The regulation of arsenic metabolism in Rhizobium sp. NT-26

Paula Corsini Madeira

University College London

Thesis submitted for the degree of Doctor of Philosophy

2016

I, Paula Corsini Madeira, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Date:

Signed:

Abstract

Arsenic (As) is a toxic metalloid and a major contaminant in terrestrial and aquatic environments. The two soluble forms, arsenite (As^{III}) and arsenate (As^{V}) are toxic to most organisms. A range of phylogenetically distant bacteria are able to oxidize As^{III} to the less toxic form, arsenate As^{V} using the periplasmic arsenite oxidase (AioBA). The two-component signal transduction system AioS/AioR and the As^{III}-binding periplasmic protein AioX are required for As^{III} oxidation and are involved in the transcriptional regulation of the *aioBA* operon. Most As^{III} oxidisers can also reduce As^{V} to As^{III} *via* the As (Ars) resistance system. The focus of this work was to understand the regulation of genes involved in As^{III} oxidation and As resistance together with those involved in phosphate metabolism in the facultative chemolithoautotrophic As^{III} oxidiser NT-26 grown under different conditions. Gene expression was studied by quantitative PCR in cells grown heterotrophically with and without As^{III} or As^{V} in late-log and stationary phases. qPCR was optimised and suitable reference genes were chosen. The expression of genes involved in phosphate transport, sensing As and the genes *aioX, aioS, aioR* (As^{III}-sensing and regulation) and *aioB, aioA* (As^{III} oxidation) and *cytC* (cytochrome *c*) were also analysed in NT-26 grown heterotrophically in the presence or absence of As^{III} or As^{V} at different growth stages (i.e., late-log and stationary phases). To understand gene expression in more detail the transcription start sites (TSS) upstream of *aioX*, *aioB* and *aioA* were identified using the 5'RACE technique. Promoter function was then confirmed by cloning the putative promoter regions upstream of a promoterless β–galactosidase gene and enzyme activity measured in NT-26 grown heterotrophically with and without As^{III} or As^{V} at different growth stages. Two functional promoters were found to control the expression of the *aioXSR* operon, one for its constitutive expression and another for expression in stationary phase. Surprisingly a TSS was identified upstream of *aioA* and the presence of a functional promoter was validated.

Acknowledgements

I would like to sincerely thank Dr. Joanne Santini for giving me the opportunity to do my PhD at her lab, and for continuously encouraging and challenging me to progress. I greatly appreciate the time and patience dedicated to me and all of your effort to make our workplace pleasant. I have learned so much during the last 4 years and I have you to thank for this.

Thanks to my thesis committee – Snezana Djordjevic and Kate Bowers for all the comments, critiques and patience.

This work could not have been done without the support of Tom O., Cameron, Linda, Pia, Tom W, Ben, Kenny and Thushi who made the work environment pleasant and fun to be in. Not to mention all the help with corrections, scientific inputs and discussions.

A big cheers to my London family for all the love, support, friendship, parties, distractions and the much needed booze relief for the hardest moments. To my dear Hollybush family, all the love. From Thursday to Thursday and one step at a time, you helped me immensely through this process. Big cwtch to Steve for all the love and support. From the fresh coffee in the morning to the glass of wine in the evening; you mastered the art of being awesome.

Meus amigos queridos, vocês me deram força pra continuar e energia pra sacudir a poeira. Vocês são meu maior orgulho e inspiração, obrigada por existirem, pelas visitas, pelas bobeiras e por todo o amor que essse oceano todo nao conseguiu inundar. Nao vou citar nomes porque não quero corer o risco de esquecer do passa, da gabi, da bel, da robertinha, da livi, da jaxpion, do patrico, da jo, da livia, do deco. Não quero ser injusta? Melhor não tentar.

Minha familia amada, esta vitória é nossa. Obrigada por todo o suporte incondicional, mesmo não sabendo muito bem no que eu estava me avernturando, vocês sempre me apoiaram e torceram por mim. Eu não teria chegado na metade se não fosse pela confiança que vocês sempre depositaram na minha capacidade. Agora eu vou deixar de lero porque eu tenho que imprimir isso aqui, tô atrasada!

Table of Contents

List of Figures

List of tables

List of Abbreviations

AAA+ ATPases Associated with diverse cellular Activities

AioBA arsenite oxidase

Ap Ampicillin

Arr respiratory arsenate reductase

Ars arsenic resistance

ATP Adenosine triphosphate

bp base pair (nucleotide)

CI confidence interval

CTAB hexadecyltrimethylammonium bromide

dH_2O Distilled water

DMSO dimethylsulphoxide

DNA deoxyribonucleic acid

dsDNA Double stranded DNA

DTT Dithiothreitol

EDTA ethylenediaminetetraacetic acid

F Forward

gDNA genomic DNA

GSP gene specific primer

HK Histidine kinase

IHF Integration host factor

kB kilo bases

LB lysogeny broth

Min Minutes

mM millimolar

mRNA messenger RNA

MSM Minimal salts medium

Mw Molecular weight

NT-26 *Rhizobium* sp. str NT-26

OD Optical density

ONPG O-nitrophenyl-β-D-Galactoside

PCR Polymerase chain reaction

Pi Inorganic Phosphate

qPCR quantitative real time PCR

Ref Reference gene for qPCR

RFU Relative fluorescence units

RNA Ribonucleic acid

RNAP RNA Polymerase

Rox^{V} Roxarsone

Rpm Rotations per minute

RR response regulator

rRNA ribosomal ribonucleic acid

RT Reverse transcriptase

SD Shine-Dalgarno sequence

SOC super optimal catabolite repression

Tat twin-arginine translocation

Tc tetracycline

TSS transcription start site

UPM universal primer

UTR untranslated region

UV ultra-violet

WHO world health organisation

YE Yeast extract

1. General Introduction

1.1 Environmental arsenic contamination

Arsenic (As) is a naturally occurring element that has been in the spotlight for both its infamous use as a poison and its pharmaceutical appeal. Arsenic was widely used as a poison in the Middle Ages and is therefore known as the "King of Poisons" and "Poison of Kings" for its potency, discretion and its use in killing members of the ruling class (Vahidnia et al. 2007). As reached political importance for being the poison of choice of the Borgia and Medici families and later was associated with the assassination of Napoleon Bonaparte in 1851 (Cullen, 2008). An As detection method was developed in 1836 which led to a reduction in its use as a poison (Marsh, 1836). In the 19th century women used As powder for cosmetic purposes believing that it would give "beauty and freshness" to the skin. Victorian women used an As solution, developed by Thomas Fowler in 1786, to achieve a rosy cheek as a consequence of the skin capillaries being damaged (Bentley & Chasteen 2002). Fowler's solution was an arsenic-containing tonic used to treat malaria, asthma and syphilis (Rohe 1896) and in 1845 it started being used to treat leukaemia. In 1910 Paul Ehrlich introduced the As based drug Salvarsan that became known as the "magic bullet" which was the first effective treatment against syphilis and marked the beginning of chemotherapy (Antman 2001; Thorburn 1983). As based drugs are still used with success to treat acute promyelocytic leukaemia (Antman 2001; Dilda & Hogg 2007; Chen et al. 2011).

As is located below phosphorous in the group 15 of the periodic table, with an atomic number of 33. As constitutes 0.00015% of the Earth's crust and is the 52nd most abundant element (Nordstrom 2012) and is found in four oxidation states: arsine (As^{-III}), elemental (As^{0}), arsenite (As^{III}) and arsenate (As^{V}), with As^{III} and As^{V}

being more common in nature (Rosen 2002). As can also be found as organic derivatives, (Table 1.1).

Table 1.1: Inorganic and organic As species

Arsenic species	**Chemical structure**
Inorganic	
Arsine	AsH_3
Arsenite	AsO_2H
Arsenate	AsO_4H_3
Organic	
Methylarsine (MMA)	CH_3AsH_2
Dimethylarsine (DMA)	$(CH_3)_2AsH$
Trimethylarsine (TMA)	$(CH_3)_3As$
Trimethylarsine oxide (TMAO)	$(CH_3)_3AsO$
Tetramethylarsonium ion (TMA^+)	$(CH_3)_4As^+$
Arsenocholine	$(CH_3)_3As^+CH_2CH_2OH$
Arsenobetaine	$(CH_3)_3As^+CH_2COOH$

Adapted from Páez-Espino et al (2009).

Approximately 99% of naturally occurring As occurs in rocks and minerals and the remaining arsenic is soluble and found in oceans, rivers, soil and sediments (Bissen and Frimmel, 2003). Arsenopyrite (FeAsS) is the most abundant arsenic-containing mineral although there are more than 200 arsenic-containing minerals, such as realgar (AsS), orpiment (As_2S_3), olivenite (Cu_2OHAsO_4), cobaltite (CoAsS), proustite (Ag_3AsS_3) and rammelsbergite (NiAs) (Bissen & Frimmel 2003; Mandal & Suzuki 2002). Volcanic action is responsible for the release of arsine in the environment and low-temperature volatilization and erosion of arsenic-containing rocks also contribute to the release of As in the atmosphere (Bissen & Frimmel 2003; WHO 2001). Inorganic arsenic of geological origin is also found in groundwater and represents a problem when this water is used for human consumption. The natural

accumulation of As in groundwater is a significant problem in Bangladesh where >40 million people are presumed to be under risk according to the World Health Organization (WHO) (WHO 2010). High concentrations of As in groundwater were also reported in Argentina, Chile, Mexico, China, Hungary, West Bengal (India) and Vietnam (Smedley & Kinniburgh 2002).

Anthropogenic sources of As in the environment include a variety of industrial processes, for example: the production of glass, plastics, pharmaceuticals and chemicals, electronics and semiconductor materials, mining, and also the use of pesticides and herbicides containing As (Angerer *et al.*, 1991). The smelting of copper, burning of fossil fuels and use of wood preservatives containing As are amongst the activities that release As into soil and freshwater environments (Bissen & Frimmel 2003). In 2000, the cumulative total of arsenic produced from anthropogenic industrial activity was estimated at 4.53 million tonnes, of which greater than 46% was produced by the coal and petroleum industries (Han et al. 2003).

According to the WHO, As is one of the top 10 chemicals of major public health concern with approximately 50 million people being exposed to concentrations of As in the water higher than 50 µg/litre, 5 times higher than the safe limit determined by WHO. As^{III} and As^{V} are both toxic to living organisms although they have different mechanisms of toxicity. As^{III} is known to bind to sulfhydryl groups of proteins, inhibiting their function and enter the cells using mainly the glycerol transport channel (Meng et al. 2004). As^{V} is toxic due to its ability to enter cells using the phosphate transport system, where it can uncouple oxidative phosphorylation by replacing phosphate in ATP. As^{V} is 100 times less toxic than As^{III} (Rosen 2002). The

symptoms of chronic As intake in humans is characterized by skin lesions with hyperpigmentation, hyperkeratosis, and hypopigmentation (Melkonian et al. 2012). The accumulation of As in the body is known to indirectly induce different types of cancer (Hunt et al. 2014; Cohen et al. 2016).

1.2 Microbial arsenic metabolism and resistance

Although As is toxic to most organisms, a range of phylogenetically diverse bacteria are able to survive and even thrive in As-contaminated environments. The biotic oxidation of As^{III} and reduction of As^{V} has been observed in many environments and plays an important role in the global cycling of arsenic (Stolz and Oremland, 2003). The activity of As-metabolising bacteria can affect the mobility and speciation of As (Rhine et al. 2005) and it is believed that the accumulation of As in Bangladesh and West Bengal groundwater is partially the product of the oxidation of arsenic by microorganisms although this mechanism is not well understood (Osborne et al. 2015; Nickson et al. 2000).

As metabolism in bacteria was first observed in 1918 with the description of a bacterium able to both oxidise and reduce As^{III} and As^{V} respectively, named *Bacterium arsenoxydans*, isolated from cattle dipping tank fluids in South Africa (Green 1918). Since then, phylogenetically diverse prokaryotes (Newman et al. 1998; Santini & Vanden Hoven 2004), including members of Proteobacteria and Gram-positive bacteria, have been reported to oxidise As^{III} to the less toxic As^{V} (Oremland & Stolz 2003). As^{V} reduction to As^{III} can either occur as part of a detoxification system (i.e., the arsenic resistance system, Ars) or as part of a respiratory process where As^{V} is used as the terminal electron acceptor and organic (e.g., acetate) or inorganic (e.g.,

hydrogen) compounds as electron donors (Newman et al. 1998; Oremland et al. 2002; Vanden Hoven & Santini 2004; Santini & Vanden Hoven 2004). All these processes have an impact on the cycling of arsenic in the environment. The respiratory process involves the presence of a respiratory As^{V} reductase (Arr) enzyme which is found in diverse bacteria (Krafft & Macy 1998; Malasarn et al. 2004). As^{V} reduction can also be a detoxification mechanism, in which after As^{V} enters the cell using phosphate transporters (PstS or Pit) it is reduced in the cytoplasm by the As^{V} reductase, ArsC, followed by the extrusion of As^{III} from the cell (Rosen 2002; Rosen et al. 1995; Gladysheva et al. 1994). This process, present in many prokaryotes, is known as the As resistance system (Ars) and will be explained in more detail in sections 1.4 and 1.6. As^{III} oxidation has been reported to occur both aerobically and anaerobically with oxygen, nitrate or chlorate being used as electron acceptors (Vanden Hoven & Santini 2004; Sun et al. 2010).

Also part of the As biotransformation and cycling is the methylation and demethylation of As by microorganisms (Qin et al. 2006). Environmental inorganic As can be converted to soluble and gaseous methylated species as part of a detoxification process that involves the reduction of As^{V} followed by the methylation of As^{III} directed by the As^{III} S-adenosylmethionine methyltransferase (ArsM) (Oremlandb & Stolz 2003; Qin et al. 2006). Although methylated As^{III} species (MAs^{III}) are more toxic than both inorganic As^{III} and As^{V} it is believed that they can disperse more easily and therefore this process can be beneficial to the organisms (Qin et al. 2006; Zhu et al. 2014). Homologues of ArsM are found in all kingdoms, being a widespread mechanism of As detoxification. Another environmentally important mechanism of As cycling is the degradation of pentavalent arsenicals (either MAs^{V} or

roxarsone [Rox^{V}]) accumulated in the environment due to the use of herbicides and pesticides for agriculture and animal husbandry. The MAs^{V} are reduced to the more toxic MAs^{III} followed by demethylation by the ArsI enzyme to generate inorganic As^{III} (Yoshinaga & Rosen 2014). The demethylation process is done by the As-C bond lyase, member of Ars, ArsI (Yoshinaga & Rosen 2014; Nadar et al. 2016). It is not known how MAs^{V} is reduced to MAs^{III} prior to the demethylation process although it is believed that it is part of a two-step pathway coordinated by different bacterial species in the community (Yoshinaga et al. 2011). The third process of the Ars detoxification process is conducted by methylarsenite oxidase (ArsH) which confers resistance to MAs^{III} by oxidizing it to the less toxic MAs^{V} (J. Chen et al. 2015). These three processes are likely to be taking place in the environmental microbial communities (Yang & Rosen 2016) creating a cycle that maintains the balance between organic and inorganic As species in the environment. The oxidation, reduction, methylation and demethylation processes coordinated by bacteria in the environment in the response to As are summarised in figure 1.1:

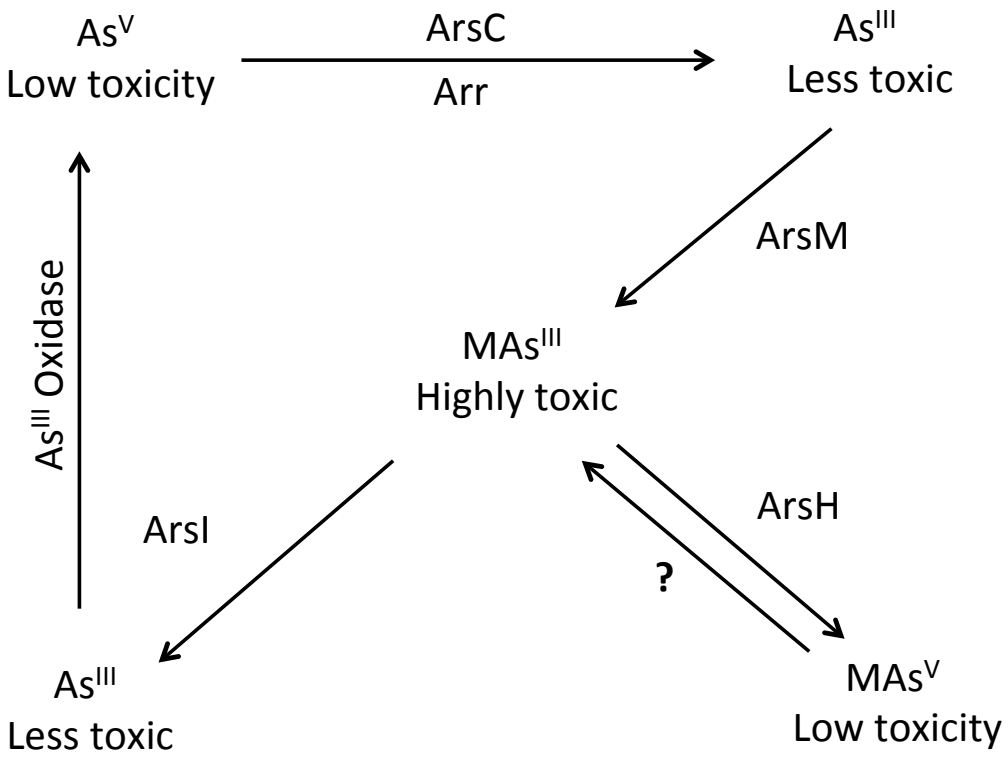


Figure 1.1: Summary of oxidation, methylation and demethylation processes coordinated by bacteria in response to As. The enzyme ArsC, ArsM, ArsH and ArsI are part of the As resistance system and are responsible for organoarsenical production and detoxification. As^{V} is reduced by ArsC or Arr to the more toxic As^{III}. Some microbes have ArsM to transform As^{III} into the considerably more toxic (and carcinogenic) MAs^{III}. Other microbes have ArsI C-As lyase, to cleave the methyl group from MAs^{III}, forming inorganic As^{III}. Since As^{III} is less toxic than MAs^{III}, this is considered a detoxification process. Other bacteria have an ArsH oxidoreductase that oxidizes MAs^{III} to relatively nontoxic pentavalent MAs^{V}, also a detoxification process. The reduction of MAs^{V} to MAs^{III} is known to occur through a yet unidentified mechanism. The oxidation of As^{III} to As^{V} is a process that occurs in the periplasm and will be explained in more detail in the next section (section 1.3) of this chapter. Adapted from (Yang & Rosen 2016).

1.3 Mechanism of As^{III} oxidation

Another mechanism to facilitate the metabolism of As in microorganisms is the As^{III} oxidation, first described in a *Bacillus* in 1918 (Green, 1918). It was first thought to be a detoxification mechanism but has since been shown to be a respiratory mechanism in many organisms. As^{III} can enter the cell *via* the glycerol uptake channel GlpF, and

possibly other mechanisms, (Meng et al. 2004) and its toxicity is related to its ability to bind to sulphydryl groups of proteins, inhibiting their function (Hughes 2002). The biotic oxidation of As^{III} has been observed in many environments, and plays an important role in the global cycle of As by converting As^{III} to the less toxic and less mobile As^{V} (Stolz et al. 2002). As^{III} is oxidized to As^{V} in the periplasm by the As^{III} oxidase (Aio), a heterodimeric enzyme composed of a large subunit AioA that has a molybdenum atom (Mo) at its active site and 3Fe-4S cluster, and a small subunit (AioB) which contains a Rieske-type 2Fe-2S cluster (Figure 1.2) (Ellis et al. 2001; Warelow et al, 2013). When As^{III} is oxidised to As^{V}, it reduces the Mo and the two electrons are passed *via* the 3Fe-4S and Rieske clusters to an electron acceptor – cytochrome *c* or azurin – the Mo oxidation state is restored by hydrolysis (Ellis et al. 2001; Anderson et al. 1992; Santini et al. 2007). Aio is exported from the cytoplasm across the membrane by the twin-arginine translocation (Tat) pathway directed by the AioB subunit (Santini et al. 2004). The Tat pathway is a common mechanism found in prokaryotes to export folded molybdoenzymes and Rieske proteins from the cytoplasm (Natale et al. 2008).

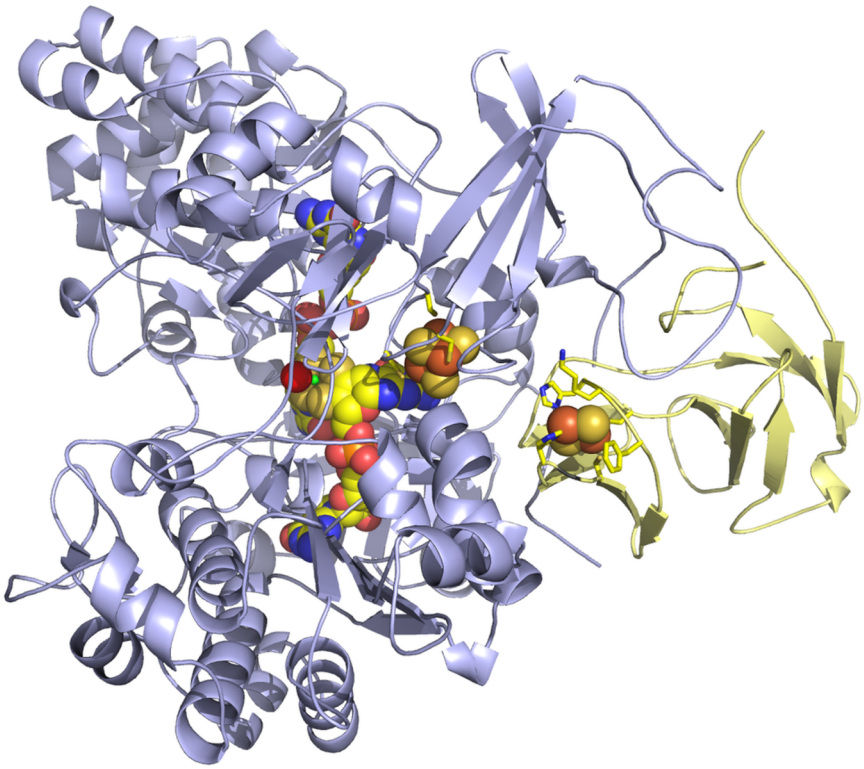


Figure 1.2. The heterodimeric structure of the AioBA from NT-26. The Aio consists the bigger subunit AioA corresponding to the α - chain (pale blue) and the smaller subunit AioB, the β - chain (pale yellow). The pterin co-factor, 3Fe-4S, 2Fe-2S clusters are shown as space filling spheres. Residues, which ligate the clusters, are shown as sticks. Atoms are coloured iron orange, sulphur dark yellow, carbon bright yellow, molybdenum green, phosphorus bright orange, oxygen red, nitrogen blue. From Warelow *et al,* 2013.

As^{III}-oxidising bacteria have been isolated from a diverse range of arsenic contaminated environments and samples (Osborne & Santini, 2012) and As^{III} oxidation was observed in both aerobes and anaerobes organisms (Stolz et al. 2006). Aerobic As^{III} oxidation can occur chemolithoautotrophically - where the organism does not require an organic source of carbon for growth and gains energy by using As^{III} as the electron donor - or heterotrophically - where the bacteria require an organic source of carbon for growth. When the As^{III} occurs heterotrophically it is considered to be a detoxification mechanism (Osborne & Santini, 2012), however, it has been shown that organisms grown heterotrophically can also gain energy from As^{III} oxidation, as it is the case for the *Rhizobium* sp. str. NT-26 (vanden Hoven and

Santini, 2004).

Microbial As^{III} oxidation can also happen by coupling the oxidation of As^{III} to the reduction of nitrate (Zargar et al. 2010) as described for the chemolithoauthotroph haloalkaliphile *Alkalilimnicola ehrlichii* str. MLHE-1 (Oremland et al. 2002; Zargar et al. 2012) or coupling the As^{III} oxidation with chlorate reduction as described for the authotrophs *Dechloromonas* sp. strain ECC1-pb1 and *Azospira* sp. strain ECC1-pb2 (Sun et al. 2010). Also As^{III} can also be used as an electron donor for anoxygenic photosynthesis (Mccann et al. 2016). The As^{III} oxidase (Arx), a periplasmic protein so far only detected in anaerobic As^{III} oxidation, is phylogenetically unrelated to the Aio and can operate in a bidirectional manner *in vitro* (Richey et al. 2009; Oremland et al. 2009). Arx is thought to have evolved from the Arr due to the high sequence identity (30 %) between them (Richey et al. 2009). Although Arx has only been experimentally shown to oxidise As^{III} in MLHE-1, Arx homologs have been found in the As^{III}-oxidisers *Ectothiorhodospira* sp. PHS-1, *Ectothiorhodospira* sp. MLP2 (Kulp et al., 2008), *Azoarcus* sp. DAO1 (Rhine *et al.* 2007) and *Halorhodospira halophila* SL1 (Zargar *et al*. 2012).

1.4 Mechanisms of As resistance

The Ars system for As detoxification is widely distributed throughout all three domains of life (Ellis et al. 2006; Jackson & Dugas 2003) and the genes encoding for the Ars can be either plasmid or chromosome borne (Diorio et al. 1995).

The minimum elements found in the Ars system are the ArsR, ArsB and ArsC proteins. ArsC is an As^{V} reductase responsible for the reduction of As^{V} to As^{III} in the cytoplasm (Guangyong Ji & Silver 1992). ArsB is a membrane protein that functions as an As^{III}

pump using the cell membrane potential force (Tisa & Rosen 1990). ArsR is responsible for the regulation of the *arsRBC* operon by repressing the transcription of this operon in the absence of As^{III} (Jianhua Wu & Rosen 1993; Wu & Rosen 1991). In addition to ArsR, ArsB and ArsC, the ArsA and ArsD are also commonly found, forming the ArsRDABC system. The ArsD acts as a repressor, along with ArsR, and ArsA is an ATPase that provides energy to ArsB making the As^{III} efflux more efficient (Rosen et al. 1988). Organisms that possess ArsAB are resistant to higher concentrations of As^{III} due to the higher efficiency of ArsAB compared to ArsB alone (Dey & Rosen 1995). ArsD has also been described as a chaperone, binding to the ArsBA pump and transferring As^{III} to ArsA (Rosen et al. 2011; J Wu & Rosen 1993).

Some exceptions to this system can be found, as shown for *Sinorhizobium meliloti* where the ArsB is absent and the As^{III} efflux is controlled by the aquaglyceroporin AqpS (Yang et al. 2005). The As^{III} efflux can also be done by the Acr3 pump, whose homologues have been described in every member of the tree of life, different from ArsB that is not found in Eukarya (Mukhopadhyay et al. 2002). A mechanism of As^{III} transport and As^{V} reduction was found in actinobacteria isolated from both soil and marine environments that consists of the ArsC fused to an aquaglyceroporin, resulting in an optimized channel for As^{III} extrusion that is able to reduce As^{V} (Wu et al. 2010). Also part of the Ars system, are the ArsI, ArsM and ArsH described previously (Section 1.2, Figure 1.1) as being involved in the methylation and demethylation of As elements (Yang & Rosen 2016).

1.5 The *aio* gene cluster and regulation of its expression.

Despite the significant progress obtained from studies on the metabolism of As^{III} in various organisms, understanding the regulation of the *aio* gene cluster, that encodes the Aio, is limited. The genes *aioB* and *aioA* (*aioBA*) have been identified in many phylogenetically diverse arsenite oxidisers (Van Lis et al. 2013), but they seem to be regulated differently.

Our understanding of *aio* gene regulation started with the suggestion that the AioSR is a two-component system responsible for the regulation of *aioBA* in *A. tumefaciens* sp. str. 5A (Kashyap et al. 2006). AioSR were confirmed as a histidine kinase and response regulator pair, in *in vitro* analysis where AioS was shown to autophosphorylate and transfer the Pi to AioR (Sardiwal et al. 2010). Two-component signal transduction systems are common to many prokaryotes and involve a histidine kinase that senses the stimulus and autophosphorylates its histidine domain, subsequently transferring this phosphoryl group to a conserved aspartate residue of the cognate response regulator (RR) (Figure 1.3). This phosphorylation can cause conformational changes of the C-terminal portion of the RR resulting in activation of this domain, which allows it to bind to the promoter region of the regulated gene (figure 1.3).

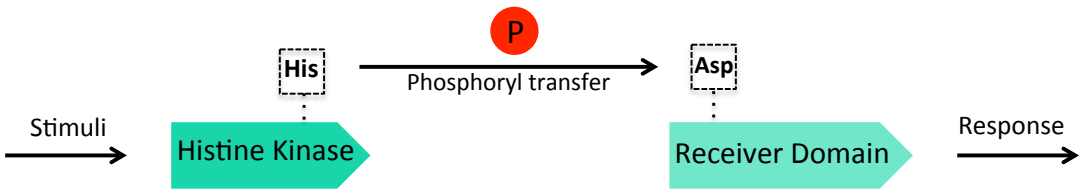


Figure 1.3: Two-component system organization with the histidine kinase and the receiver domain

AioS is a membrane bound histidine kinase that phosphorylates the cognate response regulator AioR, activating it by specific conformational changes (Sardiwal et al. 2010). AioR has a Helix-turn-helix DNA-binding motif at its C-terminal domain, and also has an ATPase domain that belongs to the AAA+ superfamily (ATPases Associated with various cellular Activities) (Figure 1.4). The AAA+ ATPase transcription activators are required for the transcription activation regulated by the σ^{54} transcription initiation factor (Gao and Stock, 2009).

Figure 1.4: AioR domain organisation. REC – receiver domain, AAA+ - ATPases Associated with a variety of cellular Activities, Db – DNA-binding domain.

For all the As^{III}-oxidisers described possessing *aioS* and *aioR* genes so far, the σ^{54} promoter sequence was found upstream of the *aioBA* genes (Koechler *et al*., 2010). In *Herminiimonas arsenicoxydans* ULPAS-1, *Agrobacterium tumefaciens* 5A and *Rhizobium* sp. str. NT-26 it has been observed that in addition to the AioSR transcription control, the expression of *aio* depends on the expression of *rpoN*, encoding the RNA polymerase sigma factor N (RpoN) also known as sigma factor (σ^{54}) (Koechler *et al*., 2010; Kang *et al.,* 2012; Andres et al. 2013). The conserved σ^{54} promoter motif was found upstream of *aioBA* in NT-26 (Santini et al. 2007), *H. arsenicoxydans* ULPAS-1 (Koechler et al. 2010) and *Thiomonas arsenitoxydans* (Moinier *et al.,* 2014) and the putative AioR-binding site upstream the σ^{54} promoter consensus sequence of *aioBA* was described for *Rhizobiales* and *Burkholderiales* (Andres et al. 2013). The presence of both the σ^{54} conserved sequence and the AioR-

binding site upstream of *aioBA,* supports the idea that AioR interacts with σ^{54} to initiate the transcription of *aioBA* (Figure 1.5).

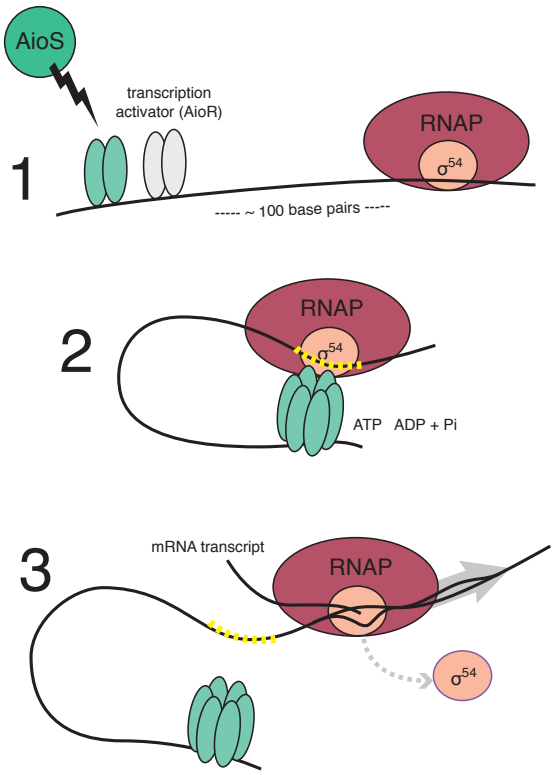


Figure 1.5: AioR interaction with sigma54 promoter upstream of *aioBA.* The common mechanism described for the activation of σ^{54} by AAA+ ATPase transcription activators is exemplified here with AioR being the transcription activator. 1 - AioR binds on the DNA in a conserved region upstream of where σ^{54} binds; 2 - When AioR is phosphorylated and active it forms an hexamer (or heptamer) and forms a loop to interact with σ^{54}; 3 - The interaction between the active AioR and σ^{54} leads to the activation of the transcription of *aioBA.*

The NT-26 AioS doesn't contain any cysteine residues which one would assume would be required for As^{III} binding (Sardiwal et al. 2010). AioX is suggested to be involved in As^{III} binding and *aioBA* regulation after RT-PCR analysis showed that *aioX* is in operon with *aioSR* in *Achromobacter* sp. str. SY8 (Cai et al. 2009). Liu et al., has

shown that AioX is a periplasmic protein, capable of binding As^{III} and required for As^{III} oxidation in *A. tumefaciens* 5A (Liu et al. 2012). It was suggested that by binding As^{III}, AioX would undergo a conformational change to facilitate interaction with AioS and therefore lead to the activation of AioR and expression of *aioBA* operon (Liu et al. 2012). The inactivation of the *aioX*, which is located upstream of *aioSR*, results in a loss of As^{III} oxidation in NT-26 and in *A. tumefaciens* 5A (Andres *et al.*, 2013; Liu et al. 2012). *aioXSR* are only found in operon in *Proteobacteria*; *aioXSR* can either be positioned in the same direction as *aioBA,* as it is the case for *Alphaproteobacteria* or in the opposite direction as observed in the *Betaproteobacteria* (Sardiwal et al. 2010).

The genes *cytC* and *moeA*, encoding for a *c*-type cytochrome and a molybdenum cofactor, respectively, can be found downstream of *aioBA* in some *Proteobacteria* such as NT-26, T. *arsenitoxydans* 3As*, A. tumefaciens* 5A and *A. albertimagni,* according to observations made based on the As^{III}-oxidase genome sequences available (Slyemi & Bonnefoy 2012; Van Lis et al. 2013; Hao et al. 2012; Trimble et al. 2012; Arsène-Ploetze et al. 2010, Slyemi *et al.,* 2013). In *A. tumefaciens* 5A, it was demonstrated that *aioB, aioA, cytC* and *moeA* are transcribed together (Kashyap *et al*., 2006) and in NT-26 *aioA* and *cytC* were shown to be transcribed in operon, but only when NT-26 is grown in the presence of As^{III} (Santini et al. 2007).

For some As^{III} oxidisers such as *Polaromonas* sp. str GM1*, Halomonas* sp. str. HAL1, *Pseudomonas* sp. str. TS44, *Chloroflexus aurianticus* and *Pseudomonas xanthomarina* sp. str. S11*, aioX, aioS* and *aioR* are not present and therefore AioX, AioR and AioS do not control the expression of *aioBA* (Cai et al., 2009; Koechler et al. 2015; Lin et al., 2012; Osborne et al. 2010), suggesting that *aioBA* is either constitutively expressed or the regulation of *aioBA* can involve other elements. A new transcription factor, AioF,

has been described for *T. arsenitoxydans* 3As (Moinier et al. 2014). AioF is a ArsR/SmtB metalloid responsive regulator that is capable of binding As^{III} and As^{V}, and suggested to be positively regulating the expression of *aioBA* either by improving the interaction between the AioR and σ^{54} or by playing a role in transcription interference between *aioX* and *aioB* (Moinier et al. 2014). The *aioBA* and *aioXSR* in *T. arsenitoxydans* 3A are positioned in opposite directions and share the same promoter region, with *aioX* transcription start site positioned inside *aioB*. However, *aioBA* and *aioXSR* are not activated by the same promoters, with the *aioBAF* operon being under the control of the σ^{54} promoter and the *aioXSR* genes under control of a σ^{70} promoter (Moinier et al. 2014). In *T. arsenitoxydans* 3As *aioBA* expression is induced in the presence of As^{III} while the expression of *aioXRS* is not differentially expressed in presence of As^{III} or As^{V} (Moinier et al. 2014, Slyemi et al. 2013).

Another mechanism that been linked to the regulation of *aioBA* expression is the quorum sensing. Quorum sensing is a "communication mechanism" used by bacteria to assess the population of its species (and others) in a given environment and allow specific response to the environmental cues (Miller & Bassler 2001). Quorum sensing has been linked to *aioBA* regulation in *Chloroflexus aurantiacus* as it has a *luxR* gene upstream of the *aioBA* genes and LuxR is a known transcription activator which responds to cell density (Van Lis et al. 2013). Quorum sensing has also been suggested to be involved in regulation of the *aioBA genes* in *A. tumefaciens* 5A after the auto inducer ethyl acetate was shown to be required for the up-regulation of *aioBA* in the absence of As^{III} in the growth medium (Kashyap et al. 2006). It has also been suggested for *A. tumefaciens* 5A that the phosphate/malate antiporter DctA, previously described as capable of substituting phosphate for As^{V} (Furbank et al.

1990), can be involved in the regulatory network controlling the expression of *aioBA* as the disruption of *dctA* affected the expression levels of *aioBA* genes in presence of As^{III} (Kang et al. 2014) suggesting that the regulation of As^{III} oxidation can be part of a more complex regulation pathway and it can be controlled by not only As^{III}, but also phosphate and quorum sensing.

1.6 Regulation and expression of the *ars* gene cluster

The *ars* genes - *arsR, arsC* and *arsB1* - are widespread in nature and appear to be generally co-transcribed in the different organisms studied (Páez-Espino et al. 2009). The core *ars* genes found in most organisms, consist of *arsR, arsC* and *arsB1*, encoding the ArsR regulator, As^{V} reductase and the As^{III} efflux pump, respectively (Xu et al. 1998). The ArsRBC from *Staphylococcus xylosus* has been shown to confer resistance to As^{III} and As^{V} to *E. coli* and *Bacillus* previously susceptible to either As^{III} or As^{V} (Rosenstein et al. 1992).

The expression of the *ars* operon is controlled by ArsR, which in the absence of As directly binds to the -35 element upstream of the *ars* promoter, repressing the transcription of the operon. In the presence of As^{III}, the ArsR undergoes a conformational change that leads to its release from the DNA, allowing the expression of the *ars* genes (Wang et al. 2004; Wang et al. 2009; Cai et al. 1998). The inactivation of the *arsR* gene results in constitutive expression of the *ars* operon as shown in both *Acidiphilium multivorum* AIU 301 (Suzuki et al. 1998) and *S. xylosus* (Rosenstein et al. 1992). For some microorganisms like *S. xylosus* and *Acidithiobacillus ferrooxidans* the *ars* operon was also induced in the presence of As^{V} (Butcher &

Rawlings 2002; Rosenstein et al. 1992), varying from that observed in *E. coli* where As^{V} was unable to dislodge ArsR (Jianhua Wu & Rosen 1993).

ArsR was shown to be involved in the regulation of the operon containing the respiratory As^{V} reductase gene (*arr*) in *Shewanella* sp. str. ANA-3 (Murphy & Saltikov 2009). In *A. tumefaciens* 5A, ArsR was described controlling the expression of the genes *phoB* and *pstS*, which encode a phosphate transcriptional regulator and a phosphate transporter, respectively (Kang, Heinemann, et al. 2012), however, one of the four copies of *arsR* found in *A. tumefaciens* 5A codes for an ArsR that is not responsive to As^{III} (Kang et al. 2016). For *Pseudomonas putida* KT2440, bearing two different copies of *arsR*, the different ArsR confer divergent resistance to As (Páez-Espino et al. 2015). The presence of multiple copies of *arsB1* and *arsC* was found in *P. aeruginosa* and found to confer increased resistance to As (Cai et al. 1998). Based on what has been described in the literature regarding the regulation of the *ars* operon and the different Ars response to the presence of As^{III}, it is possible to say that the response to As in microorganisms is complex and involve several different elements.

1.7 Phosphate metabolism and gene regulation

Phosphate and arsenate are analogues and phosphorus is located above As in the periodic table, in the group 15. It has been reported that phosphate and As metabolism can be interconnected. In *A. tumefaciens* 5A and *H. arsenicoxydans* ULPAS-1, the presence of As^{III} can affect the expression of genes *pstS* and *phoB,* involved in phosphate transport and response to low inorganic phosphate (Pi) concentrations, respectively (Kang et al. 2012; Cleiss-Arnold et al. 2010).

Pi plays key roles in cells as it is involved in energy metabolism, protein phosphorylation and is a major structural component of nucleic acids and membrane phospholipids (Yuan *et al*, 2006). Pi is also involved in transcriptional control of genes involved in chemotaxis, cell division, quorum sensing and motility (Pratt *et al*, 2009).

The Pi concentration in the bacterial cell is sensed by and controlled by the two-component system PhoR and PhoB, where PhoR is the histidine kinase that senses changes in Pi concentrations and phosphorylates the response regulator PhoB (Hoch & Silhavy, 1995). PhoB is a transcriptional activator that, when phosphorylated, activates the transcription of more than 50 genes as described for *E. coli* and *Caulobacter crescentus* (Yang et al. 2012; Lubin et al. 2016). PhoB binds to the promoter region at a conserved sequence known as the PhoBox, located usually close to the -35 region upstream of the PhoB-regulated genes (Makino *et al*., 1996; Yuan *et al,* 2006). Most of the identified PhoBoxes contain two conserved 7 nucleotide direct repeats of 5'-CTGTCAT-3' separated by a conserved 4 nt spacer in the middle (Makino *et al*, 1989). The PhoB when bound to the PhoBox interacts with the σ^{70} subunit of RNA polymerase to control transcription initiation (Makino *et al*, 1993).

PhoR plays a role in activation of PhoB, but can also prevent PhoB phosphorylation or even dephosphorylates it when the Pi concentration in the cell increases (Wanner *et al*., 1996). The response to Pi concentration can also involve the Pi-binding protein PstS (high affinity Pi transporter) that is saturated when Pi is in excess and forms a complex with PhoU (negative regulatory protein of the Pho regulon) and PhoR, inhibiting PhoR activity. After Pi starvation period, PhoR can dephosphorylate PhoB, returning to the inactive state (Figure 1.6) (Wanner, 1993; Lamarche, 2007; Hsieh and Wanner, 2010).

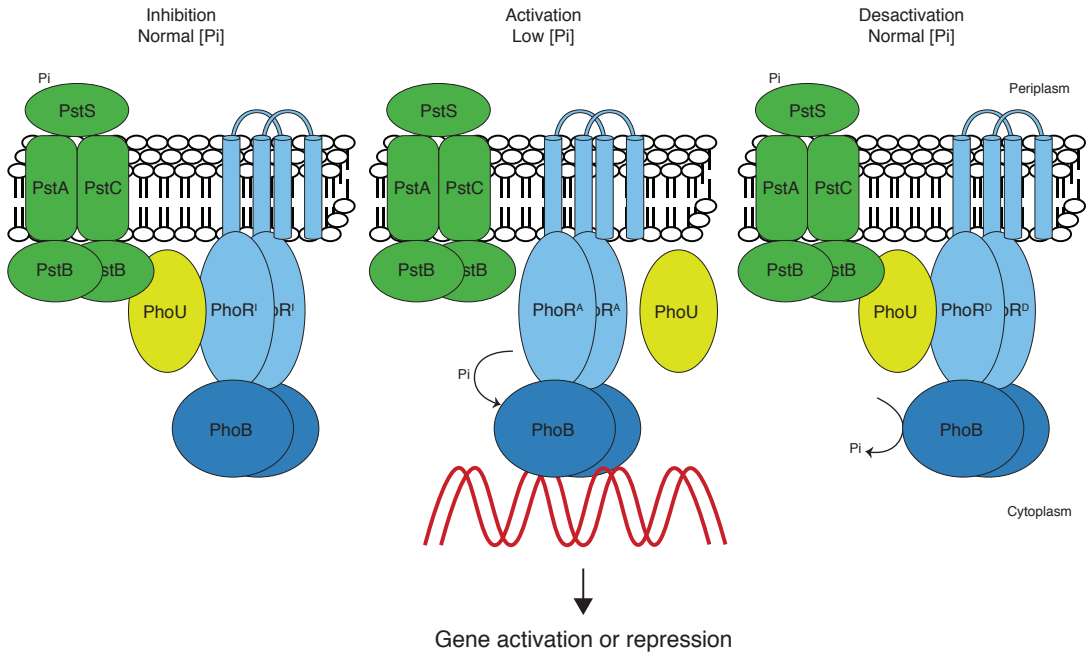


Figure 1.6: Model for transmembrane signal transduction by environmental Pi. When Pi is in excess, the Pi binding protein PstS is saturated and forms a closed complex along with PhoU and PhoR leading to formation of $PhoR^{I}$, which interferes with phosphorylation of PhoB. Under Pi limitation, the PhoU is released and $PhoR^{A}$ activates PhoB through phosphorylation. After the period of Pi limitation is over, the $PhoR^{D}$ dephosphorylates the PhoB, forming the closed complex again. Formation of $PhoR^{D}$ requires an increased amount of PhoU or PstB in addition to excess Pi. Adapted from Hsieh and Wanner (2010).

