UNIVERSITEIT • STELLENBOSCH • UNIVERSITY

jou kennisvennoot • your knowledge partner

The Identification and Characterisation of the Arsenic Resistance Genes of the Gram-positive bacterium, *Sulfobacillus thermosulfidooxidans* VKM B-1269^{T}

Jacobus Arnoldus van der Merwe

Thesis presented in partial fulfillment of the requirements for the degree of Master of Science at the University of Stellenbosch

Supervisor: Professor Douglas E. Rawlings

March 2007

I, the undersigned, hereby declare that the work contained in this thesis is my own original work unless otherwise referenced or acknowledged and that I have not previously, in its entirely or part, submitted in to any university for a degree.

…….. ……..

J.A. van der Merwe Date

ABSTRACT

The arsenic resistance operon (*ars* operon) of the Gram-positive, iron-oxidizing, acidophilic, moderately thermophilic bacterium, *Sulfobacillus thermosulfidooxidans* VKM B-1269^T (*Sb. t.* VKM B-1269^T), was isolated and characterised. The *ars* operon was chromosomally located and consisted of an *arsR* (codes for a transcriptional regulator) and an *arsB* (codes for a membrane located arsenic/antimony efflux pump). The *arsRB* genes were transcribed in the same direction. An *arsC* (codes for an arsenate reductase), usually associated with *ars* operons, was absent from this *ars* operon. PCR and Southern-hybridization experiments revealed that no *arsC*, representative of either the Grx/GSH or Trx ArsC families was present in the genome of *Sb. t.* VKM B-1269^T. An interesting feature of the *ars* operon was the presence of a gene encoding a 525 amino acid (60.83 kDa) kumamolisin-As precursor located upstream of the *arsRB* operon. The intergenic region between the termination end of the kumamolisin-As precursor gene and the transcriptional start of the *arsR* gene was only 77 bp, suggesting that this *ars* operon might consist of three genes. RT-PCR analysis showed that the *ars* operon of *Sb. t.* VKM B-1269^T, was not co-transcribed with the kumamolisin-As precursor gene in its native *Sulfobacillus* host.

The *ars* operon of *Sb. t.* VKM B-1269^T did not complement an *Escherichia coli* arsenic sensitive mutant. mRNA transcript analysis and promoter expression studies confirmed that processes involved in the production of functional proteins from the *ars* operon transcript were likely to be responsible for the inability of the *arsRB* operon of *Sb. t.* VKM B-1269^T to confer resistance to arsenic in the heterologous *E. coli* host.

Eight *Sulfobacillus* strains isolated from different geographical areas were subjected to amplified ribosomal DNA restriction enzyme analysis (ARDREA) using the restriction endonuclease *Eco*1015 (*Sna*BI) and revealed that they could be divided into the proposed *Sulfobacillus* spp. subgroup I and subgroup II, respectively (Johnson *et al*., 2005). The presence, distribution and relatedness of the *ars* genes among members of genus *Sulfobacillus* was determined. Phylogenetic sequence comparisons revealed two clearly defined *arsB* clusters within genus *Sulfobacillus* and showed that the *arsB* of a specific *Sulfobacillus* sub specie is distinctive of that specific *Sulfobacillus* sub specie. Futhermore, sequence analysis of the isolated *arsB* homologue fragments from the respective *Sulfobacillus* spp. showed that

four distinctive profiles could be identified based on differences in the location of restriction endonuclease recognition sites.

OPSOMMING

Die arseen weerstandbiedendheidsoperon (*ars* operon) van die Gram-positiewe, ysteroksiderende, asidofiliese, matige termofiliese bakterium, *Sulfobacillus thermosulfidooxidans* VKM B-1269^T (*Sb. t.* VKM B-1269^T), was geïsoleer en gekarakteriseer. Die *ars* operon was op die chromosoom geleë en het uit 'n *arsR* (kodeer vir 'n transkripsionele reguleerder) en 'n *arsB* (kodeer vir 'n membraan geleë arseen/timien uitskeidings pomp) bestaan. Die *arsRB* gene word in dieselfde rigting getranskribeer. 'n *arsC* (kodeer vir 'n arsenaat reductase), wat gewoontlik geassosïeer word met *ars* operons, was afwesig van hierdie *ars* operon. PKR en Southern-hibridisasie eksperimente het aangedui dat geen *arsC*, verteenwoordigend van beide die Grx/GSH of Trx ArsC families, nie teenwoordig was in die genoom van *Sb. t.* VKM B-1269^T, nie. 'n Interressante eienskap van hierdie *ars* operon was die teenwoordigheid van 'n geen wat stroom-op van die *arsRB* operon geleë is en 'n 525 amino suur (60.83 kDa) kumamolisin-As voorloper kodeer. Die intergeniese gedeelte tussen die terminerings einde van die kumamolisin-As voorloper en die transkriptionele begin van die *arsR* geen was slegs 77 bp, wat voorgestel het dat die *ars* operon moontlik uit drie gene bestaan. RT-PKR analiese het bewys dat die *ars* operon van *Sb. t.* VKM B-1269^T, nie geko-getranskribeer word met die kumamolisin-As voorloper in sy oorspronklike *Sulfobacillus* gasheer nie.

Die *ars* operon van *Sb. t.* VKM B-1269^T, het nie 'n *Escherichia coli* arseen sensitiewe mutant gekomplimenteer nie. mRNA transkrip-analiese en promoter uitdrukkings eksperimente het bevestig dat prosesse wat betrokke is in die produksie van funksionele proteïene vanaf die *ars* operon transkrip, moontlik vir die onvermoë van die *arsRB* operon van *Sb t.* VKM B-1269^T verantwoordelik was om weerstandbiedendheid teen arseen in die heteroloë *E. coli* gasheer te verleen.

Agt *Sulfobacillus* stamme wat geïsoleer is vanuit verskillende geografiese areas, was onderhewig aan geamplifiseerde ribosomale DNA restriksie-ensiem-analiese (ARDREA) deur gebruik te maak van restriksie endonuklease *Eco*1015 (*Sna*BI) en het aangedui dat hulle in die voorgestelde *Sulfobacillus* spp. subgroup I en subgroup II ingedeel kan word (Johnson *et al*., 2005). Die aanwesigheid, verspreiding en verwantskappe van die *ars* gene tussen lede van genus *Sulfobacillus* was bepaal. Filogenetiese DNA volgorde vergelykings het aangedui dat

twee duidelik definïeerbare *arsB* groepe van mekaar onderskei kan word en dat die *arsB* van ‘n spesifieke *Sulfobacillus* sub spesie uniek tot daardie spesifieke *Sulfobacillus* subspesie is. Bykomend, DNA volgorde analiese van die geïsoleerde *arsB* homoloog fragmente van die *Sulfobacillus* spp. het gewys dat vier unieke profiele, op grond van verskille in die ligging van restriksie ensiem herkenning setels, geïdentifiseer kan word.

ACKNOWLEDGEMENTS

The following dissertation was fulfilled during the time from January 2004 until December 2006 in the Department of Microbiology, University of Stellenbosch, South Africa. During this time many persons made invaluable contributions to this project. I wish to express my sincere gratitude and appreciation to the following persons/institutions for their invaluable contributions to the successful completion of this study:

My supervisor, Prof. D.E. Rawlings, for his unending patience and guidance throughout this project. His experience, skill, determination and quenchless passion for science made it a privilege to study under his supervision. Your encouragement and belief in me brought out the best in me and kept me focused and motivated during difficult times.

My co-supervisor, Dr. S.M. Deane, for her invaluable support and encouragement throughout the duration of this project. Your endless enthusiasm has truly been an inspiration and great motivation to me. Thank you also for proof-reading my thesis and correcting my many mistakes.

My sincere thanks also go to all my colleagues of the Biomining Research Group, especially Dr. Marla Tuffin and Lonnie van Zyl, for your ever-willingness to help. I am greatly indebted to you for your valuable guidance and discussions during the completion of this work. Thank you for your daily support, your ideas and insight into my project and most of all, your patience.

To all the many friends (local and abroad) I am privileged to have. Thank you all for the countless great times we shared in the past. A special mention to Ike James, Maurits Perold, Andre Kotzé, Nic Northcote and the rest of my brothers at Van der Stel Cricket Club. Your encouragement and support throughout the duration of this work is more appreciated than you will know.

I received endless love and support from my whole family throughout my studies. Thank you for the interest you showed in my work and the continuous support which inspired me to make each and everyone of you proud. I am truly blessed to have great parents that

continuously encourage me and keep me motivated. Thank you for your love and absolute belief in my abilities when times were tough during the course of this study. Your relentless support and prayers throughout my years of study, especially when I was writing this thesis, was the driving force behind striving to always give my absolute best.

Finally, our Heavenly Farther for blessing me with the mental ability, strength and perseverance to be able to complete this project.

This work was funded by grants from the National Research Foundation (NRF) (Pretoria, South Africa), BHP Billiton Minerals Technology (Randburg, South Africa) and the University of Stellenbosch. Prof. D.B. Johnson kindly provided seven of the *Sulfobacillus* isolates tested during the course of this study.

This work is dedicated to my brother, Pieter, and my parents, Kobus and Estelle, who have shown me the true meaning of courage.

TABLE OF CONTENTS

TITLE PAGE ..... **i**

CERTIFICATION ..... **ii**

ABSTRACT ..... **iii**

OPSOMMING ..... **v**

ACKNOWLEDGEMENTS ..... **vii**

LIST OF FIGURES ..... **xii**

LIST OF TABLES ..... **xvii**

ABBREVIATIONS ..... **xviii**

CHAPTER 1: GENERAL INTRODUCTION 1

CHAPTER 2: THE ISOLATION, SEQUENCING AND ANALYSIS OF THE ARSENIC RESISTANCE GENES OF *SULFOBACILLUS THERMOSULFIDOOXIDANS* VKM B-1269^{T} 38

CHAPTER 3: THE EXPRESSION AND REGULATION OF THE ARSENIC RESISTANCE GENES OF *SULFOBACILLUS THERMOSULFIDOOXIDANS* VKM B-1269^{T} 64

CHAPTER 4: THE DISTRIBUTION AND EVOLUTIONARY RELATIONSHIP OF ARSENIC RESISTANCE GENES WITHIN GENUS *SULFOBACILLUS* 85

CHAPTER 5: GENERAL DISCUSSION 112

APPENDIX 1: Description of additional clones and primers used during the sequencing of pStArs1 119

APPENDIX 2: Annotated sequence obtained from pStArs1 120

REFERENCES 128

LIST OF FIGURES

Figure 1.1: The migration of arsenic during bacterial oxidation of arsenopyrite 10

Figure 1.2: Arsenical transport and detoxification pathways in prokaryotes..... 13

Figure 1.3: The genes and products of the arsenic resistance (*ars*) operons of *Staphylococcus aureus* plasmid pI258 and *Escherichia coli* plasmid pR773..... 15

Figure 1.4: The dual mode of energy coupling of the arsenic transport systems of *E.coli* plasmid R773 23

Figure 1.5: Model of the As(III)/Sb(III)-stimulated ArsA ATPase of the *E. coli* plasmid R773 27

Figure 1.6: The proposed reaction mechanism pathway of the $ArsC_{ec}$-family..... 30

Figure 1.7: The proposed reaction mechanism pathway of the $ArsC_{sa}$-family 31

Figure 1.8: A model of the metalloregulatory circuit of ArsR and ArsD, using the *ars* operon of plasmid R773 as an example..... 36

Figure 2.1: An autoradiograph following Southern-hybridization indicating the 6kb-8kb fragments targeted for mini-genebank construction 46

Figure 2.2: Southern-hybridization of pStArs1 probed against genomic DNA of *Sb. thermosulfidoxidans* VKM B-1269^T (*Sb. t.* VKM B-1269^T)..... 47

Figure 2.3: The physical and genetic map of pStArs1 49

Figure 2.4: Phylogenetic tree of a selection of different ArsB homologues from several known *ars* operons 51

Figure 2.5: Multiple sequence alignment comparison of a selection of different ArsB homologues to the *Sb. t.* VKM B-1269^T ArsB..... 52

Figure 2.6: A profile of the putative trans-membrane spanning domains and polypeptide loops of the ArsB proteins of (A) *Sb. t.* VKM B-1269^T, (B) *S. aureus* plasmid pI258 and (C) *E. coli* plasmid R773, as generated by the web-based TMHMM program 53

Figure 2.7: Phylogenetic tree of ArsR proteins from well documented *ars* operons..... 54

Figure 2.8: Multiple sequence alignment comparison of different ArsR homologues to the *Sb. t.* VKM B-1269^T ArsR 56

Figure 2.9: Secondary structure profiles of ArsR proteins of (A) *Sb. t.* VKM B-1269^T, (B) *E. coli* plasmid R773, (C) *B. halodurans* and (D) *S. aureus* plasmid pI258, as predicted by the PsiPRED web-based program 57

Figure 2.10: Multiple nucleotide sequence alignment comparison of representatives of the Trx ArsC family in order to identify shared concerved sequences 59

Figure 3.1: Binding positions of primers used for RT-PCR analysis..... 70

Figure 3.2: The sequence of the putative promoter region of the kumamolisin-As precursor gene 71

Figure 3.3: The sequence of the putative promoter region of *arsR*..... 72

Figure 3.4: The autoradiograph showing differences in mRNA expression levels under uninduced and induced conditions..... 75

Figure 3.5: The quantitative indication of the mRNA expression levels of pStArs1 under uninduced and induced conditions, as determined by the UVIgeltec software program..... 76

Figure 3.6: An ethidium bromide-strained agarose gel of RT-PCR products showing that the kumamolisin-As precursor gene is co-transcribed with the *ars* genes of pStArs1 in the heterologous *E. coli* host 77

Figure 3.7: An ethidium bromide-stained agarose gel of RT-PCR products showing that the *arsR* of pStArs1 contains a functional promoter..... 78

Figure 3.8: An ethidium bromide-stained agarose gel of RT-PCR products showing that the *ars* operon of *Sb. t.* VKM B-1269^{T} is not co-transcribed with the kumamolisin-As precursor gene in its native *Sulfobacillus* host 79

Figure 3.9: β-galactosidase activity of the promoter-*lacZ* fusions of the kumamolisin-As precursor (pMCStKumppr800) and *arsR* (pMCStArsRpr)..... 80

Figure 4.1: Computer software-generated restriction enzyme maps used in the analysis and differentiation of the 16S rDNA genes of acidophilic bacteria 94

Figure 4.2: An 1% agarose gel showing the respective DNA banding patterns after ARDREA was performed on mesophilic and moderately thermophilic Gram-positive, iron-oxidizing acidophilic bacteria 95

Figure 4.3: Restriction enzyme maps generated by computer of digested 16S rDNA genes for the analysis and differentiation of *Sulfobacillus* spp..... 97

Figure 4.4: An 1% agarose gel showing the DNA banding patterns corresponding to the *Sulfobacillus* spp., subgroup I and *Sulfobacillus* spp., subgroup II profile after ARDREA was performed on the respective *Sulfobacillus* isolates..... 98

Figure 4.5: A diagram showing the binding positions of Ferro arsBfwd and Ferro aerBrev with respect to the *arsB* of *Sb. t.* VKM B-1269^{T} 99

Figure 4.6: The identification and distribution of putative *arsB* homologous within the respective *Sulfobacillus* isolates..... 100

Figure 4.7: Phylogenetic trees of *Sulfobacillus* spp. based on their identified *arsB* homologues or their respective 16S rDNA nucleotide sequences..... 102

Figure 4.8: Computer software-generated restriction enzyme maps showing the four main profiles of the identified and sequenced sections of the respective *Sulfobacillus* spp. *arsB* homologues 103

Figure 4.9: PCR amplification of genomic DNA of the *Sb. t.* isolates with primer set KaraArsBfwd and KaraArsBrev 104

Figure 4.10: PCR amplification of the genomic DNA of the *Sb. thermosulfidooxidans* (*Sb. t.*) isolates to determine layout similarities with respect to their putative *ars* operons 105

Figure 4.11: Trans-alternating field electrophoresis preformed on the genomic DNA of the *Sb. t.* isolates to determine the location of their respective *ars* operons 107

LIST OF TABLES

Table 1.1: Bacterial arsenic resistance (*ars*) operons that have been molecularly characterized 16

Table 2.1: Bacterial strains, plasmids and PCR primers used in this study (Chapter 2)..... 40

Table 2.2: The open reading frames and features of construct pStArs1 48

Table 2.3: The upper level of some heavy metals concentrations where metabolic activity has been reported in the listed neutrophilic and acidophilic microorganisms 60

Table 3.1: Bacterial strains, plasmids and PCR primers used in this study (Chapter 3)..... 66

Table 4.1: Bacterial strains, plasmids and PCR primers used in this study (Chapter 4)..... 89

Table 4.2: The *Sulfobacillus* isolates representative of the two *Sulfobacillus* spp., subgroups 109

ABBREVIATIONS

α	alpha
~	approximately
β	beta
∞	eternity
>	more then
aa	amino acid
A	adenosine
ADP	adenosine 5'-diphosphate
Amp	ampicillin
ANDREA	amplified ribosomal DNA restriction enzyme analysis
ATP	adenosine 5'-triphosphate
bp	base pairs
°C	degrees Celsius
C	cytosine
C-terminal	carboxyl-terminus
CTAB	hexadecyltrimethyl ammonium bromide
Cys	cysteine
dH_20	distelled water
DIG	dioxigenin-11-dUTP (DIG-dUTP)
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
g	gram(s)
G	guanine
G+C	guanine:cytosine ratio
H	hour(s)
H_2SO_4	sulfuric acid
His	histidine
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase pair(s) or 1000bp
kDa	kilo Daltons

LA	Luria Bertani agar
LB	Luria Bertani broth
M	molar
mA	milli-ampere
MBD	metal binding domain
mg	milligrams
MIC	minimal inhibitory concentration
ml	milliliters
mm	millimeters
mM	millimolar
mRNA	messenger ribonucleic acid
N-terminal	amino terminus
NBD	nucleotide binding domain
NCBI	National Center of Biotechnology Information
O/N	over night
O/P	operator/promoter region
OD_{600}	optical density at 600 nanometers
ORF	open reading frame
p	plasmid
PCR	polymerase chain reaction
PFGE	pulse field gel electrophoresis
pH	potential of hydrogen
RBS	ribosome binding site
rDNA	ribosomal deoxyribonucleic acid
RFLP	β restriction fragment length polymorphism
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
s	second(s)
S	Svedberg unit
SDS	sodium dodecyl sulfate
Ser	serine
SET	sucrose EDTA buffer
spp.	several species
SSC	saline-sodium citrate
T	thymine
T_{opt}	optimum growth temperature
TBE	Tris-borate EDTA buffer
TE	Tris EDTA buffer

Tris	Tris (hydroxymethyl) aminomethane
Trp	tryptophan
UV	ultraviolet
v/v	volume/volume
V	volts
w/v	weight/volume

CHAPTER 1

GENERAL INTRODUCTION

TABLE OF CONTENTS

1.1 Introduction to biomining 2

1.2 Microbial diversity in biomining environments 4

1.3 Characteristics of genus *Sulfobacillus* 7

1.4 The chemical and biological properties of arsenic 9

1.5 Arsenic resistance mechanisms in microorganisms 11

1.5.1 The molecular genetics of efflux systems involved in bacterial arsenic resistance 12

1.5.2 The general structure of bacterial arsenic resistance (*ars*) operons 14

1.6 Variations in the structure of bacterial *ars* operons 16

1.7 The characteristics and function of the proteins present in bacterial *ars* operons

1.7.1 Arsenic/antimony anion-translocating proteins 22

1.7.1.1 ArsB: Membrane-associated arsenite/antimony efflux pump 24

1.7.1.2 ArsA: Arsenite/antimony-stimulated ATPase 26

1.7.2 ArsC: Arsenate reductase 28

1.7.3 The regulation of *ars* operons 32

1.7.3.1 ArsR: Primary trans-acting repressor 32

1.7.3.2 ArsD: Secondary trans-acting repressor 34

1.8 Aim of thesis 37

1.1 Introduction to biomining

The need to develop an alternative for traditional mining methods in the metal-extraction industry has increased in recent years. Mining companies have been looking for new mining methods to recover metals from ores containing low-grade deposits and to extract small quantities of metals left after traditional physical-chemical processing of high-grade mineral ores. The main reason for this is that it is not economically viable to recover metals from low concentration mineral deposits by using traditional mining processes (Rawlings and Silver, 1995). The microbial-aided decomposition and solubilization of mineral compounds is a naturally occurring phenomenon (Ehrlich, 1997). This natural ability of certain microorganisms has been successfully implemented into commercial mining processes and has subsequently had a major impact on the economical recovery of mineral values from low-grade deposits and depleted high-grade ores (Rawlings, 2002).

Biomining is a general term that comprises of both microbial-dependent bioleaching and biooxidation processes. Bioleaching is generally accepted as the conversion of an insoluble metal (usually a metal sulfide) into a soluble form (usually a metal sulfate), whereafter the metal is extracted from water. An example of this type of process is the conversion of copper-containing minerals such as covellite (CuS) or chalcocite (Cu_2S) into soluble copper sulfate. Biooxidation commonly refers to the extraction of minerals as solid, insoluble residues. During this process, microbial activity changes the ultra-structure of the mineral, thereby enhancing the accessibility of chemicals used for recovery purposes. An example of this type of process is the removal of arsenic, iron and sulfur from gold-containing arsenopyrite ores. The gold that remains in the mineral is more accessible to subsequent extraction with cyanide treatment (Brierley, 1997; Suzuki, 2001; Rawlings, 2005).

Biomining is a well established economically important biotechnological practice with distinctive advantages over traditional mining operations. Besides the fact that biomining operations are economically advantageous for the recovery of small quantities of mineral

deposits, these microbially-based processes are in general more environmentally friendly compared to traditional mining methods. Traditional mining methods consume large amounts of energy during the roasting or melting of mineral ores and may lead to the production of sulfur dioxide and other environmentally harmful gaseous emissions. Mine tailings and waste products produced as a result of traditional mining methods may be biologically leached when exposed to air or rain, resulting in unwanted acid mine drainage and metal pollution. In addition, the shorter construction time, low-cost maintenance and operational simplicity of biomining processes have further contributed to the increased implementation of biomining in the mining industry (Rawlings *et al*., 2003; Rawlings, 2005; Valenzuela *et al*., 2006).

Two main types of commercial-scale microbially-assisted mineral degrading processes are currently employed. They are namely irrigation-type and stirred-type processes. Irrigation-type leaching involves the percolation of acidic leaching solutions through crushed ore or concentrates that are placed in columns, heaps or dumps. The dump/heap is irrigated with raffinate, an iron- and sulfate-rich recycled wastewater from which the metal (e.g. copper) has been removed. A consortium of microorganisms growing on the surface of the mineral in the dump/heap will produce the ferric iron and acid that will ultimately convert insoluble copper sulfides to soluble copper sulfate. The copper sulfate-containing leach solution is removed from the base of the dump/heap, whereafter the copper is recovered by means of solvent extraction and electrowinning (Schnell, 1997). Irrigation-type processes are mainly applied to extract metals from low-grade ores that are not suited for smelting or the production of concentrates. The most extensively recovered metal with irrigation type leaching methods is copper (Rawlings *et al*., 2003). Irrigation-type leaching reactors are relatively cheap to construct and to operate, but unfortunately have some minor drawbacks. Recent developments in heap-leaching technology focus on improved inoculation and distribution of microbial species within the dump/heap, more effective oxygen diffusion and better heat and pH management (Rawlings 2002; Rohwerder *et al*., 2003; Rawlings, 2005). The use of stirred tank leaching processes greatly increases the rate and efficiency of mineral biooxidation in comparison with irrigation type processes. Stirred-tank leaching processes employ a

series of highly aerated, continuous-flow bioreactors to recover minerals of interest. Finely ground mineral or concentrate is added to the first tank together with dissolved ammonia- and phosphate-containing fertilizers. The mineral suspension subsequently flows through a series of pH and temperature controlled bioreactors in which the mineral leaching occurs. Since the oxidation of minerals is an exothermic process, elevated temperature levels may develop in bioreactors where mineral decomposition is rapid. In order to maintain favorable microbial growth conditions, temperature levels are regulated with large volumes of air being blown through each bioreactor. Additional cooling mechanisms may also be employed. Large agitators ensure that the mineral solids remain in suspension and ensure efficient flow into the next bioreactor (Okibe and Johnson, 2004; Rawlings 2002, 2005). Although the majority of commercial-scale operations employ stirred tank bioreactors in pretreatment processes for the recovery of gold from gold containing pyrite/arsenopyrite concentrates, processes for the extraction of cobalt, copper and nickel have recently been developed (Rawlings *et al*., 2003; Briggs and Millard, 1997; Dew and Miller, 1997).

1.2 Microbial diversity in biomining environments

Although irrigation-type and stirred tank-type processes have considerable differences, one feature they have in common is that neither is conducted under sterile conditions. Unlike other commercial fermentation processes, no effort is made to maintain a sterile setting, as the environment in which the consortium of biomining microorganisms operate is inhospitable to most other organisms. An additional reason for this is that continuous selection of microorganisms that oxidize minerals more efficiently will create optimized microbial populations (Rawlings, 2002). In general, the types of organisms present in irrigation-type operations are similar to those found in stirred-tank operations, although the proportions of the microbes is dependent on the mineral being decomposed and the conditions under which the different operations are conducted (Rawlings, 2005).

Microbial biomining communities are composed of a vast variety of microorganisms which participate in a complex system of microbial interactions and nutrient flow

processes. The different types of microorganisms that have been isolated from commercial biomining operations share several physiological features. They are all chemolithoautotrophic and have the ability to use ferrous iron and/or reduced inorganic sulfur compounds as electron donors. The oxidation of sulfur compounds results in subsequent sulfuric acid production in the environment. Therefore these microorganisms are all acidophilic and capable of growing in low pH (pH 1.4-2.0) surroundings. They grow autotrophically by fixing CO_2 and primarily prefer to use O_2 as an electron accepter. An additional feature these microorganisms have in common is that they harbor heavy metal resistance mechanisms that enable them to be remarkably tolerant to a wide range of metal ions (Krebs *et al*.,1997; Norris, 1997; Rawlings 2005).

Because biomining processes are carried out across several temperature gradients, microorganisms can be divided into different groups on the basis of their optimum temperature of growth. Three groups have been recognized: mesophiles (T_{opt} at 20°C-40°C), moderate thermophiles (T_{opt} at 40°C-60°C) and extreme thermophiles (T_{opt} > 60°C) (Johnson, 1998). In biomining processes that operate at 40°C or less, the most prominent microorganisms are considered to be a consortium of Gram-negative γ-proteobacteria. They include the iron- and sulfur-oxidizing *Acidithiobacillus ferrooxidans*, the sulfur-oxidizing *Acidithiobacillus caldus* and the iron-oxidizing leptospirilli, *Leptospirillum ferrooxidans* and *Leptospirillum ferriphilum* (Goebel and Stackebrandt, 1994; Norris, 1997; Rawlings, 1997; Coram and Rawlings, 2002). Types of moderately thermophilic microorganisms that have been isolated from operational stirred-tank processes include several *At. caldus*-like and *Leptospirillum*-like species, iron- and sulfur-oxidizing eubacteria representative of the Gram-positive genera *Acidimicrobium*, *Alicyclobacillus, Ferromicrobium* and *Sulfobacillus* (Clark and Norris, 1996; Johnson and Roberto, 1997; Norris, 1997; Okibe and Johnson, 2004) as well as several members of the archaeal genus *Ferroplasma* (Edwards *et al*., 2000; Golyshina *et al.*, 2003). Biomining consortia operating at temperatures > 60°C are dominated by iron and sulfur-oxidizing species of the archaeal genera *Acidianus*, *Sulfolobus* and *Metallosphaera* (Fuchs *et al*., 1995; Norris, 1997; Norris *et al.*, 2000; Rawlings, 2005; Mikkelsen *et al*., 2006).

The diversity of moderate thermophilic and extreme thermophilic microorganisms in commercial biomining processes has been less well documented in the past, as for many years only mesophilic bacterial species were considered to be important. Conducting bio-oxidation processes at elevated temperatures (> 40°C) has several substantial benefits over biomining processes occurring in the vicinity of 40°C. One significant advantage is that the biochemical processes responsible for the decomposition of minerals occur at higher rates in surroundings with elevated temperature levels (Okibe *et al*., 2003). As previously mentioned, the bio-oxidation of minerals is an exothermic process which may lead to an increase in temperature levels within bioreactors, creating unfavorable growth conditions for mesophilic bacteria. Performing bio-oxidation processes with thermophilic and extreme thermophilic microorganisms will be more economical, as the costs of cooling mechanisms used to regulate the temperature fluctuations caused by exothermic processes will be reduced (Okibe and Johnson, 2004). Furthermore, several minerals are more efficiently recovered at higher temperatures. The extraction of copper from chalcopyrite is the most notable example (Norris *et al*., 2000).

The discovery of moderate thermophilic and extreme thermophilic microorganisms with potential metal leaching abilities suitable for use in biomining processes is rapidly growing. This is partly because of an increase in the number of environments (similar to those of commercial biomining conditions) being screened, partly because of an increase in the vast variety of minerals being tested, and most importantly, because of new immunological (immunofluorescence and dot immunoassays) and molecular (DNA-DNA hybridization, PCR amplification and sequencing of 16S rDNA, pulsed-field gel electrophoresis (PFGE) and fluorescence *in situ* hybridization (FISH)) techniques being implemented to screen for the presence of suitable candidates (Brierley and Brierley, 1997). Bacteria representative of genus *Sulfobacillus* could have considerable potential for use in commercial bio-oxidation of mineral ores and concentrates at elevated temperatures. They have been identified and isolated from a range of thermal acidic environments, such as geothermal areas (Brierly *et al*., 1978; Ghauri and Johnson, 1991; Atkinson *et al*., 2000), self-heating mineral ores and spoil dumps (Golovacheva and Karavaiko, 1978; Marsh and Norris, 1983; Vartanyan *et al*., 1986; Robertson *et al*., 2002;

Kinnunen *et al*., 2003), commercial bio-mining operations (Dopson and Lindström, 2003; Okibe *et al*., 2003) and environments with acid mine drainage (Brierley and Brierley, 1997; Baker and Banfield, 2003).

1.3 Characteristics of genus *Sulfobacillus*

Members of the genus *Sulfobacillus* fall within the low G+C Gram-positive division of the bacterial firmicutes lineage (Baker and Banfield, 2003). The genus *Sulfobacillus* includes Gram-positive, spore-forming, non-motile acidophilic moderate thermophiles with a growth temperature optimum of 40°C-60°C. They have a highly versatile metabolism and can grow autotrophically (utilizing ferrous iron, sulfide-containing mineral compounds and reduced inorganic sulfur as sole energy sources), heterotrophically (utilizing organic carbon and energy sources such as glucose, casein hydrolysate and yeast extract) and mixotrophically (simultaneously using organic and inorganic substances as sources of energy and carbon). Optimal growth of *Sulfobacillus* spp. occurs in mixotrophic conditions where reduced sulfur compounds, in inorganic forms (e.g. tetrathionate or pyrite) or organic forms (e.g. cysteine), and CO_2 together with glucose or yeast extract are utilized as sources of energy and carbon, respectively. Furthermore, *Sulfobacillus* spp. are facultative anaerobes, using ferric iron as an electron acceptor in the absence of oxygen (Bridge and Johnson, 1998; Hallberg and Johnson, 2001; Rawlings, 2002; Yahya and Johnson, 2002). Another distinctive feature of genus *Sulfobacillus* is the unique fatty acid composition of the lipids comprising their membranes. Sulfobacilli membranes contain branched chain, anteiso fatty acids, distinguishing them from the majority of living organisms which produce straight-line saturated and unsaturated fatty acids using short-chain acetyl-CoA esters as primers and malonyl-CoA for chain elongation. This characteristic of Sulfobacilli makes them a member of an exclusive group of bacteria whose membranes consist of branched and alicyclic fatty acids. This group of bacteria comprises only 10% of known bacterial species and characteristically use branched short-chain carboxylic fatty acids to synthesize higher-branch-chain fatty acids in lipid production. Furthermore, Sulfobacilli are also capable of synthesizing ω-cyclohexyl-α-oxyundecanoic fatty acids, a

phenomenon previously only detected in the acidothermophilic *Alicyclobacillus acidocaldarius* and *A. acidoterrestris* and the mesophilic *Curtobacterium pusillum* (Oshima and Ariga, 1975; Suzuki *et al*., 1981; Kaneda, 1991; Tsaplina *et al*., 1994).

Only two species of genus *Sulfobacillus* were initially recognized. Strains *Sb. thermosulfidooxidans* and *Sb. acidophilus* could be distinguished from each other by using a combined approach of detecting differences in their physiological characteristics, differences in their growth rates, differences in cell biomass yields during heterotrophic growth conditions and by comparing their ability to grow autotrophically in the presence of iron and sulfur (Norris *et al*., 1996). The genomic DNA of *Sb. thermosulfidooxidans* and *S. acidophilus* has a guanine-cytosine content (mol% G+C) of 48-50 and 55-57, respectively. Several other moderate thermophiles with *Sulfobacillus*-like characteristics have been isolated in the recent past and only some of them could be distinguished from the two named *Sulfobacillus* spp. on the basis of their mol% G+C content (Norris, 1997). Recently, two other species, *Sulfobacillus sibiricus* (Melamud *et al*., 2003) and *Sulfobacillus thermotolurans* (Bogdanova *et al*., 2006) were proposed and validated. Molecular techniques based on microbial genotype analysis facilitate a more effective approach to accurate strain identification and taxonomical classification and have subsequently contributed to the recognition of previously unclassified *Sulfobacillus* spp. Johnson and coworkers (Johnson *et al*., 2005) have shown that amplified ribosomal DNA restriction enzyme analysis (ARDREA) can be implemented to successfully distinguish between moderately thermophilc *Sulfobacillus*-like isolates at a species level. Information obtained from this highly reliable procedure, indicated that the tested *Sulfobacillus*-like isolates could conveniently be divided into two major subgroups based on differences in patterns after electrophoretic separation of digested amplified ribosomal DNA fragments. Johnson *et al*. proposed that *Sulfobacillus*-like isolates could be divided into *Sulfobacillus* sub-group I, containing *Sb. thermosulfidooxidans/Sb. montserratensis*-like isolates, and *Sulfobacillus* sub-group II, containing *Sb. acidophilus*/*Sb yellowstonensis*-like isolates. The prospect for the application of ARDREA as a tool to identify and discriminate between newly discovered *Sulfobacillus*-like isolates could be of considerable industrial importance in the near future.

1.4 The chemical and biological properties of arsenic

The name Arsenic is derived from the Greek word *arsenikon*, which means "yellow orpiment". The isolation of arsenic from arsenic containing compounds was first reported by Albertus Magnus in 1250 A.D. Arsenic is the 33rd element on the periodic table and shares chemical properties with other group V elements phosphorous (P) and antimony (Sb). Arsenic is classified as a metalloid or semi-metal, as it exhibits both metallic and non-metallic characteristics. Arsenic is widely distributed in natural environments. It is usually associated with metal containing ores in the form of arsenopyrite (FeAsS), but low concentrations of arsenic may also be found in the earth's atmosphere and water. Arsenic can be stable in the environment in any of four oxidation states: arsine (As (-III)), metallic (As (0)), arsenate (As(V)) and arsenite (As (III)). These oxidation states of arsenic are interconvertable and the speciation between them is mainly dependent on the redox condition and pH of the environment. Arsenite (As(III)) is considerably more toxic than arsenate (As (V)) (Knowles and Benson, 1983). Arsenate (As (V) as $H_2AsO_4^-$ ($2.5<pH<7$) and $HAsO_4^{2-}$ ($7<pH<12$)) occurs as the predominant form of arsenic in aqueous aerobic environments, whereas arsenite (As (III) as H_3AsO_3 ($0<pH<10$) and $H_2AsO_3^-$ ($10<pH<12$)) will be present in higher concentrations in anoxic environments. Elemental arsenic and gaseous arsine will rarely be encountered in nature (Inskeep *et al*., 2002). The fate of arsenic during spontaneous microbial oxidation of arsenopyrite in industrial biomining operations is of considerable interest. The microbially catalyzed oxidation of arsenopyrite produces dissolved Fe(II), arsenic as As(III) and sulfur as either S(VI) or S(0). Subsequent oxidation of Fe(II) to Fe(III) and S(0) to S(VI) is facilitated by leaching microorganisms present in the leaching solution. As(III) is then further oxidized to As(V) by oxygen, Fe(III), or other medium components (chemicals, metabolites or biomass components). These reactions are constantly in competition with each other and are strongly influenced by the availability of Fe(III) and the concentration and oxidation state of arsenic present in the leaching solution (Pol'kin *et al*., 1975; Shrestha, 1988; Barrett *et al*., 1993; Panin *et al*., 1993; Breed *et al*., 1996). The oxidation of arsenopyrite produces mixtures of iron-containing precipitating compounds like ferric arsenate ($FeAsO_4$) and jarosite ($KFe_3(SO_4)_2(OH)_6$). The formation

of these compounds is dependent on the extent of FeAsS oxidation that has occurred, the type of leaching organisms, pH, temperature and the ionic composition of the leaching solution (Mandl *et al*., 1992; Tuovinen *et al.,* 1994).

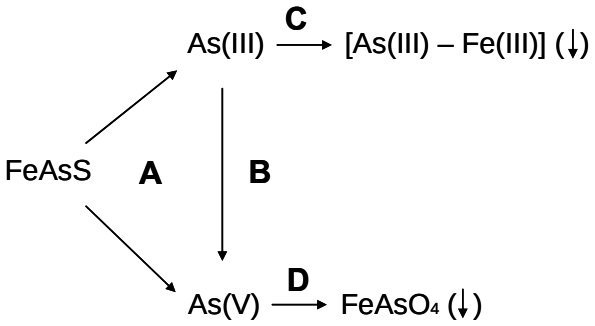


Figure 1.1: **The migration of arsenic during bacterial oxidation of arsenopyrite. (A)** the activity of acidophilic chemolithotrophic leaching bacteria; **(B)** As(III) may be further oxidized to As(V) by oxygen, Fe(III) or other medium components; **(C)** the pH-dependent adsorption of As(III) by the formation of an iron-containing precipitating compound; **(D)** As(V) is transferred to $FeAsO_4$ in a pH dependent reaction, where-after it precipitates. Symbol (↓) indicates precipitation. Adapted from Mandl *et al*., 1992.

The poisonous properties of arsenic have been known for centuries (Azcue and Nriagu, 1994). Due to the poisonous nature of this element, it has been extensively used for agricultural (herbicides, pesticides, insecticides, fungicides, wood preservatives and vine killer), industrial (manufacture of glassware), medical (treatment of some forms of leukemia and myelomas (Roboz *et al*., 2000)) and toxicological purposes. In 1908 the Nobel Prize in medicine was awarded to Paul Ehrlich for the discovery of the arsenical Apräparat 606 compound, "Salvarsan", which is used for the treatment of syphilis and sleeping sickness (Silver *et al*., 2002). The toxicity of arsenic to microorganisms is primarily due to its ability to act as a soft metal ion, forming strong bonds with reactive thiolates of cysteine residues and imidazolium nitrogens of histidine residues present in proteins. If these residues are located within the active sites of vital enzymes, binding of arsenite will cause changes in the conformation of that enzyme and will ultimately inhibit catalytic or biological activity (Oremland and Stolz, 2003; Rosen, 2002). Arsenate is a molecular analogue of phosphate and may interfere with the cellular uptake of phosphate. Uptake of phosphate and arsenate into cells is facilitated by the *Pit* and *Pst* phosphate

transport systems. Furthermore, arsenate has the ability to inhibit oxidative phosphorylation and other cellular processes that involve phosphate (Tamaki and Frankenberger, 1992).

1.5 Arsenic resistance mechanisms in microorganisms

Microorganisms require the presence of certain transition metals, heavy metals and metalloids to perform important biochemical functions. Both essential (calcium, cobalt, chromium, copper, iron, potassium, magnesium, manganese, sodium, nickel and zinc) and nonessential (aluminum, arsenic, cadmium, lead, mercury and silver) metals are toxic to microorganisms at elevated concentrations. High levels of nonessential heavy metals and metalloids are increasingly found in microbial habitats, due to natural and industrial processes. Microorganisms have therefore evolved different mechanisms to tolerate high levels of heavy metals and metalloids in their immediate environment, providing them with a competitive selective advantage. Microorganisms may possess one or a combination of six different metal resistance mechanisms: (1) efflux of the toxic metal out of the cell; (2) enzymatic detoxification; (3) exclusion by a impermeable barrier; (4,5) intra- or extra-cellular sequestration and (6) reduction in the sensitivity of cellular targets to the metal (Bruins *et al*., 2000; Dopson *et al*., 2003). Microorganisms have evolved different resistance mechanisms to tolerate the harmful effects of arsenical compounds. Reported resistance mechanisms include the conversion of stable arsenic compounds to gaseous species (archaea, bacteria), methylation of arsenic or arsenate (archaea, bacteria) (Bentley and Chasteen, 2002), oxidation of arsenite to less toxic arsenate (e.g. *Alcaligenes faecalis*) (Anderson *et al*., 1992), selecting phosphate uptake pathways that do not transport arsenate effectively (e.g. the cyanobacterium, *Anabaena variabilis*), the over production of intracellular thiols (e.g. the protozoan, *Leishmania*) and sequestration in vacuoles (e.g. fungi) (Cervantes *et al*., 1994; Rosen, 1999; Stolz *et al*., 2002; Tamaki and Frankenberger, 1992). The best characterised, and probably the most widespread arsenic detoxification system in microorganisms, is a mechanism whereby intracellular arsenate is converted to arsenite and extruded out of the cell via carrier-mediated membrane transport proteins. This system is controlled by a cluster of genes, called the

arsenic resistance (*ars*) operon (Cervantes *et al*., 1994; Rosen,1999; Mukhopadhyay *et al*., 2002; Silver *et al*., 2002).

1.5.1 The molecular genetics of efflux systems involved in bacterial arsenic resistance

Although variation in components comprising bacterial efflux systems may exist in different bacterial species, common themes are (1) cellular uptake of As(V) and As(III); (2) the reduction of As(V) to As(III) by arsenate reductases; and (3) the extrusion of As(III) out of the cell. The transport of arsenate into bacterial cells is mediated by the Pit and Pst phosphate transport systems. During periods of phosphate abundance, arsenate will enter cells by means of the constitutively expressed, nonspecific Pit system. During times of phosphate starvation, the carefully regulated, more specific Pst system is induced (Nies and Silver, 1995). The Pst system discriminates between phosphate and arsenate 100 times better than the Pit system. It has been reported that the inactivation of the Pit system in favor of the Pst system may lead to greater arsenate resistance in microorganisms (Willsky and Malamy, 1980; Cervantes *et al*., 1994). The single-gene product Pit system relies on proton motive force for phosphate/arsenate transport, while the multi-component Pst system uses ATP-hydrolysis to facilitate phosphate/arsenate translocation (Silver and Walderhaug, 1992). Uptake of arsenite into bacterial cells is probably facilitated by glycerol transport proteins. Meng *et al*., (2004) recently showed that a aqua-glyceroporin (GlpF) mediates transport of arsenite and antimony into *Escherichia coli*. Members of the aqua-glyceroporin family are multifunctional channels that are responsible for the transport of neutral organic substances such as glycerol and urea (Mukhopadhyay *et al*., 2002; Rosen, 2002). Cytosolic arsenate is then reduced to arsenite by the product of the *arsC* gene, arsenate reductase. Reduction of arsenate is facilitated by a pathway consisting of a cascade of metabolic intermediates, with reduction of either thioredoxin (Trx) or glutaredoxin/glutathione (Grx/GSH) supplying the initial energy for the process. Arsenite is subsequently extruded from the cell by two basic transport systems: membrane potential-driven transporters (e.g. ArsB and AseA (both present in the case of *Bacillus subtilis*) and ArsM) or by As(III)-translocating

ATPase transporters (e.g. ArsAB). The characteristics and functions of the genes comprising the *ars* operon will be discussed in detail later.

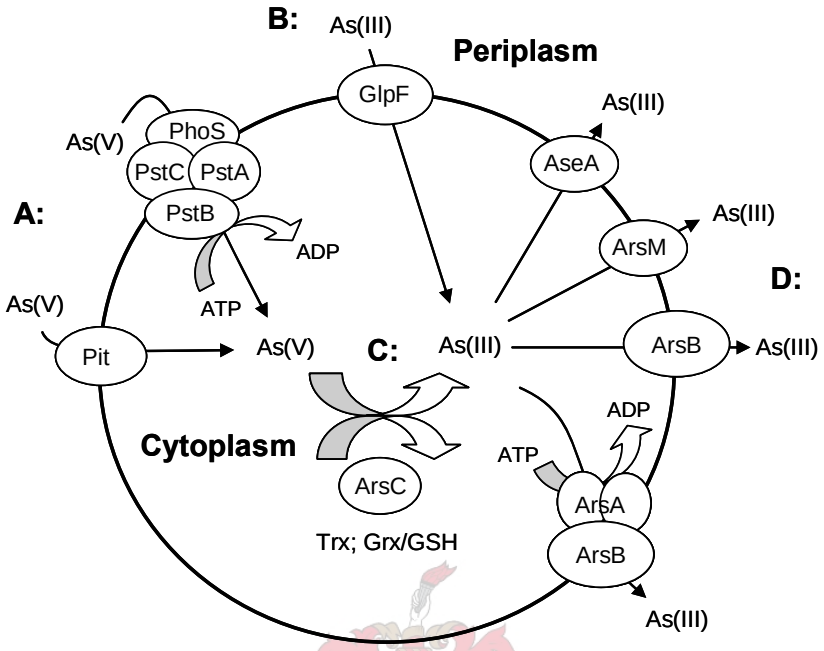


Figure 1.2: **Arsenical transport and detoxification pathways in prokaryotes. (A)** Arsenate uptake in *E. coli* is facilitated by two phosphate transport systems: the membrane potential-coupled Pit phosphate uptake system and the multi-component ATP-coupled Pst phosphate uptake system. In the case of gram-positive bacteria, it has been hypothesized that phosphate and arsenate enter the cytoplasm by means of two similar membrane transporters (Silver *et al*., 1981). **(B)** Arsenite transport to the cytoplasm of *E. coli* is mediated by the aqua-glycoprotein channel, GlpF (Meng *et al.*, 2004). **(C)** Once inside the cytoplasmic space, arsenate is reduced to arsenite by ArsC, using either the Trx or the Grx/GSH coupled pathway. **(D)** Arsenite is then extruded from the cells by two types of arsenite transporters: the membrane potential-driven transporters ArsB, AseA (both present in the case of *Bacillus subtilis*) and ArsM or the As(III)-translocating ATPase ArsAB transporter.

