTGCGTACGA
GGCCTGTTTCGCGGTTCTTAACGACCCATACTCGCTCTTCTGCTTGCTAGACAAACTGAT
GCCGATCTTGCTCGATGCGGCAGATGACGTACTGGCTGAGGCGGCAAAAGGCACGTTCGA
AGCCGCGTTGCTCG
>784222..784364_NC_011002
ATTCAGTGGCACCGCATCTTAGACGAGCAGGCACTGCGTCCGCCGCGCAGCGGTTCCTGG
CCCAAAGATATGGTATGCATATTCCGGAAGGATTCACTTTCGTTCGTCCGCACGAGAGGG
GGCACGCTGGCGACGAGGAGCGG
>784509..784696_NC_011002
CAGTCGGCAAACAGGGACGGCGATGGCGGTGTCGATCTCTATGCTGTGGGCGAGGACAAT
CAATCGTGGATCGCTCAGTGCAAATGTTGGGCACAGCATCGCACCGTGGGGCCAGAGGTC
GTGCGCGAACTTTACGGCGCCATCGCCCTTGCGGATAAAGGGGCGTCGTCAAAAAGTCGC
GGCATCCT
>784816..784973_NC_011002
CCAGCGCTATGTGCCAGTGACGTCAATAGTCCTATGACGACCTCAGGTTCCGATGACTGC

TGGCGTCAAGTTGCGCATACGCTCGCCGGTGAACGCCGCGTAAAGGTCCGTTGTCTATGC
GTAGCTGACGCAAGAGCGTCCACTTCAGCTCCCCCAAC
>786704..786860_NC_011002
GAACTTGGACGTCGACACGATTTCGGCGGCAATTGCCAGCCTTGGTAAGACTTCCCTTAT
TACACGGGAAACCACAGATGCGGGCGAGGTTTACGCGTTGGGCACATCAATCCGGGAACT
GTTGCGTGTACAGCCCAGAAATCTGAAAGTGCGGGCC
>788081..788216_NC_011002
CGGTTAAGTGTCTCTGCCGCCTCGCCCTTGAATCCAAGCGGCGAGGCGGCTGACGCGATA
CGCCCGCCGCGATTTTACTTGCGTCCGCATTGGACATAGCGACCGGTTTCAAGCGGCAGC
ACGAGGCAAACATTTG
>792765..792921_NC_011002
GCCGCTCGCCTTAATTTGCTGGTTGTGAAGCGTCTAGAGCTTGTTTCTCGATGCGGGAGG
TCTAGCCATCGGCACTTTGCCGCCGATGGCGTTACGAGCTCGATGCCAGTTGTAGTGGTG
CTGCCAACTGGCCCAGGCTTCGTTGCGATGGGCCGAG
>793199..793331_NC_011002
CGTGCTTGGCGAGCCAGCGCAAGTACAGCTCCATCTGACTCTTGTACTCGGCCTTGAAGG
AACCGAGCTTCAGTTCGATCGCGACGAGGCGTTTGAGCTTGCGGTTGTAGAACAACAGGT
CGAGATAAAAGTC
>794674..794801_NC_011002
GTGTCCCTGTCCGATTCGACCGGCAAGCCAAAGTCCGGATTCGCCGCGCCTCTCGACGAA
CGCTAAAAAGACACTAAAAATCACGGCGCGAACAAGTGCTCGGTAAACATCGGCGGCTCA
CCCGCTCG
>806108..806248_NC_011002
GCGCTCCATCGTGATGCATTGCTCGACCGGACACACATGCATGCAAAGATTGCACCCGAC
GCACGCCGAATCGACCACTTCGAAATGCCGCACGCCGTCCTTCGTCGCCGTGATCGCCTG
GTGCGAGGTGTCCTCGCACGC
>810233..810398_NC_011002
ATCCTGTCCGATTGGACAGGTTTTAGACGATCAAAATGACCAGCGCAACGTCTTTCTTGC
GCGACGCGATGGAGCGAGAAACGTGCCACCGCAACGCAGCATCGCGGCGGAATCGCGGCA
TCCGGATGCGTGCGCGCGACGCGAATCAGCAGCTAGAATGAAGCCC
>818118..818267_NC_011002
GCCAACAGAACCAATCTTGGAAGAAGCAGTAATTTTGAGTTGTGTGTATCACACACGAGA
ATTCAACATGTCGCTTTGGATCGCGACCGGTGCTTTCGCACGGCGCCGATCTACAAGGCA
GACTCGTTATAGGATCAAGCCTTACGGGCA
>823381..823569_NC_011002
AGGATCAGCTCGCTGCCATCAAGCGCCCACACTTATCGGCTGTTAATTTTTAAAGAGCAT
GTCTGCGAGGGATCTTGAACCCCTCAGCAGCGCTGCGTTTTCAGCAGCAGAGAAGTGAGA
TTATGAACCGTGTTTCGCAGCTCGTCAACAACTTTCTGAACGACTTCGTTGCGACTGCGG
GGCTCAACT
>830182..830326_NC_011002
AAGGTTTCCGGCTTGCGATGGCCGAACACCGTGACGTTGTCGGGCAGCGTCAGCTTCTTC
GTCATCTCGCTTTCGGCATCGTGGTTGCCGCCGAGGACGAACGCACGGATCCCGGCCTTG
CGCAGGCGGCCCATCTGCTGGCCGA
>865330..865475_NC_011002
TGCGGCTGCACGGTGACGACGCCGACTTCCGGCGTTTGCGGCGGCGGTGCCGATTCTTTT
TTCCCGCACGCGGCCAGGAAAACGGCAGCCGTTGCGGCAGTGATTAAGCGGTATGGAACC
CGTTCGACGCGCATGGAGCGACCTCT
>866766..866907_NC_011002
CAGTAGCAGTCCTGATTCTTTTCCCGCGATGCGCGGCGATGACGCGCGCGGCAAGGCCGG
CGGCGGCCGCGCATCACGCGGACGGTGTTGCCCGCATCATCCGGGCACGCTGAAACCGCA
TTTCCTTCAGCAGGTGCAATGG
>869202..869370_NC_011002
GAAGGAGGCGCGGCCGGAGCGCGCGCATGCGATGAAGCAGGTCGAACCGATTCTAAAAGG
GGGCCGCTGCGCGCCGAACCAACGATGTTCGATATCGAAAGCAAGCGTTTTCATAAGGGC

AATGCCGATGGCACGGGCATCGTTGTCAGGAACGATCGGCGCACATGCG
>875625..875793_NC_011002
CGACGAAGGTAGCAGGTTATGGGAGAAAACGGCGGTTGTGCTGCAAGGCACCACGAATCC
TCCCGACCATCCCGATTCCGCGCGTTTGCCCCCATGGATGGGTACCTTCGTGGTGGCAAA
TCTTGTGCCGGGTTTGCGTGGTCGCCCAGGAAAGGTAGCGGCTTTT