In the absence of PhoR, other Pi-independent controls can act on the Pho regulon in a cross-regulation mechanism. For example, the sensor protein CreC, involved in carbon catabolism, can induce PhoB activity (Wanner & Wilmes-Riesenberg, 1992) and when acetyl phosphate is in excess, it can also directly phosphorylate PhoB (Wanner, 1993). It has been shown that other non-cognate histidine kinases like QseC, ArcB, KdpD, BaeS and EnvZ can also activate PhoB, in the absence of PhoR (Lamarche *et al.,* 2008) indicating that Pi metabolism has a complex regulatory network.

1.8 Overview of regulation of gene expression

Prokaryotic gene expression is primarily controlled at the level of transcription (Browning & Busby, Stephen J 2004; Wösten 1998). The RNA synthesis is controlled by a regulator that interacts with a specific regulatory DNA element. The transcription of DNA to RNA by the RNA Polymerase (RNAP) occurs in the cytoplasm and is followed by the translation of the RNA to protein, by the ribosome (Kaberdin & Bläsi 2006; Laursen et al. 2005). The regulation at the transcriptional level allows for the control of the amount and type of protein synthesised by the cell at certain time, avoiding unnecessary energy waste and optimizing the process (Browning & Busby, Stephen J 2004; Desroche et al. 2005).

The transcription initiates with the RNAP binding to the promoter region (usually around the -10 and -35 positions, with respect to the transcription start site) upstream of a gene; the recognition of the binding site to the DNA is directed by the sigma factors of the RNAP. Sigma factor binds to conserved regions within the promoter (Gross et al. 1998; Qiu et al. 2013). When a particular gene needs to be expressed, the sigma factor will bind upstream this gene at the sigma factor binding consensus sequence in the promoter region and direct the RNAP to the correct place to initiate transcription. This process can be additionally regulated by the presence of negative regulators that bind to the promoter region, making the assembly of the sigma factor/RNAP complex impossible. On the other hand, the presence of activators can release the negative regulators and/or recruit the RNA Polymerase allowing for the transcription to commence.

Among the mechanisms that can interfere with transcription initiation is the two-component system that consists of a histidine kinase, responsible for sensing the signal and phosphorylates its cognate response regulator in the presence of the stimulus. The response regulator can act either as a positive regulator (Lin et al. 2015; Blanco et al. 2012) or a negative regulator (Foreman et al. 2012; Kato et al. 2003), depending on the system . That is a form of control in response to concentration of nutrients and metals as seen for AioXRS (response to As in the cell, section 1.5) and PhoRB (response to Pi concentration in the cell, section 1.7). Another form of control involves the quorum sensing mechanism, used by prokaryotes to assess the populations of its species (and others) in a given environment, and control the gene expression accordingly (Miller & Bassler 2001). The regulation by quorum sensing involves the presence of autoinducers, molecules produced and secreted by bacteria that can be sensed when accumulated in the medium, indicating increased density of cells. Bacteria are able to sense autoinducers from their own species and from others, allowing "communication" between different organisms and species. The autoinducers can then bind to enhancers, repressors and even affect the activation of two-component systems, adding another layer to the already complex regulation of gene expression in prokaryotes.

1.9 Aims of this study

The overall objective of this study was to gain a better understanding of the mechanisms involved in the regulation of genes involved in arsenic metabolism in *Rhizobium* sp. str. NT-26.

1 – The first aim was to optimise the use of qPCR in NT-26 to study the expression of

genes related to the metabolism of phosphate and resistance to arsenic in response to the presence of As^{III} and As^{V} in the growth medium and growth phase.

2 - The second aim was to monitor the expression of the *aio* genes in NT-26 grown in the presence or absence of As^{III} or As^{V} grown at different growth phases using qPCR.

3 – The third objective was to elucidate which transcriptional regulation processes are involved in the control of *aioX, aioB* and *aioA* expression in NT-26 by determining the transcription start site of each gene and monitoring the activity of their promoters.

Chapter 2

Quantitative real-time PCR (qPCR) analysis of gene expression in NT-26 grown in the presence and absence of As^{III} and As^{V}

2.1 Introduction

Arsenic (As) contamination in the environment poses a huge risk to life due to the high toxicity of arsenite (As^{III}) and arsenate (As^{V}), the two soluble forms of As. As^{III} has a high affinity for the sulphydryl group of proteins and can enter cells using the same channel used by glycerol. As^{V} is a molecular analogue of inorganic phosphate (Pi) and competes with it in cellular reactions. Many organisms are able to survive in As^{V} contaminated environments due to the presence of an As resistance system (Ars).

The facultative chemolithoautotrophic arsenite oxidizer, *Rhizobium* sp. str. NT-26, was isolated from the Granites goldmine in Northern Territory, Australia (Santini et al. 2000). NT-26 is a α-Proteobacteria motile rod with two sub terminal flagella and is able to gain energy from arsenite oxidation under both chemolithoautotrophic and heterotrophic conditions (Figure 2.1) (Santini *et al*., 2000). NT-26 is able to resist to a concentration of 0.5M As^{V} and possesses genes encoding for both the Ars and As^{III} oxidation (Andres et al. 2013).

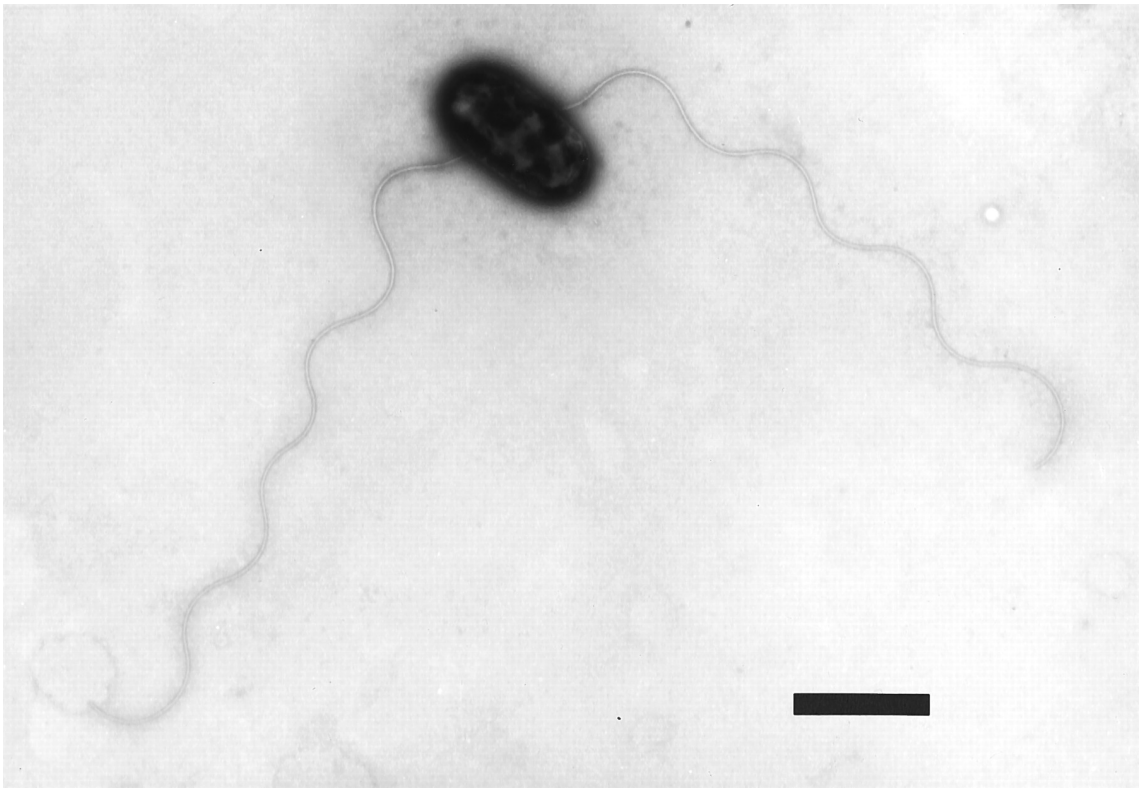


Figure 2.1: Electron micrograph of NT-26 grown in minimal medium plus arsenite (5 mM). The bar represents 1 µm.

2.1.1 Arsenic resistance system in NT-26

NT-26 has a 4.2 Mbp chromosome and two plasmids, one of which is a mega plasmid (322.264 Kb). The Ars genes in NT-26 are found both in the chromosome and the mega plasmid. The genes *arsR2a, arsR2b, arsH2, arsC2a, arsC2b, arsB2* and *arsA* are found in the chromosome and the genes *arsR1a, arsC, arsC1b* and *arsB1* are located in the mega plasmid (Figure 2.2). The expression of *arsR1a, arsC* and *arsB1*, located near the phosphate metabolism genes *pstS* and *phoB* will have their expression monitored by qPCR in this chapter.

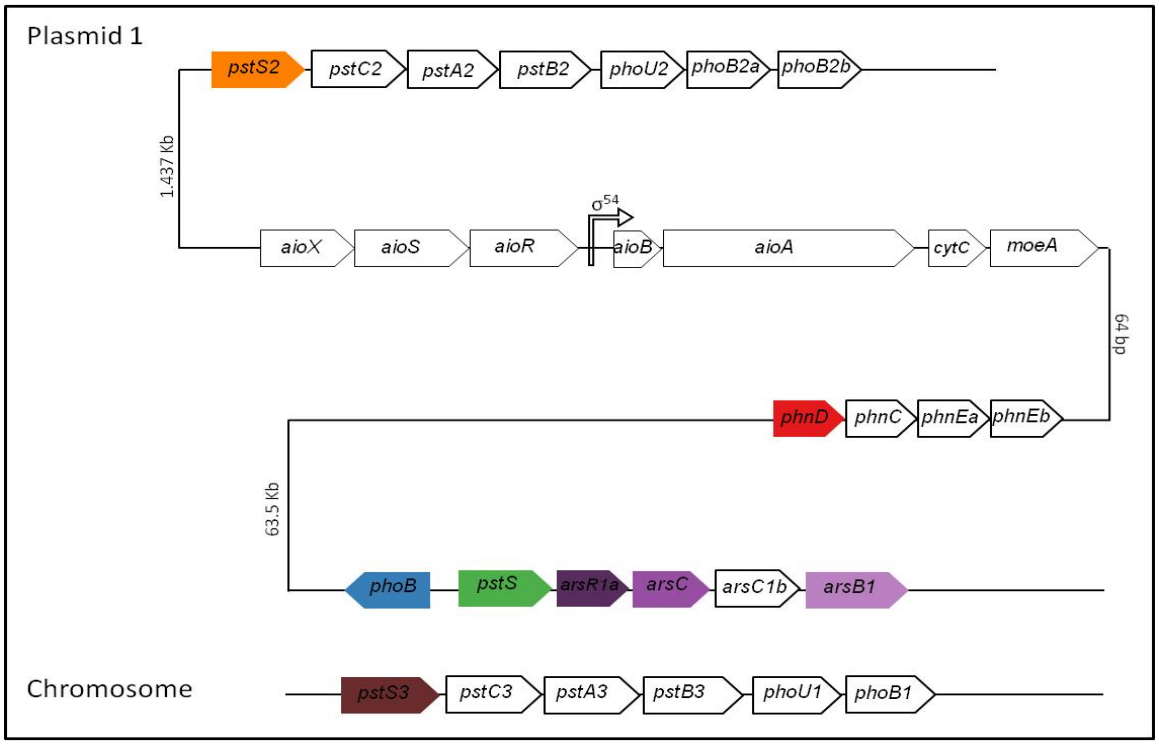


Figure 2.2: Gene arrangement in NT-26 highlighting the *ars* gene cluster *asrR1a, arsC* and *arsB1,* phosphate transporter genes *pstS, pstS2* and *pstS3,* phosphonate transporter gene *phnD,* and the Pi response gene *phoB.* The genes highlighted in the different colours will have their expression measured by qPCR.

2.1.2 Phosphate and Arsenic metabolism in NT-26

NT-26 has three different *pstS* encoding for the high affinity phosphate transporter (PstS) in different locations (Figure 2.2), one *pstS* is in the mega plasmid, near the As^{V} resistance genes *arsR1a, arsC, arsC1b* and *arsB1*, and a *phoB*; a *pstS2* in the plasmid and near genes involved in Pi sensing and metabolism including *pstC2, pstA2, pstB2, phoU2, phoB2a, phoB2b*; and another one in the chromosome – *pstS3*, also followed by genes related to Pi metabolism: *pstC3, pstA3, pstB3, phoU1* and *phoB1* (Figure 2.2). The protein sequence identity between the different PstS found in NT-26 is low, indicating that they are not duplications of the same gene (Table 2.1). The position of the *pstS* in the chromosome and plasmid, near genes either related to Pi metabolism, As resistance or As^{III} oxidation could be indicative of their relationship. It has been suggested based on *in silico* analysis that *pho, pstS* and the phosphonate transporter

phnD are often clustering with *ars* or *aioBA* genes in As^{III} oxidisers, reinforcing the idea that the phosphate metabolism and As metabolism are connected (Li et al. 2013)

.

Table 2.1. Protein sequence identity between the different PstS found in NT-26

Ptn Identity	**PstS**	**PstS2**	**PstS3**
PstS	100%	**32%**	**27%**
PstS2	**32%**	100%	**46%**
PstS3	**27%**	**46%**	100%

Transcriptomic analysis of NT-26 growing heterotrophically showed that in the presence of As^{III}, *phoR* was up-regulated by 1.32-fold, and proteomic data showed that PstS and ArsC were up-regulated by 3.7-fold and 22-fold, respectively (Andres *et al*., 2013). It is not known how *phoR, phoB, pstS, ars* genes are regulated in response to As^{V} or As^{III} in NT-26.

2.1.3 Real time PCR

To better understand the effect of As^{III} and As^{V} on the regulation of genes related to As resistance and Pi metabolism (Figure 2.2), qPCR was selected as a suitable method. When optimized to the organism and conditions of the study, qPCR can be extremely useful to observe and compare the expression of the selected genes in the conditions tested. To analyse the gene expression in an organism, it is important to rely on a sensitive and accurate method, capable of detecting changes in the expressions levels of the genes tested in different conditions.

qPCR is a powerful tool as it combines PCR with fluorescent techniques (Gibson et al. 1996, Heid et al. 1996) allowing for the detection of the amplification product as the

reaction occurs, instead of relying on the detection of the amplicon at the end of the PCR cycle as happens in conventional PCR (Freeman et al. 1999).

The detection of the PCR products by qPCR is done by detecting fluorescent signal emitted when a DNA-binding dye binds to the double strand of DNA (dsDNA) as it is formed. As the amount of DNA increases in the reaction, the fluorescent signal also rises and is detected by a specialized thermal cycler (Higuchi et al. 1992, Higuchi et al. 1993, Lee et al. 1993, Livak et al. 1995). SYBR Green I is a DNA-binding dye that binds to the minor groove of the dsDNA (Morrison et al. 1998), when in solution it only displays a background fluorescence that is increased by 1000-fold upon binding dsDNA (Dragan et al. 2012). The downside of SYBR Green I is that it binds non-specifically to dsDNA, as opposed to probes designed to bind to a specific sequence (e.g. TaqMan probe). The presence of non-specific binding can interfere with the fluorescence level and affect the accuracy of quantification. To avoid this pitfall the primers for qPCR need to be carefully designed and the melting curves of each qPCR reaction need to be analyzed to check the specificity of the reaction (Ririe et al. 1997).

After the fluorescence detection by the qPCR machine, the fluorescence levels need to be normalized to avoid errors. qPCR machines usually come with their own software that will correct for the baseline fluorescence signal, emitted by SYBR Green I alone, prior to binding the double stranded DNA. This correction is individual for each machine and dye. The fluorescence values are then translated to expression values that correspond to the PCR cycle on which the fluorescence was first detected (Cq value). This Cq value is different for each gene tested as different genes will have different expression levels.

When comparing multiple genes and conditions using qPCR it is important to take into consideration that each gene is expressed with some variation even when the condition (or treatment) doesn't change. To correct for this variation, it is imperative to use reference (Ref) genes to normalize for this variation. Ref genes are defined as any of the genes that are constitutively expressed at a relatively constant level with low expression variation between samples and conditions to be tested. The variation observed on the expression of validated Ref genes is used as a normalization factor, and internal control, when comparing the variation in expression of all the genes and conditions of the experiment (Vandesompele et al. 2002).

There are no published studies selecting or validating Ref genes in *Rhizobia* or *Agrobacteria* grown heterotrophically with yeast extract or in the presence of either As^{III} or As^{V}. Before any analysis of the gene expression in NT-26 could be carried out using qPCR, suitable Ref genes needed to be selected based on their expression stability in all conditions to be tested.

2.1.4 Aims

To understand how As^{III} and As^{V} regulate the expression of genes related to As resistance and Pi metabolism in NT-26, a quantitative real time PCR (qPCR) was used to measure the expression levels of *phnD, phoB, pstS, arsR1a, arsC, arsB2*, *pstS2* and *pstS3* in NT-26 grown in the presence and absence of As^{III} or As^{V}.

The aim of the first part of this chapter was to optimize qPCR to observe the changes on gene expression in NT-26 in the conditions tested. Different Ref were designed and tested to be used to normalize for variation in gene expression in NT-26.

The aim of the second part of this chapter was to use qPCR and the normalized Ref genes to observe the expression of As resistance genes and Pi metabolism genes in NT-26 grown heterotrophically with yeast extract and addition of either As^{V} or As^{III} until late-log and stationary phases.

2.2. Materials and Methods

2.2.1 Bacterial strain

The strain used in this work was *Rhizobium* sp. str NT-26 selected for spontaneous Rifampicin resistance (Vanden Hoven & Santini 2004).

2.2.2 Growth of NT-26

The strain NT-26 was grown heterotrophically in 10 ml of minimal salts medium (MSM) (Table 2.2) with 0.04% (w/v) yeast extract (YE) (Oxoid^{TM}) at 28^{o}C and shaking at 150 rpm (Santini et al. 2000). For experiments with As^{III} or As^{V}, final concentrations of 5 mM As^{III} (pH 8) or 5 mM As^{V} (pH 7) were added to the MSM medium, respectively.

Table 2.2. Components of minimal salts medium (MSM)

Component	Concentration (mg/l)
Solution A *	
$Na_2SO_4.10H_2O$	70
KCl	50
$MgCl_2.6H_2O$	40
CaCl2.2H2O	50
Solution B **	
KH_2PO_4	170
KNO_3	154
$(NH_4)_2SO_4$	100
$NaHCO_3$ *	0.024
SL8 trace elements **	
EDTA disodium	5.2
$FeCl_2.4H_2O$	1.5
ZnCl2	0.07
$MnCl_2.4H_2O$	0.1
H_3BO_3	0.062
$CoCl_2.6H_2O$	0.192
$CuCl_2.2H_2O$	0.017
$NiCl_2.6H_2O$	0.024
$Na_2MoO_4.2H_2O$	0.036

*Filter sterilised into a sterile bottle

**Adjust to pH7 and store unsterilised at 4^{o}C

2.2.3 Growth experiments

NT-26 growth experiments were conducted in 1 L glass bottles with 200 ml MSM containing 0.04% yeast extract in the presence and absence of either 5 mM As^{III} or 5 mM As^{V}. Experiments were started with 5% (v/v) inoculum of NT-26 grown overnight in the respective medium. All cultures were incubated at 28°C under aerobic

conditions with shaking (150 rpm). Aliquots were taken periodically and the cell density (OD_{600}) was measured at an absorbance of OD_{600}nm in a spectrophotometer (Thermo Scientific NanoDrop 2000c).

2.2.4 RNA isolation in NT-26

Total RNA was isolated from NT-26 grown heterotrophically in 10 ml MSM containing 0.04% YE, 0.04% YE and 5 mM As^{III}, and with 0.04% YE and 5 mM As^{V}. RNA was isolated from each condition in both late-log and stationary phases of growth (Section 2.3.1, table 2.4). The SV Total RNA Isolation System (Promega) was used to isolate the RNA following the manufacturer's instructions and the resulting RNA was further purified to remove contaminating DNA using the DNA-*free* Kit™ (Ambion). RNA was quantified and purity was confirmed by an *A*260/*A*280 ratio >1.8 using a Nanodrop ND-1000 (Thermo Scientific). RNA was stored at -80 °C until use.

To ensure the expression studies were statistically significant, RNA was isolated from five biological replicates for each of the conditions tested (Table 2.3). It has been suggested that the impact of biological variation can be made smaller by using a high number of biological replicates (Kitchen et al. 2010). This study used five biological replicates for each condition tested, and the qPCR was performed using triplicates of each sample. To further minimize the variation and handling error, all RNA was isolated on the same day using the same reagents.

Table 2.3. RNA purity and yield from all NT-26 cultures.

Sample	Growth phase	Code	OD_{600}	[RNA] in ng/uL	260/280
YE 0.04%	Late log	1	0.123	111.7	2.03
YE	Late log	2	0.128	166.7	2.02
YE	Late log	3	0.14	164.1	2.03
YE	Late log	4	0.132	160.8	2.01
YE	Late log	5	0.14	151.7	2
YE	Stationary	6	0.199	89.9	1.87
YE	Stationary	7	0.172	144.3	1.94
YE	Stationary	8	0.169	182.9	1.91
YE	Stationary	9	0.169	176.3	1.89
YE	Stationary	10	0.171	156.5	1.94
AsIII	Late log	11	0.138	212	2.02
AsIII	Late log	12	0.12	90.2	2
AsIII	Late log	13	0.126	167.2	2.02
AsIII	Late log	14	0.125	159.5	1.97
AsIII	Late log	15	0.126	156.1	2.05
AsIII	Stationary	16	0.219	153.9	1.94
AsIII	Stationary	17	0.247	87.5	1.89
AsIII	Stationary	18	0.219	85.7	1.89
AsIII	Stationary	19	0.214	144.2	1.9
AsIII	Stationary	20	0.231	132.5	1.91
AsV	Late log	21	0.111	67.7	2.04
AsV	Late log	22	0.123	83.3	2.03
AsV	Late log	23	0.112	50.1	1.97
AsV	Late log	24	0.12	49.5	1.97
AsV	Late log	25	0.098	53.8	1.98
AsV	Stationary	26	0.207	117	1.84
AsV	Stationary	27	0.206	84.4	1.85
AsV	Stationary	28	0.191	124.2	1.86
AsV	Stationary	29	0.194	123.4	1.89
AsV	Stationary	30	0.21	95.3	1.94

2.2.5 cDNA synthesis

cDNA was synthesised using the RevertAid Premium First Strand cDNA Synthesis kit (Thermo Scientific) and the quantitative PCR (qPCR) First Strand cDNA Synthesis protocol was followed using Random primers provided and according to the manufacturer's specifications. From each RNA preparation, 2 cDNA synthesis reactions were performed using 1 µg of total RNA in each 20 µl reaction. The expected final concentration of cDNA after the reaction was completed was 50 ng/µl. Reverse transcriptase was added to one of the reactions, and water in the place of reverse transcriptase to the other reaction as a negative control.

To minimize the variation between samples and guarantee robustness and reproducibility of the results, all the cDNA was prepared on the same day and the same amount of RNA was used (1 µg) in each reaction. As the same kit was used and the samples were prepared following the same protocol, the efficiency of cDNA synthesis is expected to be similar.

2.2.6 Quantitative polymerase chain reaction (qPCR)

qPCR reactions were performed using the PikoReal 96 Real-Time PCR System (Thermo Scientific) and the DyNAmo™ ColorFlash SYBR® Green qPCR Kit (Thermo Scientific). The reactions were prepared to a final volume of 10 µl and dispensed in a 96-well PCR plate. Each reaction mixture consisted of 5µl 2x master mix (containing hot-start DNA polymerase, SYBR® Green I, PCR buffer, 5mM $MgCl_2$, dNTP mix), 0.2 µl of 10 µM of each primer (forward and reverse), 0.6 µl of molecular biology grade water (Sigma) and 4 µl of cDNA (diluted in water to 5 ng/µl) to make 2 ng/µl final concentration of cDNA per reaction. The PCR reaction cycle used started with 7 minutes incubation at 95 °C, followed by 40 cycles of 95 °C for 10 s, 54 °C for 15 s and

72 °C for 30 s. A melting curve was prepared for each reaction following the PCR by increasing the temperature stepwise by 0.2 °C every 1 s starting at 54 °C and finishing at 95 °C. The baseline and quantification cycle (Cq) of each reaction was automatically determined using PikoReal Software version 2.1 (Thermo Scientific). Negative controls were prepared by adding water instead of cDNA for each primer set and a reverse transcriptase-negative control was also included to confirm the absence of genomic DNA in each RNA preparation. The qPCR reactions were carried out in triplicate.

2.2.7 qPCR primer design

Oligonucleotide primers were designed for NT-26 genes using the Beacon Designer™ (Premier Biosoft) to amplify PCR amplicons between 75-113 bp with melting temperatures (T_m) between 55 and 65 °C and a GC content of 50 – 60% (Table 2.4). All the primers were checked for hairpins, self-dimers and hetero-dimers using OligoAnalyzer 3.1 (Integrated DNA Technologies Inc). All primers were synthesised by Eurofins MWG Operon (Ebersberg, Germany).

Table 2.4 Primers designed for qPCR in NT-26

Primer	Sequence (5' → 3')	Gene product	Gene name	Locus tag[a]	Amplicon size (bp)
arsB1	F: ATTCCGCTGACGATCTAC R: GATAATCCGTACCCACTGA	Arsenite efflux pump	*arsB1*	NT26_p10119	76
arsC	F: GATCCGAAGTCGAAATTC R: TCAGGAATGCCGTAATAC	Arsenate reductase	*arsC*	NT26_p10117	75
arsR1a	F: AACAGGAACAAGCTATTCTC R: TTGACAAGCAACCGAAAC	DNA-binding transcriptional repressor	*arsR1a*	NT26_p10116	76
glnA	F: GAGTTCTTCCACTTCATC R: GGTAGGAGTTGACATAGG	Glutamine synthetase	*glnA*	NT26_0267	82
gltA	F: GACGGACGAGTACTTCAT R: ATCGCCTTCAGAGTGATG	Citrate synthase	*gltA*	NT26_1612	76
gyrB	F: CATTGCTTCCCTGCTGAT R: GCGAGATAGAGGTGATTCC	DNA gyrase subunit B	*gyrB*	NT26_1762	75
phoB	F:ATCTCTATGTTGATGAACGAA R: GGGTTTGTTGGAGAAGTT	Phosphate regulon transcriptional regulator	*phoB*	NT26_p10113	77
phnD	F: ATTTCAGTGGCTTTCGTC R: GAATGTCGGCAATCAGAT	Phosphonate/phosphite ABC transporter	*phnD*	NT26_p10036	91
pnp	F: CGAAGACGAGCAGTATGT R: TAGGGCGGGAAGTTGTAG	Polynucleotide phosphorylase	*pnp*	NT26_3961	78
pstS	F: AAGACGACAACATCACGAT R: TCAGTCAGGTAGGCAAGA	Phosphate-binding protein	*pstS*	NT26_p10115	79
pstS2	F: CGAGAACGACAACCTTATC R: GTAGAGGAACGAGTAGCC	Phosphate ABC transporter	*pstS2*	NT26_p10024	75
pstS3	F: GCGGAATATGCCTTTACC R: AGTTCTTGTAGGGATTGTC	Phosphate-binding protein	*pstS3*	NT26_1219	113
recA	F: CTGACCGCCTCTATCTCA R: AACATGACGCCGATCTTC	Recombinase A	*recA*	NT26_1894	77

F: Forward primer. R: Reverse primer. [a]Locus tag number for NT-26 gene sequence on the NCBI Server (http://www.ncbi.nlm.nih.gov) accession numbers FO082822.1, FO082820.1 and FO082821.1 (Andres et al. 2013) .

2.2.8 PCR amplification efficiency

Different reaction efficiencies in a qPCR experiment can affect the overall relative expression result. An efficient reaction is dependent on the primer and sample quality. To determine if the qPCR assays were optimised, a 5-fold dilution series of cDNA (25 ng, 5 ng, 1 ng, 0.2 ng and 0.04 ng of cDNA per 10 µl reaction), made from RNA isolated from a culture grown without As^{III}, in late-log phase and with a starting concentration of 25 ng, was prepared and used as the template in the qPCR. The reaction was carried out for each primer pair. A standard curve for each primer pair was created by plotting the Cq values against the logarithm of the cDNA dilutions. The PCR amplification efficiencies for each primer set tested were calculated from the slope of the corresponding standard curve (Efficiency = $10^{(-1/slope)}$-1 x 100%).

2.2.9 Data analysis

PikoReal Software (Thermo Scientific) was used to generate amplification plots that represent the accumulation of product over the duration of the entire PCR reaction, and to determine the Cq-value, when the amplification can be detected by the qPCR machine. At this point the fluorescence generated by the SYBR green binding to the double strand of DNA will be higher than the threshold line that corrects for background fluorescence, giving the Cq value. The fluorescence threshold is specific to the qPCR machine and the PikoReal Software automatically calculates it. The Cq values and PCR amplification efficiency were input into qBase+ (www.biogazelle.com/qbase) (Hellemans *et al*. 2007) a data analysis software that can be used for analysis of gene expression data from qPCR experiments. This software uses the expression of Ref genes to normalise expression levels of target

genes, based on the calculation of a normalization factor that takes in account the variation observed on the expression of the Ref genes in all conditions tested that is considered a normal variation. qBase+ software also calculates the expression level of each gene by taking in account the efficiency of each primer used; and calculates the relative expression of each gene in the conditions tested, by comparing the expression levels observed in two different treatments (i.e. NT-26 grown with YE alone compared to NT-26 grown with addition of As^{III} or As^{V}; NT-26 grown until late-log compared to stationary phase). The analyses of variance (ANOVA) were used to estimate and test interactions between the expression of each gene in the conditions tested. ANOVA allows to group samples and compare their expression against a chosen reference (i.e. expression values for NT-26 grown until late-log being reference when compared to stationary phase). The significance value (p-value) indicates whether the two conditions tested can be considered different (p-value lower than 0.05) or not different (p-value > 0.05). Another statistical test applied using qBase + software was the Spearman correlation test (Spearman 1904) that ranks the genes that were expressed in a similar fashion taking in consideration all conditions tested. Genes correlated (r-value > 0.7) are not necessarily related with regards to their functions and interactions, the results obtained by Spearman correlation test only provides an indication of which genes respond similarly to the conditions tested in this chapter.

2.3. Results

In order to understand how and if As^{III} and As^{V} regulate the expression of the *ars* genes (*arsR1a, arsC* and *arsB1*) and the phosphate metabolism genes (*phnD, phoB, pstS, pstS2* and *pstS3*) in NT-26, RNA was isolated from different growth conditions and two different growth phases and qPCR was used to quantify gene expression. To be able to use qPCR to compare the genes expression it was necessary to undertake normalization steps, such as Ref gene selection and primer optimization, specific to NT-26 and to the conditions tested to ensure robustness and reproducibility of the results.

The aim of this chapter was to understand how each gene responded to the presence of As^{III} and As^{V} and whether the growth phase had any effect on the transcription levels of the aforementioned genes.

2.3.1. Growth experiments for *Rhizobium* sp. strain NT-26

To determine how long it takes for NT-26 to reach late-log and stationary phases, a growth curve was prepared with NT-26 grown in MSM with 0.04% YE alone and with As^{III} or As^{V} (Figure 2.3).

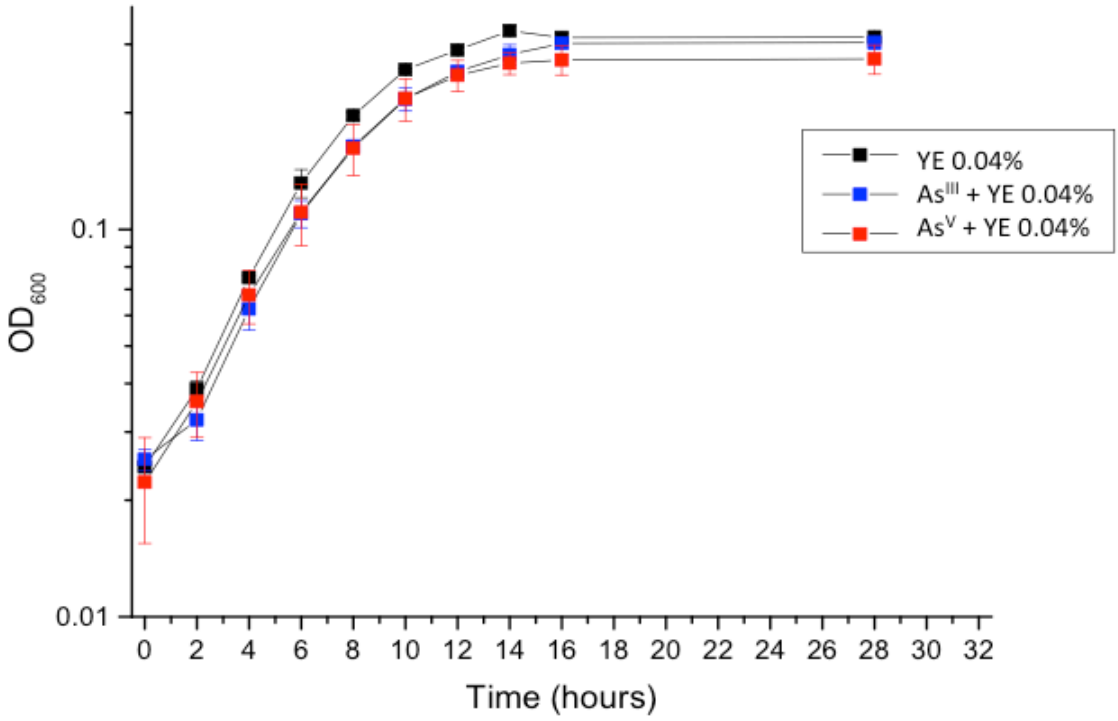


Figure 2.3. Growth curves of *Rhizobium* sp. str. NT-26. Cells were grown aerobically at 28 °C in MSM in the presence of yeast extract 0.04% (■), 5 mM As^{III} (■), 5 mM As^{V} (■). Data points are the mean of three independent cultures and the error bars are shown.

The generation times for NT-26 grown with 0.04% YE (2.53 h) and As^{III} (2.51 h) are consistent with previous observations by Santini *et al,* 2000. The generation time of the As^{V} culture was 2.5 h.

2.3.2 qPCR preparation for expression studies in NT-26

qPCR preparation was done following specific qPCR guidelines (Bustin et al. 2009) to guarantee reproducibility and the reliability of the results. The primers were designed specifically for NT-26 (refer to materials and methods section 2.6) and the most stable Ref genes for the conditions tested were used as reference/control in the expression studies.

2.3.2.1 Determining the qPCR primer efficiency

Before the Ref gene could be selected, the amplification specificity of each primer and the PCR efficiency for each primer pair needed to be checked. In an efficient reaction the amount of product detected by the qPCR machine (Cq) is relative to the amount of cDNA added to the reaction and when the Log of the quantity of cDNA is plotted against the Cq value to make a standard curve, a line is observed as exemplified in Figure 2.4 for the *recA* primers. All the primer efficiencies are summarized in table 2.5.

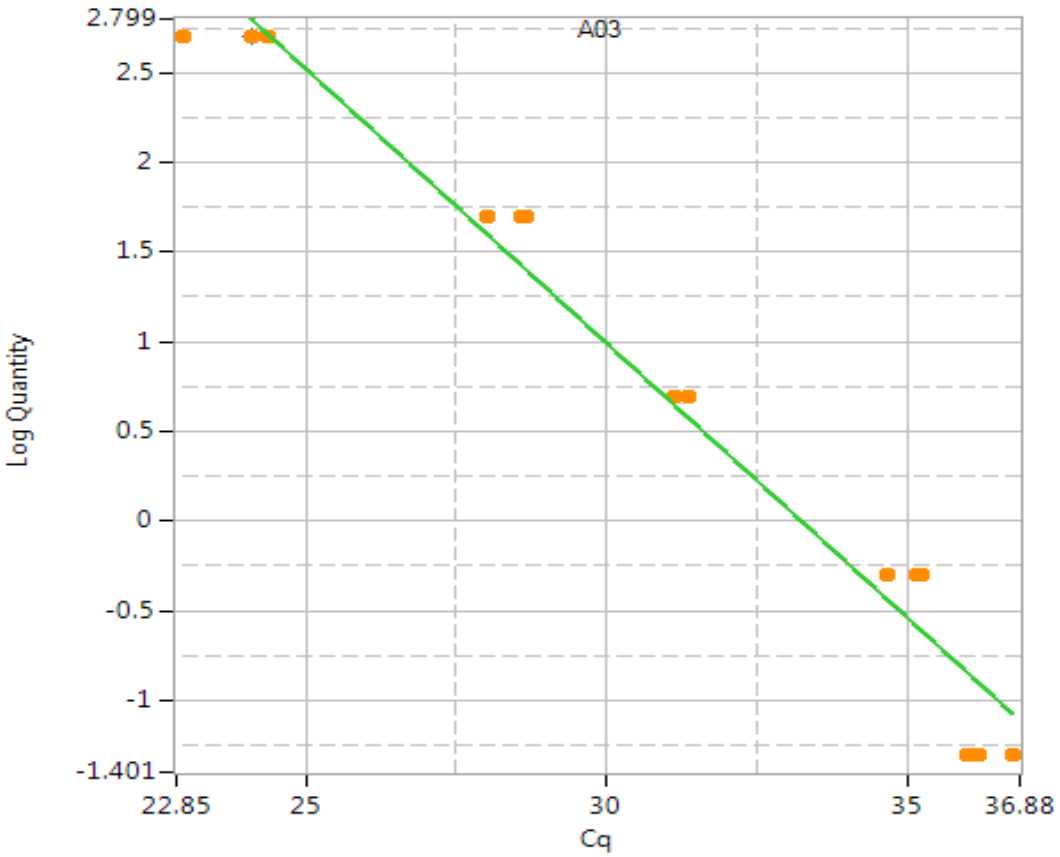


Figure 2.4. Standard curve for *recA* primer pair. The standard curve was created by plotting the Cq values against the logarithm of the cDNA dilutions. The efficiency of *recA* primer is 106%. The orange dots represent the Cq value observed for each one of the 3 replicates tested for each one of the 5 cDNA dilutions (250, 50, 10, 2, 0.4 ng cDNA).

Table 2.5. Primer efficiency calculated based on the standard curve for each primer pair.

Primer	Efficiency (%)
arsB	90.3
arsC	95.2
arsR1a	100.1
glnA	102.8
gltA	95.9
gyrB	98.7
phnD	106.7
phoB	104.3
pstS	100
pstS2	103
pstS3	100
recA	106

To check the amplification specificity by qPCR, it was important to observe the amplification plot generated during the reaction. Each reaction generates an amplification plot, which represents the accumulation of the PCR product during the reaction. When cDNA is added to a qPCR reaction and the primer specifically anneals to the cDNA, the SYBR green binds to the double strand of DNA generating

fluorescence that can be detected as the product amplifies. When enough amplicon is generated to yield a detectable fluorescent signal, the quantification cycle (Cq) can be determined (Higuchi *et al*. 1993). The Cq value is determined from the amplification plot by detecting the point in the PCR cycle at which the qPCR machine can detect fluorescence from the SYBR Green I binding to dsDNA, after correcting for background fluorescence. The example in Figure 2.5 shows the amplification plot with a Cq value of 27 when using the *gyrB* primer pair and 25 ng of cDNA in triplicate.

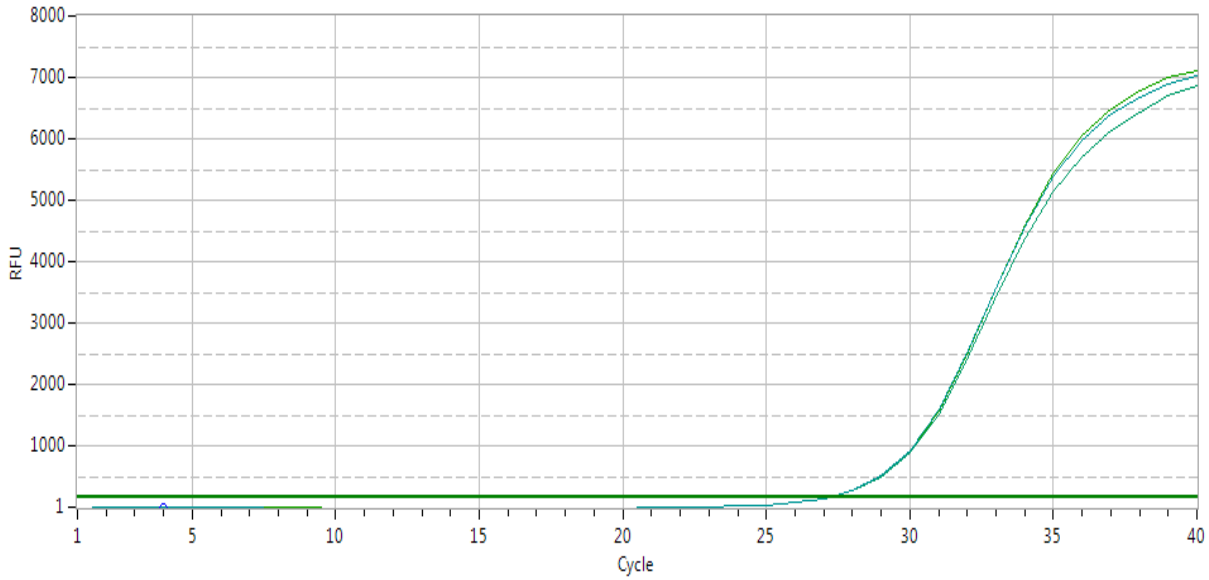


Figure 2.5. Representative amplification plot showing a Cq of 27 when using the *gyrB* primer pair and 25 ng of cDNA in triplicate. The relative fluorescence units (RFU) are shown on the y-axis, and the PCR cycle number is show on the x-axis. The green line represents the threshold.

To confirm no non-specific PCR products were amplified, product melt curves were also analysed. A single peak in the melt curve is indicative of one single product being amplified and no secondary amplification products or primer dimers are present as shown in the Figure 2.6. In the case of genomic DNA contamination or non-specific binding of the primer, multiple peaks would be observed. Non-specific binding was

observed for *pnp* primer pair that was excluded from the analysis and its efficiency couldn't be calculated.

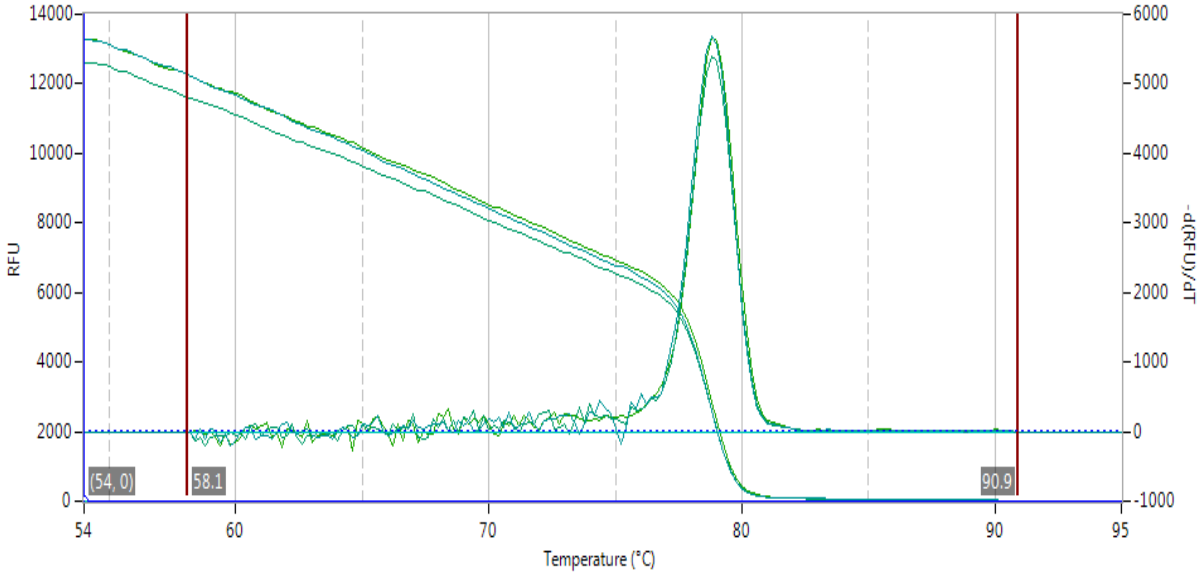


Figure 2.6: Representative melt curve graph with SYBR Green. The example shown is the *gyrB* primer pair. A melt curve was performed following PCR on the amplified product by gradual increase in temperature, from the annealing temperature (54 °C) to 95 °C while monitoring the fluorescence value every 0.2 °C. As the temperature increased, the double stranded product becomes single stranded, resulting in SYBR Green unbinding and a decrease in fluorescence. The peak is used to determine the melting point.