The toxicity of arsenic to micro-organisms consequently dependents on several endogenous factors (e.g. the presence of genes capable of encoding membrane-associated oxyanion uptake and efflux pumps) and exogenous factors (e.g. the redox potential and pH of the environment influence the oxidation state and mobility of arsenic) (Silver *et al*.,

2002). Although microorganisms employ a number of mechanisms to cope with arsenic toxicity, several bacteria that benefit from the presence of arsenic have recently been discovered. Bacteria classified as dissimilatory arsenate reducers have the ability to utilize arsenate as a thermal electron acceptor in anaerobic respiration, while some other bacteria are capable of using arsenite as the electron donor for chemoautotrophic growth (Jackson *et al*., 2003).

1.5.2 The general structure of bacterial arsenic resistance (*ars*) operons

The presence of arsenical compounds in the environment selects and maintains microbes possessing genetic determinants which confer resistance to arsenic (*ars* genes). The *ars* genes are widely distributed in microorganisms and are usually located on the chromosome, plasmids or transposable elements. Although the number of the genes and the gene layout within the *ars* operon varies, two of the most commonly encountered forms consist of five genes (*arsRDABC*) and three genes (*arsRBC*), respectively. The five gene *ars* operon, located on plasmid R773 of *E. coli*, initially described by Hedges and Baumberg in 1973, is the most thoroughly studied *ars* system and has to date only been found on plasmids of Gram-negative bacteria (R773 and R46 of *E. coli* and pKW301 of *Acidiphilium multiforum*) (Chen *et al*., 1985; Bruhn *et al*., 1996; Suzuki *et al*., 1998). The three gene *ars* operon, discovered by Novick and Roth in 1968, is found on plasmids of Gram-positive bacteria (*Staphylococcus aureus* (pI258); *Staphylococcus xylosus* (pSX267)) (Novick and Roth, 1968; Ji and Silver, 1992b; Rosenstein *et al*., 1992) and on the chromosomes of the Gram-negative bacteria *E. coli*, *Pseudomonas aeroginosa* and *Pseudomonas fluorescens* (Carlin *et al*., 1995; Cai *et al*., 1998; Prithivirajsingh *et al*., 2001). The *ars* operons are transcribed from a single operator/promoter region. The *arsR*, *arsB* and *arsC* gene products of both operons have similarities in sequence and function (Figure 1.3). The *arsR* and *arsD* genes encode *trans*-acting regulatory proteins. ArsR is an inducer-sensitive transcriptional repressor and controls the basal level of *ars* operon expression, while ArsD is an inducer-independent transcriptional repressor and is thought to regulate the upper level of *ars* operon expression. The *arsB* encodes a membrane-associated arsenite/antimony efflux pump that uses the difference in

membrane potential to extrude arsenite/antimony from the cell. ArsB can physically associate with the gene product of *arsA*, an arsenite/antimony-stimulated ATPase, to mediate efflux of arsenite/antimony in an ATP-dependent process. The ArsAB ATPase complex is much more efficient at arsenite extrusion than ArsB alone. The final gene, *arsC*, encodes an arsenate reductase that converts intracellular arsenate to arsenite, which in turn acts as the substrate for the ATP-hydrolysis to facilitate arsenite export from the cell (Cervantes *et al*., 1994; Silver, 1996).

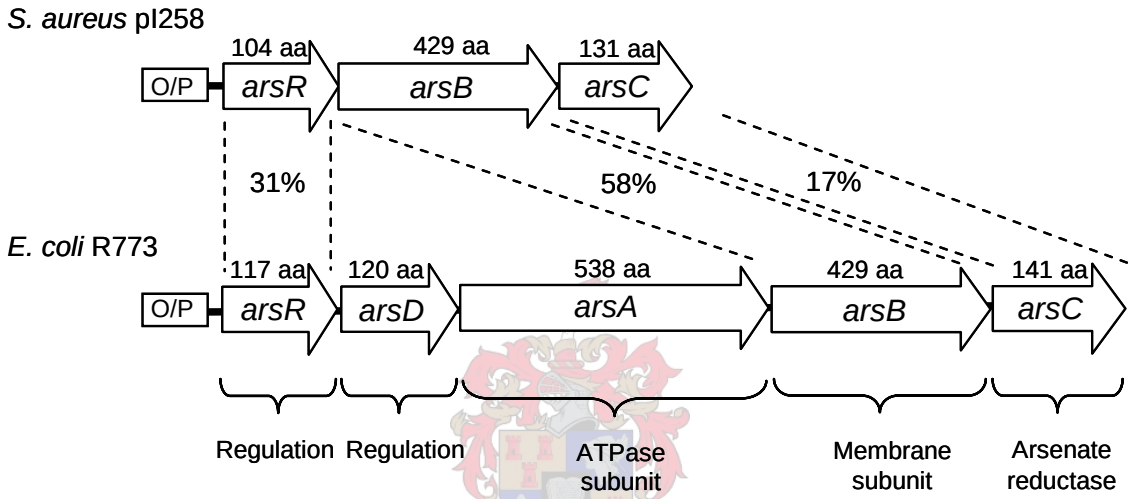


Figure 1.3: **The genes and products of the arsenic resistance (*ars*) operons of *S. aureus* plasmid pI258 and *E. coli* plasmid R773.** The alignment of the arsenic resistance genes (arrows) with amino acid (aa) sizes of predicted product sizes (above genes). The percent identities between the aa products are indicated and the functions of the arsenic resistance genes are shown below. Both the *arsRBC* and *arsRDABC* operons are transcribed from a single operator/promoter site. O/P indicates the putative operator/promoter region. Adapted from Silver, 1996.

1.6 Variations in the structure of bacterial *ars* operons

Recent advances in the development of molecular techniques and computer software programs have contributed to the generation of substantial volumes of sequencing information from genome sequencing projects. It has become apparent that *ars* gene homologues are widely distributed in *Bacteria*, *Archaea* and also in some *Eukarya*, indicating that the *ars* operon system is a ubiquitous mechanism by which microorganisms obtain resistance to the toxic effects of arsenicals. Sequence information has revealed that the layout and transcription of the *ars* genes may differ from the conventional five and three gene *ars* operons. The *ars* operons that have been molecularly characterised are listed in Table 1.1.

Table 1.1: **Bacterial arsenic resistance operons (*ars* operons) that have been molecularly characterised**

Organism	Gram staining	Operon location	Operon structure	Reference	Accession No.
Bacteria:					
Acidophilium multivorum	-	plasmid, pKW301	*arsRDABC*	Suzuki *et al*., 1998	AB004659
Acidithiobacillus caldus	-	chromosome	*arsCRB*	Kotze *et al*., 2006	DQ810790
Acidithiobacillus caldus	-	transposon, Tn*AtcArs*	*arsRCDADA (orf7)(*CBS*)B*	Tuffin *et al.,* 2004	AY821803
Acidithiobacillus ferrooxidans	-	chromosome	*arsCRBH*	Butcher *et al*., 2000	AF173880
Bacillus subtilis	+	SKIN element	*arsR(yqcK)BC*	Sato and Kobayashi, 1998	D84432
Bacillus subtilis	+	chromosome	*aseRA*	Moore *et al*., 2005	*NC_000964
Chromobacterium violaceum	-	chromosome	*arsRBC*	Carepo *et al*., 2004	*NC_005085
Escherichia coli	-	chromosome	*arsRBC*	Carlin *et al*., 1995	X80057

Escherichia coli	-	plasmid, R773	*arsRDABC*	Chen *et al*., 1985	J02591
Escherichia coli	-	plasmid, pR46	*arsRDABC*	Bruhn *et al*., 1996	U38947
Lactobacillus plantarum	+	Plasmid, pWCFS103	*arsRDAB; D2*	Van Kranenburg *et al*., 2005	CR377166
Leptospirillum ferriphilum	-	chromosome	*arsRC*(fused*)B*	Tuffin *et al*., 2006	[#]N/S
Leptospirillum ferriphilum	-	transposon, Tn*LfArs*	*arsRCDA(*CBS*)B*	Tuffin *et al*., 2006	DQ057986
Pseudomonas aeruginosa	-	chromosome	*arsRBC*	Cai *et al*., 1998	AF010234
Pseudomonas fluorescens	-	chromosome	*arsRBC*	Prithivirajsingh *et al*., 2001	AF047036
Serratia marcescens	-	plasmid, pR478	*arsRBCH*	Ryan and Colleran, 2002	AJ288983
Staphylococcus aureus	+	plasmid, pI258	*arsRBC*	Ji and Silver, 1992b	M86824
Staphylococcus xylosus	+	plasmid, pSX267	*arsRBC*	Rosenstein *et al*., 1992	M80565
Streptomyces sp. Strain FR-008	+	linear plasmid, pHZ227	*arsRBOCT*	Wang, 2006	DQ231520
Synechocystis sp. Strain PCC6803	-	chromosome	*arsBHC; H*	López-Maury *et al*., 2003	*BA000022.2
Yersiniae enterocolitica	-	plasmid, pYV	*arsHRBC*	Neyt *et al*., 1997	U58366
Archae:					
Ferroplasma acidermanus		chromosome	*arsRB*	Gihring *et al*., 2003	*NZ_AABC04 000026
Halobacterium sp. strain NRC-1		plasmid, pNRC100	*arsMR2; arsADRC*	Ng *et al*., 1998	AF016485

[#] N/S: the nucleotide sequence number of the *ars* operon has not been submitted to the GenBank database.

* the specific accession number of the *ars* operon sequence is not available, but the genome sequence accession number of the referred microorganism is being presented.

Several atypical genes have been found to be associated with *ars* operons. Neyt and coworkers (Neyt *et al*., 1997) reported the presence of a novel gene, *arsH*, on plasmid pYV of *Yersiniae enterocolitica*. It was shown to be divergently transcribed from an *arsRBC* operon. The expression of *arsH*, either in *cis* or in *trans*, was essential to confer resistance to arsenic in *Y. enterocolitica*. The influence of *arsH* on arsenic resistance in *Y. enterocolitica* is surprising, as the *arsRBC* operon of *Y. enterocolitica* alone is sufficient enough to confer arsenic resistance in *E. coli* and staphylococci. The *arsH* of pYV, encoding a 26.4 kDa protein, has no ATP-binding motif, no hydrophobic domain nor any other recognizable motif or domain. The *arsH* of pYT shows 82% amino acid sequence identity to a putative *arsH* present on the IncH12 plasmid, R478. This 272 kb plasmid was originally isolated from *Serratia marcescens* and sequence analysis revealed the presence of an *ars* operon with a layout similar to that of the *ars* operon found on pYV. The removal of the R478 *arsH* subsequently resulted in total loss of resistance to arsenic. It has been suggested that ArsH may be involved in a secondary regulation control mechanism, either for the entire *ars* operon, acting as a putative binding site for another regulatory protein, or for the control of expression of a particular gene such as *arsB* (Ryan and Colleran, 2002). However, two recently identified *arsH* homologues, located on the chromosomes of *Acidithiobacillus ferrooxidans* and cyanobacterium *Synechocystis* sp. Strain PCC 6803, do not seem to play an essential role in conferring resistance to arsenic. *At. ferrooxidans* contains an *ars* operon consisting of an *arsCRBH*, with *arsCR* and *arsBH* being divergently transcribed with respect to each other. The role of *arsH* expression in arsenic resistance in *At. ferrooxidans* is unclear as the expression of *arsH* in an *E. coli* host was not required to confer resistance to arsenic (Butcher *et al*., 2000). The *ars* operon of cyanobacterium *Synechocystis* sp. Strain PCC 6803 consists of a co-transcribed *arsBHC* operon and is regulated by a separately transcribed *arsR* homologue (sll1957). As in the case of *At. ferrooxidans,* the expression of *arsH* was not required for arsenic tolerance (López-Maury *et al*., 2003).

The sequence of the archaeon, *Halobacterium* sp. Strain NRC-1 megaplasmid pNRC100 revealed the presence of a unique cluster composed of several *ars* gene homologues. An operon encoding an *arsR2* and a putative methyltransferase (*arsM*) was identified. The

two genes are co-transcribed and the expression of *arsR2* seems to be constitutive. Deletion of the *arsM* resulted in sensitivity to arsenite. It is believed that *arsM* is responsible for the transfer of methyl groups from S-adenosylmethionine (Adomet) to intracellular As(III), resulting in the formation of mono-, di- and tri-methylated arsenic species. The methylated arsenic species will then subsequently move down the created concentration gradient, to the outside of the cell (Wang *et al*., 2004; Qin *et al*., 2006). This will ultimately lead to a decrease in intracellular arsenite, indicating a novel mechanism of arsenic resistance involving a putative As(III) S-adenosylmethyltransferase. Another putative *arsM* homologue is present in *Rhodopseudomonas palustris*. The expression of this *arsM* in an arsenic sensitive *E. coli* mutant resulted in increased arsenic resistance (Qin *et al*., 2006).

Another set of novel genes has been associated with the *ars* operon found on the linear plasmid pHZ227, originally isolated from *Streptomyces* sp. strain FR-008. The *ars* operon of pHZ277 is arranged in unusual configuration, with *arsR1*, *arsB* and *arsO* constituting one operon and *arsC*, together with *arsT*, the other. The two *ars* gene clusters are divergently transcribed. Deletion of the *ars* gene cluster of pHZ227 in *Streptomyces* sp. strain FR-008 resulted in sensitivity to arsenic. The addition of the *ars* gene cluster into arsenic-sensitive Streptomyces hosts resulted in increased tolerance to arsenicals. A construct containing only *arsR1*, *arsB* and *arsC*, was sufficient to confer resistance to arsenate and arsenite, suggesting that the absence of *arsO* (encoding a putative flavin-binding monooxygenase) and *arsT* (encoding a putative thioredoxin reductase) did not significantly effect arsenic resistance in *Streptomyces* sp. strain FR-008. In one exceptional case, it was reported that the expression of *arsT* was required for arsenate resistance in the mutant strain *Streptomyces lividans* TK24 (Wang *et al*., 2006).

An interesting variation of the *ars* operon was isolated from the legume symbiont, *Sinorhizobium meliloti*. This *ars* operon contains four genes: *arsR*, the *arsB* is replaced by a gene encoding an aqua-glyceroporin-like channel (*aqpS*), *arsC* and *arsH*. The presence of AqpS in this operon is interesting, since aquaglyceroporin-like channels are usually associated with transport of arsenic and antimonite into cells. Disruption of *aqpS*

showed an increase in arsenite resistance but not arsenate resistance, while disruption of *arsC* showed increased arsenate sensitivity. AqpS and ArsC together, contribute to a novel arsenate detoxification pathway. This mechanism implies that intracellular arsenate is converted to arsenite by ArsC, leading to a concentration gradient of arsenite in the cell relative to the outside of the cell. This will ultimately lead to downhill transport of arsenic to the outside of the cell through the AqpS channels. (Yang *et al*., 2005).

The *ars* operon of *Bacillus subtilis*, located on the *skin* (sigK insertion) element, shows homology to the conventional three gene *ars* operon, but has an additional gene (*orf2*) located between the *arsR* and *arsB* genes (Sato and Kobayashi, 1998). The *orf2*, renamed as *yqcK* (Moore and Helmann, 2005), shows 32% homology to a cadmium-inducible gene (*cadI*) situated in front of a putative *arsRBC* operon of *Mycobacterium tuberculosis*. It has been reported that *cadI* homologous are located adjacent to or in *ars* operons and may be involved in arsenic and cadmium detoxification reactions (Hotter *et al*., 2001). The contribution of *yqcK* to arsenical resistance in *B. subtilis* is still unclear. *B. subtilis* is also host to another chromosomally located *ars* operon, containing two typical bacterial *arsRB* homologues, called *aseR* and *aseA* (formerly known as *ydeT* and *ydfA*, respectively). An unlinked *arsC* gene *(yusI*) transcribes an ArsC-related protein. Experiments performed by Moore *et al.*, (2005) established that expression of AseA on its own, has no significant contribution to arsenic resistance in *B. subtilis*.

Analysis of the complete genome of the archaeon, *Ferroplasma acidarmanus*, revealed another putative two gene operon that shows homology to *arsR* and *arsB* respectively. An *arsA*-like gene was also identified, but was located apart from the putative *arsRB* operon. No *arsC* or *arsD* homologues were present in the chromosome of *F. acidarmanus.* The fact that *F. acidarmanus* shows resistance to high levels of both arsenate and arsenite, together with the absence of both *arsC* and genes encoding phosphate transport systems, is very intriguing. This strongly suggests that *F. acidarmanus* contains an atypical arsenic resistance pathway, yet to be discovered (Gihring *et al*., 2003).

Plasmid pWCFS103 of *Lactobacillus plantarum* harbors an atypical *ars* gene cluster. The cluster consists of co-transcribed *arsRD1AB* and another *arsD2*-like gene expressed on its own. The pWCFS103 *ars* gene cluster conferred resistance to arsenite and arsenate, although no *arsC* was present. The *arsC* appeared to be present on the chromosome of *L. plantarum* WCFS1. This makes the layout of the *L. plantarum ars* gene cluster unique compared to those of other bacteria where the *arsC* is typically associated with *ars* operons (Van Kranenburg *et al*., 2005).

Fused *ars* gene homologues within *ars* operons have also been reported. The chromosomally located *ars* operon of *M. tuberculosis* contains an a*rsB*/*arsC* fused into one continuous open reading frame, encoding a 498-residue hypothetical polypeptide. No functional analysis of this protein has been published (Hotter *et al*., 2001; Sato and Kobayashi, 1998). In another case, the *ars* operon (*Lfars*) present on the chromosome of *Leptospirillum ferriphilum* contains a fused *arsR/arsC* gene preceding an *arsB* homologue. *Lfars* conferred poor resistance to arsenate and arsenite in both *E. coli* and *L. ferriphilum*, probably due to poor promoter expression and regulation of *Lfars* (Tuffin *et al*., 2006).

High levels of arsenate and arsenite resistance in *L. ferriphilum* are conferred by a second *ars* operon (Tn*LfArs*) situated on a transposable element. Tn*21*-like *tnpA* (transposase) and *tnpR* (resolvase) genes flank a group of *ars* genes with an unusual layout (ars*RCDA*(a gene coding a CBS-domain-containing protein)*B*) (Tuffin *et al*., 2006). Tn*LfArs* showed high similarity to another transposable element containing a series of *ars* genes (Tn*AtcArs*), previously isolated from a highly arsenic-resistant strain of *Acidithiobacillus caldus*. Tn*AtcArs* ars genes are also flanked by Tn*21*-like *tnpA* and *tnpR* genes, but are in the atypical order *arsRCDADA*(*orf7*)(a gene encoding a CBS domain-containing protein)*B*. The *orf7* encodes a putative NADH-like oxidoreductase. Both Tn*LfArs* and Tn*AtcArs* were transpositionally active in *E. coli* (de Groot *et al*., 2003; Tuffin *et al*., 2005).

1.7 Characteristics and function of the proteins present in bacterial *ars* operons

1.7.1 Arsenic/antimony anion-translocating proteins

Bacterial resistance to arsenical and antimonial compounds is mediated by the active extrusion of toxic oxyanions As(III) and Sb(III) from the cells. The extrusion of the metalloid oxyanions is mediated by two different energy-dependent transport mechanisms. Knowledge about the energetics of these transport mechanisms has come from *in vivo* and *in vitro* studies performed on the *ars* operon of plasmid R773 in an *unc* strain of *E. coli*, defective in the H+-translocating ATPase (F_0F_1) that catalyses the equilibrium between ATP and the electrochemical proton gradient. Unfortunately, no *unc* mutant strain of *S. aureus* is available to determine the nature of the energetics involved in the extrusion of arsenical and antimonial compounds in Gram-positive bacteria. The ArsB of *S. aureus* plasmid pI258 has superimposable hydrophatic profiles corresponding to the ArsB of *E. coli* plasmid R773, suggesting the energetics involved in transport might be similar to those detected in *unc* mutant *E. coli* cells (Dey and Rosen, 1995; Dou *et al.*, 1992).

From the aggregate of results, a model was proposed which implies that the arsenic/antimony anion-translocating system exhibits a dual mode of energy coupling which is dependent on the composition of the protein subunits comprising the transport complex. In bacteria containing the three gene *arsRBC* operon, in which the *arsA* is not expressed, resistance is conferred by means of carrier-mediated efflux via the membrane-spanning ArsB, where energy is supplied by the membrane potential of the cell. In bacteria harboring the five gene *arsRDABC* operon, where ArsA is co-expressed with ArsB, extrusion is catalyzed by an As(III)/Sb(III)-stimulated translocating ATPase composed of the catalytic ArsA ATPase subunit and the integral membrane protein ArsB. Immunoblotting binding experiments performed in *unc* mutant *E. coli* cells demonstrated that the integral membrane ArsB is a prerequisite for the association of ArsA to the

bacterial inner membrane (Tisa and Rosen, 1990; Dey *et al.*, 1994; Dey and Rosen, 1995; Kuroda *et al*., 1997).

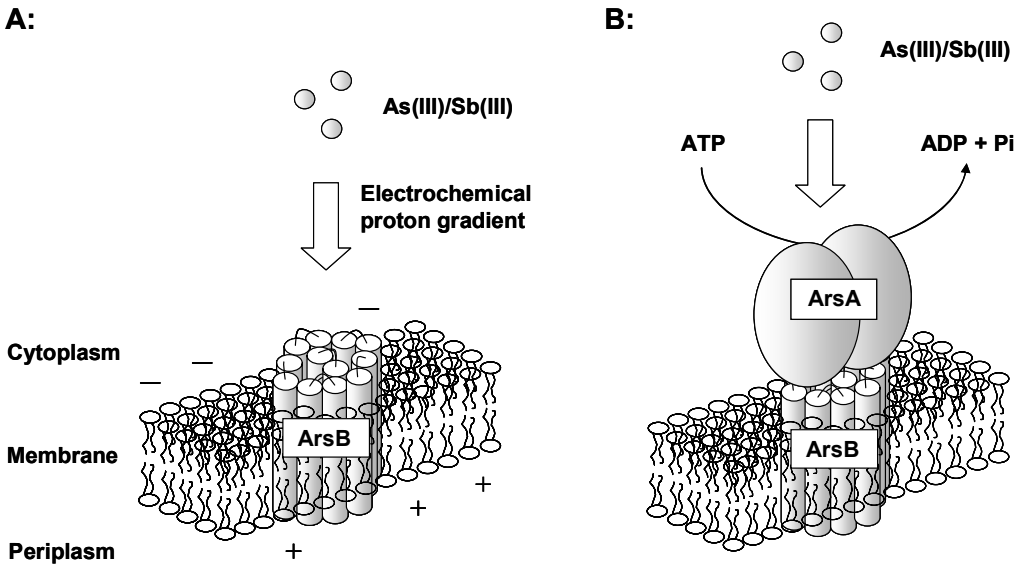


Figure 1.4: **The dual mode of energy coupling of the arsenic transport systems of *E. coli* plasmid R773.** The extrusion of arsenical and antimonial compounds from bacterial cells is mediated by two different transport systems. **(A)** The ArsB, a membrane-associated efflux pump, uses the electrochemical proton gradient as driving force to extrude intracellular arsenite/antimony. **(B)** When ArsA, an arsenite/antimony-stimulated ATPase, is co-expressed with ArsB, the ArsA can physically associate with ArsB to form an ArsAB ATPase complex which mediates efflux of intracellular arsenite/antimony in an ATP-dependent process. The ArsAB ATPase complex is much more efficient at arsenite/antimonite extrusion than ArsB alone. Adapted from Kuroda *et al.*, 1997.

The question arises concerning the importance of ArsA in bacterial *ars* operons, as *ars* operons not expressing ArsA still confer resistance to arsenical and antimonial compounds. One possibility is that it would expand the range of substrates recognized by the transport complex, allowing for the resistance to additional compounds. Another possibility is that ArsAB is a more effective resistance mechanism. Differences in levels of arsenic resistance were detected in *unc* mutant strains of *E. coli* expressing just *arsB* and strains expressing both *arsA* and *arsB*. *E. coli* cells expressing the *arsB* gene conferred an intermediate level of arsenite resistance compared with cells expressing both *arsA* and *arsB*. Broër *et al*. showed that the co-expression of the *E. coli* plasmid R773

ArsA with the ArsB of *S. aureus* plasmid pI258, which has 58% identity to the *E. coli* plasmid R773 ArsB, resulted in an increased resistance to arsenic (Broër *et al*., 1993; Dey and Rosen, 1995). Several other advantages to utilizing an ATPase should also be considered. Since ATP levels drop more slowly than the membrane potential under conditions of stress, an ATPase would be a more effective and suitable mechanism to extrude toxic compounds from the cell. Furthermore, ArsB function is limited to the equilibrium potential determined by the environment. In contrast, an ATPase would be able to maintain a low intracellular substrate concentration, independent of the external environment (Tisa and Rosen, 1990; Tsai *et al*., 1997; Rensing *et al*., 1999).

1.7.1.1 ArsB: Membrane-associated arsenite/antimony efflux pump

Researchers have previously been unsuccessful in performing biochemical analysis on the ArsB polypeptide due to its low levels of expression. In order to obtain sufficient amounts of ArsB polypeptide for visualization, Francisco and coworkers (San Francisco *et al*., 1989) used a mini-Mu-mediated *lacZ* gene translational fusion together with the T7 RNA polymerase-promoter expression system to demonstrate that the ArsB protein is an integral membrane protein located in the inner membrane. This result is consistent with initial ArsB polypeptide hydropathy plot experiments performed by Chen *et al*., (1986) which indicated that the ArsB polypeptide contains at least 10 hydrophobic regions that could be membrane-spanning α-helices. It was proposed that a poor level of *arsB* expression is due to several factors operating at a translational level. Analysis of the *arsB* translational initiation region (TIR) revealed several features that might explain the poor expression of *arsB*. Firstly, a GC-rich sequence of 8 bp could potentially form a stable stem-loop structure. Secondly, the second codon, UUA, is a less utilized leucine codon. Thirdly, previous studies had shown that the *arsB* mRNA transcript is not stable, with rapid degradation occurring within the *arsB* open reading frame (Tisa and Rosen, 1990; Dou *et al*., 1992).

Two unrelated families of bacterial membrane-associated arsenite/antimony efflux pumps have been identified to date (Rosen, 2002). The ArsB of *E. coli* plasmid R773 is a

member of the best studied family of arsenite/antimony efflux pumps (ArsB family), which is found in the majority of bacteria. The ArsB of plasmid R773 is an integral membrane protein composed of 429 amino acid residues (45.6 kDa). The topological arrangement of the ArsB was determined after a series of in frame fusions to *blaM*, *lacZ* and *phoA* genes was made. Analysis of data from 26 fusions led to the proposal of a topological model in which the ArsB has 12 trans-membrane α- helices, with the N- and C-termini located in the cytosol. The ArsB polypeptide contains five cytoplasmic loops and six periplasmic loops. Three of the cytoplasmic loops (C1, C3 and C5) have a net positive charge, whereas five of the six periplasmic loops are either uncharged (P1 and P3) or have a net negative charge (P2, P4 and P5). Based on the suggested topological model of the ArsB polypeptide, the cytoplasmic C3 loop has 22 amino acid residues, 10 of which are charged. Furthermore, two other cytosolic loops, C1 and C4, contain 4 and 5 charged amino acid residues respectively. It is possible that these cytosolic loops are involved in the physical interaction of the ArsA with the bacterial inner membrane (Chen *et al*., 1986; Wu *et al*., 1992). The second and less studied family of bacterial arsenite efflux pumps (Acr3p family) is comprised of the ArsB encoded by the *ars* operon present on the *B. subtilis skin* element. The ArsB polypeptide contains 10 trans-membrane α-helices and was shown to transport arsenite but not antimonite (Sato and Kobayashi, 1998).

The ArsB functions as a uniporter carrier protein, obtaining its energy from the electrochemical proton gradient. However, the chemical nature of the transported species is unclear. The ArsB protein contains a single residue, Cys 369, which has been predicted to be located in the 11th trans-membrane α-helix. Site-directed mutagenesis of this single cysteine residue to a serine or alanine did not result in significant changes in level of arsenite/antimonite resistance when compared to the wild type. This strongly suggests that the transport of arsenite/antimonite by ArsB does not involve metal thiol chemistry, indicating that electrophoretical substrate transport, in response to membrane potential, is the most likely alternative. The nature of the transported species has not yet been determined. It has been postulated that the chemical form of the substrate

recognized by the ArsB pump is the arsenite/antimonite oxyanion (Chen *et al*., 1996; Rensing *et al*., 1999).

1.7.1.2 ArsA: Arsenite/antimony-stimulated ATPase

The 583 residue ArsA (63.2 kDa) of *E. coli* plasmid R773 functions as the catalytic energy transducing component of the As(III)/Sb(III)-stimulated translocating ATPase. When ArsA is purified in the absence of ArsB, ArsA is a soluble ATPase that is allosterically activated by either As(III) or Sb(III) (Rosen *et al*., 1999). The ArsA is composed of two structurally homologous halves, a N-terminal A1 (residues 1-288) and a C-terminal A2 (residues 314-583) domain, connected by a flexible linker peptide of 25 residues (residues 289-313). The two halves share 23 % identity and are most likely the result of an ancestral gene duplication and fusion (Bhattacharjee *et al*., 2000; Li and Rosen, 2000).

The crystal structure of ArsA has revealed three different domains that contribute to its catalytic properties. Each homologous half contains a nucleotide binding domain (NBD) with a consensus glycine-rich P-loop (GKGGVGKT) which binds ATP. The two NBDs are located at the interface between the two homologous halves and in close proximity to each other. Mutations in either the A1 or A2 NBD resulted in loss of As(III)/Sb(III) resistance, loss of As(III)/Sb(III) transport and loss of ATPase activity, suggesting that both NDBs are required for enzymatic activity (Karkaria *et al*., 1990; Kaur and Rosen, 1993). In addition to the NBDs, a highly conserved metalloid-binding domain (MBD) is positioned at the opposite end of each of the homologous halves with respect to the NBDs. The two MBDs are composed of three cysteine residues (Cys-113 and Cys-172 (A1); Cys-422 (A2)), two histidine residues (His-148 (A1); His-453 (A2)) and one serine residue (Ser-420(A2)). Allosteric activation of the ArsA ATPase occurs through binding of As(III) or Sb(III) to Cys-113, Cys-172 and Cys-422 (Bhattacharjee *et al*., 1995; Zhou *et al*., 2000). Transduction of the information between NBD and MBD is mediated by a 12-residue consensus sequence DTAPTGHTIRLL signal transduction domain, termed DTAP, found on each of the two homologous halves of ArsA. Tryptophan fluorescence

experiments performed on tryptophan residues Trp-141 and Trp-159 of DTAP showed that DTAP undergoes significant conformational change during ATP hydrolysis. This implies that the NBDs and MBDs are connected to each other by the physical movement in each DTAP (Zhou and Rosen, 1997; Rensing *et al*., 1999; Bhattacharjee *et al*., 2000). The mechanism of allosteric activation of the ArsA ATPase has been determined by a combination of molecular genetics and biochemical, structural and kinetic analysis. ArsA has a low level of ATPase activity. The interaction of As(III) or Sb(III) with Cys-113 and Cys-172 in the A1 subunit and Cys-422 in the A2 subunit leads to the physical movement of the two subunits toward each other. The two NBDs interact at the newly formed interface and ultimately lead to the acceleration of ATP hydrolysis. The rate limiting step in this mechanism is the isomerization between the conformations of ArsA, after ATP hydrolysis and substrate release are completed. There is no conclusive evidence that the metalloid ions which activate the hydrolytic activity of ArsA, are the same ions that are transported across the membrane via ArsB (Rosen *et al*., 1999; Bhattacharjee *et al*., 2000; Li and Rosen, 2000; Zhou *et al*., 2001).

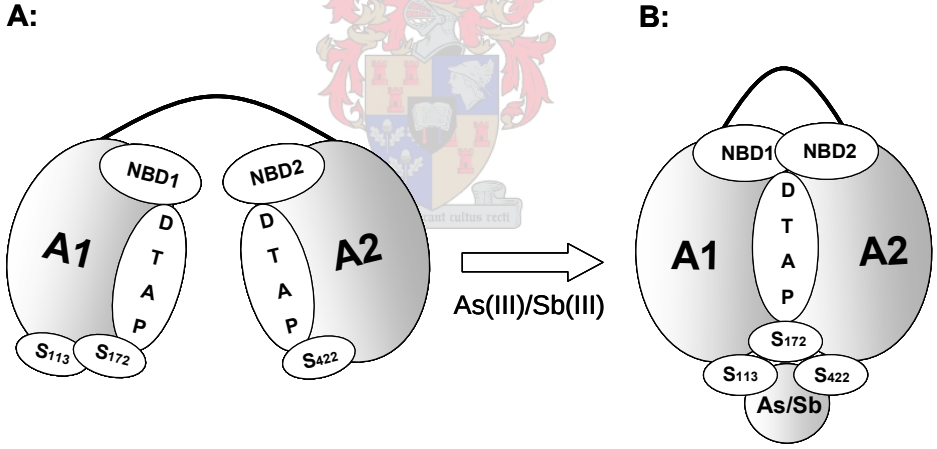


Figure 1.5: **Model of the As(III)/Sb(III)-stimulated ArsA ATPase of the *E. coli* plasmid R773. (A)** In the absence of the allosteric effectors As(III) or Sb(III), the A1 and A2 subunits of ArsA connected by a 25bp flexible linker, have a low basal rate of ATP hydrolosis. **(B)** The binding of either As(III) or Sb(III) to Cys133, Cys172 or Cys422 comprising the metalloid binding domain (MBD), results in significant conformational changes of the signal-transduction domain (DTAP). This conformational change brings the A1 and A2 subunits together, forming an interface between the two nucleotide binding domains (NBD) which results in increased levels of ATP hydrolysis. Adapted from Rensing *et al*., 1999.

1.7.2 ArsC: Arsenate reductase

The bacterial arsenate reductases are small monomeric cytoplasmic proteins that utilize cysteine thiol oxidation/reduction cycling to convert intracellular arsenate to arsenite, the substrate of the ArsB efflux pump. The question arises as to why less toxic arsenate is reduced to more toxic arsenite prior to extrusion? It is likely that arsenic extrusion mechanisms evolved earlier than the arsenate reductases. Since the early earth atmosphere was not oxidizing, arsenite was probably the predominant form of arsenic in the environment and early organisms would have evolved extrusion systems to cope with arsenite and not arsenate. Once the atmosphere became oxidizing, selective pressure arose to evolve an enzyme to facilitate the reduction of intracellular arsenate to arsenite, prior to transport by the existing arsenite extruding systems (Mukhopadhyay *et al*., 2002; Rosen, 2002). Furthermore, evolutionary development and selection of arsenate-specific efflux systems would possibly have been less feasible. The ability of bacteria to discriminate between intracellular phosphate and arsenate might be difficult, and could ultimately lead to the leakage of phosphate from the cell (Nies and Silver, 1995; Silver *et al*., 2002).

Enzymes capable of reducing arsenate to arsenite have apparently evolved independently at least three times (Mukhopadhyay *et al*., 2002). These arsenate reductases can by subdivided into families whose sequences are unrelated and whose energy-coupling mechanisms differ in detail. Two distinct families of arsenate reductases have been identified in bacteria. The families exhibit only 15% sequence homology and members representative of each family were shown to reduce arsenate to arsenite both *in vivo* and *in vitro* and additionally, confer resistance to arsenate. The first and best-studied group of arsenate reductase is referred to as the glutaredoxin/glutathione (Grx/GSH) clade and has the ArsC *of E .coli* plasmid R773 ($ArsC_{ec}$) as the prototype, while the second family of arsenate reductases is referred to as the thioredoxin (Trx) clade and has the ArsC of *S. aureus* plasmid pI258 ($ArsC_{sa}$) as the prototype. Several differences between the two bacterial ArsC families were detected from protein crystallography, enzymology and mutational studies. The 15.8 kDa $ArsC_{ec}$ (141 aa residues) contains two cysteine

residues, Cys12 and Cys106. Single mutations of each cysteine residue demonstrated that Cys106 is not required for reductase activity, while alteration at Cys12 resulted in the loss of reductase activity (Lui *et al*., 1995). Cys12 is surrounded by a triad of basic aa residues composed of Arg60, Arg94 and Arg107, which forms a binding site for oxyanions (Shi *et al*., 2003). It has been demonstrated that phosphate, sulfate (not nitrate) and arsenite (not antimonite) are competitive inhibitors for $ArsC_{ec}$ activity (Gladysheva *et al*., 1994). The 14.8 kDa $ArsC_{sa}$ (131 aa residues) is related to a family of low-molecular weight proteins called tyrosyl phosphate phosphatases and exhibits low-level phosphatase activity (Zegers *et al*., 2001). It contains three cysteine residues, Cys 10 (equivalent to Cys12 of $ArsC_{ec}$), Cys82 and Cys89. Mutational analysis showed that all three cysteine residues were essential for reductase activity in *S. aureus* (Messens *et al*., 1999). In contrast, phosphate and nitrate (not sulfate) stimulates the activity of $ArsC_{sa}$, whereas arsenite, antimonite and tellurite have an inhibitory effect (Ji *et al*., 1994). The *Saccharomyces cerevisiae* Arr2p is the only currently identified eukaryotic arsenate reductase and represents the third convergent evolutionary family of arsenate reductases (Mukhopadhyay and Rosen, 1998).

The most striking difference between the two bacterial ArsC families is the nature of their energy coupling systems. Both bacterial ArsC families require additional proteins to assist in the reduction reaction. $ArsC_{ec}$ derives its reducing power from reduced GSH and Grx (small intracellular proteins that function as disulfide reducing agents) as a source to supply reduction potential (Oden *et al*., 1994; Martin *et al*., 2001), whereas $ArsC_{sa}$ requires reduced thioredoxin (similar to the GSH and Grx, but with different substrate specificity) for reductase activity (Ji and Silver, 1992; Ji *et al*., 1994). This strongly suggests that the Grx/GSH–coupled and Trx-coupled ArsC arsenate reductases represent two families that might have evolved independently to perform a similar chemical reaction. Based on data obtained from previous research, models describing the detailed catalytic mechanisms of both bacterial ArsC families, with interesting similarities and differences, have been proposed.

The following mechanism for the enzymatic reduction of arsenate by $ArsC_{ec}$ was suggested by Martin *et al.*, (2001) and Demel *et al.*, (2004): In the **first step**, arsenate binds non-covalently via the thiolate of Cys-12 and the three basic residues Arg60, Arg94 and Arg107 and subsequently leads to the formation of a thioarsenate binary adduct with the release of a water molecule. In the **second step**, GSH attacks the thioarsenate binary adduct and a {ArsC Cys-12}S-As(V)-S{glutathione} tertiary complex is formed and a water molecule is released. This reaction requires a free thiol on GSH and ArsC to proceed. Glutathione reacts only after arsenate interacts with the active binding site. During the **third step**, arsenate is reduced to a dihydroxy monothiol As(III) intermediate and the binding of Grx results in the formation of a water molecule and a mixed disulfide GrxS-SG compound, which would be recycled by glutathione reductase utilizing another equivalent of GSH. In the **fourth step**, the arsenite-ArsC bond is hydrolyzed, causing the enzyme-A(III) complex to dissociate with the release of arsenite. ArsC returns to its original conformation, which completes the cascade of events responsible for the enzymatic reduction of arsenate. The proposed pathway and intermediates are shown in Figure 1.6.

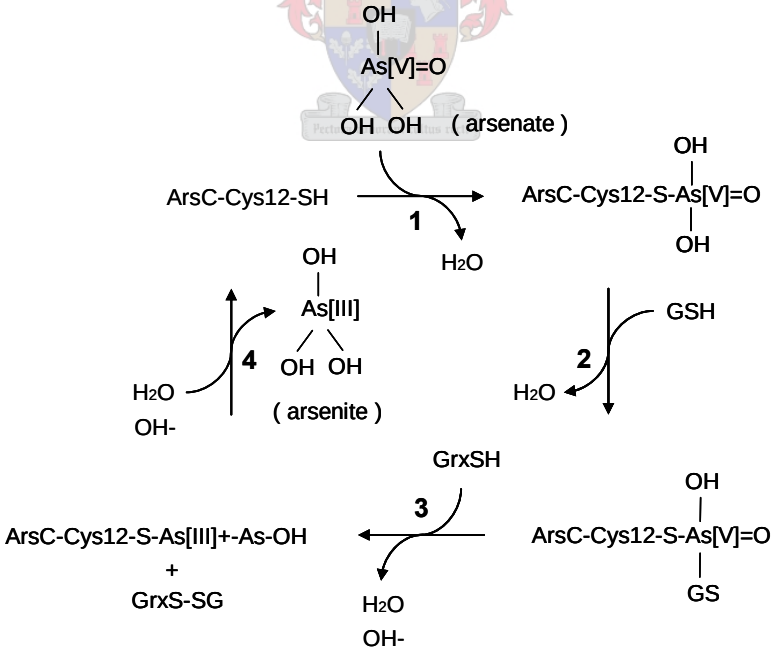


Figure 1.6: **The proposed reaction mechanism pathway of the $ArsC_{ec}$-family.** Details of the mechanism are described in the text. Adapted from Mukhopadhyay *et al.*, 2002.

Based on research done by the respective research groups of Bennet, Messens and Zegers, several features of the $ArsC_{sa}$-family arsenate reduction mechanism are essentially different from that observed in the $ArsC_{ec}$-family. The following catalytic mechanism was proposed: In the **first step** of the reaction, the arsenate substrate binds via the thiolate of Cys-10. A hydroxyl is subsequently protonated to a water molecule which leaves the arsenate substrate to facilitate the formation of a covalent Cys-10-$HAsO_3^-$ intermediate. The **second** and **third steps** involve the three essential cysteines (Cys-10, Cys82 and Cys89) in a triple cysteine redox relay system to form the Cys-82-Cys-89 disulfide bond and the release of an arsenite ion. In the **fourth step**, ArsC is regenerated by thioredoxin that reduces the Cys82-Cys89 disulfide bond, enabling the ArsC to participate in the next cycle of enzymatic arsenate reduction (Bennet *et al*., 2001; Messens *et al*., 2002; Zegers *et al*, 2001). The suggested pathway and intermediates are illustrated in Figure 1.7.

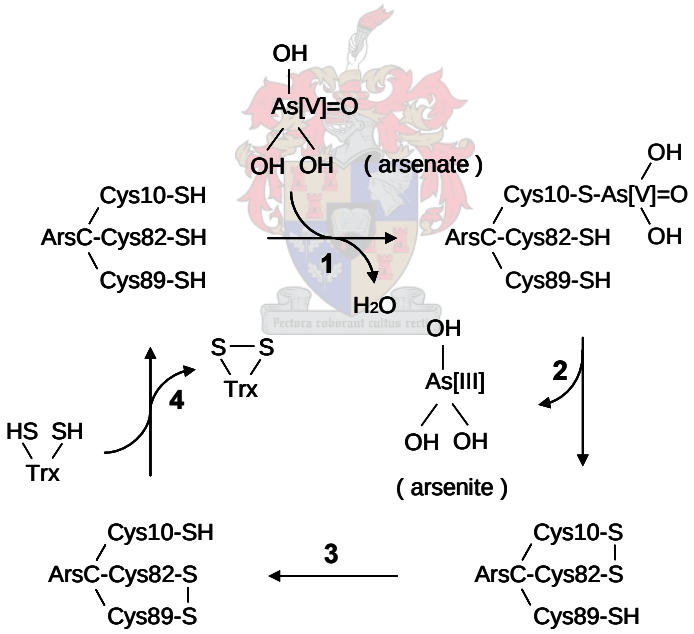


Figure 1.7: **The proposed reaction mechanism pathway of the $ArsC_{sa}$-family.** Details of the mechanism are described in the text. Adapted from Mukhopadhyay *et al*., 2002.

1.7.3 The regulation of *ars* operons

Although several reports have been documented on the regulatory abilities of the two inducer-responsive *trans*-acting repressors (*arsR* and *aseR*) on the *ars* operons of *Bacillus subtilis*, most of the knowledge about the regulation of *ars* operon expression has been obtained from studies conducted on the *ars* operon of *E. coli* plasmid R773. It was shown that transcriptional regulation of the *ars* operon of plasmid R773 involves the association of two different inducer-responsive *trans*-acting regulatory proteins, ArsR and ArsD, with a single promoter sequence (Tsai *et al*., 1997; Silver *et al*., 2002; Moore *et al*., 2005; Moore and Helmann, 2005).

1.7.3.1 ArsR: Primary *trans*-acting repressor

The 117 amino acid residue ArsR (13.18 kDa) of plasmid R773 is a *trans*-acting cytoplasmic polypeptide that negatively regulates the basal level of expression of the *ars* operon, including its own synthesis (San Francisco *et al*., 1989; Wu and Rosen, 1991). ArsR is a member of the prokaryotic SmtB/ArsR family of metalloregulatory transcriptional repressors that respond to a variety of metals including As(III), Sb(III), Cd(III) and Zn(III). It has been postulated that members constituting the SmtB/ArsR family of transcriptional regulators contain at least three domains: a metal binding domain, a DNA-binding domain and a dimerization domain (Xu and Rosen, 1997; Busenlehner *et al*., 2003).

Although members of the SmtB/ArsR family of transcriptional regulators show low sequence similarity to each other (e.g. the ArsR repressors from *S. aureus* plasmids pI258 and pSX267 are only approximately 30% identical to the *E. coli* chromosomal and plasmid R773 ArsR repressors), the majority of them have a highly conserved $ELC_{32}V(G/C_{34})DL$ sequence (Shi *et al*., 1994). However, one exceptional case is the atypical ArsR repressor of *Acidithiobacillus ferrooxidans*, which contains a unique conserved GX(L/I)A sequence (Butcher and Rawlings, 2002). These conserved sequences are associated with the N-terminal of the ArsR polypeptide and are believed to

form a portion of the domain involved in metal ion sensing. In order to identify the metal binding domain in ArsR, Shi and coworkers (Shi *et al*., 1994) selected three mutants, which showed inability to respond to the inducer but retained expression, for analysis. Mutants C32Y, C32F and C34Y, containing alterations in the cysteine residues of the highly conserved sequence, retained the ability to bind to the *ars* promoter, but had reduced inducer response *in vivo* and *in vitro*. This result indicated that the cysteine pair Cys-32 and Cys-34 comprise part of the inducer binding domain of the ArsR protein. However, the inducers As(III) and Sb(III) are frequently in a three coordinate conformation, suggesting that a third aa residue might be involved in inducer binding in ArsR. In addition to Cys-32 and Cys-34, ArsR contains three other cysteine residues, Cys-37, Cys-108 and Cys-116. However, it was previously shown that Cys-108 and Cys-116 were not required for repressor function, but site-directed mutagenesis of Cys-37, revealed that Cys-37 was required for metal binding (Wu and Rosen, 1991). It was postulated that all three cysteine residues are capable of forming soft metal bonds between the sulfur thiolates and trivalent arsenic, but the binding of only two cysteine residues, in any combination, will lead to conformational change that results in the dissociation of ArsR from the DNA (Shi *et al*., 1996).