B. Extracted sequences of only intergenic regions from ChIP-Seq peaks with P value > 0.0005

>34503..34658_NC_011000
TTCCACATCCTGATTGTTCTGTTGCAGGCATTCATTTTCATGATGCTGACGCTGGTGTAT
CTCGGCCAGGCGCACGACAAGCACTAAGCGCGGCGTGCAAGAAAGAGTTTCCGTTTTAGT
TTTTTAAATCTCAGTTCCAAGTCTTTTCACAAAGGA
>61713..61898_NC_011000
GGACGATCTGGGCTGTACTGGCCCATGACCGGCCGTACCAAAAGGGTTACGTGAGCGTGA
AGCCGGCCTGATCGGTGAATGCCACGTTGAAGATCAACGTTTAACACTGGGTGAACGTCG
AAAGGTTGCGCAGGAATCGAGTGTGATGACAAGCCAGGTAGGACCGGGACTCGCTAAACC
TGAATC
>136340..136504_NC_011000
TCGTTCATTGCTGGCCGCCTCAGAGCAGGATTTCGGCGCACCAAAACAAAACTTGAGGGG
CACGTGGAGATTCGGCCCGATTTTTTATCCTTCTTTGTCATTTCGGACCGGCCCGGCACG
GCGCCGGATTAAACGGGGGCAGAAACCGGCCCGATGCGGCTCGGG
>139291..139424_NC_011000
GCGGGCGAGGCAGACGGTCATGGCGAGCGAGTAAAAATGCGAGTAAGAATCATTCGCATT
GCGTTAAAATGTCGCGCATCGTATCAAAACCCGCCAACCGCGCAACAATTCCATGCCGCG
TACAAGGGCATCCG
>146554..146750_NC_011000
AATCCGCCTTGTCGGAAAAATTCCTACACGGTTTACAGACAAATCCTGCACGGCATACGC
CACTTCGCTGATACCGGCATAAATGCCGTTTGGCCAAAATTCGCCCTGTAAGACTATAAA
CGAGCCGGCTGAGCGTCCAGTGCGCACTGACCGAGCAAACGTTTTCATAACATGAATGGC
CAAAGCAAAGCTCTTTA
>159157..159313_NC_011000
CCCCGTTCCCGCCGCCGTTTTCCCGTGTCCGCCGCCGATGCGCCGCGCGGCGTGATCCTG
CGTCGAATTGACACGTGGCCACACTCCGCAGGCATACTACGCGTCAGCAGCAACGTAGCG
TCATTCGTAAGATTCATAAGTTGGCAGGCGGCGGGGG
>179296..179450_NC_011000
ATGTGACGTGACGTCGCGGGTCAGTGATGAGTCCGTTTAGCAAACCGCGTCATGTGGTTT
TCGCAAAACTGTCGTCCGATGAAAGCTTTGATTTTTTCGAATGCAGGGTAATTGCGTATG
TTGATAGACCTAATACTCAATCGCTAATGTCCTAA
>203026..203161_NC_011000
ACATCGCATGTGTTTTCGGAATCCCGTGCGATGTCGCATAAACTATTGATTTTAAAGGAA
AAGTAGGTTTTCGTCAAATTTCGCCAGATGATCTGCGCAGATGCGATGCGTCATAGTGCG
CCGTTCACACTAGCGA
>253951..254114_NC_011000
CAGTTCGATCCTGCCCATGGGCACCAGCAGTAAAAAGTCAGTGTCTTGTTTGCGCGCGGC
GCAACCTGTGAAATTCCTTTTGGGAGTTGAAAATGGCCAAGGAAAAGTTTGAGCGGACCA
AGCCGCACGTGAACGTTGGTACGATTGGTCACGTTGACCACGGC
>257756..257891_NC_011000
GCCAACGCAGATGGCGAACCCGAAAAAGTTTTGCAGTGGTGAAGCCGCAGGCCGCGAAGC
GATTCGCGGCGTGAGGTGGAAATACTCCTAACGAGGTCGGACGCCGTTGTTGAACGAGGT
ACGTGAGGTTTCGGCC
>365924..366070_NC_011000
ATCCCCGACATTGCACCGTTCAACTATTAAGCACGCCCGCGTCAGGCGAGCATCCGGAAA
TTTTTTCCGCCATCCGGGCGGCGCGTCGTGCAACCAGTAAAACTTGCGCATTCCGGTCGA

AAGAAAAACCGCTGCGAGCCCGCAGCG
>400302..400447_NC_011000
TACGGCTAATCGGTCTGGTGGTCAACCGTATTCCCATGCTCTCAGGGTTATTGCGGAGAT
CGGGTTTACAAAGCTTGCAAGCGATGCGAATTTTTTGTAGGGGCCGATCGATTTTGGCGC
CACGTTCGAGGTGCAACCGTTAAACT
>452189..452342_NC_011000
ATATGATTTTCTGGAGGATTAAAATTAATTTTGAGCGCTATCTTTGTCATAATTTTCTGG
AAATAATTGATGTTGGCGTGTGGTGAACGTTTGGGATACTATTTTAGGTTCGGCATATGA
CGTCAAAAAACACGGGGCTGCTTATTTATTTCGA
>562335..562477_NC_011000
GGTATGAACGGCCTCGGATGGCACAACGGATCTTACCGAAACCCGTTGCGCGCCGCGCCC
GCCAGAATTGACAGGCCGCGTCAGGCTACAGGGGATTCATGACGGGCAAAAGTCGCGATA
AACCCTTGTGCCGCAAGGCGTGT
>567233..567372_NC_011000
CTTTTCGCGCTTCGATAAACCGTGAACAATGGGATCACGCGCAACGCTCCGCCGAAAGCG
GCGCGGATCGGGCATTTGCCGGCTCGCCAGCGGGCCGGGTTTCAAGATTCAAGAGTGGGG
AACCATCATGCATATCGGGT
>596436..596589_NC_011000
AACGCGGGGAAACGGTGTGTCGCCCCTTCGCGTTCGCACCCGAATTGTTTAGTCGTCGAA
CGAAACATGCCGCGAAATGGACGGAATTGTCGCTCGCGCGCGTGGTTGCATTTTTTACCC
ACACTAAAAATCGGCAAAAACCCGGCGGGCGTAT
>657042..657190_NC_011000
CGAGCCAATAACGGCGTCAAATGTTTTTTAATGGTCCTGCATCAGGGCTTTCCCGAGCAG
TGCGAACAAAAGATTCGCAAAAGGAATAATCGGCATGCCGTTCATTATTTCGAACGGCTT
TCAGGGTTTTCCCCGCTGATGTGCGGCAT
>812763..812898_NC_011000
TAATGCGTCAGCTTGTCGTCCCGGGGTCCCCGCAATACGGGGCGACGAGTCCGATGCAGG
CATGCATCGCATGCCATTCGCGTGAGGCGGCGATGGAAGCAGCGAGGGTTTTGGCGGAGA
CGGCGGACAGCGAACC
>919933..920067_NC_011000
GTGCCGGATCGCCGCGTCATGTCAATCCACCGAGTAATGAACGTGTCATCCGATGTAAAC
ATCTGGTAATTTTCTGCCGGATTATGCAAAATCGCCGCTCCCAAATACAACTATCGAGGG
ATTCCGGGATGCAAC
>1147412..1147552_NC_011000
TTGGTTACTTTTCTTTGGCAAGACAAAGAAAAGTGACCGCTGCCCCGCGCAGGGGCGACG
CTAATAGACCGACACGATCACAGGTGCTTCACGCGACCGGAAGCCGACCACAGAAAAGCC
GGAAAAAGCCCCAAAATACCG
>1246576..1246713_NC_011000
ATACGAGGTATACAACCGGCGGCACGCGTATCGACAACGAAACTGACGCGGGGCGAAACA
ACTCGGCCCGACAGTGCAAAAAACGCTCCAATCGACTTTCCTAGAATCGATGATTTTAGA
GAAATTCCGTCAATATTT
>1704707..1704879_NC_011000
GCAGTCGTGATTGTGTCGAACCGTAACCCACTATAGACGTTCCGGCGCATTTTCGCGATC
CGGCCGAATGGCCAAAGCCGTTTCGTTACGTCGCGTTACGACTCTCGCAGCCCCGTATCA
CCTCGGCGCGGGCGGCGCGCTTACATTGGTTGCGTGAGCCGGGCGGGCCGGCT
>3241596..3241765_NC_011000
CCAAAGCTTTACCAGTGACAGCAAGAAACCCGTGACAGCTCAAGGCTTCGCGGGTTTTTT
ACTTCCAGCAACGCCAAAGTTTCACCATCGACAGCCGGGGGTATTTGGGGGTATTTTCGG
GGGTATCAGACATACCCCCGTGCACAGATACCCCCATGCCACTGACCGAC
>3244006..3244151_NC_011000
TCCCTTGCCGGGTTGGCGCGGTTCCCACACAGGGGGCTTTGCGTGAAGGGACGCAATGGG
TATCAAAGAGGACCTTGCCGAGCGTAAGGCCGGCTTCGTCAGTTTGCGCGCGCTGCTGGA
AGACATATGCGCCCATGAGAACGTGA
>3244246..3244373_NC_011000