All the primers pairs, except from *pnp*, used for Ref gene selection and expression studies had their specificity validated after analysing the Cq values and melting curve of each primer pair.

2.3.2.2 Determination of the expression stability of reference gene candidates in NT-26

To perform qPCR comparing gene expression in different samples it is important to normalize the reactions and compensate for small variations within the tested samples and that was done by using validated Ref genes. The genes that encode for glutamine synthetase (*glnA*), citrate synthetase (*gltA*), DNA gyrase subunit B (*gyrB*),

polynucleotide phosphorylase (*pnp*) and recombinase A (*recA*) were selected to be tested as Ref genes in NT-26 grown with YE and in addition of either As^{III} or As^{V} until late-log and stationary phases. These genes have been described as housekeeping genes as they encode for proteins with essential functions in the cell (Martens et al. 2008). qPCR was carried out using cDNA obtained from all the RNA samples (table 2.3) and the variation of the expression values of *recA, glnA, gltA* and *gyrB* were analysed to find the most stable genes (i.e. those that showed no variation between the six different conditions). The *pnp* primer pair was excluded from the analysis as it formed dimers and bound non-specifically to NT-26 cDNA.

The program GeNorm (Bustin et al. 2009; Vandesompele et al. 2002) was used to compare and test the candidate Ref genes. GeNorm's algorithm is based on the principle that the expression ratio of two ideal Ref genes should be identical in all samples, independent of the treatment or condition, so the Ref gene expression stability measure (M value) is calculated as the average pairwise variation of a particular gene with all other control genes included in the analysis. High gene expression variability results in high M values and indicates low expression stability, stable genes should present a M value lower than 1.5 (Vandesompele et al. 2002). According to GeNorm analysis the M value for *gltA* was 0.965 and *gyrB* was 0.902, for *glnA* was 0.892, and *recA* was 1.535. As *recA* had M values higher than 1.5 it was excluded and after that, the M values for *gltA, gyrB* and *glnA* were 0.652, 0.668 and 0.724, respectively. When the M values are higher than 0.5, the use of multiple Ref genes is recommended for a more accurate normalisation (Hellemans et al. 2007).

The selection of stable Ref genes for NT-26 is critical to this study, as no information is available in the literature comparing and validating Ref genes used in gene

expression analyses in *Rhizobium* or *Agrobacterium* species for the conditions tested here. All the expression studies in NT-26 were performed using *glnA, gyrB* and *gltA* as Ref genes whenever the conditions YE, As^{III} and As^{V} were used.

2.3.3 Analysis of gene expression in NT-26 measured by qPCR

To determine if genes related to arsenic resistance and Pi transport (Figure 2.2) are differentially expressed in response to the presence of 5mM As^{III} or 5mM As^{V}, qPCR analyses were undertaken. The primers used were designed to amplify genes involved in As resistance (*arsR1a, arsC, arsB1*), phosphate transport (*pstS, pstS2* and *pstS3*), phosphonate transport (*phnD*) and response to low concentrations of Pi (*phoB*) (Figure 2.2).

The expression values obtained by qPCR that are being presented in this chapter were normalized to the Ref genes and compared to the expression observed for NT-26 grown with YE alone in late log-phase, unless stated otherwise.

2.3.3.1 Relative expression in NT-26 grown with As^{III} measured by qPCR

It has been shown for *A. tumefaciens* 5A and *H. arsenicoxydans* ULPAS-1 that the presence of As^{III} in the growth medium can induce the expression of the *pstS, phnD* and *phoB* genes (Cleiss-Arnold et al. 2010; Kang et al. 2012) and *arsR1a* has been described as a regulator of the operon containing *arsR1a, arsC* and *arsB1* allowing the transcription of these genes in presence of As^{III} (Jianhua Wu & Rosen 1993). Although it is not know how and if the expression of *phnD, phoB, pstS, arsR1a, arsC, arsB1*, *pstS2* and *pstS3* genes respond to the presence of As^{III} in NT-26.

The first aim was to understand if the presence of 5 mM As^{III} in the NT-26 growth medium would affect the expression of the target genes when compared to YE alone. The samples used in this analysis were obtained in the late-log phase of growth. The qPCR reactions were normalized using the Ref genes (Figure 2.7).

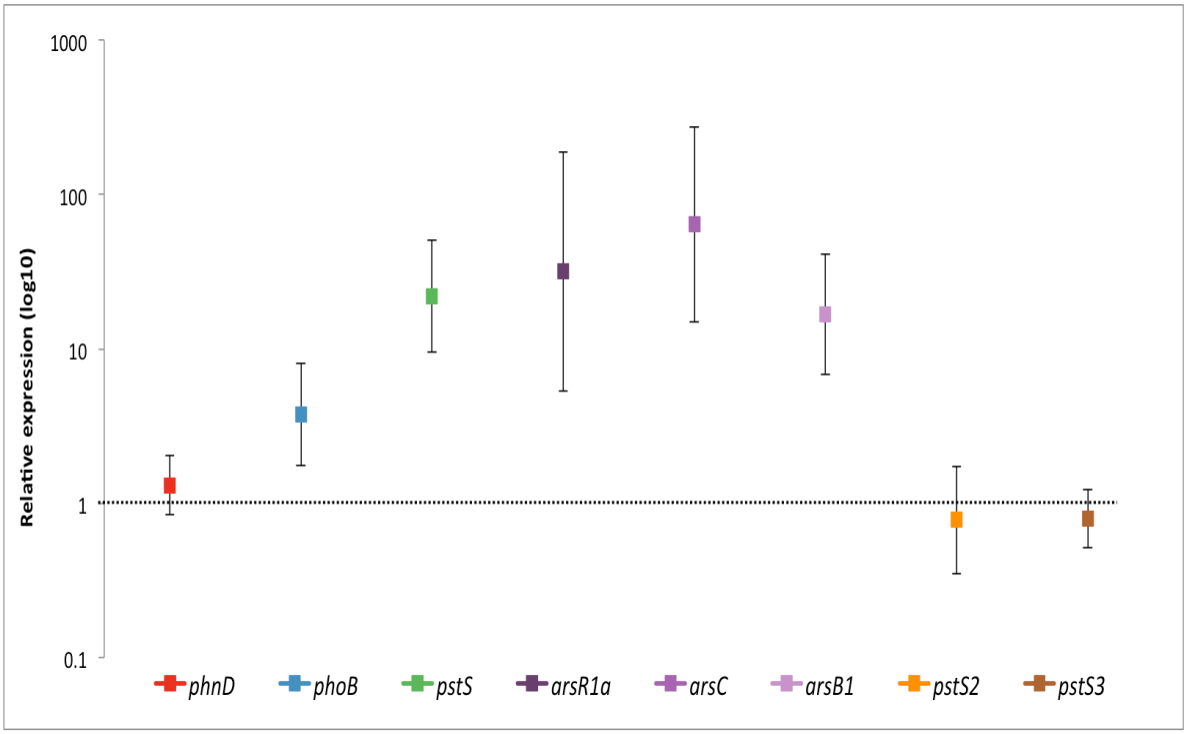


Figure 2.7 Relative expression analysis using qPCR and the software qBase + to calculate the expression ratios on the basis of the NT-26 YE samples vs As^{III} in the heterotrophic growth medium. The dotted line illustrates the threshold where the expression is not different when the two conditions are compared. All the genes above the line are up-regulated in the presence of As^{III} when compared to YE alone. The error bars show the 95% upper and lower confidence intervals.

The expression of the genes *phnD, pstS2* and *pstS3* are not affected by the presence of As^{III} in the medium compared to YE alone, however *phoB, pstS, arsR1a, arsC* and *arsB1* are all up-regulated in the presence of As^{III} by an average of 3.75-, 22-, 31-, 64- and 16-fold respectively.

2.3.3.2 Relative expression in NT-26 grown with As^V

It has been suggested that As^V can play a role in the regulation of phosphate metabolism genes due to its similarity with phosphate (Kang et al. 2014; Kang, et al. 2012). Although, it is not known if that is true for NT-26, nor if the presence of As^V will have any effect on the expression of either Ars or Pi metabolism genes.

The effect of As^V on the expression of *phnD, phoB, pstS, arsR1a, arsC, arsB1*, *pstS2* and *pstS3* was also analysed by qPCR. The relative expression was calculated by comparing the expression levels of genes in NT-26 grown until late-log phase with 5 mM of As^V with those of NT-26 grown with YE (Figure 2.8).

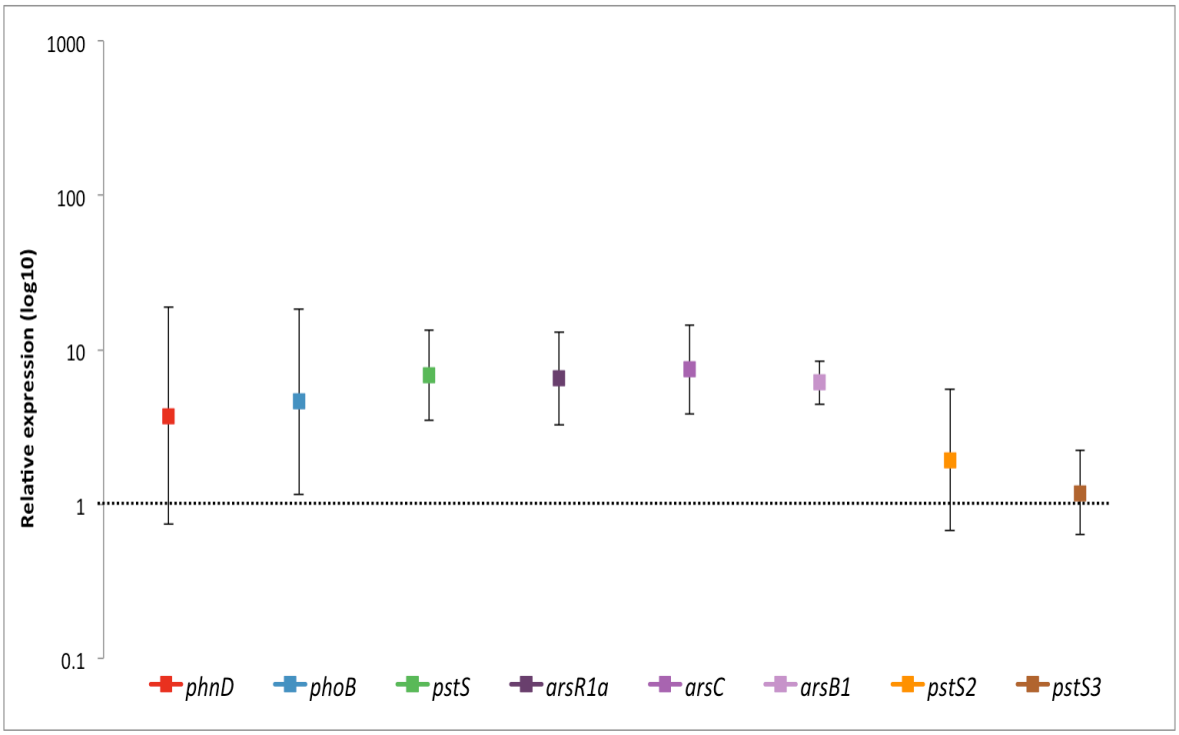


Figure 2.8. Relative expression analysis using qPCR and the software qBase+ to calculate the expression ratios on the basis of the NT-26 YE samples vs NT-26 with As^{V} in the heterotrophic growth medium. The dotted line illustrates the threshold where the expression is not different when the two conditions are compared. All the samples above the line are being up-regulated in the presence of As^{III} when compared to YE alone. The error bars show the 95% upper and lower confidence intervals; any error bar crossing the threshold line indicates that the expression in the two conditions tested is not significantly different.

qPCR analysis showed that the expression of *phnD, phoB, pstS2* and *pstS3* genes were not significantly up or down-regulated following growth with As^{V} after checking the confidence interval statistics. Although *phoB* is slightly above the threshold line, it is not significantly different when compared the two conditions with p-value of 0.14 (higher than the p–value < 0.05, considered for statistical significant differences). *pstS, arsR1a, arsC* and *arsB1* were up-regulated by 6.8-, 6.5-, 7.4- and 6.1-fold respectively.

2.3.3.3 Growth phase comparison

It is known that genes can be differentially expressed depending on the bacterial growth, with genes up-regulated in exponential phase and others up-regulated in stationary phase of growth (Chang et al. 2002). Understanding when the genes are expressed could provide some clues about the way a cell regulates its metabolism. To better understand the expression of *phnD, phoB, pstS, arsR1a, arsC, arsB1*, *pstS2* and *pstS3* in NT-26, qPCR analysis was undertaken with samples obtained from NT-26 grown with YE alone until late-log and stationary phases (table 2.3) and the expression levels of each gene were compared (Figure 2.9).

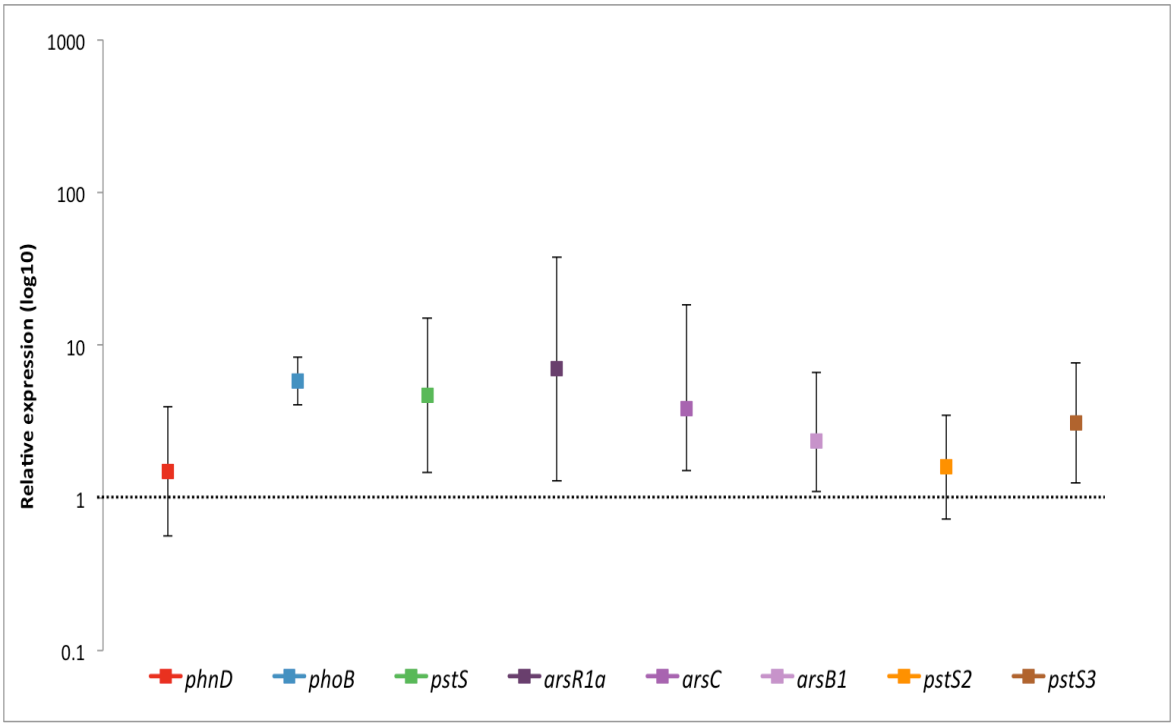


Figure 2.9. Relative quantities of the tested genes in NT-26 samples grown with YE until stationary phase compared to late log phase. The dotted line illustrates the threshold where the expression is not significantly different when comparing the two conditions ($p>0.05$). All the samples above the line were up-regulated in stationary phase of growth compared to late-log phase. The error bars show the 95% upper and lower confidence intervals; any error bar crossing the threshold line indicates that the expression in the two conditions tested is not significantly different.

It is noticeable that in stationary phase the genes *phoB, pstS, arsR1a, arsC, arsB1* and *pstS3* were up-regulated by an average of 5.8-, 4.7-, 7-, 3-, 2.3 and 3-fold, respectively. This indicates that these genes are more expressed in stationary phase than in late-log phase of growth. No difference in expression between the samples obtained in late-log or stationary phases was observed for *phnD* and *pstS2*.

When As^{III} is added to the growth medium, the genes were expressed slightly different compared to medium with YE alone, from what was observed previously (Figure 2.7). When comparing NT-26 grown with As^{III} until late log with stationary phase, the only gene up-regulated was *phoB* by 6-fold as observed on figure 2.10, below:

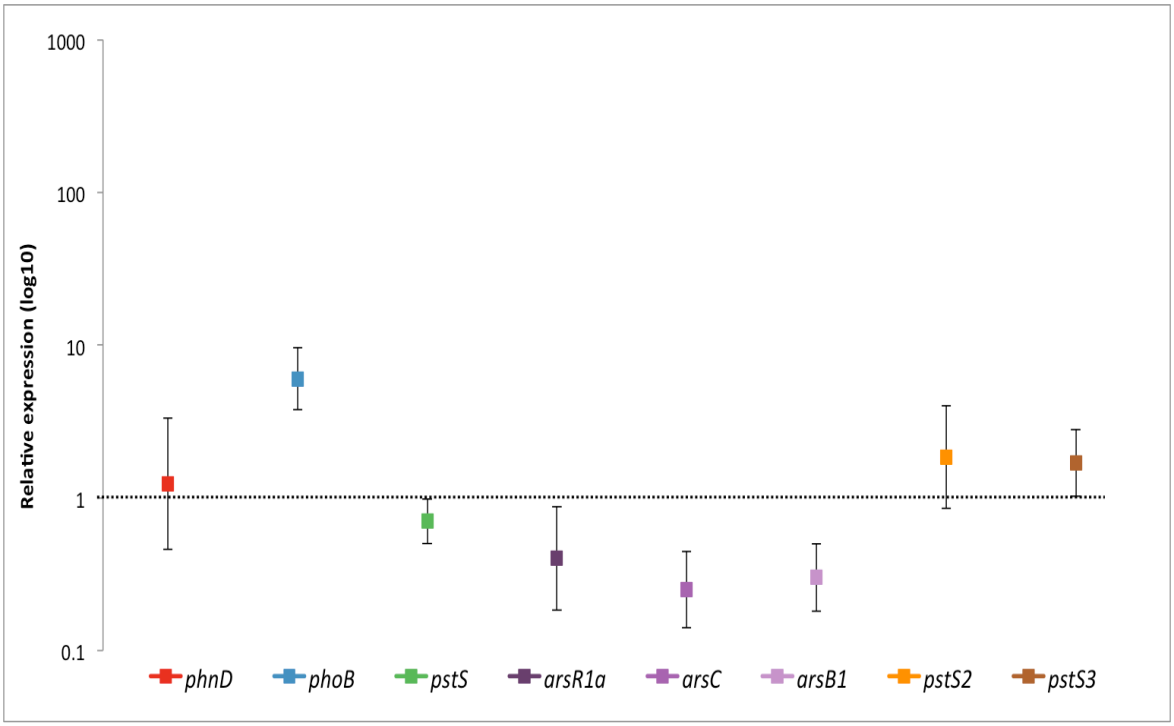


Figure 2.10. Relative quantities of the tested genes in NT-26 samples grown with As^{III} until stationary phase compared to late log phase. The dotted line illustrates the threshold where the expression is not significantly different when comparing the two conditions ($p>0.05$). All the samples above the line were up-regulated in stationary phase of growth compared to late-log phase. The error bars show the 95% upper and lower confidence intervals; any error bar crossing the threshold line indicates that the expression in the two conditions tested is not significantly different.

The genes *phnD, pstS, arsR1a, arsC, arsB1, pstS2* and *pstS3* do not show significant variation ($p>0.05$) in expression when comparing late-log with stationary phase of growth in media containing As^{III} (Figure 2.10).

The same comparison was made for NT-26 grown with As^{V} until late-log and stationary phases (Figure 2.11). It is important to keep in mind that As^{V} is reduced to As^{III} in the cytoplasm and ArsB then extrudes As^{III} from the cell; therefore, during growth, the concentration of As^{V} in the media will decrease while the concentration of As^{III} will increase. The effect of As^{V} in stationary phase might not correspond to the presence of As^{V} but to the accumulation of As^{III}.

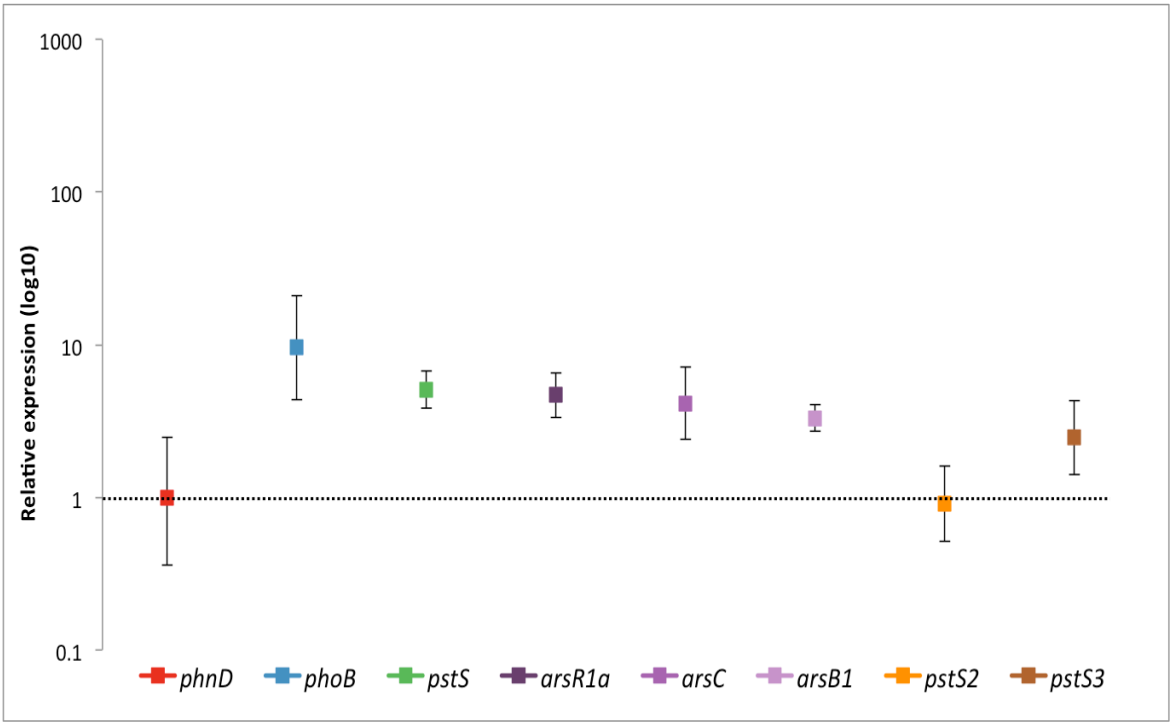


Figure 2.11. Relative quantities of the tested genes in NT-26 samples grown with As^V until stationary phase compared to late log phase. The dotted line illustrates the threshold where the expression is not significantly different when comparing the two conditions ($p>0.05$). All the samples above the line were up-regulated in stationary phase of growth compared to late-log phase. The error bars show the 95% upper and lower confidence intervals; any error bar crossing the threshold line indicates that the expression in the two conditions tested is not significantly different.

The genes *phoB, pstS, arsR1a, arsC* and *arsB1* were up-regulated in stationary phase when NT-26 was grown with As^V by an average of 9.6-, 5.1-, 4.7-, 4.1- and 3.3-fold respectively. The expression of *phnD, pstS2* and *pstS3* did not vary in stationary phase in comparison to late-log phase of growth in medium containing As^V.

2.3.3.4 Correlation between the gene expression

Apart from having different levels of expression under the different conditions tested, some genes have a similar expression pattern indicating that they are required at the same time or that they respond to a similar stimulus (when part of the same stimulon), behaving as a group. That happens for genes in the same operon but not necessarily for genes in different operons. To identify which genes behave as

a group, correlation tests were undertaken. Correlation values can vary from 1 to -1 where 1 is the maximum value for positive correlation and -1 is found when two genes behave in the opposite way under the same conditions, therefore they are negatively correlated. The correlation is significant when the p value is <0.05 and the r value is > 0.7 (Spearman 1904) .

Taking in to consideration all the conditions tested – NT-26 with As^{III}, As^{V}, YE alone in late-log and stationary phases, a Spearman correlation test (Spearman 1904) was performed and the results are shown in Table 2.6. An example of two genes whose expression was highly correlated is shown on the Figure 2.12 for the genes *arsC* and *pstS* and an example of two genes whose expression was low correlated is shown in Figure 2.13 for *pstS3* and *arsB1*.

Table 2.6: Correlation values (r) between two genes expressed under all conditions tested. All values presented in bold are statistically significant ($p<0.05$) and have a R value higher than 0.7:

target x	target y	correlation value (r)	significance (p-value)
arsC	***pstS***	**0.938**	**8.45E-07**
arsB1	***arsC***	**0.912**	**2.40E-06**
arsB1	***pstS***	**0.896**	**2.72E-06**
arsB1	***arsR1a***	**0.77**	**1.34E-05**
phoB	***pstS***	**0.768**	**1.34E-05**
arsR1a	***pstS***	**0.737**	**3.17E-05**
arsR1a	***arsC***	**0.726**	**4.07E-05**
phnD	*phoB*	0.657	4.01E-04
phnD	*pstS2*	0.632	7.65E-04
arsR1a	*phoB*	0.587	2.27E-03
phoB	*pstS2*	0.58	2.46E-03
phoB	*arsC*	0.573	2.59E-03
arsB1	*phoB*	0.571	2.59E-03
phoB	*pstS3*	0.527	6.20E-03
phnD	*pstS*	0.519	6.89E-03
pstS2	*pstS3*	0.452	2.25E-02
phnD	*pstS3*	0.422	3.47E-02
arsB1	*phnD*	0.418	3.49E-02
pstS3	*pstS*	0.355	8.09E-02
phnD	*arsC*	0.346	8.55E-02
arsR1a	*pstS2*	0.335	9.37E-02
pstS2	*pstS*	0.285	1.60E-01
arsB	*pstS3*	0.274	1.73E-01
arsR1a	*phnD*	0.265	1.82E-01
arsB1	*pstS2*	0.249	2.04E-01
arsR1a	*pstS3*	0.229	2.38E-01
pstS3	*arsC*	0.16	4.11E-01
pstS2	*arsC*	0.112	5.52E-01

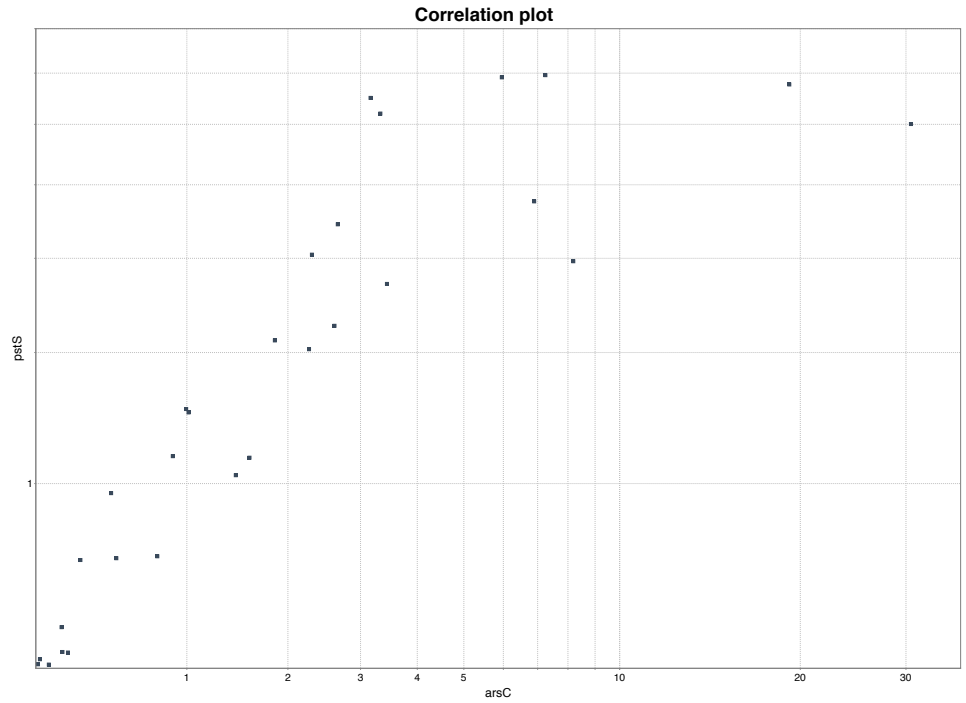


Figure 2.12. Correlation plot of the genes *pstS* and *arsC*. The expression levels of *arsC* is shown on x-axis and were plotted against *pstS* in y-axis. The r-value for *pstS* and *arsC* is 0.938 with a p-value < 0.05.

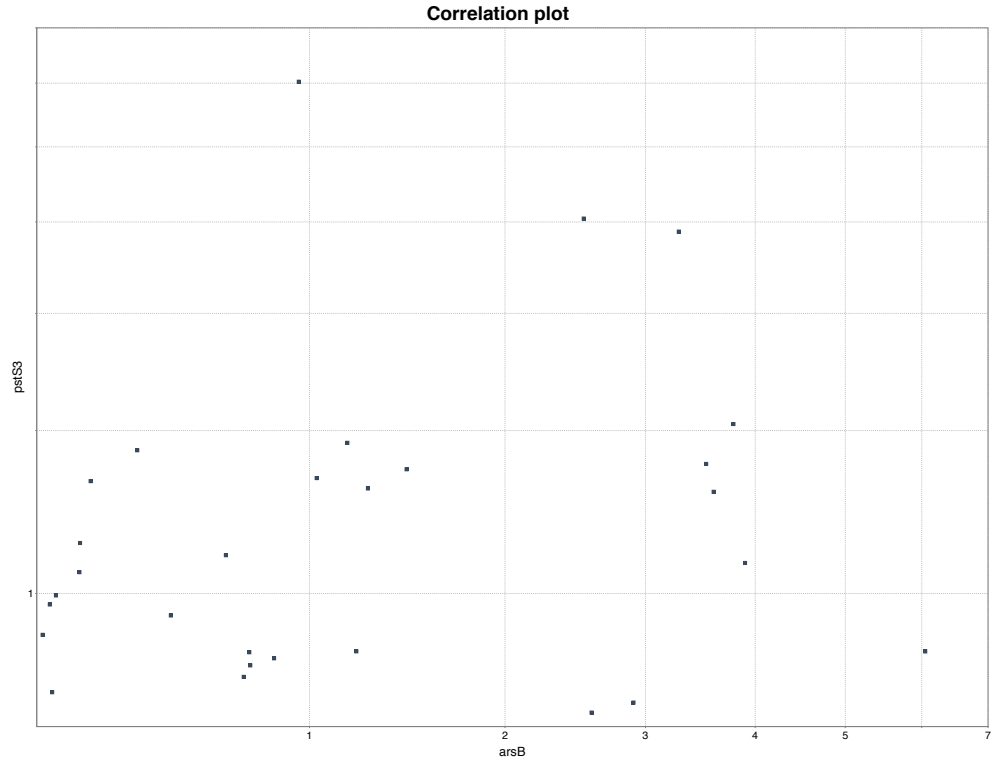


Figure 2.13. Correlation plot of the genes *pstS3* and *arsB1*. The expression levels of *arsB*1 is shown on x-axis and were plotted against *pstS3* in y-axis. The r-value for *pstS3* and *arsB1* is 0.274 with a p-value = 0.173.

As expected, the genes involved in As resistance – *arsR1a, arsC* and *arsB1* – respond similarly to either the presence or absence of As^{III} and As^{V} and have their expression levels fluctuate in a similar fashion with correlation values higher than 0.7 and were statistically significant ($p<0.05$). The expression of the phosphate transporter gene *pstS* is also correlated to the *ars* genes, but only the *pstS* closest to the *ars* operon and not *pstS2* or *pstS3.* The gene *phoB* has its expression correlated to *pstS* but not to the *ars* genes.

2.4. Discussion

In this chapter, qPCR was used to monitor and compare the expression levels of the genes *phnD, phoB, pstS, arsR1a, arsC, arsB1*, *pstS2* and *pstS3* in NT-26 grown in different conditions - YE alone, As^{III} and As^{V} - until late-log and stationary growth phases. qPCR is a useful technique to observe and compare gene expression and normalisation steps, specific to the organism and conditions tested, need to be followed to allow reproducibility and significance to the results obtained.

2.4.1 Real time PCR – Reference gene selection and sample preparation

For the qPCR experiments presented here, the SYBR Green fluorescence detection method was chosen as it provided a better cost-benefit when compared to TaqMan probes as no specific probe needed to be designed for each gene to be tested, making the experimental process simpler and cost effective. It has been reported that in qPCR using TaqMan probes that the expression level could be overestimated (Cao & Shockey 2012), although another study claims that there was no significant difference in the sensitivity and reproducibility of the results when comparing the two methods (Maeda et al. 2003). For SYBR Green to be considered a good method,

it is extremely important that the primers are designed to amplify products between 75-220 bp (Taylor et al. 2010) and that the melting curve is analysed thoroughly to detect any non-specific binding. All the primers used in this chapter were checked for specific binding and the amplification efficiencies varied and were within the acceptable range – from 80-110% (Taylor et al. 2010; Bustin et al. 2009).

The selection of Ref genes is the first step taken to normalize qPCR and make the analysis specific for NT-26 and the conditions tested. Normalization using Ref genes compensates for variations in sample amount, extraction yield, reverse transcription efficiency and RNA quality. A good set of Ref genes should be stably expressed in all conditions tested. Five different genes encoding for proteins that are essential were chosen, as it was believed to be constitutively expressed in NT-26 as they encode for proteins with housekeeping functions. The expression level of the genes that encode for glutamine synthetase (*glnA*), citrate synthetase (*gltA*), DNA gyrase subunit B (*gyrB*), polynucleotide phosphorylase (*pnp*) and recombinase A (*recA*) were monitored in NT-26 grown with As^{III}, As^{V} and YE alone until late-log and stationary phases. The optimal Ref genes were chosen based on their expression levels and the results were ranked using the computational program GeNorm (Galisa et al. 2012; Vandesompele et al. 2002; Vandesompele et al. 2009; Bustin et al. 2009), a highly cited algorithm (10,000 citations according to Google Scholar, to date) to determine the most stable Ref gene between the candidates. The genes *glnA, gltA* and *gyrB* had lower expression stabilities with M values of 0.724, 0.652 and 0.668 respectively. According to Vandesompele et al, 2002, stable genes should have a M value lower than 1.5 and when the M values are higher than 0.5 more than two Ref genes should be used (Hellemans et al. 2007).

2.4.2 Gene expression analysis using qPCR in NT-26 grown with As^{III}

The expression levels of the genes *phnD, phoB, pstS, arsR1a, arsC, arsB1*, *pstS2* and *pstS3* in NT-26 were monitored using qPCR. The relative expression results were obtained comparing the expression levels for the samples with As^{III} with the expression in the samples containing YE alone (As^{III} expression value/YE expression value), and the results were corrected by taking the Ref genes variation in account to normalize for "normal fluctuations" in gene expression and inter sample variability.

Summarized in Table 2.7 are the relative expression values for all the samples analysed and presented in the figure 2.7. The table shows the confidence interval (95 % low – 95% high) and the p-values calculated to know if the expression values in the to conditions tested are significantly different (p-value < 0.05) or not (p-value > 0.05).

Table 2.7. Relative expression values for samples grown until late-log phase with As^{III} compared to YE:

	As^{III}	YE	
gene	**Expression (95% low - 95% high)**		**p-value**
phnD	**1.308** (0.838 - 2.042)	1 (0.563 - 1.777)	0.5258
phoB	**3.754** (1.756 - 8.026)	1 (0.746 - 1.34)	0.01488
pstS	**22.011** (9.539 - 50.793)	1 (0.825 - 1.212)	0.01488
arsR1a	**31.625** (5.341 - 187.279)	1 (0.649 - 1.541)	0.01488
arsC	**64.212** (14.96 - 275.619)	1 (0.566 - 1.767)	0.01488
arsB1	**16.714** (6.784 - 41.182)	1 (0.78 - 1.283)	0.01488
pstS2	**0.779** (0.349 - 1.738)	1 (0.779 - 1.283)	0.8413
pstS3	**0.793** (0.514 - 1.223)	1 (0.718 - 1.394)	0.4221

The genes *phoB, pstS, arsR1a, arsC* and *arsB1* were up-regulated in NT-26 when it was grown with 5mM As^{III}, the expression values for those genes in presence of As^{III} are statistically different when compared to YE alone. The statistical significance is measured by the p-value, when p-value < 0.05 it means that the probability of the

expression being equal in the two conditions tested is lower than 5% (Hung et al. 1997). The genes *phnD, pstS2* and *pstS3* are neither up- or down-regulated when NT-26 was grown with As^{III}, meaning that their expression values were similar in the two conditions tested (As^{III} vs YE).

The genes *arsR1a, arsC,* and *arsB1* were expected to be up-regulated in NT-26 grown heterotrophically with YE and As^{III} as ArsR was described in *E. coli* as a transcriptional repressor that regulates the expression of the *arsR1a, arsC* and *arsB1* operon in the presence of As^{III} (Rosen 1999; Carlin et al. 1995; G. Ji & Silver 1992). In *E. coli,* ArsR acts to prevent transcription of its own operon by binding to the DNA in the regulatory region, in presence of As^{III}, ArsR releases the DNA to allow the transcription to occur (Rosenstein et al. 1992).

The up-regulation of *phoB* and *pstS* in the presence of As^{III} is in accordance to what has been observed in the As^{III}-oxidisers *Halomonas* sp. HAL1, *H. arsenicoxydans* ULPAS-1 and *A. tumefaciens* 5A (Cleiss-Arnold et al. 2010; Chen et al. 2015; Kang et al. 2012; Kang et al. 2014). In *A. tumefaciens* 5A it was shown that PhoB is required for normal transcription of arsenite oxidase genes (*aio* genes) and that *pstS* and *phoB* can be regulated by ArsR1 in response to As^{III} (Kang et al. 2012). It is important to point out that PhoB regulated genes have a consensus sequence (CTGTCAT-N_4-CTGTCAT) in the promoter region known as a PhoBox (Makino et al. 1988). The PhoBox was not identified upstream of the *ars* operon*, pstS* or *aio* genes. This could indicate that if PhoB is participating in the regulation of those genes in NT-26, then this regulation might be indirect and PhoB binds elsewhere on the genome that can interfere with the activation of those genes or the consensus PhoBox sequence is different in NT-26.

The expression of genes *pstS2, pstS3* and *phnD* were not altered when NT-26 was grown with As^{III} in addition to YE, while *pstS* and *phoB* were up-regulated in the same conditions. This result suggests that *pstS* and *phoB* might play a role on the ability of NT-26 to resist to As. This would make sense if the proximity of *pstS* and *phoB* to the *ars* genes is taken into consideration as the proximity between genes often indicates relationship (Kang et al. 2014). The genes *pstS2, pstS3* and *phnD* are organized in operons containing several other phosphate metabolism genes, suggesting that their function is related to phosphate and phosphonate metabolisms, respectively (Figure 2.2).

2.4.3 Gene expression analysis using qPCR in NT-26 grown in presence of As^{V}

The results obtained using qPCR to monitor the expression of *phnD, phoB, pstS, arsR1a, arsC, arsB1*, *pstS2* and *pstS3* grown heterotrophically until late-log phase with YE and As^{V} differed from those observed for NT-26 grown until late-log phase with YE and As^{III}. The relative expression values for NT-26 samples grown until late-log phase with As^{V} compared to YE alone, first shown in Figure 2.8, are summarised in Table 2.8. The table shows the confidence interval (95 % low – 95% high) and the p-values calculated to know if the expression values in the two conditions tested (presence of As^{V} and YE *versus* YE alone) are significantly different.

Table 2.8. Relative expression values for samples grown until late-log phase with YE and As^{V} compared to YE alone:

	As^{V}	YE	
gene	Expression (95% low - 95% high)		p-value
phnD	**3.73** (0.742 - 18.762)	1 (0.563 - 1.777)	0.1919
phoB	**4.601** (1.153 - 18.369)	1 (0.746 - 1.34)	0.1481
pstS	**6.868** (3.49 - 13.484)	1 (0.825 - 1.212)	0.01587
arsR1a	**6.505** (3.26 - 12.98)	1 (0.649 - 1.541)	0.01587
arsC	**7.451** (3.836 - 14.473)	1 (0.566 - 1.767)	0.01587
arsB1	**6.123** (4.424 - 8.476)	1 (0.78 - 1.283)	0.01587
pstS2	**1.932** (0.672 - 5.55)	1 (0.779 - 1.283)	0.1919
pstS3	**1.165** (0.631 -2.217)	1 (0.718 - 1.394)	1

The genes *pstS, arsR1a, arsC* and *arsB1* were up-regulated when As^{V} was present along with YE, compared to YE alone. When comparing this result to the one obtained for As^{III} (Figure 2.6 and Table 2.7), it is noticeable that the expression levels of *pstS, arsR1a, arsC* and *arsB1* with As^{III} are higher than with As^{V}. This result makes sense for the *ars* genes, as the ArsR1a has been described in *E. coli* as a repressor that responds to As^{III} instead of As^{V} and in presence of As^{III} it allows the expression of *arsC* and *arsB1*.

PstS, PstS2 and PstS3 are Pi transporters and because As^{V} is an analogue of phosphate, it might be expected for As^{V} to participate in the regulation of *pstS, pstS2* or *pstS3*. On the other hand, as mentioned before (section 2.3.4.1), As^{III} has been shown in *A. tumefaciens* 5A to activate the transcription of *pstS1* through ArsR1 activation but not the other *pstS* present on the genome (Kang et al. 2012). According to the results presented here in this study this might be happening in NT-26 as well, although, further studies need to be done to better understand the role PstS plays in NT-26 and if ArsR1a is indeed regulating the expression of *pstS* and *phoB.* The genes *phoB, phnD, pstS2* and *pstS3* were neither up- nor down-regulated in NT-26 grown

with As^{V} in addition to YE, and apart from *phoB,* this result is similar to what was observed when As^{III} was present in addition to YE in the growth medium. Their expression values for samples grown with YE in addition of As^{V} are not significantly different from the expression values observed for the samples obtained on medium containing YE alone, especially when analysing the standard deviation of the expression values of YE and As^{V} as presented on Table 2.8.

2.4.4 qPCR in NT-26 grown until late-log and stationary phases.