Located adjacent to the metal binding box is a putative helix-turn-helix motif characteristic of known DNA binding regions. A mutant (H50Y), showing constitutive levels of *ars* operon expression and responding poorly to inducers, was isolated and analysed by Shi *et al*., (1994). His50 was shown to be localized within the second putative helix, suggesting that this region might be involved in DNA recognition and association. The specific DNA binding site of ArsR was identified by DNAseI footprint analysis which revealed that the DNA binding site consists of a region of imperfect dyad symmetry that spans from -64 to -40 relative to the transcriptional start of the *ars* operon mRNA. It is known that RNA polymerase occupies promoter sequences from about -50 to +20 relative to the transcriptional start site, so the binding of ArsR to its DNA binding site will therefore result in the blockage or interference in the initiation of the transcriptional process (Wu and Rosen, 1991; 1993a).

Further analysis of the R773 ArsR binding site with hydroxyl radical foot-printing revealed that the repressor protected two small regions of 4 bp each. These 4 bp protected regions were separated by a stretch of 10 bp, indicating that the active form of the ArsR is most likely a homodimer (Wu and Rosen, 1993b). Similar symmetrical dyad sequences have also been found within the promoter regions of putative ArsR DNA recognition and binding areas in the *ars* operons of Gram-positive bacteria (Ji and Silver, 1992b; Rosenstein *et al*., 1994; Sato and Kobayashi, 1998). The presence of symmetrical dyad binding sequences in both Gram-negative and Gram-positive bacterial *ars* operon promoter areas implies the possible existence of a dimerization domain in ArsR. The dimerization properties of the R773 ArsR were investigated by using a yeast two-hybrid system which analyses protein-protein interactions. Data indicated that the sequences required for dimerization are contained in residues 9-89, suggesting that a core sequence of approximately 80 residues is sufficient to perform the regulatory activities of the ArsR repressor: metal recognition, DNA binding and dimerization (Xu *et al*., 1996, Xu and Rosen, 1997)

1.7.3.2 ArsD: Secondary *trans*-acting repressor

The *arsD* of the *ars* operon of plasmid R773 encodes a second cytoplasmic *trans*-acting regulatory polypeptide involved in controlling the upper level of expression of the *ars* operon. Initial studies to elucidate the function of the 26.44 kDa homodimeric ArsD (120 amino acid residue per polypeptide monomer) were performed by Wu and Rosen, 1993b. They found that the introduction of a frameshift mutation into the *arsD* gene resulted in increased levels of expression of the downstream *ars* genes. Interestingly, mutants expressing the *ars* operon with the defective *arsD* gene showed sensitivity to arsenite, even though the *arsABC* genes were transcribed. The co-expression of a wild-type *arsD* gene in *trans* with the *ars* operon containing the mutated *arsD* gene caused reduced expression of the downstream genes to wild-type levels. It appears that the down-regulation of the *ars* operon is dependent on the amount of intracellular ArsD present. This implies that the expression of *arsD* is required for resistance when the transcription of *arsABC* is controlled by the *ars* operon promoter. Unlike the ArsR repressor protein,

the ArsD repressor protein was shown to be inducer-independent as the addition of As(III) or Sb(III) did not prevent the decreased expression levels of the *ars* operon.

Although ArsR and ArsD share no sequence homology, DNaseI footprinting analysis showed that ArsD binds as a homodimer to the same operator/promoter region usually occupied by ArsR. However, the ArsD binds with two orders of magnitude lower affinity compared to ArsR, suggesting that ArsR will bind to the promoter site preferentially during *ars* operon expression. The R773 ArsD contains eight cysteine residues, more than any of the other proteins encoded by the *ars* operon. Six of the cysteines residues are located on each monomer and are organized in vicinal pairs, Cys-12-Cys13, Cys-112-Cys113 and Cys-119-Cys-120. The role of these cysteine pairs in metalloid binding was determined. Mutation or deletion of the Cys-119-Cys-120 pair had no effect on the repression or metalloid responsiveness of ArsD. In contrast, introduced mutations in the Cys-12-Cys13 and Cys-112-Cys113 pairs resulted in the loss of inducer response *in vivo* and *in vitro*, but had no influence on the ability of ArsD to repress *ars* operon expression, an indication that these cysteine pairs are required for ArsD activity *in vivo*. However, as previously mentioned, the inducers As(III) and Sb(III) are in a three coordinate conformation, suggesting that a third ligand is probably involved in the binding process. Li and Rosen (2002) have postulated an ArsD/inducer-binding model where the third co-ordination is to a hydroxyl group. Furthermore, *in vivo* experiments demonstrated that ArsR has a higher affinity to its inducer compared to ArsD (Li *et al*., 2001; 2002).

As illustrated in Figure 1.8, the level of *ars* gene expression is maintained within a narrow range by a homeostatically regulated circuit formed by the *trans*-acting repressor proteins ArsR and ArsD. From an aggregate of results, the following model for the regulated circuit was proposed by Chen and Rosen (1997): **(A)** In the absence of the inducer As(III) or Sb(III), constitutively expressed ArsR binds to the operator/promoter site of the *ars* operon, repressing transcription. **(B)** In the presence of low levels of As(III)/Sb(III), the ArsR-As(III)/Sb(III) complex dissociates from the operator/promoter site, resulting in the transcription of the *ars* operon. **(C)** Sufficient increase in the concentration of the low affinity DNA binding protein ArsD, leads to the binding of

ArsD to the operator/promoter site and transcription is repressed again. This is of great importance, as this prevents the expression of integral membrane ArsB proteins at levels that might become toxic for the cell. **(D)** In the presence of high levels of As(III)/Sb(III), the ArsD-As(III)/Sb(III) complex dissociates from the O/P site, leading to a further increase in *ars* operon expression. As discussed above, these two *trans*-acting repressors form an integral regulatory circuit with ArsR controlling the basal level of *ars* operon expression, while ArsD is responsible for controlling the upper level of *ars* operon expression (Wu and Rosen, 1993b).

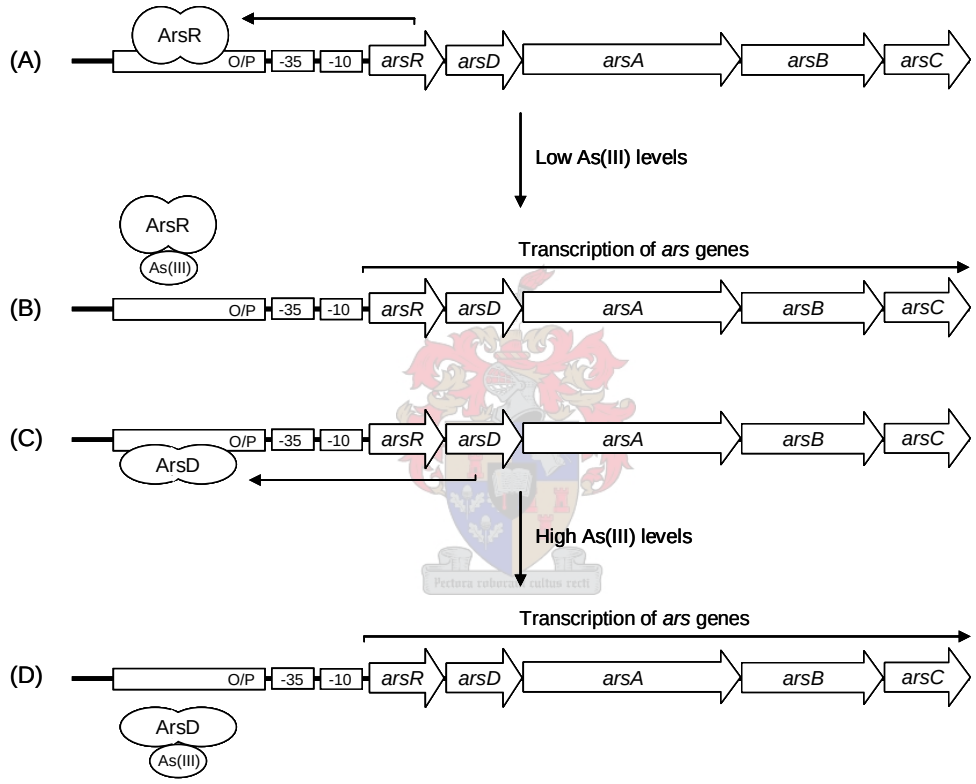


Figure 1.8: **A model of the metalloregulatory circuit of ArsR and ArsD, using the *ars* operon of *E. coli* plasmid R773 as an example.** Details of the mechanism are described in the text. Adapted from Chen and Rosen, 1997.

In addition to the regulatory role of ArsD in *E. coli* plasmid R773 *ars* operon expression, Lin and coworkers (Lin *et al*., 2006) have proposed that the ArsD of plasmid R773 could act as a metallochaperone which sequesters intracellular arsenite/antimony and delivers it to ArsA, the catalytic subunit of the ArsAB efflux ATPase. They found that ArsD interacts with ArsA with low affinity in the absence of metalloid and high affinity when

the metalloid is bound to ArsD. Furthermore, it was shown that ArsD binds to the metalloid with a higher affinity in comparison to ArsA. This will enable the ArsD to scavenge the cytosol for free metalloids for delivery to the ArsA, allowing the ArsAB efflux ATPase to regulate intracellular arsenite/antimony concentrations by extrusion.

1.8 Aim of Thesis

Moderately thermophilic organisms, like bacteria representative of genus *Sulfobacillus*, will increasingly play a bigger role in commercial bio-oxidation processes occurring at elevated temperatures. During these processes, microbially catalyzed oxidation of arsenopyrite (FeAsS) containing ores may lead to the production of high and toxic levels of arsenical compounds within biooxidation tanks. Bacteria have developed arsenic efflux mechanisms to tolerate the harmful effects of arsenical compounds. Components of these arsenic efflux systems are encoded by genes located within *ars* operons.

Little is known about the genetic systems and regulatory mechanisms involved in arsenic resistance in members of genus *Sulfobacillus*. The aim of this project was to identify and characterise the arsenic resistance genes of *Sulfobacillus thermosulfidooxidans* VKM B-1269^T (type strain VKM B-1269^T=AT-1^T=DSM 9293^T) (Golovacheva and Karavaiko, 1978). This strain, originally from an ore of the Nikolaev copper-zinc-pyrite deposit of Eastern Kazakhstan in Russia, was the first reported case of a bacterium representative of genus *Sulfobacillus* to be isolated and cultured. The isolation and sequence analysis of the *ars* operon would lay the foundation to determine how closely related and widely distributed the *ars* genes are among members representative of genus *Sulfobacillus*. Furthermore, expression and regulatory studies will be performed on the *ars* operon in both its native host and a heterologous *Escherichia coli* host. An attempt to address these questions has led to the work presented in the dissertation to follow.

CHAPTER 2

THE ISOLATION, SEQUENCING AND ANALYSIS OF THE ARSENIC RESISTANCE GENES OF *SULFOBACILLUS THERMOSULFIDOOXIDANS* VKM B-1269^T

TABLE OF CONTENT

2.1 Introduction 39

2.2 Material and Methods

2.2.1 Bacterial strains, plasmids and PCR primers 40

2.2.2 Media and growth conditions 40

2.2.3 DNA isolation, techniques and analysis 42

2.2.4 Sulfobacilli media and growth conditions 42

2.2.5 Total DNA isolation from Sulfobacilli 43

2.2.6 Detection of constructs containing the *ars* operon from *Sb. thermosulfidooxidans* VKM B-1269^T (*Sb. t.* VKM B-1269^T) by using colony hybridization 44

2.2.7 Polymerase chain reaction (PCR) 44

2.3 Results

2.3.1 Identification and isolation of the *ars* operon of *Sb. t.* VKM B-1269^T 45

2.3.2 Restriction mapping and sequence analysis of the *ars* operon of *Sb. t.* VKM B-1269^T 46

2.3.3 Detection of putative *arsC* gene homologues in *Sb. t.* VKM B-1269^T 58

2.3.4 Minimum inhibitory arsenic concentrations of *Sb. t.* VKM B-1269^T 59

2.4 Discussion 60

2.1 Introduction

Biomining is a well established biotechnological practice which implements microbial-aided decomposition and solubilization processes for the recovery of certain mineral compounds. Although the majority of these industrial processes are carried out by mesophilic microbial communities (T_{opt} being 20-40°C), considerable interest has been shown in understanding moderate thermophilic microorganisms, as they will increasingly play bigger roles in commercial biomining processes operational at elevated temperatures. Conducting commercial biochemical processes at higher temperatures (>40°C) improves the rate of mineral biooxidation and has substantial economical advantages over conventional operations carried out in the vicinity of 40°C. Bacteria representative of genus *Sulfobacillus* have been identified as possible members of microbial consortia in biomining plants that function at elevated temperatures. These bacteria are Gram-positive, acidophilic, moderate thermophiles and employ a highly versatile metabolism.

The microbially catalyzed oxidation of arsenopyrite (FeAsS) containing ores leads to the production of arsenical compounds and may result in an increase in levels of arsenic within biooxidation tanks. Bacteria have evolved arsenic efflux mechanisms to tolerate the harmful effects of arsenical compounds. Components comprising these arsenic efflux systems are encoded by genes located within *ars* operons. These *ars* operons are usually found on chromosomes, plasmids and transposable elements and may differ from each other with respect to the gene layout and number of genes. Two of the most commonly encountered forms of *ars* operon arrangement consist of five genes (*arsRDABC*) and three genes (*arsRBC*), respectively. The five gene *ars* operon has been found on plasmids of Gram-negative bacteria (plasmids R773 and pR46 of *E. coli* and pKW301 of *Acidiphilium multivorum*) (Chen *et al*., 1985; Bruhn *et al*., 1996; Suzuki *et al*., 1998), while the three gene *ars* operon is associated with plasmids of Gram-positive bacteria (*Staphylococcus aureus* (pI258); *Staphylococcus xylosus* (pSX267)) (Ji and Silver, 1992b; Rosenstein *et al*., 1992) and the chromosomes of Gram-negative bacteria *E. coli*,

Pseudomonas aeroginosa and *Pseudomonas fluorescens* (Carlin *et al*., 1995; Cai *et al*., 1998; Prithivirajsingh *et al*., 2001).

Here we report the isolation and sequence analysis of the *ars* operon of *Sb. t.* VKM B-1269^T. We showed that this *ars* operon consists of only two *ars* genes (*arsR* and *arsB*, respectively), making it only the second reported bacterial *ars* operon consisting of two *ars* genes. We further showed that representatives of either the Grx/GSH and Trx ArsC families appear to be absent in the genome of *Sb. t.* VKM B-1269^T. The *ars* operon of *Sb. t.* VKM B-1269^T was not functional in *E. coli* and could not confer resistance to arsenate or arsenite in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant.

2.2 Material and Methods

2.2.1 Bacterial strains, plasmids and PCR primers.

The bacterial strains, plasmids and primers used in this study are described in Table 2.1.

2.2.2 Media and growth conditions.

E. coli cells were grown aerobically on Luria Bertani agar plates or in Luria Bertani broth at 37°C as described by Sambrook *et al.*, 1989. Ampicillin was added at a concentration of 100 µg/ml when required.

Table 2.1: **Bacterial strains, plasmids and PCR primers used in this study**

Strains, plasmids or primers	Genotype or description	Reference or source
Strains		
Escherichia coli		
DH5α	*F'/endA1 hsdR17* ($r_K^-m_K^+$) *supE44 thi-1recA1 gyrA* (Nal^r) *relA1 Δ(lacZYA-ArgF) U169 (Φ80dlac Δ(lacZ) M15)*	Promega Corp.
XL1-Blue	F'::Tn*10* proA$^+$*B*$^+$ *lacI*^{q} *Δ(lacZ) M15/recA1 endA1 gyrA96* (Nal^r) *thi hsdR17*($r_K^-m_K^+$) *supE44 relA1 lac*	Promega Corp.

GM41dam$^-$	*Hfr-H dam-3 thi-1 rel-1*	Valarie Mizrahi (University of Witwatersrand)
ACSH50Iq	*rpsL* Δ(*lac –pro*) Δ*ars*::*cam*	Butcher and Rawlings, 2002
Sulfobacillus thermosulfidooxidans VKM B-1269^T	Russia (copper/zinc-pyrite deposit)	Golovacheva and Karavaiko, 1978
Plasmids		
pBluescript SK$^+$	Ap^r; *lacZ'*; ColE1 replicon vector	Stratagene
p*Eco*R252	Ap^r; *Eco*RI inactivation cloning vector	Zabeau and Stanley,1982
pGEM-T®	Ap^r; T-tailed PCR product cloning vector	Promega
pKK223-3	Ap^r; *tac* promoter ColE1 replicon vector	Pharmacia Biotech
pUCBM21	Ap^r; *lacZ'*; ColE1 replicon vector	Boehringer-Mannheim
pStArs1	Ap^r; 5300 bp *Eco*RI-*Pst*I fragment, containing the *ars* operon of *Sb. t.* VKM B-1269^T, cloned into pUCBM21 digested with *Eco*RI *and Pst*I	This study
pUCarsCdeg	Ap^r; amplified 276 bp *ars*C fragment of pI258, using primer set prArsCfwd deg and prArsCrev deg, cloned into pUCBM21 digested with *Bam*HI and *Xba*I	This study
Primers[#]		
Ferro arsBfwd	5'-GTTIGCCAACGAIGGIGCGGC-3'	This laboratory
Ferro arsBrev	5'-ACATGCAICCAGAGCAGIGTIGC-3'	This laboratory
prKaraArsBfwd (*Eco*RI)	5'-GTGAGAATTCGATACAACCTCTGTCCCG-3'	This study
prKaraArsBfwd (*X*baI)	5'-GCTCTCTAGACCGATGAAAGAAGTGC-3'	This study
prArsCfwd deg (*Bam*HI)	5'-AGGATCCCGTAGCCAAATGGCTGAAG-3'	This study
prArsCrev deg (*Xba*I)	5'-ATCTAGAGCTGGATCATCAAAWCCCCAATG-3'	This study

Ap^r: ampicillin resistance

[#] restriction endonuclease sites incorporated into primers are indicated in parenthesis and are underlined in the primer sequence

2.2.3 DNA isolation, techniques and analysis.

Restriction endonuclease digestions, gel electrophoresis, small-scale plasmid preparation, ligation reactions and Southern-blot hybridization were performed using standard methods (Sambrook *et al.*, 1989). Large-scale plasmid preparation was done by means of the alkaline lysis procedure whereafter plasmid DNA was purified by CsCl/ethidium bromide equilibrium centrifugation (Current Protocols in Molecular Biology) or by using the Nucleobond AX100 system (Macherey-Nagel). DNA fragments to be used for cloning purposes were extracted from agarose gels with the GFX^{TM} kit (Amersham BioSiences). Clones to be sequenced were isolated and purified from 5 ml LB O/N culture using the High Pure Plasmid Isolation Kit (Roche Molecular Biochemicals). Sequencing of constructs was carried out by using the dideoxy chain termination method and was executed by an ABI PRISM^{TM} 377 automated DNA sequencer. Sequences were analysed using several software programmes, but mainly by the PC based DNAMAN (version 4.1) package (Lynnon Biosoft). The gapped-BLAST program of the National Center of Biotechnology Information (NCBI) at http://www.ncbi.nih.nlm.gov (Altshul *et al.,* 1997) was used for comparison searches. Alignments and the subsequent construction of the phylogenetic trees were done by using the Multiple Sequence Alignment tool in DNAMAN. The labeling of probes, hybridization and detection was conducted by using the dioxigenin-dUTP nonradioactive DNA labeling and detection kit (Roche Molecular Biochemicals). When Southern-blot hybridization was performed with a homologous DNA probe, prehybridization and probing were done at 42ºC and the buffer B washes were conducted at 65 ºC. In the cases where a heterologous DNA probe was utilized in Southern-blot hybridization, the prehybridization and probing steps were carried out at 37ºC and the buffer B washes performed at 52 ºC

2.2.4 Sulfobacilli media and growth conditions.

Sulfobacilli cells were grown aerobically under autotrophic conditions in ferrous sulfate ($FeSO_4$) media. The $FeSO_4$ media consisted of the following sterile stock solutions: 10X basal salts media 10.0 (v/v), pH 2.5, containing $(NH_4)_2SO_4$ 1.25 (m/v) and $MgSO_4$ 0.5 (m/v); 0.5M $FeSO_4$ 5.0 (v/v), pH > 1.3; 1000X trace elements 0.1 (v/v); tetrathionate ($K_2S_4O_6$) 0.5 (v/v) and Tryptone soy broth 0.025 (w/v). The pH of the stock solutions

were adjusted with concentrated H_2SO_4. The pH of the $FeSO_4$ media, containing the respective sterile stock solutions, was then adjusted with concentrated H_2SO_4 to pH ~2.2. Cultures were inoculated in flat-bottomed flasks and incubated at 37°C or 40°C on shakers. Sulfobacilli were kept in screw-top bottles containing autoclaved iron pyrite and $FeSO_4$ media for long term storage. Prior to total DNA isolation, cultures were routinely grown for ~2 days until the media became light orange in color (pH 2.5-2.7). In order to prevent excessive formation of jarosite ($KFe_3(SO_4)_2(OH)_6$), culture media pH were again adjusted to pH ~2.2 with concentrated H_2SO_4 (~0.3 (v/v)). Cultures were grown for a further 2 days whereafter they were re-inoculated in fresh ferrous sulfate ($FeSO_4$) media, now containing only 0.5M $FeSO_4$ 0.05 (v/v), pH > 1.3. Total DNA was isolated soon after the media became turbid (OD_{600} 0.06-0.10).

2.2.5 Total DNA isolation from Sulfobacilli.

Sulfobacilli cells were harvested from the ferrous sulfate ($FeSO_4$) media by centrifugation at 10000 rpm for 30 minutes. The cell pellet was resuspended in acidified water (pH 1.8) and subjected to several low speed (2000 rpm) and high speed (9000 rpm) centrifugation spins. These additional washing steps were performed in order to remove excess jerosite from the solution. The resulting cell pellet was then washed thrice with TE buffer (pH 8) in order to neutralize the media. The washed cells were resuspended in TE buffer whereafter they were lysed at 37°C in the presence of protease K (1 mg/ml). Additional 50 µl 10% SDS (0.8% g/v) was added to the lysate and it was incubated at 55°C for 30 minutes. Proteins were precipitated by using 0.5 volumes of 7.5 M $(NH)_4OAc$ (pH 7.5), whereafter the DNA was precipitated by the addition of 2.5 volumes of 100% ethanol (Current Protocols in Molecular Microbiology). The clean DNA pellet was resuspended in TE buffer and stored at 4°C.

An alternative method for total DNA isolation from *Sb. t.* VKM B-1269^T was used to obtain DNA for the assembly of a size selective mini gene bank. Genomic DNA was isolated using 500 µl 10% SDS (0.5% w/v) followed by incubation with 50 µl of 20 mg/ml Protease K. The lysate was treated with CTAB/NaCl solution (10% CTAB in 0.7 M NaCl) and purified by CsCl/Ethidium bromide equilibrium centrifugation (Current Protocols in Molecular Biology).

2.2.6 Detection of constructs containing the *ars* operon from *Sb. thermosulfidooxidans* VKM B-1269^T by using colony hybridization.

Electro competent DH5α cells were transformed with recombinant DNA and were plated on selective media. Colonies were then randomly selected, restreaked in a grid-like pattern on selective media and incubated O/N at 30°C. After incubation, colonies were replica plated on fresh selective media using a replica plating fork. Duplicate plates were made and contained 48 colonies per plate. One set of the replica plates contained Hybond N^+ membrane discs soaked in sterile dH_2O. The membrane discs were placed on the plates and the colonies were grown on the membrane O/N at 30°C. Membranes were then removed and subsequently placed (colony side up) for 10 minutes on a pad of filter paper soaked in denaturation solution. Membranes were removed and then placed (colony side up) for 5 minutes on a pad of filter paper soaked in neutralization solution. This neutral step was repeated, using a fresh pad of filter paper. After a 5 minute wash in 2X SSC, the membrane discs were placed (colony side up) for 30 minutes onto Whatmann 3MM filter paper to air dry. The transferred DNA was fixed with a UV-light for 5 minutes. After cross linking, the membranes were washed O/N at 65°C in 3X SSC/0.1% SDS solution. Hybridization was performed on the membranes, using a 560 bp *arsB* fragment, amplified from genomic *Sb. t.* VKM B-1269^T DNA with primer set prKaraArsBfwd and prKaraArsBrev, as probe to detect constructs containing the *arsB* gene. The composition of the stock solutions is described in The DIG System User's Guide for Filter Hybridization (Roche Molecular Biochemicals).

2.2.7 Polymerase chain reaction (PCR).

PCR amplifications were performed with the GoTaq® Flexi DNA Polymerase kit (Promega). The PCRs were carried out with a Hybaid PCR Sprint cycler. After the initial denaturation step of the double-stranded template DNA, lasting 60 s at 94°C, 25 cycles of denaturation (30 s at 94°C), an annealing step of 45 s and a final variable extension step at 72°C was performed. A cooling step of ∞ at 4°C completed the reaction. Alterations in the annealing temperatures and elongation times were dependant on the primer sets used.

2.3 Results

2.3.1 Identification and isolation of the *ars* operon of *Sb t.* VKM B-1269^T.

In order to identify putative *arsB* gene(s) within the genome of *Sb. t.* VKM B-1269^T, total DNA was isolated and amplified by making use of degenerate *arsB* primers Ferro arsBfwd and Ferro arsBrev. A PCR product of 850 bp was obtained and subsequently sequenced in both directions. Sequence alignment of this 850 bp PCR product to several other known *arsB* genes in the GenBank database, indicated the presence of certain conserved regions among the compared isolates. A set of primers (prKaraArsBfwd and prKaraArsBrev) was then designed to amplify a 560 bp fragment within the original 850 bp PCR product, to obtain a DNA fragment containing most of the conserved regions present within the various *arsB*'s.

In order to isolate the flanking regions of the *arsB* of *Sb. t.* VKM B-1269^T, a size selective mini genebank was assembled. Bacterial genomic DNA was cut with randomly selected restriction endonucleases. DNA fragments were separated on a 0.8% agarose gel and subjected to Southern-hybridization, using the 560 bp *arsB* PCR fragment as a probe. The rationale behind this step in the experiment was to determine DNA fragments containing the *arsB* gene within a certain fragment size range. A positive hybridization signal was obtained in every lane of the agarose gel, indicating the presence of the *arsB* in *Sb t.* VKM B-1269^T (Figure 2.1). Fragments digested with endonuclease *Pst*I and *Eco*RI were targeted for the construction of the mini genebank. Digested genomic DNA was run on an agarose gel and *Pst*I-*Eco*RI fragments within the 6 kb-8 kb range were cut out, purified and cloned into pUCBM21.

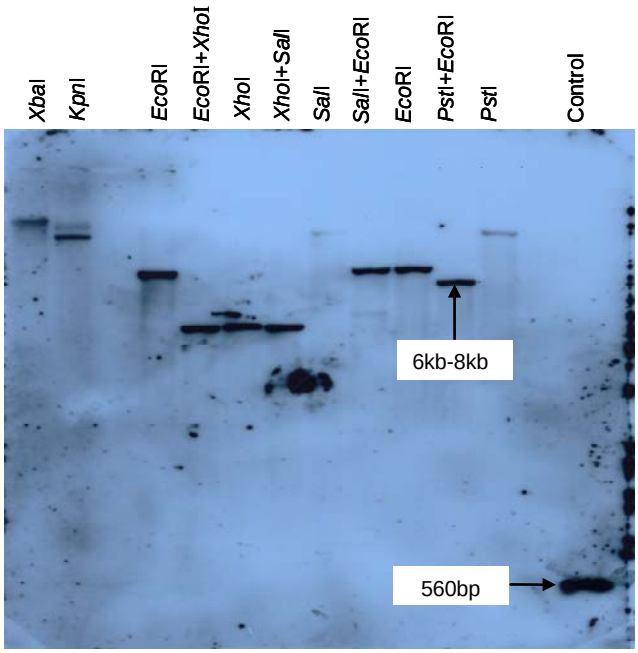


Figure 2.1: An autoradiograph following Southern-hybridization indicating the 6 kb-8 kb fragments targeted for mini-genebank construction. The result of the southern-hybridization of *Sb. t.* VKM B-1269^{T} genomic DNA, following single digestion (with *Xba*I, *Kpn*I, *Eco*RI, *Xho*I, *Sal*I and *Pst*I) and double digestion (with *Eco*RI+*Xho*I, *Xho*I+*Sal*I and *Pst*I+*Eco*RI) and probing with the 560 bp *arsB* fragment of *Sb. t.* VKM B-1269^{T}.

In an attempt to select for fragments containing the *ars* genes, constructs were transformed and tested for complementation in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant. The *ars* operon present in *Sb. t.* VKM B-1269^{T} did not confer arsenic resistance to the mutant on selective LA plates containing 0.5.mM sodium arsenite, suggesting that the *ars* was not functional in the mutant cells. Constructs containing 6 kb-8 kb *Pst*I-*Eco*RI fragments were therefore subjected to colony-hybridizations and three positive signals were obtained from a total of 3000 colonies screened. The recombinant plasmid DNA isolated from the colonies was cut with *Pst*I and *Eco*RI, and a construct yielding an insert of 5300 bp was chosen for further study.

2.3.2 Restriction mapping and sequence analysis of the *ars* genes from *Sb. t.* VKM B-1269^{T}.

The entire isolated 5300 bp fragment, called pStArs1, was subsequently mapped and completely sequenced in both directions. The complete annotated sequence of pStArs1 is

shown in Appendix 2. To confirm that *Sb. t.* VKM B-1269^T was the source of pStArs1, a Southern-hybridization experiment was performed. Two DNA fragments (1.3 kb *Hin*dIII-*Hin*dIII and 1.1 kb *Hin*dIII-*Hin*dIII fragments) covering an internal region of pStArs1 were labeled and used together as a probe against isolated *Sb. t.* VKM B-1269^T genomic DNA cut with restriction endonucleases *Bam*HI and *Hin*dIII, respectively. The Southern-hybridization experiment revealed signals that corresponded to the predicted banding patterns of pStArs1. Positive controls included in this experiment were pStArs1 digested with *Hin*dIII and subclones HH1.3 and HH1.1 digested with *Hin*dIII, respectively. Although the 1.3 kb and 1.1 kb *Hin*dIII-*Hin*dIII doublet in lane 3 (Figure 2.2 B) is not easily distinguishable as two separate fragments, the mini gene-bank was made by digesting *Sb. t.* VKM B-1269^T genomic DNA with *Pst*I and *Eco*RI, which lie either side of these two *Hin*dIII fragments. It is unlikely that a rearrangement could have occurred internal to these sites as sometimes occurs with partial *Sau*3A gene-banks. This result confirmed that pStArs1 originated from *Sb. t.* VKM B-1269^T and that pStArs1 is present in a single copy within the *Sb. t.* VKM B-1269^T genome.

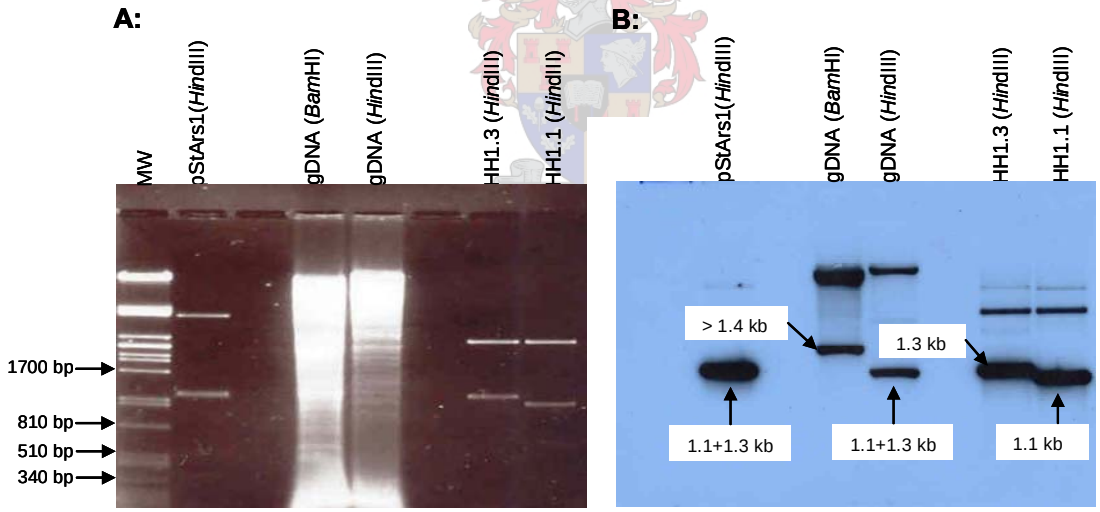


Figure 2.2: **Southern-hybridization of pStArs1 probed against genomic DNA of *Sb. t.* VKM B-1269^T. (A)** An 0.8% agarose gel showing genomic DNA cut with *Bam*HI (gDNA (*Bam*HI)) and *Hin*dIII (gDNA (*Hin*dIII)) respectively. The positive controls used in this experiment were pStArs1 cut with *Hin*dIII (pStArs1(*Hin*dIII)), subclone HH1.3 cut with *Hin*dIII (HH1.3(*Hin*dIII)) and subclone HH1.1 cut with *Hin*dIII (HH1.1(*Hin*dIII)). **(B)** Autoradiograph after hybridization using a probe containing both 1.1 kb and 1.3 kb *Hin*dIII-fragments of pStArs1. The corresponding predicted product sizes are indicated. MW indicates the molecular weight marker.

Mapping and sequencing of pStArs1 revealed five putative open reading frames. The predicted amino acid sequences of each of the identified ORFs were compared to sequences of proteins within the GenBank database, using the BLASTX program from NCBI (Figure 2.3). The characteristics of the proteins that showed highest similarity to the predicted protein products of the putative ORFs of pStArs1 are shown in Table 2.2.

Table 2.2: **The open reading frames and features of construct pStArs1**

Putative protein or ORF	Size (aa; kDa)	Most related protein and predicted size (aa; kDa)	Identity/similarity (%)	NCBI accession number
ORF1	277[#]	Glycosyl transferase domain protein	53/72	NP_952896.1
ORF2	525; 57.75	Kumamolisin-As precursor (553; 60.83)	44/62	BAC41257.1
arsR	115; 12.65	*Bacillus subtilis* AseR (111; 12.1)	50/76	CAB12340.1
arsB	440; 48.40	*Staphylococcus aureus* ArsB of plasmid pI258 (429; 47.19)	53/67	P30329
ORF5	313[#]	Nucleoside-diphosphate sugar epimerases	58/76	ZP_0030003.1

[#] these open reading frames are truncated with only the region of the protein present on pStArs1 being shown.

BLAST comparison search results identified homologues to the *arsR* (codes for a transcriptional regulator) and *arsB* (codes for an arsenite efflux pump) genes, but unlike most other *ars* operons, in which an *arsC* is located downstream of the *arsB*, this putative *ars* operon did not contain an *arsC* homologue. The *arsR* and *arsB* genes are translated in the same direction.

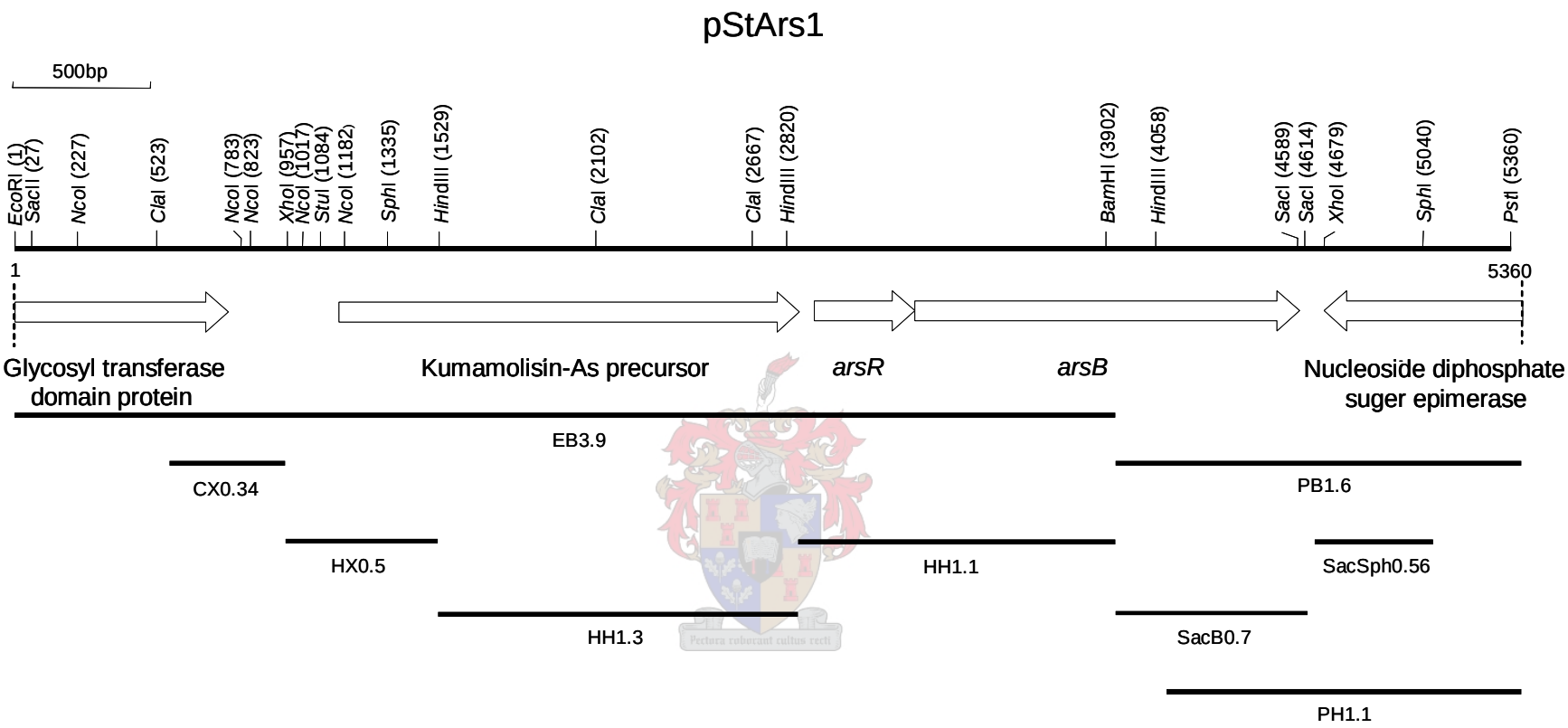


Figure 2.3: **The physical and genetic map of pStArs1.** The 5300 bp fragment containing the *ars* operon of *Sb. t.* VKM B-1269^{T}, was mapped and sequenced in both directions. More details concerning the subclones and primers used for sequencing are given in Appendix 1.

An interesting feature was the presence of a gene encoding a 525 aa Kumamolisin-As precursor. This putative proteinase was located upstream from *arsR*, with only a 77 bp intergenenic region between the termination end of the kumamolisin-As precursor and the translational start of the *arsR* gene. The 440 aa ArsB of *Sb. t.* VKM B-1269^T has the highest homology with the 429 aa ArsB of *Staphylococcus aureus* pI258 and a putative 430 aa ArsB from *Staphylococcus haemolyticus*, strain JCSC1435, exhibiting 51.5% and 51.1% sequence identity, respectively. Phylogenetic analysis of different ArsB homologues showed that the ArsB of *Sb. t.* VKM B-1269^T falls outside the group of ArsB's comprising the typical bacterial cluster of arsenite efflux pump homologues (Figure 2.4). The ArsB of *Sb. t.* VKM B-1269^T was positioned between the bacterial ArsB cluster and the archaeal ArsB of *F. acidarmanus*. Furthermore, these respective ArsB's showed 16% sequence identity to the only member of the Acr3p family of bacterial arsenite efflux pumps, located on the *skin* element of *B. subtilis*.

Using the web-based soft-ware program TMHMM (which predicts putative trans-membrane spanning domins; http://www.cbs.dtu.dk/services/TMHMM) the ArsB of *Sb. t.* VKM B-1269^T is predicted to contain 11 trans-membrane spanning regions. The location of the putative trans-membrane spanning regions of different ArsB homologues from several known *ars* operons was compared with respect to the predicted 11 trans-membrane spanning domains of the *Sb. t.* VKM B-1269^T ArsB, but no significant sequence differences within those areas could be detected (Figure 2.5). A profile constructed by the TMHMM program shows that the putative trans-membrane spanning domains and polypeptide loops of the *Sb. t.* VKM B-1269^T ArsB show similarities to that of the ArsB of both *S. aureus* plasmid pI258 and *E. coli* plasmid R773 (Figure 2.6).

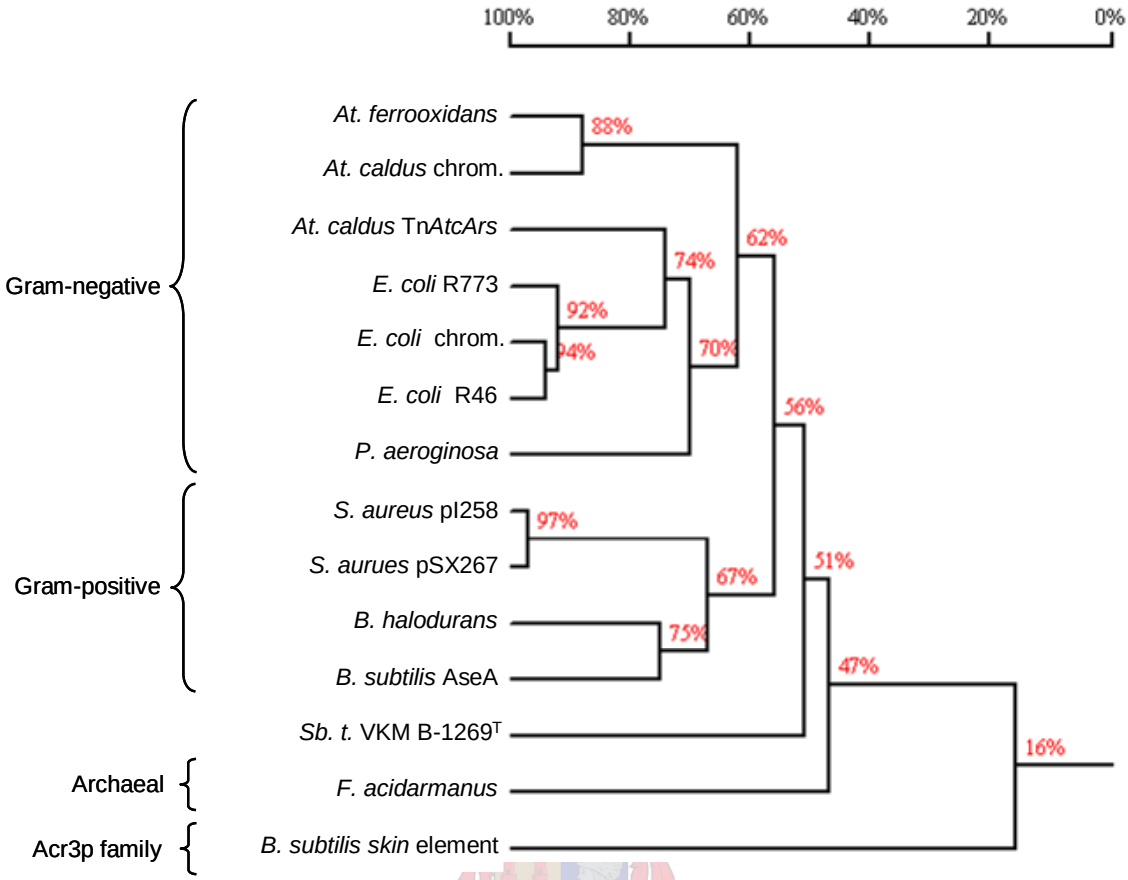


Figure 2.4: **Phylogenetic tree of a selection of different ArsB homologues from several known *ars* operons.** Accession numbers: *At. ferrooxidans*, AAF69238; *At. caldus* chromosome, ABG81354; *At. caldus* Tn*AtcArs*, AAX35679; *E. coli* plasmid R773, ARB1_ECOLI; *E. coli* chromosome, NP_417959; *E. coli* plasmid R46, NP_511239; *P. aeroginosa*, NP_250968; *S. aureus* pI258, P30329; *S. aureus* pSX267, Q01255; *B. halodurans*, NP_243865; *B. subtilis* AseA, NP_388415.1; *F. acidarmanus*, ZP_00610443.1; *B. subtilis skin* element, BAA06969.

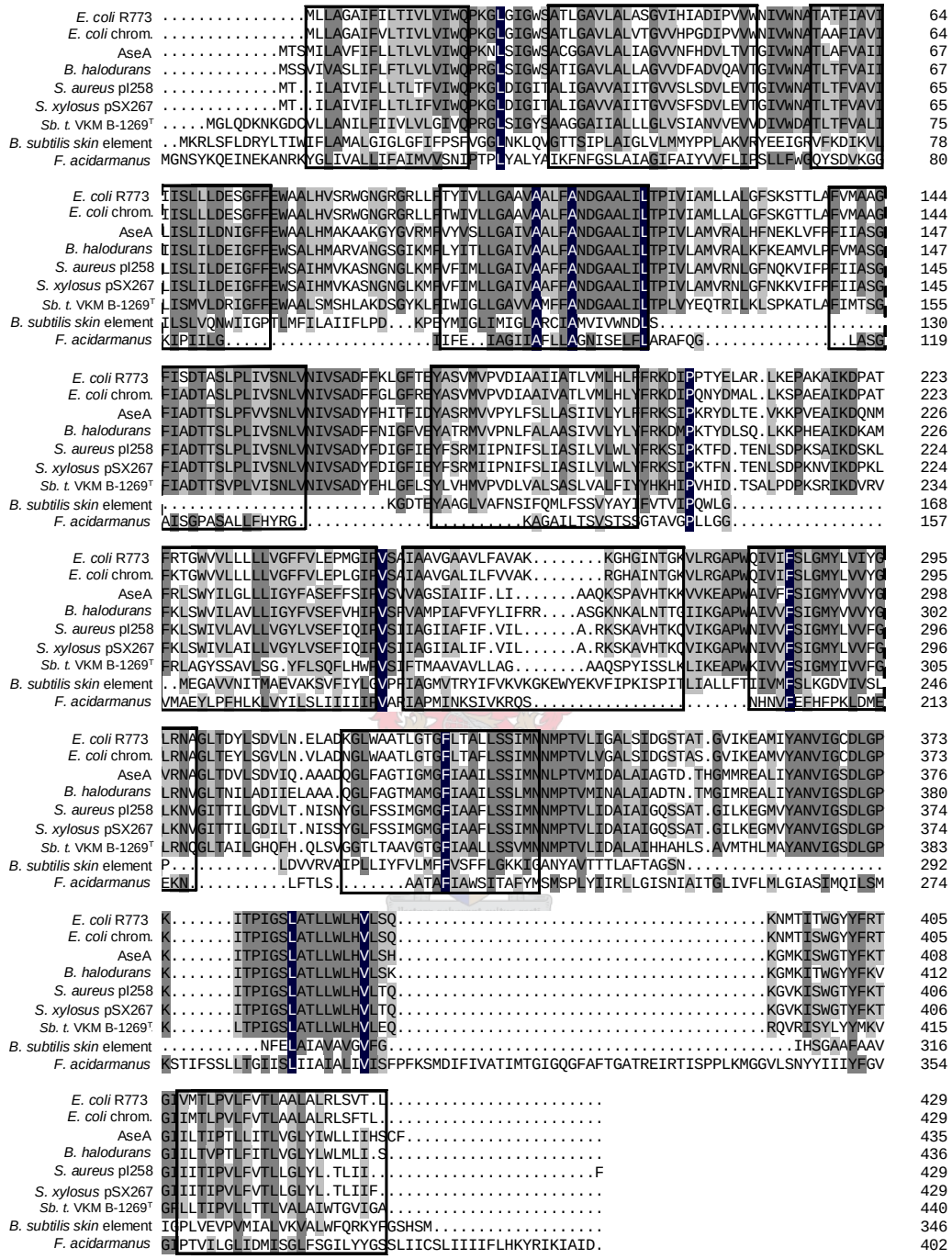


Figure 2.5: **Multiple sequence alignment comparison of a selection of different ArsB homologues to the *Sb. t.* VKM B-1269^{T} ArsB.** The putative trans-membrane spanning domains of the compared ArsB proteins with respect to the *Sb. t.* VKM B-1269^{T} ArsB were determined by the web-based TMHMM program and are indicated as boxed areas. The accession numbers of the ArsB proteins are given in Figure 2.4.