CGTTCACGGTATTACCGGATCGCGCTCGACGACCGCGTGGCGCGCGCTTCAACACATCGT
CCGCGAAGGAAAATTCGCAACATCATGGGAGCTCAATCCTGATGACATTCCGTTCTAGCG
AGTACGAC
>3245237..3245374_NC_011000
TACCTTGAGCGAAGCGCAGCGAAAGGCGGTGGAACTGGTCGCCTGCCCCGTTGACCGTAA
TCCTGCCGCCAAAAGACAACCCTGATGGGTGGGTGTGCGTACCCTGAGGGTCACAGTCAT
CAATGCAAGAACCATTAG
>3246856..3246992_NC_011000
GGCGTGGAAACCCGTGGACCATAGGCGGACGAACCGCCGCACACGCGGCTATTTTTTCGT
TTGCGCGCTTCTGCACGCCAGTTTTCTATGGCGGGCCGGGCGAGGGCCAGCTTCGGCTGG
GCCGGTTTCCTATGGGC
>3247147..3247344_NC_011000
CTGTCCTCGAGATCGTGTCGGCAACTCTGCGCGCCGCCGCGCTCGCGCCGACCGTCGATG
CTGCGCTCGACACCTGCGGCGACGCCCTCGCCCGCGTGGCTGACCTTGCTCGTGCGGAGG
TGCGCCATGGCTGACAAGATCACGACGCTCGATGACGCCCTGAGCGACAGTATCGATAGC
GCCAGCTGTTCCAAGGAA
>3247579..3247713_NC_011000
AATCACCATCGTTGAGAGGCTTCGCTCGGACACTCGGTGGAACGTCACGTTGCTCGGCGA
TCCCGCCTTTGGTTTCGCGCGCAACGGAGCGCCCGTCATCACCAACAATTTCGCGTTCCG
CGACTCGTCGCTGGT
>3247833..3248013_NC_011000
GATTCGCTGCTGGCCGACCCGCTACCGGTCGCCTCCAAGTTGGCAGCCGAGACGGGCGCA
TTGTTTCCCTTTCGTTGCGGGACACGCTGGCGTGCGCCGACCGCGTCGACACATTGGGAT
TGGTCCGTATCGTGTCAGCGCTGGAACACAAAAGCGCTAACGGCTCCCTTCTCGTCGACG
A
>3252472..3252600_NC_011000
TACGAGTTTGGCCTAGGCATCAAGGATCAGCCACAGAATCGATTCGATGCGCTGCATAAC
AAACTAAACAATCCGACCGGAGAATACGGCTTCTATCGCGTCATGATTGCAGAACACCGG
GGTGACACC
>3253014..3253166_NC_011000
GTTTTGAATATCGATCACGTCACCGGCTTTCGCTTGACCAGTCGGCGCCACAAACACGTC
ATTGGACGATTGCTTTGGCGGTGCGGCAATAGCCAAGGTCGTTGCAAACAGTGCGGCGAC
GGCAACAATTTTTCTCATCATTTAGGTGCTCAA
>3253231..3253381_NC_011000
GCGTGAAACTTGGTCTGCGCTGCCTGGGGTAATCGAATCCTTCAACCCGGATGCGATGAC
GTGTACGGTTCAGCCGGCAATCAAATTGCGGTCGCGTGACGAGCAAGGTCAACTCAGCAG
CCTTATGCTTCCGCTGCTGGTCGATTGCCCC
>3253510..3253673_NC_011000
CTCGACGCGCGTTCACGATCTGTCTGACGGCATGGTGTTCGTCGGCTTCCGATCCCGGCC
GCGAGTGCCGGGTGCAATCAGCACGACGACCACTCAGCTCCGCACCGACGACGGTTCGAT
GCACGTCGAGCTCGACCCGGCCGGCGGCGTGGTGTCGCTGAAGG
>3253802..3253937_NC_011000
GCGAACGGCAAGCGCGTCGACGACACGCACACGCACCCCGACGCCCAGGGCGGCAACACC
GGCCCGGTCAACTAATGGAGGCATGATGATCGACGGACTGATTAGCGGGAAACTCTACGG
CACGGCAGCCGAGCGC
>3254478..3254652_NC_011000
TCGCCCGCAAGGGGGTAGACATGGCCTGCATTGAACCCGAGGCAGTTGAGCCGCAGCGCC
TCGTCCCGCTCGACACGTGGATTTACGAGAACCGCTACATGCAACGGAAGCGGGAGCCGT
GACGAGCAAGGGCAATGCATGGCGACTGATCGAACCGGGTTGCACGGAAGCATTG
>3297325..3297457_NC_011000
ACTCGATAAGATCAAAAAAGCTATTTTTATAGCCGTTACCCGAAATCCCGAAACATTGCG
CCAGCACCCGCACCAGCACGCCGCCCGAGGACGGAGTGCTGACTTTATAAATTGGACAAT
CCGGGTGCGTCAA
>3375805..3375977_NC_011000

GCACTCACGGTCCGCACCAAGAAAAAAGGACACGCCCTCCGTCACCGGGAAAGCGTGTCC
TAAAACCAGTGCCAGGGAGTGCGCCGTCAGCGCCGGCACCGGTCAACCTTGCGTCGCGCG
AGCCGATCGCTCAGCTACCCTGCGCCACTTGCCAGGCCTCGTACAGATCGGGC
>3456900..3457078_NC_011000
TACCCCACCCATGTTTTTGCAACCGCACAATGGAATCAATTCCATTGCAAACAACCATTC
AACGGTTCGCGGCAATATGAATTCATTTCACGAATTCGCACGCGCACGAAATCGATGACG
AAGAATCACCCCGTGAAGCCCGCCCACCGGGGCGCCTTCACTTCGAACGGACGAACGAT
>3474261..3474410_NC_011000
GTACCCGCGATTTCGATTTTACGCAAGCCGCCCGTTTTCTCAGGTTGACGGATATCGGTT
CCGCATTTAACAATTGTCTGCGGAATCGAAAAATCTCATTTTCGATACGAGAGCGAACTG
GGGCCGCGCGATGCGTTTTCGACGTTGCGT
>3496591..3496745_NC_011000
GCCTATGAACTTGCCGTTCTGATTCGATACTAAGCAAGGTTCGGCGTCAGGCATGGCCGG
CGCGCCGCGCCGCACGCACGCCGGACACCCTTGCCGGGCACCCCCGAAGGGACGTTACTC
AAACCTCTATGAACACTTCTCAAGACCTGTCGATC
>3504662..3504804_NC_011000
TTACGTGAGCGTGAAGCCGGCCTGATCGGTGAATGCCACGTTGAAGATCAACGTTTAACA
CTGGGTGAACGTCGAAAGGTTGCGCAGGAATCGAGTGTGATGACAAGCCAGGTAGGACCG
GGACTCGCTAAACCTGAATCGTG
>35295**32**..3529665_NC_011000
TATTGACACATCAGATCGAGATTTCTCGACAAATAAGAAAGAAAACAAAACAAACTCCGG
CGCCGACGTTGCAATTAAGCCACGTCGATACGAAATGCACGCCGCACTCCTTGACAGACG
CAAATTGACCGGAT
>3532227..3532381_NC_011000
TGACAATCTCATCCTCCTCGGCCAAAACCCCTTTCCACAAGGTATTATCAAGATCGGTGA
CAATCAGTTTTATCGGATTTTTTGATTTCAAAATCCTTAGCGCCTGATCAACGCGCGAAA
TCAATGCCAGATAAAACTCCGAGCTATCGTTGGCA
>3671381..3671556_NC_011000
TTTATCGATAATCGACGCCCATTCGAATCAGTAGACTCCGCCCATCCTCGGTGTAACTAA
ACACGATTGGCGCGACACGCAGCCATAGCCGCCGCACGCCCCTTCGGATTATTTCCCCAC
CGCGGAATGCGAATTCGCCGAATGGCGGCGAATGCCACGTCAAAAGAAGAAAGAGG
>3718958..3719122_NC_011000
CTGTTAAATCAAGTAGTTGACGACCAACAAAGGCGTTAGCGTAGCGTTATTGGCCGCGCG
AATCAATGTCCGTTTGGGAATGGAAAGCCTCGCGATGCCGCGAATTCCGCATTTTTTACC
GAAATCGCACTACATTTATCTTACATAAGATAGAATGCGTGCGGC
>3737727..3737897_NC_011000
CGCGAGCCCGTCAGCTTTGAAAAATTTTGAAATATTTGGCGTGCAAGCGGTGCAGCGCAC
ACCTTTCAGGAATGCTTCGATTCGCGAAACGTTGCATTTCACGGCGCTTTCAGCGACTTT
CGCGTTTACCCGCGGCGGCTAAAAAAGCACGGATTTCGTGCGACGCCGGTC
>3820929..3821084_NC_011000
GTGAGCCCGTTCGCCGAGCCTTTGCATCCTGGATGCTGGCGATTCTACAGACATGACGGC
TCACGTATCGCTCGAGCACTTCACAGCGTTCCGCTTCGTTCATCCAGCCGCTGCACGCAG
TGCATGTCCTGAGCGGAGCACAGTCGCCGGACTGTG
>7278..7422_NC_011001
CCCCAACCCCTGAACACCGACCGGCAACCGGATCCGGTGTCATGTGAAACAAAGGCCGAA
ATTCGCGGTTCGCTGGAAAATGCTTGCAAGACAAGGACTTGGCTGACTTTCTTCGCTTTG
TTTCACGGAATTCCGAGCTTCCTTC
>24329..24498_NC_011001
GTTGCCGTTGTGCGATGCCGCATCTGGAATTAATTGTTGCCGATTTCGAATAATTTGCGG
CGTAAATATTGAGATGAATGAGACAAACGGCCTGGACGGCGTGCCGGCCGGGCCATGGTC
GCACGCATCGATGACGCCATCGTGTGAAAACGGTGAGTCTTTACGGGAGA
>90660..90811_NC_011001
CGTTCGGCATTGAATTGAGCAAATAGATTCGGAATCCGAATCTATGGAGCGGGAATTTCG
CTTGACGAAGCAGCACCCGACGTTTTTGGCGTCGCGCGTCACACGCGCGGGAACAGCCCG