To observe the gene expression over the course of growth and understand whether the growth phase plays any role on the regulation of As resistance system genes (*arsR1a, arsC* and *arsB1)* or phosphate metabolism (*phnD, phoB, pstS, pstS2* and *pstS3)*, NT-26 was cultivated in presence of YE and with addition of either As^{III} or As^{V} and until late-log and stationary phases and the expression of the genes was monitored by qPCR. The results presented in Figures 2.8, 2.9 and 2.10 are summarized in the Tables 2.9, 2.10 and 2.11, below with the respective confidence intervals and p-values:

Table 2.9. Relative expression values for samples grown with YE until late-log compared to stationary:

	Late-log	Stationary	
gene	**Expression (95% low - 95% high)**		**p-value**
phnD	1 (0.563 - 1.777)	**1.49** (0.563- 3.942)	0.5983
phoB	1 (0.746 - 1.34)	**5.831** (4.056 - 8.383)	0.04545
pstS	1 (0.825 - 1.212)	**4.686** (1.459 - 15.055)	0.02778
arsR1a	1 (0.469 - 1.541)	**7.002** (1.292 - 37.953)	0.02778
arsC	1 (0.566 - 1.767)	**3.855** (0.81 - 18.356)	0.0455
arsB1	1 (0.78 - 1.283)	**2.352** (0.837 - 6.61)	0.02778
pstS2	1 (0.779 - 1.283)	**1.59** (0.723 - 3.493)	0.08333
pstS3	1 (0.718 - 1.394)	**3.093** (1.248 - 7.661)	0.01389

Table 2.10. Relative expression values for samples grown with As^{III} until late-log compared to stationary:

	Late-log	Stationary	
gene	**Expression (95% low - 95% high)**		**p-value**
phnD	1 (0.641 - 1.561)	**1.24** (0.46 - 3.338)	0.6905
phoB	1 (0.468 - 2.138)	**6.003** (3.759 - 9.587)	0.02778
pstS	1 (0.433 - 2.308)	**0.701** (0.502 - 0.979)	0.2393
arsR1a	1 (0.169 - 5.922)	**0.4** (0.184 - 0.87)	0.2393
arsC	1 (0.233 - 4.292)	**0.25** (0.141 - 0.445)	0.1333
arsB1	1 (0.406 - 2.464)	**0.302** (0.181 - 0.503)	0.05556
pstS2	1 (0.448 - 2.232)	**1.857** (0.854 - 4.035)	0.2393
pstS3	1 (0.48 - 1.542)	**1.698** (1.024 - 2.815)	0.1333

Table 2.11. Relative expression values for samples grown with As^{V} until late-log compared to stationary:

	Late-log	Stationary	
gene	**Expression (95% low - 95% high)**		**p-value**
phnD	1 (1.999 -5.03)	**1.126** (0.485 - 2.614)	0.8413
phoB	1 (0.25 - 3.992)	**9.582** (4.352 - 21.097)	0.0127
pstS	1 (0.509 - 1.963)	**5.102** (3.844 - 6.773)	0.0127
arsR1a	1 (0.501 - 1.996)	**4.723** (3.339 - 6.563)	0.0127
arsC	1 (0.515 - 1.942)	**4.132** (2.402 - 7.109)	0.0127
arsB1	1 (0.722 - 1.384)	**3.31** (2.715 - 4.035)	0.0127
pstS2	1 (0.348 - 2.873)	**0.905** (0.515 - 1.592)	0.8413
pstS3	1 (0.526 - 1.902)	**2.456** (1.411 - 4.275)	0.07407

Analysing the data together gives a better view of the expression of each gene in stationary phase in the presence or absence of As^{III} and As^{V}. The gene *phoB* was up-regulated in stationary phase under all conditions tested, indicating that although As^{III} and As^{V} cause up-regulation of this gene in late-log phase, it is also subjected to another regulatory mechanism that leads to its up-regulation in stationary phase; this is confirmed by the fact that in the absence of either As^{III} or As^{V} it is still possible to see up-regulation of *phoB* by 5.8-fold. PhoB is a regulator active in response to low concentrations of Pi in the medium. It has been observed that in stationary phase of

growth, the cell enters a starvation period due to a decrease of nutrients in the growth medium, leading to the activation of the expression of genes responsible for the uptake of essential nutrients and compounds such as Pi (Marzan & Shimizu 2011). *pstS3* is more present in stationary phase when NT-26 is grown with YE alone (Table 2.9) as *pstS3* is up-regulated by 3-times compared to late-log phase; this result suggests that PstS3 can be used for the uptake of Pi when the cell goes through a starvation period, although more experiments need to be done to confirm this, especially as *pstS2* is not up-regulated in stationary phase and *pstS* is only up-regulated in stationary phase in presence of As^{V} and YE alone. It is not clear how NT-26 uses its phosphate transporters PstS, PstS2 and PstS3 or if all of them are specific to the transport of Pi. It seems that they have different functions as *pstS* has increased expression in presence of As^{III} and As^{V}. As shown in table 2.1, the low sequence identity between PstS, PstS2 and PstS3 contributes to the hypothesis that PstS, PstS2 and PstS3 might have different functions and their expression responds to different stimuli in NT-26.

The up-regulation of *pstS* in medium containing As^{V} can also be explained by the fact that ArsC reduces As^{V} to As^{III} in the cytoplasm followed by ArsB pumping As^{III} out of the cell, leading to an accumulation of As^{III} in the medium in stationary phase, when NT-26 would have already reduced most of the As^{V} that was present in the media at the start. However, the concentration of neither As^{III} nor As^{V} was measured in the NT-26 growth medium over the course of growth and without that, is not possible to draw any conclusions about the accumulation of As^{III} following As^{V} reduction in stationary phase. The presence of As^{III} up-regulates the expression of *pstS*, as well as *arsR1a, arsC* and *arsB1* that are also up-regulated in stationary phase in media

containing As^{V}. These four genes are not up-regulated in stationary phase in media containing As^{III} and this is probably due to all the As^{III} in the media already being oxidized to As^{V} and shown by Santini *et al* for NT-26 (Santini et al. 2000), suggesting that it is As^{III} and not As^{V}, responsible for the up-regulation of those genes. More experiments need to be done to confirm this hypothesis, the concentration of As^{III} and As^{V} in the growth media during the course of growth needs to be measured to confirm that in stationary phase the As^{III} was already oxidized to $As^{V,}$ as shown previously for NT-26 (Santini et al. 2000), and to determine if the As^{V} was reduced to As^{III} as suggested here.

2.4.5 Correlation between gene expression levels

The expression levels of *phnD, phoB, pstS, arsR1a, arsC, arsB1*, *pstS2* and *pstS3* vary in different ways according to the condition tested - with As^{III}, As^{V} and growth phase. Some genes showed similar expression levels in response to the same conditions, indicating that their expression is correlated (Table 2.6), meaning that whenever one gene is up-regulated, down-regulated or not different, the other genes will also be.

The expression of the genes *pstS, arsR1a, arsC* and *arsB1* are all correlated; That can indicate two different things: 1- they are in the same operon, meaning they are co-transcribed and under the control of the same promoter in the conditions tested or 2 – they are not in an operon but their promoters respond to the conditions tested in this work, in a similar fashion. The first possibility make sense as the genes are positioned together and facing the same direction (Figure 2.1) but, this being the case, it is challenging to explain how *pstS* and *phoB* are also correlated, when *phoB* is not correlated to the expression of the other genes on the operon. The second

possibility adds more complexity to the regulation of these genes as it would require the presence of one promoter controlling the expression of *pstS* responsive to As^{III}, As^{V} and YE in stationary phase and another one (or more) controlling its expression in other conditions and related to *phoB.* The expression of *phnD, pstS2* and *pstS3* are not correlated on the conditions tested here; this suggests that although *pstS* and *phoB* are both thought to be involved in response to Pi concentration in the cell, they seem to be adapted to respond to As^{III} and As^{V} in NT-26. *pstS* follows the same expression pattern of the *ars* genes but the same is not observed for *pstS2* and *pstS3*. Suggesting that *pstS* is more related to As resistance system and could be being used as a As^{V} transporter while PstS2 and PstS3 kept their function and act as Pi transporters. The different roles of PstS, PstS2 and PstS3 can also be expected if observed their low identity, with PstS2/PstS3 sharing 46% identity, PstS/PstS2 sharing 32% identity and PstS/PstS3 sharing 27% as summarized in table 2.1. The proximity of *pstS* and *phoB* to the *ars* genes in the plasmid might be another indicator that these genes evolved to be part of – or to contribute to – the As^{V} resistance system.

PstS is a high affinity phosphate transporter, and different studies have shown that although it can transport As^{V}, its affinity for Pi can be 500 - to 850 times higher than for As^{V} (Wang et al. 2015; Elias et al. 2012). The affinity of PstS for phosphate increases when the concentration of As^{V} is in excess over Pi and that happens more often for *pstS* located near the *ars* genes (Michael et al. 1980; Elias et al. 2012). That suggests that PstS adapts to As^{V} rich environment by optimizing the Pi uptake, given that Pi is essential for the cell (Westheimer 1987), and the presence of a *pstS* near the *ars* genes can be an indicative of a mechanism that evolved to support NT-26 survival in As^{V} rich medium. PstS increased specificity for Pi in As^{V} rich medium has

been observed in both non As^{V}-sensitive species such as *Escherichia coli* and *P. fluorescens* and As^{V}-resistant bacteria such as *Halomonas* sp. GFAJ-1 (Elias et al. 2012). More studies need to de done to compare the specificity of PstS, PstS2 and PstS3 to Pi and As^{V} in NT-26 to be able to understand how each one of the transporters function in NT-26. The role Pi plays in the regulation of the *pstS, ars, phoB, pstS2, pstS3* and *phnD* also needs to be studied to elucidate how the As^{V} resistance is linked to the response to Pi concentration in NT-26.

It is unclear why *phoB* is present in the vicinity of *ars* genes. PhoB could be a trans acting regulator, responsive to As^{III} and stationary phase - due to the decrease of Pi -, controlling the expression of genes elsewhere in the genome. It would be useful to observe if the PhoB is active, where it is binding in NT-26, and its relationship with the other PhoB present in NT-26 (PhoB1, PhoB2a, PhoB2b as shown in Figure 2.2).

2.5 Chapter summary

- The genes *glnA, gyrB* and *gltA* were selected as Ref genes because they were stably expressed in NT-26 when it was grown with As^{III}, As^{V} and YE alone and in late-log and stationary phases. They were used to normalize the qPCR.
- The genes *pstS, arsR1a, arsC* and *arsB1* were up-regulated when NT-26 was grown in the presence of As^{III}, As^{V} and in YE alone in stationary phase. Their expression are highly correlated ($r > 0.7$, p-value < 0.05) indicating that they can be transcribed under the control of the same promoter or at least respond to the same stimuli.
- The presence of *phoB* and *pstS* near the *ars* genes and their similar expression patterns indicate that these genes may be involved on As resistance in NT-26.

- *pstS2* and *pstS3* are not responsive to the presence of As^{III} or As^{V} in the growth media, contrary to what was seen for *pstS,* indicating that they perform different functions in NT-26.
- *phoB* is more highly expressed in presence of As^{III} and As^{V} (in stationary phase), but *ars* genes have no PhoBox on their promoter region, suggesting that PhoB can act controlling a transcriptional regulator which controls the *ars* genes.

Chapter 3

Regulation of arsenite oxidation genes in NT-26

3.1 Introduction

The facultative chemolithoautotrophic arsenite oxidizer, NT-26 is able to grow chemolithoautotrophically with As^{III} as the electron donor, oxygen as the electron acceptor, and carbon dioxide or bicarbonate as the carbon source. NT-26 can gain energy from As^{III} oxidation under both chemolithoautotrophic and heterotrophic conditions (Santini *et al*., 2000).

3.1.1 As^{III} oxidase genes in NT-26

NT-26 oxidizes As^{III} using a periplasmic arsenite oxidase (Aio), which is a heterodimer composed of a large subunit AioA that has a molybdenum atom at its active site and 3Fe-4S cluster, and a small subunit (AioB) which contains a Rieske-type 2Fe-2S cluster (Santini & vanden Hoven, 2004; Santini *et al.*, 2007). The AioA and AioB subunits are encoded by the genes *aioA* and *aioB* which are part of a cluster also containing the genes *cytC* (encoding a periplasmic *c*-type cytochrome) and *moeA1* (encoding a molybdenum cofactor biosynthesis gene), which are transcribed together only when NT-26 is grown with As^{III} (Santini *et al.*, 2007) (Figure 3.1). The periplasmic *c*-type cytochrome encoded by *cytC* can act as an electron acceptor, contributing to the Aio activity, however, the *cytC* is not essential for the As^{III}-oxidation (Santini et al. 2007). NT-26 also has a putative As^{III}-binding protein AioX and a two-component system AioS and AioR that respond to the presence of As^{III} in the medium and positively regulate the expression of the As^{III} oxidase genes (Sardiwal et al. 2010) (Figure 3.1).

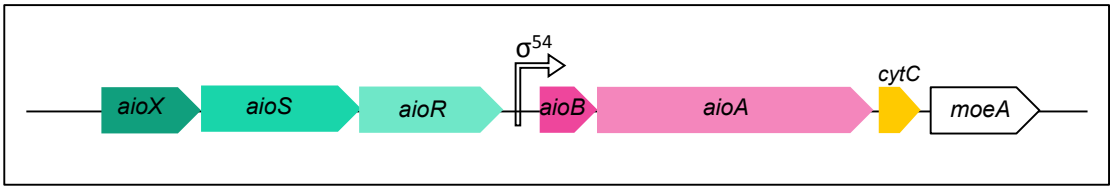


Figure 3.1: Organisation of the NT-26 *aio* gene cluster. *aioX*, As^{III}-binding protein gene; *aioS,* sensor histidine kinase gene; *aioR,* transcriptional regulator gene; *aioB,* As^{III} oxidase small subunit gene; *aioA,* As^{III} oxidase large subunit gene; *cytC,* cytochrome *c* gene; *moeA,* molybdenum cofactor biosynthesis gene. (Santini & Vanden Hoven 2004).

3.1.2 Regulation and expression of the *aio* gene cluster in NT-26

Despite the significant progress obtained from studies on the metabolism of As^{III} in various organisms, understanding the regulation of the *aio* gene cluster is somewhat limited. The genes *aioB* and *aioA* have been identified in many phylogenetically diverse arsenite oxidisers (Osborne & Santini 2012), but seem to be regulated differently.

The genes *aioS* and *aioR* are part of a cluster with *aioB, aioA, cytC* and *moeA1* (Figure 3.1). *aioS* and *aioR* are constitutively expressed and transcribed as a unit, in the conditions tested, in NT-26 grown with either YE alone or addition of As^{III} (Santini et al. 2000), different from *aioB* and *aioA,* that are only expressed in presence of As^{III}. *cytC* and *moeA* are both constitutively expressed, but also seen transcribed in unit with *aioBA* when in presence of As^{III} (Santini et al. 2000; Santini et al. 2007).

Studies in NT-26 where *aioS* and *aioR* were mutated showed that these genes are essential for As^{III} oxidation by providing transcriptional regulation of the *aioBA* (Sardiwal *et al*., 2010). AioR has the sequence motif essential to bind and activate the

σ^{54}, found upstream *aioB* (Santini et al. 2007; Sardiwal et al. 2010) and the conserved region where AioR binds in the promoter region was also identified upstream *aioB* and σ^{54} (Andres et al. 2013) (Figure 1.5). Transcriptome analysis in NT-26 showed that induction of *aioSR* occurs after prolonged exposure to As^{III} (6-8 hours) and this induction is a specific response to As^{III} (Andres et al. 2013). The transcription regulation of the genes *aioB* and *aioA* by the two-component AioR and AioS is dependent on AioX, as the inactivation of the *aioX* gene, results in a loss of As^{III} oxidation in NT-26 (Andres *et al.*, 2013).

3.1.3 Aims

The aim of this chapter was to observe the effect of the addition of either As^{III} or As^{V}, in the growth medium containing YE, on the expression of the genes involved with As^{III} sensing - *aioX, aioS* and *aioR* - and As^{III} oxidation - *aioB, aioA* and *cytC* in NT-26. For that, RT-PCR and qPCR techniques were used. The effect of the growth phase on the expression of the aforementioned genes was also assessed using qPCR. The objective is to understand how NT-26 responds to the presence of As^{III} and As^{V} in the growth media and to compare the regulation of these genes with what is known for others As^{III} oxidizers.

3.2 Materials and Methods

3.2.1 Growth conditions

NT-26 was grown in MSM with an addition of 0.04% (v/v) YE in the absence and presence of 5mM As^{III} or 5mM As^{V} (as described on Section 2.2.2). The cells were grown until late-log (OD_{600} from 0.100 to 0.140) and stationary phases (OD_{600} from 0.170 to 0.24) identical to Section 2.2.3.

3.2.2 RNA isolation

RNA from NT-26 cells was obtained as described previously (Section 2.2.4), and the purity and concentration of the RNAs isolated are the same as stated in Chapter 2 in Section 2.3.2.

3.2.3 RT-PCR

The reverse transcriptase-PCR (RT-PCR) was used to confirm the transcription of *aioX, aioS, aioR, aioB* and *cytC* and co-transcription of *aioX, aioS* and *aioR,* genes using the Access RT-PCR system kit (Promega) in accordance with the manufacturer's instructions. For each primer pair tested, two reactions were prepared using the same amount of total RNA per reaction (26 ng). The reverse transcriptase was added to one of the reactions, while molecular grade water was added to another reaction to test for DNA contamination and to be used as a negative control.

3.2.4 RT-PCR primer design

The primer pairs used in RT-PCR reactions to assess the transcription of *aioX, aioS, aioR, aioB* and *cytC* were designed to amplify a fragment inside the target gene in NT-26. The primers to analyse the co-transcription of *aioX and aioS* and *aioS and aioR*

were designed to amplify the 3'-end of the first gene (*aioX* and *aioS*, respectively) and the 5'-end of the next gene (*aioS* and *aioR*, respectively). The primer sequences are listed in the table 3.1:

Table 3.1. RT-PCR primers sequence and details:

Product	Sequence (5' - 3')	Amplicon (bp)	Locus tag[a]
aioX	F: GCGGTACCGTGATCGGCGGAACTCTG R: GCCTGCAGTCATCCCAGCCTCCGCACG	858	NT26v4_p10026
aioS	F: GCGGATCCCTATGATCTGCTCGACCGTAC R: GCGAATTCTGCTCATGCACGTCAATGTCT	880	NT26v4_p10027
aioR	F: GCGGATCCCTCGAAGATGATCCGATCAT R: GCGAATTCGCTGCATGACGCCAATCTCG	697	NT26v4_p10028
aioB	F: GCCTGCAGATGTCACGTTGTCAAAACATGGTCG R: GCCTGCAGTTATAGAACGTTGGACAGACGGCCG	544	NT26v4_p10029
cytC	F: CGAATTCATGCGGAAACTGTTTTCGATCG F: GCGAATTCTTACGGGGTGCTGAATGTCTTGAG	407	NT26v4_p10031
aioX and aioS	F: AGGCTTCTTCCGCAAATCGT R: ACGACCATCATTAGCAGGGC	452	NT26v4_p10026 NT26v4_p10027
aioS and aioR	F: TAAACCTGGTTCTGAACGCGA R: TCGAAAGCATCTGTCCCGTC	479	NT26v4_p10027 NT26v4_p10028

F: Forward primer.

R: Reverse primer.

[a]Locus tag number for NT-26 gene sequence in the NCBI Server (http://www.ncbi.nlm.nih.gov). The NT-26 genome sequence accession numbers are FO082822.1, FO082820.1 and FO082821.1 (Andres et al. 2013) .

3.2.5 Agarose gel electrophoresis

The DNA fragments obtained after the RT-PCR were mixed with 1 X concentration of DNA loading buffer (Promega) and separated by agarose gel electrophoresis. The samples were applied to a 0.8% (w/v) agarose gel dissolved in 1 X TBE buffer (89 mM Tris; 89 mM boric acid; 2 mM EDTA; pH 8). 0.5 µg/ml of ethidium bromide was added for the DNA visualisation under UV light.

To determine the size of DNA fragments a 1 kb Plus DNA ladder (Fermentas) was

used as a molecular weight standard which contained the following fragment sizes: 20.0, 10.0, 7.0, 5.0, 4.0, 3.0, 2.0, 1.5, 1.0, 0.7, 0.5, 0.4, 0.3, 0.2 and 0.075 kb.

3.2.6 cDNA synthesis

The cDNA used in the qPCR was prepared using the qPCR First Strand cDNA Synthesis protocol from the RevertAid Premium First Strand cDNA Synthesis kit (Thermo Scientific) and all the steps and samples used here were the same as described in section 2.2.5.

3.2.7 qPCR reactions and primer design

The qPCR reactions were performed as described in section 2.2.6 with primers designed to amplify fragments from 75-113 bp and T_m between 55 and 65 °C and CG content of 50 – 60% (table 3.2), similar to what was done in section 2.2.7. The OligoAnalzyer 3.1 Software (Integrated DNA Technologies Inc) was used to check for the likelihood of hairpins, self-dimers and hetero-dimers formation and the primers were synthesised by Eurofins MWG Operon (Ebersberg, Germany).

Table 3.2: Primers designed for qPCR in NT-26

Primer	Sequence (5' → 3')	Gene product	Gene name	Locus tag[a]	Amplicon size (bp)
aioA	F: TGGCAGACAGAAGAAGAC R: TGGCAGACAGAAGAAGAC	Arsenite oxidase large subunit	*aioA*	NT26_p10030	75
aioB	F: TTCCACGATCTGTCCTCA R: AAGACCGAGAAGTGACCA	Arsenite oxidase small subunit	*aioB*	NT26_p10029	87
aioR	F: CAAGGACCTCTTCTTCAG R: GAGCCAGAGAATGTCTTC	Two-component response regulator	*aioR*	NT26_p10028	78
aioS	F: TGCTAATGATGGTCGTCAG R: GTTCAGATGCTGCTCCTG	Two-component system signal transduction	*aioS*	NT26_p10027	80
aioX	F: GTCTAATGACCTCGAAGTG R: TTATGAGCTGAACCTCCT	Putative periplasmic arsenite-binding protein	*aioX*	NT26_p10026	77
cytC	F: CAAGGACCTCTTCTTCAG R: GAGCCAGAGAATGTCTTC	*c*-type cytochrome c_{552}	*cytC*	NT26_p10031	79

F: Forward primer.R: Reverse primer.

[a]Locus tag number for NT-26 gene sequence in the NCBI Server (http://www.ncbi.nlm.nih.gov). The NT-26 genome sequence accession numbers are FO082822.1, FO082820.1 and FO082821.1 (Andres et al. 2013) .

3.3 Results

The presence of As^{III} and As^{V} in the growth medium is known to affect the expression of genes in NT-26, especially genes involved in As^{III} oxidation and As^{V} reduction – as seen in Chapter 2 and previously described by (Andres et al. 2013; Santini et al. 2007). The aim of this study was to further investigate the effect of As^{III} and As^{V} on the regulation of *aioX, aioS, aioR, aioB, aioA* and *cytC.* This was monitored using RT-PCR and quantified using qPCR. The transcription level of each gene was quantified using RNA isolated from NT-26 grown in a medium containing YE in the absence and presence of As^{III} or As^{V} until late log and stationary phases.

3.3.1 Gene expression in NT-26 measured by RT-PCR

The expression of the genes *aioB, aioA* and *cytC* in NT-26 in response to As^{III} was first studied by Santini et al., 2007 using RT-PCR. It was shown that *aioB* and *aioA* are co-transcribed in presence of As^{III} in the growth medium and also that *aioA* and *cytC* are co-transcribed in presence of As^{III}, although *cytC* was also shown to be transcribed in absence of As^{III} suggesting that it is constitutively expressed (Santini et al. 2007). RT-PCR was also used by Sardiwal et al. (2010) to show that *aioS* and *aioR* are expressed in NT-26 growth with YE alone, suggesting constitutive expression, and that *aioR* and *aioB* can not be transcribed as a unit, even in presence of As^{III} (Sardiwal et al. 2010).

3.3.1.1 Transcription of *aioX, aioS, aioR, aioB, aioA* and *cytC* when NT-26 was grown in the absence and presence of As^{III}

To observe the expression of *aioX, aioS, aioR, aioB* and *cytC* genes under two different growth conditions, RT-PCR was performed using RNA isolated from NT-26 grown to late-log phase in the presence and absence of As^{III}. The correct size

transcripts of *aioX, aioS, aioR* and *cytC* were detected when NT-26 was grown in both the presence and absence of As^{III} whereas *aioB* expression was only detected when NT-26 was grown in the presence of As^{III} (Figure 3.2). No PCR products were visualized in the negative controls, which did not contain reverse transcriptase, confirming that the samples were free from DNA contamination (Figure 3.2, negative controls).

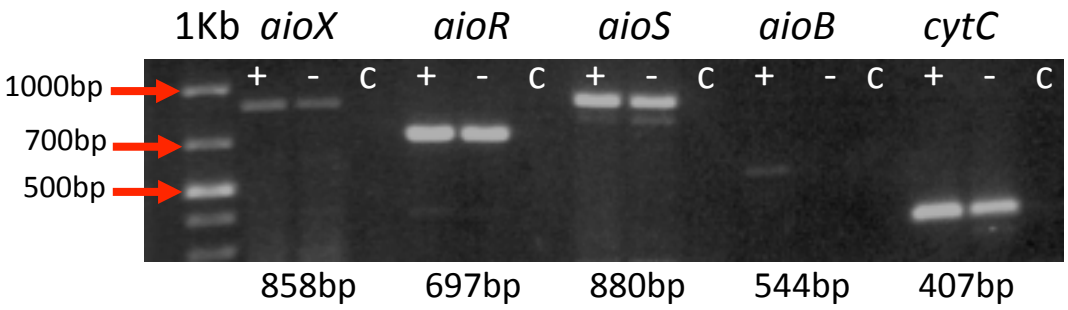


Figure 3.2: RT-PCR analysis of the *aioX, aioS, aioR, aioB* and *cytC* genes. RNA was isolated from NT-26 grown until late log phase with and without As^{III}. The sizes of the molecular markers are shown on the left and the sizes of the expected band are shown below the lanes. **+**: NT-26 grown with As^{III}; **-**: grown with YE alone and **C**: negative control.

3.3.1.2 Co-transcription of *aioX, aioS* and *aioR* when NT-26 was grown in the absence and presence of As^{III}

To determine whether the *aioX, aioS* and *aioR* genes are co-transcribed, and therefore, part of the same operon, two sets of primers were designed, the first was to amplify the 5' region of *aioS* and 3' region of *aioX*, the second was used to amplify the 5' portion of *aioR* and the 3' portion of *aioS* (Figure 3.3 A). RT-PCR was performed using the RNA isolated from NT-26 grown with and without As^{III} (figure 3.3).

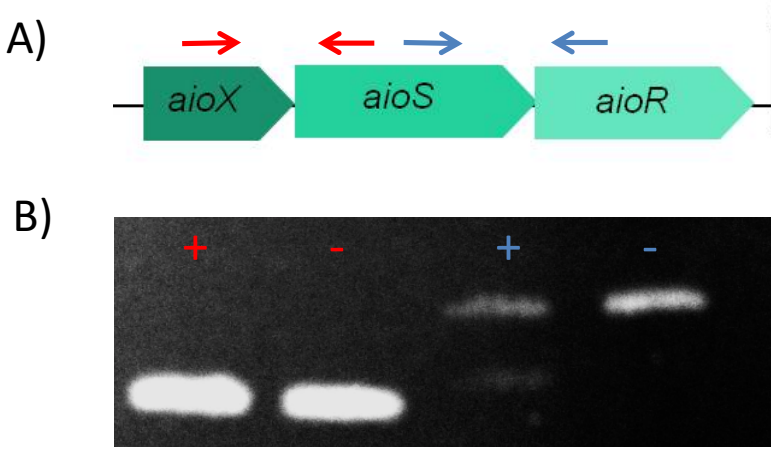


Figure 3.3: RT-PCR analysis of the co-transcription of *aioX* - *aioS*, and *aioS*- *aioR*. A) Positions of the primers designed to amplify *aioX/aioS* (red) and *aioS/aioR* (blue) B) Products of RT-PCR using *aioX/aioS* and *aioS/aioR* primers and RNA isolated from NT-26 grown until late log phase with and without As^{III}. **+**: RNA isolated from NT-26 grown with As^{III}; **-**: RNA isolated from NT-26 grown without As^{III}.

This result shows that *aioX, aioS* and *aioR* are co-transcribed in both the presence and absence of As^{III}, which means that they are part of the same operon and are under the control of the same promoter(s). The band corresponding to *aioS* and *aioR* in presence of As^{III} (Figure 3.3 B, in blue) is slightly less bright when compared to YE alone, as the same amount of RNA was used in both reactions that could indicate that less product is formed in presence of As^{III}. Although it is important to mention that RT-PCR is not a quantitative method and only qualitative conclusions should be taken (presence *vs* absence of an amplification product). No PCR products were obtained when the RT step was absent and only DNA polymerase was used in the reaction, confirming that the samples were free from DNA contamination.

3.3.2 Gene expression in NT-26 measured by qPCR

To determine the effect of As^{III} or As^{V} on the expression of the As^{III} oxidation related genes in NT-26, qPCR was used. The primers used in the qPCR were designed to amplify the genes *aioX, aioS, aioR, aioB, aioA* and *cytC* (Figure 3.1).

The expression values obtained by qPCR presented in this chapter were normalised to the Ref genes chosen and validated in Chapter 2 and the same qPCR protocol was followed (Chapter 2, section 2.2.7).

3.3.2.1 qPCR primer test

As qPCR is a sensitive method and SYBR Green I binds to any dsDNA, it is important that the primers used on each reaction are specific to the gene and are carefully designed and tested.

The efficiency of the *aioX, aioS, aioR, aioB, aioA* and *cytC* primers was calculated as described in Chapter 2 (section 2.2.8) and is as follows:

Table 3.3: qPCR primer efficiency

Primer	Efficiency (%)
aioX	80
aioS	97.3
aioR	99.8
aioB	82.1
aioA	100.8
cytC	103.3

As the efficient was greater than 80% all the primers were considered suitable for qPCR. Their different amplification efficiencies will be taken into account when comparing the expression levels of the different genes tested (Hellemans et al. 2007).

The melting curve of each primer pair was analysed using the Piko Real Software (section 2.2.9) to ensure that no hairpins or dimers were formed during the reaction and also to check for gDNA contamination and non-specific binding. All the primer pairs checked were satisfactory and were used to observe the relative expression of NT-26 genes grown in the presence of As^{III}, As^{V} and at different growth phases.

3.3.2.2 Relative expression in NT-26 grown in the presence of YE and As^{III}

The presence of As^{III} in addition to YE in NT-26 growth medium, induced the up-regulation of genes related to As^{V} reduction and Pi metabolism, as seen in Chapter 2. The presence of As^{III} is also known to induce the expression of *aioB, aioA* and *cytC* in NT-26, according to the results presented on section 3.3.1.1 (Figure 3.2) that are in accordance to the literature (Santini et al. 2007). However, there are neither quantitative comparisons between the expression levels of these genes, nor a study assessing the expression of *aioX* in NT-26 grown with As^{III}.

To better understand how the *aioX, aioS, aioR, aioB, aioA* and *cytC* genes are expressed in presence of As^{III} in NT-26, qPCR was used and the expression levels were normalised to the Ref genes. The level of expression observed in presence of As^{III} was compared with the values obtained for NT-26 grown with YE alone, both until late log phase (Figure 3.4).

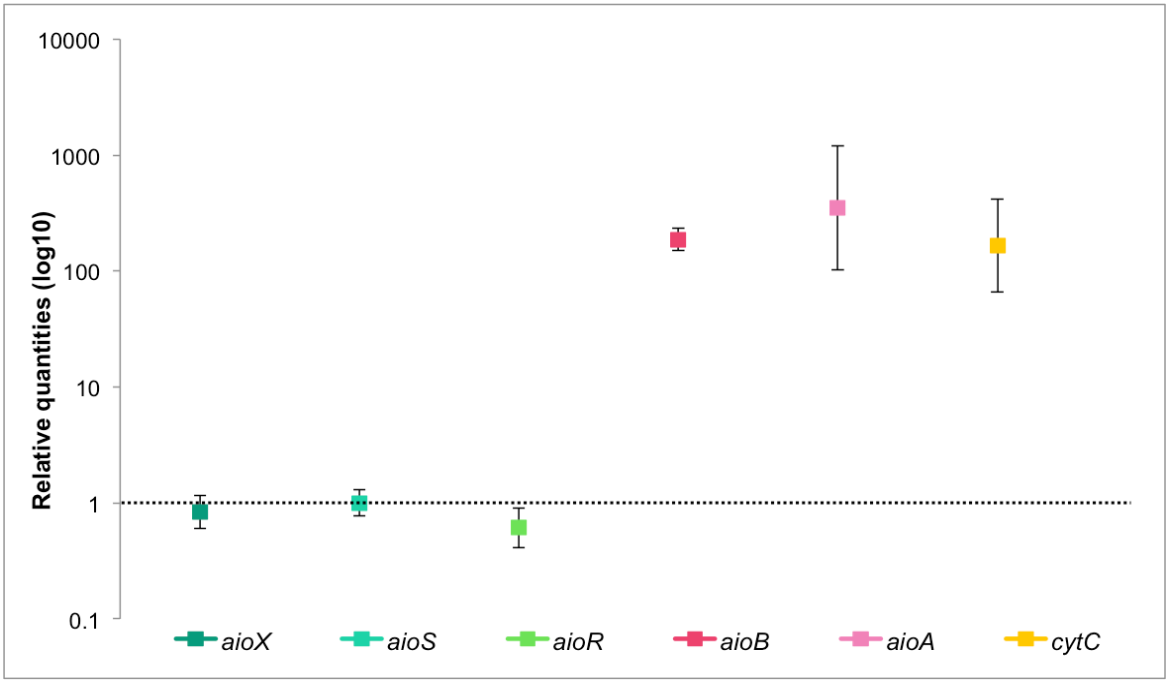


Figure 3.4: Relative expression analysis using qPCR and the software qBase+ to calculate the expression ratios on the basis of the NT-26 YE samples *vs*. NT-26 with YE and As^{III} in the growth medium. The dotted line illustrates the threshold where the expression is not different when the two conditions are compared. All the samples above the line are up-regulated in the presence of As^{III} when compared to YE alone. The error bars show the 95% upper and lower confidence intervals.

The expression of the genes *aioX, aioS* and *aioR* were not affected by the presence of As^{III} in the medium - they are constitutively expressed, meaning that they are equally expressed in both conditions. The genes *aioB, aioA* and *cytC* were up-regulated in presence of As^{III} by 186-, 349- and 166-fold, respectively.

3.3.2.3 Relative expression in NT-26 grown in the presence of YE and As^{V}

As well as understanding the effect the presence of As^{III} has on the regulation of *aioX, aioS, aioR, aioB, aioA* and *cytC*, the effect of the presence of As^{V} was also assessed. As^{V} is reduced to As^{III} in the cytoplasm, so As^{V} might directly or indirectly induce the expression of these genes involved on As^{III} sensing and oxidation in NT-26.

The effect of As^{V} on the expression of *aioX, aioS, aioR, aioB, aioA* and *cytC* was analysed by qPCR. The relative expression was calculated comparing the expression levels of genes in NT-26 grown with As^{V} as opposed to NT-26 grown with YE alone (Figure 3.4).

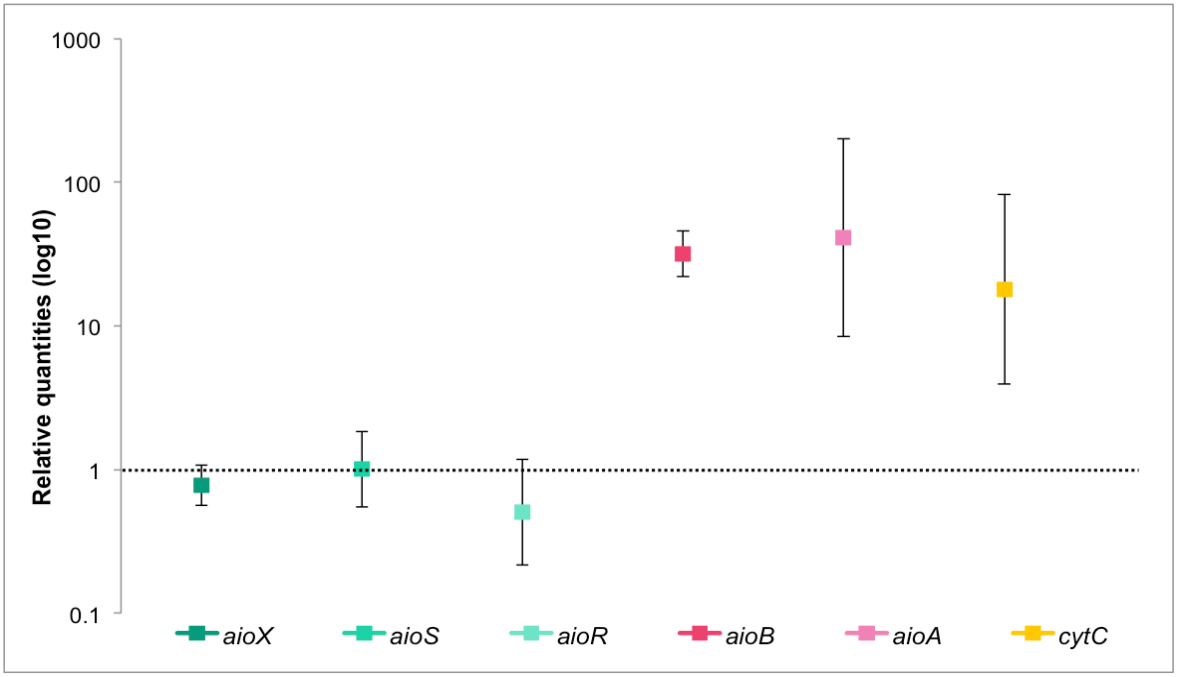


Figure 3.5: Relative expression analysis using qPCR and the software qBase+ to calculate the expression ratios on the basis of the NT-26 YE samples *vs*. NT-26 with As^{V} in the heterotrophic growth medium. The dotted line illustrates the threshold where the expression is different when the two conditions are compared. All the samples above the line are being up-regulated in the presence of As^{V} when compared to YE alone. The error bars show the 95% upper and lower confidence intervals.

The analysis of qPCR results showed that the expression of *aioX, aioS,* and *aioR* genes was not significantly affected by the presence of As^{V}, similar to what was observed for As^{III}. As for *aioB, aioA, cytC*, they were all up-regulated by an average of 31-, 41- and 18-fold, respectively.

3.3.2.4 Effect of growth phase on expression of *aioX, aioS, aioR, aioB, aioA* and *cytC* in NT-26

As demonstrated in Chapter 2 (section 2.3.4.3) the genes *phoB, pstS arsR1a, arsC, arsB1* and *pstS3* are up-regulated in stationary phase indicating that the genes responsive to As^{III} and As^{V} can also respond to stationary phase in NT-26. That suggests that they are required when the cell is about to enter a starvation or stress period.

To elucidate the effect of growth phase on the expression of *aioX, aioS, aioR, aioB, aioA* and *cytC* qPCR analysis were undertaken with samples obtained from NT-26 grown in media with YE alone until late-log and stationary phases and the results were compared (figure 3.6)

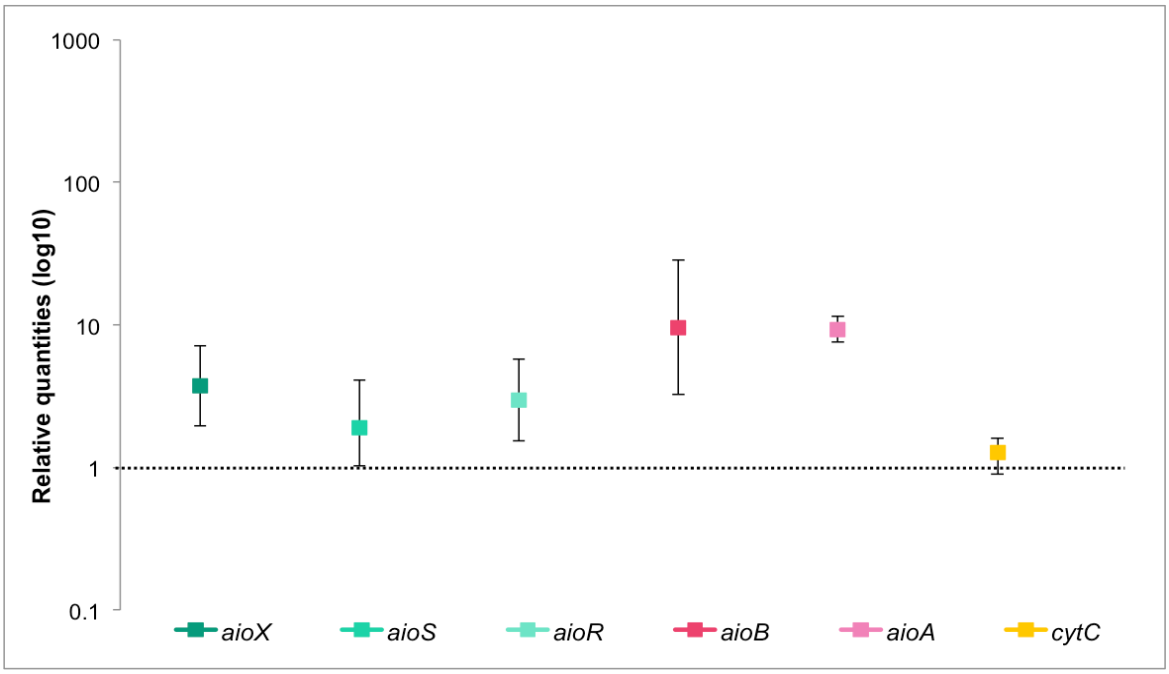


Figure 3.6: Relative quantities of the tested genes in NT-26 samples grown with YE until stationary phase compared to late log phase. The dotted line illustrates the threshold where the expression is not significantly different when comparing the two conditions ($p>0.05$). All the samples above the line were up-regulated in stationary phase of growth compared to late-log phase. The error bars show the 95% upper and lower confidence intervals; any error bar crossing the threshold line indicates that the expression in the two conditions tested is not significantly different.

When NT-26 was grown with YE alone, the expression of *aioX, aioR* and *aioS* were up-regulated in stationary phase compared to late-log phase by 3.74-, 2.97-, 1.9-fold, respectively. The same was observed for *aioB* and *aioA* with an increase in 9.57- and 9.33- fold respectively. *cytC* was not up-regulated in stationary phase compared to late-log. Similar results were seen when the expression of these genes was compared in NT-26 grown in presence of As^{III} until late log and stationary phases (figure 3.7).

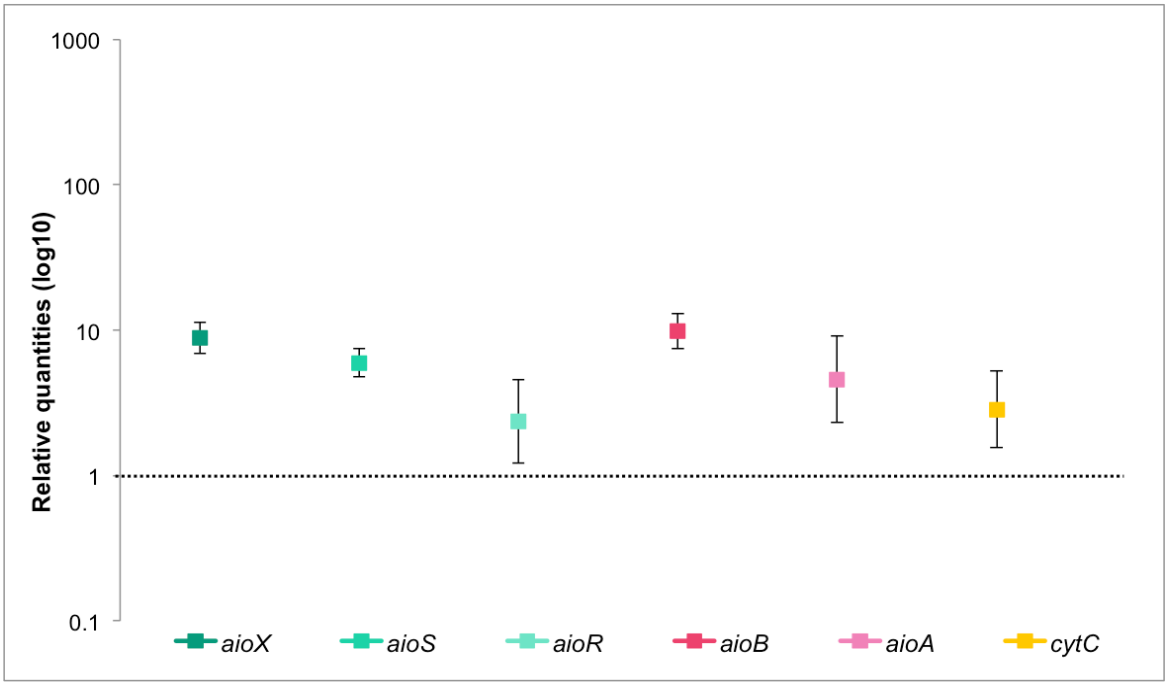


Figure 3.7: Relative quantities of the tested genes in NT-26 samples grown with As^{III} until stationary phase compared to late log phase. The dotted line illustrates the threshold where the expression is not significantly different when comparing the two conditions ($p>0.05$). All the samples above the line were up-regulated in stationary phase of growth compared to late-log phase. The error bars show the 95% upper and lower confidence intervals; any error bar crossing the threshold line indicates that the expression in the two conditions tested is not significantly different.

Although the genes *aioX, aioS* and *aioR* were not up-regulated in response to the presence of As^{III} (figure 3.4), the growth phase appears to have an effect on the expression of these genes as they are up-regulated by 8.81-, 2.35- and 5.95-fold, respectively. The genes *aioB, aioA* and *cytC* are also up-regulated in stationary phase by 9.8-, 4.59- and 2.85-fold, respectively. These results suggest that the stage of growth plays a role on the regulation of *aioX, aioS, aioR*, *aioB, aioA* and *cytC* in NT-26.