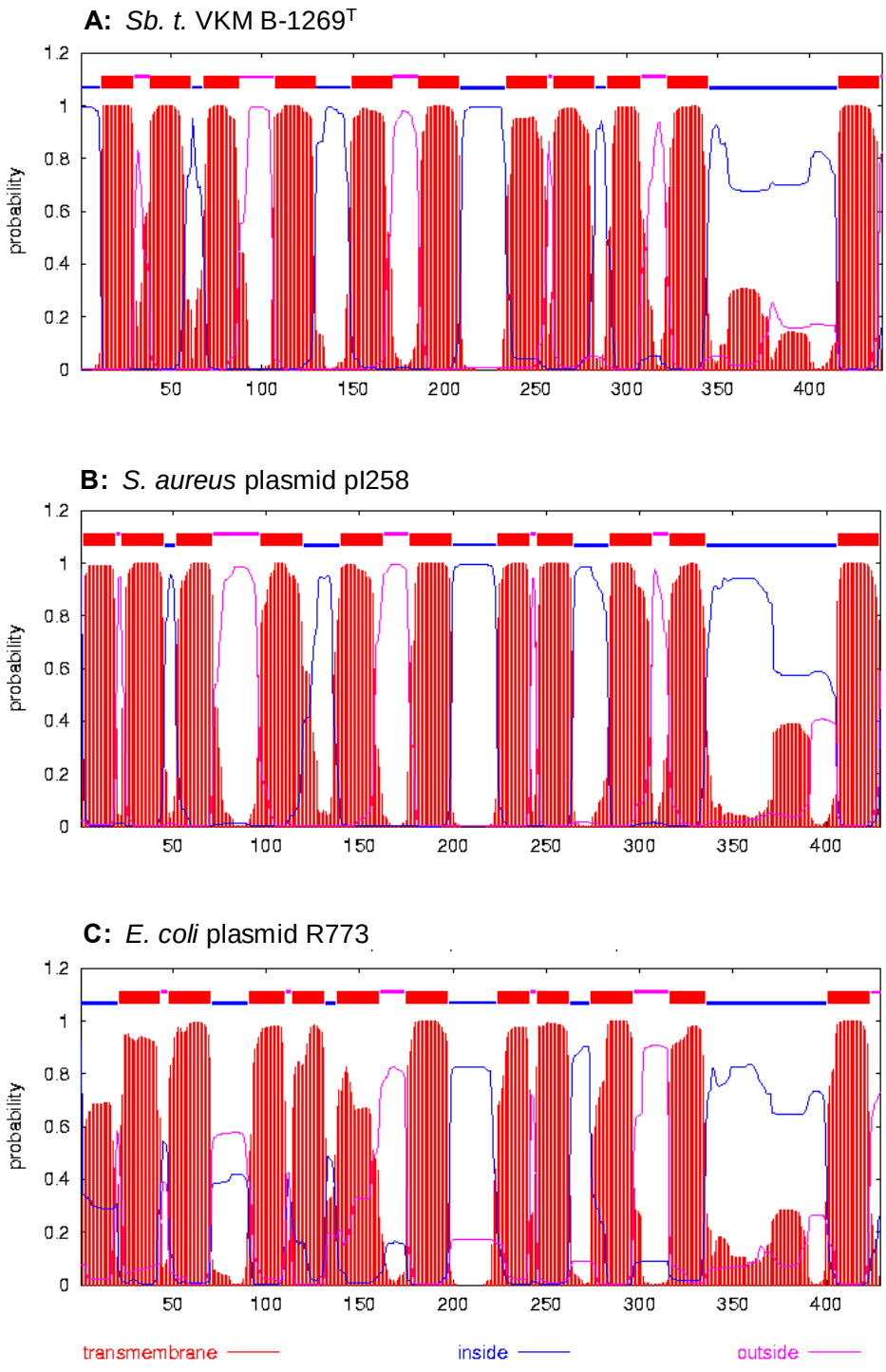


Figure 2.6: **A profile of the putative trans-membrane spanning domains and polypeptide loops of the ArsB proteins of (A) *Sb. t.* VKM B-1269^T, (B) *S. aureus* plasmid pI258 and (C) *E. coli* plasmid R773, as generated by the web-based TMHMM program.**

The BLAST results of the ArsR of *Sb. t.* VKM B-1269^{T}, indicated that this protein was most closely related to the *B. subtilis* AseR and the *B. halodurans* ArsR, sharing 34.5% and 29.5% sequence identity, respectively. Although members of the Smt/ArsR family of transcriptional regulators share low sequence identity in general, phylogenetic comparisons of the ArsR of *Sb. t.* VKM B-1269^{T} to a wide selection of ArsR homologues, indicates that it groups with the *B. halodurans* ArsR and *Bacillus subtilis* AseR, and clearly within the Gram-positive cluster (Figure 2.7).

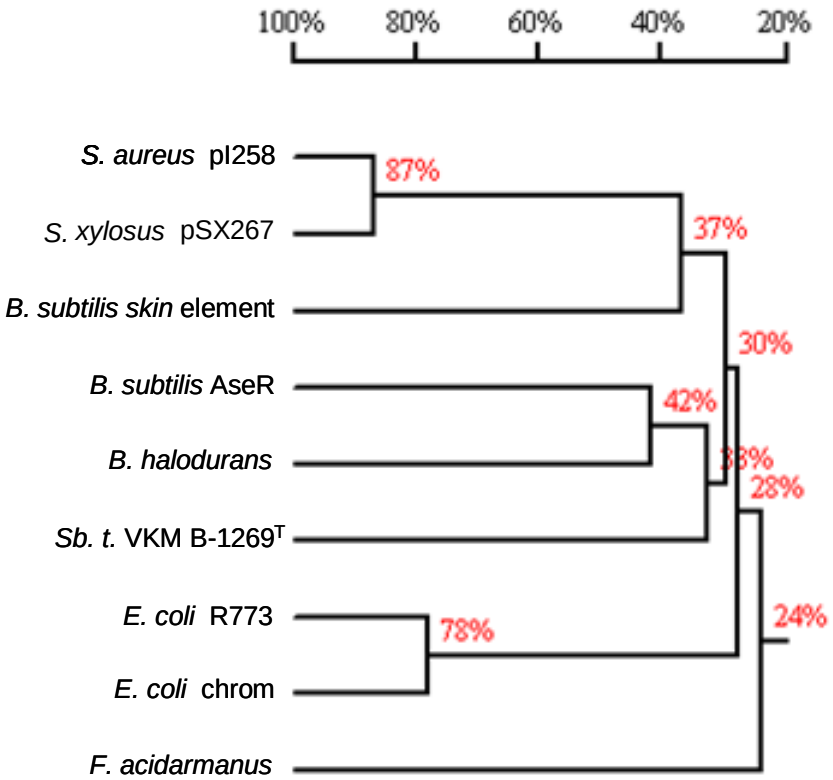


Figure 2.7: **Phylogenetic tree of ArsR proteins from well documented *ars* operons.** Accession numbers: *E. coli* chromosome, AAC76526 ; *E. coli* plasmid R773, CAA34168; *S. aureus* pI258, AAA25636; *S. xylosus* pSX267, AAA27587; *B. halodurans*, NP_243865.1; *B. subtilis* AseR, CAB12340.1; *F. acidarmanus*, ZP_00610452.1; *B. subtilis skin* element, BAA06967.

The Smt/ArsR family of metalloregulatory transcriptional repressors contains five α-helices and two β-sheets arranged as α1-α2-α3-αN-β1-β2-α5, forming DNA binding helix-turn-helix motifs and several distinct metal binding domains (Cook *et al*., 1998; Busenlehner *et al*., 2003). Comparative biochemical and spectroscopic studies performed by Eicken and coworkers (Eicken *et al*., 2003) revealed that members of the Smt/ArsR family of transcriptional regulators contain one or two structurally distinct metal binding sites, denoted α3N and α5 respectively (Figure 2.8). The highly conserved $ELC_{32}V(G/C_{34})DL$ sequence is present in the α3N metal binding site, as part of the proposed α3-turn-αR DNA binding motif. The α5 metal binding site is located in the α5 helix and contains a LVAYLTENCC conserved sequence (Figure 2.8; 2.9). The Smt/ArsR family of transcriptional regulators could be divided into two separate subgroups on the basis of differences in the location of their metal binding sites. One subgroup uses only the $ELC_{32}V(G/C_{34})DL$ binding motif (e.g. ArsR of *E. coli* plasmid R773), while the second subgroup (e.g. the chromosome-located ArsR of *At. ferrooxidans*) uses a conserved GX(L/I)A metal binding motif adjacent to the conserved $ELC_{32}V(G/C_{34})DL$ sequence and the LVAYLTENCC metal binding motive located on the α5 helix (e.g. ArsR of *At. ferrooxidans*) (Shi *et al*., 1994; Butcher and Rawlings, 2002; Busenlehner *et al*., 2003).

The multiple sequence alignment of different ArsR homologues to the *Sb. t.* VKM B-1269^T ArsR revealed the presence of an EYCNCEF sequence within the projected area of the conserved $ELC_{32}V(G/C_{34})DL$ sequence. In addition, no GX(L/I)A or LVAYLTENCC sequences were evident from the sequence alignment comparisons. Secondary structure prediction of the *Sb. t.* VKM B-1269^T ArsR (conducted by the web-based program PsiPRED v2.4; http://www.predictprotein.org) showed that the ArsR contained five α-helices and two β-sheets arranged as α1-α2-α3-αN-β1-β2-α5 (Figure 2.9).

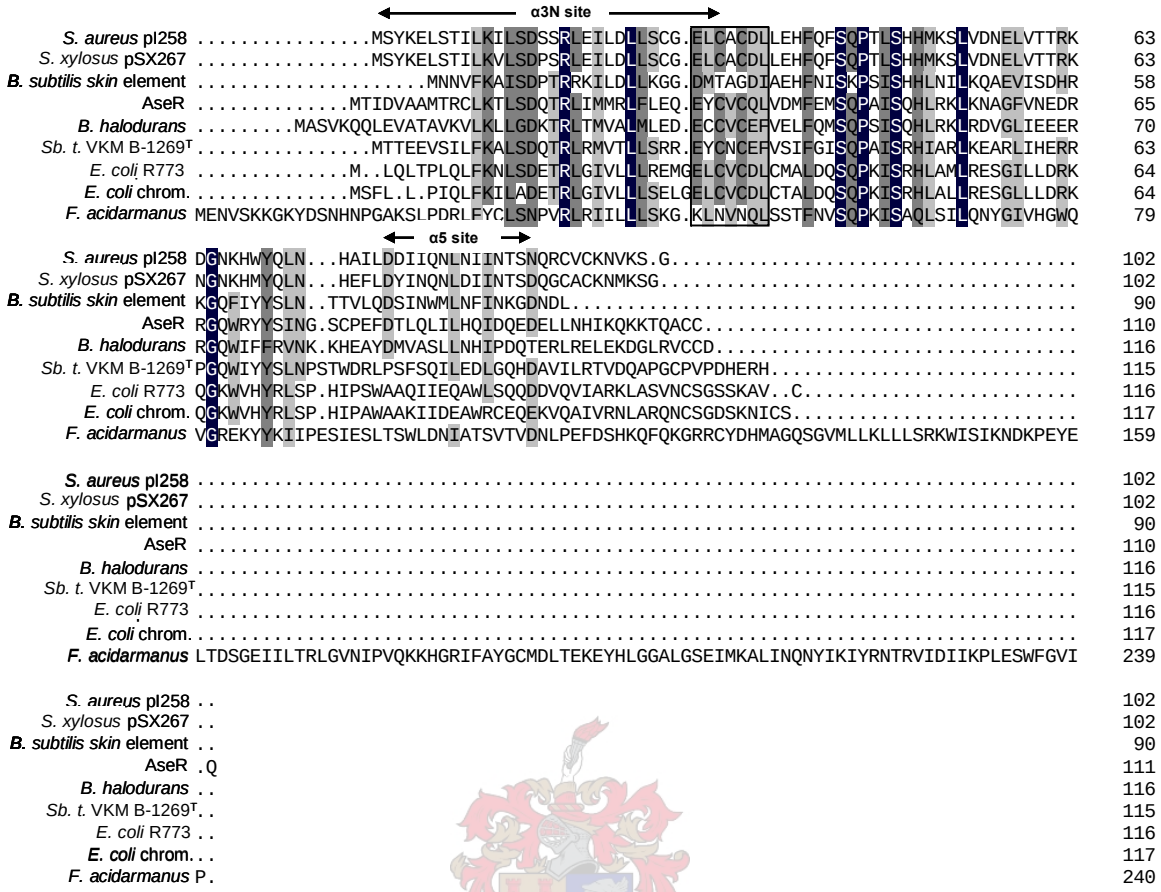


Figure 2.8: **Multiple sequence alignment comparison of different ArsR homologues to the *Sb. t.* VKM B-1269^T ArsR**. The proposed α3N and α5 metal binding regions are indicated above the alignment, with the conserved $ELC_{32}V(G/C_{34})DL$ consensus sequence located within the boxed area in the α3N site. The accession numbers of the ArsR proteins are given in Figure 2.7.

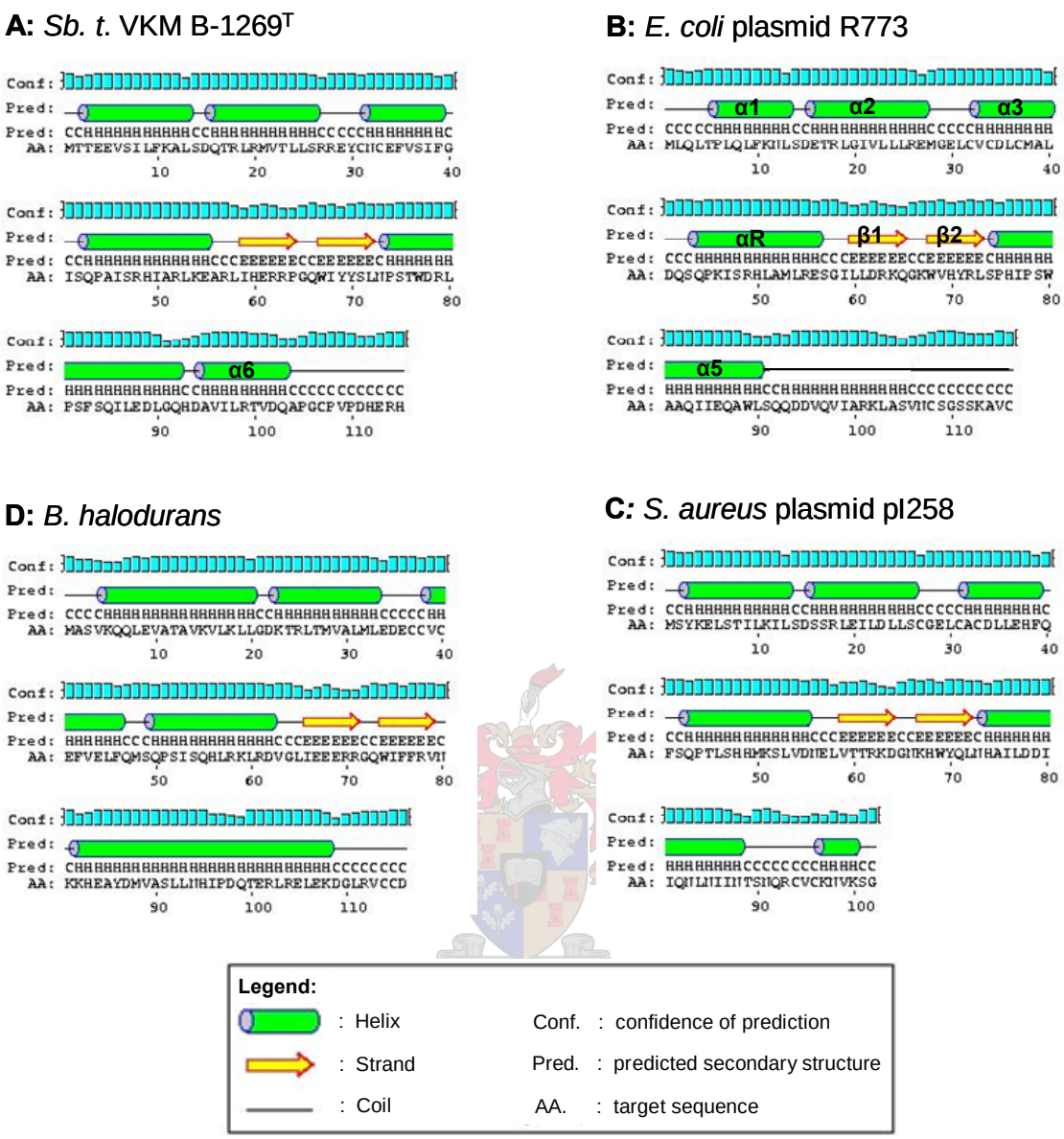


Figure 2.9**: Secondary structure profiles of ArsR proteins of (A) *Sb. t.* VKM B-1269^T; (B) *E. coli* plasmid R773; (C) *B. halodurans* and (D) *S. aureus* plasmid pI258, as predicted by the PsiPRED web-based program.** The order of arrangement of the five α-helices and the two β-sheets are indicated on the *E. coli* plasmid R773 profile, with the additional α-helix (α6) indicated on the *Sb. t.* VKM B-1269^T profile.

2.3.3 Detection of putative *arsC* gene homologues in *Sb. t.* VKM B-1269^T.

No representatives of either the Grx/GSH or Trx ArsC family was present in the proximity of the *arsRB* operon of *Sb. t.* VKM B-1269^T. In the quest to identify an ArsC, the sequences of representatives of the Trx ArsC family were aligned in order to design ArsC degenerate PCR primers (prArsCfwd deg and prArsCrev deg) within conserved regions (Figure 2.10). If a putative *arsC* gene were to be present in the genome of *Sb. t.* VKM B-1269^T, the sequence would most likely be closest to those encoding the Trx family of ArsC proteins, as they are usually associated with Gram-positive bacterial *ars* operons. DNA amplification using primers prArsCfwd deg and prArsCrev deg with *Sb. t.* VKM B-1269^T genomic DNA was not successful even after the annealing temperature times, elongation times and $MgCl_2$ concentrations were altered (results not shown). The predicted product was obtained when the *S. aureus* plasmid pI258 was used as a positive control during this experiment.

In addition, several Southern-hybridization experiments were performed in order to detect a putative ArsC protein. DNA fragments representative of the Grx/GSH ArsC family (a subclone of the *E. coli* plasmid R773 *arsC*) and the Trx ArsC family (a fragment amplified from the *S. aureus* plasmid pI258 *arsC* and the *arsC* of *At. caldus TnAtcArs*, respectively) were labeled and probed against isolated genomic DNA of *Sb. t.* VKM B-1269^T. No hybridization signal was obtained, indicating the absence of an *arsC* homologue (results not shown). Plasmid pUCarsCdeg was used as a positive control during this experiment.

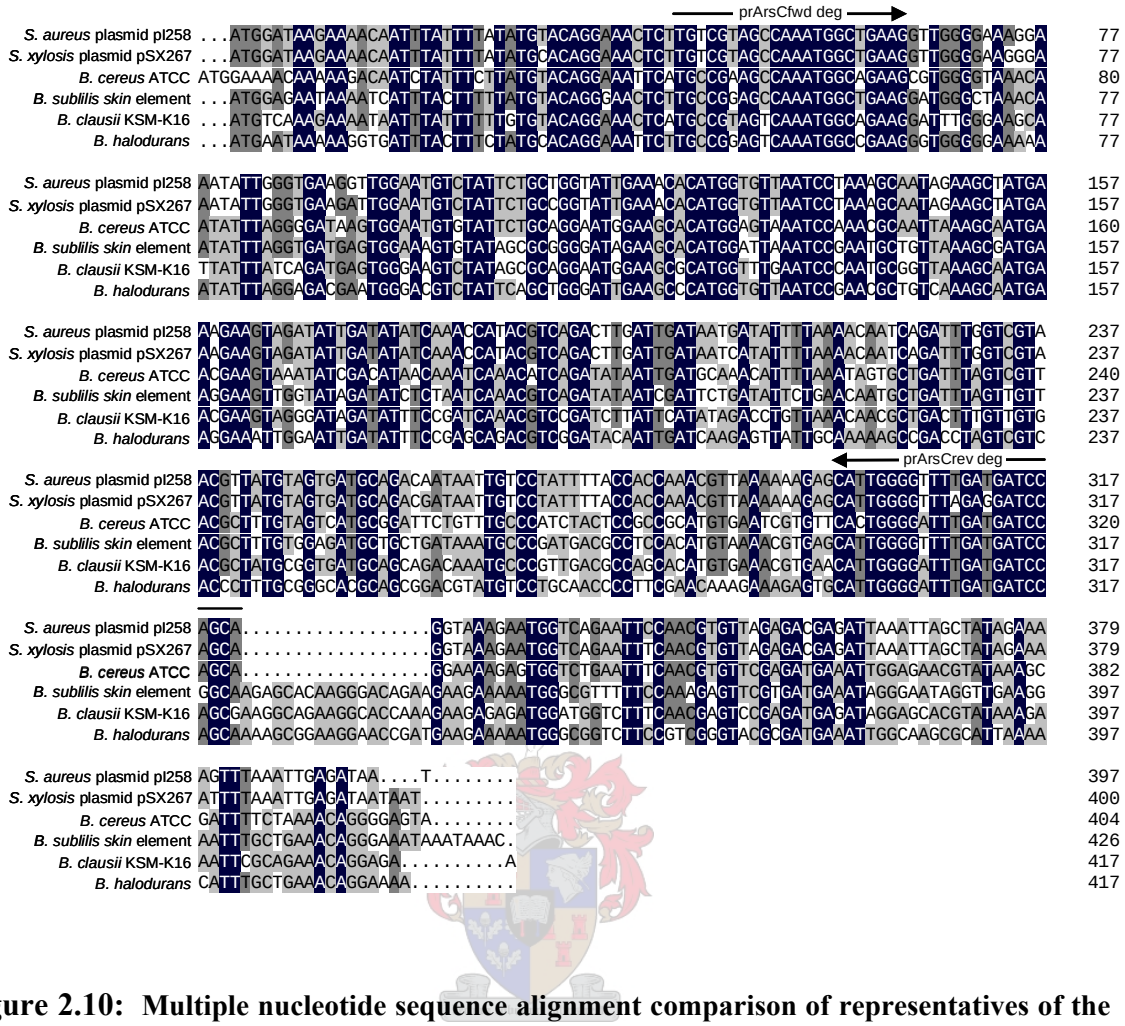


Figure 2.10: Multiple nucleotide sequence alignment comparison of representatives of the Trx ArsC family in order to identify shared conserved sequences. Accession numbers: *S. aureus* pI258, A53641; *S. xylosus* pSX267, C49102; *B. cereus* ATCC14579, NP_832893.1; *B. subtilis skin* element, NC_000964.2; *B. clausii* KSM-K16, AP006627.1; *B. halodurans*, NP_243864.1. The positions of the degenerate primers are indicated.

2.3.4 Minimum inhibitory arsenic concentrations of *Sb. t.* VKM B-1269^T.

Sb. t. VKM B-1269^T was tested for its ability to grow in the presence of different As(V) and As(III) concentrations in comparison with other acidophilic microorganisms (Table 2.3). Cells were inoculated in flat-bottom shake flasks containing growth media and 5, 10, 20, 30 and 50 mM As(V), and 50, 100, 250, 500 and 1000 μM As(III), respectively. An iron-containing precipitate formed after only 30 minutes of incubation in flasks

containing either As(V) or As(III). The $FeSO_4$ concentration and pH values of the Sulfobacilli growth media were reduced in an attempt to limit the rate of precipitate formation, but did not have any significant effect. The formation of this iron-containing precipitate or the presence of arsenic had a significant influence on the ability of the *Sb. t.* VKM B-1269^T cells to grow, making it difficult to draw any conclusions about the tested As(V) and As(III) MIC values. No precipitate formed in uninoculated controls, suggesting that the presence of *Sb. t.* VKM B-1269^T cells in the growth media played a fundamental role in the formation of the precipitate.

Table 2.3: **The upper level of some heavy metal concentrations where metabolic activity has been reported in the listed neutrophilic and acidophilic microorganisms**. Adapted from Dopson *et al*., 2003.

Microorganism	Maximum metal concentration whereby metalbolic functioning continues (mM)				
	As (III)	Cu (II)	Zn (II)	Cd (II)	Ni (II)
Neutrophilic bacterium					
E. coli	4	1	1	0.5	1
Acidophilic bacterium					
At. ferrooxidans	84	800	1071	500	1000
Sb. thermosulfidooxidans	*ND	6	43	*ND	5
Acidophilic archaea					
F. acidarmanus	13	16	*ND	9	*ND

* ND: Not determined

2.4 Discussion

Moderately thermophilic bacteria representative of the genus *Sulfobacillus*, have great potential of being present in biomining processes that operate at elevated temperatures (40°C-55°C). The microbial catalysed oxidation of arsenopyrite ores and concentrates produces large quantities of arsenic in the leaching solution. It was thus of interest to gain information about the molecular biology of genus *Sulfobacillus* with respect to

resistance to arsenical compounds. The *ars* operon of *Sb. t.* VKM B-1269^T has been isolated and consists of an *arsR* (codes for a transcriptional regulator) and *arsB* (codes for a arsenite efflux pump). The *arsRB* were transcribed in the same direction. Although this *ars* operon seems minimalist in comparison with the commonly encountered *arsRBC* and *arsRDABC* forms, two other cases of *ars* operons containing this two-gene configuration have been reported to exist in the bacterium *B. subtilis* (Rosen,1999) and the archaeon *Ferroplasma acidarmanus* (Gihring *et al*., 2003). An interesting feature about the *ars* operons of *Sb. t.* VKM B-1269^T was the presence of a gene encoding a 525 aa (60.83 kDa) kumamolisin-As precursor located upstream of the *arsR* gene. Kumamolysin is generally expressed as a 64 kDa precursor, which is autocatalytically converted to a 43 kDa active enzyme. Kumamolisin is a member of a rapidly growing family of endopeptidases. The association of kumamolisin with the *ars* genes of *Sb. t.* VKM B-1269^T and its possible involvement with respect to arsenic resistance is unclear and needs to be determined.

Results obtained from BLAST searches of the NCBI database showed that the ArsB of *Sb. t.* VKM B-1269^T has highest homology to the ArsB on plasmid pI258 of *S. aureus*, with 51.5% aa sequence identity. Phylogenetic comparisons indicated that the *Sb. t.* VKM B-1269^T ArsB has a much longer branch length with respect to the bacterial cluster of ArsB's, suggesting that this protein is not closely related to the aligned bacterial ArsB proteins (shares only 56% sequence identity with the bacterial ArsB homologues). The fact that this ArsB is unexpectedly positioned between the bacterial and archaeal ArsB groups strongly suggests that *Sb. t.* VKM B-1269^T contains an *arsB* with a different ancestory to other bacteria (Figure 2.4). This might explain the inability of this *ars* operon to confer resistance to arsenate or arsenite in *E. coli* ACSH50^{Iq}. A web-based software program that predicts trans-membrane spanning domains, profiles suggest that the ArsB of *Sb. t.* VKM B-1269^T contains 11 trans-membrane regions, similar to those of *S. aureus* pI258 and *E. coli* plasmid R773 (Figure 2.6).

Phylogenetic comparisons of the ArsR of *Sb. t.* VKM B-1269^T to other ArsR homologues showed that it groups with the *B. halodurans* ArsR and *B. subtilis* AseR and forms a clear

separate subgroup within the Gram-positive bacterial ArsR cluster (Figure 2.7). However, the *Sb. t.* VKM B-1269^T ArsR is the least similar to the other two members of this subgroup, as it contains the longest branch-length within the cluster. As was mentioned earlier, the SmtB/ArsR family of transcriptional regulators can be divided into two separate subgroups on the basis of differences in the location of their metal-binding sites, composed of conserved sequence regions on the α3 and α5 helices of the protein. Although the conserved $ELC_{32}V(G/C_{34})DL$ sequence usually acts as the metal binding motive on the α3 helix, multiple sequence alignment of the *Sb. t.* VKM B-1269^T ArsR to other ArsR proteins revealed that the ArsR contained an EYCNCEF sequence in the proximity of its α3 helix (Figure 2.8). A Cys-32 and Cys-34 residues were identified within this consensus sequence and may probably be involved in the metal binding process, but unlike the *E. coli* plasmid R773 $ELC_{32}V(G/C_{34})DL$ sequence, no Cys-37 representative was present in the immediate vicinity of the EYCNCEF sequence (Shi *et al*., 1994; 1996). Furthermore, no GX(L/I)A sequence, usually located adjacent to the $ELC_{32}V(G/C_{34})DL$ sequence or the α5 associated LVAYLTENCC sequence was detected, showing that the ArsR of *Sb. t.* VKM B-1269^T is a member of the subgroup within the SmtB/ArsR family of transcriptional regulators represented by the *S. aureus* CadC, *Listeris momcytogenes* CadC and the *E. coli* ArsR. Secondary structure of the *Sb. t.* VKM B-1269^T ArsR was predicted by using the web-based program PsiPRED v2.4 and showed that the ArsR contained five α-helices and two β-sheets arranged as α1-α2-α3-αN-β1-β2-α5 as proposed by Busenlehner and coworkers (Busenlehner *et al*., 2003). *Sb. t.* VKM B-1269^T ArsR contained an additional α-helix (α6) on the C-terminal side of the protein in comparison to the secondary structure profile of the *E. coli* plasmid R773 ArsR (Figure 2.9).

The absence of an *arsC* (arsenate reductase) in this operon was surprising as *arsC* genes are usually associated with *ars* operons. The gene product of *arsC* plays a fundamental role in the resistance mechanism of arsenic as it reduces arsenate As(V) to arsenite As(III). The possibility that an *arsC* homolog might be located somewhere else in the *Sb. t.* VKM B-1269^T genome could not be excluded as $ArsC_{ec}$ homologues lacking other *ars* genes have been reported in *Haemophilus influenzae* and *Neisseria gonorrhoeae* (Rosen,

1999). PCR and Southern-hybridization experiments revealed that no ArsC, representative of either the Grx/GSH or Trx ArsC families was present in the genome of *Sb. t.* VKM B-1269^T.

The ability of this *arsRB* operon to confer resistance to As(V) and As(III) in *Sb. t.* VKM B-1269^T was examined by attempting to calculate MICs in liquid batch cultures. Unfortunately several problems were encountered during the course of these experiments due to the formation of an iron-containing precipitate. The formation of this iron-containing compound had significant influence on the *Sb. t.* VKM B-1269^T cell growth. The microbial-aided oxidation of arsenopyrite involves a complex network of integral biochemical reactions that may lead to the formation of iron-containing compounds like ferric arsenate ($FeAsO_4$) and jarosite ($KFe_3(SO_4)_2(OH)_6$) (Mandl *et al*., 1992; Tuovinen *et al*., 1994). The rate of production of these iron-containing precipitates is dependent on variables such as the oxidation state of the chemicals present in the solution, metabolite production and the presence of certain biomass components (Breed *et al*., 1996). No significant conclusions with respect to the ability of the *Sb. t.* VKM B-1269^T *ars* operon to confer resistance to arsenical compounds could be drawn. If the precipitation problem could be solved by using alternative growth media, determining to what extent this *arsRB* operon contributes to arsenic resistance within *Sb. t*. VKM B-1269^T will still be problematic. This is largely due to the lack of suitable arsenic sensitive *Sb. t.* mutants with which to perform comparative analysis. In addition, the current limited knowledge with respect to the different arsenic resistance tolerance levels in members of genus *Sulfobacillus* limits preliminary comparisons of MIC's in liquid batch cultures.

CHAPTER 3

THE EXPRESSION AND REGULATION OF THE ARSENIC RESISTANCE GENES OF *SULFOBACILLUS THERMOSULFIDOOXIDANS* VKM B-1269^T

TABLE OF CONTENT

3.1 Introduction 65

3.2 Materials and Methods

3.2.1 Bacterial strains, plasmids and PCR primers 66

3.2.2 Media and growth conditions 66

3.2.3 DNA isolation, techniques and analysis 68

3.2.4 Total mRNA extraction and purification 69

3.2.5 RT-PCR analysis of mRNA 69

3.2.6 Polymerase chain reaction (PCR) 70

3.2.7 Construction of the promoter-*lacZ* reporter constructs 70

3.2.8 β-galactosidase assays 72

3.3 Results

3.3.1 The *St. t.* VKM B-1269^T *ars* operon does not complement the *E. coli* ACSH50^{Iq} arsenic sensitive mutant 73

3.3.2 Expression analysis of the *ars* operon of *Sb. t.* VKM B-1269^T in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant 73

3.3.3 Expression analysis of the *ars* operon of *Sb. t.* VKM B-1269^T in *Sb.*. 78

3.3.4 Reporter-gene studies of the putative promoter regions of the kumamolisin-As precursor gene and *arsR* in *E. coli* 80

3.4 Discussion 81

3.1. Introduction

Recent advances in the fundamental understanding and development of genetic tools have contributed immensely to the successful expression of recombinant proteins in heterologous hosts. The successful heterologous expression of proteins and enzymes involves two major steps: (1) The introduction of the foreign DNA into the host cells and (2) the nature of factors effecting the stability and expression of foreign DNA for protein synthesis in the chosen expression system (Rai and Padh, 2001). Although many alternative organisms and expression systems are being used for recombinant protein expression, *E. coli* offers a reliable prokaryotic system for the production of foreign proteins.

E. coli is an attractive host for the production of heterologous proteins due to the vast knowledge about its genetics, physiology and complete genomic sequence. In addition, *E. coli* can grow rapidly to high densities in simple and inexpensive media. However, in spite of the extensive knowledge on the physiology and molecular biology of *E. coli*, there is no guarantee that every gene can be expressed in a full-length and biologically active form in this bacterium (Baneyx, 1999). This is due to several factors which may include unique and subtle structural features of the gene sequence, the stability and translational efficiency of mRNA, the effectiveness of protein folding, the degradation of the protein by host cell proteases, differences in codon usage between the foreign gene and the native *E. coli* and the potential toxicity of the protein to the host. These factors are varied and at times poorly understood (Jana and Deb, 2005).

Plasmid-determined resistances to arsenic and antimony have been found in both Gram-negative and Gram-positive bacteria. The plasmid-encoded arsenic/antimony resistance mechanism seems essentially the same in staphylococcal species and *E. coli*, with both organisms governing an inducible resistance to arsenite, arsenate and antimony (Silver *et al*., 1981). Expression studies performed on the *ars* genes of staphylococcal plasmids pI258 and pSX267 have been successfully conducted in *E. coli.* Although similarities with respect to *ars* operon expression and function in the heterologous host were shown,

quantitative differences in induced levels of resistance were detectable (Rosenstein *et al*., 1992; Broër *et al*., 1993).

Here I confirm that *arsRB* mRNA transcripts were produced in the heterologous *E. coli* host, despite the fact that the *arsRB* of *Sb. t*. VKM B-1269^{T} could not confer resistance to arsenate or arsenite in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant. We showed that the kumamolisin-As precursor gene is co-transcribed with the *arsRB* genes in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant. In contrast to the heterologous *E. coli* host, the *arsRB* operon of *Sb. t.* VKM B-1269^{T} was not co-transcribed with the kumamolisin-As precursor gene in its native *Sulfobacillus* host. Promoter expression studies performed in *E. coli* revealed the relative activity strengths of the kumamolisin-As precursor and *arsR* promoters.

3.2 Materials and Methods

3.2.1 Bacterial strains, plasmids and PCR primers.

The bacterial strains, plasmids and primers used in this study are described in Table 3.1.

3.2.2 Media and growth conditions.

E. coli cells were grown aerobically on LA plates or in LB at 37°C as described by Sambrook *et al.*, 1989. Amp (100 μg/ml) and was added for selection when required.

TABLE 3.1: **Bacterial strains, plasmids and PCR primers used in this study**

Strains, plasmids or primers	Genotype or description	Reference or source
Strains		
Escherichia coli		
DH5α	*F'/endA1 hsdR17* ($r_K^-m_K^+$) *supE44 thi-1recA1 gyrA* (Nal^r) *relA1 Δ(lacZYA-ArgF) U169 (Φ80dlac Δ(lacZ) M15)*	Promega Corp.
XL1-Blue	F'::Tn*10* $proA^+B^+$ $lacI^q$ *Δ(lacZ) M15/recA1 endA1*	Promega Corp.

ACSH50^{Iq}	*gyrA96* (Nal^r) *thi hsdR17*($r_K^-m_K^+$) *supE44 relA1 lac rpsL* Δ(*lac –pro*) Δ*ars*::*cam*	Butcher and Rawlings, 2002
Sulfobacillus thermosulfidooxidans VKM B-1269$^{T☼=}$	Russia	Golovacheva and Karavaiko, 1978
Plasmids		
pBluescript SK$^+$	Ap^r; *lacZ'*; ColE1 replicon vector	Stratagene
pUCBM21	Ap^r; *lacZ'*; ColE1 replicon vector	Boehringer-Mannheim
pMC1403	Ap^r; ColE1 replicon,	Casadaban *et al*, (1983)
pKK223-3	Ap^r; *tac* promoter ColE1 replicon vector	Pharmacia Biotech
pGEM-T®	Ap^r; T-tailed PCR product cloning vector	Promega
pStArs1	Ap^r; 5300 bp *Eco*RI-*Pst*I fragment, containing the *ars* operon of *Sb. t.* VKM B-1269^T, cloned into pUCBM21 digested with *Eco*RI and *Pst*I	This study
p*tac*ArsRB	Ap^r; 2694 bp *Cla*I(blunt)-*Pst*I fragment of pStArs1 cloned into pKK223-3 digested with *Sma*I and *Pst*I	This study
pArsRB	Ap^r; 2620 bp *Cla*I-*Xho*I fragment of pStArs1 cloned into pSK$^+$ digested with *Cla*I and *Xho*I. This construct was then digested with *Xba*I and *Kpn*I and cloned into pUCBM21 digested with *Xba*I and *Kpn*I	This study
pMCStKumpr400	Ap^r; PCR product of the putative kumamolisin-As precursor promoter (843-1255 bp) obtained with primers prKumlacfwd400/prKumlacrev, cloned into pMC1403	This study
pMCStKumppr800	Ap^r; PCR product of the putative kumamolisin-As precursor promoter (449-1255 bp) obtained with primers prKumlacfwd800/prKumlacrev, cloned into pMC1403	This study
pMCStArsRpr	Ap^r; PCR product of *arsR* promoter (2403-2881 bp) obtained with primers prArsRlacfwd/prArsRrev, cloned into pMC1403	This study

Primers[#]

prKumlacfwd400 (*Eco*RI)	5'-GTACGGAATTCCACGCTCACTTTGGCTCG-3'	This study
prKumlacfwd800 (*Eco*RI)	5'-GTACGAATTCGGGGAAGACTACCGTAAACG-	This study
prKumlacrev (*Bam*HI)	3'	This study
prArsRlacfwd (*Eco*RI)	5'-GTACGGATCCACTAGTTTCGCCTGCATCG-3'	This study
prArsRlacrev (*Bam*HI)	5'-GTACGAATTCTCAGGTATTTGCGGTCCCG-3'	This study
prArsRfwd (*Bam*HI)	5'-GTACGGATCCTCTTCGGTGGTCATTGGTCC-3'	This study
prArsRrev (*Pst*I)	5'-GTACGGATCCGATTATGAATATATCAATG-3'	This study
prKum-Asc	5'-GTACCTGCAGGATAAATAAGATATTAGC-3'	This study
prRTKum-As	5'-GGATGTGAATAATAACGGAGC-3'	This study
prArsRc	5'-GTGCAATCATAACCTGTTACC-3'	This study
prRTArsR	5'-CGTTGAGTGATCAAACAAGG-3'	This study
prArsBc	5'-CTACAGTTCTAAGAATGACCG-3'	This study
prRTArsB	5'-GCAAGATAAGAACAAAGGG-3'	This study
	5'-CCATCGAAATGAGAATAAGCG-3'	

Ap^r: ampicillin resistance

[#] restriction endonuclease sites incorporated into primers are indicated in parenthesis and are underlined in the primer sequence

3.2.3 DNA isolation, techniques and analysis.

Restriction endonuclease digestions, gel electrophoresis, small-scale plasmid preparation, ligation reactions and Southern-blot hybridization were performed using standard methods (Sambrook *et al.*, 1989). Large-scale plasmid preparation was done by means of the alkaline lysis procedure whereafter plasmid DNA was purified by CsCl/etbr centrifugation (Current Protocols in Molecular Biology) or by using the Nucleobond AX100 system (Macherey-Nagel). DNA fragments to be used for cloning purposes were extracted from agarose gels with the GFX^{TM} kit (Amersham BioSiences). Clones to be sequenced were isolated and purified from 5 ml LB O/N culture using the High Pure Plasmid Isolation Kit (Roche Molecular Biochemicals). Sequencing of constructs was carried out by using the dideoxy chain termination method and was executed by an ABI PRISM^{TM} 377 automated DNA sequencer. Sequences were analysed using several software programmes, but mainly by the PC based DNAMAN (version 4.1) package (Lynnon Biosoft). The labeling of probes, hybridization and detection was conducted by using the dioxigenin-dUTP nonradioactive DNA labeling and detection kit (Roche Molecular Biochemicals). When Southern-blot hybridization was performed with a

homologous DNA probe, prehybridization and probing were done at 42ºC and the buffer B washes were conducted at 65 ºC. In the cases where a heterologous DNA probe was utilized in Northern-blot hybridization, the prehybridization and probing steps were carried out at 37ºC and the buffer B washes performed at 52 ºC. Quantitative measurements of the Slot-blot autoradiographs were done by using the UVIgeltec Version 12.4 for Windows software program (UVItec).

3.2.4 Total mRNA extraction and purification.

Total RNA from 50 ml of mid-exponential-phase cultures of *E. coli* ACSH50I^{q} containing various plasmids was isolated as previously described (Trindale *et al*., 2003). *E. coli* ACSH50I^{q} cells were grown in LB media containing the appropriate antibiotic and 0 μM or 25 μM arsenite. Total RNA was also isolated from *Sb. t.,* VKM B-1269^{T}. Isolated RNA was transferred to a Hybond-N^{+} nylon membrane (Amersham) using a Slot-blot Minifold® II grid from Schleicher & Schuell by standard procedures (Sambrook *et al*., 1989) and hybridized according to the manufacture's recommendations, using digoxigenin labeled DNA probes specific for the kumamolisin-As precursor and *arsB* transcripts. The labeling of probes and detection was conducted by using the dioxigenin-dUTP nonradioactive DNA labeling and detection kit (Roche Molecular Biochemicals).

3.2.5 RT-PCR analysis of mRNA.

For RT-PCR, the 1st Strand cDNA synthesis kit (AMV; Roche Molecular Biochemicals) was used for cDNA synthesis and cDNA product detection. The reverse-transcriptase reactions were conducted according to the manufacturer's instructions. The PCR was performed as described above, using 1.5 μl (~125 ng) of the 20 μl (total volume) reverse-transcriptase reaction. The extension times were altered as required for the different primer pairs. Primers prRTKum-As, prRTArsR and prRTArsB were used for the synthesis of cDNA from kumamolisin-As precursor, *arsR* and *arsB* mRNA, respectively. For the PCR, the following primer sets were used in combination: prKum-Asc/prRTArsR, prKum-Asc/prRTArsB and prArsRc/prRTArsB. As a positive control for kumamolisin-As precursor, *arsR* and *arsB* cDNA synthesis, primer sets prKump-Asc/prRTKump-As, prArsRc/prRTArsR and prArsBc/prRTArsB were used in the PCR,

respectively (Figure 3.1). In order to determine whether mRNA was contaminated with DNA, PCR reactions were performed with each primer pair without any AMV reverse transcriptase present in the preceding step.

3.2.6 Polymerase chain reaction (PCR).

The specifications of the performed PCR amplifications are described in section 2.2.7 of Chapter 2 (page 44).

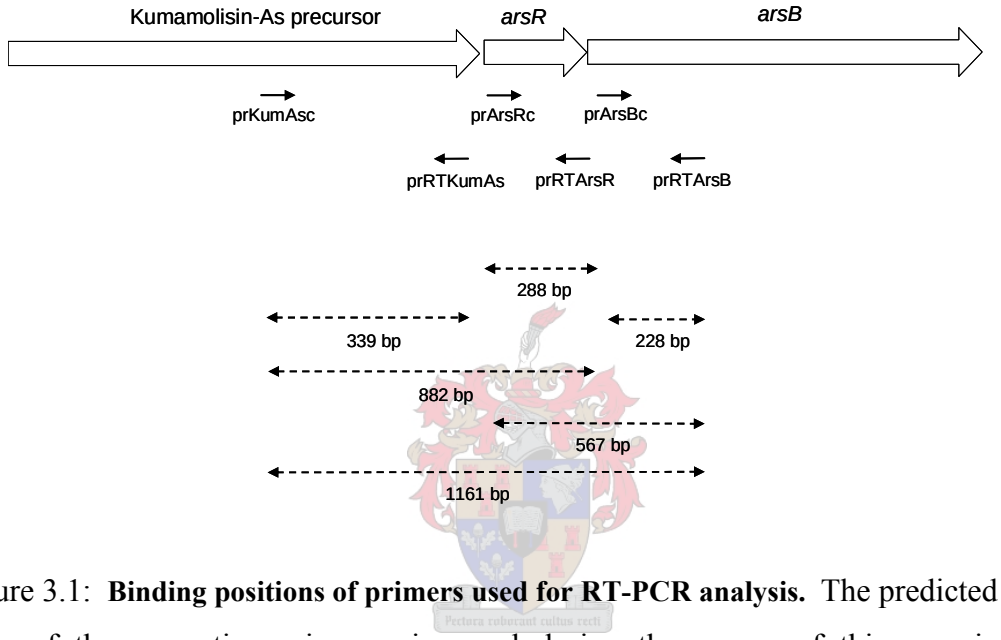


Figure 3.1: **Binding positions of primers used for RT-PCR analysis.** The predicted product sizes of the respective primer pairs used during the course of this experiment are indicated below the dashed arrows.