CAAGTCATGCGTGGCGCACGGCGATGAAGACA
>432522..432657_NC_011001
CGCAAACTTCAAACACGCTAAGGGCCTTGACATGCCGTTTTTCTCCATTATTATTGCTCA
CCACGTCACTATTGTTTGTTGCGTCATGTTTGGAGATCGATTCCGGCCCTGGTCGGACAG
CGAAAAGCGGCAGCAA
>506891..507029_NC_011001
TGGATGACGCAGCGCACGGCGATGCTGCGATGGAAGCGTGATTGAAGCGAATCGAATCAT
CGAATGACGCGCGGCGCACGCCGCGGCATTGCATTAATCGATAGGTGATAATAAAAATGA
AAGAATGATCCTGTCGATT
>1018568..1018719_NC_011001
CGTGCGAACGAATTATTCGACGGCGCGTACTTCGGTGCAGCGGAAGCGAAAAAACGTGGA
GCGGAATCGTTTGCGTGGCAGGCAATCCGTCGGATTGCTTGCATCGGACGGTATCAAGAA
CGAAAGCGCGATGTCAAACTTTTTTTGATTAT
>1173919..1174053_NC_011001
ATCATGGGTGACATGCTGCAGGGTTCGCAGTGGACGCCGCTTTCGCCCGGTGATCGTGAT
CGCCCGACGCACATCGAGGTGAAGCAGGACGTAAGGATCGAGGTGCGGGGCAGCGCCGAC
CCTGATGCGACGGCC
>2223385..2223519_NC_011001
CCTGGGCCGCCTCGCGACCACCCCGTCACATTCGTCAACAATTCCCCCACAAACGCGAAC
AAGAACGCTTCTCATTTAACAGGAAATTACATAGAATGCCGCCTGAAAACAAATCGTAAT
AATTCCCGCCGGCAT
>252177 2..2521943_NC_011001
AGGGAAAACACGTACGTTTTTGAGATGGGTGGGTGGCGATGGGGGCGCAACCGCGACGCG
GCGACAACGTTTGCTTCGGGTGCAATCGGGTTGCTATTTTGGCAAGAAGGTGGGAGCGGG
TTCCTTTCTGGGCTTGGCGTTCGGCGACGGGGATTTCCCGGTGATGCATTTT
>2720007..2720185_NC_011001
GTACCACCCCGATAACGATTCAGCCCGCTGCGTGATTTGCATTGGCCGAAAGATTGGTTC
ACGCCGCCGGCAGGTTTTCGTACGCTGAAAGTCCGAACTTTTTCCGCGCACGCCATGAAC
CGTACGATCCGCTACAAGGGCTATGAAGTCGCTCCCGCTGCCGCCCGGTTGCCGAACGG
>2739339..2739504_NC_011001
CCGACAGTGCGACAGGTTACACAACGCGAATTACAAAATCCCAAAATCCGGATCGTTTAT
CTTTTTTGCCTAAACAGTTTTATCCGAATCGGCTTTACCCGAATTTGCTCGTCGCGCAAA
TTGGGGCGCCGCCAATTGATCAGATAAATACCGTGCCTCGAGGTGC
>2790921..2791083_NC_011001
CGCCCGGCACGGCATCGATGTGCCAAAGTGGAAGTGTCAGCCACCACTTAACCAGAGGGA
ACGCTCATCATGAAGGTTACGACCGTTGGTATCGATCTGGCAAAGAACGTCTTTCAAATC
CACGGCGTCGACGCGCACGGCAAAGCCGTCCTCAAGAAGCAGA
>2912664..2912801_NC_011001
AATAATCCGGCGATAATCCCGACACGCGTTTATCAACTTGTCACAACGCGCGGCCCGATG
CGCCAATGCCAGCATCCGTTGCGCGCGGCGAAACGATTTGACTTATCCTGCTCCATTCCA
CCTGAATTCGAATTCCAT
>3025995..3026135_NC_011001
TTTCTGTTCTCTGTCAAACGTAACCAATCGCAAATGACATCACGCGGTCGCAATCCGCGG
CCCTCGCGTGCGCCGGCCCGTGCGGCGGGATTTTTTTTCGGGTTGGCCCGCGACTATTGG
CGTGCGATAAAGGGTTTTCCC
>3069020..3069154_NC_011001
AAGCGCTTCGCCCGATCGTGGTCTCCTCAGCCGGTCATTTCGGGATCAATAACCGCTCTT
CTTCTTGCACGCGAAGAGAAGATTTCGCACGGCGTCATCCGTCGATTGCTGGATCACGTC
GAAATCTCCGTTGCA
>3070082..3070253_NC_011001
TCGAAGCTCTTACCACGATTCAGGTTTAGCGAGTCCCGGTCCTACCTGGCTTGTCATCAC
ACTCGATTCCTGCGCAACCTTTCGACGTTCACCCAGTGTTAAACGTTGATCTTCAACGTG
GCATTCACCGATCAGGCCGGCTTCACGCTCACGTAACCCTTTTGGTACGGCC
>3136490..3136633_NC_011001