A different result was obtained when the gene expression was quantified in NT-26 grown until stationary phase in presence of As^{V} (figure 3.8). The genes *aioX, aioS* and *aioR* were up-regulated in stationary phase by 25-, 16- and 7.5-fold, respectively.

aioB was up-regulated by 3.7-fold and *aioA* and *cytC* gene expression in stationary phase were not significantly different when compared to late-log phase.

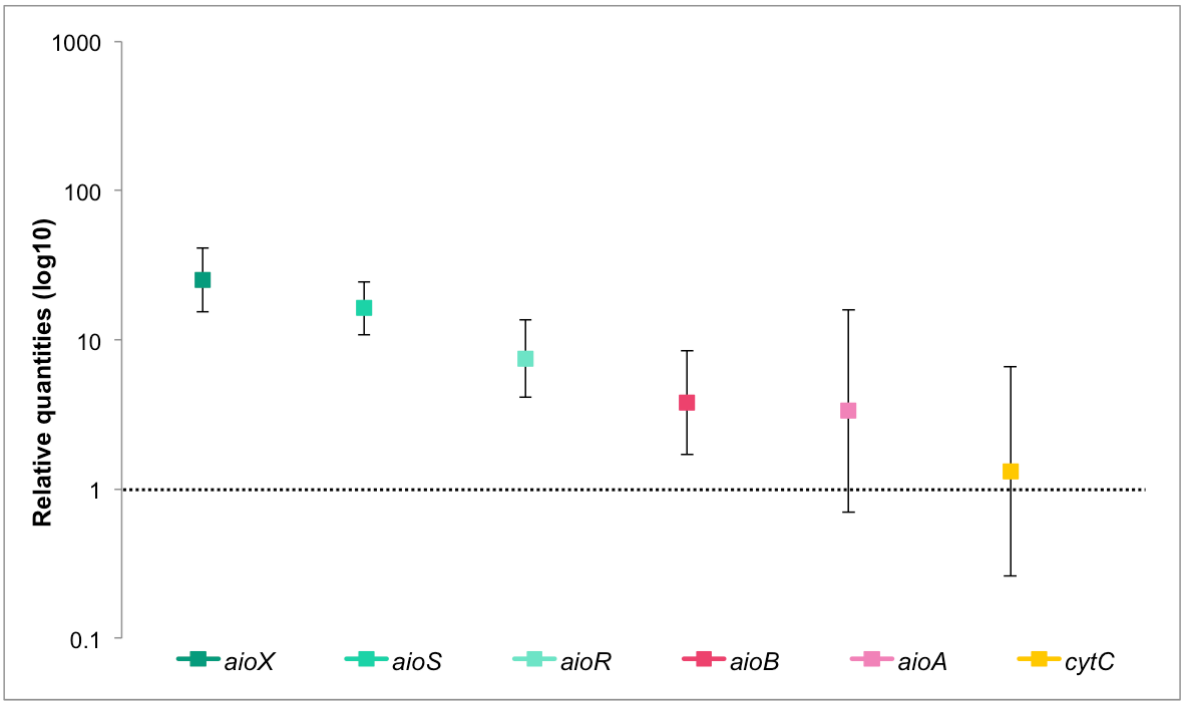


Figure 3.8: Relative quantities of the tested genes in NT-26 samples grown with As^{V} until stationary phase compared to late log phase. The dotted line illustrates the threshold where the expression is not significantly different when comparing the two conditions ($p>0.05$). All the samples above the line were up-regulated in stationary phase of growth compared to late-log phase. The error bars show the 95% upper and lower confidence intervals; any error bar crossing the threshold line indicates that the expression in the two conditions tested is not significantly different.

According to these results, the growth phase is important in the regulation of *aioX, aioS* and *aioR* as these genes were up-regulated in stationary phase and were constitutively expressed in late-log (Figure 3.5) both in presence of As^{V}.

These results can suggest that the up-regulation in stationary phase could be related to a general stress response in the cell due to the decrease of nutrients or due to

quorum sensing as shown for *A. tumefaciens* 5A (Kashyap *et al*., 2006). These hypotheses require further analyses.

3.3.2.5 Correlation between gene expression

To compare the expression patterns of *aioX, aioS, aioR, aioB, aioA* and *cytC* genes under all the conditions tested a correlation analysis were undertaken using the Spearman correlation test (Spearman, 1904) (Table 3.4). A positive correlation has a r-value ranging between 0.7 and 1, which means that whenever a gene is up-regulated, your correlated pair will also be, as illustrated in Figure 3.10 for *aioB* and *aioA* (r-value: 0.96). Genes can also be negatively correlated, when that occurs, the r-value is ranging from -0.7 to -1 and that means that they always behave in opposite ways, whenever one gene is up-, the other will be down-regulated in the conditions tested. The expression of two genes will be considered highly correlated if the r-value is higher than 0.7 (or -0.7) and p-value is lower than 0.05 (Hinkle, 1979).

Table 3.4: Correlation values (r) between two genes expressed under all conditions tested. All values presented in bold are statistically significant ($p<0.05$) and have a r-value > 0.7:

target x	target y	Correlation value (r)	Significance (p-value)
cytC	***aioA***	**0.957**	**0.00E+00**
aioA	***aioB***	**0.96**	**0.00E+00**
cytC	***aioB***	**0.938**	**1.51E-07**
aioS	***aioX***	**0.923**	**3.74E-07**
aioR	***aioX***	**0.824**	**2.85E-06**
aioS	***aioR***	**0.795**	**3.33E-06**
aioS	*aioB*	0.383	8.00E-02
aioS	*aioA*	0.363	9.17E-02
aioB	*aioX*	0.312	1.56E-01
aioA	*aioX*	0.277	2.05E-01
cytC	*aioS*	0.201	3.87E-01
cytC	*aioR*	-0.14	5.74E-01
cytC	*aioX*	0.126	5.82E-01
aioA	*aioR*	0.054	8.33E-01
aioR	*aioB*	0.016	9.31E-01

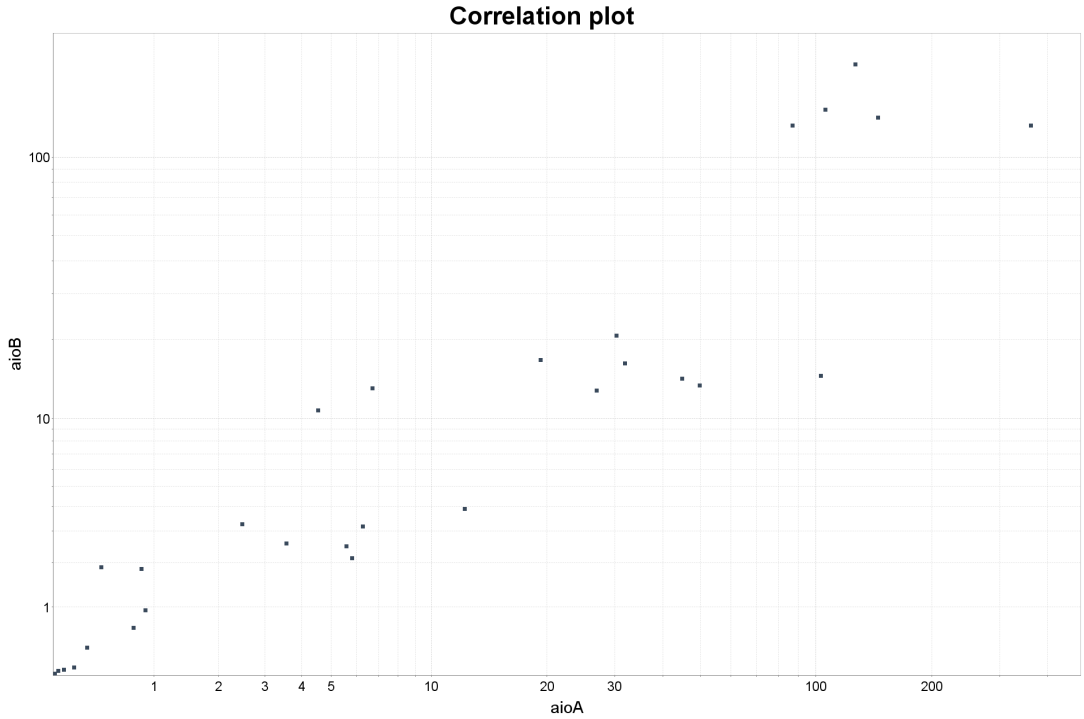


Figure 3.9: Correlation plot obtained when *aioA* (x axis) and *aioB* (y axis) expression levels are compared when NT-26 was grown in presence of As^{III}, As^{V} or with YE alone until late-log and stationary phases of growth. The r-value for these genes is 0.96 and it is statistically significant ($p<0.05$).

As expected from previous observations (figures 3.4 to 3.8), *aioX, aioS* and *aioR* behave in a similar way, i.e., when one gene is up-regulated the others will also be. This coordinated behaviour is probably due to the presence of one (or more) promoter upstream of *aioX* controlling the expression of these genes as they are transcribed in the same transcriptional unit. Similarly, and as expected, the expression of *aioB* and *aioA* are highly correlated (r: 0.96) as these genes are also co-transcribed and their expression is under the control of the sigma54 (σ^{54}) promoter as previously described (Sardiwal et al. 2010). The expression of *cytC* is correlated with *aioB* (r: 0.938) and *aioA* (r: 0.957) and although it was shown that they can be co-transcribed in the presence of As^{III} (Santini et al. 2007) they can be differently expressed as shown in the figures 3.6 and 3.8.

3.4 Discussion

In this chapter RT-PCR and qPCR were used to study the expression of *aioX, aioS, aioR, aioB, aioA* and *cytC* in NT-26 grown when grown heterotrophically in the presence of As^{III} and As^{V} and in two growth phases, late-log and stationary phase.

3.4.1 Effect of As^{III} on gene expression in NT-26

The NT-26 genes *aioS* and *aioR* were previously shown by RT-PCR to be constitutively expressed but no results on *aioX* have been published (Sardiwal et al. 2010). As shown in section 3.3.1.1, *aioX, aioS* and *aioR* are co-transcribed.

In this chapter it was also shown, using qPCR, that As^{III} has no effect on the transcription of *aioX, aioS* and *aioR* confirming that these genes are constitutively expressed in NT-26. This result obtained for NT-26 is similar to what has been

reported for *Thiomonas arsenitoxydans* 3As, where the presence or absence of As^{III} had no effect on the expression of *aioX* (Slyemi et al. 2013). AioX, AioS and AioR are involved in the regulation of *aioBA* when As^{III} is present. *aioX* codes for a periplasmic protein that has been suggested to be involved on As^{III} sensing (Liu et al. 2012; Cai et al. 2009), is essential for the As^{III} oxidation in NT-26 (Andres et al. 2013), and is likely to be involved on AioS ability to sense and respond to As^{III}. In presence of As^{III} AioS autophosphorylates itself and transfers the phosphoryl group to the response regulator AioR (Sardiwal et al. 2010), which binds to a conserved σ^{54} promoter region upstream *aioB,* activating the transcription of *aioB* and *aioA* (Sardiwal et al. 2010; Koechler et al. 2010). The genes *aioS* and *aioR* were also shown to be up-regulated in presence of As^{III} and their expression was also related to growth phase, according to results presented based on transcriptomics analyses in *H. arsenicoxydans* ULPAS-1 (Cleiss-Arnold et al. 2010).

It has been previously shown in NT-26 that As^{III} is required in the growth medium for expression of the *aioA* and *aioB* genes. It was observed for *A. tumefaciens* 5A that *aioB* and *aioA* are transcribed in absence of As^{III} when in late-log (Kashyap et al. 2006), and in *H. arsenicoxydans* ULPAS-1*, aioB* is expressed in presence of As^{III} but only in the late phase of growth (Cleiss-Arnold et al. 2010). In this study the qPCR results show that when NT-26 was grown in the presence of As^{III} in late-log phase, *aioB* and *aioA* were up-regulated by 186- and 349-fold respectively (Table 3.5 and Figure 3.4). It was known from RT-PCR data that *aioBA* are only expressed in NT-26 in presence of As^{III} (Santini et al. 2007), these genes are expected to be up-regulated as they are needed for As^{III} oxidation, but no quantitative result has been presented until now.

Table 3.5. Relative expression values for *aioX, aioS, aioR, aioB, aioA* and *cytC* obtained from NT-26 samples grown until late-log phase with YE and As^{III} compared to YE alone:

	YE	As^{III}	
gene	**Expression (95% low - 95% high)**		**p-value**
aioX	1 (0.831 - 1.204)	**0.834** (0.599 - 1.161)	0.4221
aioS	1 (0.738 - 1.356)	**1** (0.768 - 1.301)	0.8413
aioR	1 (0.531 - 1.884)	**0.61** (0.412 - 0.904)	0.2513
aioB	1 (0.642 - 1.557)	**186.953** (149.503 - 233.786)	0.01488
aioA	1 (0.501 - 1.998)	**349.352** (102.114 - 1195.20)	0.01488
cytC	1 (0.598 - 1.673)	**166.27** (66 - 418.874)	0.01488

Santini et al (2007) have shown that *cytC* is constitutively expressed and is expressed in the presence of As^{III} in the same operon as *aioB* and *aioA* and *moeA1*. In absence of As^{III} only *cytC* and *moeA1* are transcribed together suggesting that there is another promoter upstream of *cytC*. The qPCR results presented in this chapter are in accordance to that, although it is valid to point out that *cytC* was up-regulated by 166-fold when NT-26 was grown in the presence of As^{III}, compared to when NT-26 is cultivated with YE alone (Table 3.5, Figure 3.4), and that it is different from what was observed using microarray analysis (Andres et al. 2013), where it was believed that *cytC* is not differentially expressed in the presence of As^{III}. Discordance between microarray and qPCR results is not uncommon, qPCR is capable of detecting smaller changes in expression (less than 2-fold) and it uses a normalization method based on Ref genes. Dallas et al (2005) observed poor correlations for 13–16% of genes tested in his study using two different methods of normalization for microarray and comparing the results with qPCR, emphasising the importance for attention in interpreting gene expression data and the need for validation of microarray data (Dallas et al. 2005, Tulpan et al. 2012).

3.4.2 Effect of As^{V} on gene expression in NT-26

To understand how NT-26 responds to As^{V}, the expression levels of *aioX, aioR, aioS, aioB, aioA* and *cytC* were analysed in NT-26 by qPCR and the expression values obtained by qPCR are summarized in the table 3.6, below:

Table 3.6. Relative expression values for *aioX, aioS, aioR, aioB, aioA* and *cytC* obtained from NT-26 samples grown until late-log phase with YE and As^{V} compared to YE alone:

	YE	As^{V}	
gene	**Expression (95% low - 95% high)**		**p-value**
aioX	1 (0.831 - 1.204)	**0.781** (0.565 - 1.078)	0.1481
aioS	1 (0.738 - 1.356)	**1.007** (0.551 - 1.841)	1
aioR	1 (0.531 - 1.884)	**0.507** (0.217 - 1.184)	0.361
aioB	1 (0.642 - 1.557)	**31.81** (11.119 - 45.746)	0.01587
aioA	1 (0.501 - 1.998)	**41.258** (8.465 - 201.098)	0.01587
cytC	1 (0.598 - 1.673)	**18.062** (3.949 - 82.617)	0.01587

It was confirmed that the genes *aioX, aioR* and *aioS* are constitutively expressed when NT-26 was grown in the presence of As^{V} as was the case when NT-26 was grown in the presence of As^{III}. *aioB, aioA* and *cytC* were up-regulated by 31-, 41- and 18-fold, respectively. When comparing this result with the one obtained when NT-26 was grown in presence of As^{III} it is noticeable that the presence of As^{V} does not induce the expression of these genes as much as As^{III} (186-, 349- and 166-fold, respectively). This can be due to AioX not being able to sense As^{V} as it does for As^{III}, compromising the activation of AioS and phosphorylation and activation of AioR, and consequently, the regulation of *aioBA*. The up-regulation of *aioBA* and *cytC* observed when NT-26 is grown with As^{V} can also be a response to the As^{III} present due to the As^{V} reduction by the ArsC in the cytoplasm and not a direct response to As^{V}, however, to be able to affirm that, the concentration of As^{V} and As^{III} in the medium need to be measured during NT-26 growth, from early-log to stationary phase.

3.4.3 Effect of growth phase on gene expression

Genes are differentially expressed at different stages of growth, responding to the growth phase, nutrients availability and environmental cues (Navarro Llorens et al. 2010). The gene expression response to growth phase can be related to a change in the cell density and is known as quorum sensing mechanism of regulation, usually activated when the cells reach stationary phase (Waters & Bassler 2005). Quorum sensing is a form of community behaviour in response to cell density and accumulation of signalling molecules (autoinducers) produced by bacteria. When the bacterium senses the presence of autoinducers in the medium, it activates its response accordingly and that allows it to respond to specific conditions (Waters & Bassler 2005). To better comprehend the expression pattern of NT-26 *aioX, aioS, aioR, aioB, aioA* and *cytC* genes, NT-26 was grown until late-log and stationary phases and the expression levels of these genes were monitored using qPCR and the results shown in Figure 3.6, 3.7 and 3.8 are summarized in the table 3.7, below:

Table 3.7. Relative expression values for *aioX, aioS, aioR, aioB, aioA* and *cytC* obtained from NT-26 samples grown with YE until late-log compared to stationary phase:

	Late-log	Stationary	
gene	Expression (95% low - 95% high)		p-value
aioX	1 (0.831 - 1.204)	**3.749** (1.96 - 7.17)	0.04545
aioS	1 (0.738 - 1.356)	**1.907** (0.88 - 4.129)	0.04938
aioR	1 (0.531 - 1.884)	**2.979** (1.54 - 5.764)	0.02778
aioB	1 (0.642 - 1.557)	**9.579** (3.237 - 28.347)	0.02778
aioA	1 (0.501 - 1.998)	**9.336** (7.552 - 11.543)	0.04545
cytC	1 (0.598 - 1.673)	**1.266** (0.995 - 1.612)	0.73

Table 3.8. Relative expression values for *aioX, aioS, aioR, aioB, aioA* and *cytC* obtained from NT-26 samples grown with YE and As^{III} until late-log compared to stationary phase:

	Late-log	Stationary	
gene	Expression (95% low - 95% high)		p-value
aioX	1 (0.718 - 1.392)	**8.81** (6.90 - 11.25)	0.0277
aioS	1 (0.768 - 1.302)	**5.951** (4.764 - 7.433)	0.0277
aioR	1 (0.675 - 1.48)	**2.25** (0.1.22 - 4.54)	0.05556
aioB	1 (0.8 - 1.251)	**9.864** (7.465 - 13.033)	0.0277
aioA	1 (0.2392 - 3.421)	**4.594** (2.314 - 9.121)	0.03704
cytC	1 (0.397 - 2.519)	**2.856** (1.56 - 5.229)	0.03704

Table 3.9. Relative expression values for *aioX, aioS, aioR, aioB, aioA* and *cytC* obtained from NT-26 samples grown with YE and As^{V} until late-log compared to stationary phase:

	Late-log	Stationary	
gene	Expression (95% low - 95% high)		p-value
aioX	1 (0.724 - 1.381)	**25.22** (15.427 - 41.229)	8.57E-06
aioS	1 (0.547 - 1.828)	**16.271** (10.778 - 24.574)	4.09E-05
aioR	1 (0.429 - 2.333)	**7.531** (4.135 - 13.718)	0.00183
aioB	1 (0.695 - 1.438)	**3.785** (1.696 - 8.447)	0.0101
aioA	1 (0.205 - 4.874)	**3.35** (0.701 - 16.008)	0.2041
cytC	1 (0.219 - 4.574)	**1.319** (0.262 - 6.636)	0.738

Expression of *aioX, aioS* and *aioR* were up-regulated in stationary phase when NT-26 was grown with YE alone. Up-regulation of these genes was also seen in presence of As^{III} or As^{V} in stationary phase (Figure 3.6, 3.7 and 3.8 and Tables 3.7, 3.8 and 3.9). Based on RT-PCR results presented in this chapter, and on the qPCR results obtained in late-log, it was confirmed that *aioX, aioS* and *aioR* are expressed at basal levels, and no up-regulation was seen. However, *aioX, aioS* and *aioR* genes are all up-regulated in stationary phase, suggesting that they could be under the control of more than one promoter responding to different stimuli. This hypothesis will be addressed on Chapter 4 as it is not possible to prove the existence of more than one promoter based on the fluctuations of gene expression alone. Quorum sensing might be involved on the regulation as in stationary phase the cell density is higher than in late-log, facilitating the action of autoinducers. The up-regulation of *aioX, aioS,* and *aioR* in stationary phase can also indicate that AioX, AioS and AioR might be required when the cell is entering a starvation or stress period. In stationary phase the nutrients start to be scarce, especially when the bacteria are grown in the lab in a non-continuous culture. It is also important to keep in mind that NT-26 is a facultative chemolithoautotroph and the experiments presented in this chapter were with NT-26 grown heterotrophically, it is possible that we will observe expression of genes required for a chemolithoautotroph growth, which might be the case for *aioX, aioS* and *aioR* up-regulation in stationary phase.

Growth phase also affected the expression of *aioB* and *aioA* that were up-regulated in stationary phase by 9.57- and 9.33-fold, respectively when grown without As^{III} or As^{V}. However, *cytC* was not affected and is not up- neither down-regulated in stationary phase compared to late-log in medium containing YE alone. The up-

regulation in response to growth phase is also described for *aioB* and *aioA* in the As^{III} oxidizer *A. tumefaciens* 5A grown without As^{III} (Kashyap et al. 2006). It was suggested for *A. tumefaciens* that quorum sensing is involved on the up-regulation of *aioBA* in absence of As^{III}, in late-log phase (compared to early-log) as it requires the presence of the autoinducer ethyl acetate (Kashyap et al. 2006). In *Chloroflexus aurantiacus*, the *aioB* and *aioA* genes are located downstream the gene encoding for LuxR, a transcription activator responsible for the activation of quorum sensing related genes (Van Lis et al. 2013b). NT-26 has no *luxR* upstream of *aioB* and more studies are required to elucidate whether quorum sensing and As^{III} response are linked in NT-26.

When NT-26 was grown with As^{III} until stationary phase *aioB, aioA* and *cytC* were all up-regulated by 9.8-, 4.59- and 2.85-, respectively. That suggests that although *cytC* doesn't respond to growth phase without As^{III}, its promoter might respond to the presence of As^{V} as in stationary phase as all the As^{III} in the media was already oxidized to As^{V} (Santini et al. 2007). This hypothesis needs to be tested by measuring the concentration of As^{III} and As^{V} at the same point where the RNA was taken (section 2.3.2). Interestingly, when NT-26 is grown until stationary phase with As^{V}, *aioA* and *cytC* are not significant up-regulated (p-value >0.05), while *aioB* is up-regulated in this condition (Table 3.7), showing for the first time that *aioB* and *aioA* might not be always co-transcribed. It is not clear why NT-26 would need *aioB* to be transcribed when *aioA* is not. *aioA* and *cytC* are known to be co-transcribed in presence of As^{III} and late-log phase in NT-26 (Santini et al. 2007) and it was thought that *aioB* would be in operon with *aioA* and consequently with *cytC*, although the results presented here suggest that *aioB* can be transcribed without *aioA* and *cytC.* Another explanation would be *aioA* and *cytC* RNA degradation after transcription,

and therefore, lower detection of *aioA* and *cytC* transcripts by qPCR. More studies are needed to identify the existence of promoters between *aioB* and *aioA* (addressed on Chapter 4) and between *aioA* and *cytC* and to elucidate the difference in expression of these genes in stationary phase. The figure 3.10 below summarises the transcription of *aioB, aioA* and *cytC* in all conditions tested and highlights the hypothesis that *aioA* and *cytC* are not always in operon with *aioB* or *aioA* and *cytC* RNA is being degraded after transcription.

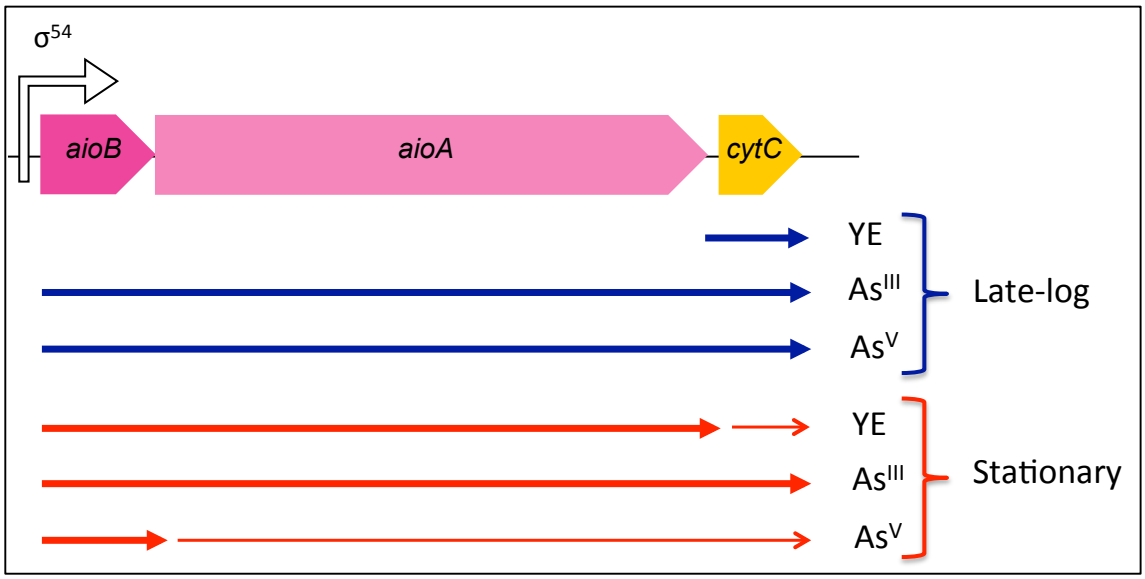


Figure 3.10: Summary of *aioB, aioA* and *cytC* transcription in NT-26 grown in presence of As^{III}, As^{V} or with YE alone and until late-log and stationary phases. The position of the genes is shown. The thicker lines represent the transcripts up-regulated on the condition tested. The thinner lines represent the presence of transcript on the condition tested based on qPCR results.

3.5 Chapter overview

- *aioX, aioS* and *aioR* are not differentially expressed in the presence of As^{III} or As^{V}.

- *aioX, aioS* and *aioR* are up-regulated in stationary phase, compared to late-log phase.

- *aioB* and *aioA* are both up-regulated in stationary phase in the absence of As^{III} or As^{V} suggesting that quorum sensing might be involved in the regulation of these genes.

- *aioB* expression is up-regulated with As^{V} in stationary phase but the same wasn't observed for *aioA*, suggesting that *aioB* and *aioA* might not be always in an operon or that *aioA* might be degraded after transcription through an RNA processing mechanism (from 3' -> 5').

- *cytC* is always expressed in basal levels in the cell and it is up-regulated in the presence of As^{III} and As^{V}.

- *cytC* expression pattern is highly correlated to *aioB* and *aioA* and although it has its own promoter, it can also be transcribed as an operon with *aioB* and *aioA.* The promoter controlling the expression of *cytC* still has to be identified.

Chapter 4

Transcription start site and promoter identification of aioX, aioB and aioA in NT-26

4.1 Introduction

To better understand the regulation of *aioXSR* and *aioBA,* it is important to identify which promoter or promoters control their expression and under which conditions these promoters are active. The first step to identify a promoter, is to define the position of the transcription start site (TSS) and then try and identify the promoter upstream of the identified TSS by using experimental and computational approaches based on the similarity of the putative promoter region with promoters identified for other bacteria. The only attempt to identify the TSS of the *aioXSR* and *aioBA* operons was made using 5' rapid amplification of cDNA ends (RACE) technique in *T. arsenitoxydans* sp. str. 3As (Moinier et al. 2014). A single TSS was identified upstream of *aioX* and a sigma 70 (σ^{70}) promoter region was found. However, the regulation of *aioX* in *T. arsenitoxydans* 3As is known to be different from NT-26 as in *T. arsenicoxydans* 3As, *aioX* and *aioB* are positioned facing opposite directions, share the same promoter region, and *aioBA* is regulated by AioF. AioF is a transcriptional regulator that binds to As^{III} and As^{V} and *aioF* is transcribed in the same operon as *aioB* and *aioA* in *T. arsenitoxydans* 3As and is absent in NT-26 (Moinier et al. 2014). In Alphaproteobacteria that have *aioXRS,* both the *aioXRS* and *aioBA* operons are transcribed in the same direction while in the Betaproteobacteria they are transcribed in opposite directions based on the As^{III} oxidisers studied so far (Sardiwal et al. 2010).

As seen in Chapter 3 the regulation of the *aioXSR* genes in response to As^{III}, As^{V} and growth phase in NT-26 is different to what has been observed for other bacteria such as *A. tumefaciens* 5A and *H. arsenitoxydans* ULPAS-1. Based on the results obtained it is proposed that expression of the *aioXSR* operon is either controlled by two different

promoters, one that promotes constitutive transcription and another that promotes transcription dependent on growth phase - likely to be more active in stationary phase, or has one single promoter and two regulators, to respond to the different stimuli. Also based on the results obtained in Chapter 3, it was suggested that *aioB* and *aioA* were up-regulated in stationary phase, when in the absence of either As^{III} or As^{V}, suggesting the presence of two promoters, one that is positively regulated by As^{III} and another that is regulated by growth phase, or it has one promoter and two regulators. Although a σ^{54} consensus sequence has previously been observed upstream of the *aioB gene* (Sardiwal et al. 2010; Andres et al. 2013), the TSS has not been determined experimentally. The 5'RACE technique will be used to determine the TSS upstream of *aioX* and *aioB*. The region upstream of the identified TSS which should contain the promoter will be cloned upstream of l*acZ* gene in the promoterless vector pPHU234 to test the activity of the promoters in NT-26 grown in the presence or absence of As^{III} or As^{V} and from early-log to stationary phase.

4.1.1 Transcriptional regulation of gene expression

The regulation of gene expression starts at the transcriptional level, where DNA is transcribed to RNA by RNA polymerase. RNA transcription starts at specific sites on the DNA upstream of the genes to be transcribed at sites known as promoters. The RNA polymerase is formed by four different subunits: α, β, β´and σ. The α β β´ together form the core of RNA polymerase and the σ factor is required to identify the promoter region upstream of the gene to be transcribed and directs the RNA polymerase to the correct site (Travers & Burgess 1969).

The specificity of the promoter region to the σ factor is crucial for the transcription of the genes. The presence of multiple σ factors and promoter regions ensure the

transcription of the right genes, whenever they are required. This makes the transcription of genes more efficient as RNA polymerase is limited and its use is controlled (Ishihama 2000). The study and identification of σ factors was mostly done in *E. coli* which has six different σ factors (σ^{54}, σ^{S} [σ^{38}], σ^{H} [σ^{32}], σ^{F} [σ^{28}], σ^{E} [σ^{24}], and σ^{FecI}) (Gross et al. 1998) which respond to different stimuli and the housekeeping σ^{70} (Ishihama 2000). The σ^{70} family controls the expression of housekeeping genes whereas the other σ factors control the response to specific stimuli and compete with σ^{70} for RNA polymerase. The promoter region upstream of a gene will have a conserved region required to bind the specific σ factor to control the expression of this gene, according to the cell requirements.

The first nucleotide to be transcribed by the RNA polymerase to form the mRNA is known as the +1 nucleotide and is the transcription start site (TSS) of the mRNA. The region between the TSS and the first codon to be translated is known as the untranslated region (UTR). The ribosome binds the UTR, at the Shine-Dalgarno sequence (SD), to initiate the translation of the gene and this region can contain regulatory elements, such as riboswitches or regulatory RNAs (rRNA). To understand how a gene is regulated it is crucial that its TSS is determined to allow the localization of the promoter region; identification of the σ factor controlling it; and to localize the binding regions of transcription factors. It is also important to determine the UTR as it can contain post-transcriptional regulatory regions such as riboswitches or small RNA target regions.

4.1.2 Aims

The three main objectives of this chapter are: 1- to test the hypothesis that the *aioXSR* operon is regulated by more than one promoter; 2- to test the hypothesis that

aioA gene is regulated by one promoter in addition to σ^{54}; 3- to test the activity of the σ^{54} promoter upstream of *aioB* and the putative promoter upstream of *aioX* in NT-26 grown in presence and absence of As^{III} or As^{V} to different growth stages.

To meet these aims, the TSS needed to be identified. The activity of the putative promoters needed to be monitored over the course of growth in NT-26 to identify any possible differential activity in response to growth phase. If active, the promoter needed to be identified by looking at conserved regions already described for σ factors in other bacteria.

4.2 Materials and Methods

4.2.1 Bacterial strains and vectors

Bacterial strains and vectors used in this chapter are shown in table 4.1.

Table 4.1. Strains and vectors used in this chapter:

Strain/Vector	Characteristics	Reference
Strains:		
Rhizobium NT-26	Spontaneous Rifampicin resistant mutant	(Santini & Vanden Hoven 2004
Escherichia coli DH5α	ϕ80dlacZΔM15, Δ(*lacZYA-argF*)U169, *deoR, recA1, endA1, hsdR17*(rk-, mk+), *supE44, thi-1, gyrA96, relA1*	Hanahan 1983
E. coli S17-1 λpir	*recA, thi, pro*, hsd^R, M+RP4: 2-Tc:Mu: Km Tn7 λpir	Simon et al. 1983
Stellar™ competent cells (*E.coli* HST08)	F–, *endA1, supE44*, thi-1, *recA1, relA1, gyrA96, phoA*, Φ80d *lacZ*Δ M15, Δ (*lacZYA - argF*) U169, Δ (*mrr - hsdRMS - mcrBC*), Δ*mcrA*, λ–	Takara Bio Inc., Shiga, Japan
Vectors:		
pPHU234	Broad-host-range *lacZ* fusion vector, Tc^R, *oriV, oriT, lacZ, trfA, trfB,* 21.6 kb	Hübner et al. 1991
pRACE	In-phusion cloning site, Amp^R, *lacZα, ori,* 2.69 kb	Takara Bio Inc., Shiga, Japan (Based on Norrander et al. 1983)

4.2.2 Stock solutions

Stock solutions were prepared at the concentrations shown in Table 4.2. The solutions were sterilised using a 0.22 µm filter (Millex – GS, Millipore) and dispensed into 1.5 ml plastic tubes. They were either stored at -20 °C or used immediately as stated.

Table 4.2. Filter sterilized stock solutions

Solution Concentration (mg/mL)	Concentration (mg/mL)	Solvent	Storage
Ampicillin (Ap)	50	MilliQ water	-20 °C
Tetracycline (Tc)	1	70% ethanol	-20 °C
Rifampicin (Rf)	100	Methanol	Made fresh

4.2.3 Media

4.2.3.1 Medium for NT-26

NT-26 was grown in McCartney bottles containing 10 ml MSM media containing 0.04 % YE with and without 5 mM As^{III} or 5mM As^{V} as described in Section 2.2.2.

For the selection of transconjugants, NT-26 was grown heterotrophically on plates containing MSM, 0.08% YE, 1.5% (w/v) purified agar (Oxoid), Rif (100 µg/ml) and Tc (10 µg/ml) at 28 °C for at least 3 days.

4.2.3.2 Medium for *E. coli*

E. coli was cultured in either lysogeny broth (LB) liquid medium or solid LB medium with addition of agar. Super optimal catabolite repression (SOC) broth was prepared and used for *E. coli* recovery after electroporation (section 4.2.4.10). The constituents of LB (liquid and solid) and SOC media are shown in Table 4.3.

Table 4.3. *E. coli* growth media constituents:

Media	Composition
LB broth	1% (w/v) tryptone*; 0.5% (w/v) yeast extract#; 1% (w/v) NaCl; pH 7
LB agar	1% (w/v) tryptone*; 0.5% (w/v) yeast extract#; 1% (w/v) NaCl; pH 7; 1.5% Difco™ agar*
SOC	2% (w/v) tryptone*; 0.5% (w/v) yeast extract#; 0.05% (w/v) NaCl. 10 mM $MgCl_2$ and 20 mM glucose (final concentration) were added after autoclaving.

* Becton, Dickinson and Company, Sparks, USA

Oxoid

4.2.4 Molecular techniques

4.2.4.1 PCR

The PCR reactions were carried out using either BioTAQ DNA Polymerase (Bioline Reagents Ltd, United Kingdom) for routine PCR, or Phusion High Fidelity DNA polymerase (Finnzymes) where cloning and sequencing were required.

For BioTAQ reactions, NH_4 reaction buffer (final concentration 1x), dNTPs (final concentration 0.2 mM), $MgSO_4$ (final concentration 3 mM), primers (1 ng/ul), BioTAQ DNA polymerase (Bioline Reagents Ltd) (1 unit/µl), 1 µl of template DNA (200 - 500 ng/µl) and PCR grade water were mixed together and made up to 25 µl final volume.

For reactions where Phusion DNA polymerase was required: GC buffer (final concentration 1 x) (Finnzymes), dNTPs (final concentration 0.2 mM), 10% DMSO, 1 U of Phusion DNA Polymerase and 1 µl of template DNA (200 - 500 ng/µl) were added and made up to 50 µl with PCR grade water.

A thermocycler (Eppendorf, Germany) was used to carry out the reactions. PCR

products were visualised by agarose gel electrophoresis (Section 3.2.5) to verify if the amplification was successful and the product was of the correct size.

4.2.4.3 PCR purification

The QIAquick PCR purification kit was used to purify PCR products following the manufacturer's protocol (QIAGEN, Crawley, UK); DNA was eluted for 3 minutes in a volume of 30 µl of nuclease free water, pre-heated at 60 °C.

4.2.4.4 Plasmid DNA isolation

Plasmid pPHU234 and recombinants were isolated from 3 ml overnight cultures of *E. coli* DH5α (Table 4.1) grown in LB containing 10 µg/ml Tc. The QIAprep Spin Miniprep Kit (QIAGEN) was used according to the manufacturer's instructions and the plasmid DNA was eluted in 30 µl of nuclease free water and incubated for 3 minutes at room temperature. The purified plasmid DNA concentration was quantified using a nanodrop spectrophotometer (Thermo Scientific NanoDrop 2000c), according to the manufacturer's specifications.

4.2.4.5 Genomic DNA isolation

NT-26 genomic DNA (gDNA) was obtained from 10 ml NT-26 cultures grown with MSM and YE as described in Section 4.2.3.1. The Wizard® Genomic DNA purification kit (Promega) was used according to the manufacturer's instructions. The gDNA concentration was estimated using a nanodrop spectrophotometer (Thermo Scientific NanoDrop 2000c).

4.2.4.6 DNA isolation from agarose gel

PCR amplicons, restriction digested fragments and plasmid DNA were electrophoresed in an agarose gel and visualised with a UV transilluminator. The band of interest was excised from the gel using a clean scalpel. The DNA was isolated using the QIAquick Gel Extraction Kit (Qiagen), according to the manufacturer's specifications. The concentration of the purified product was determined using a nanodrop spectrophotometer.

4.2.4.7 Restriction endonuclease digestion

PCR purified DNA, plasmid and genomic DNA, were digested using restriction endonucleases in a total reaction volume of 20 µl at 37°C for 3 h to overnight. The reactions contained a 1 X concentration of enzyme specific buffer and 1 U of enzyme (Promega) per µg of DNA.

4.2.4.8 DNA ligation using T4 DNA ligase

For DNA ligations using T4 DNA ligase (Promega), 0.5 U T4 DNA ligase was added to the reaction with 1 X ligase buffer (Promega). After being restriction digested and gel purified, the DNA (insert) and pPHU234 plasmid were mixed with the buffer and T4 DNA ligase to a final volume of 10 µl. The reactions were incubated at 4 °C overnight then microdialysed on 0.02 µm dialysis paper (Millipore, Watford, UK) for 45 minutes. The dialysed ligation reaction was kept at -20 °C until used.

4.2.4.9 Preparation of *E. coli* electrocompetent cells

For the preparation of electrocompetent cells, a protocol was based on that of Ausubel et al. (1987) was followed. *E. coli* DH5α was grown overnight in 10 mL LB

broth at 37 °C with shaking at 180 rpm. The overnight culture was diluted in fresh LB medium (1 ml in 200 ml LB) in a 1 L Schott bottle and the cells were grown at 37 °C with shaking at 180 rpm until the culture reached an $OD_{600\ nm}$ of 0.5-0.7. The cells were left on ice for 30 min. The cells were harvested by centrifugation at 1852 x *g* for 10 min at 4 °C, the supernatant discarded and the pellet suspended in 200 ml sterile ice cold dH_2O and incubated on ice for 10 min. The cells were centrifuged again for 10 min at 5143 x *g* at 4 °C, the supernatant discarded and the cells again suspended in cold dH_2O, and left on ice for 10 min. The cells were centrifuged for the third time, the supernatant discarded and the cells suspended in approximately 40 ml of ice cold 10% glycerol. The cells were incubated on ice for 30 min and then centrifuged for 10 min at 5142 x *g*. The pellet was suspended in 0.4 ml of ice cold 10% glycerol and dispensed into sterilized microcentrifuge tubes (100 µl per tube). The cells were snap frozen in liquid nitrogen and stored at -80 °C until needed. All centrifugation steps were carried out using a Beckman Coulter Avanti J-26XP centrifuge (Beckman Coulter, USA) and pre-chilled centrifuge bottles.

4.2.4.10 Electroporation and colony selection

Electrocompetent *E. coli* DH5α cells (Section 4.2.4.9) were thawed on ice. DNA was added to the electrocompetent cells and the mix was incubated on ice for 10 min along with the appropriate electroporation cuvettes (0.2 cm electrode gap, Bio-Rad, UK). Cells were shocked using a Bio-Rad *E. coli* pulser (Bio-Rad) for 1 pulse at 2.5 kV. 1 mL of SOC (Table 4.3) was added to allow the cells to recover and the mix was transferred to a 25 mL sterile plastic tube. The *E. coli* DH5α cells were then incubated at 37 °C with shaking at 180 rpm for 1 hour. After recovery, serial ten-fold dilutions in SOC were prepared and 100 µl volumes of each dilution were spread onto LB agar

plates (Table 4.3) supplemented with 10 µg/ml of Tc, when stated. The plates were incubated overnight at 37 °C.

4.2.4.11 Transformation using $CaCl_2$ method

E. coli str. S17-1 λpir cells (Table 4.1) were transformed with pPHU234 by itself or pPHU234 ligated to the putative promoters using a quick $CaCl_2$ method. 5 ml of *E. coli* str. S17-1 λpir culture was grown in LB (Table 4.3) to mid-log phase (OD_{600nm} 0.5-0.7) and harvested by centrifugation at 10,000 x *g* for 5 min. The cells were suspended in 1 ml ice-cold 10 mM $CaCl_2$ and transferred to a 1.5 ml tube. 500 µl of ice-cold 75 mM $CaCl_2$ was gently mixed to the suspension. The cells were centrifuged at 10,000 x *g* for 5 min, the supernatant discarded and the cell pellet suspended in 100 µl of ice-cold 75 mM $CaCl_2$. 300 ng DNA was added to the cells and incubated on ice for 2 h. The cells were streaked onto LB agar (Table 4.3) containing 10 µg/ml Tc (Table 4.2) and the plates were incubated overnight at 37 °C.

4.2.5 Determination of TSS using SMARTer RACE 5´ Kit

The determination of the TSS of *aioX, aioB* and *aioA* in NT-26 was done using the SMARTer RACE 5´ Kit from Clonetech (Takara Bio Inc., Shiga, Japan) following the manufacturer's specifications. All the reagents, cells, enzymes and vector mentioned in this section were provided with the kit apart from the gene specific primers designed for *aioX* and *aioA* (section 4.2.5.1). Figure 4.1 shows an overview of how the SMARTer 5´RACE system works to synthesise the 5´ready cDNA and Figure 4.2 exemplifies the 5´RACE PCR.