3.2.7 Construction of the promoter-*lacZ* reporter constructs.

The putative promoter regions of the kumamolisin-As precursor gene (Figure 3.2) and *arsR* (Figure 3.3) were amplified by PCR using the following sets of primers: prKumlacfwd400/prKumlacrev (843-1255 bp) and prArsRlacfwd/prArsRrev (2403-2881 bp), respectively. The PCR products were digested with *Bam*HI and *Eco*RI and ligated into the promoterless *lacZ* reporter gene of pMC1403, yielding constructs pMCStKumpr400 and pMCStArsRpr, respectively. Fusions were confirmed by DNA sequencing.

Figure 3.2: **The sequence of the putative promoter region of the kumamolisin-As precursor gene.** The position of the binding sites of the respective primers used to create the in-frame *lacZ* transcriptional-translational fusions are indicated by arrows, with the arrow head showing the direction of primer elongation. The putative ribosome binding site of the kumamolisin-As precursor gene is indicated by a dotted line, while the translational start codon is boxed.

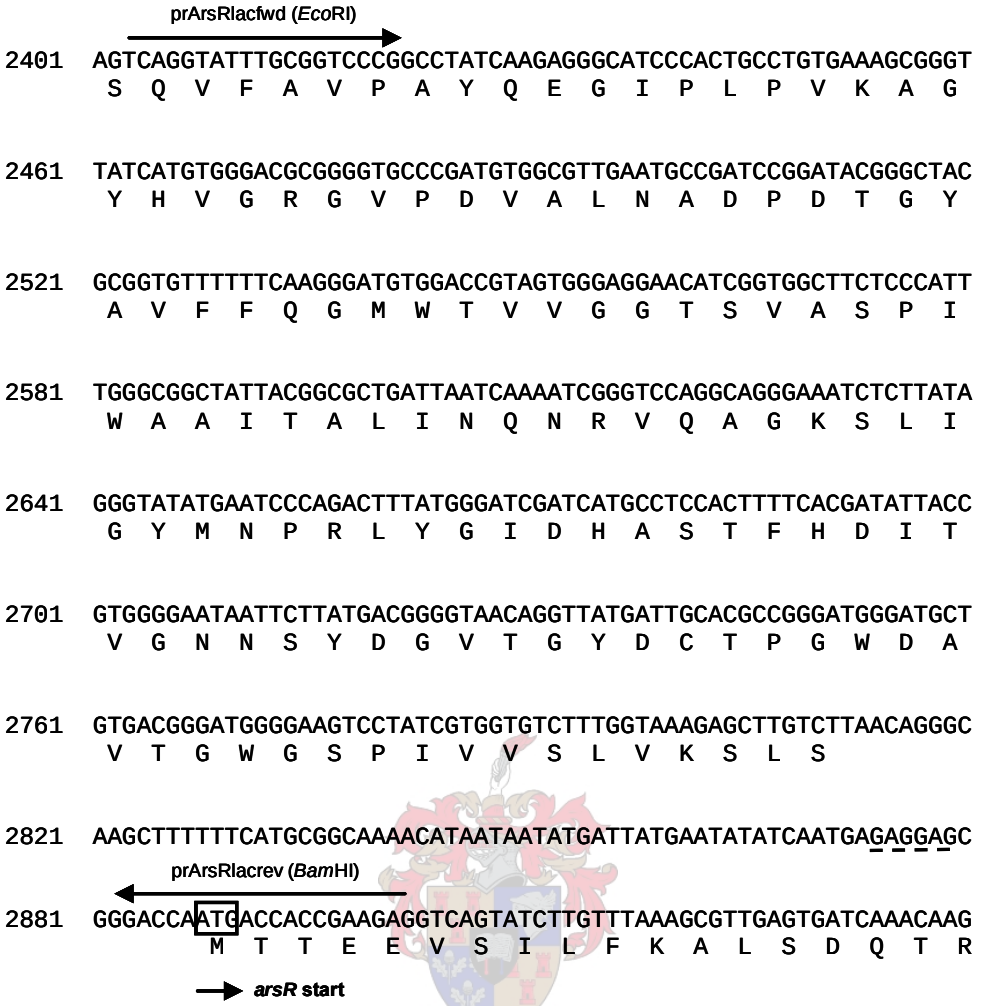


Figure 3.3: **The sequence of the putative promoter region of *arsR*.** The position of the binding sites of the respective primers used to create the in-frame *lacZ* transcriptional-translational fusions are indicated by arrows, with the arrow head showing the direction of primer elongation. The putative ribosome binding site of the *arsR* is indicated by a dotted line, while the translational start codon of the *arsR* is boxed.

3.2.8. β-galactosidase assays.

Cultures to be assayed were grown O/N in 5 ml LB medium containing the appropriate antibiotics. The O/N cultures were then diluted 1:100 into fresh 5 ml LB medium containing the same antibiotic and were incubated for 4 hours (OD_{600} 0.4-0.6). A 500 µl assay volume was used to measure the β-galactosidase activity according to the method of Miller (Miller, J.H., 1972). The β-galactosidase assays were performed in triplicate during each experiment and all experiments were repeated three times.

3.3 Results

3.3.1 The *Sb. thermosulfidooxidans* VKM B-1269^T (*St. t.* VKM B-1269^T) *ars* operon does not complement the *E. coli* ACSH50^{Iq} arsenic sensitive mutant.

In the initial attempt to identify and isolate the *ars* operon of *Sb. t.* VKM B-1269^T, constructs assembled in the size-selective mini genebank were transformed and tested for complementation of the *E. coli* ACSH50^{Iq} arsenic sensitive mutant. The *ars* operon present in *Sb. t.* VKM B-1269^T did not confer arsenite resistance to the mutant on selective LA plates containing 0.5 mM sodium arsenite, suggesting that the *ars* operon was not functional in the heterologous host. As was mentioned earlier, the expression of a full-length and biologically active protein in a heterologous host depends on factors that influence transcription, translation and protein folding. In order to test whether lack of transcription was a problem affecting the intracellular expression, the *ars* operon was cloned behind the *tac* promoter system (p*tac*ArsRB) and transformed in *E. coli* ACSH50^{Iq}. No complementation in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant was conferred by p*tac*ArsRB when tested on selective plates containing 0.5 mM arsenite.

3.3.2 Expression analysis of the *ars* operon of *Sb. t.* VKM B-1269^T in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant.

In an attempt to further investigate why the *ars* operon of *Sb. t.* VKM B-1269^T did not complement the *E. coli* arsenic sensitive mutant, it was important to determine whether the *ars* operon was transcribed in the heterologous host. Total RNA was extracted from *E. coli* ACSH50^{Iq} cells containing different constructs under uninduced (0 µM arsenite) and induced (25 µM arsenite) conditions. Slot-blot hybridization analysis was performed to detect differences in expression levels of the *ars* operon mRNA transcript caused by the addition of the inducer. Autoradiographs after hybridization using a 336 bp *Nco*I-*Hin*dIII kumamolisin-As precursor fragment and the 560 bp *arsB* fragment of *Sb. t.* VKM B-1269^T as probes, gave an indication of the changes in the expression levels of the mRNA transcripts in *E. coli* ACSH50^{Iq}. Similar hybridization signals were obtained

when the 336 bp kumamolisin-As precursor probe and 560 bp *arsB* probe were used against the total RNA extracted from *E. coli* ACSH50Iq, containing pStArs1, indicating that a mRNA transcript covering both the kumamolisin-As precursor gene and the *ars* genes was expressed (Figure 3.4 A+B). Furthermore, the mRNA transcript appeared to be constitutively expressed even under uninduced conditions. An increase in the intensity of the hybridization signal when total RNA was isolated under induced conditions, indicated that there was an increase in the expression levels of this mRNA transcript when As(III) was present.

RNA Slot-blot analysis of pArsRB (without the promoter of the kumamolisin-As precursor gene) revealed very poor expression levels with respect to those obtained in pStArs1. The lack of a sufficient amount of RNA could possibly give an explanation for these poor hybridization signals, but hybridization using the 1502 bp 16S rDNA probe of *E. coli* confirmed the presence of RNA at suitable concentrations (Figure 3.4 C). As to be expected, no hybridization signal was obtained when the 336 bp kumamolisin-As precursor probe was used, as construct pArsRB did not contain the sequence of the kumamolisin-As precursor corresponding to the labeled probe. The results obtained from RNA analysis suggest that the mRNA transcript produced by construct pStArs1 was read from a promoter located upstream of *arsR*. The removal of this promoter in construct pArsRB significantly reduced the level of mRNA expression. In contrast, the mRNA transcript produced from pArsRB showed very low levels of mRNA expression, indicating that the putative *arsR* promoter is much weaker with respect to the putative promoter positioned upstream. A quantitative indication of the mRNA expression levels of pStArs1 mRNA, as determined by the UVIgeltec software program, is presented in Figure 3.5 A+B. The very low expression levels of pArsRB mRNA could not be detected by the UVIgeltec software program, and could therefore not be quantitatively determined.

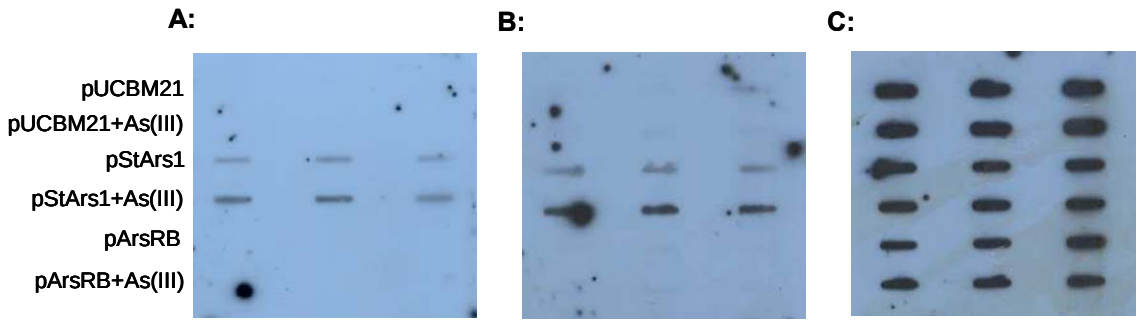
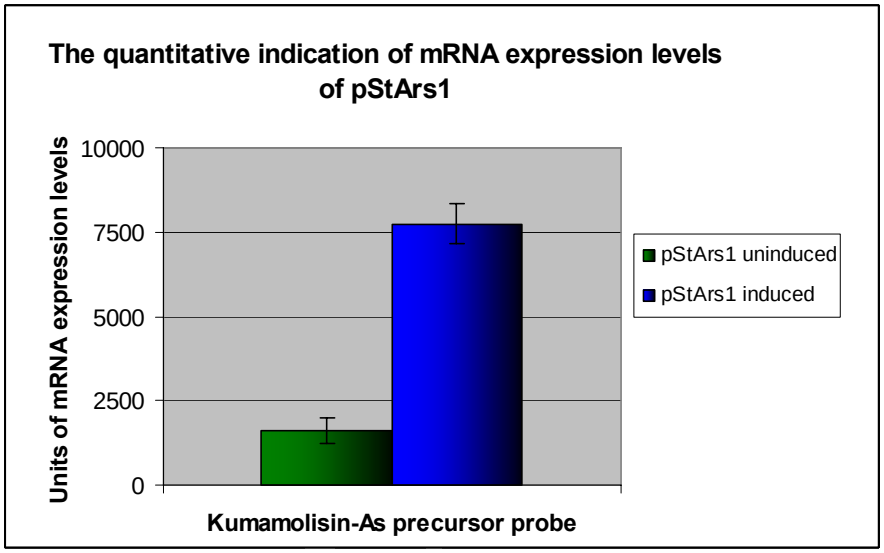


Figure 3.4: **The autoradiograph showing differences in mRNA expression levels under uninduced and induced conditions.** Northern-hybridization was performed using **(A)** a 336 bp kumamolisin-As precursor fragment and **(B)** a 560 bp *arsB* fragment of *Sb. t*. VKM B-1269^{T} as probes. **(C)** Hybridization performed with a 1502 bp 16S rDNA probe of *E. coli* acted as the positive control. Total RNA isolated from *E. coli* ACSH50^{Iq} transformed with pUCMB21 acted as the negative control. The slots were prepared in triplicate with RNA extracted from the same *E. coli* culture.

An interesting feature of the *ars* operon of *Sb. t*. VKM B-1269^{T} was the presence of a gene encoding a 525 amino acid kumamolisin-As precursor with only a 77 bp intergenenic region between the termination end of the kumamolisin-As precursor and the translational start of the *arsR* gene. Cases where putative proteinases are associated with *ars* genes have not been reported to date and it was decided to confirm whether the kumamolisin-As precursor was co-transcribed with the *ars* operon. *E. coli* ACSH50^{Iq} cells, containing pStArs1, were grown in the presence of 25 μM arsenite, whereafter total RNA was extracted. RT-PCR performed on pStArs1 mRNA revealed that all three open reading frames were co-transcribed in *E. coli*. Fragments corresponding to estimated sizes of 882 bp (product of prKum-Asc and prRTArsR showing that the kumamolisin-As precursor and *arsR* are linked), 567 bp (product of prArsRc and prRTArsB showing that *arsR* and *arsB* are linked) and 1161 bp (product of prKum-Asc and prRTArsB showing that prKum-Asc and *arsB* are linked) were obtained (Figure 3.6). This confirmed that the kumamolisin-As precursor gene, *arsR* and *arsB* were transcribed as one mRNA transcript, originating from a promoter located in front of the kumamolisin-As precursor gene.

A:



B:

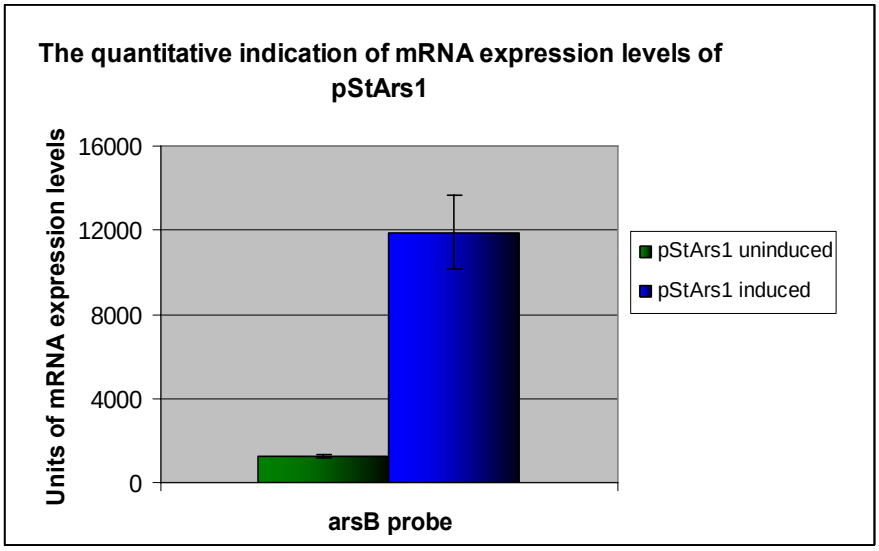


Figure 3.5: **The quantitative indication of the mRNA expression levels of pStArs1 under uninduced and induced conditions, as determined by the UVIgeltec software program.** Northern-hybridization was performed on the mRNA, using **(A)** a 336 bp kumamolisin-As precursor fragment and **(B)** a 560 bp *arsB* fragment of *Sb. t.* VKM B-1269^{T} as probes. The error bars represent the standard deviations of the triplicate readings of the different Northern-hybridization signals obtained from RNA isolated from pStArs1, as calculated by the UVIgeltec software program.

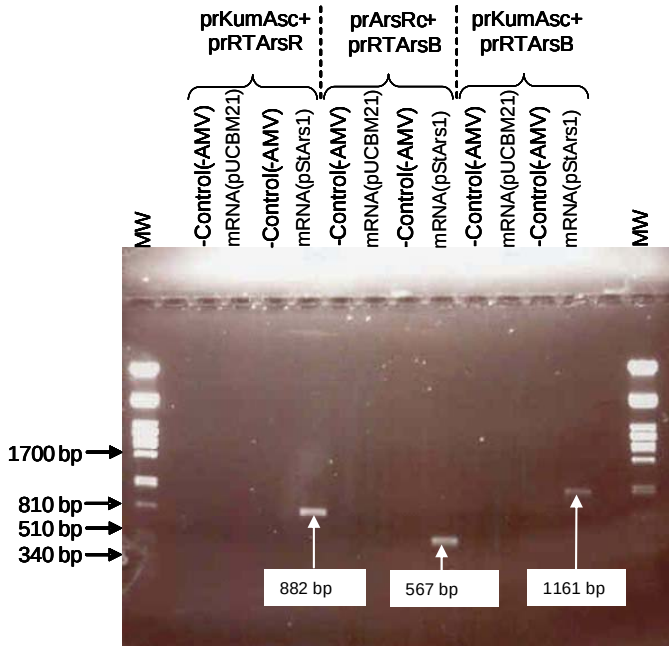


Figure 3.6: **An ethidium bromide-stained agarose gel of RT-PCR products showing that the kumamolisin-As precursor gene is co-transcribed with the *ars* genes of pStArs1 in the heterologous *E. coli* host**. The three different primer sets used are indicated. Messenger RNA isolated from *E. coli* ASCH50^{Iq} containing pUCBM21 acted as a negative control (mRNA (pUCBM21)) and was present in the first two lanes of each primer set. RT-PCR reactions with no AMV present acted as additional negative controls (-Control (-AMV)). Messenger RNA extracted from *E. coli* cells transformed with pStArs1 is shown as mRNA (pStArs1). MW indicates the molecular weight marker.

Analysis performed on the mRNA transcribed by *E. coli* ACSH50^{Iq} cells containing pStArs1 indicated that the kumamolisin-As precursor gene plays a role in the transcription of *arsR* and *arsB* in *E. coli*. Since a mRNA transcript was produced in *E. coli* ACSH50^{Iq} that extended from the kumamolisin-As precursor gene to the *arsB*, we wished to determine whether *arsRB* had its own promoter and if it could be expressed in the absence of the promoter located in front of the kumamolisin-As precursor gene. In order to do this, mRNA was isolated from *E. coli* ACSH50^{Iq} transformed with pArsRB that was grown in media containing 25 μM arsenite. RT-PCR performed on the mRNA transcript of pArsRB revealed an amplified product corresponding to a predicted size of 567 bp when primers prArsRc and prRTArsB were used together but no product was obtained using primers prKum-Asc and prRTArsR. This is an indication that *arsR*

contains an active promoter from which an mRNA transcript, with a length stretching at least to *arsB*, was produced from pArsRB (Figure 3.7).

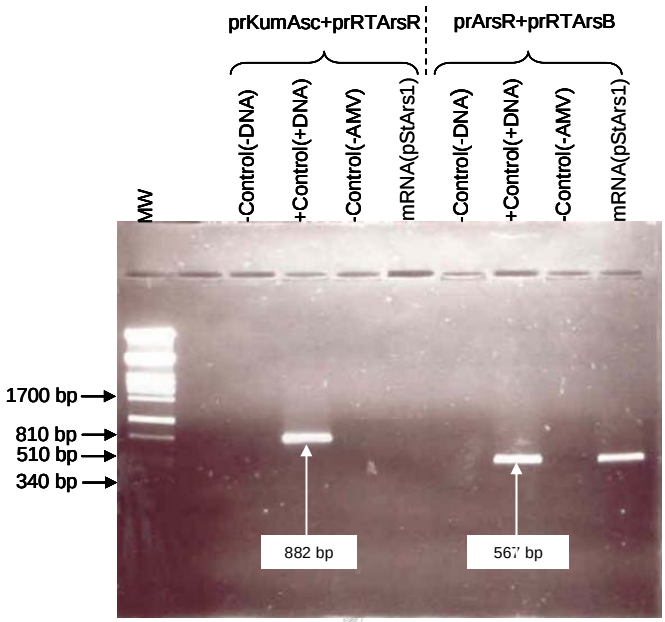


Figure 3.7: An ethidium bromide-stained agarose gel of RT-PCR products showing that the *arsR* of pStArs1 contains a functional promoter. The two different primer sets used are indicated. RT-PCR reactions with no AMV present acted as a negative control (-Control(-AMV)). In addition, the absence of pStArs1 DNA was used as another negative control (-Control-(DNA)). The presence of pStArs1 DNA acted as a positive control (+Control(+DNA)) in the PCR reaction. Messenger RNA extracted from *E. coli* cells transformed with pStArs1 is shown as mRNA (pStArs1). MW indicates the molecular weight marker.

3.3.3 Expression analysis of the *ars* operon of *Sb. t.* VKM B-1269^T in *Sulfobacillus*.

Although heterologous expression in *E. coli* is easier to investigate, we wished to study the expression of the *ars* operon in its natural *Sulfobacillus* host and determine whether any similarities were shared with respect to the expression of the *ars* operon of pStArs1 in the heterologous host and native *Sulfobacillus* host. To do this, total RNA was extracted from *Sb. t*. VKM B-1269^T cells grown under uninduced (0 μM arsenite) and induced (25 μM arsenite) conditions. Unfortunately, after numerous extraction attempts, the yields of total RNA isolated from *Sb. t*. VKM B-1269^T was not sufficient to subject it to RNA Slot-blot analysis. RT-PCR could be performed on total RNA extracted from *Sb.*

t. VKM B-1269^T grown in the presence of arsenite. RT-PCR analysis showed that no 882 bp or 1161 bp product was formed when primer set prKum-Asc/prRTArsR or prKum-Asc/prRTArsB was used. However, the 567 bp amplified PCR product was obtained when primers prArsRc and prRTArsB were used. This confirmed that the *arsR* and *arsB* genes were co-transcribed in *Sulfobacillus* and that the kumamolisin-As precursor gene and the *ars* genes of pStArs1 were transcribed as two separate mRNA transcripts. Although no mRNA transcript, linking the kumamolisin-As precursor gene and the *ars* operon was detected in *Sulfobacillus*, it was important to determine whether the kumamolisin-As precursor gene was indeed transcribed in *Sulfobacillus*. RT-PCR analysis using primers prKum-Asc and prRTKum-As, which were designed to anneal internally to the kumamolisin-As precursor gene, resulted in the amplification of the predicted 339 bp product, showing that the kumamolisin-As precursor was transcribed in *Sulfobacillus* (Figure 3.8).

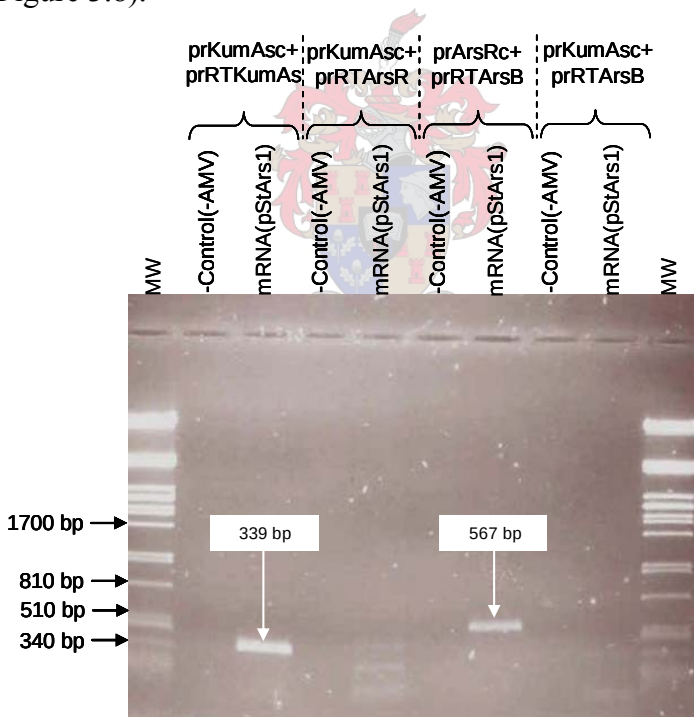


Figure 3.8: **An ethidium bromide-stained agarose gel of RT-PCR products showing that the *ars* operon of *Sb. t.* VKM B-1269^T is not co-transcribed with the kumamolisin-As precursor gene in its native *Sulfobacillus* host.** The four different primer sets used are indicated. RT-PCR reactions with no AMV present acted as a negative control (-Control(-AMV)). Messenger RNA extracted from *E. coli* cells transformed with pStArs1 is shown as mRNA (pStArs1). MW indicates the molecular weight marker.

3.3.4 Reporter-gene studies of the putative promoter regions of the kumamolisin-As precursor gene and *arsR* in *E. coli.*

Promoter expression studies of the kumamolisin-As precursor gene and *arsR* were conducted in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant using a promoterless *lacZ* fusion construct pMC1403 fused to the regions upstream of the kumamolisin-As precursor and *arsR* genes. Preliminary β-galactosidase assays performed with construct pMCStKumppr400, spanning a 412 bp region upstream of the kumamolisin-As precursor gene, revealed very low promoter activity (2.6 Miller units). An additional construct, pMCKumppr800, was then obtained by amplifying a 806 bp putative promoter region of the kumamolisin-As precursor gene, using the primer set prKumlacfwd800/prKumlacrev. pMCStKumppr800 was constructed as described in section 3.2.7. When transformed into *E. coli* ACSH50^{Iq}, pMCStKumppr800 and pMCStArsRpr gave β-galactosidase activity of 18 and 256 Miller units, respectively (Figure 3.9).

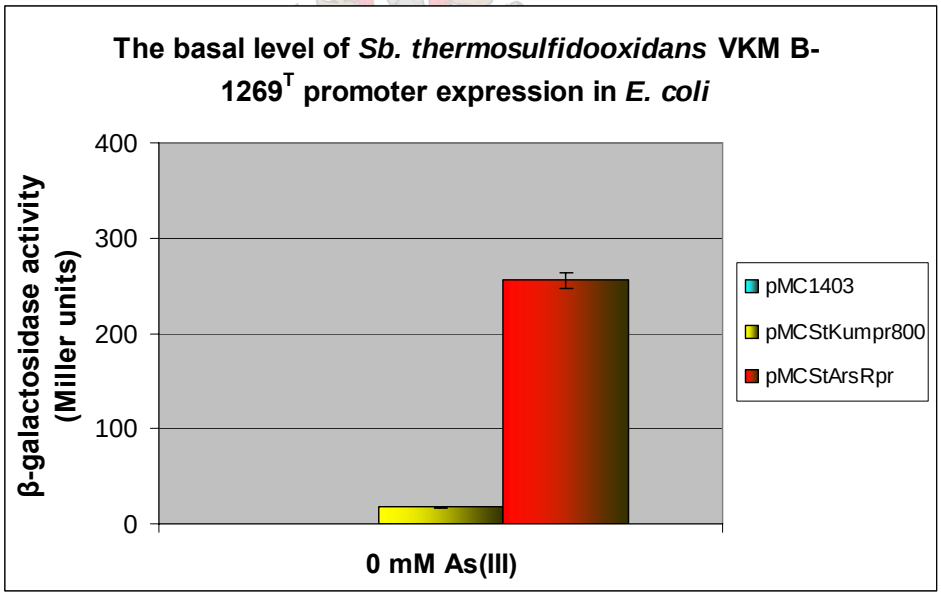


Figure 3.9: **β-galactosidase activity of the promoter-*lacZ* fusions of the kumamolisin-As precursor (pMCStKumppr800) and *arsR* (pMCStArsRpr).** The promoterless *lacZ* fusion construct pMC1403 was used as the negative control. The β-galactosidase assays were performed in triplicate and the error bars represent the standard deviations of the triplicate readings of the conducted β-galactosidase assays.

This result was different from what was found from the Slot-blot expression results, where the hybridization signal obtained for *arsRB* (mRNA transcribed by pArsRB) was considerably weaker than when the kumamolisin-As precursor gene was present upstream of *arsRB* (mRNA transcribed by pStArs1).

3.4 Discussion

E. coli is one of the most widely used hosts for the production of heterologous proteins. The production of active/functional proteins in *E. coli* is challenging and a number of problems have been encountered which show limitations in the general usefulness of *E. coli* as a host system. The biggest problem is lack of knowledge about factors affecting the stability and expression of full-length and biologically active heterologous proteins in *E. coli* (Baneyx, 1999; Rai and Padh, 2001). Successful heterologous expression of staphylococcal *ars* genes in *E. coli* has been reported (Rosenstein *et al*., 1992; Broër *et al*., 1993).

The *ars* operon of *Sb. t.* VKM B-1269^{T} was transformed in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant to test for complementation in the presence of arsenite. The *ars* operon of *Sb. t*. did not complement the *E. coli* ACSH50^{Iq} arsenic sensitive mutant, suggesting that the heterologous host was unable to produce biologically active *ars* gene products. In an attempt to determine whether the lack of *ars* gene transcription was involved in the inability of the *ars* operon to confer resistance and to eliminate factors which might be responsible for lack of *ars* gene transcription, the *ars* operon of *Sb. t.* VKM B-1269^{T} was cloned behind the *tac* promoter system. This construct was still unable to confer resistance to *E. coli* ACSH50^{Iq} in the presence of arsenic. This is a strong indication that factors involved in transcriptional processing do not contribute to the lack of *Sb. t*. VKM B-1269^{T} *ars* operon function in *E. coli* ACSH50^{Iq}.

Slot-blot analysis performed on mRNA isolated from *E. coli* ACSH50^{Iq} grown in uninduced and induced conditions confirmed that stable mRNA transcripts were produced in the heterologous *E. coli* host. The mRNA transcript produced from pStArs1

was constitutively expressed and showed increased levels of expression under induced conditions. Furthermore, this mRNA transcript appeared to be transcribed from a putative promoter positioned in front of the kumamolisin-As precursor gene, and stretched to the *arsB*. Because an intergenic region of only 77 bp was separating the open reading frames of the kumamolisin-As precursor gene and *arsR*, the presence of a functional *arsR* promoter had to be investigated. Very low levels of mRNA transcription from pArsRB (not containing the putative promoter of the kumamolisin-As precursor gene) were detected, confirming *arsR* promoter activity in the heterologous *E. coli* host. The mRNA transcript produced from the *arsR* promoter appeared to be constitutively expressed, irrespective of the absence of inducer. The significant difference of the hybridization signal strength when pStArs1 and pArsRB mRNA was probed with *arsB* suggests that *arsRB* is transcribed from a much weaker promoter in comparison to the one positioned upstream of the kumamolisin-As precursor gene. Very low levels of mRNA transcription from pArsRB were detected in the heterologous *E. coli* host, irrespective of the presence of arsenic. The expression levels of the mRNA under uninduced and induced conditions were quantitatively calculated by the UVIgeltec software program. Analysis revealed a 4.8 fold increase (1625 to 7735 units) in mRNA expression levels when the 336 bp kumamolisin-As precursor fragment of *Sb. t.* VKM B-1269^{T} was used as a probe during hybridization. In the case where the 560 bp *arsB* fragment of *Sb. t*. VKM B-1269^{T} acted as a probe, a 9.5 fold increase (1255 to 11897 units) in mRNA expression levels were obtained. The stability of the mRNA transcript, the presence of the mRNA transcript produced by the *arsR* promoter (pArsRB) and differences in the concentration of labeled fragments used during the hybridization process, could possibly have resulted in the observed differences in the quantitatively determined increase in expression levels when the different probes were used. The very low expression levels of pArsRB mRNA could not be detected by the UVIgeltec software program, and was therefore not quantitatively determined.

RT-PCR analysis was performed in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant to determine whether the kumamolisin-As precursor gene was indeed co-transcribed with the *ars* operon. Amplified fragments corresponding to the predicted PCR products with

estimated sizes of 882 bp, 567 bp and 1161 bp confirmed that the kumamolisin-As precursor gene was co-transcribed with *arsRB* in *E. coli* ASCH50^{Iq}. In addition, RT-PCR showed that a mRNA transcript stretching from *arsR* to *arsB* was produced from pArsRB in *E. coli* ACSH50^{Iq}. This was conclusive evidence that *arsR* did have a functional promoter. All these findings were consistent with results obtained during Slot-blot analysis where a 336 bp kumamolisin-As precursor fragment and a 560 bp *arsB* fragment of *Sb. t*. VKM B-1269^{T} were used as probes. To investigate if any similarities were shared with respect to the expression of the *ars* operon of pStArs1 in the heterologous host and native *Sulfobacillus* host, RT-PCR was performed on total RNA extracted from *S. t*. VKM B-1269^{T}. In contrast to the heterologous *E. coli* host, analysis revealed that the *ars* operon of *Sb. t.* VKM B-1269^{T} was not co-transcribed with the kumamolisin-As precursor gene in its native *Sulfobacillus* host. The *arsR* and *arsB* of *Sb. t.* VKM B-1269^{T} were co-transcribed in the native *Sulfobacillus* host.

β-galactosidase reporter gene studies of the kumamolisin-As precursor gene and the *arsR* gene promoters was performed in *E. coli* ACSH50^{Iq} to determine relative promoter activity strength. β-galactosidase assay results were contradictory to what we expected. Slot-blot expression studies showed that the putative promoter of the kumamolisin-As precursor gene was much stronger in comparison to the promoter of *arsR*. In contrast, β-galactosidase assays revealed that the promoter of *arsR* resulted in a β-galactosidase activity of 256 Miller units, with the kumamolisin-As precursor promoter producing only 18 Miller units of β-galactosidase activity in the *E. coli* ACSH50^{Iq} arsenic sensitive mutant. It has to be emphasized that β-galactosidase assays performed using the promotorless *lacZ* fusion construct are influenced on both a transcriptional and translational level, while Slot-blot analysis results are affected on a transcriptional level alone.

Initiation of translation of *E. coli* mRNA requires a ribosome binding site, also called the Shine Dalgarno site, complementary to the 3' end of the 16S rDNA and of consensus 5'-UAAGGAGG-3'. Studies have shown that the spacing of the ribosome binding site and the initiation AUG codon plays a crucial role in the effective translation of genes

(Baneyx, 1999). Furthermore, differences in codon position and recognition between bacterial hosts have a significant impact on heterologous protein production (Jana and Deb, 2005). In the light of these findings it can not be excluded that factors involved in the translational process within *E. coli* reacted differently to the respective putative ribosome binding sites of the kumamolisin-As precursor gene and the *arsR* gene. β-galactosidase assays indicated that the *arsR* promoter-fusion had more β-galactosidase activity in comparison to the kumamolisin-As precursor promoter-fusion, so it is therefore evident that the ribosome binding site of *arsR* initiated translation of the β-galactosidase gene more effectively, although levels of mRNA transcribed by the *arsR* promoter were much lower in comparison to the mRNA transcribed by the promoter of the kumamolisin-As precursor gene.

It is clear from an aggregate of results that promoter activity, mRNA expression and mRNA instability are not decisive factors contributing to the inability of the *ars* operon of *Sb. t.* VKM B-1269^{T} to compliment the heterologous *E. coli* ACSH50^{Iq} arsenic sensitive mutant. This inability of the *Sb. t.* VKM B-1269^{T} *ars* operon is likely to be caused by processes occurring after the production of the mRNA transcript. The fact that many fundamental aspects of *E. coli* physiology remain to be uncovered and that only a small amount of information has been exploited for practical purposes will continue to encourage research to optimize this microorganism for heterologous protein expression.

CHAPTER 4

THE DISTRIBUTION AND EVOLUTIONARY RELATIONSHIP OF ARSENIC RESISTANCE GENES WITHIN GENUS *SULFOBACILLUS*

TABLE OF CONTENT

4.1 Introduction 87

4.2 Materials and Methods

4.2.1 Bacterial strains, plasmids and PCR primers 89

4.2.2 Media and growth conditions 89

4.2.3 DNA isolation, techniques and analysis 91

4.2.4 Sulfobacilli media and growth conditions 91

4.2.5 Total DNA isolation from Sulfobacilli 92

4.2.6 Polymerase chain reaction (PCR) 92

4.2.7 Amplified ribosomal DNA restriction enzyme analysis (ARDREA) of Sulfobacilli 92

4.2.8 Pulsed-field gel electrophoresis (TAFE) 93

4.3 Results

4.3.1 ARDREA of the *Sulfobacillus* spp. corresponds to the respective Gram-positive mesophilic and moderately thermophilic bacterial profiles 93

4.3.2 ARDREA indicates that *Sulfobacillus* spp. can be divided into two separate subgroups within genus Sulfobacillus 96

4.3.3 The identification and distribution of other *arsB* homologues within the two *Sulfobacillus* spp. subgroups 99

4.3.4 The *Sb. thermosulfidooxidans* (*Sb. t.*) isolates share common features with respect to *arsRB* gene orientation and layout 104

4.3.5 The *arsRB* operon of *Sb. t*. VKM B-1269^{T} is chromosomally located 106

4.4 Discussion 108

4.1 Introduction

The ability to identify and differentiate between bacterial species is of considerable environmental and biotechnological importance. Methodologies currently employed for bacterial strain identification and taxonomical classification can be divided into two broad categories, namely microbiologically-based (phylotypically-based) and molecular biologically-based methods. Microbiologically-based methods are basic techniques and provide limited information with respect to the differentiation and taxonomical classification of bacteria. Molecular biologically-based methods comprise a broad range of techniques and include DNA reassociation kinetics, separation of total DNA samples based on buoyant density differences associated with the %A+T and %G+C content, DNA-DNA hybridization and the sequencing and analysis of 16 rDNA and other genes to characterise bacterial strains. Several important criteria must be evaluated in assessing the utility of a particular typing method for bacterial strain identification and taxonomical classification: (1) all organisms within a species must be typeable by the method used, (2) the method must have high levels of differentiation and (3) the method must be reproducible (Olive and Bean, 1999).

The notable shortcomings of microbiologically-based methods in accurate bacterial strain identification and differentiation to a genus or species level, have subsequently led to an increase in the employment of molecular biologically-based methods for bacterial typing purposes. Typing methods based on microbial DNA analysis are highly sensitive and allow for a high degree of specificity, resulting in the gain of a great deal of information (Olive and Bean, 1999; Johnson *et al*., 2005). One technique that has been used in this context is amplified ribosomal DNA restriction enzyme analysis (ARDREA), also known as RFLP (restriction fragment length polymorphism). In this method, standard restriction digestion analysis is performed on PCR amplified rDNA genes (typically that code for 16S or 23S rRNA), whereafter differences in DNA banding patterns can be detected after electrophoretic separation on a regular agarose gel has been performed (Spiegelman *et al*., 2005).

Johnson and coworkers (Johnson *et al*., 2005) reported that ARDREA can be implemented to successfully distinguish between mesophilic and moderately thermophilc *Sulfobacillus*-like isolates at a species level. From an aggregate of results, they proposed that *Sulfobacillus*-like isolates could be divided into *Sulfobacillus* subgroup I, containing *Sb. thermosulfidooxidans/Sb. montserratensis*-like isolates, and *Sulfobacillus* subgroup II, containing *Sb*. *acidophilus*/*Sb yellowstonensis*-like isolates, respectively. The prospect for the application of ARDREA as a tool to identify and discriminate between different *Sulfobacillus*-like isolates is of considerable industrial importance.

Genetic determinants (*ars* operons) which confer resistance to arsenical compounds are widely distributed within the domain *Bacteria* (Eubacteria) and are usually present on the chromosome, plasmids or transposable elements. The distribution of *ars* operons within different iron-oxidizing, acidophilic bacterial species within the same genus is not well reported. It has been shown that some strains of genus *Leptospirillum* and *Acidithiobacillus* contain several sets of *ars* genes, while other members within the same genus contain at least a set of chromosomal *ars* genes (de Groot *et al*., 2003; Tuffin *et al*., 2004; 2006). The referred strains were isolated from different geographical areas and it is therefore a strong indication that an environment containing arsenical compounds selects and maintains microbes possessing these additional *ars* genes.

Here we confirmed that five additional *Sulfobacillus* isolates can conveniently be divided into separate subgroups by using the molecular biologically-based method, ARDREA, as determined by Johnson and coworkers (Johnson *et al*., 2005). Furthermore, we identified and partially sequenced *arsB* homologues present within the respective *Sulfobacillus* spp. and compare the phylogeny obtained using these sequences with ARDREA and 16S rDNA sequence analysis. These newly discovered *arsB*'s seem to be spread throughout the members of genus *Sulfobacillus*, irrespective of the subgroup they represent. In addition, it was shown that the *Sb. t.* VKM B-1269^{T} *arsRB* operon is chromosomally located.

4.2 Materials and Methods

4.2.1 Bacterial strains, plasmids and PCR primers.

The bacterial strains, plasmids and primers used in this study are described in Table 4.1.

4.2.2 Media and growth conditions.

E. coli cells were grown aerobically on LA plates or in LB at 37°C as described by Sambrook *et al.*, 1989. Amp was added for selection at a concentration of 100 µg/ml when required.

Table 4.1 **Bacterial strains, plasmids and PCR primers used in this study**

Strains, plasmids or primers	Genotype or description	Reference or source
Strains		
Escherichia coli		
DH5α	*F'/endA1 hsdR17* ($r_K^- m_K^+$) *supE44 thi-1recA1 gyrA* (Nal^r) *relA1 Δ(lacZYA-ArgF) U169 (Φ80dlac Δ(lacZ) M15)*	Promega Corp.
XL1-Blue	F'::Tn*10* $proA^+B^+$ $lacI^q$ *Δ(lacZ) M15/recA1 endA1 gyrA96* (Nal^r) *thi hsdR17*($r_K^- m_K^+$) *supE44 relA1 lac*	Promega Corp.
$ACSH50^{Iq}$	*rpsL* Δ(*lac –pro*) Δ*ars*::*cam*	Butcher and Rawlings, 2002
Sulfobacillus acidophilus		
$ALV^{☼=}$	Warwickshire, England (Coal spoil)	Marsh and Norris., 1983
$THI^{☼=}$		
Sulfobacillus montserratensis	Iceland (Thermal Spring)	Norris *et al.*, 1996
$L15^{☼}$	Montserrat Island	Yahya *et al.*, 1999
$Y0017^{☼=}$	Yellowstone National Park, U.S.A. (Thermal Spring)	
Sulfobacillus thermosulfidooxidans		

G2☼		
VKM B-1269$^{T☼=}$	Montserrat Island	Johnson *et al*., 2005
MT13$^{=}$	Russia	Golovacheva and Karavaiko, 1978
Sulfobacillus yellowstonensis YTF1☼	Commercial bioleaching operation, R.S.A	Schutte, 2004 (M. Sc thesis)
	Yellowstone National Park, U.S.A. (Thermal Spring)	Johnson *et al*., 2001

Plasmids

pGEM-T®	Ap^{r}; T-tailed PCR product cloning vector	Promega
p*St*Ars1	Ap^{r}; 5300 bp *Eco*RI-*Pst*I fragment, containing the *ars* operon of *Sb. t.* VKM B-1269^{T} cloned into pUCBM21 digested with *Eco*RI *and Pst*I	This study

Primers[#]

Ferro arsBfwd	5'-GTTIGCCAACGAIGGIGCGGC-3'	This laboratory
Ferro arsBrev	5'-ACATGCAICCAGAGCAGIGTIGC-3'	This laboratory
prKaraArsBfwd (*Eco*RI)	5'-GTGAGAATTCGATACAACCTCTGTCCCG-3'	This study
prKaraArsBrev (*X*baI)	5'-GCTCTCTAGACCGATGAAAGAAGTGC-3'	This study
prKaraArsB2fwd	5'-GCATTGGTTTTTTCGAATGGGC-3'	This study
prKaraArsB2rev	5'-CGTCAACAGCACCGGAATCG-3'	This study
prKum-Asc	5'-GGATGTGAATAATAACGGAGC-3'	This study
prRTKum-As	5'-GTGCAATCATAACCTGTTACC-3'	This study
prArsRc	5'-CGTTGAGTGATCAAACAAGG-3'	This study
prRTArsR	5'-CTACAGTTCTAAGAATGACCG-3'	This study
prArsBc	5'-GCAAGATAAGAACAAAGGG-3'	This study
prRTArsB	5'-CCATCGAAATGAGAATAAGCG-3'	This study
fDD2	5'-CCGGATCCGTCGACAGAGTTTGATCITGG CTCAG-3'	Lane, 1991
rPP2	5'-CCAAGCTTCTAGACGGITACCTTGTTACG ACTT-3'	Lane, 1991

Ap^{r}: ampicillin resistance

[#] restriction endonuclease sites incorporated into primers are indicated in parenthesis and are underlined in the primer sequence

☼ *Sulfobacillus* strains kindly supplied by Prof. D.B. Johnson

= *Sulfobacillus* species classification done within this study

4.2.3 DNA isolation, techniques and analysis.

Restriction endonuclease digestions, gel electrophoresis, small-scale plasmid preparation, ligation reactions and Southern-blot hybridization were performed using standard methods (Sambrook *et al.*, 1989). Large-scale plasmid preparation was done by means of the alkaline lysis procedure whereafter plasmid DNA was purified by CsCl/etbr equilibrium centrifugation (Current Protocols in Molecular Biology) or by using the Nucleobond AX100 system (Macherey-Nagel). DNA fragments to be used for cloning purposes were extracted from agarose gels with the GFX™ kit (Amersham BioSiences). Clones to be sequenced were isolated and purified from 5 ml LB O/N culture using the High Pure Plasmid Isolation Kit (Roche Molecular Biochemicals). Sequencing of constructs was carried out by using the dideoxy chain termination method and was executed by an ABI PRISM™ 377 automated DNA sequencer. Sequences were analysed using several software programmes, but mainly by the PC based DNAMAN (version 4.1) package (Lynnon Biosoft). The labeling of probes, hybridization and detection was conducted by using the dioxigenin-dUTP nonradioactive DNA labeling and detection system (Roche Molecular Biochemicals). When Southern-blot hybridization was performed with a homologous DNA probe, prehybridization and probing were done at 42ºC and the buffer B washes were conducted at 65 ºC. In the cases where a heterologous DNA probe was utilized in Southern-blot hybridization, the prehybridization and probing steps were carried out at 37ºC and the buffer B washes performed at 52 ºC.

4.2.4 Sulfobacilli media and growth conditions.

The details of Sulfobacilli media and growth conditions used are as previously described in section 2.2.4 of Chapter 2 (page 42).