TGTTCGCCTCCTGAATGGACCGGGCCTGCCACGCCGTCCGCGGCCGTTTCCCCGGTCACA
CGTGACACACGATACACCGAATGACACTTTGTGTCATGCGTTACAATCGGTTTTTCGAAG
GACACCTGGCGCCCCATGACGATT
>4443..4616_NC_011002
TGCTCTACCATTGAATTCCCTTGAGCCTCCAATCGAGTACGGGGCGCACCGATATTTAGC
GCCAGTTGTCACGTGACAACGCACACGTCGGTCGCGCGATTCATGGAATGGTGGGATTCG
GAGTAGATAGAAGTTGACGCCGGGCATTCGTGGGCCGGTAATCGCAGCGCGACC
>11987..12129_NC_011002
TTTCGACCGTGCTGATGGAACGTAAATAATACGCATCATCGATTAATCCGGTGCGTCAGT
CAGATAATTCGTTTTACGAAATGATGGCGTGAAAAACGTCGATCGACGCGCGCGCGGTCC
CGTGTGCCCATCCCGACAAAACG
>36502..36690_NC_011002
GCGTACCAGACGCATACTCGGCCAGCCTTCCAGATGGTTACTGACATGTTGCGCAAGGCT
TTGCGTCGCCTGCAACCCGACGAAAGACCCATGTTGCACTCTGATCAAGGCTGGCATTAT
CAAATGCCCGGCTGGCGGCTCATGCTCGAGCAACGCTCGTTGGCGCAAAGCATGTCGCGT
AAAGGCAAT
>88283..88413_NC_011002
CGTGCATCGCATGGCATGTGCGTCGTCGATGCCGATCGCGTCAATGCGGATCTGCTCGCC
TTTTTGCAATCGTGACGCGGGAACTTGCGCGACTCACACACTTCGAGCGGGACGAACGGA
ACCTGGAGAGC
>122093..122278_NC_011002
CGGCGCACGCCGGATGCCTGTAATTAGGGGATTTCCCGCGTTTGCTTCGTTTTTATTGAT
CGCTGCTTGTCGCGCTGTTTTTCTTCTTGCATGCTTGTATAGACAAATTCCTGAGGCGGC
TTGCTGCACCGAAGGGATGGCGAACTTCAACGCAAGGAGAAATCTCGTGAATGGCAAGCG
CATCAA
>130355..130492_NC_011002
TGCATGCGTACGCGGCGTGCCGCCCGTTGCCGGCTTGGTTCAAGCTGCCAGTGACGCACG
ACGTCGTTTAGTGCGCATCGCGCAAATCACGCCGGCTCAGGTGGACGATTCACCTATTTG
ACAGCCCCGATCGGCGGA
>135581..135723_NC_011002
TCGTGATCGCGATCATGTCGTTCGCGTTCTTTGCGCAGGGCGTCGCGTCGGCATCGTGGT
CACTGATCGGCGACATCGCGCCGCGCAGCATGCTCGGGGTCACGGGCGGCGCGGTCAACT
TCGTCGGCAACCTGTCTGGTATC
>138427..138569_NC_011002
TTCGCCAACGGCTGCCGTTTCAGGCTGCGGTCATCGCGATTGCCGCGGGGCGGCATCGGC
TGCATGGCGCGGACTGAAAGCTTGGCATTCAGCGAAATCGCCGAAACATGACGCTCTCGT
GAATCATCGGATAAATCGCTGTA
>149887..150045_NC_011002
GTCCGGAAATATCCCTCGATTACGTCCGCTTTTCTCATTCGTCGTCGATTTTTTGGCGCA
CCACACTTTCTTCCAGTCGTGCCCCGCCGGCAGCGCCGGACGCACGCGTGTTGAAGTCAT
GGACGCCACCCACGTGAGGACAGCATGGACAACGACAGG
>204538..204677_NC_011002
GACTGGTCTTTGCCGAATTCAATTGCACGCGACGCGGGGTTTCCCGGGAACAAGCCGTCG
CGATCGACATTGCAGGCGACGAACAGAGTTGGCGCAAAGATCGACATCGTCACCATTGCG
TGCAAACCTGAGTGTTGCAA
>205769..205901_NC_011002
CTTCAGGTTGTGTCCATACATGATAGCGGTTTCAACCGGAAATGATAAAGAAAATTGCGC
CGACTACGAATTCCATTGCCGATATGGCAATGCACGACTGCTTCGAACACCTGAGAAGTG
GTTTGACTTGAAT
>209614..209811_NC_011002
GCCGCCACTTGCGACAACCTTTCGTTCCGCGCCACCCGCGCCACCCGCGCCACTCGCGCA
GGCCGCCATCCGCCGTCTGCGTCAACGCCGATTCCACGGCCGCCGTGTCAAAAACGACCA
TTACGTCAGACGAGAAACGGTAACAAACGTGGTATTTCGATCAAACCCGCATCGGCAATT
TGAACAATGAGGCCGATT

>333449..333583_NC_011002
CGGCCGTTTGATGTTGTGGTGCCTGCGGATCGAATCGCTCTGAAATGGCGTTGTCCGGTG
AATTCATTGTTGAAGAAAAAATCGCGTGATGCAAACCACGCACGGCCGGCCGGCATTTAA
TGTGTTCAGCGATTC
>357264..357459_NC_011002
GCAGATCCGGTAAAAAATCGGCATCCGGGAATTTGAGTTTTTTGCCTTTCTTAATAATTT
GCGTGATTAATTTAATAAGAATCGCTTTCATCCGGAATGAAAAGTAGTTAATGAAGATGA
AAGAGTGTTGAGTTTTGCCAAAAAAATGTCGGCCCCATGGAAAACGTTTGCGAGTCGGGC
CTATATTTGACTCGCG
>370809..370953_NC_011002
TCAACGCTTTGCCATATCGGGCAATCCGCGAACGATTCGCGCCCCCTTCGCGGCGGCAGG
GTCCACGCGTCATCGCTCGAGCTGTTCCGCGGTTGTCCAACCGAAATGAGGTTTCTCATG
AAAACGAAAGCGAAAGCGAACGTGA
>390461..390594_NC_011002
TTTGGCAAAGTATTTATCCGCTCGATAAAAGATTTTGAATGACGGCGCCAGCGCGAAATC
GTTTGCGGAATGTGTGTATTTTGCGTAAGCCTTTGAATAATATAAATTTTACAATCTATA
AAGAATGCCGGCTT
>400289..400485_NC_011002
AAATCCGCGATTGCACATGCACGCTGCTGCTAAGCGCCTCTGCCGCGCAACTTTATGACG
CACGCCGATTAATTTATCGGTTTATATCGCTTGGGATTGCCGGATCGAGGGAACAGAATT
TCCGTTGCCGGAACCCGGCTGCAGCGGGGCCGCATGCGCTCGACGCCGCGACGGTCGCGA
AGGCGCCTGCCCGGTTC
>422485..422621_NC_011002
TCGGGCACGGCGGGCGGCCTCGCGCCGACATCTGTGTTCCAAACAGTACATTTGCGCGGC
AATTTAATGCAAATGAGAGGCATTTGCAATCGTAATTTTGAGATTGGGAGGGCGAAGCGA
AATGGCGGGCCGCACGG
>466699..466882_NC_011002
CGTTTGGGACGCCGGCTGCAATGGCGCCGGCGCGGGTCAAATTCTCGTTAACGGCACGGT
TAACAAAAAATCGAGGCCACGCGTGGTGGCCCGCCGACGATGGCGGCATCGCACGGCGAC
GCAAACGTATGCGGCCTCGCAGCGAGATGCCCGATCGCCATCCGGACCTGCGCAGCGTTT
GCCT
>500495..500632_NC_011002
TACGAAGCACGGAGAATGTCATGTCGGAGTCCCAATCCGTCGACACGATCGGCGCGCTCG
CCCATTTCATTCGCGCGGCGCGAATCCAGCAGGGGTTCACCCGGGACGAACTCGCGAACG
CGACCGGGCTGTCGCCGA
>629782..629939_NC_011002
ACCACTCCAGCAAAAACCGGGCCATTGCCGTCAAACGCTTGCCACGCCGGCACGAGACGG
AAATTTCAGCGAACCGCTCGGCGTGCCGAATTCCTCATTTCGGTGCACGAACCGGCGCCA
TCGGTAACGGTTGCCCTCGCAGCGCGCGCCGGCGCACC
>682721..682913_NC_011002
CATGACAGCCGCTGTTCTTTCATGCAAACCTCGTTTGCGAAACTGGCCAATTCCCGTTCA
TTCTCGCAACGCGTACCGTATCAGGCAAATCCCGGCGCAAGCCGCCTTCACCGGCACACC
ATGAGGAACGCCGCCCATGCACACGCTGATCGTTCATTGCCATCCCGAACCCCGCTCGTT
CAACGCCGCGCTA
>687343..687498_NC_011002
GAGTAAACGGCGATACATGGGTATTTTCCTCACGTTTCATTGGCGCCGTATTGCTTCGCG
CGCTCGCCCGCGAGCATGCCCCCGGCCGGTGCGGGGCGCGGCGAACATCGTCCGCGGGCC
CGCGTGCCAAGGCTCTCTCAGGTTCTCTTTTCGGCT
>708654..708802_NC_011002
CGCGGCACAATGGGCTACGGTTGCTCCACATGTCGCTTCCCGATTTCGCGCACGCGAAAA
TCACCGTCCGCCCTTCTATCGCACCGTGCCGTTCTTGCAGTGTCCACGTCACATGCCTTT
GCTCGCATTGCACAATGCGGGGCGGCCGC
>710870..711047_NC_011002
CGACGTTTATCTCGCGCGTGACGATGCGGCGAGCACAGTCTCGATCCGCCTATCTGGCTC