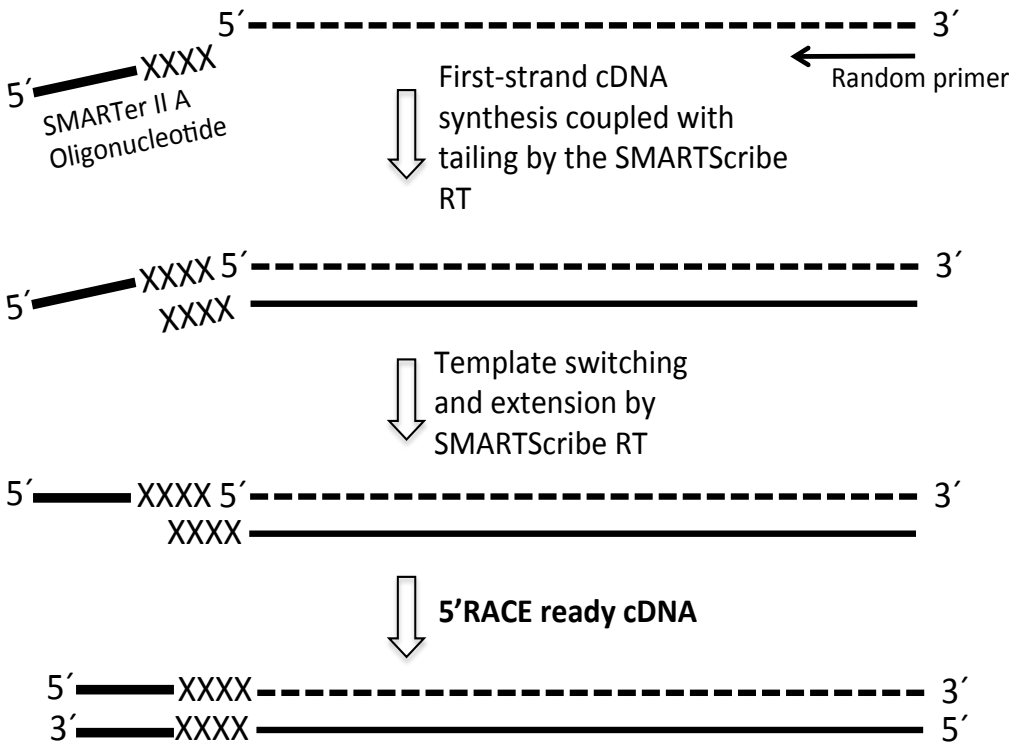


Figure 4.1. Mechanism of SMARTer cDNA synthesis. The random primer anneals at the RNA to synthesize the first-strand cDNA. The SMARTScribe Reverse Transcriptase (RT) adds a few nucleotides to the 3´end of the first strand cDNA when it reaches the 5´end of the RNA. The SMARTer II A Oligonucleotide anneals to the 3' tail of the newly formed cDNA and serves as an extended template for SMARTScribe RT.

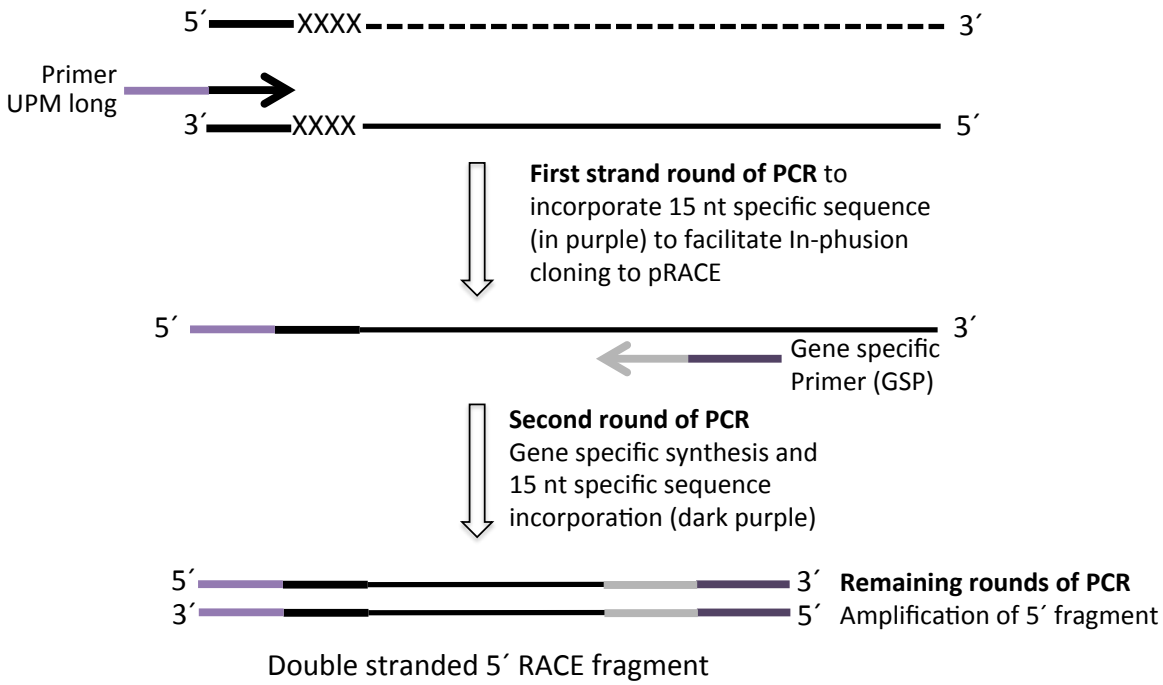


Figure 4.2. Overview of the 5´RACE PCR. The universal primer (UPM) anneals at the 5´ portion of the cDNA, in the first round of PCR to incorporate 15 nt with specific sequence to the newly formed fragment. The GSP primer is designed to bind and amplify specifically to gene of interest and to add 15 nt to with specific sequence to facilitate In-phusion cloning to the pRACE vector.

4.2.5.1 Primer design for SMARTer RACE 5´

According to SMARTer RACE 5´ kit specifications, two sets of primers specific to the target gene – *aioX* and *aioA* - were designed (Table 4.4). The gene specific primer (GSP) and the nested GSP (NGSP) were designed to be specific to bind either *aioX* or *aioA* in the reverse orientation of the gene, be 23–28 nt long to ensure specific annealing, 50–70% GC and have a Tm ≥65 °C. Both primers had 15 nt overlaps with the pPRACE vector, to allow In-phusion cloning, by adding the sequence GATTACGCCAAGCTT to the 5´ends of both GSPs and NGSP sequences. The GSP binding on *aioX* was designed to amplify a fragment with a minimum size of 914 bp

(Figure 4.3). The GSP binding on *aioA* was designed to amplify all the way to *aioB,* located upstream *aioA* – generating an amplicon with a minimum size of 917 bp - and also the TSS of *aioA,* if existent – generating an amplicon with the minimum size of 377 bp (Figure 4.4).

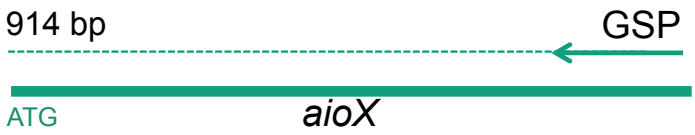


Figure 4.3. Position of the *aioX* GSP designed to amplify 914 bp of *aioX* in a reaction with UPM primer (Clontech).

Figure 4.4. Position of the *aioA* GSP designed to amplify 917 bp of a fragment containing a small portion of *aioA* and the entire *aioB*. In case of *aioA* having its own TSS, the *aioA* GSP will amplify a band of minimum 377 bp when used in a reaction with the UPM (Clontech).

Table 4.4. Primer design for TSS determination of *aioX* and *aioA* using 5´RACE:

Primer	Sequence (5'-3')	GC%	Tm (°C)	Product size (bp)
***aioX* GSP**	GATTACGCCAAGCTTCATCCCAGCCTCCGCA CGCGTTCCA	60	66	> 914
***aioX* NGSP**	GATTACGCCAAGCTTTCGGGTGAGCACCGA GCGTCCGAGA	68	66	> 828
***aioA* GSP**	GATTACGCCAAGCTTTCGGTGAGGCGCTGC TGCTGGGTGT	60	>75	> 917
***aioA* NGSP**	GATTACGCCAAGCTTCCAGGCGTCGCTTTCC GCCTGCTGCT	61	>75	> 674

4.2.5.2 RNA samples used for TSS determination

The NT-26 RNA samples used to determine the TSS of *aioX* and *aioB* were obtained from NT-26 grown without As^{III} until late-log (named sample 3, with OD_{600} 0.140) and stationary phase (named sample 10, with OD_{600} 0.171) and with NT-26 grown with As^{III} until late-log (named sample 15, obtained from OD_{600} 0.126) and stationary phases (named sample 19, obtained from OD_{600} 0.214). The RNA samples were the same used in Chapter 2 and 3 for the qPCR work that were obtained as described in Chapter 2, Section 2.2.4, Table 2.4.

4.2.5.3 5'RACE-ready cDNA synthesis

To generate full-length cDNAs, the SMARTScribe Reverse Transcriptase was used in the reverse transcription reactions. The terminal transferase activity of the SMARTScribe RT adds a few additional nucleotides to the 3´end of the first-strand of cDNA when it reaches the 5´end of the RNA (Figure 4.1).

For each 20 µl 5´RACE ready cDNA synthesis reaction, 5 µl RNA (150 ng/µl) was mixed with 1 µl random primer mix (10x stock concentration) and 5 µl of molecular grade water in a microcentrifuge tube and incubated for 3 minutes at 72 °C, followed by 2 minutes at 42 °C. After the incubation the following was added to the tube: 1 µl of SMARTer II A Oligonucleotide, 4 µl first strand buffer, 0.5 µl Dithiothreitol (DTT) (100 mM), 1 µl dNTPs (20 mM), 0.5 µl RNAse inhibitor (20 U/µl) and 2 µl of SMARTScribe Reverse transcriptase (100 U). The microcentrifuge tube was incubated at 42 °C for 90 minutes followed by 10 minutes at 72 °C. 10 µl of Tricine EDTA buffer (10 mM Tricine and 1 mM EDTA, stock concentration) was added to the 20 µl 5´RACE ready cDNA sample that was then stored at -20 °C until used.

4.2.5.4 5´Rapid Amplification of cDNA Ends PCR (5'RACE)

To generate 5´cDNA, the 5´RACE protocol was followed according to the manufacturers specification (Clontech, Takara Bio Inc., Shiga, Japan). For a 50 µl reaction, 2.5 µl 5´RACE ready cDNA sample was mixed with 5 µl Universal Primer Mix (UPM), 1 µl *aioX* or *aioA* GSP (10 µM) (Table 4.4), 15.5 µl H_2O, 25 µl SeqAmp buffer and 1 µl SeqAmp DNA Polymerase in a PCR tube. The reaction was mixed and put into a thermocycler (Eppendorf, Germany) for 25 cycles of 94 °C for 30 seconds, 68 °C for 30 seconds and 72 °C for 2 min.

5 µl of the resulting PCR was separated in an agarose gel by electrophoresis (Section 3.2.5) In the absence of a band with the expected size, or the presence of multiple bands, a nested PCR using the product of the 5´RACE PCR could be carried out. The nested PCR can be performed using either *aioX* NGSP or *aioA* NGSP (Table 4.4) and the UPM primer (provided with the kit). 5 µl of the 5´ RACE PCR was diluted with 245 µl tricine EDTA buffer. The PCR steps were repeated using 5 µl of the diluted PCR, 1 µl NGSP and 1 µl UPM short. The PCR tube was put into a thermocycler for 20 cycles of the aforementioned PCR conditions.

The PCR products were separated on an agarose gel by electrophoresis and the expected sized bands (bigger than 914 bp for *aioX* fragments, bigger than 917 for *aioB* and bigger than 377 bp but smaller than 917 bp for *aioA* TSS) were extracted from the gel using the reagents and instructions provided with the kit.

4.2.5.5 In-phusion cloning

To clone the 5´ RACE product, 7 µl of the gel purified RACE product (section 4.2.5.4, above) was mixed with 1 µl linearized pRACE vector and 2 µl In-phusion HD master

mix in a PCR tube. This mixture was incubated for 15 minutes at 50 °C and transferred to ice immediately after according to the manufacturer's specifications.

The In-phusion reaction mixture was transformed into Stellar competent cells (table 4.1) using the following protocol: in a 14 ml round bottom tube, 2.5 µl of the In-phusion reaction was mixed with 50 µl Stellar chemically competent cells previously thawed on ice. The cells were left on ice for 30 min before being heat shocked for 45 seconds at 42 °C, followed by incubation on ice for 2 minutes. SOC medium (table 4.3) was added to a final volume of 500 µl and the reaction was incubated at 37 °C for 1 hour with shaking at 180 rpm. 100 µl of 1/20 serial dilution were spread on LB agar containing 100 µg/ml of ampicillin (table 4.3). The plates were left overnight at 37 °C and 5-10 independent colonies were selected and streaked onto the same medium.

To confirm the presence of insert in the selected colonies, a colony PCR was performed. This was done by mixing a colony with 30 µl of PCR grade water and incubating the mixture at 99 °C for 10 minutes. 1 µl of this mix was added to the PCR (section 4.2.4.1) along with M13 forward primer - to anneal on the vector - and a *aioX* or *aioA* GSP reverse primer. The *aioX* transformants containing inserts bigger than 914 bp were selected for sequencing and the same was done for *aioB* fragments bigger than 917 bp or in between 377 -917 bp for *aioA* TSS.

4.2.5.6 Sequencing

To identify the sequence of the 5´portion of mRNA, the plasmid containing the insert was sequenced by GATC Biotech (Germany) using LIGHTrun™ Sanger technology (GATC Biotech AG). The samples were prepared as directed by GATC in volumes of 10 µl containing 5 µl of a 80-100 ng/µl plasmid and 5 µl of 5 pmole/µl of the specific

primer. The primers used were the M13 Forward to anneal on the pRACE vector end and the *aioX* GSP to anneal on the *aioX* specific end as illustrated in Figure 4.5. The primers used to sequence the fragment containing *aioB* and *aioA* are shown in Figure 4.6.

Figure 4.5. Scheme of the primers M13 Forward and GSP used to sequence *aioX* UTR. The green lines correspond to the *aioX* insert with its UTR and the black line represents the pRACE vector.

Figure 4.6. Scheme of the primers M13 Forward and GSP used to sequence *aioB* and *aioA* UTR. The pink line corresponds to the *aioB* and *aioA* insert with its UTR and the black line represents the pRACE vector ends.

MEGA 6.0 (Tamura et al. 2013) was used to analyse the chromatogram obtained from GATC Biotech. The sequences obtained from the different clones were aligned to either the *aioX* (accession number: NT26v4_p10026), *aioB* (accession number: NT26v4_p10029) or *aioA* (accession number: NT26v4_p10030) sequences using ClustalW (Thompson et al. 1994).

4.2.6 Determination of promoter activity in NT-26

4.2.6.1 Amplification of putative *aioX, aioB* and *aioA* promoter fragments

To test the activity of the putative promoters upstream of *aioX, aioB* or *aioA,* fragments corresponding to the putative promoter region needed to be amplified and inserted into the vector pPHU234 (Table 4.1) downstream of a promoterless *lacZ* gene (Figure 4.16). Eight different primers were designed to amplify the NT-26 putative promoter regions upstream of *aioX*, one pair of primer was designed to amplify the region that comprises the σ^{54} promoter upstream *aioB* (Figure 4.14), and another pair was designed to amplify the putative promoter upstream of *aioA* (Figure 4.15). The restriction sites *Bam*HI and *Pst*I (Promega) were added to the 5´ends of the forward and reverse primers, respectively (Table 4.5). Three forward primers (named p1.1, p1.2 and 1.3) were designed to amplify the promoter 1 region of *aioX* and to be used in conjunction with the same reverse primer (p1 R) (Figure 4.11). Three more forward primers were designed to amplify the promoter region 2 (named p2.1, p2.2 and p2.3), with another reverse primer (p2 R) (Figure 4.13).

Table 4.5. Primers designed to amplify the putative *aioX*, *aioB* and *aioA* promoter regions.

Primer	Sequence (5'-3')	Target	Tm (°C)	Expected product size*
p1.1 F	GGATCCGAGGCAGGCATAGGAGGTTT	*aioX* promoter 1	68	235 bp
p1.2 F	GGATCCTCTTCATCCTGTTCATGCAATCG	*aioX* promoter 1	66.7	127 bp
p1.3 F	GGATCCTATAGAGGACGTAAATTCAGGCG	*aioX* promoter 1	66.7	74 bp
p1 R	CTGCAGTTCCAGAGCCGTCACTCAGA	*aioX* promoter 1	68	-
p2.1 F	GGATCCGCTTATACCTCGCCACTGCC	*aioX* promoter 2	69.5	219 bp
p2.2 F	GGATCCGGCCTCCAATAAGGCAGGAA	*aioX* promoter 2	68	150 bp
p2.3 F	GGATCCGGCGCAGCTAAAGCTTACTT	*aioX* promoter 2	66.4	66 bp
p2 R	CTGCAGTATGCCTGCCTCGGACCAAT	*aioX* promoter 2	68	-
p*aioB* F	GCCGGATCCAGGCCAGAAAATTCAACGCC	*aioB promoter*	71	51 bp
p*aioB* R	CGGGCTGCAGGGAGGATCTGCAATCGTTGTG	*aioB promoter*	73	-
p*aioA* F	GGATCCTCGGTCTTCGACCCTGAAA	*aioA promoter*	66	114 bp
p*aioA R*	CTGCAGACGCCTTCGGCAAAGATGTC	*aioA promoter*	68	-

*Expected product size when the Forward and Reverse (F and R) primer pairs for the same promoter are used together.

4.2.6.2 Transfer of pPHU234 and recombinant versions into NT-26

The vector pPHU234 harbouring either the *aioX, aioB* or *aioA* promoter regions was transferred to the Rif resistant NT-26 by bacterial conjugation. *E. coli* str S17-1 λpir (Rif sensitive) containing either the empty pPHU234 or the pPHU234::P*aiox*1, pPHU234::P*aiox*2, pPHU234::P*aioB* or pPHU234::P*aioA* promoter (donor) was grown in 10 ml LB until it reached OD_{600nm} 0.5-0.7, the NT-26 (recipient) was grown in MSM with 0.04% YE until mid- to late-log phase (OD_{600nm} 0.07 - 0.14). The cells were mixed in a proportion of 1:1 (based on the OD_{600nm}) and filtered onto a sterile 25 mm 0.22 µm filter (Millipore). The filter was placed facing upwards on a LB agar plate (Table 4.3) and incubated overnight (16 to 20 hours) at 28 °C. The filter was then washed in 2 ml MSM and the suspension plated onto MSM agar containing 0.08% yeast, 100 µg/ml Rif and 5 µg/ml Tc. The plates were incubated at 28 °C until

single colonies were visible (~ 5 days) and then subsequently purified by streaking onto the same solid selective medium at least 3 times. Single colonies were inoculated into 10 ml MSM containing 0.04% yeast extract and 5 µg/ml Tc at 28 °C with and without As^{III} and As^{V} as described in section 2.2.1. The liquid cultures were sub-cultured at least 5 times before the promoter studies were undertaken.

4.2.6.3 β-galactosidase assay to measure promoter activity

To quantify the promoter activity, β-galactosidase activity was measured according to the Zhang & Bremer (1995) protocols. The solutions used were prepared according to the Zhang & Bremer (1995) specifications (Table 4.6). The β-galactosidase assay measures the transcription of the *lacZ* gene that encodes the β-galactosidase. The β-galactosidase is able to cleave the ONPG substrate releasing o-nitrophenol that has a detectable yellow colour. As the pPHU234 has no promoter upstream of the *lacZ*, the β-galactosidase activity is a measure of the activity of the promoter inserted in the pPHU234.

Table 4.6. Solutions used on β-galactosidase assay:

Solution	Component	Concentration
Permeabilization solution	Na_2HPO_4	50 mM
	KCl	20 mM
	$MgSO_4$	2 mM
	CTAB (hexadecyltrimethylammonium bromide)	0.8 mg/ml
	Sodium deoxycholate	0.4 mg/ml
	β -mercaptoethanol	5.4 µl/ml
Substrate solution	Na_2HPO_4	60 mM
	NaH_2PO_4	40 mM
	O-nitrophenyl-β-D-Galactoside (ONPG)	4 mg/ml
	β-mercaptoethanol	2.7 µl/ml
Stop solution	Na_2CO_3	1 M

NT-26 containing pPHU234 – and the recombinant versions of the plasmid - was grown in MSM containing 0.04% YE, 5 µg/ml Tc with and without 5 mM As^{III} or 5 mM As^{V} overnight at 28 °C with shaking at 150 rpm. The culture was diluted 1:20 in fresh MSM medium containing 0.04% YE and 5 µg/ml Tc (0.5 ml in 10 ml MSM) in a McCartney bottle and the cells were grown at 28 °C with shaking at 150 rpm until the culture reached the desired phase of growth.

A 3 ml aliquot of NT-26 culture was taken at early-log (OD_{600nm} between 0.030 and 0.058), mid-log (OD_{600nm} between 0.070 and 0.098), late-log (OD_{600nm} between 0.115 and 0.140) and stationary phases of growth (OD_{600nm} higher than 0.200). 1 ml was used to measure the OD_{600nm} and the remaining 2 ml were centrifuged in a 14 ml round bottom tube at 3134 x g for 10 minutes at 4 °C. The supernatant was discarded and the cells were suspended in 20 µl fresh MSM. 20 µl of the suspended cells were added to a microcentrifuge tube containing 80µl permeabilization solution and the tubes were kept at room temperature until the end of the time course experiment. After all the samples were taken they were incubated at 28 °C for 30 min. 600 µl of substrate solution (pre-warmed at 28 °C) was then mixed with the cells and the time of addition was recorded for each sample. The samples were kept at 28 °C until a yellow colour was visible, at this point 700 µl of stop solution was added and the time recorded again. After the stop solution was added to the last sample, the tubes were centrifuged at 3134 x *g* for 10 minutes and the supernatant was transferred to a glass cuvette. The absorbance of the yellow colour was measured at OD_{420nm} in a spectrophotometer. β-galactosidase activity was calculated using the Miller units equation, based on the initial OD_{600nm} of the culture, the time it took for the β-

galactosidase reaction to occur and the final OD_{420} as shown in the following equation:

$$\text{Miller units: } 1000 * \frac{\text{OD420nm}}{(\text{OD600nm culture} * \text{cells volume (20 µl)} * \text{reaction time in min}}$$

According to Miller, a fully induced promoter has more than 1000 Miller units (Miller, 1972), however this value can vary for different organisms. No β-galactosidase activity is expected with the negative control (i.e. pPHU234 alone) or if the entire putative promoter is absent. Without a promoter upstream the *lacZ* gene, the β-galactosidase cannot be encoded and the ONPG is not cleaved to release o-nitrophenol. Without o-nitrophenol, no yellow colour can be detected.

4.3 Results

To better understand regulation of the *aioXSR* operon, the promoter regions upstream of these genes were studied. The possibility of more than one promoter came from results in Chapter 3, which showed that the operon was up-regulated in stationary phase. The regulatory region upstream of *aioB* was also studied to confirm that *aioB* is only under the control of σ^{54} promoter. The first part of the study was to determine the *aioX* and *aioB* transcription start sites and to then use this information to map the promoter regions and identify when they are functional.

4.3.1 Transcription start site identification

4.3.1.1 5´RACE *to identify the TSS upstream of* aioX *in NT-26*

To determine the NT-26 *aioX* TSS, the SMARTer RACE 5´kit was used to amplify the 5´portion of the mRNA (figure 4.1), each reaction was started with RNA obtained from NT-26 grown in the presence or absence of As^{III} and until late-log or stationary phase. RNA obtained from different conditions were used, as the existence of more than one promoter controlling the expression of *aioX* would create more than one TSS – one for each promoter.

To amplify the UTR of *aioX*, cDNA was prepared with RNA obtained from NT-26 grown without As^{III} until late-log (Sample 3) and stationary phase (Sample 10) and with As^{III} until late-log (Sample 15) and stationary phase (Sample 19). The 5´RACE PCR was performed using the 5´RACE ready cDNA along with *aioX* GSP and UPM primers (Figures 4.2 and 4.3). After applying the PCR samples in an agarose gel, it was expected that bands bigger than 914 bp would be observed with the amplification of the *aioX* UTR.

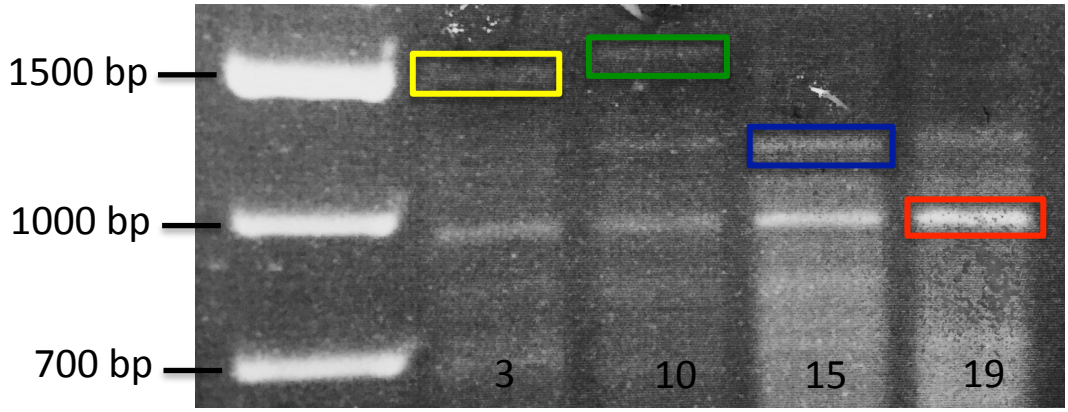


Figure 4.7. Agarose gel photo to confirm amplification of 5´UTR region of *aioX*. The PCR products from samples 3 (YE late-log), 10 (YE stationary), 15 (As^{III} late-log) and 19 (As^{III} stationary) are shown. The sizes of the molecular markers are shown on the left. The yellow (1500 bp), green (approximately 1600 bp), blue (1200 bp) and red (950 bp) squares highlight the bands that were excised from the gel to identify the UTR upstream of *aioX*.

Four different sized bands were observed on the gel (Figure 4.7, yellow, green, blue and red boxes) greater than 914 bp and contained putative *aioX* TSS. A band corresponding to 950 bp was seen in all the samples and this band from sample 19 was excised for identification (Figure 4.7, red box). A band corresponding to approximately 1200 bp was observed for the samples 10, 15 and 19 and the band corresponding to sample 15 was extracted (Figure 4.7, blue box). Single bands of about 1500 bp and greater than 1500 bp were also visualised in samples 3 and 10, respectively; both were excised and analysed further (Figure 4.7, yellow and green boxes, respectively).

The four different sized bands were purified from the gel, ligated to the linearized pRACE vector and transformed into Stellar competent cells. Colony PCR was used to identify recombinant plasmids using *aioX* GSP and M13 Forward primers that anneal to *aioX* and the vector, respectively.

Recombinant plasmids carrying the four different sized bands were isolated and the clones sequenced.

4.3.1.2 *aioX* UTR identification

The sequences corresponding to the 1500 bp, 1600 bp, 1200 bp and 950 bp putative UTRs were aligned against *aioX* to assess specificity. The sequences obtained from the three different clones from the 1500 bp sample were not specific to *aioX*, the fragment amplified was similar to a fragment from the NT-26 chromosome due to mispriming and this sample was not considered for *aioX* UTR identification. The same result was obtained for the 1600 bp putative UTR as the fragment amplified also corresponds to a sequence from the chromosome and the sample was discarded from the UTR analyses. The use of an universal primer (UPM) in the PCR reaction increase the chances of mispriming even though the *aioX* GSP was designed to bind specifically to *aioX*. The sequence of the samples corresponding to the 1200 bp and 950 bp bands in figure 4.5 were specific to *aioX* and were considered for further UTR identification.

The four samples corresponding to the red PCR band were aligned together and along with *aioX* specific sequencing to facilitate the search for the ATG of *aioX*. The sequence upstream of the *aioX* ATG was aligned against the pRACE and UPM specific sequences to exclude the region specific to the vector and primer, respectively. The final sequence was consistent between the four different clones. The UTR was identified and the TSS is the first 5´nucleotide of the UTR. These were named UTR 1 and TSS 1 (Figure 4.8).

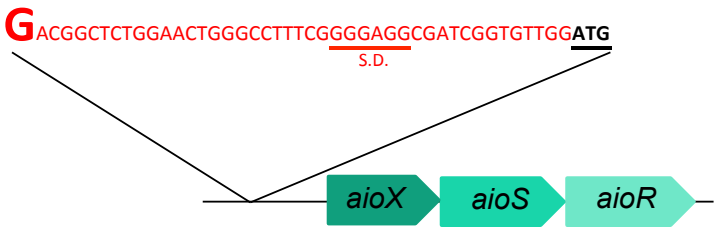


Figure 4.8. TSS 1 and UTR 1 upstream of *aioX*. The UTR sequence can be seen in red and the TSS 1 is in bold and capitalized. The Shine Dalgarno (SD) conserved sequence is underlined.

The same was done for the blue samples and a different UTR was identified. UTR 2 is highlighted in blue, with the TSS 2 shown in bold as seen in Figure 4.9, below:

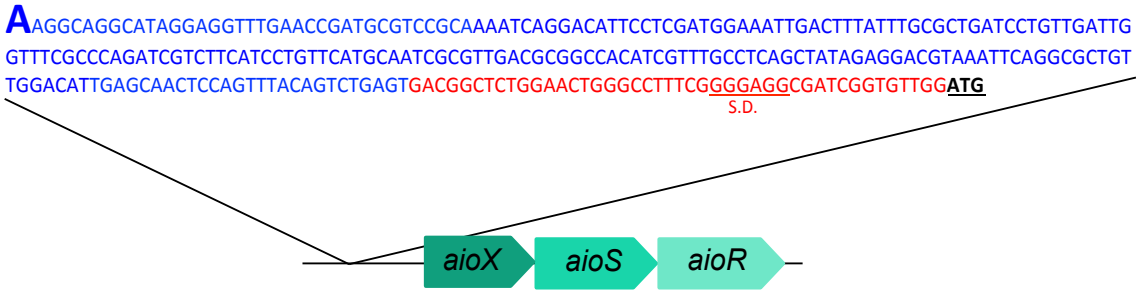


Figure 4.9. TSS 2 and UTR 2 upstream of *aioX*. The UTR 1 sequence can be seen in red and the TSS 2 is in blue, bold and capitalized. The SD conserved sequence is underlined.

4.3.1.3 5´RACE to identify the TSS upstream *aioB* and *aioA* in NT-26

To confirm the position of the TSS upstream of *aioB* in NT-26 and to determine whether *aioB* is controlled by more than one promoter, the 5'RACE technique was used. The RNA sample used (Sample 19 [Chapter 2, section 2.2.4]) was isolated from NT-26 grown with As^{III} to stationary phase, as it is known from Chapter 3 that the abundance of *aioB* mRNA is higher. The *aioA* GSP designed to determine the *aioB* TSS anneals to the start of the *aioA* gene (Figure 4.4). In the case of *aioA* having its own TSS, two bands would be expected in a gel after the 5'RACE PCR: one larger than 917

bp for a single mRNA with *aioB* and a small part of *aioA*, and another one larger 377 bp for only a portion of the *aioA* mRNA. If *aioB* and *aioA* are transcribed as a single transcriptional unit, one or more bands would be observed corresponding to the portion of *aioA* and *aioB* transcript in one operon at more than 917 bp as this is the distance between the *aioA* GSP and the *aioB* ATG (Figure 4.10).

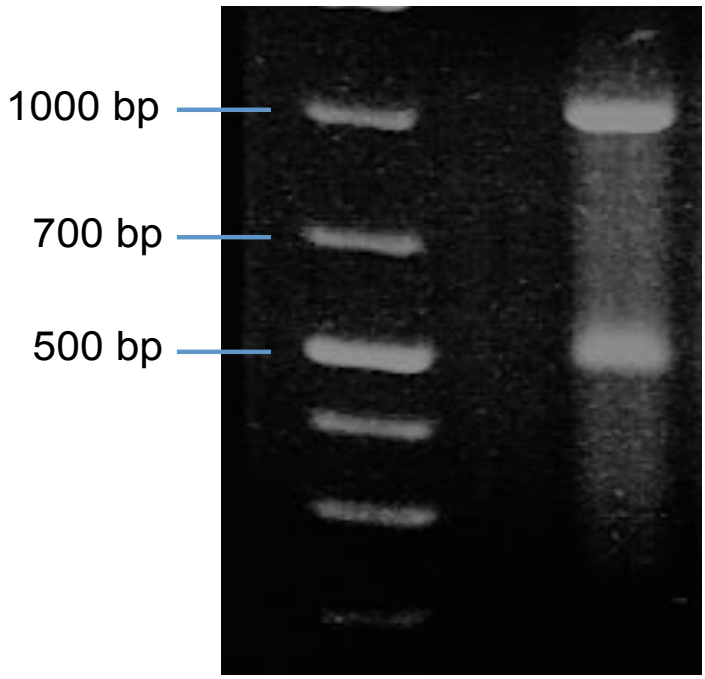


Figure 4.10. Agarose gel photo to confirm amplification of 5' region of *aioB* and portion of *aioA* mRNA. The sizes of the molecular markers are shown on the left. The upper band is the expected size for an mRNA containing *aioB* and a portion of *aioA* (> 917) and the smaller band is the expected size for an mRNA containing a portion of *aioA* only (> 377).

Two different bands corresponding to two sizes were observed, the largest (1 kb) corresponds to the expected size for a fragment containing *aioB* and portion of *aioA* and the smaller one seem to be from the portion of *aioA* and its own TSS (500 bp). The two bands were purified from the gel, ligated to the pRACE vector and transformed into Stellar competent cells. Transformants (5 of each) were grown in liquid LB and the plasmids were isolated and sent for sequencing to identify the TSSs.

The RNA sample 15 (Sample 15 [Chapter 2, section 2.2.4]), isolated from NT-26 grown with As^{III} to late-log was also used to observe if a different sized band would be amplified. However, the same two bands were amplified - one with 1 kb and another with 500 bp - and only the bands obtained from the reaction with the Sample 19 were used on the following experiments.

4.3.1.4 *aioB* and *aioA* UTR identification

To identify the *aioA* UTR the primers *aioA* GSP and the M13 Forward primers were used for sequencing (Figure 4.6). Three different clones containing the putative *aioA* smaller band (500 bp) were sequenced and the sequences were aligned against the primers M13 Forward and *aioA* GSP and the *aioA* sequence. All the sequences were specific to *aioA* and the identified TSS and UTR are shown in Figure 4.11. The SD sequence was identified upstream of the *aioA* ATG and indicates that *aioA* might indeed have a functional promoter controlling its expression (Figure 4.11).

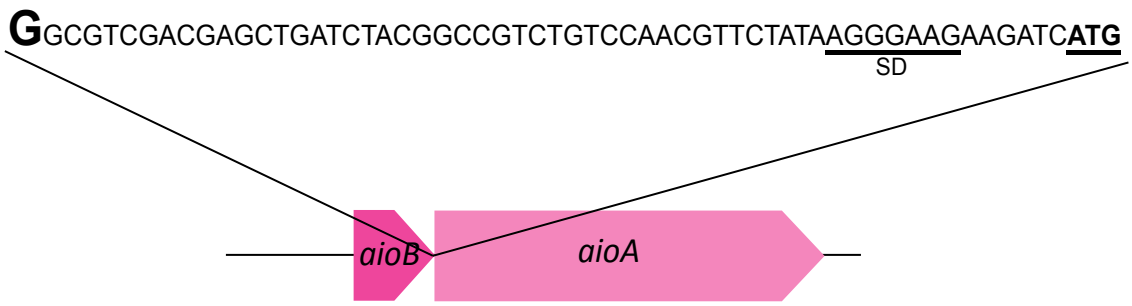


Figure 4.11. TSS and UTR upstream of *aioA*. The TSS is shown in bold and the UTR is shown in between the TSS and the first transcribed codon - ATG, underlined and in bold. The SD conserved sequence is underlined.

To identify the TSS upstream of *aioB*, the primers *aioA* GSP and M13 Forward primer were used to sequence the pRACE vector carrying the 1 kb fragment (Figure 4.6). Five different clones were sequenced and aligned to the *aioB* sequence to locate the ATG of *aioB* and the M13 Forward primer to locate the sequence specific to the vector. The UTR upstream of *aioB* was identified, the TSS and the SD conserved sequence are shown in the figure 4.12, below.

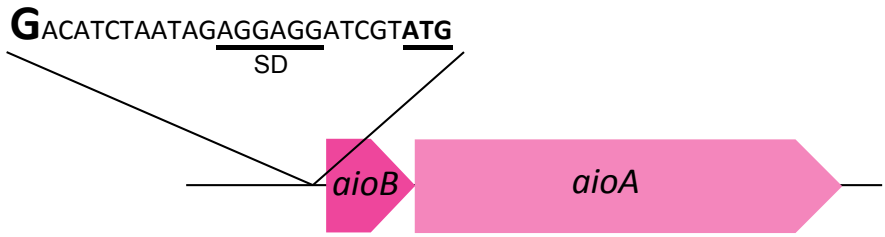


Figure 4.12. TSS and UTR upstream of *aioB*. The TSS is in bold. The Shine Dalgarno (SD) conserved sequence is underlined.

4.3.2 *aioX, aioB* and *aioA* promoter activity in NT-26

The presence of two different TSS upstream of *aioX* supports the idea that expression of the *aioX, aioS* and *aioR* genes is controlled by two different promoters, as previously suggested in Chapter 3. Promoters regions are most commonly found flanking the -10 and -35 positions upstream of the TSS which is designated as the +1 position. To be able to localise the promoters upstream of *aioX* and to verify that they are functional, the region upstream of the TSS was cloned upstream of a promoterless *lacZ* gene in plasmid pPHU234 (figure 4.13). The plasmid was transferred into NT-26 by conjugation and then β-galactosidase activity measured in NT-26 cells grown with and without AsIII and AsV at different growth phases.

The presence of a TSS upstream of *aioB* is expected as *aioB* and *aioA* have been shown to be co-transcribed and their transcription to be under control of the σ^{54} promoter. The σ^{54} consensus sequence is located at exactly the -12 and -24 positions upstream the TSS identified for *aioB* (Figure 4.15). The presence of a promoter upstream of *aioA* was unexpected as it was thought that *aioB* and *aioA* are always co-transcribed. To check the functionality of this new putative promoter upstream of *aioA* and the σ^{54} promoter upstream of *aioB*, the regions upstream of each TSS were cloned and tested as described for the putative *aioX* promoters.

4.3.2.1 Cloning the putative promoter regions upstream of *aioX, aioB* and *aioA* into pPHU234

To determine where a regulator could bind upstream of *aioX* and if the TSS determined by 5' RACE is a true TSS, three different sized DNA fragments were amplified upstream of both TSS 1 and TSS 2. The different sized fragments were amplified to try to narrow search area for the promoter. The fragments p1.1, p1.2 and p1.3 were amplified for TSS 1 containing 235 bp, 127 bp and 74 bp, respectively. The fragments p2.1, p2.2 and p2.3 were amplified for TSS 2 with 219 bp, 150 bp and 66 bp, respectively (Figure 4.13).

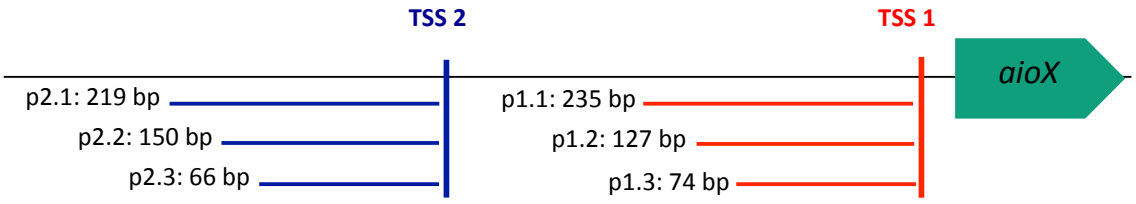


Figure 4.13. Position and size of the fragments amplified upstream the TSS 1 and TSS 2 of *aioX* to test the promoter activity.

To determine *aioB* σ^{54} promoter activity and test if *aioA* is a functional promoter, the region upstream the TSS of *aioB* and *aioA* were also amplified. The primers were designed to amplify a 51 bp fragment upstream *aioB* containing the σ^{54} promoter consensus region and named PaioB (Figure 4.14). For *aioA* putative promoter test, the primers were designed to amplify a 114 bp fragment upstream the TSS, named PaioA (Figure 4.15). All the primers used to amplify the *aioX, aioB* and *aioA* fragments were designed by adding a *Bam*HI restriction site for the forward primer and a *Pst*I site for the reverse primer (Table 4.5) to allow ligation to the pPHU234 vector (Figure 4.16).

AGGCCAGAAAATTCAACGCCCGCTCCCTGGCACAACGATTGCAGATCCTCCTTT**G**ACA
TCTAATAGAGGAGGATCGTATG

Figure 4.14. P*aioB* fragment amplified upstream the TSS of *aioB* to test σ^{54} promoter activity in NT-26. The amplified 51 bp fragment is highlighted in yellow; the σ^{54} promoter consensus sequence is underlined; the TSS is in bold and capitalized; and *aioB* start codon ATG is highlighted in green.

CTCGGTCTTCGACCCTGAAAAGGGCGGCCAGCAGGTTTGGGGTCAGGCCACGCAGAACC
TGCCGCAATACGTGCTCCGCGTCGCCGACAATGGCGACATCTTTGCCGAA**G**GCGTCGA
CGAGCTGATCTACGGCCGTCTGTCCAACGTTCTATAAGGGAAGAAGATCATG

Figure 4.15. P*aioA* fragment amplified upstream the TSS of *aioA* to test the putative promoter activity in NT-26. The amplified 114 bp fragment is highlighted in yellow; the TSS is in bold and capitalized; and *aioB* stop codon TAA is highlighted in green; and the *aioA* start codon ATG is highlighted in pink.

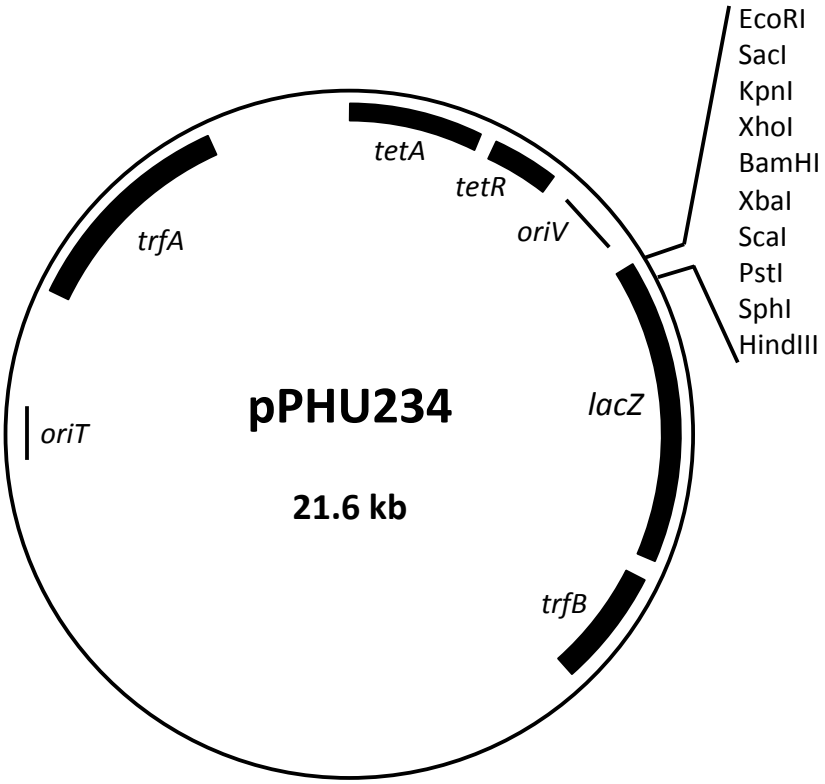


Figure 4.16. Map of the plasmid pPHU234 showing the multiple cloning site upstream the promoterless *lacZ.*

The PCRs were performed using the specific primers, Phusion Polymerase and gDNA from NT-26 in a final volume of 50 µl, of which 5 µl was applied to an agarose gel and observed under UV-light to check if the right sized amplicon was obtained (Figure 4.17 for *aioX* PCR and Figure 4.18 for *aioB* and *aioA* PCR).

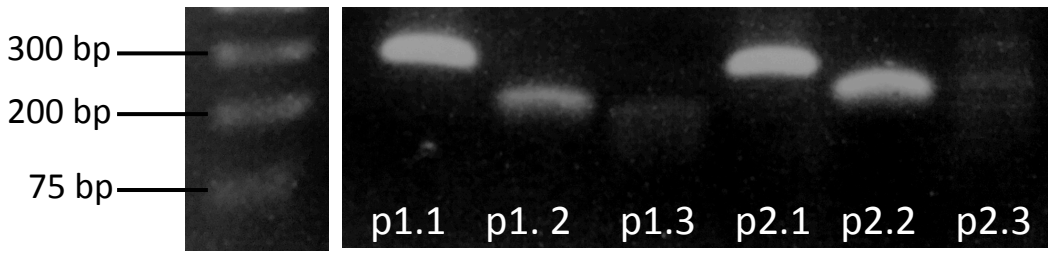


Figure 4.17. Agarose gel to confirm amplification of promoter fragments p1.1, p1.2, p1.3 p2.1, p2.2 and p2.3. The size of the molecular markers are shown on the left.