4.2.5 Total DNA isolation from Sulfobacilli.

Sulfobacilli cells were harvested from $FeSO_4$ media by centrifugation at 10000 rpm for 30 minutes. The cell pellet was resuspended in acidified water (pH 1.8) and subjected to several low speed (2000 rpm) and high speed (9000 rpm) centrifugation spins. These additional washing steps were performed in order to remove excess ferric oxide from the media. The resulting cell pellet was then washed thrice with TE buffer (pH 8) in order to neutralize the media. Small scale extraction of Sulfobacilli chromosomal DNA was preformed by using the High Pure PCR Template Preparation Kit™ (Roche Molecular Biochemicals).

4.2.6 Polymerase chain reaction (PCR).

The specifications of the performed PCR amplification reactions are described in section 2.2.7 of Chapter 2 (page 44).

The 16S rRNA genes of each of the tested isolates were subjected to PCR amplification, using the primer set fDD2 and rPP2. Primers fDD2 and rPP2 are derivatives of the eubacterial primers 27f and 1492r, which are specific to positions 8 to 27 and 1493 to 1510 of the *E. coli* 16S rRNA gene, respectively (Lane, 1991). PCR was carried out on 200 ng of isolated chromosomal DNA. Double-stranded template DNA were denatured at 94°C for 60s, where-after 35 cycles of a denaturation step (30 s at 94°C), an annealing step (45 s at 53°C) and a final variable extension step (90 s at 72°C) were performed.

4.2.7 Amplified ribosomal DNA restriction enzyme analysis (ARDREA) of Sulfobacilli.

The restriction enzymes used for the differentiation between the *Sulfobacillus* strains, include *Xcm*I (from New England Biolabs), *Eco*721 and *Eco*1015 (both from MBI Fermentas). Digestions of the amplified 16S rRNA genes were conducted according to the manufacturers directions, whereafter DNA fragments were separated using a 2% (w/v) agarose gel in 1X TBE buffer. Agarose gels were subjected to electrophoresis at 60 V and subsequently visualized under UV light.

4.2.8 Pulsed-field gel electrophoresis (TAFE).

Sulfobacilli cells were harvested from $FeSO_4$ media**,** washed thrice in mineral salts media and resuspended in SET buffer (50 mM Tris, 2 mM EDTA, 25% Sucrose, adjusted to pH 8). The cell suspension was adjusted with SET buffer to obtain an approximate optical cell density of 2 at OD_{600}. Proteinase K (1 mg/ml) was added to the cell suspension and incubation at 37°C was carried out for 5 minutes. An equal volume of 2% LMP agarose (Seaplague, FMC Bioproducts) was then added to the cell suspension and applied to a 2x2x25 mm mold in order to create "plugs" containing embedded cells. The plugs were extracted from the mold and washed in a petridish containing 1.6 ml TE buffer and 400 µl 10 % SDS until plugs became clear. The embedded cells were lysed by incubating the plugs in ESP buffer (0.5M EDTA, pH 8; 1% SDS; Proteinase K (1 mg/ml)) for 30 minutes at 37°C and then 16 hours at 50°C. The ESP solution was then replaced with fresh ESP solution and the plugs were incubated for a further 16 hours at 50°C. The Proteinase K was inactivated by washing the plugs in TE_{50} with Pefabloc (Roche Molecular Biochemicals) for 12 hours before it was incubated O/N at 4°C. The gel plugs were subjected to *Aspergillus oryzae* S1 nuclease (Sigma) and treated according to Barton *et al*., (1995). Trans-alternating field electrophorsis (TAFE) was performed using a Beckman GeneLine apparatus. Electrophoresis was carried out in an 1% agarose gel at 150 mA and 12°C for 15 h with a pulse interval of 13 seconds. After completion, the gel was subjected to electrophoresis for 1 hour with a pulse interval of 4 seconds.

4.3 Results

4.3.1 ARDREA of the *Sulfobacillus* spp. corresponds to the respective Gram-positive mesophilic and moderately thermophilic bacterial profiles.

In order to differentiate between iron-oxidizing Gram-positive mesophilic acidophilic bacteria and iron-oxidizing Gram-negative mesophilic acidophilic bacteria, Johnson *et al*., (2005) proposed that the PCR amplified 16S rDNA genes of the respective isolates

should be digested with the restriction endonuclease *Xcm*I. In the case of iron-oxidizing Gram-positive mesophilic acidophilic bacteria, a distinctive two DNA fragment pattern, one fragment corresponding to 200 bp and the other to >1250 bp, would be obtained (Figure 4.1 A). Similarly, discrimination between iron-oxidizing Gram-positive moderately acidophilic bacteria, representing genus *Sulfobacillus*, and other iron-oxidizing, acidophilic moderately thermophilic bacterial species could be obtained by digesting the PCR amplified 16S rDNA genes of the respective isolates with the restriction endonuclease *Eco*721. The *Sulfobacillus* banding pattern should contain three DNA fragments, corresponding to 100 bp, 300 bp and 1100 bp (Figure 4.1 B).

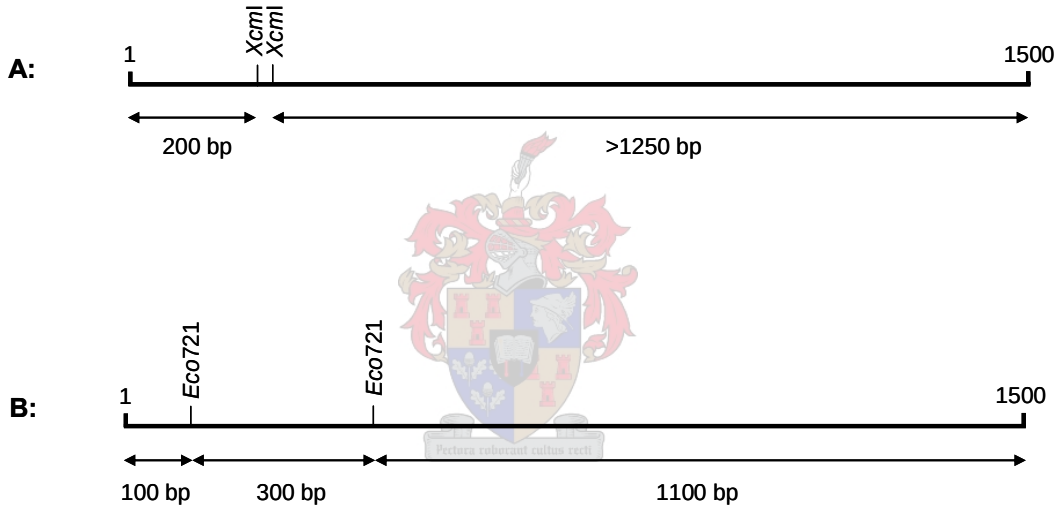


Figure 4.1: **Computer software-generated restriction enzyme maps used in the analysis and differentiation of the 16S rDNA genes of acidophilic bacteria: (A)** Gram-positive, iron-oxidizing, mesophilic acidophiles 16S rDNA digested with *Xcm*I and **(B)** Gram-positive, iron-oxidizing, moderately thermophilic acidophiles 16S rDNA digested with *Eco*721. The sizes of the computer-generated restriction enzyme maps are based on the assumption that the DNA to be analysed is 1500 bp in length. Depending on the exact position of the primers used, a slight difference in the size of the actual fragments may be detected on an agarose gel. Adapted from Johnson *et al*., 2005.

To confirm that all isolates were indeed Sulfobacilli, the 16S rDNA genes of isolates *Sulfobacillus acidophilus*, strain ALV and THI (*Sb. a.* ALV; TH1), *Sulfobacillus montserratensis,* strain L15 and YOO17 (*Sb. m*. L15; YOO17), *Sulfobacillus thermosulfidooxidans,* strain G2, VKM B-1269^{T} and MT13 (*Sb. t.* G2; VKM B-1269^{T}; MT13) and *Sulfobacillus yellowstonensis*, strain YTF1 (*Sb. y*. YTF1) were amplified with primer set rDD2 and rPP2. In the cases of *Sb. m*. L15 and *Sb. m* YOO17, the 16S rDNA genes were digested with *Xcm*I and *Eco*721, respectively (Figure 4.2 A). The amplified 16S rDNA genes of the other *Sulfobacillus* isolates were digested with *Eco*72I (Figure 4.2 B).

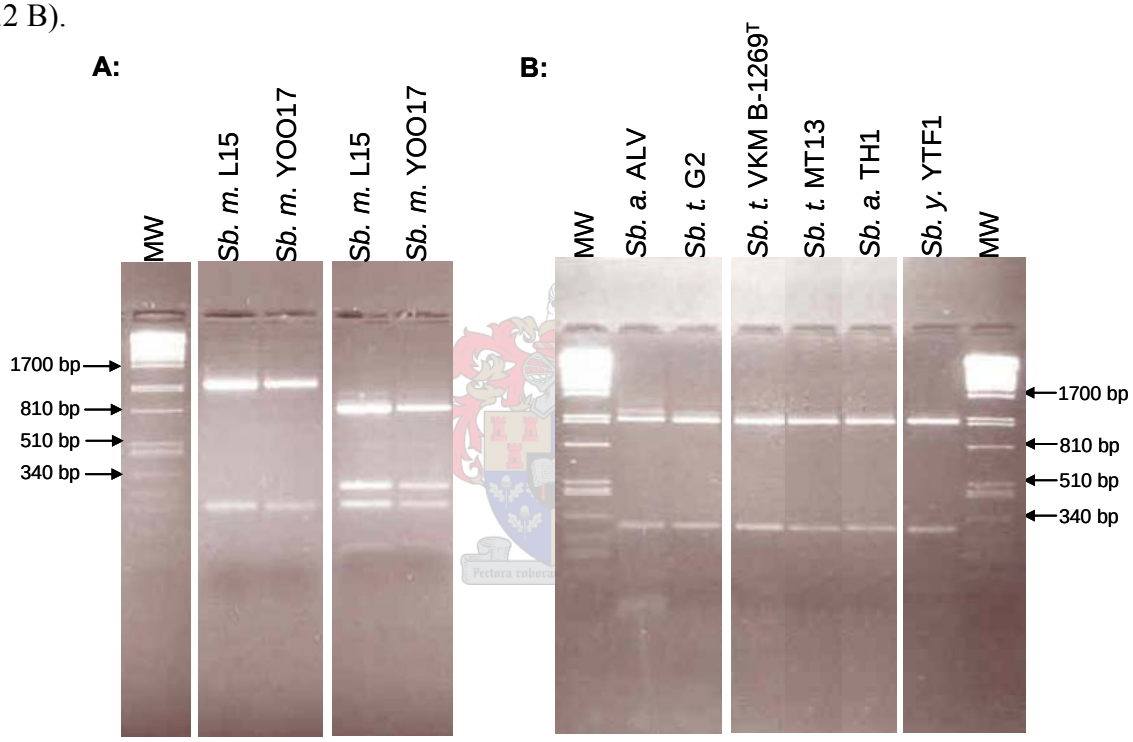


Figure 4.2: **An 1% agarose gel showing the respective DNA banding patterns after ARDREA was performed on mesophilic and moderately thermophilic Gram-positive, iron-oxidizing, acidophilic bacteria. (A)** The amplified 16S rDNA genes of the mesophilic Gram-positive, iron-oxidizing, acidophiles *Sb. m*. L15 and *Sb. m*. YOO17 were digested with *Xcm*1 (lanes 2 and 3) and *Eco*72I (lanes 4 and 5), **(B)** while the amplified 16S rDNA genes of the moderately thermophilic Gram-positive, iron-oxidizing acidophiles were digested with *Eco*721. MW indicates the molecular weight marker.

ARDREA of these *Sulfobacillus* strains confirmed the banding patterns found by Johnson *et al*. (2005) using mostly different isolates. Johnson and coworkers (Johnson *et al*.,

2005) proposed that discriminatory DNA band patterns could be obtained by performing ARDREA with *Xcm*I and *Eco*72I to distinguish between mesophilic (using *Xcm*I) and moderately thermophilic (using *Eco*721) Gram-positive acidophiles. When the amplified 16S rDNA gene of the mesophilic Gram-positive acidophile *Sb. m*. L15 was digested with *Xcm*I, the predicted DNA fragments of 200 bp and >1250 bp were obtained (Figure 4.2 A(lane 2)). An identical result was also obtained after the 16S rDNA gene of *Sb. m*. YOO17 was digested with *Xcm*I (Figure 4.2 A(lane 3)). Because isolate *Sb. m.* YOO17 has previously not been classified as a mesophilic or moderately thermophilic Gram-positive acidophile on the basis of ARDREA results, the 16S rDNA gene of *Sb. m*. YOO17 was subsequently digested with *Eco*721 (Figure 4.2 A(lane 5)). DNA fragments of estimated sizes 800 bp, 200 bp and 350 bp were obtained which did not correspond to the moderately thermophile DNA band profile (1100 bp, 300 bp and 100 bp) as proposed by Johnson and coworkers (Johnson *et al*., 2005). Furthermore, ARDREA performed with *Eco*721 on *Sb. m.* L 15 revealed similar results (Figure 4.2. A(lane 4)). The accuracy of these results was confirmed by 16S rDNA sequence analysis using the PC based DNAMAN (version 4.1) package (Lynnon Biosoft). Restriction enzyme maps of the 16S rDNA genes of *Sb. m*. L15 (accession nr. AY007663) and *Sb. m*. YOO17 (accession nr. AY140239) revealed DNA fragment sizes corresponding to those obtained by ARDREA. Digestion of the amplified 16S rDNA genes of the moderately thermophilic Gram-positive acidophiles with *Eco*721 yielded the predicted DNA fragment profile containing bands of 300 bp and 1100 bp in size. The 100 bp fragment predicted to be present within these digested DNA fractions was not visible on the 1% agarose gel.

4.3.2 ARDREA indicate that *Sulfobacillus* spp. can be divided into two separate subgroups within genus *Sulfobacillus*.

In addition, Johnson and coworkers (Johnson *et al*., 2005) also showed that *Sulfobacillus*-like isolates could conveniently be divided into two major subgroups, based on differences in patterns after their respective amplified 16S rDNA genes were digested with *Eco*1015 (*Sna*BI). Johnson *et al*., (2005) proposed that subgroup I, containing *Sb.*

t./Sb. m.-like isolates would reveal a two-fragment DNA banding pattern, corresponding to 550 bp and 950 bp, respectively, after digestion with *Eco*1015 (*Sna*BI) (Figure 4.3 A), while no *Eco*1015 (*Sna*B1) digestion site is located on the 16S rDNA genes of *Sulfobacillus* subgroup II, containing *Sb. a./Sb y.*-like isolates, revealing an uncut 1500 bp DNA fragment (Figure 4.3 B).

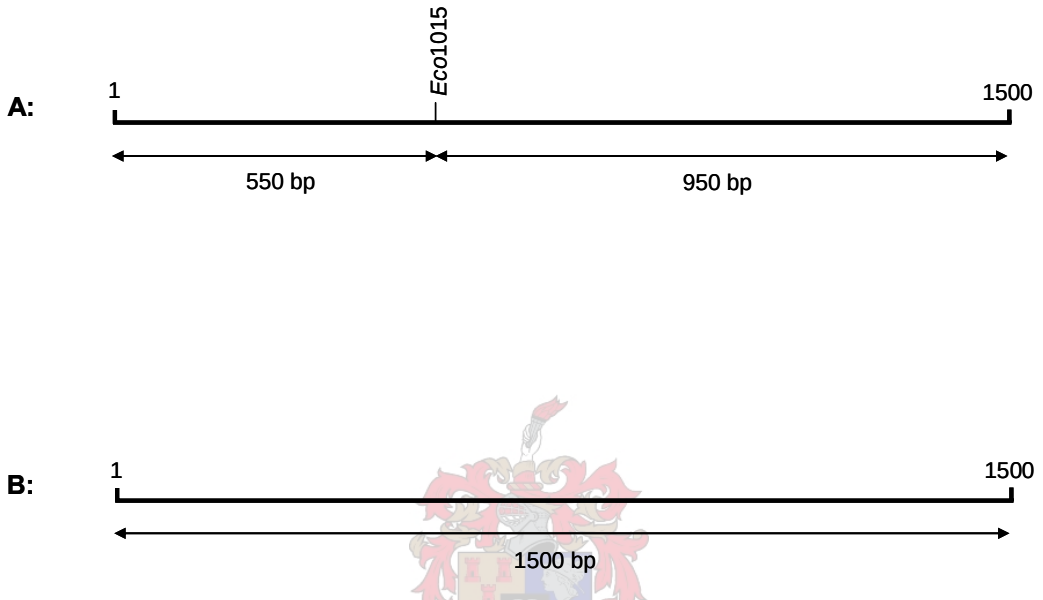


Figure 4.3: **Restriction enzyme maps generated by computer of digested 16S rDNA genes for the analysis and differentiation of *Sulfobacillus* spp. (A)** Mesophilic and moderately thermophilic *Sulfobacillus* spp., subgroup I 16S rDNA digested with *Eco*1015 (*Sna*BI) and **(B)** mesophilic and moderately thermophilic *Sulfobacillus* spp., subgroup II 16S rDNA digested with *Eco*1015 (*Sna*BI). The sizes of the computer-generated restriction enzyme maps are based on the assumption that the DNA to be analysed is 1500 bp in length. Due to this reason, a slight difference in the size of the actual fragments may be observed on an agarose gel. Adapted from Johnson *et al*., (2005).

The 16S rDNA of the respective *Sulfobacillus* isolates described in section 4.3.1 was amplified using primer set rDD2 and rPP2 and subsequently digested with *Eco*1015 (*Sna*BI). The DNA fragments were separated by means of electrophoresis to obtain distinctive banding profile representative of each of the proposed *Sulfobacillus* spp. subgroups (Figure 4.4).

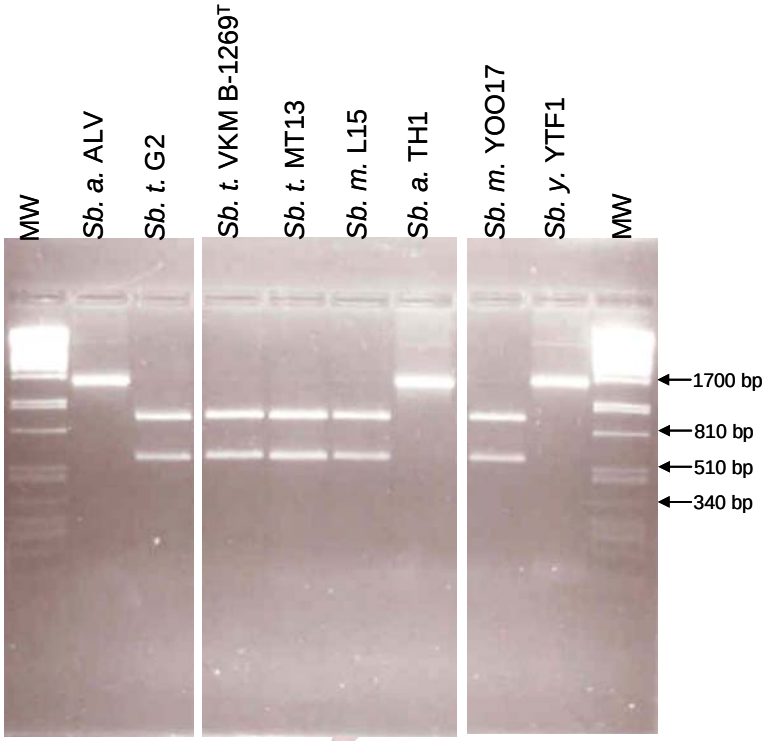


Figure 4.4: **An 1% agarose gel showing the DNA banding patterns corresponding to the *Sulfobacillus* spp., subgroup I and *Sulfobacillus* spp., subgroup II profile after ARDREA was performed on the respective *Sulfobacillus* isolates.** The amplified 16S rDNA of the *Sulfobacillus* spp. was digested with *Eco*1015 (*Sna*BI) to obtain the DNA banding patterns representative of the *Sulfobacillus* spp., subgroup I and *Sulfobacillus* spp., subgroup II profiles, respectively. MW indicates the molecular weight marker.

ARDREA of these *Sulfobacillus* isolates revealed banding patterns which corresponded to the profiles of the two *Sulfobacillus spp* subgroups as determined by Johnson *et al*., (2005). All the *Sb. t./Sb. m.*-like isolates had two DNA fragments, corresponding to 550 bp and 950 bp, respectively, while all the *Sb. a./Sb y.*-like isolates had a 1500 bp DNA fragment after their respective amplified 16S rDNA was digested with *Eco*1015 (*Sna*BI). Therefore the results obtained for the *Sulfobacillus* isolates *Sb. a.* ALV, TH1; *Sb. m.* YOO17 and *Sb. t.* VKM B-1269^T, MT13, for which ARDREA analysis has not been reported previously, were in agreement with results reported by Johnson *et al*., (2005).

4.3.3 The identification and distribution of ArsB homologues within the two *Sulfobacillus* spp. subgroups.

An important aspect of this study was to determine how widely distributed the *ars* genes are among members representative of genus *Sulfobacillus*. Furthermore, sequence analysis of any *ars* genes produced phylogenetic relationships that agreed with the phylogeny obtained using the 16S rDNA sequence data. Genomic DNA of the *Sulfobacillus* isolates was subjected to PCR amplification with the degenerate *arsB* primer set Ferro arsBfwd and Ferro arsBrev. These primers were designed to anneal to specific conserved regions within the putative *arsB* genes, resulting in the amplification of a DNA fragment corresponding to approximately 850 bp. Variations in the size of the amplified DNA fragment are dependent on the location of the primer target binding regions of the respective putative *arsB* genes. In order to identify any putative amplified *arsB* homologues, especially where more than one band was obtained, Southern-hybridization was performed using the 560 bp *Sb. t.* VKM B-1269^T *arsB* probe (Figure 4.5). Results are shown in Figure 4.6.

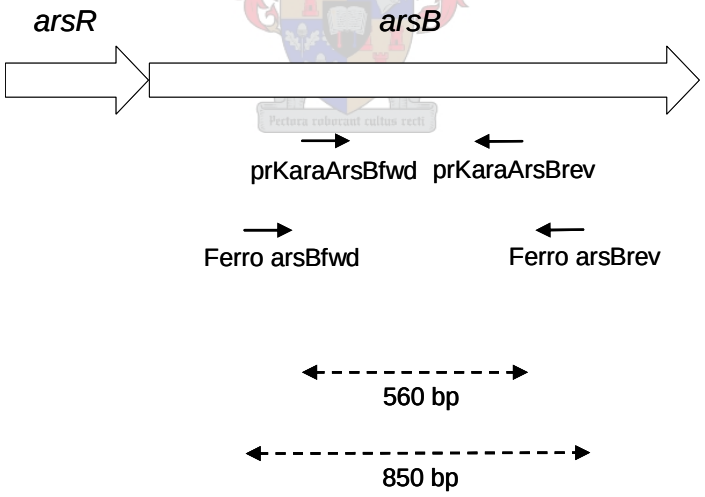


Figure 4.5: **A diagram showing the binding positions of Ferro arsBfwd and Ferro arsBrev with respect to the *arsB* of *Sb. t.* VKM B-1269^T.** Putative *arsB* genes were identified by means of Southern-hybridization using an amplified 560 bp *Sb. t.* VKM B-1269^T *arsB* fragment. This fragment was obtained using primer pair prKaraArsBfwd/prKaraArsBrev. The predicted product sizes of the primer pairs are indicated at the bottom.

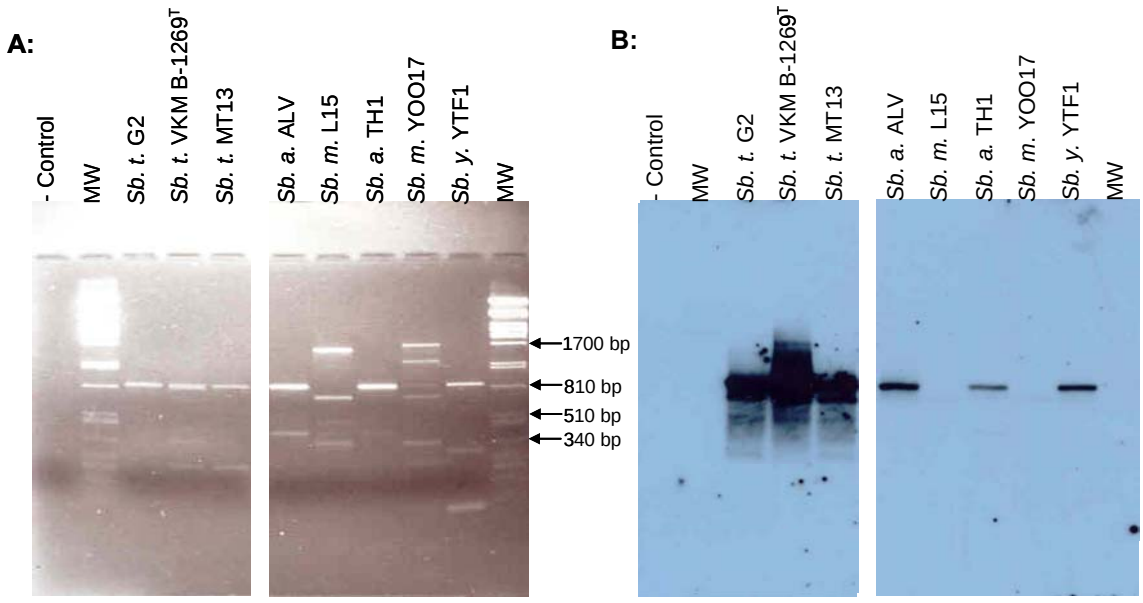


Figure 4.6: **The identification and distribution of putative *arsB* homologues within the respective *Sulfobacillus* isolates. (A)** Putative *arsB* genes of the *Sulfobacillus* spp. was amplified from total genomic DNA by using the degenerate primer set Ferro arsBfwd and Ferro arsBrev. **(B)** Autoradiograph after Southern-hybridization using the 560 bp *arsB* fragment of *Sb. t*. VKM B-1269^{T} as a probe. Minus template DNA acted as the negative control (lanes indicated as – Control), while the molecular weight marker is shown as MW.

PCR amplification of the genomic DNA of the *Sulfobacillus* spp. with the degenerate *arsB* primers resulted in an 850 bp DNA fragment in the majority of the tested *Sulfobacillus* isolates. Exceptions to this were observed in the case of *Sb. m.* L15 and *Sb. m.* YOO17. Although several DNA fragments were amplified during the PCR of these two isolates, none of these fragments were in the vicinity of the predicted 850 bp product (Figure 4.6 A). Southern-hybridization using the 560 bp *arsB* fragment of *Sb. t*. VKM B-1269^{T} as a probe, confirmed that all the amplified 850 bp products of the various *Sulfobacillus* spp. samples were *arsB* homologues. No hybridization signal was observed in the cases of *Sb. m.* L15 and *Sb. m*. YOO17 (Figure 4.6 B). The identified putative *arsB* fragments of the respective *Sulfobacillus* isolates were subsequently cloned and sequenced.

Phylogenetic comparisons of the different sequenced *arsB* homologues fragments of the *Sulfobacillus* spp. revealed that two clearly defined separate *arsB* clusters occurred within genus *Sulfobacillus* (Figure 4.7 A). Within the *arsB* cluster of *Sulfobacillus* spp. representing subgroup I, the *arsB* homologues shared high sequence similarity. Of interest is that the *arsB* fragment isolated from *Sb. t*. MT13, which came from South Africa, is very similar to the two *arsB* fragments of the other *Sb. t*. strains isolated from Russia (*Sb. t.* VKM B-1269^{T}) and Montserrat Island (*Sb. t.* G2). The isolated *arsB* fragments of the *Sb. a.* spp. show 99.8% (100%) sequence homology, with the *Sb. y.* YTF1 being the least similar member of this cluster representing members of the *Sulfobacillus* spp. subgroup II. The *arsB* genes that were grouped in each cluster was in agreement with the 16S rDNA sequence based groups. However, the percentage nucleotide sequence identity between the *arsB* gene sequences of each cluster (63%) was considerably lower than the 16S rDNA sequence identities (90%). Therefore, the *arsB* gene sequences are considerably less conserved than the 16S rDNA sequences (Figure 4.7 B). ARDREA of *arsB* genes could be used should one wish to identify the phylogenetic group to which an *arsB* gene belongs without sequencing the gene. Whether such an ARDREA analysis is of any value in a gene that has a higher rate of evolution than the 16S rDNA gene would require the confirmation of *arsB* ARDREA groups using more isolates of genus *Sulfobacillus*. The various *arsB* homologue fragments isolated from the *Sulfobacillus* spp. revealed four main profiles which can be identified on the basis of differences in the location of restriction enzyme sites located within these fragments (Figure 4.8).

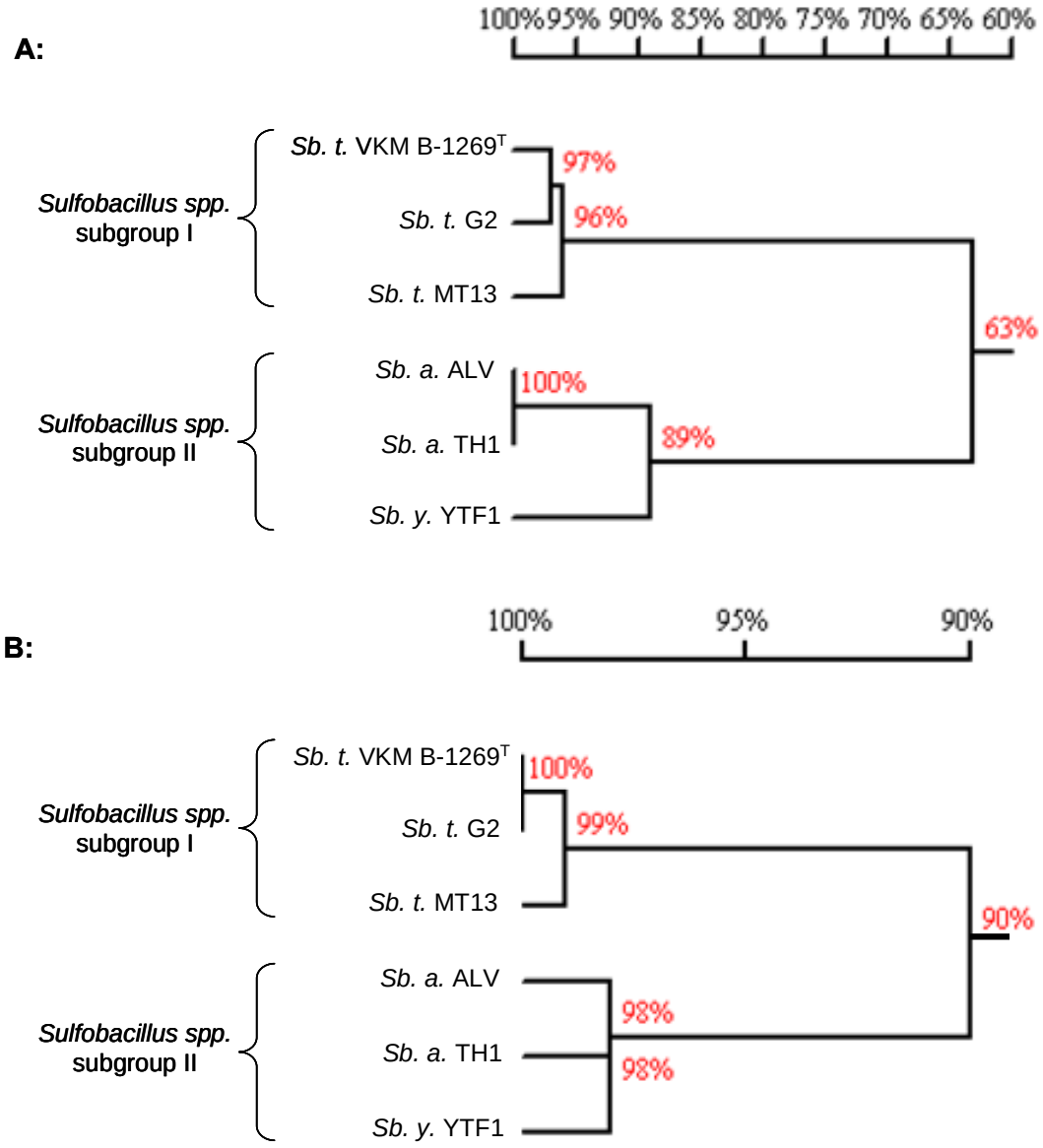


Figure 4.7: **Phylogenetic trees of *Sulfobacillus* spp. based on their identified *arsB* homologues or their respective 16S rDNA nucleotide sequences. (A)** Phylogenetic comparison of the *arsB* homologue fragments of the respective *Sulfobacillus* spp. **(B)** The phylogenetic comparison of the 16S rDNA nucleotide sequences of the relevant *Sulfobacillus* spp. in which a putative *ars* operon was identified. Accession numbers of the 16S rDNA nucleotide sequences that have been submitted to the GenBank database: *Sb. t.* G2 (AY140233); *Sb. t.* VKM B-1269^T (X91080) and *Sb. y.* YTF1 (YTF1AY007665). The 16S rDNA nucleotide sequences of the other *Sb.* isolates were reported in the M. Sc. thesis dissertation of M. Schutte, 2004.

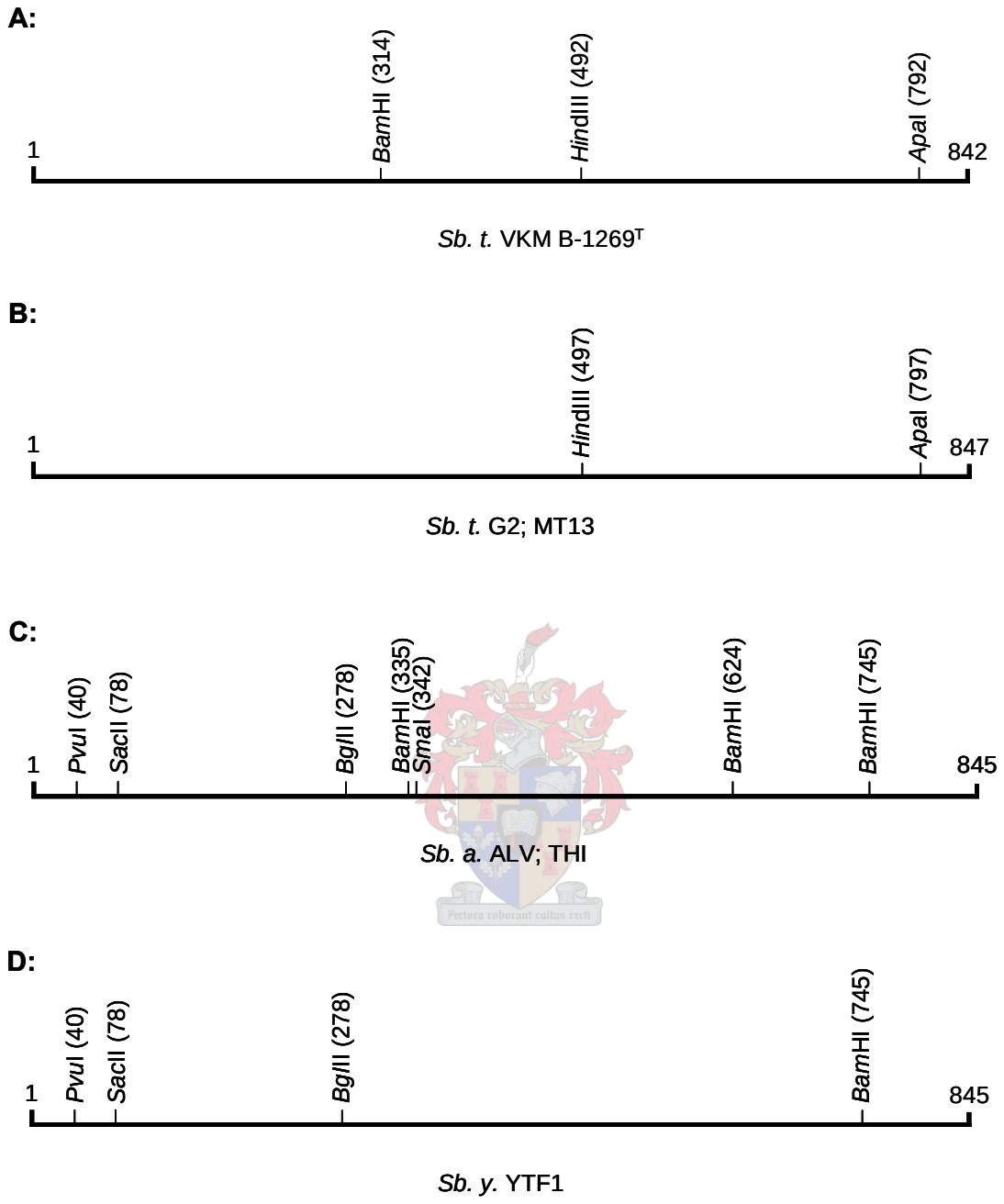


Figure 4.8: **Computer software-generated restriction enzyme maps showing the four main profiles of identified and sequenced sections of the respective *Sulfobacillus* spp. *arsB* homologues. (A)** *Sb. t.* VKM B-1269^{T}; **(B)** *Sb. t.* G2 and MT13; **(C)** *Sb. a.* ALV and TH1; **(D)** *St. y.* YTF1.

4.3.4 The *Sb. t.* isolates share common features with respect to the *arsRB* gene orientation and layout.

From the aggregate of results obtained from experiments performed in section 4.3.3, it is not only clear that the *Sb. t.* isolates are evolutionarily very similar on the basis of 16S rDNA gene comparisons, but also that the *arsB* homologues isolated from them shared high levels of sequence similarity. It was therefore decided to determine whether the layout of the various *Sb. t. ars* operons was similar and if they shared any common features with respect to the open reading frames present on the 5300 bp pStArs1 fragment originally isolated from *Sb. t.* VKM B-1269^T. Total genomic DNA of the *Sulfobacillus* isolates was subjected to PCR amplification using primers prKaraArsBfwd and prKaraArsBrev (Figure 4.9).

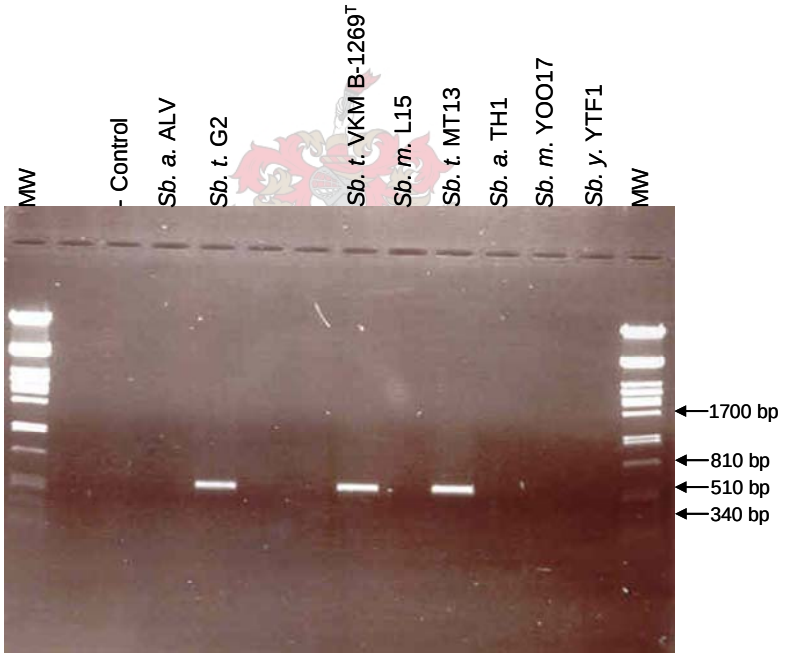


Figure 4.9: **PCR amplification of genomic DNA of the *Sb. t.* isolates with primer set prKaraArsBfwd and prKaraArsBrev.** Minus template DNA acted as the negative control (lane indicated as –Control), while the molecular weight marker is shown as MW.

The predicted 560 bp PCR product was only present in the *Sb. t.* isolates. This was expected as the *arsB* fragments isolated from the *Sb. t.* strains, including *Sb. t.* VKM B-1269^T, were shown to share a high percentage sequence identity. Because no 560 bp

PCR product was amplified in the other *Sulfobacillus* isolates, the attention was focused on determining the structural characteristics of the putative *ars* operons of the *Sb. t.* isolates G2 and MT13. Several primers were designed in order to determine orientation and layout similarities of the putative *ars* operons of *Sb. t.* G2 and MT13 with respect to the sequenced *ars* operon of *Sb. t.* VKM B-1269^{T}. Genomic DNA of the three *Sb .t.* isolates was subjected to PCR amplification, using primer sets prKumAsc + prRTArsR, prArsRc + prRTArsB and prKumAsc + prRTArsB, respectively. The pStArs1 *arsRB* operon and the predicted product sizes of the relevant primer pairs are shown in Figure 3.1 of Chapter 3 (page 70).

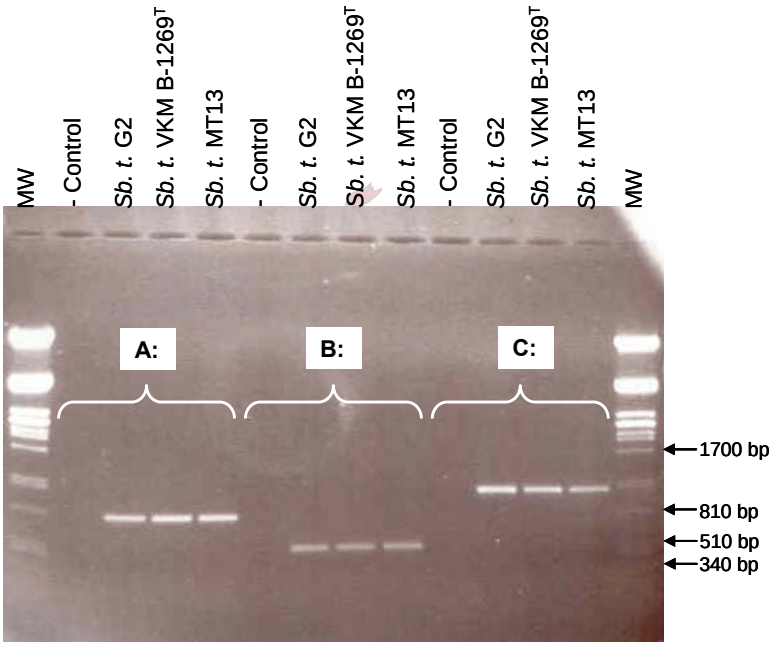


Figure 4.10: **PCR amplification of the genomic DNA of the *Sb. t.* isolates to determine layout similarities with respect to their putative *ars* operons.** Primer sets used were **(A)** prKumAsc + prRTArsR **(B)** prArsRc + prRTArsB and **(C)** prKumAsc + prRTArsB, respectively. Minus template DNA acted as the negative control (lanes indicated as –Control), while the molecular weight marker is shown as MW.

The predicted 882 bp PCR product (amplified by primers prKumAsc + prRTArsR), the 567 bp PCR product (amplified by primers prArsRc + prRTArsB) and the 1161 bp PCR product (amplified by prKumAsc + prRTArsB) were detected in all three of the

respective *Sb. t*. isolates (*Sb. t*. G2; VKM B-1269^T and MT13) (Figure 4.10). Therefore, the *ars* operons of the *Sb. t*. isolates are very similar in structure.

4.3.5 The *arsRB* operon of *Sb. t.* VKM B-1269^T is chromosomally located.

Trans-alternating field electrophoresis was performed on the genomic DNA of the *Sb. t*. isolates to determine similarities in terms of the location of their respective *ars* operons. The TAFE gel was subjected to Southern-hybridization using either a labeled 560 bp *arsB*, 288 bp *arsR* or an 1456 bp 16S rDNA fragment from *Sb. t*. VKM B-1269^T as probe (Figure 4.11). Genomic DNA of isolate *Sb. m*. YOO17 was included to act as a negative control for the *Sulfobacillus* spp. subgroup I *ars* operon probes.

Pulsed-field gel electrophoresis revealed the presence of intact chromosomal DNA in the wells and background smears comprised of degraded chromosomal DNA in all the *Sb. t*. isolates. In addition, plasmids within *Sb. t.* VKM B-1269^T (estimated sizes of 80 kb and 6 kb), *Sb. t.* MT13 (estimated size 170 kb) and *Sb. m*. YOO17 (estimated sizes 120 kb and 50 kb) were identified. The characteristics of these newly identified plasmids do not fall within the scope of this research project and will therefore not be discussed further. These plasmids did not hybridize to the 560 bp *arsB* probe of *St. t*. VKM B-1269^T after Southern-hybridization was performed, which strongly suggests that the respective *ars* operons harbored by these *Sulfobacillus* isolates are chromosomally located. As to be expected, no hybridization signals were found when any of the probes representative of *Sulfobacillus* spp. subgroup I *ars* operons were hybridized against the genomic DNA of isolate *Sb. m*. YOO17. Probing with the 1465 bp 16S rDNA probe of *Sb. t.* VKM B-1269^T confirmed the presence of *Sb. m*. YOO17 genomic DNA within the TAFE gel (Figure 4.11 C). In contrast, hybridization signals were obtained in all the *Sb. t*. isolates G2, VKM B-1269^T and MT13, using the 560 bp *arsB*, 288 bp *arsR* and 1456 bp 16S rDNA of *Sb. t.* VKM B-1269^T as probes and shared similarities with respect to each other (Figure 4.11). Irrespective of whether *arsR*, *arsB* or 16S rDNA fragments were used as probes, all of the Southern-hybridization gave very similar results. This was to be expected if all three genes are present on the chromosome of these *Sb. t.* isolates.