CATTTCGCGTGACGATGCAAAAGTGGTCGACAGACGGATCCAGAAGTTTGCACACGATCA
TTGCAGCGCCGGCGCCATCCTGTGGCGCGGATGGAACAATTCGTGGAAATGGTTCGTG
>712628..712787_NC_011002
CATGCAGGTCGAAGCCCGCTGGCCCCAATGGAATCCGCTGACTCATGCCTCATTTTTGCC
ACGGACCGTGATGTCCAGAAGACGATGCAGGTTTGATGAGTCCCGGTCTTACCTGTTTGC
CATTACAACACTACCTTGCGCAACCTTGGACTACAGATCG
>713043..713195_NC_011002
GGCAGGAAGTGATGGAAGTGTTGACGGGCCCTGAGCGTCGGCGGCGCTGGACGGCGGAGC
AGAAGCTGGCGATGGTTCGCGAGAGTTTCGAGCCGGGGAAATCGGTTTCGATGGTCGCGC
GCCAGCACGGCGTGAATCCGAACCAACTGTTCC
>714169..714365_NC_011002
TCGGCATCAGCAAGAGAGGCGATCGATACCTGCGCACGCTGATGATTCATGGCGCACGGT
CCGTCATGAAGCACGCAAAAGACCCGGGTTCCTGGGCAGTGCAGTTGAGCATGCGGCGTC
CGCTCAATGTCGTAGTCGTTGCACTGGCCAACAAGATCGCGCGGACGATCTGGGCTCTGC
TAGCTCATGGCCGCTCA
>716403..716546_NC_011002
CATCGAAGAAGTAACGACTGGCTCAGTTTTACTTTGCGCGATCAGGCGCGAAGTGGCTCA
GATTTCAAATGCGTTTGACAACATCACGGTCCGTGGCAAAAATGAGGCATGAGTCAGCGG
ATTCCATTGGGGCCAGCGGGCTTC
>748152..748291_NC_011002
ACGTTCGACCCCGCTTTTCTCGTCAAACTCGCTTGGATGTGCGCGGCGAAATTTGGCCGT
TATTTTCGTCTTGCCAATATTTGAATCACCGTGCAACAGCAGGCACGGCATACGTTCGCG
GCGGGGCGTCTCAAAGAGGC
>748460..748598_NC_011002
GAAGGCCTGAACTGACGTCGCGTAGTCGACCGCCTCAGCCTCCTGCTCCTTCACGTCATG
TCTGTACGGGACAATCCTGTCGCGCAACTCTTCCAATGGGACGATCTTTTTGCGTGACCG
TCGACGTGCACGGGCCGTT
>751453..751614_NC_011002
TCGTCAGACGCGTTAGCGGAATTAAAGGACACGTATAACGCGTTAGCAAAATAAACGGAC
ATCCGATCGTGCGCGCGTTGCCAAAACGCAACGCATGAAAAGACGCGCGAGACGGCAGCA
ATGCCTTATCTCGCGGGGCCGTGGCCGTGTCTTACATTAGCG
>754523..754670_NC_011002
CGCCACCATCGCCACGATCGTCACGGTGTTCTTTGCCGGTTATTGCATCGAACGTAATCT
GCCCGAACTGGCGGAAAGCGACGCGGTCGGCTCTTTTATGCAAGCGCTGCGCGTTCTCTG
AAGCACCATCGCTCGGTTGCTAGTCGAT
>784816..784973_NC_011002
CCAGCGCTATGTGCCAGTGACGTCAATAGTCCTATGACGACCTCAGGTTCCGATGACTGC
TGGCGTCAAGTTGCGCATACGCTCGCCGGTGAACGCCGCGTAAAGGTCCGTTGTCTATGC
>810233..810398_NC_011002
ATCCTGTCCGATTGGACAGGTTTTAGACGATCAAAATGACCAGCGCAACGTCTTTCTTGC
GCGACGCGATGGAGCGAGAAACGTGCCACCGCAACGCAGCATCGCGGCGGAATCGCGGCA
TCCGGATGCGTGCGCGCGACGCGAATCAGCAGCTAGAATGAAGCCC
>869202..869370_NC_011002
GAAGGAGGCGCGGCCGGAGCGCGCGCATGCGATGAAGCAGGTCGAACCGATTCTAAAAGG
GGGCCGCTGCGCGCCGAACCAACGATGTTCGATATCGAAAGCAAGCGTTTTCATAAGGGC
AATGCCGATGGCACGGGCAT