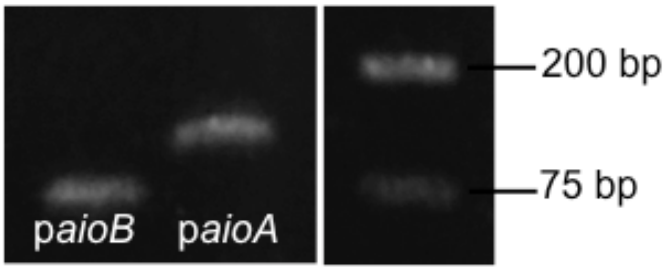


Figure 4.18. Agarose gel confirming the amplification of *paioB* and *paioA* promoter fragments. The sizes of the molecular markers are shown on the left.

The bands corresponding to the aioX TSS 1 were amplified with the correct size, although the p1.3 is barely visible. The aioX TSS 2 samples were successfully amplified except p2.3, which didn't provide a right sized band. This sample was returned to the thermocycler for 5 more PCR cycles after which it was possible to observer a right sized band in the agarose gel (not shown). The bands corresponding to PaioB and PaioA were also amplified with the correct size and the remaining PCR products were PCR purified. All eight samples and the vector, pPHU234, were digested for 3 hours 37 °C with *Bam*HI and *Pst*I, run in an agarose gel, purified from the gel and run again to confirm the recovery of the right sized bands and linearized pPHU234 (Figure 4.19).

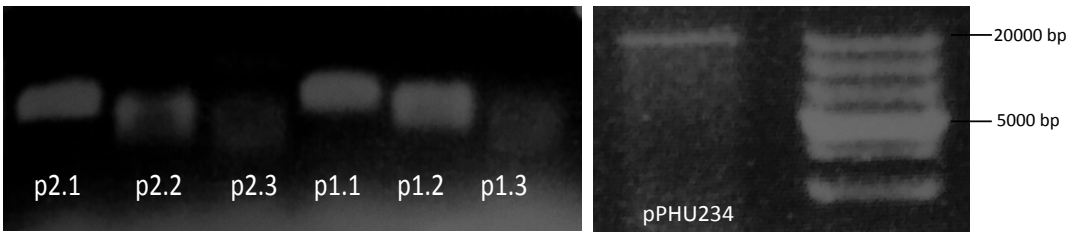


Figure 4.19. Agarose gel showing the p21, p22, p23, p11, p22, p23 and pPHU234 recovered after the digestion and gel extraction.

The linearized vector pPHU234 was ligated to each of the promoter fragments to form: pPHU234::Paiox1.1, pPHU234::Paiox1.2 and pPHU234:1.3 , pPHU234::Paiox2.1, pPHU234::Paiox2.2, pPHU234::Paiox2.3, pPHU234::PaioB and pPHU234::PaioA.

The product of each one of the eight ligations was transformed by electroporation into E. coli DH5α competent cells. Each transformation was spread on LB plates containing 10 μg/μl of Tc for clone selection. Individual colonies were selected and the presence of insert was confirmed by colony PCR (not shown). No clones were obtained for pPHU234::Paiox1.3 and pPHU234::Paiox2.3 which is probably due to the small concentration of insert used in the reaction. The bands corresponding to the p1.3 and p2.3 PCR products and recovery after digestion in the agarose gel were very light (Figures 4.13 and 4.14) and, due to their small size (smaller than 100 bp), it was known that the ligation into a big vector (bigger than 20000 bp) could be challenging.

The clones with the right insertion (pPHU234::PaioX.1.1, pPHU234::PaioX1.2, pPHU234::PaioX2.1, pPHU234::PaioX2.2, pPHU234::PaioB and pPHU234::PaioA) were grown in liquid LB medium containing 10 μg/μl Tc overnight. The vectors carrying the inserts were isolated and their concentration quantified using a NanoDrop spectrophotometer.

As the aim was to test the promoter activity in NT-26 grown under different conditions, the putative promoter fragments needed to be expressed in NT-26. The pPHU234::PaioX1.1, pPHU234::PaioX1.2, pPHU234::PaioX2.1, pPHU234::PaioX2.2, pPHU234::PaioB and pPHU234::PaioA were transformed into *E. coli* str S17-1 λpir using the quick $CaCl_2$ method. The *E. coli* str S17-1 λpir is a donor cell and therefore, able to transfer the vector to NT-26 via conjugation. The transformation of

pPHU234::PaioX.1.1, pPHU234::PaioX1.2, pPHU234::PaioX2.1, pPHU234::PaioX2.2, pPHU234::PaioB and pPHU234::PaioA vectors into *E. coli* str S17-1 λpir were successful as colonies were obtained on a LB 10 μg/μl Tc plate grown overnight. To have a negative control for the promoter activity test, the pPHU234 without any insert was also successfully transformed into *E. coli* str S17-1 λpir using the quick $CaCl_2$ method.

E. coli str S17-1 λpir (donor) carrying either pPHU234, pPHU234::PaioX1.1, pPHU234::PaioX1.2, pPHU234::PaioX2.1, pPHU234::PaioX2.2, pPHU234::PaioB, pPHU234::PaioA and NT-26 (recipient) were mixed and filtered for the conjugation to occur. After the conjugation and isolation of NT-26 single colonies carrying pPHU234, pPHU234::PaioX1.1, pPHU234::PaioX1.2, pPHU234::PaioX2.1, pPHU234::PaioX2.2, pPHU234::PaioB, pPHU234::PaioA the promoter activity was tested.

4.3.2.2 β-galactosidase assay in NT-26 to measure promoter activity

To determine whether the promoter fragments contain functional promoters which are active in NT-26, the 7 NT-26 transconjugants carrying either pPHU234, pPHU234::PaioX1.1, pPHU234::PaioX1.2, pPHU234::PaioX2.1, pPHU234::PaioX2.2, pPHU234::PaioB, pPHU234::PaioA, were grown heterotrophically with YE with and without As^{III} or As^{V}. Samples were taken at early-log (OD_{600} 0.030 – 0.058), mid-log (OD_{600} 0.07 – 0.098), late-log (OD_{600} 0.115 – 0.140) and stationary (OD > 0.200) phases to measure the β-galactosidase activity. The intention was to observe and compare the activity under all the different conditions. If the fragments PaioX1.1, PaioX1.2, PaioX2.1, PaioX2.2, PaioB or PaioA contained a promoter active in NT-26 on the conditions tested, it would be possible to see as the presence of β-galactosidase

activity indicates transcription of *lacZ* due to activation of each of the promoters tested (Li et al. 2012; Miller 1972).

The NT-26 transconjugant with pPHU234 without insert was used as a negative control. pPHU234 has no promoter upstream of the *lacZ* and without transcription of *lacZ* there is no β-galactosidase activity. To confirm that this would be a suitable negative control, the activity of NT-26 without pPHU234 was also measured and compared to the NT-26 transconjugant. The reaction was only done once and both had a small and similar β-galactosidase activity. This proves that the negative control used is a suitable one and that pPHU234 without any insert does not induce β-galactosidase and the small activity means that the substrate is beginning to auto-hydrolyse.

NT-26 transconjugants were grown in 10 ml MSM medium containing 7.5 µg/µl Tc and 0.04% YE with addition of As^{III} and As^{V}, when stated. Samples were taken when the cells reached early-log, mid-log, late-log and stationary phases of growth to observe the activity of the promoters over the course of growth. All the samples were mixed with permeabilization solution, followed by substrate solution that contains the substrate for the β-galactosidase. In presence of an active promoter, the solution develops a yellow colour, as exemplified in Figure 4.20, below.

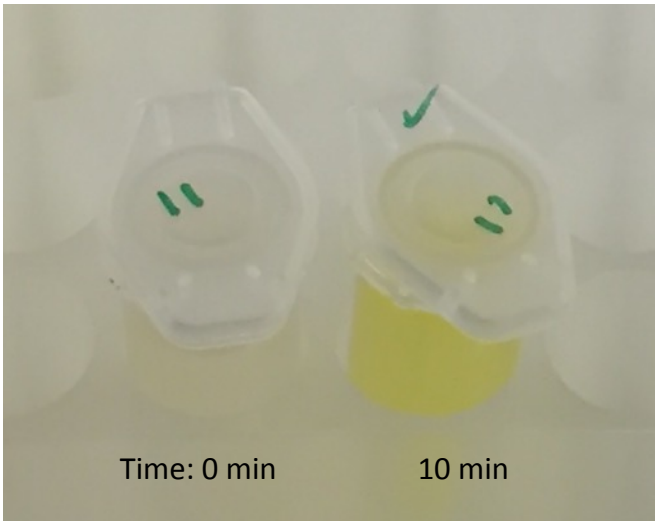


Figure 4.20. Yellow colour development indicating β-galactosidase activity in NT-26 pPHU234:P*aioX*1.1. The samples shown correspond to NT-26 pPHU234:P*aioX*1.1 grown in the presence of As^{III} upon addition of the substrate solution and after a 10 minute incubation at 28 °C.

The absorbance of the yellow colour was measured for each sample and the β-galactosidase activity was calculated taking in consideration the initial amount of cells (OD_{600}), the time taken for the yellow colour to develop and the intensity of the colour (OD_{420}) (Miller 1972). The β-galactosidase units were plotted against growth phase to observe if there is a relationship between the growth phase and the activity of the putative promoter.

The β-galactosidase activity for the putative promoters PaioX1.1 and PaioX1.2 in NT-26 grown in presence or absence of As^{III} or As^{V} was measured and the results presented are the average of two independent experiments (Figure 4.21).

It is possible to conclude from figure 4.21 that both PaioX1.1 and PaioX1.2 contain a promoter region active in NT-26 based on the activity of β-galactosidase. The largest fragment (PaioX1.1, 235 nt), had a higher β-galactosidase activity than the smaller one (i.e., PaioX1.2, 127 nt). This result suggests that the smaller fragment (i.e.,

PaioX1.2) is missing an important element for full activity. The presence of either As^{III} or As^{V} doesn't seem to influence the activity of the promoter, indicating that the PaioX1, and consequently the *aioX, aioS* and *aioR* operon, does not respond to presence of As^{III} or As^{V} as seen on Chapter 3. However, when analysing the Figure 4.21, the activity of PaioX1.1 in mid-log and presence of As^{III} seems to be slightly higher than As^{V} and YE alone but more replicates need to be tested to observe if this difference is significantly different. When observing the activity of the promoter during the course of growth, it is noticeable, especially for PaioX1.1 that the promoter is always active, without any abrupt change that could indicate a direct effect of growth phase. From what can be seen in figure 4.21, the promoter seems to be constitutively active and the increase of β-galactosidase is due to the accumulation of β-galactosidase over the course of time.

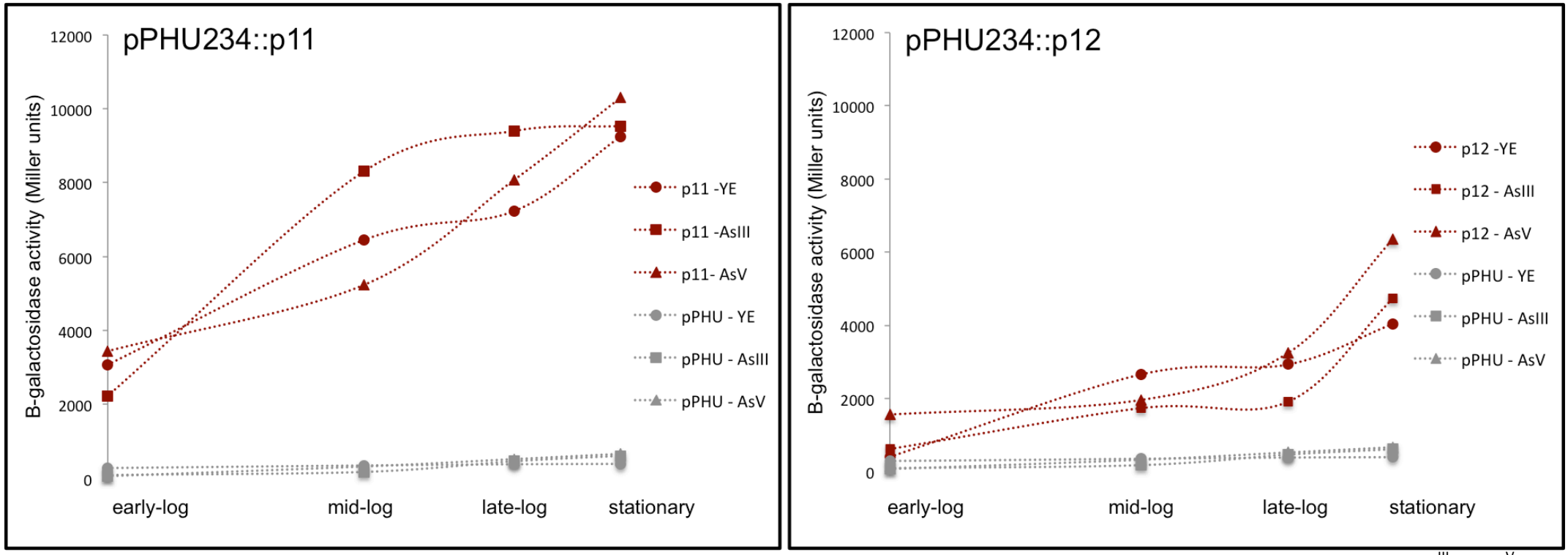


Figure 4.21. β-galactosidase activity for the putative promoters PaioX1.1 and Paiox1.2 in NT-26 grown heterotrophically with YE with and without As^{III} or As^{V} over the course of growth. The negative control samples (i.e., the promoterless pPU234 plasmid) are shown in light grey. The NT-26 transconjugants with pPHU234::PaioX1.1 and pPHU234::PaioX1.2 are shown in red. The results plotted here correspond to the average of two independent experiments.

The activity of the putative promoters PaioX2.1 and PaioX2.2 in NT-26 grown with YE in addition of either As^{III} or As^{V} were measured (Figure 4.22). Figure 4.22 shows that NT-26 carrying the pPHU234::PaioX2.1 had higher β-galactosidase activity than pPHU234::PaioX2.2. This suggests that the region required for full promoter activity is in the region between 150 nt and 219 nt upstream *aioX* TSS 2. It appears that although the promoter is active from early-log to late-log phases, an increased activity was detected in stationary phase. This result is in accordance with the hypothesis, first presented in Chapter 3, that more than one promoter controls expression of *aioX, aioS* and *aioR* and that one of them has increased activity in stationary phase. From figure 4.22 it is also possible to observe that the presence of As^{III} or As^{V} does not induce the activity of either PaioX2.1 or PaioX2.2.

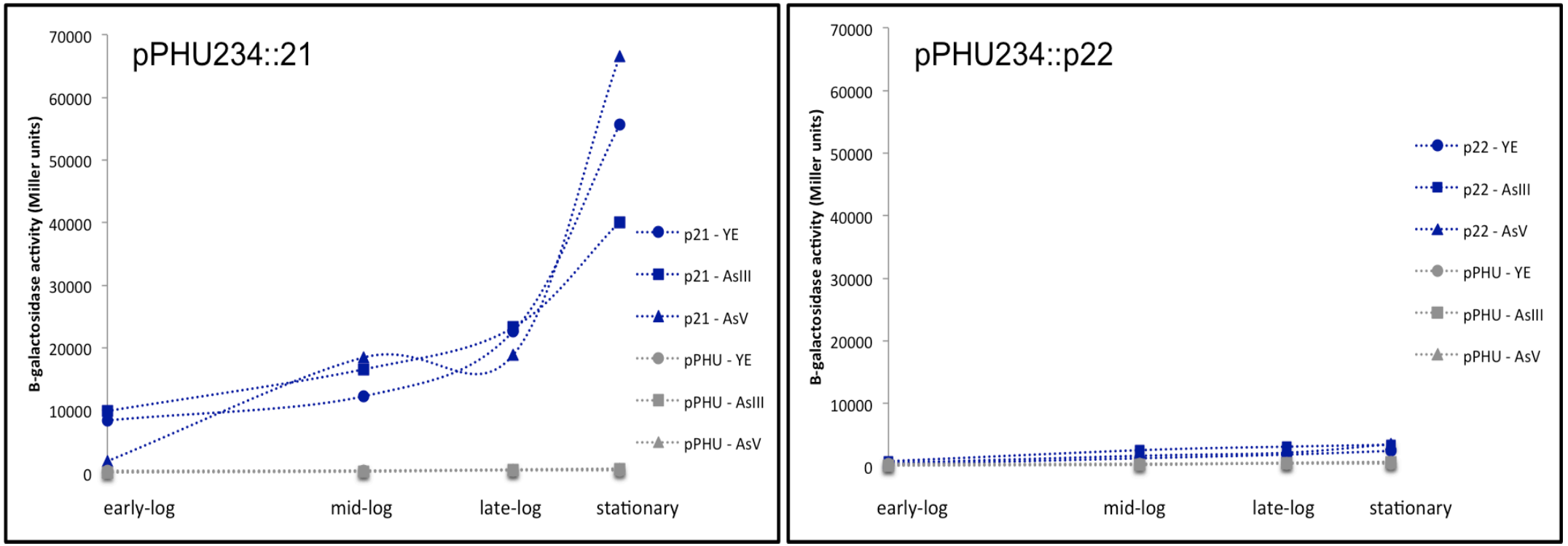


Figure 4.22. β-galactosidase activity for the putative promoters PaioX2.1 and Paiox2.2 in NT-26 grown heterotrophically with YE with and without As^{III} or As^{V} over the course of growth. The negative control samples (i.e., the promoterless pPU234 plasmid) are shown in light grey. The NT-26 transconjugants with pPHU234::PaioX2.1 and pPHU234::PaioX2.2 are shown in blue. The results plotted here correspond to the average of two independent experiments.

The activity of the putative promoters PaioB and PaioA in NT-26 grown in presence or absence of As^{III} or As^{V} was measured and are shown in Figure 4.23. NT-26 transconjugants carrying the pPHU234::PaioB showed no activity when compared to the negative control, which is probably due to the fact that the promoter region cloned only contains the region where σ^{54} binds and not the region required for AioR binding which unsurprisingly appears to be essential for promoter activity.

The NT-26 transconjugants carrying the pPHU234::PaioA had activity when NT-26 was grown with As^{III} during the whole course of growth; activity in NT-26 grown with As^{V} increasing from mid-log to stationary phase; and in NT-26 grown with YE alone the promoter showed increased activity from mid-log. In early-log phase the activity detected was comparable with the negative control, for the 3 conditions tested. These results confirm that the fragment found upstream *aioA,* and considered a putative promoter, is active and functional in NT-26.

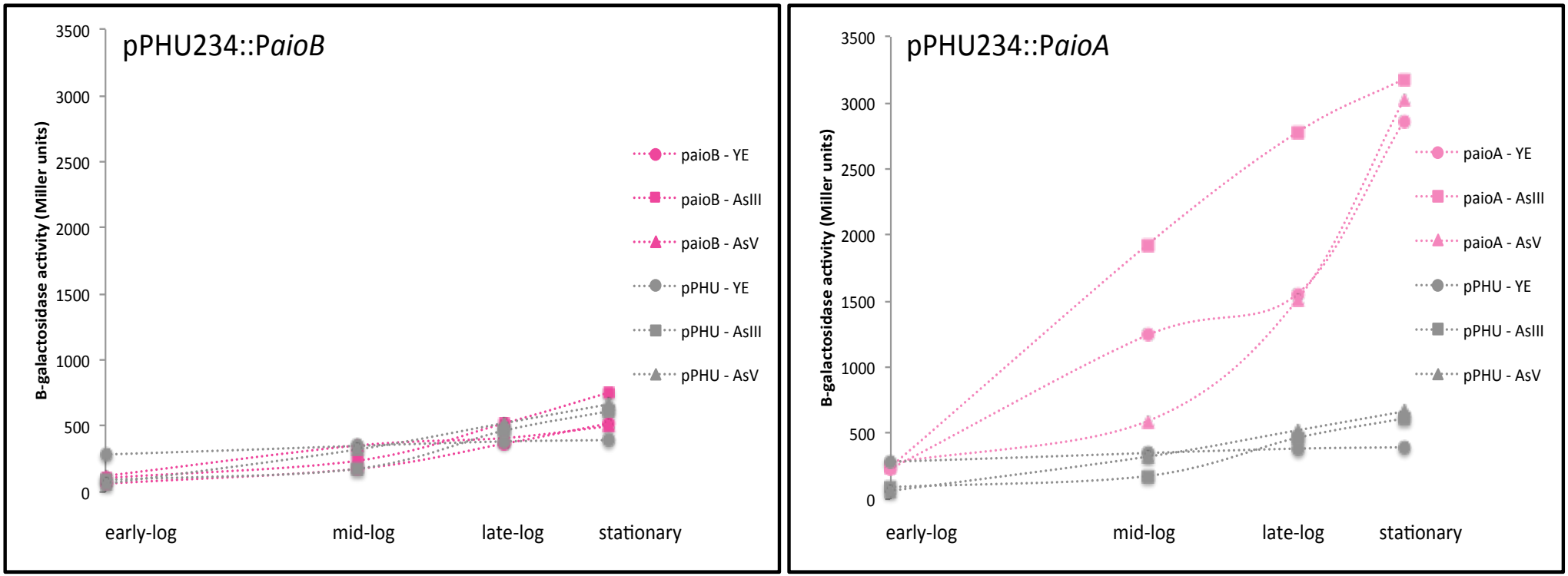


Figure 4.23. β-galactosidase activity for the putative promoters PaioB and PaioA in NT-26 grown heterotrophically with YE with and without As^{III} or As^{V} over the course of growth. The negative control samples (i.e., the promoterless pPU234 plasmid) are shown in light grey. The NT-26 transconjugants with pPHU234::PaioB are shown in dark pink and pPHU234::PaioA are shown light pink. The results plotted here correspond to the average of two independent experiments.

4.3.2.3 Analysis of promoter regions PaioX2.2, PaioX1.2 and PaioA to find conserved promoter sequences

Using promoter prediction softwares like BPROM and Virtual Footprint, both containing a database of promoters sequences that can be used to find a similar pattern in the provided DNA sequence (containing the putative promoter), no promoters upstream *aioX* or *aioA* could be identified. The search for the *aioX* and *aioA* putative promoters were then done by eye, based on the published consensus sequences for the most common sigma factor, the σ^{70}. A consensus sequence for the σ^{70} promoter [TTGAC] and [TATA] was identified upstream of the TSS 1 (Figure 4.24).

GAGGCAGGCATAGGAGGTTTGAACCGATGCGTCCGCAAAATCAGGACATTCCTCGAT
GGAAATTGACTTTATTTGCGCTGATCCTGTTGATTGGTTTCGCCCAGATCGTCTTCA
TCCTGTTCATGCAATCGCG**TTGAC**GCGGCCACATCGTTTGCCTCAGC**TATA**GAGGAC
GTAAATTCAGGCGCTGTTGGACATTGAGCAACTCCAGTTTACAGTCTGAGTGACGGC
TCTGGAACTGGGCCTTTCGGGGAGGCGATCGGTGTTGGATG

Figure 4.24. σ^{70} conserved region upstream of *aioX* TSS 1. The σ^{70} conserved region is shown in bold. The light red highlights the portion corresponding to PaioX1.2 and the dark red highlights the PaioX1.1. The region downstream the TSS 1 is highlighted in light grey and the aioX first codon is in dark grey.

The search was done looking for the conserved regions described for *E. coli* [TTGACA] [TATAAT] (Harley & Reynolds 1987) and the region described for *R. etli* [CTTGAC] [TATNNT] (Ramírez-Romero et al. 2006). It is also possible that the σ^{70} binding site is [TgGACA-16-TAcAgT] as underlined in figures 4.24, that is positioned at -10 and -35 relative to the transcriptional start site.

The region containing PaioX2 was analysed by eye and a conserved region for a RpoE2 promoter in *R. meliloti* (GGAAC-a14gcgTTt7g) was identified (Sauviac et al. 2007) (Figure 4.25).

GCTTATACCTCGCCACTGCCCAGGGTGAATACACAGCAGCGAAATCCGGCAAAAGGGA
AGAGGTGTGAAGGCCTCCAATAAGGCAGGAATCACGCAGCTCTTAGCTACGGATGGAG
CAGGACGATCGGCCCGGAGTGTGCAGTG**GGAAC**CGCGGGCGCAGCTAAAGC**TT**ACTTG
CAG**G**AATTGCAAAAAACGGTAAACCATTGGTCCGAAGGCAGGCATAGGAGGTTTGAAC
CGATGCGTCCGCAAAATCAGGACATTCCTCGATGGAAATTGACTTTATTTGCGCTGAT
CCTGTTGATTGGTTTCGCCCAGATCGTCTTCATCCTGTTCATGCAATCGCGTTGACGC
GGCCACATCGTTTGCCTCAGCTATAGAGGACGTAAATTCAGGCGCTGTTGGACATTGA
GCAACTCCAGTTTACAGTCTGAGTGACGGCTCTGGAACTGGGCCTTTCGGGGAGGCGA
TCGGTGTTGGATG

Figure 4.25. RpoE2 conserved region upstream *aioX* TSS 2. The RpoE2 conserved regions are shown in bold. The light blue highlights the portion corresponding to PaioX2.2 and the dark blue highlights the PaioX2.1. The region downstream the TSS 2 is highlighted in light grey and the *aioX* first codon is in dark grey.

It was not possible to find any conserved promoter region upstream *aioA* using the computational tools BPROM and Virtual Footprint and neither comparing the putative *aioA* promoter region with the conserved region described for σ^{54}, σ^{70} or RpoE. More work need to be done to identify the putative promoter that could possibly be regulating the expression of *aioA*. To confirm that the *aioA* putative promoter is being expressed in full, along with the *lacZ*, and the β-galactosidase activity measured corresponds to the activity of the fragment cloned, RT-PCR was performed using RNA from NT-26 PaioA grown with YE and As^{III} until stationary phase. For this reaction, primers that anneal specifically in the 5' portion of p*aioA* and 3' portion of the *lacZ* gene were used. A band with the expected size was obtained, confirming that *paioA* is being expressed along with *lacZ* (Figure not shown). This fragment needs to be sequenced to confirm that *paioA* was cloned in full and no mutation was added in the process generating an artefact that could be masking the real activity (or lack of activity) of this putative promoter.

4.4 Discussion

To understand the transcriptional regulation of the *aioXSR* operon and of *aioB* and *aioA* in NT-26 it was essential to identify and characterize the promoter regions, which are located upstream of TSSs. The promoter region is where transcription factors bind to regulate the transcription of a gene or operon. Promoter regions are always located upstream of the TSS.

As shown in chapter 3, the genes *aioX, aioS* and *aioR* form an operon and are constitutively expressed, transcribed in the presence or absence of As^{III} or As^{V} and during late-log and stationary phases. However, they are up-regulated in stationary phase suggesting that their expression is under the control of more than one promoter. Also in chapter 3, the hypothesis that *aioA* can be controlled by more than one promoter was raised based on its up-regulation in stationary phase, in absence of either As^{III} or As^{V}.

To test for these hypotheses the promoter – or promoters – controlling the expression of *aioXSR* operon and *aioB* needed to be identified and its activity validated. This was done by first identifying the TSS, and consequently the UTR, upstream of *aioX, aioB* and *aioA* using the 5'RACE method. To validate the presence of promoters upstream of the TSS the putative promoter regions were cloned upstream of *lacZ* in a promoterless vector and the activity tested under different conditions and different growth stages.

4.4.1 TSS identification upstream of *aioX*

To identify and study the promoter region upstream of the *aioXSR* operon it was essential to determine the position where transcription starts. RNA was isolated from NT-26 grown in the presence and absence of As^{III} until late-log and stationary phases.

Although an aioX GSP specific primer was used, two of the four fragments obtained in the 5' RACE PCR were due to mispriming. That could be explained by the fact that an universal primer UPM was used along with the *aioX* GSP primer, reducing the specificity of the reaction and allowing amplification of non-specific regions. Based on the GSP position inside *aioX*, it was expected that a fragment containing the TSS of *aioX* to be bigger than 914 bp, in case of specific amplification. All the bands analysed were bigger than 914 bp although only two were specific to *aioX.* The two fragments that were found to be specific to NT-26 *aioX,* after the sequencing analyses, contained two different UTR sequences, and therefore two TSS, named TSS 1 for the proximal and TSS 2 for the distal one.

Five different clones for both TSS 1 and TSS 2 were analysed and they presented the exact same sequence, confirming that the UTRs were sequenced in full. This result is in accordance with the observations from Chapter 3 where it was considered, for the first time, the possibility of an *aioXSR* operon being regulated by more than one promoter as this operon is constitutively expressed in NT-26 but also up-regulated in stationary phase.

4.4.2 *aioXSR* Promoter identification and activity

The identification of the TSS *per se* does not confirm the presence of an active promoter, nor is it possible to define which promoter is responding to constitutive

expression or to growth phase. To determine the presence of an active promoter the putative promoter regions were cloned in pPHU234 and expressed in NT-26.

The putative promoter regions PaioX1.1 and PaioX1.2 induced β-galactosidase activity in NT-26 indicating they contain an active promoter. The activity of PaioX1.2 was lower than PaioX1.1, the activities of PaioX1.1 or PaioX1.2 were not affected by the presence of As^{III} or As^{V} and the activity of PaioX1.1 was constant over the course of growth (Figure 4.21). This result together with the qPCR from Chapter 3, shows that *aioXSR* is expressed in a basal level and a housekeeping promoter is controlling this operon. It is not uncommon for two-component systems to contain a constitutive promoter in addition to an auto-regulated promoter (Bijlsma & Groisman 2003; Soncini et al. 1995; Mitrophanov et al. 2010). The σ^{70} is known as the housekeeping promoter and controls the expression of the genes essential for general transcription in exponentially growing cells (Paget & Helmann 2003) and as shown in Figure 4.21, a consensus sequence for a σ^{70} promoter was found upstream *aioX*.

It is not possible to conclude at the moment why PaioX1.1 has higher activity then PaioX1.2 when the σ^{70} is found in the region covered by the PaioX1.2 fragment. More work needs to be done to confirm which sigma factor binds upstream of the TSS 1 and the binding site. The use of in silico tools to align the regulatory regions of housekeeping genes in NT-26 could help to determine the conserved region specific to NT-26. Site-directed mutagenesis could be used to alter the [TTGAC] and [TATA] motifs to determine what residues are essential for promoter activity. Another option would be to use PaioX1.2 in a mobility shift assay with different sigma factors from the σ^{70} family and RNA polymerase to elucidate which one is binding the promoter

region upstream of TSS 1 (Mantovani 1994; Ramírez-Romero et al. 2006; Aramaki & Fujita 1999). This could be combined with DNAse footprinting to confirm the promoter binding region.

The results presented in Chapter 3 and the TSS determinations suggest the presence of a 2nd promoter upstream of *aioX*. The β-galactosidase activity of the fragment PaioX2.1 was higher than PaioX2.2, and PaioX2.1 seems to contain a strong promoter due to its higher activity, consequence of high transcription of *lacZ*. The activity was not influenced by the presence of As^{III} or As^{V} in the growth medium. It was possible to observe that although PaioX2.1 activity was detected from early log-phase, activity peaked in stationary phase (Figure 4.22). This indicates that this region may contain a promoter that is responsive to growth phase. This result reinforces the hypothesis drawn in Chapter 3 that the *aioXSR* operon is controlled by two promoters; one for the basal expression and another for growth rate regulation. In *R. meliloti* the sigma factor RpoE2 binds to a conserved region [GGAACN-17n-cgTT-9n-G] similar to the one found in the PaioX2.2 region of NT-26 [GGAAC-16n-gcTT-9n-G] (Figure 4.24) (Sauviac et al. 2007). RpoE2 is an extracytoplasmic sigma factor (ESF) (Heimann 2002) and it responds to osmotic stress and stationary phase following carbon or nitrogen starvation in *R. meliloti* (Costanzo & Ades 2006). NT-26 has a putative *rpoE* gene in the chromosome in a identical synteny to *R. meliloti* with a *phyR*-like gene adjacent to *rpoE* (Bastiat et al. 2010). PhyR is often involved in the regulation of RpoE2 (Bastiat et al. 2010; Sauviac et al. 2007). In NT-26 the RpoE2 binding conserved region was also found upstream of the catalase gene *katA* that has been reported as being regulated by RpoE in *R. meliloti* (Sauviac et al. 2007) and was up-regulated in NT-26 in response to As^{III} (Andres et al. 2013). The quorum sensing and flagella regulator gene

qseB (Sperandio et al. 2002) also contains the RpoE2 conserved binding sequence in NT-26 and was shown to be up-regulated with As^{III} (Andres et al. 2013). The QseB participates in the flagellum and regulation motility network *via* quorum sensing mechanisms (Clarke & Sperandio 2005; Sperandio et al. 2002) and NT-26 is a motile bacteria with two subterminal flagella (Santini et al. 2000). The motility in NT-26 has been reported as being a response to As^{III} and the flagella genes *fliF, flaA* and *fla* are up-regulated in NT-26 in presence of As^{III}, as well as *qseB* (Andres et al. 2013). However, the pPHU234::P*aioX*2.1 is not more active in the presence of As^{III} when compared to YE alone (Figure 4.22), indicating that although the RpoE2 seems to be involved on the regulation of *aioXSR* operon as well on the regulation of *qseB* and *katA* expression*,* the expression of these genes respond to different stimuli, with *qseB* and *katA* being up-regulated in response to As^{III} while the *aioXSR* operon is up-regulated in stationary phase. It might be the case that RpoE2 is a positive regulator for one of the genes while being a negative regulator for the others.

The higher activity observed for *P*aioX2.1 in stationary phase and the presence of the RpoE2 conserved region, contributes to the hypothesis that the promoter region 2 upstream of the *aioXSR genes* contains the regions required for growth phase response and, possibly, quorum sensing regulation as first suggested in Chapter 3. However, it is not clear why the *P*aioX2.1 is more active than *P*aioX2.2 when the *rpoE2* binding region is located within *P*aioX2.2 (Figure 4.25). The conserved region upstream those genes is shown in Figure 4.26, below:

aioX: TG**GGAAC**CGCGGGCGCA----------GCTAAAGC**TT**ACTT--GCAG**G**AATT
qseB: CG**GGAAC**ATTGCATTTCT----------CCGCGGCATACGAGG**TT**GCG-GTGAT**G**CCGA
katA: TT**GGAAC**CGGAGTGAGATGGTCAAGCCGCAACCTAGAACGG**TT**CCAGATTTCACCCA
rpoE2: GT**GGAAC**TTTTTTTCCGTT----------GCCGCA**TT**GTTTCCTCAC**G**ATAT

Figure 4.26. RpoE2 conserved region upstream *aioX, qseB, katA* and *rpoE2* in NT-26. The conserved regions are in bold and underlined. The alignment was done manually based on the sequence upstream each gene. Accession numbers: *aioX* (NT26v4_p10026), *rpoE2* (NT26v4_2861), *katA* (NT26v4_0389), *qseB* (NT26v4_0655).

The results presented in this chapter add another layer of complexity to the regulation of the *aioXSR* genes in NT-26; they show the presence of two different active promoter regions upstream of the *aioXSR* operon that seem to be controlled by RpoE2 and σ^{70}. However, further studies need to be done to confirm the involvement of these proteins in regulating the *aioXSR* operon. The only published work on the promoter region of the *aioXSR* operon was made by Moinier *et al*. when the *aioX* TSS in *T. arsenitoxydans* 3As was identified. The presence of a consensus sequence for σ^{70} was identified upstream of the TSS and it was also found that AioF participates in the regulation of *aioB* in *T. arsenitoxydans* 3As (Moinier et al. 2014). The presence of a σ^{70} upstream of *T. arsenitoxydans* 3As *aioX* is in accordance with the results presented in this chapter for NT-26. However, the genes *aioX* and *aioB* are divergent and NT-26 has no homologue of AioF indicating that the regulation of *aioXSR* in NT-26 varies from that of *T. arsenitoxydans* 3As apart from the involvement of σ^{70}. The σ^{70} conserved region found in *T. arsenitoxydans* 3As (cTGACg-(17 bp)-TAcgcT) is also different from the one found in NT-26 (TTGAC-(23)-TATA).

In silico analysis needs to be performed to determine the RpoE2 and σ^{70} conserved regions specific to NT-26 and to locate these upstream of other genes in the NT-26

genome. The σ^{70} controlling the basal expression of *aioXSR* still needs to be identified.

4.4.3 TSS identification upstream *aioB* and *aioA* and σ^{54} promoter activity in NT-26

Growth of NT-26 in the presence of As^{III} results in the up-regulation of *aioB* and *aioA* and the expression of these genes are regulated by the two component signal transduction system proteins AioS and AioR (Sardiwal et al. 2010a). AioR is thought to bind to a conserved region upstream of *aioB* (Andres et al. 2013) interacting with the σ^{54} allowing for the initiation of transcription. The positions of the σ^{54} consensus motif upstream of *aioB* was first described in NT-26 (Santini & Vanden Hoven 2004). This σ^{54} consensus sequence is also found in the As^{III}-oxidisers that possess AioS and AioR, identified so far (Koechler et al. 2010), such as *A. tumefaciens* 5A, *Thiomonas* sp., *Achromobacter* sp., *Rhodoferax ferrireducens, Ochrobactrum tritici* and *H. arsenicoxydans* ULPAS-1 (Koechler et al. 2010). In this chapter, the TSS upstream of *aioB* was identified using 5'RACE and the σ^{54} consensus sequence was found to be exactly at the expected -12 and -24 regions as predicted (Figures 4.11 and 4.14). No other TSS was identified suggesting that *aioB* is under the control of the σ^{54} promoter alone. The up-regulation of *aioB* and *aioA* in stationary phase, as observed in Chapter 3, can either be explained by the activation of the AioS and AioR by a stimulus other than As^{III}, by the presence of another σ^{54} response regulator responsive to growth phase and capable of binding upstream *aioB*. These hypotheses need to be tested further.

To test the activity of the σ^{54} promoter region upstream *aioB,* NT-26 containing pPHU234::P*aioB* was grown in presence and absence of As^{III} or As^{V}. The β-

galactosidase activity was measured over the course of growth and the result compared to the activity obtained for the negative control. No difference was observed between the expression of pPHU234::P*aioB* and the negative control (i.e., pPHU234) in NT-26 under all the conditions tested (Figure 4.23). As the P*aioB* fragment only contains the σ^{54} consensus sequence and not the putative AioR-binding sequence the β-galactosidase gene was not transcribed and therefore activity not detected. Next time a fragment containing the binding sites for both AioR and σ^{54} need to be cloned downstream of *lacZ* in pPHU234 and expression in response to As^{III}, As^{V} and growth phase determined.

When the 5'RACE was performed to determine the TSS upstream of *aioB*, one unexpected TSS upstream *aioA* was identified. The presence of a TSS upstream of *aioA* suggests it is not always co-transcribed with *aioB.* To test whether the putative *aioA* promoter is functional, the region upstream of the TSS was cloned into pPHU234 and its activity determined by measuring the β-galactosidase activity. It was observed that P*aioA* is functional when compared to the negative control (Figure 4.23). P*aioA* is more expressed in presence of As^{III} and the promoter seems to be active from late-log to stationary phase. When grown with YE alone or in presence of As^{V} the activity starts to increase after late-log-phase and is possible to notice a peak in activity when NT-26 reaches stationary phase. The nature and exact location of the *aioA* promoter needs to be determined. The two computational tools (BPROM and Virtual Footprint) used to look into this promoter region do not have a promoter database specific to *Rhizobium,* making the search for promoter conserved regions harder. Although the details of *aioA* promoter were not determined this is the first

time that *aioA* has been shown to be under control of its own promoter, which is very surprising.

To confirm the hypothesis presented in this chapter, the promoter activity in NT-26 needs to be measured at least one more time to have at least three replicates to confirm the results and give – or not – significance to the difference in activity observed between As^{III}, As^{V} and YE samples.

4.5 Chapter summary

- Two TSSs were identified upstream of the *aioXSR* operon using 5'RACE and designated TSS 1 and TSS 2.
- The existence of two active promoters – *P*aioX1 and *P*aioX2 - regulating the *aioXSR* operon were confirmed.
- *P*aioX1 and *P*aioX2 don't respond to As^{III} or As^{V} but more replicates need to be tested to confirm this result.
- *P*aioX1, which may have a σ70 promoter region, is involved in the constitutive expression of the *aioXSR* genes.
- *P*aioX2 may be a RpoE2-like promoter and is more active in stationary phase.
- The RpoE2 promoter is involved in growth phase regulation and consensus sequences also found usptream of the *qseB* quorum sensing gene, the *kat* catalase gene and the *rpoE2* gene in NT-26.
- The TSS upstream of *aioB* confirmed the σ^{54} consensus sequence at the -12 and -24 positions.
- A TSS was found upstream of *aioA* suggesting the presence of an active promoter yet to be confirmed and identified.

- The putative *aioA* promoter responds to presence of As^{III} from mid-log to stationary phase; it was also active when NT-26 was grown until stationary phase in presence of As^{V} and in absence of both As^{V} and As^{III}.

Chapter 5

Concluding remarks

5.1 Thesis Overview

Arsenic (As) is a toxic metalloid and a major contaminant in terrestrial and aquatic environments. The two soluble forms, arsenite (As^{III}) and arsenate (As^{V}) are toxic to most organisms. A range of phylogenetically distant bacteria are able to oxidize As^{III} to the less toxic form, arsenate As^{V} using the periplasmic arsenite oxidase (AioBA). The two-component signal transduction system AioS/AioR and the As^{III}-binding periplasmic protein AioX are required for As^{III} oxidation and are involved in the transcriptional regulation of the *aioBA* operon. Most As^{III} oxidisers can also reduce As^{V} to As^{III} *via* the As (Ars) resistance system.

The main focus of this work was to understand the regulation of genes involved in As^{III} oxidation and As resistance together with those involved in phosphate metabolism in the facultative chemolithoautotrophic As^{III} oxidiser NT-26 grown under different conditions. Gene expression was studied by quantitative PCR in cells grown heterotrophically with and without As^{III} or As^{V} in late-log and stationary phases. qPCR was optimised and suitable reference genes were chosen to analyse gene expression in NT-26 (Chapter 2). The expression of genes involved in phosphate transport, sensing and As resistance were compared in NT-26 samples obtained with and without As^{III} or As^{V} in late-log and stationary phases. The expression of the genes *aioX, aioS, aioR* (As^{III}-sensing and regulation) and *aioB, aioA* (As^{III} oxidation) and *cytC* (cytochrome *c*) were also analysed in NT-26 grown heterotrophically in the presence or absence of As^{III} or As^{V} at different growth stages (i.e., late-log and stationary phases) (Chapter 3). To understand gene expression in more detail the transcription start sites (TSS) upstream of *aioX*, *aioB* and *aioA* were identified using the 5'RACE technique. Promoter function was then confirmed by cloning the putative promoter

regions upstream of a promoterless β–galactosidase gene and enzyme activity measured in NT-26 grown heterotrophically with and without As^{III} or As^{V} at different growth stages. Two functional promoters were found to control the expression of the *aioXSR* operon, one for its constitutive expression and another for expression in stationary phase. Surprisingly a TSS was identified upstream of *aioA* and the presence of a functional promoter was validated (Chapter 4).

5.2 General conclusions

5.2.1 Regulation of genes involved in As resistance and phosphate metabolism in NT-26

Expression studies in NT-26 using qPCR and samples obtained from NT-26 grown heterotrophically in the presence or absence of As^{III} or As^{V} showed that the expression of the Pi transporter *pstS* and the phosphate regulatory gene *phoB* are correlated with the expression of genes encoding for the As resistance system, *arsR1a, arsC* and *arsB1.* The genes *pstS* and *phoB* are co-localized with the *ars* operon. Expression of all five genes was up-regulated in the presence of both As^{III} and As^{V}. It was subsequently observed that the up-regulation of these genes in the presence of As^{V} was probably due to the presence of As^{III} as a by-product of As^{V} reduction by ArsC. ArsR is known to de-repress the transcription of the *ars* operon in the presence of As^{III} (Jianhua Wu & Rosen 1993). In *A. tumefaciens* 5A ArsR was shown to bind to the intergenic region between *pstS* and *phoB* thereby regulating their expression (Kang, Heinemann, et al. 2012). The up-regulation of *pstS* and *phoB* in response to the presence of As^{III} has also been shown for *H. arsenicoxydans* ULPAS-1(Cleiss-Arnold et al. 2010) and in *A. tumefaciens* GW4 the *pstS1* gene was also up-regulated with As^{III} (Kang et al. 2014). The up-regulation of *phoB* suggests that PhoB

could act as a trans acting regulator in response to As^{III}, possibly being activated by AioR through phosphorelay as suggested by Kang and collaborators (Kang et al. 2012). The regulation of *ars* genes in NT-26 can't be directly regulated by PhoB as no Phobox was found upstream of the *ars* genes, indicating that the regulation is likely to happen indirectly. The regulation of *pstS* expression in response to As^{III} has been shown as being PhoB dependent in *A. tumefaciens* 5A reinforcing the idea that the As^{III} and Pi metabolism in bacteria can be inter-connected (Kang et al. 2012; Kang et al. 2014). The position of the genes *pstS* and *phoB* in the NT-26 plasmid, near *ars,* indicates that these genes may be associated with As resistance in NT-26 as has been observed for other As^{III}-oxidizers (Li et al. 2013). According to Li et al. the conserved clustering of *phoB, pstS, ars* and *aio* genes suggests relationship between their gene expression or metabolism (Li et al. 2013). In NT-26 the up-regulation of *arsR1a, arsC* and *arsB1* in stationary phase in the absence of As^{III} and As^{V} may be due to a reduction of Pi in the medium, and indirect regulation by PhoB in response to Pi stress in the cell. The expression of the genes *pstS2* and *pstS3* were not affected by the presence of As^{III} or As^{V} in the growth medium, indicating that they are probably dedicated to Pi transport and metabolism.