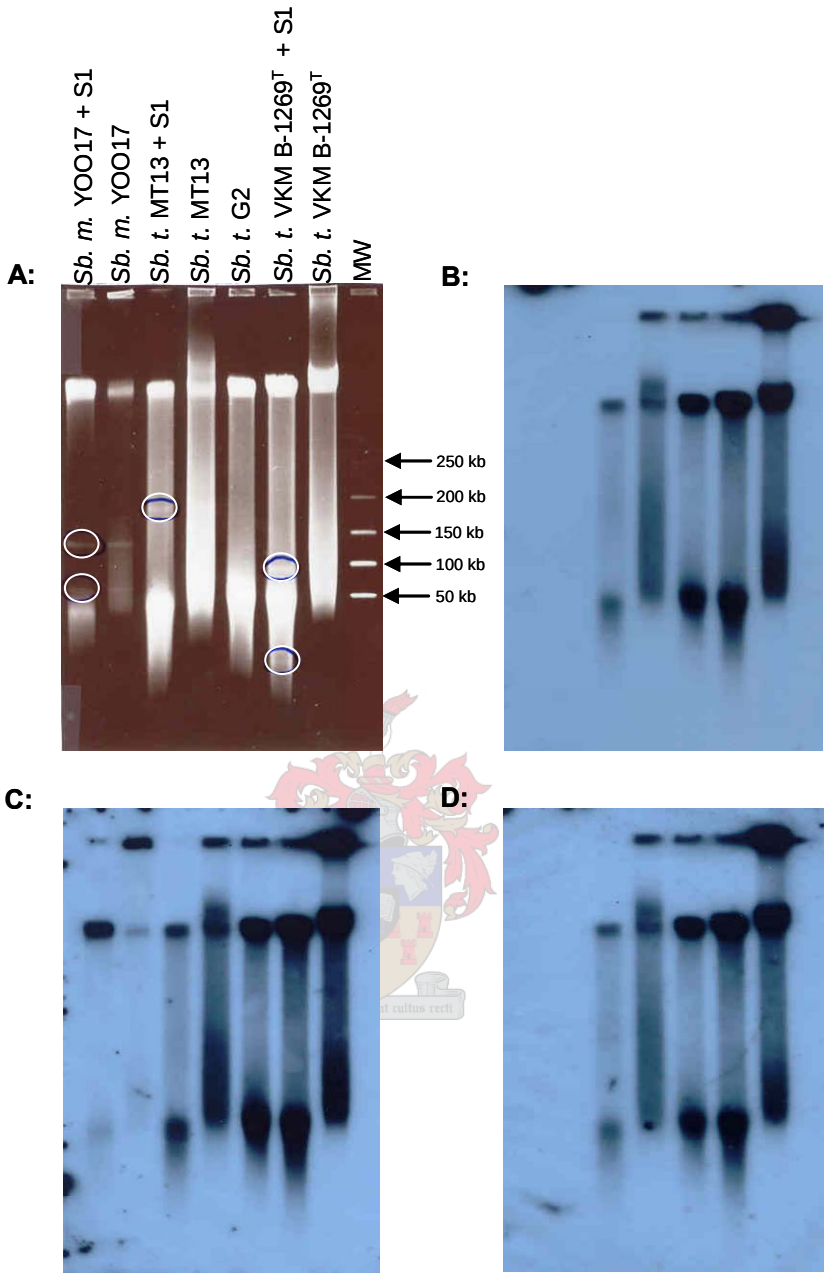


Figure 4.11: **Trans-alternating field electrophoresis performed on the genomic DNA of the *Sb. t.* isolates to determine the location of their respective *ars* operons. (A)** A TAFE gel containing the uncut genomic DNA of the tested *Sulfobacillus* spp. showing the presence of putative plasmids, indicated by white circles, within certain of the tested strains. Autoradiographs after Southern-hybridization using the **(B)** 560 bp *arsB*, **(C)** 1456 bp 16S rDNA and **(D)** 288 bp *arsR* probes of *Sb. t.* VKM B-1269^T, respectively. Genomic DNA subjected to S1 nuclease is indicated. The MW indicates the molecular weight marker. The TAFE gel was kindly performed by Dr. S.M. Deane.

4.4 Discussion

The identification and differentiation of bacterial species has been accomplished by a number of different approaches (Krieg, 1988). Strain identification and taxonomical classification with microbiologically-based (phylotypically-based) methods lack the ability to accurately make genus and species affiliations of bacterial isolates. The implementation of modern molecular biologically-based techniques facilitates a more effective approach with respect to accurate bacterial strain sub-typing. One technique that has successfully been utilized in this context is amplified ribosomal DNA restriction enzyme analysis (ARDREA), also known as RFLP (restriction fragment length polymorphism). The operating principle of this method is that differences in the 16S rDNA sequences of different species will create differences in the location of sites for various restriction enzymes, revealing a unique fingerprint for each species or strain after electrophoretic separation on a regular agarose gel. ARDREA is therefore a rapid method that does not require sequencing of the 16S rDNA gene (Johnson *et al*., 2005; Spiegelman *et al*., 2005).

There is sufficient diversity in 16S rDNA gene sequences at a species level so that microorganisms isolated from different sources at different times and from different geographical regions may be differentiated into subgroups within specific species (Olive and Bean, 1999). In the light of this statement, Johnson and coworkers (Johnson *et al*., 2005) applied an ARDREA protocol to successfully distinguish between *Sulfobacillus*-like isolates at a species level. The *Sulfobacillus* isolates used in my study were subjected to this ARDREA protocol and revealed that they could be divided into the proposed *Sulfobacillus* spp. subgroup I and subgroup II, respectively. The observation that *Sulfobacillus* isolates different from those used by Johnson and coworkers (Johnson *et al.*, 2005) could effectively be divided into the suggested *Sulfobacillus* spp. subgroups, indicate the robustness of ARDREA using the restriction endonuclease *Eco*1015 (*Sna*BI). The isolates representative of the respective *Sulfobacillus* spp. subgroups, determined during the course of this study, are indicated in Table 4.2.

Table 4.2: The *Sulfobacillus* isolates representative of the two *Sulfobacillus* spp., subgroups.

***Sulfobacillus* spp., subgroup I**
Sb. thermosulfidooxidans G2; VKM B-1269^{T}; MT13
Sb. montserratensis L15; YOO17
***Sulfobacillus* spp., subgroup II**
Sb. acidophilus ALV; TH1
Sb. yellowstonensis YTF1

ARDREA analysis using *Xcm*I was performed on isolate *Sb. m*. L15 and revealed the mesophilic Gram-positive acidophilic bacteria profile suggested by Johnson *et al*., (2005). In addition, ARDREA analysis performed on isolate *Sb. m*. YOO17 showed a DNA band profile corresponding to the suggested mesophilic Gram-positive acidophilic fragment pattern when *Xcm*I was used, but no similarities were found to the proposed moderately thermophilic Gram-positive acidophilic fragment pattern when *Eco*721 was utilized. ARDREA performed with *Eco*721 on *Sb. m.* L 15 revealed similar results with respect to those found in case of *Sb. m* YOO17, suggesting that *Sb. m.* YOO17 can be classified as a mesophilic Gram-positive acidophile on the basis of ARDREA results obtained when using *Xcm*I and *Eco*721 together. Although it was considered in the past that genus *Sulfobacillus* contained exclusively moderately thermophilic Gram-positive acidophilic bacteria, it is now clear that the genus may also include mesophilic Gram-positive, acidophilic bacteria isolates such as *Sb. m.* L15 and *Sb. m.* YOO17. ARDREA analysis performed with *Eco*1015 (*Sna*BI) on *Sb. m*. L15 and *Sb*. *m.* YOO17 revealed that both isolates had the DNA band profile which corresponded to members of the *Sulfobacillus* spp. subgroup I (Johnson *et al*., 2005).

An important question that needed to be addressed during the course of this study was to determine how widely distributed and related the *ars* genes are among genus *Sulfobacillus*. Degenerate *arsB* primers were used to amplify a predicted 850 bp

fragment from genomic DNA of the tested *Sulfobacillus* isolates. Although several additional fragments were detected, the predicted 850 bp PCR fragment was present in the majority of these *Sulfobacillus* spp. The *Sb. m*. isolates L15 and YOO17, however, gave more amplified fragments than was expected. Because degenerate primers were used for putative *arsB* amplifications and identification, the possibility could not be excluded that an amplified fragment of a size other than 850 bp fragment, might possibly be a putative *arsB*. Southern-hybridization by using the 560 bp *arsB* probe of *Sb. t*. VKM B-1269^{T} was used to confirm the presence of putative *arsB*'s within the *Sulfobacillus* isolates. A hybridization signal was obtained in all the cases where the amplified 850 bp fragment was present, confirming the presence of *arsB* homologues. No hybridization signals were obtained in the case of the *Sb. m*. strains L15 and YOO17, suggesting the absence of any *arsB*'s with sequence homology to the designed degenerate primers and the 560 bp *arsB* probe of *Sb*. *t*. VKM B-1269^{T}. Because *ars* genes are wide spread, the mesophilic *Sulfobacillus* isolates *Sb m*. L15 and YOO17 would be expected to have *ars* operons but these are apparently likely to be different from those found in the moderately thermophilic *Sulfobacillus* isolates.

Subsequent phylogenetic sequence comparisons of the *arsB* homologues of the *Sulfobacillus* spp., revealed two clearly defined separate *arsB* clusters within genus *Sulfobacillus*. There was an average of 63% sequence identity between the clusters, with each cluster being comprised of either *Sulfobacillus* spp. subgroup I or subgroup II. To determine the evolutionary relatedness of these *arsB* fragments with respect to their species classification, phylogenetic comparisons of their respective 16S rDNA sequences were done and a profile almost identical to the *arsB*'s profile was obtained. This is conclusive evidence that the isolated *arsB* fragment of a specific *Sulfobacillus* sub specie is distinctive and characteristic to that of the *Sulfobacillus* sub specie. they represent. It is therefore unlikely that the *arsB*'s of the *Sulfobacillus* spp. were obtained by means of recent horizontal gene transfer due to selective pressures caused by arsenic present in their respective immediate environments. Sequence analysis of the isolated *arsB* homologue fragments from the *Sulfobacillus* spp. reveal that four distinctive profiles

could be identified based on differences in the location of restriction endonuclease recognition sites.

The *Sb. t.* isolates were shown to be evolutionarily very similar on the basis of 16S rDNA gene comparisons and also that the *arsB* homologues isolated from them shared high levels of sequence homology. The products obtained through PCR amplification showed that the open reading frames coding for the kumamolisin-As precursor, the *arsR* and the *arsB* are present in *Sb. t.* G2 and *Sb. t.* MT13 and that they are similar in orientation and layout with respect to the 5300 bp pStArs1 fragment originally isolated from *Sb. t.* VKM B-1269^{T}. Because *ars* operons may be located on chromosomes, plasmids or transposons, it was important to determine where these putative *ars* operons were situated within the genome of the respective *Sb. t.* isolates. Trans-alternating field electrophorsis was performed on the genomic DNA of the *Sb. t*. isolates (*Sb. t.* G2, VKM B-1269^{T} and MT13), followed by Southern-hybridization. Hybridization signals indicated that the respective *ars* operons present within all three tested *Sulfobacillus* isolates were chromosomally located.

CHAPTER 5

GENERAL DISCUSSION

TABLE OF CONTENT

5.1 The research in perspective 113

5.2 Future prospects 117

5.1 The research in perspective

The constant exposure to selective pressures caused by continuous environmental changes will force the individuals of any microbial population to adapt to these changes to ensure their own survival and proliferation. The presence of arsenical compounds in the environment selects and maintains microbes possessing genetic determinants (*ars* genes) which confer resistance to arsenic. It is known that *ars* genes are widely distributed in *Bacteria*, *Archaea* and *Eukarya* and are usually located on chromosomes, plasmids or transposable elements. Two of the most commonly encountered forms of *ars* operon arrangements consist of five genes (*arsRDABC*) and three genes (*arsRBC*), respectively. The five gene *ars* operon has been found on plasmids of Gram-negative bacteria (plasmids R773 and pR46 of *E. coli* and pKW301 of *Acidiphilium multivorum*), while the three gene *ars* operon is associated with plasmids of Gram-positive bacteria (*Staphylococcus aureus* (pI258); *Staphylococcus xylosus* (pSX267)) and the chromosomes of Gram-negative bacteria *E. coli*, *Pseudomonas aeroginosa* and *Pseudomonas fluorescens*. These *ars* operons are transcribed from a single operator/promoter region. With the advent of genome sequencing projects, it has become apparent that the layout and transcription of the *ars* genes may differ from the conventional five and three gene *ars* operons (Table 1.1).

Bacteria representative of genus *Sulfobacillus* could have considerable potential for use in commercial bio-oxidation of mineral ores and concentrates at elevated temperatures (>40°C). During the microbial-aided mineral solubilization of arsenopyrite (FeAsS) containing ores, toxic arsenical compounds are leached into the biooxidation tanks. Little is known about the genetic systems and regulatory mechanisms conferring resistance to arsenical compounds in members of genus *Sulfobacillus*. This is the first study to describe the isolation and sequencing of the *ars* genes from the chromosome of a Gram-positive acidophile. The *ars* operon of *Sulfobacillus thermosulfidooxidans* VKM B-1269^T (*Sb. t.* VKM B-1269^T) consists of an *arsR* (codes for a transcriptional regulator) and *arsB* (codes for an arsenite efflux pump), with the *arsRB* transcribed in the same direction. Although the *arsRB* operon seems simple in comparison with other *ars*

operons, this two gene configuration is not unique. Two other cases of *ars* operons containing this two-gene configuration have been reported to exist in the bacterium *B. subtilis* (Rosen,1999) and the archaeon *Ferroplasma acidarmanus* (Gihring *et al*., 2003). Rosen, (1999) has speculated that two-gene *arsRB* operons would potentially give resistance to only trivalent metalloid salts, indicating that the *arsRB* gene configuration potentially represents an early stage in the evolutionary development of *ars* operons. In order to confer broad-range resistance to arsenical compounds, the presence of an *arsC* (codes for an arsenate reductase) gene is required. An *arsC* was not associated with the *arsRB* operon of *Sb. t.* VKM B-1269^T, but the possibility that an *arsC* homolog might be located somewhere else in the *Sb. t.* VKM B-1269^T genome could not be excluded as $ArsC_{ec}$ homologues lacking other *ars* genes have been reported in *Haemophilus influenzae* and *Neisseria gonorrhoeae* (Rosen, 1999). PCR and Southern-hybridization experiments suggested that no ArsC, representative of either the Grx/GSH or Trx ArsC families, was present in the genome of *Sb. t.* VKM B-1269^T. The ability of the *arsRB* operon to confer resistance to As(V) and As (III) in its native *Sulfobacillus* host was examined by calculating the minimum inhibitory arsenic concentration (MICs) in liquid batch cultures. Unfortunately several problems were encountered in the attempt to accurately determine this. The formation of iron-containing compounds and arsenic precipitation in the $FeSO_4$ media during cell growth, made it difficult to draw any conclusions about the tested As(V) and As(III) values. If an alternative growth media could be developed to limit the precipitation problem, determining to what extent this *arsRB* operon contributes to arsenic resistance within *Sb. t*. VKM B-1269^T will still be problematic. This is largely due to the lack of suitable arsenic sensitive *Sb. t.* mutants with which to perform comparative analysis.

The *arsRB* operon of *Sb. t*. VKM B-1269T was expressed in its native *Sulfobacillus* host. Analysis of total mRNA revealed that the *ars* genes were co-transcribed from a single operator/promoter region, located upstream of the *arsR*. The *ars* operon of *Sb. t.* VKM B-1269^T was unable to complement an *E. coli* arsenic sensitive mutant in the presence of arsenic. Although *E. coli* are generally satisfactory as a prokaryotic expression system for heterologous protein production purposes, the production of active/functional proteins

in *E. coli* is challenging. Significant problems in post-translational modification of heterologous proteins in *E. coli* are frequently encountered. It was clear from mRNA transcript analysis and promoter expression studies that one or more processes (mRNA transcript processing, mRNA transcript translational efficiency, the effectiveness of protein folding and transport) may be responsible for the inability of the *arsRB* operon of *Sb. t.*, strain VKM B-1269^T, to compliment the arsenic sensitive heterologous *E. coli* host.

The lipid composition of cytoplasmic membranes is very complex and differs strongly between bacteria. The fluidity and permeability properties of the cell membrane are largely determined by their lipid composition and may be adjusted in response to changes in environmental parameters such as temperature, pH and pressure (Konings *et al*., 2002). Acidophilic bacteria are forced to maintain a steep proton gradient across the cytoplasmic membrane to ensure survival in extreme pH environments. This can only be realized with a specific membrane composition (Driessen *et al*., 1996). Successful expression of acidophilic Gram-negative bacterial *ars* genes in the neutrophilic Gram-negative *E. coli* heterologous host have been reported (Butcher *et al*., 2000; Tuffin *et al*., 2004; Kotzé *et al*., 2006; Tuffin *et al*., 2006). This implied that the membrane-associated acidophilic *arsB* was effectively incorporated in the neutrophilic *E. coli* cytoplasmic membrane, despite differences in their cytoplasmic lipid composition and organization. The unique fatty acid composition of the lipids comprising the membranes of Sulfobacilli cells may contribute to the formation of multilayer structures within the cytoplasmic membranes of these bacteria. Duna and co-workers (Duda *et al*., 2001) have shown that the cytoplasmic membrane of *Sb. t.,* strain VKM B-1296^T, contains multilayer regions in the form of flat lamellar inverted lipid membranes. These multilayer structures are likely to be involved in processes related to bacterial cell growth and differentiation, but may also play a significant role in the oxidation of sulfur containing compounds. The involvement of these multilayer structures in membrane-associated transport systems remains unclear (Taylor and Wirsen, 1997). Cases where Gram-positive *ars* genes were successfully expressed in a heterologous *E. coli* host, have been reported (Rosenstein *et al*., 1992; Broër *et al*., 1993). Although certain post-translational modifications are probably

beyond the reach of *E. coli*, the nature of the physical structure of the *Sb. t*. VKM B-1269^{T} ArsB and the defective incorporation of this ArsB into the *E. coli* cytoplasmic membrane, could possibly contribute to the inability of the *arsRB* operon of *Sb. t.* VKM B-1269^{T}, to confer resistance to arsenic in an arsenic sensitive *E. coli* mutant.

The application of amplified ribosomal DNA restriction enzyme analysis (ARDREA) as a tool to accurately identify and discriminate between newly discovered *Sulfobacillus* isolates at a species level was demonstrated by Johnson *et al*., (2005). Eight *Sulfobacillus* strains, isolated from different geographical areas, were subjected to ARDREA using the restriction endonuclease *Eco*1015 (*Sna*BI). The observation that *Sulfobacillus* isolates, different from those used by Johnson and coworkers (Johnson *et al*., 2005) could effectively be divided into the suggested *Sulfobacillus* spp. subgroup I, containing *Sb. thermosulfidooxidans*/*Sb. montserratensis*-like isolates and subgroup II, containing *Sb. acidophilus*/*Sb. yellowstonensis*-like isolates, indicated the robustness of ARDREA as a molecular typing method for previously unclassified *Sulfobacillus* isolates.

One objective of this study was to detect the presence of *ars* genes in other *Sulfobacillus* isolates and to determine how widely distributed and related these newly discovered *ars* genes were among genus *Sulfobacillus*. Several *arsB* homologues were identified and isolated from *Sulfobacillus* isolates classified as moderately thermophilic. In contrast, no *arsB* homologues were identified in the mesophilic *Sb. m.* isolates L15 and YOO17 when PCR, using the same degenerate primers, and Southern-hybridization experiments were performed. It is important to note that sections of the *arsB* homologues were initially amplified by degenerate *arsB* primers. These primers were designed to anneal internally within the putative *arsB* homologues, which resulted in the amplification of only an 850 bp *arsB* area of the whole *arsB* homologue. Phylogenetic sequence comparisons of the *arsB* fragments revealed two clearly separate *arsB* clusters within genus *Sulfobacillus*. The two clusters shared 63% sequence identity and showed that the *arsB* of a specific *Sulfobacillus* sub specie is distinctive to that specific *Sulfobacillus* sub specie. Phylogenetic comparisons of the *Sulfobacillus* 16S rDNA sequences and their respective *arsB* fragment sequence revealed the evolutionary relatedness of the *arsB* fragments with

respect to their specific species classification. The 16S rDNA and *arsB* phylogenetic profiles were very similar to each other, suggesting that it was unlikely that the *arsB* of a *Sulfobacillus* spp. was obtained by means of horizontal gene transfer due to selective pressures caused by the presence of arsenic in their respective habitats. Sequence analysis of the isolated *arsB* homologue fragments from the *Sulfobacillus* spp. revealed that four distinctive profiles could be identified based on differences in the location of restriction endonuclease sites. These genetic maps could be potentially used as a guide to identify and discriminate between newly discovered *Sulfobacillus arsB* homologues.

5.2 Future prospects

The *arsRB* operon of *Sb. t*. VKM B-1269^T was not functional in the heterologous *E. coli* arsenic sensitive mutant due to possible problems encountered during the production of functional proteins from the *ars* operon transcript. Future work based on our preliminary findings could involve Western-blot analysis on the cytosolic and membrane protein fractions isolated from the *E. coli* arsenic sensitive mutant, using ArsB-specific antibodies for detection purposes. This would give us a clear indication whether the *arsB* is indeed translated in the heterologous *E. coli* host, and if so, if the translated *arsB* polypeptide is associated with the Gram-negative cytoplasmic membrane. Recommended future investigations will be the transformation of the *arsRB* operon of *Sb. t*. VKM B-1269^T in a heterologous Gram-positive host. *Bacillus subtilis* has come to be regarded as the experimental model for Gram-positive bacteria. As was mentioned earlier, *Bacillus subtilis* harbors two different *ars* operons (Rosen, 1999). One of these *ars* operons, comprised of *aseR* and *aseA*, was the first reported bacterial *ars* operon with a two-gene configuration. Although *aseA* shares only 54% sequence identity with the *arsB* of the *Sb. t*. VKM B-1269^T, *ars* operon, it will be interesting to determine if the *arsB* will complement the *B. subtilis* arsenic sensitive mutants HB5007 (CU1065 *arsR-C::tet*;*aseA*::cm) and HB5004 (CU1065 *aseA::cm*) (Moore *et al*., 2005).

Determining the role of the *arsRB* operon of *Sb. t.* VKM B-1269^T with respect to arsenic resistance in its native host, presents considerable challenges. No significant conclusions

with respect to the ability of the *Sb. t.* VKM B-1269^{T} *ars* operon to confer resistance to arsenical compounds could be drawn, due to the formation of iron-containing compounds and arsenic precipitation in the $FeSO_4$ media during cell growth. Future work will have to include the development of growth media that limits precipitation formation. Futhermore, an in depth study to develop a mechanism for the transfer of genetic material into Sulfobacilli is needed to create arsenic sensitive *Sulfobacillus* spp. knockout mutants. These *Sulfobacillus* mutants could be important for comparative analysis proposes and facilitate a means to better understand the expression and regulatory processes involved in the *ars* operon of *Sb. t*. VKM B-1269^{T} and other *Sulfobacillus* spp. *ars* operons.

APPENDIX 1: Description of additional clones and primers used during the sequencing of pStArs1.

Plasmids or primers	Description of construct
Plasmids	
pEB3.9	Ap^r; (3.9 kb *Eco*RI-*Bam*HI fragment of pStArs1 cloned into pUCBM21 digested with *Eco*RI and *Bam*HI)
pPB1.6	Ap^r; (1.6 kb *Bam*HI-*Pst*I fragment of pStArs1 cloned into pUCBM21 digested with *Bam*HI and *Pst*I)
pCX0.34	Ap^r; (0.34 kb *Cla*I-*Xho*I fragment of pStArs1 cloned into pBluescript SK^+ digested with *Cla*I and *Xho*I)
pHX0.5	Ap^r; (0.5 kb *Hin*dIII-*Xho*I fragment of pStArs1 cloned into pBluescript SK^+ digested with *Hin*dIII and *Xho*I)
pHH1.3	Ap^r; (1.3 kb *Hin*dIII-*Hin*dIII fragment of pStArs1 cloned into pUCBM21 digested with *Hin*dIII)
pHH1.1	Ap^r; (1.1 kb *Hin*dIII-*Hin*dIII fragment of pStArs1 cloned into pUCBM21 digested with *Hin*dIII)
pSacB0.7	Ap^r; (0.7 kb *Bam*HI-*Sac*I fragment of pStArs1 cloned into pBluescript SK^+ digested with *Bam*HI and *Sac*I)
pPH1.1	Ap^r; (1.1 kb *Hin*dIII-*Pst*I fragment of pStArs1 cloned into pUCBM21 digested with *Hin*dIII and *Pst*I)
pSacSph0.56	Ap^r; (0.56 *Sac*I-*Sph*I fragment of pStArs1 cloned into pUCBM21 digested with *Sac*I and *Sph*I)
Primers	
pArsRBrev	GCTATGGACACGAGTCCG
prCX0.34fwd	GCAGATGCCGGAATT
prHX0.5fwd	CCGCATTTCCTACAAA
prKumAsfwd	CCTTTGGCAAACTTGGTGG

Ap^r: ampicillin resistance

APPENDIX 2: Annotated sequence obtained from pStArs1.

Shown below is the double stranded sequence of the *Sb. thermosulfidooxidans* VKM B-1269^{T} chromosomal fragment contained in pStArs1 determined during the course of this study. Protein translations of the three complete open reading frames encoded on this fragment have been shown below the DNA sequence with the name and direction of transcription of each open reading frame labeled in grey boxes with arrows. Incomplete open reading frames are not indicated. The translational start codons of the open reading frames are boxed and the putative ribosomal binding sites of the kumamolisin-As precursor gene and *arsR* are indicated by a dotted line. Restriction endonuclease recognition sites are indicated by bold type and labeled above the DNA sequence. The sequence to which primers were designed is underlined and labeled.

```
        EcoRI                          SacII
1       GGAATTCGTTTTAGTGGTCAATTCCAGCCGCGGCCAACAGTCAGCTCGTCAAGCCGTCAA
        CCTTAAGCAAAATCACCAGTTAAGGTCGGCGCCGGTTGTCAGTCGAGCAGTTCGGCAGTT

61      GCGAATAATGGATGTTCTCTTCTCATTTTTGGGGATCATCGTGACGTTGCCCATTATGGC
        CGCTTATTACCTACAAGAGAAGAGTAAAAACCCCTAGTAGCACTGCAACGGGTAATACCG

121     GCTGGTGGCATTGTTAATCAAATGGGACAGTCCTGATGGCCCCGTCTTTTTTGTTCAACA
        CGACCACCGTAACAATTAGTTTACCCTGTCAGGACTACCGGGGCAGAAAAAACAAGTTGT

                                                              NcoI
181     ACGGGTTGGTCTTAACGGGCGCTTATTTCCCTGTTTCAAGTTTCGGACCATGGTTCCCAA
        TGCCCAACCAGAATTGCCCGCGAATAAAGGGACAAAGTTCAAAGCCTGGTACCAAGGGTT

241     TGCCGAAAGCCTTAAAGAAGAAATTGCCCATTTGAATGTTATGAGTGGTCCCGTGTTTAA
        ACGGCTTTCGGAATTTCTTCTTTAACGGGTAAACTTACAATACTCACCAGGGCACAAATT

301     GGTGCCAGATGATCCCCGTGTGACCCGCATTGGAAAGTTTTTACGCAAAACGAGCCTCGA
        CCACGGTCTACTAGGGGCACACTGGGCGTAACCTTTCAAAAATGCGTTTTGCTCGGAGCT

361     TGAACTCCCCCAATTACTGAATGTCTTATTAGGTCACATGAGCTTAGTGGGACCGAGACC
        ACTTGAGGGGGTTAATGACTTACAGAATAATCCAGTGTACTCGAATCACCCTGGCTCTGG
```

prKumlacfwd800

```
421   GGCTCTTCCCTCTGAAGTCGAACGCTACGGGGAAGACTACCGTAAACGCTTATCGGTCCG
      CCGAGAAGGGAGACTTCAGCTTGCGATGCCCCTTCTGATGGCATTTGCGAATAGCCAGGC

                                                        ClaI
481   TCCTGGCATTACGTGTTTATGGCAAATTTCAGGGCGCAATGAAATCGATTCATGGTCCCT
      AGGACCGTAATGCACAAATACCGTTTAAAGTCCCGCGTTACTTTAGCTAAGTACCAGGGA

541   CTTGTTAGACTTCAAAATTCTTTTGAAAACTATTCCCGCTGTTCTTCAACAAAAAGGTGC
      GAACAATCTGAAGTTTTAAGAAAACTTTTGATAAGGGCGACAAGAAGTTGTTTTTCCACG

601   TCACTAGGCGGATAAAGCCGGAGAGGGGAGATGGCATAAGCTATTCTTCTCTCTCCACCT
      AGTGATCCGCCTATTTCGGCCTCTCCCCTCTACCGTATTCGATAAGAAGAGAGAGGTGGA

661   TTGCATCCTGAATCTCCCTATCTTATGATGGATGAAGATATTTCCAAATCATTATGAGGG
      AACGTAGGACTTAGAGGGATAGAATACTACCTACTTCTATAAAGGTTTAGTAATACTCCC

721   AGGTAGCAGATGCCGGAATTAGAGCCCCAACTCTTGCATGTGATTGGGGATGCACCTGAA
      TCCATCGTCTACGGCCTTAATCTCGGGGTTGAGAACGTACACTAACCCCTACGTGGACTT

          NcoI                                          NcoI
781   GGGCCATGGAGCGTATTTTCATTGCCTACAGGAGACCACACGGCCATGGTCTCGGTGATA
      CCCGGTACCTCGCATAAAAGTAACGGATGTCCTCTGGTGTGCCGGTACCAGAGCCACTAT

                prKumlacfwd400
841   ATTCCACGCTCACTTTGGCTCGAAATCAGCAGGCGCGATCCCTTTTCTTCTGATTTGTCA
      TAAGGTGCGAGTGAAACCGAGCTTTAGTCGTCCGCGCTAGGGAAAAGAAGACTAAACAGT

                                                              XhoI
901   CTTATTGAGCGTATTGGGCGCATGGCAATTTTGCACCGGCTACAACAAACTGGGGAACTC
      GAATAACTCGCATAACCCGCGTACCGTTAAAACGTGGCCGATGTTGTTTGACCCCTTGAG

                                                             NcoI
961   GAGACGATTGTTGTAGACACTGATGACATTCAAGAGTTATGGAAAAAGCCTGATGAACCA
      CTCTGCTAACAACATCTGTGACTACTGTAAGTTCTCAATACCTTTTTCGGACTACTTGGT

1021  TGGTATATGACCTTACGGCGCTGTGGACAGTGTCATGAGATGGTCCCTCATGGGGAAGTT
      ACCATATACTGGAATGCCGCGACACCTGTCACAGTACTCTACCAGGGAGTACCCCTTCAA

          StuI
1081  CTTGAGGCCTTGGCCAATGCTTTGCCGCCTAATTCACGGGGGCAAATTACGGTGGAAGTG
      GAACTCCGGAACCGGTTACGAAACGGCGGATTAAGTGCCCCCGTTTAATGCCACCTTCAC
```

*Nco*I

1141 CTGTGCCCATCTTGTATGGTGCAAACATCCCATGTTCTTAAC**CCATGG**GGAGTAGTTGAA
GACACGGGTAGAACATACCACGTTTGTAGGGTACAAGAATTG**GGTACC**CCTCATCAACTT

Kumamolisin-As precursor start →

prKumlacrev

1201 CGGTAGCGTTTTAGGGAGAAAGGTGGGCGGGATACGATGCAGGCGAAACTAGTGAGAGCA
GCCATCGCAAAATCCCTCTTTCCACCCGCCCTATGCTACGTCCGCTTTGATCACTCTCGT
M Q A K L V R A

1261 ACTGGACATCTATTGCCCGAGGAAACCCTTCAGGATTATCACCGCATTTCCTACAAAGAG
TGACCTGTAGATAACGGGCTCCTTTGGGAAGTCCTAATAGTGGCGTAAAGGATGTTTCTC
T G H L L P E E T L Q D Y H R I S Y K E

*Sph*I

1321 CGGACCACCCAGTGG**GCATGC**GTCTTCAAATCTCGTCACGAAACGGACCTTAAAAGACGA
GCCTGGTGGGTCACC**CGTACG**CAGAAGTTTAGAGCAGTGCTTTGCCTGGAATTTTCTGCT
R T T Q W A C V F K S R H E T D L K R R

1381 TTGGCTTACGATTCCAGCGTGTTAACGCGTGAGGAGGTCCTCCAACTTTATGGACCTGAT
AACCGAATGCTAAGGTCGCACAATTGCGCACTCCTCCAGGAGGTTGAAATACCTGGACTA
L A Y D S S V L T R E E V L Q L Y G P D

1441 CCTGATCTTATCGACCGGGCCCGTCAATGGCTAAGTCGCCACGGGGTAAGAGTGTTGAAA
GGACTAGAATAGCTGGCCCGGGCAGTTACCGATTCAGCGGTGCCCCATTCTCACAACTTT
P D L I D R A R Q W L S R H G V R V L K

*Hin*dIII

1501 CAAGATGGCTTTATTTTGTGGTTGCAAGG**AAGCTT**GGGTCAAATTGAAGAAACGTTGAAA
GTTCTACCGAAATAAAACACCAACGTTCC**TTCGAA**CCCAGTTTAACTTCTTTGCAACTTT
Q D G F I L W L Q G S L G Q I E E T L K

1561 ATCCCCTTTGGGGAAAAAGATGGACAGTTTATGCCGCTCCGTGAACCCTTGGTGCCTGAA
TAGGGGAAACCCCTTTTTCTACCTGTCAAATACGGCGAGGCACTTGGGAACCACGGACTT
I P F G E K D G Q F M P L R E P L V P E

1621 TGGCTCGCTCCCCATATTGTTGGATTTGTGGGCTTGGAAAACGTGTCTAAGTTATATCCC
ACCGAGCGAGGGGTATAACAACCTAAACACCCGAACCTTTTGCACAGATTCAATATAGGG
W L A P H I V G F V G L E N V S K L Y P

1681 CGTTTTCGATTCCCGACACACCCAGAAGAGTTGGCTAATAATGGGCAGGGGTTCTTTCCC
GCAAAAGCTAAGGGCTGTGTGGGTCTTCTCAACCGATTATTACCCGTCCCCAAGAAAGGG
R F R F P T H P E E L A N N G Q G F F P

1741 CTTGATATTCAAACGGCTTACGCGTTTCCCGCTTCGCTCAATGGTTCAGGGCTTACCATT
GAACTATAAGTTTGCCGAATGCGCAAAGGGCGAAGCGAGTTACCAAGTCCCGAATGGTAA
L D I Q T A Y A F P A S L N G S G L T I

1801 GGGCTTCTCGAATTTTCAAATGGCTTTAATCCCCAGGATGTGATGACGTTCTGGAACCAA
CCCGAAGAGCTTAAAAGTTTACCGAAATTAGGGGTCCTACACTACTGCAAGACCTTGGTT
G L L E F S N G F N P Q D V M T F W N Q

```
1861  TTTGGTATTGCTTGCCCAAATGTGAGCTTTGTATCGGTTGATGGAACCCCGAATGATCTG
      CCCGAAGAGCTTAAAAGTTTACCGAAATTAGGGGTCCTACACTACTGCAAGACCTTGGTT
       F  G  I  A  C  P  N  V  S  F  V  S  V  D  G  T  P  N  D  L

1921  GGGGTTAATGCCTACGATTTGGAGGCAACGTTGGATATTGAATGGGCAGGAGCTATGGCT
      CCCCAATTACGGATGCTAAACCTCCGTTGCAACCTATAACTTACCCGTCCTCGATACCGA
       G  V  N  A  Y  D  L  E  A  T  L  D  I  E  W  A  G  A  M  A

1981  CCTTTGGCAAACTTGGTGGTTTATGAAGCCAATGCAGGCTCGTCGGATACCTCTTTTGCT
      GGAAACCGTTTGAACCACCAAATACTTCGGTTACGTCCGAGCAGCCTATGGAGAAAACGA
       P  L  A  N  L  V  V  Y  E  A  N  A  G  S  S  D  T  S  F  A

2041  TTATCCGTTCTTAAGGCTTTGCAGTATGCTTATAACGACGTGCTTAATTGTCCCGATATT
      AATAGGCAAGAATTCCGAAACGTCATACGAATATTGCTGCACGAATTAACAGGGCTATAA
       L  S  V  L  K  A  L  Q  Y  A  Y  N  D  V  L  N  C  P  D  I

         ClaI
2101  TTATCGATCAGTTATGGGGATGGCGAAACACGGTTTCCCGTTTCTACTATGCAGGCATGG
      AATAGCTAGTCAATACCCCTACCGCTTTGTGCCAAAGGGCAAAGATGATACGTCCGTACC
      L  S  I  S  Y  G  D  G  E  T  R  F  P  V  S  T  M  Q  A  W

2161  GATATGGTGGCTCGCAATGCGGCTTTGATTGGTATGACGATCTTCGTCGCTTCTGGGGAT
      CTATACCACCGAGCGTTACGCCGAAACTAACCATACTGCTAGAAGCAGCGAAGACCCCTA
       D  M  V  A  R  N  A  A  L  I  G  M  T  I  F  V  A  S  G  D

2221  CAGGGAGCTTATGGTCTGCACGGACCCGGACGAAAAATTTGTCATGTCGATGCGCCGGCG
      GTCCCTCGAATACCAGACGTGCCTGGGCCTGCTTTTTAAACAGTACAGCTACGCGGCCGC
       Q  G  A  Y  G  L  H  G  P  G  R  K  I  C  H  V  D  A  P  A

2281  AACTCTCCGCATATGGTATCGGTAGGTGGCACGCATTTGCTCTTAAATAGCCAAGGGCAG
      TTGAGAGGCGTATACCATAGCCATCCACCGTGCGTAAACGAGAATTTATCGGTTCCCGTC
       N  S  P  H  M  V  S  V  G  G  T  H  L  L  L  N  S  Q  G  Q

                                    prKum-Asc
2341  ATAGTCGAAGAAACCGGCTGGACGGATGTGAATAATAACGGAGCATCGGGGGGAGGTATC
      TATCAGCTTCTTTGGCCGACCTGCCTACACTTATTATTGCCTCGTAGCCCCCCTCCATAG
       I  V  E  E  T  G  W  T  D  V  N  N  N  G  A  S  G  G  G  I

              prArsRlacfwd
2401  AGTCAGGTATTTGCGGTCCCGGCCTATCAAGAGGGCATCCCACTGCCTGTGAAAGCGGGT
      TCAGTCCATAAACGCCAGGGCCGGATAGTTCTCCCGTAGGGTGACGGACACTTTCGCCCA
       S  Q  V  F  A  V  P  A  Y  Q  E  G  I  P  L  P  V  K  A  G

2461  TATCATGTGGGACGCGGGGTGCCCGATGTGGCGTTGAATGCCGATCCGGATACGGGCTAC
      ATAGTACACCCTGCGCCCCACGGGCTACACCGCAACTTACGGCTAGGCCTATGCCCGATG
       Y  H  V  G  R  G  V  P  D  V  A  L  N  A  D  P  D  T  G  Y

2521  GCGGTGTTTTTTCAAGGGATGTGGACCGTAGTGGGAGGAACATCGGTGGCTTCTCCCATT
      CGCCACAAAAAAGTTCCCTACACCTGGCATCACCCTCCTTGTAGCCACCGAAGAGGGTAA
       A  V  F  F  Q  G  M  W  T  V  V  G  G  T  S  V  A  S  P  I
```

```
2581  TGGGCGGCTATTACGGCGCTGATTAATCAAAATCGGGTCCAGGCAGGGAAATCTCTTATA
      ACCCGCCGATAATGCCGCGACTAATTAGTTTTAGCCCAGGTCCGTCCCTTTAGAGAATAT
       W  A  A  I  T  A  L  I  N  Q  N  R  V  Q  A  G  K  S  L  I

                                        ClaI
2641  GGGTATATGAATCCCAGACTTTATGGGATCGATCATGCCTCCACTTTTCACGATATTACC
      CCCATATACTTAGGGTCTGAAATACCCTAGCTAGTACGGAGGTGAAAAGTGCTATAATGG
       G  Y  M  N  P  R  L  Y  G  I  D  H  A  S  T  F  H  D  I  T

                                                prRTKum-As
2701  GTGGGGAATAATTCTTATGACGGGGTAACAGGTTATGATTGCACGCCGGGATGGGATGCT
      CACCCCTTATTAAGAATACTGCCCCATTGTCCAATACTAACGTGCGGCCCTACCCTACGA
       V  G  N  N  S  Y  D  G  V  T  G  Y  D  C  T  P  G  W  D  A


2761  GTGACGGGATGGGGAAGTCCTATCGTGGTGTCTTTGGTAAAGAGCTTGTCTTAACAGGGC
      CACTGCCCTACCCCTTCAGGATAGCACCACAGAAACCATTTCTCGAACAGAATTGTCCCG
       V  T  G  W  G  S  P  I  V  V  S  L  V  K  S  L  S

      HindIII
2821  AAGCTTTTTTCATGCGGCAAAACATAATAATATGATTATGAATATATCAATGAGAGGAGC
      TTCGAAAAAAGTACGCCGTTTTGTATTATTATACTAATACTTATATAGTTACTCTCCTCG
            arsR start
                  prArsRlacrev                              prArsRc
2881  GGGACCAATGACCACCGAAGAGGTCAGTATCTTGTTTAAAGCGTTGAGTGATCAAACAAG
      CCCTGGTTACTGGTGGCTTCTCCAGTCATAGAACAAATTTCGCAACTCACTAGTTTGTTC
              M  T  T  E  E  V  S  I  L  F  K  A  L  S  D  Q  T  R


2941  GCTACGCATGGTCACCTTACTTTCCCGCCGGGAGTATTGCAACTGCGAATTTGTGTCTAT
      CGATGCGTACCAGTGGAATGAAAGGGCGGCCCTCATAACGTTGACGCTTAAACACAGATA
        L  R  M  V  T  L  L  S  R  R  E  Y  C  N  C  E  F  V  S  I


3001  TTTCGGGATTTCCCAACCGGCTATCTCTCGCCATATTGCCCGTCTCAAAGAGGCCCGTTT
      AAAGCCCTAAAGGGTTGGCCGATAGAGAGCGGTATAACGGGCAGAGTTTCTCCGGGCAAA
        F  G  I  S  Q  P  A  I  S  R  H  I  A  R  L  K  E  A  R  L


3061  AATCCATGAAAGACGGCCGGGACAATGGATTTATTATTCCTTGAATCCGTCAACGTGGGA
      TTAGGTACTTTCTGCCGGCCCTGTTACCTAAATAATAAGGAACTTAGGCAGTTGCACCCT
        I  H  E  R  R  P  G  Q  W  I  Y  Y  S  L  N  P  S  T  W  D

                                                      prRTArsR
3121  TCGTCTGCCATCCTTTTCGCAAATTCTGGAAGATTTGGGACAACATGATGCGGTCATTCT
      AGCAGACGGTAGGAAAAGCGTTTAAGACCTTCTAAACCCTGTTGTACTACGCCAGTAAGA
        R  L  P  S  F  S  Q  I  L  E  D  L  G  Q  H  D  A  V  I  L
                                                   arsB start
3181  TAGAACTGTAGACCAAGCACCTGGTTGTCCTGTGCCTGACCATGAACGACATTAATGGGA
      ATCTTGACATCTGGTTCGTGGACCAACAGGACACGGACTGGTACTTGCTGTAATTACCCT
        R  T  V  D  Q  A  P  G  C  P  V  P  D  H  E  R  H     M  G

                  prArsBc
3241  TTGCAAGATAAGAACAAAGGGGATTGTGTGTTGCTGGCTAATATCTTATTTATCATTGTT
      AACGTTCTATTCTTGTTTCCCCTAACACACAACGACCGATTATAGAATAAATAGTAACAA
        L  Q  D  K  N  K  G  D  C  V  L  L  A  N  I  L  F  I  I  V
```

```
3301  CTGGTTTTAGGTATTGTCCAGCCAAGAGGGCTGTCCATCGGCTATAGTGCGGCGGGTGGC
      GACCAAAATCCATAACAGGTCGGTTCTCCCGACAGGTAGCCGATATCACGCCGCCCACCG
       L  V  L  G  I  V  Q  P  R  G  L  S  I  G  Y  S  A  A  G  G

3361  GCGATTATCGCGTTATTACTCGGACTCGTGTCCATAGCTAATGTGGTTGAAGTCGTGGAT
      CGCTAATAGCGCAATAATGAGCCTGAGCACAGGTATCGATTACACCAACTTCAGCACCTA
       A  I  I  A  L  L  L  G  L  V  S  I  A  N  V  V  E  V  V  D

                                                    prRTArsB
3421  ATTGTCTGGGATGCGACCTTAACGTTCGTGGCGCTTATTCTCATTTCGATGGTACTGGAC
      TAACAGACCCTACGCTGGAATTGCAAGCACCGCGAATAAGAGTAAAGCTACCATGACCTG
       I  V  W  D  A  T  L  T  F  V  A  L  I  L  I  S  M  V  L  D

3481  CGCATTGGTTTTTTCGAATGGGCAGCGCTTAGCATGAGCCACCTCGCCAAAGACAGTGGA
      GCGTAACCAAAAAAGCTTACCCGTCGCGAATCGTACTCGGTGGAGCGGTTTCTGTCACCT
       R  I  G  F  F  E  W  A  A  L  S  M  S  H  L  A  K  D  S  G

3541  TACAAGCTATTTATATGGATAGGATTATTGGGGGCAGTGGTGGCCATGTTCTTTGCCAAT
      ATGTTCGATAAATATACCTATCCTAATAACCCCCGTCACCACCGGTACAAGAAACGGTTA
       Y  K  L  F  I  W  I  G  L  L  G  A  V  V  A  M  F  F  A  N

3601  GATGGGGCCGCCTTAATTCTGACACCGCTAGTCTATGAGCAAACGCGCATTTTAAAGCTG
      CTACCCCGGCGGAATTAAGACTGTGGCGATCAGATACTCGTTTGCGCGTAAAATTTCGAC
        D  G  A  A  L  I  L  T  P  L  V  Y  E  Q  T  R  I  L  K  L

                                                          prKaraArsBfwd
3661   TCTCCGAAAGCGACGTTGGCGTTCATTATGACCAGCGGTTTTATTGCCGATACAACCTCT
       AGAGGCTTTCGCTGCAACCGCAAGTAATACTGGTCGCCAAAATAACGGCTATGTTGGAGA
        S  P  K  A  T  L  A  F  I  M  T  S  G  F  I  A  D  T  T  S

3721   GTCCCGCTGGTGATTTCCAATTTGGTTAATATTGTGTCCGCCGACTATTTCCATCTCGGT
       CAGGGCGACCACTAAAGGTTAAACCAATTATAACACAGGCGGCTGATAAAGGTAGAGCCA
        V  P  L  V  I  S  N  L  V  N  I  V  S  A  D  Y  F  H  L  G

3781   TTCTTATCCTACCTGGTGCACATGGTTCCCGTCGATCTTGTGGCCTTAAGCGCGAGCTTA
       AAGAATAGGATGGACCACGTGTACCAAGGGCAGCTAGAACACCGGAATTCGCGCTCGAAT
        F  L  S  Y  L  V  H  M  V  P  V  D  L  V  A  L  S  A  S  L

3841   GTGGCATTATTCATTTATTACCACAAGCATATCCCCGTACACATTGACACCAGTGCTTTA
       CACCGTAATAAGTAAATAATGGTGTTCGTATAGGGGCATGTGTAACTGTGGTCACGAAAT
        V  A  L  F  I  Y  Y  H  K  H  I  P  V  H  I  D  T  S  A  L

         BamHI
3901   CCGGATCCCAAAAGTCGGATAAAAGACGTCCGGGTATTTCGCCTGGCTGGATATTCTTCC
       GGCCTAGGGTTTTCAGCCTATTTTCTGCAGGCCCATAAAGCGGACCGACCTATAAGAAGG
        P  D  P  K  S  R  I  K  D  V  R  V  F  R  L  A  G  Y  S  S

3961   GCTGTACTATCCGGATATTTTCTTAGCCAGTTTCTTCATTGGCCCGTCTCAATTTTTACG
       CGACATGATAGGCCTATAAAAGAATCGGTCAAAGAAGTAACCGGGCAGAGTTAAAAATGC
        A  V  L  S  G  Y  F  L  S  Q  F  L  H  W  P  V  S  I  F  T
```