C. Extracted sequences of only intergenic regions with threshold of p value>0.0001

>34503..34658_NC_011000
TTCCACATCCTGATTGTTCTGTTGCAGGCATTCATTTTCATGATGCTGACGCTGGTGTAT
CTCGGCCAGGCGCACGACAAGCACTAAGCGCGGCGTGCAAGAAAGAGTTTCCGTTTTAGT
TTTTTAAATCTCAGTTCCAAGTCTTTTCACAAAGGA
>61713..61898_NC_011000
GGACGATCTGGGCTGTACTGGCCCATGACCGGCCGTACCAAAAGGGTTACGTGAGCGTGA
AGCCGGCCTGATCGGTGAATGCCACGTTGAAGATCAACGTTTAACACTGGGTGAACGTCG
AAAGGTTGCGCAGGAATCGAGTGTGATGACAAGCCAGGTAGGACCGGGACTCGCTAAACC
TGAATC
>253951..254114_NC_011000
CAGTTCGATCCTGCCCATGGGCACCAGCAGTAAAAAGTCAGTGTCTTGTTTGCGCGCGGC
GCAACCTGTGAAATTCCTTTTGGGAGTTGAAAATGGCCAAGGAAAAGTTTGAGCGGACCA
AGCCGCACGTGAACGTTGGTACGATTGGTCACGTTGACCACGGC
>365924..366070_NC_011000
ATCCCCGACATTGCACCGTTCAACTATTAAGCACGCCCGCGTCAGGCGAGCATCCGGAAA
TTTTTTCCGCCATCCGGGCGGCGCGTCGTGCAACCAGTAAAACTTGCGCATTCCGGTCGA
AAGAAAAACCGCTGCGAGCCCGCAGCG
>596436..596589_NC_011000
AACGCGGGGAAACGGTGTGTCGCCCCTTCGCGTTCGCACCCGAATTGTTTAGTCGTCGAA
CGAAACATGCCGCGAAATGGACGGAATTGTCGCTCGCGCGCGTGGTTGCATTTTTTACCC
ACACTAAAAATCGGCAAAAACCCGGCGGGCGTAT
>812763..812898_NC_011000
TAATGCGTCAGCTTGTCGTCCCGGGGTCCCCGCAATACGGGGCGACGAGTCCGATGCAGG
CATGCATCGCATGCCATTCGCGTGAGGCGGCGATGGAAGCAGCGAGGGTTTTGGCGGAGA
CGGCGGACAGCGAACC
>1758173..1758326_NC_011000
CATTGGAATTCAGTCGCACCAATTTCAATCGTGAGATGCTGCGAAATATGACAGCTTGGG
AGTTACGCGCTACGCTCGCATTCGCGATGACAAAATGGCAATCGTATTGGGAAGCGCGAA
ATGCTTAAACTTGCCCGGACGATATCGAGCTGCC
>3241596..3241765_NC_011000
CCAAAGCTTTACCAGTGACAGCAAGAAACCCGTGACAGCTCAAGGCTTCGCGGGTTTTTT
ACTTCCAGCAACGCCAAAGTTTCACCATCGACAGCCGGGGGTATTTGGGGGTATTTTCGG
GGGTATCAGACATACCCCCGTGCACAGATACCCCCATGCCACTGACCGAC
>3244006..3244151_NC_011000
TCCCTTGCCGGGTTGGCGCGGTTCCCACACAGGGGGCTTTGCGTGAAGGGACGCAATGGG
TATCAAAGAGGACCTTGCCGAGCGTAAGGCCGGCTTCGTCAGTTTGCGCGCGCTGCTGGA
AGACATATGCGCCCATGAGAACGTGA
>3245237..3245374_NC_011000
TACCTTGAGCGAAGCGCAGCGAAAGGCGGTGGAACTGGTCGCCTGCCCCGTTGACCGTAA
TCCTGCCGCCAAAAGACAACCCTGATGGGTGGGTGTGCGTACCCTGAGGGTCACAGTCAT
CAATGCAAGAACCATTAG
>3245405..3245582_NC_011000
ACAACCCTGAGGGTTCGGCGATCCGCCGACCCTGAACCTTCAGGGTATTGCCCTGAGGGT
AGCGTCGTCGCTTTCAATGCAGCCATTGCAAACAACGCCCAACCGGGCAAAGGAACTGCA
ATGGCAAGCAATGCCGCAACCGCGCAGCCAGCCGGCGGCGCACCTCTTCCCGTTCCCT
>3246856..3246992_NC_011000
GGCGTGGAAACCCGTGGACCATAGGCGGACGAACCGCCGCACACGCGGCTATTTTTTCGT
TTGCGCGCTTCTGCACGCCAGTTTTCTATGGCGGGCCGGGCGAGGGCCAGCTTCGGCTGG
GCCGGTTTCCTATGGGC
>3254478..3254652_NC_011000
TCGCCCGCAAGGGGGTAGACATGGCCTGCATTGAACCCGAGGCAGTTGAGCCGCAGCGCC
TCGTCCCGCTCGACACGTGGATTTACGAGAACCGCTACATGCAACGGAAGCGGGAGCCGT

GACGAGCAAGGGCAATGCATGGCGACTGATCGAACCGGGTTGCACGGAAGCATTG
>3666829..3666990_NC_011000
CCGGCAGGGAAATCATCGGCCGGCGCCCCGCTCGCGCGGCGCCGGCATGGCGCGTCGTGC
AGCGCATCAGCCGATCGTTGCACCGCGGGCTACAATGCCGGATCGAATCCGAATCAGGCA
ACGCTGGAGTGCAAGCATGGAATGCAAAGTAAGCTGGATGGG
>3737727..3737897_NC_011000
CGCGAGCCCGTCAGCTTTGAAAAATTTTGAAATATTTGGCGTGCAAGCGGTGCAGCGCAC
ACCTTTCAGGAATGCTTCGATTCGCGAAACGTTGCATTTCACGGCGCTTTCAGCGACTTT
CGCGTTTACCCGCGGCGGCTAAAAAAGCACGGATTTCGTGCGACGCCGGTC
>7278..7422_NC_011001
CCCCAACCCCTGAACACCGACCGGCAACCGGATCCGGTGTCATGTGAAACAAAGGCCGAA
ATTCGCGGTTCGCTGGAAAATGCTTGCAAGACAAGGACTTGGCTGACTTTCTTCGCTTTG
TTTCACGGAATTCCGAGCTTCCTTC
>432522..432657_NC_011001
CGCAAACTTCAAACACGCTAAGGGCCTTGACATGCCGTTTTTCTCCATTATTATTGCTCA
CCACGTCACTATTGTTTGTTGCGTCATGTTTGGAGATCGATTCCGGCCCTGGTCGGACAG
CGAAAAGCGGCAGCAA
>2720007..2720185_NC_011001
GTACCACCCCGATAACGATTCAGCCCGCTGCGTGATTTGCATTGGCCGAAAGATTGGTTC
ACGCCGCCGGCAGGTTTTCGTACGCTGAAAGTCCGAACTTTTTCCGCGCACGCCATGAAC
CGTACGATCCGCTACAAGGGCTATGAAGTCGCTCCCGCTGCCGCCCGGTTGCCGAACGG
>2790921..2791083_NC_011001
CGCCCGGCACGGCATCGATGTGCCAAAGTGGAAGTGTCAGCCACCACTTAACCAGAGGGA
ACGCTCATCATGAAGGTTACGACCGTTGGTATCGATCTGGCAAAGAACGTCTTTCAAATC
CACGGCGTCGACGCGCACGGCAAAGCCGTCCTCAAGAAGCAGA
>3025995..3026135_NC_011001
TTTCTGTTCTCTGTCAAACGTAACCAATCGCAAATGACATCACGCGGTCGCAATCCGCGG
CCCTCGCGTGCGCCGGCCCGTGCGGCGGGATTTTTTTTCGGGTTGGCCCGCGACTATTGG
CGTGCGATAAAGGGTTTTCCC
>11987..12129_NC_011002
TTTCGACCGTGCTGATGGAACGTAAATAATACGCATCATCGATTAATCCGGTGCGTCAGT
CAGATAATTCGTTTTACGAAATGATGGCGTGAAAAACGTCGATCGACGCGCGCGCGGTCC
CGTGTGCCCATCCCGACAAAACG
>88283..88413_NC_011002
CGTGCATCGCATGGCATGTGCGTCGTCGATGCCGATCGCGTCAATGCGGATCTGCTCGCC
TTTTTGCAATCGTGACGCGGGAACTTGCGCGACTCACACACTTCGAGCGGGACGAACGGA
ACCTGGAGAGC
>122093..122278_NC_011002
CGGCGCACGCCGGATGCCTGTAATTAGGGGATTTCCCGCGTTTGCTTCGTTTTTATTGAT
CGCTGCTTGTCGCGCTGTTTTTCTTCTTGCATGCTTGTATAGACAAATTCCTGAGGCGGC
TTGCTGCACCGAAGGGATGGCGAACTTCAACGCAAGGAGAAATCTCGTGAATGGCAAGCG
CATCAA
>130355..130492_NC_011002
TGCATGCGTACGCGGCGTGCCGCCCGTTGCCGGCTTGGTTCAAGCTGCCAGTGACGCACG
ACGTCGTTTAGTGCGCATCGCGCAAATCACGCCGGCTCAGGTGGACGATTCACCTATTTG
ACAGCCCCGATCGGCGGA
>204538..204677_NC_011002
GACTGGTCTTTGCCGAATTCAATTGCACGCGACGCGGGGTTTCCCGGGAACAAGCCGTCG
CGATCGACATTGCAGGCGACGAACAGAGTTGGCGCAAAGATCGACATCGTCACCATTGCG
TGCAAACCTGAGTGTTGCAA
>390461..390594_NC_011002
TTTGGCAAAGTATTTATCCGCTCGATAAAAGATTTTGAATGACGGCGCCAGCGCGAAATC
GTTTGCGGAATGTGTGTATTTTGCGTAAGCCTTTGAATAATATAAATTTTACAATCTATA
AAGAATGCCGGCTT