The evidence of co-regulation of *aioS, aioR, aioB, aioA, pstS, phoB* and *ars* shown in A. *tumefaciens* 5A (Kang, Heinemann, et al. 2012) is consistent with what was observed in this work for NT-26. However, no link between Pi concentration and *aioX, aioS, aioR, aioB, aioA, pstS, phoB* and *ars* expression can be confirmed for NT-26 before further analyses. Additional experiments need to be completed to observe the expression of *pstS, phoB, aio* and *ars* genes in the presence of different concentrations of Pi to elucidate if the regulation of these genes in NT-26 is also

responding to a decrease in Pi concentration, as suggested for *A. tumefaciens* 5A, *A. tumefaciens* GW4 and *H. arsenicoxydans* ULPAS-1. The minimal concentration of Pi that induces the PhoB in NT-26 needs to be defined to understand the interplay between Pi starvation and As response. It is also crucial to look further to find which genes PhoB can be regulating in NT-26 by searching where PhoB binds in NT-26, and to analyse the relationship between the *phoB* found upstream *ars* with the other *phoB* genes present in the chromosome and plasmid.

5.2.2 Regulation of the *aioXSR* genes and promoter identification

The genes *aioX, aioS* and *aioR* were shown to be co-transcribed and constitutively expressed in NT-26 but also up-regulated in stationary phase, suggesting that more than one promoter could be controlling this operon. *aioX, aioS* and *aioR* were not up-regulated by either As^{III} or As^{V}. In *A. tumefaciens* 5A the *aioX* gene was found to be up-regulated in presence of As^{III} (Liu et al. 2012), in this thesis it was shown that NT-26's *aioX* is constitutively expressed and not up-regulated by As^{III} and this result is in accordance to what has been shown for *T. arsenitoxydans* 3As (Slyemi et al. 2013), suggesting that the regulation of *aioX, aioS* and *aioR* operon varies between organisms.

The regulatory region upstream of the *aioX, aioS* and *aioR* operon in NT-26 was studied. Two transcription start sites were found and two different active promoters were identified. The promoter nearest to *aioX* contains a putative σ^{70}-binding site. Only one other explores the regulatory region upstream of the *aioX, aioS* and *aioR* operon in *T. arsenitoxydans* 3As and in this study one TSS was identified with a consensus sequence for a σ^{70} upstream *aioX* (Moinier et al. 2014).

The analysis of the promoter furthest from *aioX* in NT-26 was shown to contain a consensus sequence for RpoE2 binding. This protein is involved in growth phase regulation in *S. meliloti* (Sauviac et al. 2007). The *aioX, aioS* and *aioR* genes were shown to be up-regulated in stationary phase in NT-26 and therefore the presence of a promoter controlling the expression of these genes in response to growth phase was anticipated. The consensus sequence for RpoE2 binding was also found upstream of the genes *qseB*, *kat* and *rpoE2.* The *kat* gene was also reported to be up-regulated when NT-26 was grown in the presence of As^{III} (Andres et al. 2013) and is known to be regulated by RpoE2 in *S. meliloti* (Sauviac et al. 2007). The *qseB* gene encodes a putative quorum sensing and flagella regulatory protein in *E. coli* (Sperandio et al. 2002) and was also shown to be up-regulated when NT-26 was grown in the presence of As^{III} in NT-26 (Andres et al. 2013). NT-26 was also shown to be more motile when grown in the presence of As^{III}, reinforcing the link presented here between the regulation of *qseB* and the *aioXSR* operon regulation by RpoE2 (Andres et al. 2013). The possibility that RpoE2 binds to the *aioX* promoter region needs to be confirmed by mobility shift experiments and exact sequence can be confirmed by DNAse footprinting. Additionally the NT-26 genome needs to be scanned to find more putative RpoE2 binding sites. The involvement of RpoE2 can also be confirmed by deleting *rpoE2* and observing the effect of the deletion on the expression of *aio* genes in NT-26.

5.2.3 Regulation of genes involved in As^{III} oxidation in NT-26 and promoter identification

Both *aioB* and *aioA* were known to be up-regulated in the presence of As^{III} in NT-26 but unexpectedly they were also up-regulated in response to stationary phase

irrespective of whether As^{III} or As^{V} are present in the growth medium. Up-regulation of *aioB* and *aioA* in response to growth phase has been previously observed in *A. tumefaciens* 5A, controlled by PhoB (Kang, Heinemann, et al. 2012) and in *H. arsenicoxydans* ULPAS-1 (Cleiss-Arnold et al. 2010). Analyses of the TSS and promoter region upstream of *aioB* in NT-26 didn't result in the identification of another TSS or promoter that could possibly be regulating the response to growth phase. The TSS found is in the position expected based on the position of the σ^{54} consensus sequence upstream of *aioB.* This suggests that the up-regulation of the *aioBA* operon in response to growth phase is not related to the presence of another promoter and instead, the control of this operon can be due to the regulation of the σ^{54} activation or could be indirect through the regulation of the regulator AioR. The σ^{54} activation could be controlled either by providing activation through AioR, in the presence of another response regulator capable of interacting with the σ^{54}, or through the integration host factor (IHF) which is required to bend the DNA between the AioR and σ^{54} binding sites. It is not uncommon for response regulators to be activated by other histidine kinases (Stock et al. 2000; Bijlsma & Groisman 2003) and the σ^{54} activation by different response regulators has also been reported previously for other organisms (Samuels et al. 2013). To confirm the activation of the σ^{54} in NT-26 more experiments need to be done, a fragment containing the σ^{54} conserved region and the AioR binding site need to be cloned into the promoterless pPHU234 plasmid to observe the activation of this promoter in NT-26 grown under different conditions testing the hypothesis that σ^{54} can be active in NT-26 grown when grown in the absence of As^{III} or As^{V} in stationary phase.

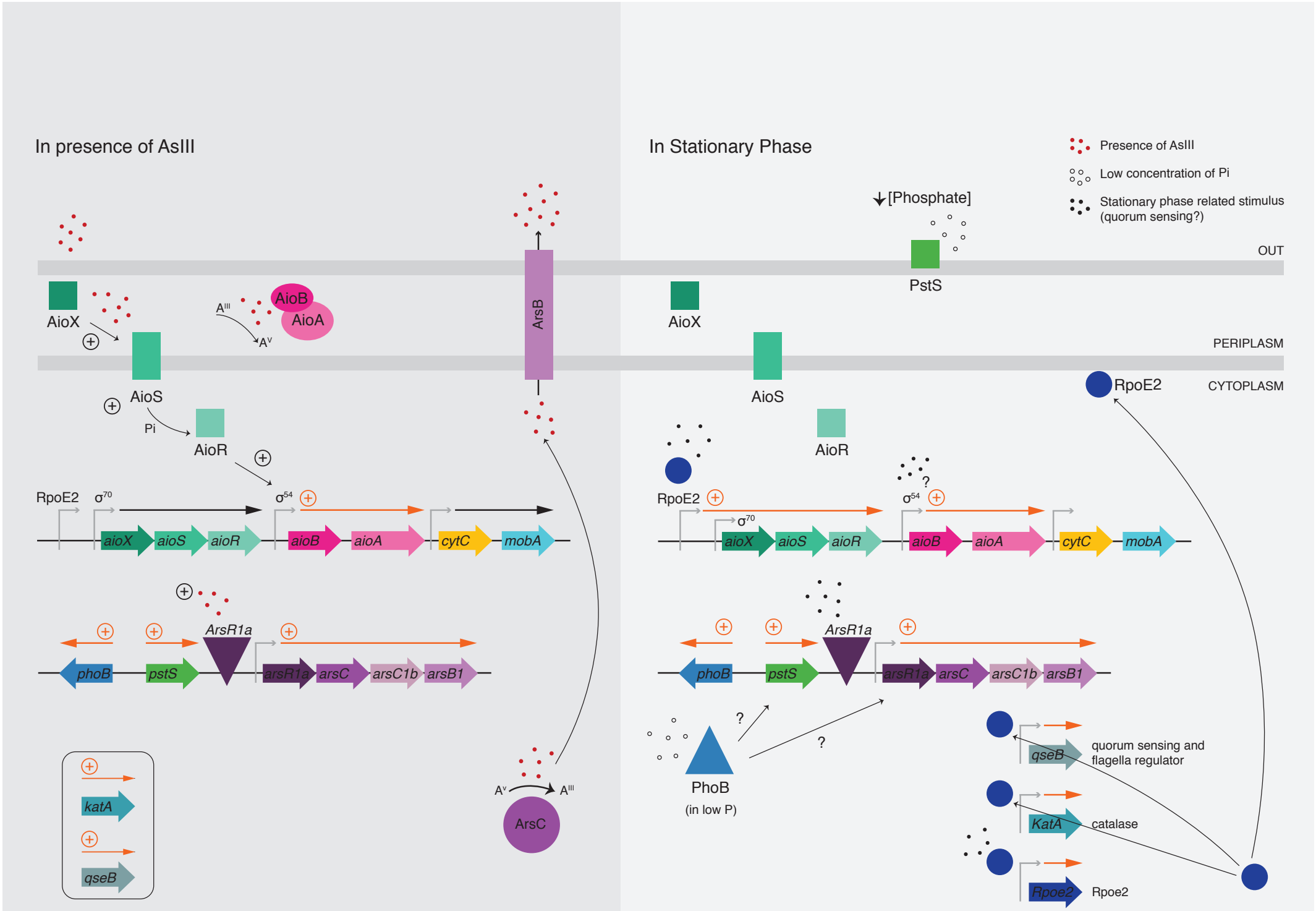
In this work it was reported for the first time that *aioB* and *aioA* might not be always co-transcribed. The analysis of qPCR results showed that *aioB* was up-regulated in stationary phase, when grown in the presence of As^{V}, but this up-regulation was not statistically significant for *aioA*. To support the hypothesis that *aioB* and *aioA* can be transcribed separately, a TSS and an active promoter were found upstream of *aioA*. The *aioA* putative promoter showed more activity when NT-26 was grown in the presence of As^{III} than with As^{V} or YE alone, in accordance to what was observed by qPCR when *aioA* was not differentially expressed in presence of As^{V}, compared to YE alone, in stationary phase. The putative promoter upstream of *aioA* needs to be identified experimentally and that can be done by using the *aioA* promoter region as a probe and incubating it with bacterial cytoplasmic extract to fish out DNA-binding proteins, then analysing the bound proteins by Mass spectrometry. Attempts to find a conserved promoter region in the putative *aioA* promoter region using bioinformatic tools were unsuccessful as no consensus sequence to known promoters was found by the various bioinformatics approaches. The putative *aioA* promoter fused to *lacZ* has to be determined to see if it is the same than the one determined in NT-26 or if it is an artefact.

The gene *cytC* encoding for cytochrome *c,* located downstream of *aioB* and *aioA* in NT-26, was shown to be constitutively expressed in accordance with previous results (Santini et al. 2007). Analyses of the *cytC* expression by qPCR also showed that *cytC* is up-regulated in the presence of As^{III} and As^{V} and according to previous results, is co-transcribed with *aioA* and *aioB* when grown under these conditions (Santini et al. 2007). That indicates that *cytC* has its own promoter that was not identified in this

study. New attempts to identify the *cytC* TSS and promoter region are needed to understand how this gene is regulated.

The up-regulation of *aioX, aioS, aioR, aioB, aioA* and *ars* genes in stationary phase in NT-26 indicates that quorum sensing might be involved in their regulation. This hypothesis was reinforced when a binding site for RpoE2 binding was found upstream of the quorum sensing regulatory gene *qseB*, indicating that the regulation of *qseB* and *aioX* in response to growth phase might be somehow connected (Figure 5.1). The involvement of quorum sensing in arsenite oxidation has been suggested previously for *A. tumefaciens* 5A (Kashyap et al. 2006) and also for *C. aurantiacus* which has a gene encoding the quorum sensing transcriptional activator LuxR, upstream of the *aioB* and *aioA* genes (Van Lis et al. 2013). More work needs to be done to understand how quorum sensing works in NT-26, for example which genes are affected by quorum sensing, what is the quorum sensing molecule and the regulators.

Based on the results presented in this thesis regarding the expression of *aioX, aioS, aioR, phoB, pstS, arsR1a, arsC* and *arsB1* in NT-26 grown in presence of either As^{III}, As^{V} and absence of both, to late-log and stationary phases, measured by qPCR; the results obtained for the promoter identification for *aioX* and *aioB*; and comparing the results with what has been published for As^{III}-oxidisers, a regulatory circuit for NT-26 was drawn to incorporate all the observations made in this thesis (Figure 5.1). The *aioA* putative promoter was not added to the figure, as it is still not identified and more work needs to be done to confirm its presence, location and function.


In presence of AsIII
In Stationary Phase
Presence of AsIII
Low concentration of Pi
Stationary phase related stimulus (quorum sensing?)
[Phosphate]
OUT
PERIPLASM
CYTOPLASM
PstS
AioX
AioB
AioA
ArsB
AioS
Pi
AioR
RpoE2
σ70
σ54
aioX
aioS
aioR
aioB
aioA
cytC
mobA
ArsR1a
phoB
pstS
arsR1a
arsC
arsC1b
arsB1
katA
qseB
ArsC
PhoB
(in low P)
quorum sensing and flagella regulator
catalase
Rpoe2

References

Anderson, G.L., Williams, J. & Hille, R., 1992. The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum-containing hydroxylase. *Journal of Biological Chemistry*, 267(33), pp.23674–23682.

Andres, J. et al., 2013. Life in an arsenic-containing gold mine: Genome and physiology of the autotrophic arsenite-oxidizing bacterium *Rhizobium* sp. NT-26. *Genome Biology and Evolution*, 5(5), pp.934–953.

Antman, K.H., 2001. Introduction: The history of arsenic trioxide in cancer therapy. *The Oncologist*, 6(90002), pp.1–2.

Aramaki, H. & Fujita, M., 1999. In vitro transcription analysis of *rpoD* in *Pseudomonas aeruginosa* PAO1. *FEMS Microbiology Letters*, 180(2), pp.311–316.

Arsène-Ploetze, F. et al., 2010. Structure, function, and evolution of the *Thiomonas* spp. genome. *PLoS Genetics*, 6(2).

Bastiat, B., Sauviac, L. & Bruand, C., 2010. Dual control of *Sinorhizobium meliloti* RpoE2 sigma factor activity by two PhyR-type two-component response regulators. *Journal of Bacteriology*, 192(8), pp.2255–2265.

Bentley, R. & Chasteen, T.G., 2002. Arsenic curiosa and humanity. *The Chemical Educator*, 7(2), pp.51–60.

Bijlsma, J.J.E. & Groisman, E.A., 2003. Making informed decisions: Regulatory interactions between two-component systems. *Trends in Microbiology*, 11(8),

pp.359–366.

Bissen, M. & Frimmel, F.H., 2003. Arsenic - A review. Part I: Occurrence, toxicity, speciation, mobility. *Acta Hydrochimica et Hydrobiologica*, 31(1), pp.9–18.

Blanco, A.G., Canals, A. & Coll, M., 2012. PhoB transcriptional activator binds hierarchically to pho box promoters. *Biological Chemistry*, 393(10), pp.1165–1171.

Browning, D.F. & Busby, Stephen J, W., 2004. The regulation of bacterial transcription initiation. *Nature Reviews Microbiology*, 2(1), pp.57–65.

Bustin, S.A. et al., 2009. The MIQE guidelines:Minimum Information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55(4), pp.611–622.

Bustin, S.A., Vandesompele, J. & Pfaffl, M., 2009. Standardization of qPCR and RT-qPCR. *OMICS*, 0, pp.3–6.

Butcher, B.G. & Rawlings, D.E., 2002. The divergent chromosomal *ars* operon of *Acidithiobacillus ferrooxidants* is regulated by an atypical ArsR protein. *Microbiology*, 148(12), pp.3983–3992.

Cai, J., Salmon, K. & DuBow, M.S., 1998. A chromosomal *ars* operon homologue of *Pseudomonas aeruginosa* confers increased resistance to arsenic and antimony in Escherichia coli. *Microbiology*, 144(10), pp.2705–2713.

Cai, L. et al., 2009. Novel gene clusters involved in arsenite oxidation and resistance in two arsenite oxidizers: *Achromobacter sp. SY8* and *Pseudomonas* sp. TS44.

Applied Microbiology and Biotechnology, 83(4), pp.715–725.

Cao, H. & Shockey, J.M., 2012. Comparison of TaqMan and SYBR green qPCR methods for quantitative gene expression in tung tree tissues. *Journal of Agricultural and Food Chemistry*, 60(50), pp.12296–12303.

Carlin, A. et al., 1995. The *ars* operon of *Escherichia coli* confers arsenical and antimonial resistance. *Journal of Bacteriology*, 177(4), pp.981–986.

Chang, D.E., Smalley, D.J. & Conway, T., 2002. Gene expression profiling of *Escherichia coli* growth transitions: An expanded stringent response model. *Molecular Microbiology*, 45(2), pp.289–306.

Chen, F. et al., 2015. Regulation of arsenite oxidation by the phosphate two-component system PhoBR in *Halomonas* sp. HAL1. *Frontiers in Microbiology*, 6(September), pp.1–9.

Chen, J., Bhattacharjee, H. & Rosen, B.P., 2015. ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone. *Molecular Microbiology*, 96(5), pp.1042–1052.

Chen, S.J. et al., 2011. From an old remedy to a magic bullet: Molecular mechanisms underlying the therapeutic effects of arsenic in fighting leukemia. *Blood*, 117(24), pp.6425–6437.

Clarke, M.B. & Sperandio, V., 2005. Transcriptional autoregulation by quorum sensing *Escherichia coli* regulators B and C (QseBC) in enterohaemorrhagic *E. coli* (EHEC).

Molecular Microbiology, 58(2), pp.441–455.

Cleiss-Arnold, J. et al., 2010. Temporal transcriptomic response during arsenic stress in *Herminiimonas arsenicoxydans*. *BMC genomics*, 11(1), p.709.

Costanzo, A. & Ades, S.E., 2006. Growth phase-dependent regulation of the extracytoplasmic stress factor, sigmE, by guanosine 3',5'-bispyrophosphate (ppGpp). *Journal of Bacteriology*, 188(13), pp.4627–4634.

Cullen WR, 2008. Arsenic and the environment. In *Is arsenic and aphrodisiac? The Sociochemistry of an Element*. pp. 287–348.

Dallas, P.B. et al., 2005. Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? *BMC genomics*, 6, p.59.

Desroche, N., Beltramo, C. & Guzzo, J., 2005. Determination of an internal control to apply reverse transcription quantitative PCR to study stress response in the lactic acid bacterium *Oenococcus oeni*. *Journal of Microbiological Methods*, 60(3), pp.325–333.

Dey, S. & Rosen, B.P., 1995. Dual mode of energy coupling by the oxyanion-translocating ArsB protein. *Journal of Bacteriology*, 177(2), pp.385–389.

Dilda, P.J. & Hogg, P.J., 2007. Arsenical-based cancer drugs. *Cancer Treatment Reviews*, 33(6), pp.542–564.

Diorio, C. et al., 1995. An *Escherichia coli* chromosomal *ars* operon homolog is

functional in arsenic detoxification and is conserved in gram-negative bacteria. *Journal of Bacteriology*, 177(8), pp.2050–2056.

Dragan, A.I. et al., 2012. SYBR Green I: Fluorescence properties and interaction with DNA. *Journal of Fluorescence*, 22(4), pp.1189–1199.

Elias, M. et al., 2012. The molecular basis of phosphate discrimination in arsenate-rich environments. *Nature*, 490(7422), pp.134–137.

Ellis, D.R. et al., 2006. A novel arsenate reductase from the arsenic hyperaccumulating fern Pteris vittata. *Plant physiology*, 141(4), pp.1544–1554.

Ellis, P.J. et al., 2001. Crystal structure of the 100 kDa arsenite oxidase from *Alcaligenes faecalis* in two crystal forms at 1.64 A and 2.03 A *Structure*, 9(2), pp.125–132.

Foreman, R., Fiebig, A. & Crosson, S., 2012. The lovK-lovR two-component system is a regulator of the general stress pathway in *Caulobacter crescentus*. *Journal of Bacteriology*, 194(12), pp.3038–3049.

Furbank, R.T., Agostino, A. & Hatch, M.D., 1990. C4 acid decarboxylation and photosynthesis in bundle sheath cells of NAD-malic enzyme-type C4 plants: Mechanism and the role of malate and orthophosphate. *Archives of Biochemistry and Biophysics*, 276(2), pp.374–381.

Galisa, P.S. et al., 2012. Identification and validation of reference genes to study the gene expression in *Gluconacetobacter diazotrophicus* grown in different carbon sources using RT-qPCR. *Journal of Microbiological Methods*, 91(1), pp.1–7.

Gladysheva, T.B., Oden, K.L. & Rosen, B.P., 1994. Properties of the arsenate reductase of plasmid R773. *Biochemistry*, 33(23), pp.7288–7293.

Green, H.H., 1918. Description of a bacterium which oxidizes arsenite to arsenate, and of one which reduces arsenate to arsenite, isolated from a cattle-dipping tank. S Afr J Sci.

Gross, C.A. et al., 1998. The functional and regulatory roles of sigma factors in transcription. In *Cold Spring Harbor Symposia on Quantitative Biology*. pp. 141–155.

Han, F.X. et al., 2003. Assessment of global industrial-age anthropogenic arsenic contamination. *Naturwissenschaften*, 90(9), pp.395–401.

Hao, X. et al., 2012. Genome sequence of the arsenite-oxidizing strain *Agrobacterium tumefaciens* 5A. *Journal of Bacteriology*, 194(4), p.903.

Harley, C.B. & Reynolds, R.P., 1987. Analysis of *E. coli* promoter sequences. *Nucleic acids research*, 15(5), pp.2343–2361.

Heimann, J.D., 2002. The extracytoplasmic function (ECF) sigma factors. *Advances in Microbial Physiology*, 46, pp.47–110.

Hellemans, J. et al., 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome biology*, 8(2), p.R19.

Vanden Hoven, R.N. & Santini, J.M., 2004. Arsenite oxidation by the heterotroph

Hydrogenophaga sp. str. NT-14: The arsenite oxidase and its physiological electron acceptor. *Biochimica et Biophysica Acta - Bioenergetics*, 1656(2–3), pp.148–155.

Hughes, M.F., 2002. Arsenic toxicity and potential mechanisms of action. *Toxicology Letters*, 133(1), pp.1–16.

Hung, H.M. et al., 1997. The behavior of the P-value when the alternative hypothesis is true. *Biometrics*, 53(1), pp.11–22.

Hunt, K.M. et al., 2014. The mechanistic basis of arsenicosis: Pathogenesis of skin cancer. *Cancer Letters*, 354(2), pp.211–219.

Ishihama, A., 2000. Functional modulation of *Escherichia coli* RNA polymerase. *Annual review of microbiology*, 54, pp.499–518.

Jackson, C.R. & Dugas, S.L., 2003. Phylogenetic analysis of bacterial and archaeal *arsC* gene sequences suggests an ancient, common origin for arsenate reductase. *BMC evolutionary biology*, 3, p.18.

Ji, G. & Silver, S., 1992. Reduction of arsenate to arsenite by the ArsC protein of the arsenic resistance operon of *Staphylococcus aureus* plasmid p1258. , 89(October), pp.9474–9478.

Ji, G. & Silver, S., 1992. Regulation and expression of the arsenic resistance operon from *Staphylococcus aureus* plasmid pI258. *Journal of Bacteriology*, 174(11), pp.3684–3694.

Kaberdin, V.R. & Bläsi, U., 2006. Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiology Reviews*, 30(6), pp.967–979.

Kang, Y.-S. et al., 2016. Regulatory Activities of Four ArsR Proteins in *Agrobacterium tumefaciens* 5A. *Applied and Environmental Microbiology*, (April), p.AEM.00262-16.

Kang, Y.-S.S. et al., 2014. Involvement of the Acr3 and DctA anti-porters in arsenite oxidation in *Agrobacterium tumefaciens* 5A. *Environmental microbiology*, 78(6), pp.5638–5645.

Kang, Y.S., Heinemann, J., et al., 2012. Integrated co-regulation of bacterial arsenic and phosphorus metabolisms. *Environmental Microbiology*, 14(12), pp.3097–3109.

Kang, Y.S., Bothner, B., et al., 2012. Involvement of RpoN in regulating bacterial arsenite oxidation. *Applied and Environmental Microbiology*, 78(16), pp.5638–5645.

Kashyap, D.R. et al., 2006. Complex regulation of arsenite oxidation in *Agrobacterium tumefaciens*. *Journal of Bacteriology*, 188(3), pp.1081–1088.

Kato, A., Latifi, T. & Groisman, E. a, 2003. Closing the loop: the PmrA/PmrB two-component system negatively controls expression of its posttranscriptional activator PmrD. *Proceedings of the National Academy of Sciences of the United States of America*, 100(8), pp.4706–4711.

Kitchen, R.R., Kubista, M. & Tichopad, A., 2010. Statistical aspects of quantitative real-time PCR experiment design. *Methods (San Diego, Calif.)*, 50(4), pp.231–6.

Koechler, S. et al., 2015. Constitutive arsenite oxidase expression detected in arsenic-hypertolerant *Pseudomonas xanthomarina* S11. *Research in Microbiology*, 166(3), pp.205–214.

Koechler, S. et al., 2010. Multiple controls affect arsenite oxidase gene expression in *Herminiimonas arsenicoxydans*. *BMC microbiology*, 10, p.53.

Krafft, T. & Macy, J.M., 1998. Purification and characterization of the respiratory arsenate reductase of *Chrysiogenes arsenatis*. *European journal of biochemistry / FEBS*, 255(3), pp.647–53.

Laursen, B.S. et al., 2005. Initiation of protein synthesis in bacteria. *Microbiology and molecular biology reviews : MMBR*, 69(1), pp.101–23.

Li, H. et al., 2013. In silico analysis of bacterial arsenic islands reveals remarkable synteny and functional relatedness between arsenate and phosphate. *Frontiers in Microbiology*, 4(NOV), pp.1–10.

Li, W. et al., 2012. Scanning assay of β-galactosidase activity. *Applied Biochemistry and Microbiology*, 48(6), pp.603–607.

Lin, M.F., Lin, Y.Y. & Lan, C.Y., 2015. The role of the two-component system BaeSR in disposing chemicals through regulating transporter systems in *Acinetobacter baumannii*. *PLoS ONE*, 10(7).

Van Lis, R. et al., 2013. Arsenics as bioenergetic substrates. *Biochimica et Biophysica Acta - Bioenergetics*, 1827(2), pp.176–188.

Liu, G. et al., 2012. A periplasmic arsenite-binding protein involved in regulating arsenite oxidation. *Environmental Microbiology*, 14(7), pp.1624–1634.

Lubin, E.A. et al., 2016. Identification of the PhoB regulon and role of PhoU in the phosphate starvation response of *Caulobacter crescentus*. *Journal of Bacteriology*, 198(1), pp.187–200.

Maeda, H. et al., 2003. Quantitative real-time PCR using TaqMan and SYBR Green for *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis*, *Prevotella intermedia*, *tetQ* gene and total bacteria. *FEMS Immunology and Medical Microbiology*, 39(1), pp.81–86.

Makino, K. et al., 1988. Regulation of the phosphate regulon of *Escherichia coli*. Activation of *pstS* transcription by PhoB protein in vitro. *Journal of Molecular Biology*, 203(1), pp.85–95.

Malasarn, D. et al., 2004. *arrA* is a reliable marker for As(V) respiration. *Science (New York, N.Y.)*, 306(5695), p.455.

Mandal, B.K. & Suzuki, K.T., 2002. Arsenic round the world: A review. *Talanta*, 58(1), pp.201–235.

Mantovani, R., 1994. An RNA polymerase II in vitro transcription system. *Methods Mol Biol*, 31, pp.289–298.

Martens, M. et al., 2008. Advantages of multilocus sequence analysis for taxonomic studies: A case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *International Journal of Systematic and Evolutionary*

Microbiology, 58(1), pp.200–214.

Marzan, L.W. & Shimizu, K., 2011. Metabolic regulation of *Escherichia coli* and its *phoB* and *phoR* genes knockout mutants under phosphate and nitrogen limitations as well as at acidic condition. TL - 10. *Microbial cell factories*, 10 VN-r(1), p.39.

Mccann, S.H. et al., 2016. Arsenite as an Electron Donor for Anoxygenic Photosynthesis : Description of Three Strains of *Ectothiorhodospira* from Mono Lake , California and Big Soda Lake , Nevada. *Life*, (Iii).

Melkonian, S. et al., 2012. Intakes of several nutrients are associated with incidence of arsenic-related keratotic skin lesions in Bangladesh. *The Journal of nutrition*, 142(12), pp.2128–34.

Meng, Y.L., Liu, Z. & Rosen, B.P., 2004. As(III) and Sb(III) uptake by GlpF and efflux by ArsB in *Escherichia coli. Journal of Biological Chemistry*, 279(18), pp.18334–18341.

Michael, H., Willsky, G.R. & Malamy, M.H., 1980. Effect of arsenate on inorganic phosphate transport in *Escherichia coli*. *Journal of bacteriology*, 144(1), pp.366–74.

Miller, J.H., 1972. *Experiments in molecular genetics*,

Miller, M.B. & Bassler, B.L., 2001. Quorum sensing in bacteria. *Annual review of microbiology*, 55, pp.165–99.

Mitrophanov, A.Y., Hadley, T.J. & Groisman, E.A., 2010. Positive autoregulation shapes response timing and intensity in two-component signal transduction systems.

Journal of Molecular Biology, 401(4), pp.671–680.

Moinier, D., Slyemi, D., Byrne, D., Lignon, S., Lebrun, R., et al., 2014. An ArsR/SmtB family member is involved in the regulation by arsenic of the arsenite oxidase operon in *Thiomonas arsenitoxydans*. *Applied and Environmental Microbiology*, 80(20), pp.6413–6426.

Morrison, T.B., Weis, J.J. & Wittwer, C.T., 1998. Quantification of lowcopy transcripts by continuous SYBR green I monitoring during amplification. *BioTechniques*, 24(6), pp.954–962.

Mukhopadhyay, R. et al., 2002. Microbial arsenic: From geocycles to genes and enzymes. *FEMS Microbiology Reviews*, 26(3), pp.311–325.

Murphy, J.N. & Saltikov, C.W., 2009. The ArsR repressor mediates arsenite-dependent regulation of arsenate respiration and detoxification operons of *Shewanella sp*. strain ANA-3. *Journal of Bacteriology*, 191(21), pp.6722–6731.

Nadar, V.S. et al., 2016. Structure of the ArsI C–As lyase: Insights into the mechanism of degradation of organoarsenical herbicides and growth promoters. *Journal of molecular biology*, 428(11), pp.2462–2473.

Natale, P., Bruser, T. & Driessen, A.J.M., 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-Distinct translocases and mechanisms. *Biochimica et Biophysica Acta - Biomembranes*, 1778(9), pp.1735–1756.

Navarro Llorens, J.M., Tormo, A. & Martinez-Garcia, E., 2010. Stationary phase in gram-

negative bacteria. *FEMS Microbiology Reviews*, 34(4), pp.476–495.

Newman, D.K., Ahmann, D. & Morel, F.M.M., 1998. A brief review of microbial arsenate respiration. *Geomicrobiology Journal*, 15(4), pp.255–268.

Nickson, R.T. et al., 2000. Mechanism of arsenic release to groundwater, Bangladesh and West Bengal. *Applied Geochemistry*, 15(4), pp.403–413.

Nordstrom, D.K., 2012. Arsenic in the geosphere meets the anthroposphere. In *Understanding the Geological and Medical Interface of Arsenic, As 2012 - 4th International Congress: Arsenic in the Environment*. pp. 15–19.

Oremland, R.S. et al., 2002. Anaerobic oxidation of arsenite in Mono Lake water and by a facultative, arsenite-oxidizing chemoautotroph, strain MLHE-1. *Applied and Environmental Microbiology*, 68(10), pp.4795–4802.

Oremland, R.S. et al., 2009. Arsenic in the evolution of earth and extraterrestrial Ecosystems. *Geomicrobiology Journal*, 26(V), pp.522–536.

Oremland, R.S. & Stolz, J.F., 2003. The ecology of arsenic. *Science*, 300, pp.939–944.

Osborne, T.H. et al., 2015. Isolation of an arsenate-respiring bacterium from a redox front in an arsenic-polluted aquifer in West Bengal, Bengal basin. *Environmental Science and Technology*, 49(7), pp.4193–4199.

Osborne, T.H. et al., 2010. Microbial oxidation of arsenite in a subarctic environment: diversity of arsenite oxidase genes and identification of a psychrotolerant arsenite oxidiser. *BMC microbiology*, 10, p.205.

Osborne, T.H. & Santini, J.M., 2012. *The Metabolism of Arsenite*,

Paez-Espino, A.D., Durante-Rodriguez, G. & de Lorenzo, V., 2015. Functional coexistence of twin arsenic resistance systems in *Pseudomonas putida* KT2440. *Environmental Microbiology*, 17(1), pp.229–238.

Páez-Espino, D. et al., 2009. Microbial responses to environmental arsenic. In *BioMetals*. pp. 117–130.

Paget, M.S. & Helmann, J.D., 2003. The σ70 family of sigma factors. *Genome Biology*, 4(1), p.203.

Qin, J. et al., 2006. Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), pp.2075–80.

Qiu, Y. et al., 2013. Characterizing the interplay between multiple levels of organization within bacterial sigma factor regulatory networks. *Nature communications*, 4, p.1755.

Ramírez-Romero, M.A. et al., 2006. The *Rhizobium etli* σ70 (SigA) factor recognizes a lax consensus promoter. *Nucleic Acids Research*, 34(5), pp.1470–1480.

Rhine, E.D. et al., 2005. Environmental microbes can speciate and cycle arsenic. *Environmental Science and Technology*, 39(24), pp.9569–9573.

Richey, C. et al., 2009. Respiratory arsenate reductase as a bidirectional enzyme.

Biochemical and Biophysical Research Communications, 382(2), pp.298–302.

Rohe, G.H., 1896. *Arsenic. In Reference-Book of Practical Therapeutics* 1st ed.,

Rosen, B.P., 2002. Biochemistry of arsenic detoxification. *FEBS Letters*, 529(1), pp.86–92.

Rosen, B.P., 1999. Families of arsenic transporters. *Trends in Microbiology*, 7(5), pp.207–212.

Rosen, B.P. et al., 1988. Molecular characterization of an anion pump. The *arsA* gene product is an arsenite(antimonate)-stimulated ATPase. *Journal of Biological Chemistry*, 263(7), pp.3067–3070.

Rosen, B.P., Ajees, A. & Mcdermott, T.R., 2011. Life and death with arsenic. *BioEssays*, 33(5), pp.350–357.

Rosen, B.P., Bhattacharjee, H. & Shi, W., 1995. Mechanisms of metalloregulation of an anion-translocating ATPase. *Journal of Bioenergetics and Biomembranes*, 27(1), pp.85–91.

Rosenstein, R. et al., 1992. Expression and regulation of the antimonite, arsenite, and arsenate resistance operon of *Staphylococcus xylosus* plasmid pSX267. *Journal of Bacteriology*, 174(11), pp.3676–3683.

Samuel M. Cohen, Aparajita Chowdhury & Lora L. Arnold, 2016. Inorganic arsenic: A non-genotoxic carcinogen. *Journal of Environmental Sciences*.

Samuels, D.J. et al., 2013. Use of a promiscuous, constitutively-active bacterial

enhancer-binding protein to define the σ^{54} (RpoN) regulon of *Salmonella typhimurium* LT2. *BMC genomics*, 14, p.602.

Santini, J.M. et al., 2000. A new chemolithoautotrophic arsenite-oxidizing bacterium isolated from a gold mine: Phylogenetic, physiological, and preliminary biochemical studies. *Applied and Environmental Microbiology*, 66(1), pp.92–97.

Santini, J.M. et al., 2007. The NT-26 cytochrome c552 and its role in arsenite oxidation. *Biochimica et Biophysica Acta - Bioenergetics*, 1767(2), pp.189–196.

Santini, J.M. & Vanden Hoven, R.N., 2004. Molybdenum-containing arsenite oxidase of the chemolithoautotrophic arsenite oxidizer NT-26. *Journal of Bacteriology*, 186(6), pp.1614–1619.

Sardiwal, S. et al., 2010a. Characterization of a two-component signal transduction system that controls arsenite oxidation in the chemolithoautotroph NT-26. *FEMS Microbiology Letters*, 313(1), pp.20–28.

Sardiwal, S. et al., 2010b. Characterization of a two-component signal transduction system that controls arsenite oxidation in the chemolithoautotroph NT-26. *FEMS Microbiology Letters*, 313(1), pp.20–28.

Sauviac, L. et al., 2007. An extracytoplasmic function sigma factor acts as a general stress response regulator in *Sinorhizobium meliloti*. *Journal of Bacteriology*, 189(11), pp.4204–4216.

Slyemi, D. et al., 2013. Organization and regulation of the arsenite oxidase operon of the moderately acidophilic and facultative chemoautotrophic *Thiomonas*

arsenitoxydans. *Extremophiles*, 17(6), pp.911–920.

Slyemi, D. & Bonnefoy, V., 2012. How prokaryotes deal with arsenic. *Environmental Microbiology Reports*, 4(6), pp.571–586.

Smedley, P.L. & Kinniburgh, D.G., 2002. A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry*, 17(5), pp.517–568.

Soncini, F.C., Vescovi, E.G. & Groisman, E.A., 1995. Transcriptional autoregulation of the *Salmonella typhimurium* phoPQ operon. *Journal of Bacteriology*, 177(15), pp.4364–4371.

Spearman, C., 1904. Spearman's rank correlation coefficient. *Amer. J. Psychol.*, 15, pp.72–101.

Sperandio, V., Torres, A.G. & Kaper, J.B., 2002. Quorum sensing *Escherichia coli* regulators B and C (QseBC): A novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in *E. coli*. *Molecular Microbiology*, 43(3), pp.809–821.

Stock, A.M., Robinson, V.L. & Goudreau, P.N., 2000. Two-component signal transduction. *Annual Review Of Biochemistry*, 69, pp.183–215.

Stolz, J.F. et al., 2006. Arsenic and selenium in microbial metabolism. *Annual review of microbiology*, 60, pp.107–130.

Stolz, J.F., Basu, P. & Oremland, R.S., 2002. Microbial transformation of elements: The

case of arsenic and selenium. *International Microbiology*, 5(4), pp.201–207.

Sun, W. et al., 2010. Anaerobic oxidation of arsenite linked to chlorate reduction. *Applied and Environmental Microbiology*, 76(20), pp.6804–6811.

Suzuki, K. et al., 1998. Expression and regulation of the arsenic resistance operon of *Acidiphilium multivorum* AIU 301 plasmid pKW301 in *Escherichia coli*. *Applied and Environmental Microbiology*, 64(2), pp.411–418.

Tamura, K. et al., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), pp.2725–2729.

Taylor, S. et al., 2010. A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. *Methods*, 50(4).

Thompson, J.D., Higgins, D.G. & Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22(22), pp.4673–4680.

Thorburn, a L., 1983. Paul Ehrlich: pioneer of chemotherapy and cure by arsenic (1854-1915). *The British journal of venereal diseases*, 59(June), pp.404–405.

Tisa, L.S. & Rosen, B.P., 1990. Molecular characterization of an anion pump. The ArsB protein is the membrane anchor for the ArsA protein. *Journal of Biological Chemistry*, 265(1), pp.190–194.

Travers, A.A. & Burgess, R.R., 1969. Cyclic re-use of the RNA polymerase sigma factor.

Nature, 222(5193), pp.537–540.

Trimble, W.L. et al., 2012. Draft genome sequence of *Agrobacterium albertimagni* strain AOL15. *Journal of Bacteriology*, 194(24), pp.6986–6987.

Vahidnia, A., van der Voet, G.B. & de Wolff, F.A., 2007. Arsenic neurotoxicity A review. *Human & Experimental Toxicology*, 26(10), pp.823–832.

Vandesompele, J. et al., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*, 3(7), p.RESEARCH0034.

Vandesompele, J., Kubista, M. & Pfaffl, M.W., 2009. Reference gene validation software for improved normalization. In *Real-time PCR: Current technology and applications.* pp. 47–64.

Wang, G. et al., 2004. Arsenic resistance in *Halobacterium* sp. strain NRC-1 examined by using an improved gene knockout system. *Journal of Bacteriology*, 186(10), pp.3187–3194.

Wang, L. et al., 2009. Identification of an arsenic resistance and arsenic-sensing system in *Campylobacter jejuni*. *Applied and Environmental Microbiology*, 75(15), pp.5064–5073.

Wang, Q. et al., 2015. Fate of arsenate following arsenite oxidation in *Agrobacterium tumefaciens* GW4. *Environmental Microbiology*, 17(6), pp.1926–1940.

Waters, C.M. & Bassler, B.L., 2005. Quorum sensing: cell-to-cell communication in

bacteria. *Annual review of cell and developmental biology*, 21, pp.319–346.

Westheimer, F.H., 1987. Why nature chose phosphates. *Science*, 235(4793), pp.1173–1178.

WHO, 2001. *Arsenic and arsenic compounds. Environmental Health Criteria 224*,

WHO, 2010. Exposure to Arsenic: A Major Public Health Concern. *Agriculture*, p.5.

Wösten, M.M.S.M., 1998. Eubacterial sigma-factors. *FEMS Microbiology Reviews*, 22(3), pp.127–150.

Wu, B., Song, J. & Beitz, E., 2010. Novel channel enzyme fusion proteins confer arsenate resistance. *Journal of Biological Chemistry*, 285(51), pp.40081–40087.

Wu, J. & Rosen, B.P., 1993. Metalloregulated expression of the *ars* operon. *Journal of Biological Chemistry*, 268(1), pp.52–58.

Wu, J. & Rosen, B.P., 1993. The *arsD* gene encodes a second trans-acting regulatory protein of the plasmid-encoded arsenical resistance operon. *Molecular microbiology*, 8(3), pp.615–23.

Wu, J. & Rosen, B.P., 1991. The ArsR protein is a trans-acting regulatory protein. *Molecular microbiology*, 5(6), pp.1331–6.

Xu, C. et al., 1998. Metalloid resistance mechanisms in prokaryotes. *Journal of biochemistry*, 123(1), pp.16–23.

Yang, C. et al., 2012. Genome-wide PhoB binding and gene expression profiles reveal

the hierarchical gene regulatory network of phosphate starvation in *Escherichia coli. PLoS ONE*, 7(10).

Yang, H. et al., 2005. Novel pathway for arsenic detoxification in the legume symbiont *Sinorhizobium meliloti* novel pathway for arsenic detoxification in the legume symbiont *Sinorhizobium meliloti*. *Society*, 187(20), pp.6991–6997.

Yang, H.-C. & Rosen, B.P., 2016. New mechanisms of bacterial arsenic resistance. *Biomedical journal*, 39(1), pp.5–13.

Yoshinaga, M., Cai, Y. & Rosen, B.P., 2011. Demethylation of methylarsonic acid by a microbial community. *Environmental Microbiology*, 13(5), pp.1205–1215.

Yoshinaga, M. & Rosen, B.P., 2014. A C·As lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 111(21), pp.7701–6.

Zargar, K. et al., 2012. ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. *Environmental Microbiology*, 14(7), pp.1635–1645.

Zargar, K. et al., 2010. Identification of a novel arsenite oxidase gene, *arxA,* in the haloalkaliphilic, arsenite-oxidizing bacterium *Alkalilimnicola ehrlichii* strain MLHE-1. *Journal of Bacteriology*, 192(14), pp.3755–3762.

Zhu, Y.-G. et al., 2014. Earth Abides Arsenic Biotransformations. *Annual Review of Earth and Planetary Sciences*, 42(1), pp.443–467.