>422485..422621_NC_011002
TCGGGCACGGCGGGCGGCCTCGCGCCGACATCTGTGTTCCAAACAGTACATTTGCGCGGC
AATTTAATGCAAATGAGAGGCATTTGCAATCGTAATTTTGAGATTGGGAGGGCGAAGCGA
AATGGCGGGCCGCACGG
>713043..713195_NC_011002
GGCAGGAAGTGATGGAAGTGTTGACGGGCCCTGAGCGTCGGCGGCGCTGGACGGCGGAGC
AGAAGCTGGCGATGGTTCGCGAGAGTTTCGAGCCGGGGAAATCGGTTTCGATGGTCGCGC
GCCAGCACGGCGTGAATCCGAACCAACTGTTCC
>716403..716546_NC_011002
CATCGAAGAAGTAACGACTGGCTCAGTTTTACTTTGCGCGATCAGGCGCGAAGTGGCTCA
GATTTCAAATGCGTTTGACAACATCACGGTCCGTGGCAAAAATGAGGCATGAGTCAGCGG
ATTCCATTGGGGCCAGCGGGCTTC
>788081..788216_NC_011002
CGGTTAAGTGTCTCTGCCGCCTCGCCCTTGAATCCAAGCGGCGAGGCGGCTGACGCGATA
CGCCCGCCGCGATTTTACTTGCGTCCGCATTGGACATAGCGACCGGTTTCAAGCGGCAGC
ACGAGGCAAACATTTG

COPYRIGHT OF CHAPTER 2

The following is quoted from JBC the journal of Biological Chemistry website:

Copyright Permission Policy

These guidelines apply to the reuse of articles, figures, charts and photos in the ***Journal of Biological Chemistry***, ***Molecular & Cellular Proteomics*** and the ***Journal of Lipid Research***.

For authors reusing their own material:

Authors need **NOT** contact the journal to obtain rights to reuse their own material. They are automatically granted permission to do the following:

- Reuse the article in print collections of their own writing.
- Present a work orally in its entirety.
- Use an article in a thesis and/or dissertation.
- Reproduce an article for use in the author's courses. (If the author is employed by an academic institution, that institution also may reproduce the article for teaching purposes.)
- Reuse a figure, photo and/or table in future commercial and noncommercial works.
- Post a copy of the paper in PDF that you submitted via BenchPress.
- Link to the journal site containing the final edited PDFs created by the publisher.

Please note that authors must include the following citation when using material that appeared in an ASBMB journal:

"This research was originally published in Journal Name. Author(s). Title. *Journal Name*. Year; Vol:pp-pp. © the American Society for Biochemistry and Molecular Biology."

COPYRIGHT OF CHAPTER 5

The following is quoted from American Society for Microbiology website, the publisher of mBio Journal

(http://journals.asm.org/site/misc/ASM_Author_Statement.xhtml)

LAST UPDATED: April 2015

ASM Journals Statement of Authors‟ Rights

Authors may republish/adapt portions of their articles

ASM also grants the authors the right to republish discrete portions of his/her article in any other publication (including print, CD-ROM, and other electronic formats) of which he or she is author or editor, provided that proper credit is given to the original ASM publication.

Curriculum vitae

Maryam Khodai-Kalaki, M.Sc.

Education

Ph.D. candidate (September 2011- present)

University of Western Ontario, London ON, Canada

Department of Microbiology and Immunology, Schulich School of Medicine and Dentistry

Thesis title: Characterization of the AtsR/AtsT global regulatory pathway in *Burkholderia cenocepacia*

M.Sc. in **Biology** (2009-2011)

University of Western Ontario, London ON, Canada, Department of Biolog

Thesis title: Identification of ERS2 a new client protein of Hsp90

M.Sc. in **Physiology** (2006-2008)

University of Tehran, Tehran, Iran, Department of Biology

Thesis title: Over-expression of pyrroline-5-carboxylate synthetase *P5CS* contributes to resistance to abiotic stress in transgenic seedlings

Publications

Peer-reviewed Journal articles:

1. **M. Khodai-Kalaki**, A. Andrade, M. Valvano. Modification of *B. cenocepacia* lipopolysaccharide with aminoarabinose and flagellin glycosylation affect

pathogenicity but not innate immune recognition in *Arabidopsis thaliana* (2015). mBio 6(3):e00679-15.doi:10.1128/mbio.00679-15.

2. **M. Khodai-Kalaki**, D. Aubert, M. Valvano. Characterization of AtsR phosphorelay system and identification of its cognate response regulator in *Burkholderia cenocepacia* (2013). J.Biol.Chem. 288:30473-30484.

3. M. Jafarzadeh*****, **M. Khodai-Kalaki***, N. Motamed, O. Noorayin. Genetic transformation of olive somatic embryos through *Agrobacterium tumefaciens* and regeneration of transgenic plants (2011). A. J. Biotech.10 (28): 5468-5475 (*****These authors contributed equally).

Conference abstract presentations:

1. **M. Khodai-Kalaki**, D. Aubert, M. Valvano. Whole genome mapping of AtsR/AtsT regulon in *B. cenocepacia* using Next Generation Sequencing. International Union of Microbiological Societies Congress (IUMS), Montreal, Canada (2014)

2. **M. Khodai-Kalaki**, D. Aubert, M. Valvano. Characterization of AtsR/AtsT signalling pathway in *B. cenocepacia*. EMBL symposium: New approaches and concepts in Microbiology. Heidelberg, Germany (2013)

3. **M. Khodai-Kalaki**, D. Aubert, M. Valvano. AtsR, a global virulence regulator as a target for antimicrobial therapy. Infection and Immunity Research Forum, London, Canada (2013)

4. **M. Khodai-Kalaki**, D. Aubert, M. Valvano. AtsR/AtsT signalling pathway negatively regulates several virulence factors in *B. cenocepacia*. London Health Research Day, London, Canada (2013)

5. **M. Khodai-Kalaki,** N. Motamed. The role of phytohormones in callus induction and somatic embryogenesis. 3rd International Conference of Biology, University of Tehran, Tehran, Iran (2008)

6. **M. Khodai-Kalaki**, M. Haghighat, N. Motamed. Comparative analytical proteomics in somatic and zygotic olive embryos. 3rd International Conference of Biology, University of Tehran, Tehran, Iran (2008)

7. **M. Khodai-Kalaki**, N. Motamed. Over-expression of *P5CS* in transgenic olive somatic embryos. 9th International Conference of Genetic Engineering, Tehran, Iran (2008)

Honours and Awards

Travel award, Department of Microbiology and Immunology, University of Western Ontario (2014)

Teaching Assistant Award of Excellence Nominee, Department of Biology, University of Western Ontario (2011)

Admission Award, Department of Biology, University of Western Ontario (2009)

Scholarships

Schulich Graduate Scholarship, Schulich School of Medicine and Dentistry, Canada (2011-present)

Ontario Student Grant (September 2011-present)

Western Graduate Research Scholarship, Department of Biology, University of Western Ontario, Canada (2009-2011)

National Graduate scholarship, University of Tehran, Iran (2006-2008)

Teaching experience

Graduate Teaching Assistant, Bacterial pathogenesis,
Department of Microbiology and Immunology, University of Western Ontario (2014)

Graduate Teaching Assistant, Regulation of gene expression
Department of Biology, University of Western Ontario (2011)

Graduate Teaching Assistant, Biology for Science
Department of Biology, University of Western Ontario (2010)

Graduate Teaching Assistant, Genomics and beyond
Department of Biology, University of Western Ontario (2010)

Graduate Teaching Assistant, Organismal Physiology
Department of Biology, University of Western Ontario (2009)

Leadership experience

- Invited member of a review committee at Schulich School of Medicine and Dentistry. To provide input in a formal review processes to examine the present state and future prospects of the School and to determine the best course of action in different categories (2014).

- Reviewer of abstract selection for 9th Infection and Immunity research forum, London ON, Canada (2014).

- Infection and Immunity organizational committee member. Reaching out for sponsors and funding agencies, inviting keynote speakers, organizing the meeting to foster intellectually stimulating dialogue among researchers (September 2013-present).

- Biology Departmental Steward in GTA union (Graduate Teaching Assistant Union) representing and speaking on behalf of Biology Graduate Teaching Assistants and representing GTA union in the Department of Biology at University of Western Ontario (September 2010- November 2011).