*Hin*dIII

4021 ATGGCGGCAGTGGCCGTGTTATTAGCAGGAGCTGCCCA**AAGTCC**TTATATCTCATCGCTC
TACCGCCGTCACCGGCACAATAATCGTCCTCGACGGGT**TTCAGG**AATATAGAGTAGCGAG
M A A V A V L L A G A A Q S P Y I S S L

4081 AAGCTTATTAAAGAGGCTCCCTGGAAAATTGTGGTCTTTTCAATTGGCATGTATATTGTC
TTCGAATAATTTCTCCGAGGGACCTTTTAACACCAGAAAAGTTAACCGTACATATAACAG
K L I K E A P W K I V V F S I G M Y I V

4141 GTTTTTGGGTTGAGAAACCAAGGGCTAACTGCCATTTTGGGCCACCAATTTCACCAGTTG
CAAAAACCCAACTCTTTGGTTCCCGATTGACGGTAAAACCCGGTGGTTAAAGTGGTCAAC
V F G L R N Q G L T A I L G H Q F H Q L

prKaraArsBrev

4201 AGTGTAGGGGGCACTTTAACAGCCGCCGTGGGCACAGGGTTTATTGCTGCACTTCTTTCA
TCACATCCCCCGTGAAATTGTCGGCGGCACCCGTGTCCCAAATAACGACGTGAAGAAAGT
S V G G T L T A A V G T G F I A A L L S

4261 TCGGTCATGAATAATATGCCGACTGTTCTAATTGATGCCTTAGCCATTCATCATGCCCAT
AGCCAGTACTTATTATACGGCTGACAAGATTAACTACGGAATCGGTAAGTAGTACGGGTA
S V M N N M P T V L I D A L A I H H A H

4321 CTTTCAGCGGTCATGACGCATTTAATGGCCTATGCTAATGTTATTGGTTCGGATTTGGGC
GAAAGTCGCCAGTACTGCGTAAATTACCGGATACGATTACAATAACCAAGCCTAAACCCG
L S A V M T H L M A Y A N V I G S D L G

4381 CCTAAACTCACACCGATTGGCTCTTTGGCTACACTGTTATGGCTTCACGTCTTAGAGCAG
GGATTTGAGTGTGGCTAACCGAGAAACCGATGTGACAATACCGAAGTGCAGAATCTCGTC
P K L T P I G S L A T L L W L H V L E Q

4441 CGCCAGGTGCGGATTAGCTATCTTTATTACATGAAAGTAGGATTTTTACTCACGATTCCG
GCGGTCCACGCCTAATCGATAGAAATAATGTACTTTCATCCTAAAAATGAGTGCTAAGGC
R Q V R I S Y L Y Y M K V G F L L T I P

4501 GTGCTGTTGACGACTTTAGTCGCTTTAGCGATTTGGACAGGTGTGATTGGAGCTTAATAT
CACGACAACTGCTGAAATCAGCGAAATCGCTAAACCTGTCCACACTAACCTCGAATTATA
V L L T T L V A L A I W T G V I G A

*Sac*I *Sac*I

4561 TGTCTTTTCGAGGATGTTGGACAACAAAA**GAGCTC**TTGGATCATATAGGCCCAA**GAGCTC**
ACAGAAAAGCTCCTACAACCTGTTGTTTT**CTCGAG**AACCTAGTATATCCGGGTT**CTCGAG**

4621 AGTAAAGGAAGCTGTCTTTGATAGGGATTATTGTCCTACCTGTTCTTTGAATAGGCAAAA
TCATTTCCTTCGACAGAAACTATCCCTAATAACAGGATGGACAAGAAACTTATCCGTTTT

*Xho*I

4681 GTTTT**CTCGAG**TCCCTCTTCTAAAGAGACTTTAGGAGACCAACCCAAATATTTTTTGGCT
CAAAA**GAGCTC**AGGGAGAAGATTTCTCTGAAATCCTCTGGTTGGGTTTATAAAAAACCGA

```
4741    TTATCAATATTCGGGCAGCGATTGGTGGGATCATCTGCTGGCAAGGGGCGAAATTCCGTG
        AATAGTTATAAGCCCGTCGCTAACCACCCTAGTAGACGACCGTTCCCCGCTTTAAGGCAC

4801    TAAAGAGGGACCCCCGCTACTTTGGCGACGGCTTGAGCAAAGTTTAAGATGGTTCGCTCT
        ATTTCTCCCTGGGGGCGATGAAACCGCTGCCGAACTCGTTTCAAATTCTACCAAGCGAGA

4861    TCCGGATTACCAATATTGAATACTTCACCAGACAAGCCGTCCCGAATAATGGCTTGATAA
        AGGCCTAATGGTTATAACTTATGAAGTGGTCTGTTCGGCAGGGCTTATTACCGAACTATT

4921    ATCCCTTCCACTTCGTCAGAAACATAGCAGAAGCTGCGTGTTTGTTGTCCATCGCCATAG
        TAGGGAAGGTGAAGCAGTCTTTGTATCGTCTTCGACGCACAAACAACAGGTAGCGGTATC

4981    ACTGTCAAGGGCTTGCCTTCTAAGGCTTGGCGAACGAAATTGGGTACGACGCGACCATCT
        TGACAGTTCCCGAACGGAAGATTCCGAACCGCTTGCTTTAACCCATGCTGCGCTGGTAGA

              SphI
5041    TCTGATTGCATGCGTGGTCCATAACAGTTAAAAAATCGTAAAATCCGCAAGTCAAGGCCA
        AGACTAACGTACGCACCAGGTATTGTCAATTTTTTAGCATTTTAGGCGTTCAGTTCCGGT

5101    AATTGGCGATGATATTCCATAGCGATGGCTTCCCCATAGCGCTTACCTTCATCATAACAA
        TTAACCGCTACTATAAGGTATCGCTACCGAAGGGGTATCGCGAATGGAAGTAGTATTGTT

5161    GCCCGTTCACCATTGGGATTCACATGGCCCCAGTAGGTTTCACTTTGAGGGGAGACTTGG
        CGGGCAAGTGGTAACCCTAAGTGTACCGGGGTCATCCAAAGTGAAACTCCCCTCTGAACC

5221    GGATCACCATACACCTCCGAAGTCGATCCCAGAATGAAACGGGCTTGAAATTTTCGGGCG
        CCTAGTGGTATGTGGAGGCTTCAGCTAGGGTCTTACTTTGCCCGAACTTTAAAAGCCCGC

5281    AGTTTCAACGCGTTTTCTGTACCAATGGAATTGACACGTAATGTTTCTAACGCAAGGCGG
        TCAAAGTTGCGCAAAAGACATGGTTACCTTAACTGTGCATTACAAAGATTGCGTTCCGCC

                          PstI
5341    CGGTAATGAATGGGAGATGCTGCAG
        GCCATTACTTACCCTCTACGACGTC
```

REFERENCES

Altshul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein data base search programs. *Nucleic Acids Res*. **35:**3308-3402.

Anderson, G.L., J. Williams and R. Hille. (1992). The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum-containing hydroxylase. *J. Biol. Chem.* **267:**23674-23682.

Atkinson, T., S. Cairns, D.A. Cowan, M.J. Danson, D.W. Hough, D.B. Johnson, P.R. Norris, N. Raven, R. Robson, C. Robinson, R.J. Sharp. (2000). A microbiological survey of Montserrat island hydrothermal biotopes. *Extremophiles* **4:**305-313.

Azcue, J.M. and J.O. Nriagu. (1994). Arsenic, historical perspectives. *In* Nriagu, J.O. (Eds.) *Arsenic in the environment*. Wiley & Sons, New York. NY, pp. 1-16.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol*. **10:**411-421.

Baker, B.J. and J.F. Banfield. (2003). Microbial communities in acid mine drainage. *FEMS Microbiol. Ecol.* **44:**139-152.

Barrett, J., D.K. Ewart, M.N. Hughes and R.K. Poole. (1993). Chemical and biological pathways in the bacterial oxidation of arsenopyrite. *FEMS Microbiol. Rev*. **11:**57-62.

Barton, B.M., G.P. Harding and A.J. Zuccarelli. (1995). A general method for detecting and sizing large plasmids. *Anal. Biochem.* **226:**235-240.

Bennet, M.S., Z. Guan, M. Laurberg and X. Su. (2001). *Bacillus subtilis* arsenate reductase is structurally and functionally similar to low molecular weight protein tyrosine phosphatases. *PNAS.* **98:**13577-13582.

Bentley, R. and T.G. Chasteen. (2002). Microbial Methylation of Metalloids: Arsenic, Antimony, and Bismuth. *Microbiol. Mol. Biol. Rev*. **66:**250-271.

Bhattacharjee, H., J. Li, M.Y. Ksenzenko and B.P. Rosen. (1995). Role of Cysteinyl Residues in Metalloactivation of the Oxyanion-translocating ArsA ATPase. *J. Biol. Chem.* **270:**11245-11250.

Bhattacharjee, H., T. Zhou, J. Li, D.L. Gatti, A.R. Walmsley and B.P. Rosen. (2000). Structure-function relationships in an anion-translocating ATPase. *Biochem. Soc. Trans*. **28:**520-526.

Bogdanova, T.I., I.A. Tsaplina, T.F. Kondrat'eva, V.I. Duda, N.E. Suzina, V.S. Melamud, T.P. Tourova and G.I. Karavaiko. (2006). Sulfobacillus thermotolerans sp. nov., a thermotolurant, chemolithotrophic bacterium. *Int. J. Syst. Evol. Microbiol.* **56:**1039-1042.

Breed, A.W., A. Glatz, G.S. Hansford and S.T.L. Harrison. (1996). The effect of As(III) and As(V) on the batch bioleaching of a pyrite-arsenopyrite concentrate. *Miner. Eng*. **9:**1235-1252.

Bridge, T.A.M. and D.B. Johnson. (1998). Reduction of Soluble Iron and Reductive Dissolution of Ferric Iron-Containing Minerals by Moderately Thermophilic Iron-Oxidizing Bacteria. *Appl. Environ. Microbiol.* **64:**2181-2186.

Brierly, C.L. and J.A. Brierly. (1997). Microbiology for the Metal Mining Industry. *In* Hurst, C.J., G.R. Knudsen, M.J. McInerney, L.D. Stetzenbach and M.V. Walker (Eds.), *Manual of Environmental Microbiology*. AMS Press, Washington D.C. pp. 830-841.

Brierley, C.L. (1997). Mining biotechnology: research into commercial development and beyond. *In* Rawlings D.E. (Eds.), *Biomining: theory, microbes and industrial processes*. Springer-Verlag, Berlin, Germany. pp. 3-17.

Brierley, J.A., P.R. Norris, D.P. Kelly and N.W. LeRoux. (1978). Characteristics of a Moderately Thermophilic and Acidophilic Iron-Oxidizing *Thiobacillus*. *Eur. J. Appl. Microbiol. Biotechnol.* **5:**291-299.

Briggs, A.P. and M. Millard. (1997). Cobalt recovery using bacterial leaching at the Kasese project, Uganda. *In International Biohydrometallurgy Symposium*, IBS97 pp. M2.4.1-2.4.12. South Aust.: Aust. Mineral Found.

Broër, S., G. Ji, A. Broër and S. Silver. (1993). Arsenic Efflux Governed by the Arsenic Resistance Determinant of *Staphylococcus aureus* Plasmid pI258. *J. Bacteriol.* **175:**3480-3485.

Bruhn D.F., J. Li, S. Silver, F. Roberto and B.P. Rosen. (1996). The arsenical resistance operon of IncN plasmid R46. *FEMS Microbiol. Lett.* **139:**149-153.

Bruins, M.R., S. Kapil and F.W. Oehme. (2000). Microbial Resistance to Metals in the Environment. *Ecotoxicol. Environ. Safety*. **45:**198-207.

Busenlehner, L.S., M.A. Pennella and D.P. Giedroc. (2003). The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic met resistance. *FEMS Microbiol. Rev*. **27:**131-143.

Butcher, B.G. and Rawlings, D.E. (2002). The divergent chromosomal operon af *Acidithiobacillus ferrooxidans* is regulated by an atypical ArsR protein. *Microbiology* **148:**3983-3992.

Butcher, B.G., S.M. Deane and D.E. Rawlings. (2000). The chromosomal arsenic resistance genes of *Thiobacillus ferrooxidans* have an unusual arrangement and confer increased arsenic and antimony resistance to *Escherichia coli. Appl. Envir. Microbiol.* **66:**1826-1833.

Cai, J., K. Salmon and M. DuBow. (1998). A chromosomal *ars* operon homologue of *Pseudomonas aeroginosa* confers increased resistance to arsenic and antimony in *Escherichia coli. Microbiology* **144:**2705-2713.

Carepo, M.S.P., J.S.N. de Azevedo, J.I.R. Porto, A.R. Bentes-Sousa, J. da Silva Bastista, A.L.C. da Silva and M.P.C. Schneider. (2004). Identification of *Chromobacterium violaceum* genes with potential biotechnological application in environmental detoxification. *Genet. Mol. Res.* **1:**181-194.

Carlin, A., W. Shi, S. Dey and B.P. Rosen. (1995). The *ars*-operon of *Escherichia coli* confers arsenical and antimonal resistance. *J. Bacteriol.* **177:**981-986.

Casadaban, M.J., A. Martinez-Arias, S.K. Shapira and J. Chou. (1983). β-Galactosidase gene fusions for analyzing gene expression in *Eschrichia coli* and yeast. *Methods Enzymol*. **100:**293-308.

Cervantes, C., G. Ji, J.L. Ramirez and S. Silver. (1994). Resistance to arsenic compounds in microorganisms. *FEMS Microbiol. Rev.* **15:**355-367.

Chen C-M., T.K. Misra, S. Silver and P.B. Rosen. (1985). Separate resistances to arsenate and arsenite (antimonite) encoded by the arsenical resistance operon of R factor R773. *J. Bacteriol.* **161:**758-763.

Chen C-M., T.K. Misra, S. Silver and P.B. Rosen. (1986). Nucleotide sequence of the structural genes for an anion pump: the plasmid-encoded arsenical resistance operon. *J. Biol. Chem*. **261:**15030-15038.

Chen, Y., S. Dey and B.P. Rosen. (1996). Soft Metal Thiol Chemistry Is Not Involved in the Transport of Arsenite by the Ars Pump. *J. Bacteriol.* **178:**911-913.

Chen, Y and B.P. Rosen. (1997). Metalloregulatory Properties of the ArsD Repressor. *J. Biol. Chem.* **272:**14257-14262.

Clark, D.A. and P.R. Norris. (1996). *Acidimicrobium ferrooxidans* gen nov, sp nov: mixed culture ferrous oxidation with *Sulfobacillus* species. *Microbiology*. **142:**785-790.

Cook, W.J., S.R. Kar, K.B. Taylor and L.M. Hall. (1998). Crystal Structure of the Cyanobacterial Metallothionein Repressor SmtB: A Model for Metalloregulatory Proteins. *J. Mol. Biol.* **275:**337-346.

Coram, N.J. and D.E. Rawlings. (2002). Molecular relationship between two groups of *Leptospirillum* and the finding that *Leptospirillum ferriphilum* sp. nov. dominates South African commercial biooxidation tanks which operate at 40°C. *Appl. Environ. Microbiol.* **68:**838-845.

De Groot, P., S.M. Deane and D.E. Rawlings. (2003). A transposon-located arsenic resistance mechanism from a strain of *Acidithiobacillus caldus* isolated from commercial, arsenopyrate biooxidation tanks. *Hydrometallurgy* **71:**115-123.

Demel, S., J. Shi, P. Martin, B.P.Rosen and B.F.P. Edwards. (2004). Arginine 60 in the ArsC arsenate reductase of *E. coli* plasmid R773 determines the chemical nature of the bound As(III) product. *Prot. Sci.***13:**2330-2340.

Dew, D.W. and D.M. Miller. (1997). The BioNIC process, bioleaching of mineral sulphide concentrates for the recovery of nickel. *In International Biohydrometallurgy Symposium*, IBS97 pp. M7.1.1-7.1.9. South Aust.: Aust. Mineral Found.

Dey, S., D. Dou and B.P. Rosen. (1994). ATP-dependent Arsenite Transport in Everted Membrane Vesicles of *Escherichia coli. J. Biol. Chem*. **269:**25442-25446.

Dey, S. and B.P. Rosen. (1995). Dual Mode of Energy Coupling by the Oxyanion-Translocating ArsB Protein. *J. Bacteriol.* **177:**385-389.

Dopson, M., and E.B. Lindström. (2003). Analysis of Community Composition during Moderately Thermophilic Bioleaching of Pyrite, Arsenical Pyrite and Chalcopyrite. *Microbial Ecol*. **48:**19-28.

Dopson, M., C. Baker-Austin, P.R. Koppineedi and P.L. Bond. (2003). Growth in sulfidic mineral environments: metal resistance mechanisms in acidophilic micro-organisms. *Microbiology*. **149:**1959-1970.

Dou, D., J.B. Owolabi, S. Dey and B.P. Rosen. (1992). Construction of a Chemeric ArsA-ArsB Protein for the Overexpression of the Oxyanion-translocating ATPase. *J. Biol. Chem.* **267:**25768-25775.

Driessen, A.J.M., J.L.C.M. van de Vossenberg and W.N. Konings. (1996). Membrane composition and ion-permeability in extremophiles. *FEMS Microbiol. Rev.* **18:**139-148.

Duda, V.I., N.E. Suzina, L.O. Severina, V.V. Dmitriev and G.I. Karavaiko. (2001). Formation of Flat Lamellar Intermembrane Lipid Structures in Microorganisms. *J. Membrane Biol*. **180:**33-48.

Edwards, K.L., P.L. Bond, T.M. Gihring and J.F. Banfield. (2000). An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science.* **279:**1796-1799.

Ehrlich, H.L. (1997). Microbes and metals. *Appl. Microbiol. Biotechnol.* **48:**687-692.

Eicken, C., M.A. Penella, X. Chen, K.M. Koshlap, M.L. VanZile, J.C.Sacchettini and D.P. Geidroc. (2003). A Metal-Ligand-mediated Intersubunit Allosteric Switch in Related SmtB/ArsR Zinc Sensor Proteins. *J. Mol. Biol.* **333:**683-695.

Fuchs, T., H. Huber, K. Teiner, S. Burggraf and K.O. Stetter. (1995). *Metallosphaera prunae*, sp. nov., a novel metal-mobilizing, thermoacidophilic archaeum, isolated from a uranium mine in Germany. *Syst. Appl. Microbiol.* **18:**560-566.

Ghauri, M.A. and D. B. Johnson. (1991). Physiological diversity amongst some moderately thermophilic iron-oxidizing bacteria. *FEMS Microbiol. Ecol.* **85:**327-334.

Gihring, T.M., P.L. Bond, S.C.Peters and J.F. Banfield. (2003). Arsenic resistance in the archaeon "*Ferroplasma acidarmanus*": new insights into the structure and evolution of the *ars* genes. *Extremophiles*. **7:**123-130.

Gladysheva, T.B., K.L. Oden and B.P. Rosen. (1994). Properties of the arsenate reductase of plasmid R773. *Biochem.* **33:**7288-7293.

Goebel, B.M. and E. Stackebrandt. (1994). Cultural and phylogenetic analysis of mixed microbial populations found in natural and commercial bioleaching environments. *Appl. Environ. Micobial.* **60:**1614-1621.

Golovacheva, R.S. and G.I. Karavaiko. (1978). A new genus of thermophilic spore-forming bacteria, *Sulfobacillus*. *Microbiology* **47:**815-822.

Golyshina, O.V., T.A. Pivovarova, G.I. Karavaiko, T.F. Kondrateva, W.R. Moore, W.R. Abraham, *et al*. (2000). *Ferroplasma acidiphilum* gen nov, sp nov, an acidophilic, autotrophic, ferrous-iron-oxidizing, cell-wall-lacking, mesophilic member of the *Ferroplasmaceae* fam nov, comprising a distinct lineage of the *Archaea*. *Int. J. Syst. Evol. Microbiol.* **50:**997-1006.

Hallberg, K.B. and D.B. Johnson. (2001). Biodiversity of acidophilic microorganisms. *Adv. Appl. Microbiol.* **49:**37-84.

Hotter, G.S., T. Wilson and D.M. Collins. (2001). Identification of a cadmium-induced gene in *Mycobacterium bovis* and *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* **200:**151-155.

Inskeep, W.P., T.R. McDermott and S. Fendorf. (2002). Arsenic (V)/(III) Cycling in Soils and Natural Waters: Chemical and Microbiological processes. *In* Frankenberger, W.T. Jr. (Eds.), *Environmental Chemistry of Arsenic*. Marcel Dekker, New York.

Jackson, C.R., E.F. Jackson, S.L. Dugas, K. Gamble and S.E. Williams. (2003). Microbial transformations of arsenite and arsenate in natural environments. *Recent Res. Devel. Microbiol.* **7:**103-118.

Jana, S. and J.K. Deb. (2005). Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Appl. Microbiol. Biotechnol*. **67:**289-298.

Ji, G. and S. Silver. (1992). Reduction of arsenate to arsenite by the ArsC protein of the arsenic resistance operon of *Staphylococcus aureus* plasmid pI258. *Proc. Natl. Acad. Sci.,* USA. **89:**9474-9478.

Ji, G. and S. Silver. (1992b). Regulation and expression of the arsenic resistance operon from *Staphylococcus aureus* plasmid pI258. *J. Bacteriol.* **174:**3684-3694.

Ji, G., E.A. Gerber, L.G. Armes, C-M. Chen, J.A. Fuchs and S. Silver. (1994). Arsenate reductase of *Stapphylococcus aureus* plasmid pI258. *Biochem.* **33:**7294-7299.

Johnson, D.B. (1998). Biodiversity and ecology of acidophilic microorganisms. *FEMS Microbiol. Ecol.* **27:**307-317.

Johnson, D.B., P. Bacelar-Nicolau, N. Okibe, A. Yahya and K.B. Hallberg. (2001). Role of pure and mixed cultures of Gram-positive eubacteria in mineral leaching. *In* Ciminelli, V.S.T., Garcia Jr., O. (Eds.), *Biohydrometallurgy: Fundamentals, Technology and Sustainable Development*, vol. **11A**. Elsevier, Amsterdam, pp. 461-470.

Johnson, D.B., N. Okibe and K.B. Hallberg. (2005). Differentiation and identification of iron-oxidizing acidophilic bacteria using cultivation techniques and amplified ribosomal DNA restriction enzyme analysis. *J. Microbiol. Methods* **60:**299-313

Johnson, D.B. and F.F. Roberto. (1997). Heterotrophic acidophiles and their roles in the bioleaching of sulfide minerals. *In* Rawlings D.E. (Eds.), *Biomining: theory, microbes and industrial processes*. Springer-Verlag, Berlin, Germany. pp. 259-279.

Kaneda, T. (1991). Iso- and anteiso-fatty acids in bacteria: biosynthesis, function and taxonomic significance. *Microbiol. Rev.* **55:**288-302.

Karkaria, C.E., C.M. Chen and B.P. Rosen. (1990). Mutagenesis of a nucleotide-binding site of an anion-translocating ATPase. *J. Biol. Chem*. **265:**7832-7836.

Kaur, P. and B.P. Rosen. (1993). Complementation between nucleotide-binding domains in an anion-translocating ATPase. *J. Bacteriol*. **175:**351-357.

Kinnunen, P.H.M., W.J. Robertson, J.J. Plumb, J.A.E. Gibson, P.D. Nichols, P.D. Fransmann and J.A. Puhakka. (2003). The isolation and use of iron-oxidizing, moderately thermophilic acidophiles from the Collie coal mine for the generation of ferric iron leaching solution. *Appl. Microbiol. Biotechnol.* **60:**748-753.

Knowles, F.C. and A.A. Benson. (1983). The biochemistry of arsenic. *Trends Biochem. Sci.* **8:**178-180.

Konings, W.N., S. Alberts, S. Koning and A.J.M. Driessen. (2002). The cell membrane plays a crucial role in survival of bacteria and archaea in extreme environments. *Antonie van Leeuwenhoek*. **81:**61-72.

Kotzé, A.A., M. Tuffin, S.M. Deane and D.E. Rawlings. (2006). Cloning and characterization of the chromosomal arsenic resistance genes from *Acidothiobacillus caldus* and enhanced arsenic resistance on conjugal transfer of the *ars* genes located on transposon Tn*AtcArs*. *Microbiology*. **152:**3551-3560.

Krebs, W., C. Brombacher, P.P. Bosshard, R. Bachofen and H. Brandl. (1997). Microbial recovery of metals from solids. *FEMS Microbiol. Rev.* **20:**605-617.

Krieg, N.R. (1988). Bacterial classification: an overview. *Can. J. Microbiol.* **34:**536-540.

Kuroda, M., S. Dey, O.I. Sanders and B.P.Rosen. (1997). Alternate Energy Coupling of ArsB, the Membrane Subunit of the Ars Anion-translocating ATPase. *J. Biol. Chem.* **272:**326-331.

Lane, D.J. (1991). 16S/23S rRNA sequencing. *In* Stackebrandt, E. and Goodfellow, M. (Eds.), *Nucleic acid techniques in bacterial systematics*. John Wiley & Sons, Chichester, United Kingdom. pp. 115-175.

Li, J. and B.P. Rosen. (2000). The linker peptide of the ArsA ATPase. *Mol. Microbiol.* **35:**361-367.

Li, S., Y. Chen and B.P. Rosen. (2001). Role of vicinical cysteine pairs in metalloid sencing by the ArsD As(III)-responsive repressor. *Mol. Microbiol.* **41:**687-696.

Li, S., B.P. Rosen, M.I. Boreges-Walmsley and A.R. Walmsley. (2002). Evidence for Cooperativity between the Four Binding Sites of Dimeric ArsD, an As(III)-responsive Transcriptional Regulator. *J. Biol. Chem.* **277:**25992-26002.

Lin, Y., A.R. Walmsley and B.P.Rosen. (2006). An arsenic metallochaperone for an arsenic detoxification pump. *PNAS*. **103:**15617-15622.

Liu, J., T.B. Gladysheva, L. Lee and B.P. Rosen. (1995). Identification of an essential cysteinyl residue in the ArsC arsenate reductase of plasmid R773. *Biochem.* **34:**13472-13476.

López-Maury, L., F.J. Florencio and J.C. Reyes. (2003). Arsenic Sensing and Resistance System in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *J. Bacteriol.* **185:**5363-5371.

Mandl, M., P. Matulová and H. Dočekalova. (1992). Migration of arsenic (III) during bacterial oxidation of arsenopyrite in chalcopyrite concentrate by *Thiobacillus ferrooxidans. Appl. Microbiol. Biotechnol*. **38:**429-431.

Marsh, R.M. and P.R. Norris. (1983). The isolation of some thermophilic, autotrophic iron- and sulfur-oxidizing bacteria. *FEMS Microbiol. Lett.* **17:**311-315.

Martin, P., S. Demal, J. Shi, T. Gladysheva, D.L. Gatti, B.P. Rosen and B.F.P. Edwards. (2001). Insights into the structure, salvation and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme. *Structure.* **9:**1071-1081.

Melamud, V.S., T.A. Pivovarova, T.V. Kolganova, G.A. Osipov, A.M. Lysenko, T.F. Kondrat'eva and G.I. Karavaiko. (2003). *Sulfobacillus sibiricus* sp nov., a New Moderately Thermophilic Bacterium. *Microbiology*. **72:**681-688.

Meng, Y., Z. Lui and B.P. Rosen. (2004). As(III) and Sb(III) Uptake by GlpF and Efflux by ArsB in *Escherichia coli. J. Biol. Chem.* **279:**18334-18341.

Messens, J., G. Hayburn, A. Desmyter, G. Laus and L.Wyns. (1999). The essentialcatalytic redox couple in arsenate reductase from *Staphylococcus aureus*. *Biochem*. **38:**16857-16865.

Messens, J., J.C. Martins, K. van. Belle, E. Brosens, A. Desmyter, M. De Gieter, J. Wieruszeski, R. Willem and I. Zegers. (2002). All intermediates of the arsenate reductase mechanism, including an intramolecular dynamic disulfide cascade. *PNAS.* **99:**8506-8511.

Mikkelsen, D., U. Kappler, A.G. McEwan and L.I. Sly. (2006). Archaeal diversity in two thermophilic chalcopyrite bioleaching reactors. *Environ. Microbiol.* **8:**2050-2055.

Miller, J.H. (1972). *Experiments in Molecular genetics*. Cold Spring Harbor Laboratory. New York: Cold Spring Harbor Press.

Moore, C.M., A. Gaballa, M. Hui, R.W. Ye and J.D. Helmann. (2005). Genetic and physiological responses of *Bacillus subtilis* to metal ion stress. *Mol. Microbiol*. **57:**27-40.

Moore, C.M. and J.D. Helmann. (2005). Metal ion homeostasis in *Bacillus subtilis*. *Curr. Opin. Microbiol.* **8:**188-195.

Mukhopadhyay, R. and B.P. Rosen. (1998). *Saccharomyces cerevisiae* ACR2 gene encodes an arsenate reductase. *FEMS Microbiol. Lett.* **168:**127-136.

Mukhopadhyay, R., B.P. Rosen, L. T. Phung and S. Silver. (2002). Microbial arsenic: from geocycles to genes and enzymes. *FEMS Microbiol. Rev*. **26:**311-325.

Neyt, C., M. Iriarte, V. Thi and G. Cornelis. (1997). Virulence and arsenic resistance in *Yersiniae*. *J. Bacteriol.* **179:**612-619.

Nies, D.H. and S. Silver. (1995). Ion efflux systems involved in bacterial metal resistance. *J. Ind. Microbiol.* **14:**186-199.

Ng, W.V., S.A. Ciufo, T.M. Smith, R.E. Bumgerner, D. Baskin, J. Faust, B. Hall, C. Loretz, J. Seto, J. Slagel, L. Hood and S. DasSarma. (1998). Snapshot of a large dynamic replicon in a halophilic archaeon: megaplasmid or minichromosome? *Genome Res*. **8:**1131-1141.

Norris, P.R., N.P. Burton and N.A.M. Foulis. (2000). Acidophiles in bioreactor mineral processing. *Extremophiles*. **4:**71-76.

Norris, P.R., D.A. Clark, J.P. Owen and S. Waterhouse. (1996). Characteristics of *Sulfobacillus acidophilus* sp. nov. and other moderately thermophilic mineral sulphide-oxidizing bacteria. *Microbiology*. **141:**775-783.

Norris, P.R. (1997). Thermophiles and bioleaching. *In* Rawlings D.E. (Eds.), *Biomining: theory, microbes and industrial processes*. Springer-Verlag, Berlin, Germany. pp. 247-258.

Novick, R.P. and C. Roth. (1968). Plasmid-linked resistance to inorganic salts in *Stapphylococcus aereus. J. Bacteriol.* **95:**1335-1342.

Oden, K.L., T.B. Gladysheva and B.P. Rosen. (1994). Arsenate reduction mediated by the plasmid-encoded ArsC protein is coupled to glutathione. *Mol. Microbiol.* **12:**301-306.

Okibe, N. and D.B. Johnson. (2004). Biooxidation of Pyrite by Defined Mixed Cultures of Moderlately Thermophilic Acidophiles in pH-Controlled Bioreactors: Signicance of Microbial Interactions. *Biotechnol. Bioeng.* **87:**574-583.

Okibe, N., M. Gericke, K.B. Hallberg and D.B. Johnson. (2003). Enumeration and Characterization of Acidophilic Microorganisms Isolated from a Pilot Plant Stirred-tank Bioleaching operation. *Appl. Environ. Micobial*. **69:**1936-1943.

Olive, D.M. and P. Bean. (1999). Principles and Applications of Methods for DNA-Based Typing of Microbial Organisms. *J. Clin. Microbiol*. **37:**1661-1669.

Oremland, R.S. and J.F. Stolz. (2003). The ecology of arsenic. *Science* **300:**939-944.

Oshima, M. and T. Ariga. (1975). Omega-cyclohexyl fatty acids in acidophilic thermophilic bacteria: studies on their presence, structure and biosynthesis using precursors lableled with stable isotopes and radioisotopes. *J. Biol. Chem.* **250:**6963-6968.

Panin, V.V., G.I. Karavaiko and S.I. Pol'kin. (1993). Mechanism and Kinetics of Bacterial Oxidation of Sulphide Minerals. *Biotechnology of Metals*, *In* Karavaiko, G.I. and Groudov S.N. (Eds.), 197, Centre of International Projects, Moscow (1985) as cited by Barrett, J. and Hughes, M.N. The Mechanism of the Bacterial Oxidation of Arsenopyrite-pyrite Mixtures: The Identification of Plant Control Parameters. *Miner. Eng.* **6:**969.

Pol'kin, S.I., V.V. Panin, E.V. Adamov, G.I. Karavaiko and A.S. Chernyak. (1975). Theory and Practice of Utilizing Microorganisms in Difficult-to-dress Ores and Concentrates. *11th International Mineral Processing Congress,* Cagliari, 901.

Prithivirajsingh, S., S. Mishra and A. Mahadevan. (2001). Detection and analysis of chromosomal arsenic resistance in *Pseudomonas fluorescens* strain MSP3. *Biochem. Biophys. Res. Commun.* **280:**1393-1401.

Qin, J., B.P. Rosen, Y. Zhang, G. Wang, S. Franke and C. Rensing. (2006). Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *PNAS*. **103:**2075-2080.

Rai, M. and H. Padh. (2001). Expression systems for production of heterologous proteins. *Curr. Sci. India* **80:**1121-1128.

Rawlings, D.E. (1997). Mesophilic, Autotrophic Bioleaching Bacteria: Description, Physiology and Role. *In* Rawlings D.E. (Eds.), *Biomining: theory, microbes and industrial processes*. Springer-Verlag, Berlin, Germany. pp. 229-245.

Rawlings, D.E. (2002). Heavy metal mining using microbes. *Annu. Rev. Microbiol.* **56:**65-91.

Rawlings, D.E. (2005). Characteristics and adaptability of iron-and sulfur-oxidizing microorganisms used for the recovery of metals from minerals and their concentrates. *Microbial Cell Fact.* **4:**1-15.

Rawlings, D.E. and S. Silver. (1995). Mining with microbes. *Biotechnology* 13:773-778.

Rawlings, D.E., D. Dew, and C. du Plessis. (2003). Biomineralization of metal-containing ores and concentrates. *Trends Biotech*. **21:**38-44.

Rensing, C., M. Ghosh and B.P. Rosen. (1999). Families of Soft-Metal-Ion-Transporting ATPases. *J. Bacteriol.* **181:**5891-5897.

Robertson, W.J., P.H.-M. Kinnunen, J.J. Plumb, P.D. Franzmann, J.A. Puhakka, J.A.E. Gibson and P.D. Nichols. (2002). Moderately thermophilic iron oxidizing bacteria isolated from a pyritic coal deposit showing spontaneous combustion. *Miner. Eng.* **15:**815-822.

Roboz, G.J., S. Dias, G. Lam, W.J.Lane, S.L. Soignet, R.P. Warrell and S. Rafii. (2000). Arsenic trioxide induces dose- and time-dependent apoptosis of endothelium and may exert and antileukemic effect via inhibition of angiogenesis. *Blood* **96:**1525-1530.

Rohwerder, T., T. Gehrke, K. Kinzler and W. Sand. (2003). Bioleaching review part A. *Appl. Microbiol. Biotech*. **63:**239-248.

Rosen, B.P. (1999). Families of arsenic transporters. *Trends Microbiol*. **7:**207-212.

Rosen, B.P., H. Bhattacharjee, T. Zhou and A.R. Walmsley. (1999). Mechanism of the ArsA ATPase. *Biochimica et Biophysica Acta (BBA)*. **1461:**207-215.

Rosen, B.P. (2002). Transport and detoxification systems for transition metals, heavy metals and metalloids in eukaryotic and prokaryotic microbes. *Comp. Biochem. Physiol*. **133:**689-693.

Rosenstein, R., P. Peschel, B. Weiland and F. Gotz. (1992). Expression and Regulation of the *Staphyloccus xylosus* Antimonite, Arsenite and Arsenate Aesistance Operon of *Staphyloccus xylosus* Plasmid pSX267. *J. Bacteriol.* **174:**3676-3683.

Rosenstein, R., K. Nikoleit and F. Gotz. (1994). Binding of ArsR, the repressor of the *Staphylococcus xylosus* (pSX267) arsenic resistance operon to a sequence with a dyad symmetry within the *ars* promoter. *Mol. Gen. Genet*. **242:**566-572.

Ryan, D. and E. Colleran. (2002). Arsenical resistance in the IncHI2 plasmids. *Plasmid* **47:**234-240.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular cloning: a Laboratory Manual*, 2nd edition. Cold Spring Harbor Laboratory. NY: Cold Spring Harbor.

San Francisco, M.J.D., L.S. Tisa and B.P. Rosen. (1989). Identification of the membrane component of the anion pump encoded by the arsenical resistance operon of R-factor R773. *Mol. Microbiol.* **3:**15-21.

Sato, T. and Y. Kobayashi. (1998). The *ars* operon in the *skin* element of *Bacillus subtilis* confers resistance to arsenate and arsenite. *J. Bacteriol.* **180:**1655-1661.

Schnell, H.A. (1997). Bioleaching of copper. *In* Rawlings, D.E. (Eds.), *Biomining: theory, microbes and industrial processes.* Springer-Verlag, Berlin, pp. 21-43.

Schutte, M. (2004). Molecular characterization of *Sulfobacillus* and related organisms. M. Sc. thesis dissertation. University of Stellenbosch.

Shi, W., J. Wu and B.P. Rosen. (1994). Identification of a Putative Metal Binding Site in a New Family of Metalloregulatory Proteins. *J. Biol. Chem*. **269:**19826-19829.

Shi, W., J. Dong, R.A. Scott, M.Y. Ksenzenko and B.P. Rosen. (1996). The role of arsenic-thiol interactions in metalloregulation of the *ars* operon. *J. Biol. Chem.* **271:**9291-9297.

Shi, J., R. Mukhopadhyay and B.P. Rosen. (2003). Identification of a triad of arginine residues in the active site of the ArsC arsenate reductase of plasmid R773. *FEMS Microbiol. Lett.* **227:**295-301.

Shrestha, G.N. (1988). Microbial leaching of refractory gold ores. *Australian Mining*. October, 48.

Silver, S. (1996). Bacterial resistance to toxic metal ions-a review. *Gene.* **179:**9-19.

Silver, S., K. Budd, K.M. Leahy, W.V. Shaw, D. Hammond, R.P Novich, G.R. Willsky, M.H. Malamy and H. Rosenberg. (1981). Inducible plasmid-determined resistance to arsenate, arsenite and antimony(III) in *Escherichia coli* and *Staphylococcus aereus*. *J. Bacteriol.* **146:**983-996.

Silver, S., L.T. Pung and B.P. Rosen. (2002). Arsenic Metabolism: Resistance, Reduction, and Oxidation. *In* Frankenberger, W.T. Jr. (Eds.), *Environmantal Chemistry of Arsenic*, Marcel Dekker, New York, NY. pp. 247-272.

Silver, S. and M. Walderhaug. (1992). Gene Regulation of Plasmid- and Chromosomal-Determined Inorganic Ion Transport in Bacteria. *Microbiol. Rev.* **56:**195-228.

Spiegelman, D., G. Whissell and C. W. Greer. (2005). A survey of the methods for the characterization of microbial consortia and communities. *Can. J. Microbiol*. **51:**355-386.

Stolz, J.F., P. Basu and R.S. Oremland. (2002). Microbial transformation of elements: the case of arsenic and selenium. *Int. Microbiol*. **5:**201-207.

Suzuki, I. (2001). Microbial leaching of metals from sulfide minerals. *Biotechnol. Adv*. **19:**119-132.

Suzuki, K., N. Wakao, T. Kimura, K. Sakka and K. Ohmiya. (1998). Expression and regulation of the arsenic resistance operon of *Acidiphilium multivorum* AIU 301 plasmid pK301 in *Escherichia coli. Appl. Environ. Microbiol.* **64:**411-418.

Suzuki, K.I., K. Saito, A. Kawaguchi *et al*., (1981). Occurrence of ω-cyclohexyl fatty acids in Curtobacterium pusillum strains. *J. Gen. Appl. Microbiol.* **27:**261-266.

Tamaki, S. and W. Frankenberger. (1992). Environmental biochemistry of arsenic. *Rev. Environ. Contam. Toxicol.* **124:**79-110.

Taylor, C.D. and C.O. Wirsen. (1997). Microbiology and ecology of filamentous sulfur formation. *Science*. **277:**1438-1485.

Tisa, L.S. and B.P. Rosen. (1990). Molecular Characterization of an Anion Pump. *J. Biol. Chem.* **265:**190-194.

Trindade M., V.R. Abratt and S.J. Reid. (2003). Induction of sucrose utilization genes from *Bifidobacterium lactis* by sucrose and raffinose. *Appl. Environ. Microbiol*. **69:**24-32.

Tsai, K., C. Hsu and B.P. Rosen. (1997). Efflux Mechanisms of Resistance to Cadmium, Arsenic and Antimony in Prokaryotes and Eukaryotes. *Zool. Stud*. **36:**1-16.

Tsaplina, I.A., G.A. Osipov, T.I. Bogdanova, T.P. Nedorezova and G.I. Karavaiko. (1994). Fatty acid composition of lipids in thermoacidophilic bacteria of the genus *Sulfobacillus*. *Microbiology*. **63:**459-464.

Tuffin, I.M., P. de Groot., S.M. Deane and D.E. Rawlings. (2005). An unusual Tn*21*-like transposon containing an *ars* operon is present in highly arsenic-resistant strains of the biomining bacterium *Acidithiobacillus caldus*. *Microbiology*. **150:**1-13.

Tuffin, I.M., P. de Groot., S.M. Deane and D.E. Rawlings. (2004). Multiple sets of arsenic resistance genes are present within highly arsenic-resistant industrial strains of the biomining bacterium, *Acidithiobacillus caldus. International Congress Series.* **1275:**165-172.

Tuffin, I.M., S. B. Hector, S.M. Deane and D.E. Rawlings. (2006). Resistance Determinants of a highly Arsenic-Resistant Strain of *Leptospirillum ferriphilum* Isolated from a Commercial Bioxxidation Tank. *Appl. Environ. Microbiol.* **72:**1-7.

Tuovinen, O.H., T.M. Bhatti, J.M. Bigham, K.B. Hallberg, O. Garcia, Jr. and E.B. Lindström. (1994). Oxidative Dissolution of Arsenopyrite by Mesophilic and Moderately Thermophilic Acidophiles. *Appl. Environ. Microbiol.* **60:**3268-3274.

Valenzuela, L., A. Chi, S. Beard, A. Orell, N. Guiliani, J. Shabanowitz, D.F. Hunt and C.A. Jerez. (2006). Genomics, metagenomics and proteonomics in biomining microorganisms. *Biotechnol. Adv.* **24:**197-211.

Van Kranenburg, R., N. Golic, R. Bongers, R.J. Leer, W.M. de Vos, R.J. Siezen and M. Kleerebezem. (2005). Functional Analysis of Three Plasmids from *Lactobacillus plantarum*. *Appl. Environ. Microbiol.* **71:**1223-1230.

Vartanyan, N.S., T.A. Pivovarova, I.A. Tsaplina, A.M. Lysenko and G.I. Karavaiko. (1986). New thermoacidophilic bacterium of the genus *Sulfobacillus*. *Microbiology.* **57:**268-274.

Wang, L., S. Chen, X. Xiao, X. Huang, D. You, X. Zhou and Z. Deng. (2006). *arsRBOCT* Arsenic Resistance System Encoded by Linear Plasmid pHZ227 in *Streptomyces* sp. Strain FR-008. *Appl. Environ. Microbiol.* **72**:3738-3742.

Wang, G., S.P. Kennedy, S. Fasiludeen, C. Rensing and S. DasSarma. (2004). Arsenic Resistance in *Halobacterium* sp. Strain NRC-1 Examined by Using an Improved Gene Knockout System. *J. Bacteriol.* **186:**3187-3194.

Willsky, G.R. and M.H. Malamy. (1980). Effect of arsenate on Inorganic Phosphate Transport in *Escherichia coli*. *J. Bacteriol*. **144:**366-374.

Wu, J. and B.P. Rosen. (1991). The ArsR protein is a *trans*-acting regulatory protein. *Mol. Microbiol*. **5:**1331-1336.

Wu, J., L.S. Tisa and B.P. Rosen. (1992). Membrane topology of the ArsB protein, the membrane subunit of an anion-translocating ATPase. *J. Biol. Chem.* **267:**12570-12576.

Wu, J and B.P. Rosen. (1993a). Metalloregulated Expression of the *ars* Operon. *J. Biol. Chem*. **268:**52-58.

Wu, J and B.P.Rosen. (1993b). The *arsD* gene encodes a second trans-acting regulatory protein of the plasmid-encoded arsenical resistance operon. *Mol. Microbiol.* **8:**615-623.

Xu, C., W. Shi and B.P. Rosen. (1996). The chromosomal *arsR* Gene of *Escherichia coli* Encodes a transacting Metalloregulatory Protein. *J. Biol. Chem*. **271:**2427-2432.

Xu, C. and B.P. Rosen. (1997). Dimerization Is Essential for DNA Binding and Repression by the ArsR Metalloregulatory Protein of *Escherichia coli*. *J. Biol. Chem.* **272:**15734-15738.

Yahya, A. and D.B. Johnson. (2002). Bioleaching of pyrite at low pH and low redox potentials by novel mesophilic Gram-positive bacteria. *Hydrometallurgy.* **63:**181-188.

Yahya, A., F.F. Roberto and D.B. Johnson. (1999). Novel mineral-oxidising bacteria from Montserrat (W.I.): physiological and phylogenetic characteristics. *In* Amils, R., A. Ballester, *Biohydrometallurgy and the Environment Toward the Mining of the 21st Century*, vol **9A**. Elsevier, Amsterdam, pp. 729-740.

Yang, H.C., J. Cheng, T.M. Finan, B.P. Rosen and H. Bhattacharjee. (2005). Novel pathway for arsenic detoxification in the legume symbiont *Sinorhizobium meliloti. J. Bacteriol.* **187:**6991-6997.

Zegers, I., J.C. Martins, R. Willem, L. Wyns and J. Messens. (2001). Arsenate reductase from *S. aureus* plasmid pI258 is a phosphate drafted for redox duty. *Nat. Struc. Biol.* **8:**843-847.

Zhou, T. and B.P. Rosen. (1997). Tryptophan Fluorescence Reports Nucleotide-induced Conformational Changes in a Domain of the ArsA ATPase. *J. Biol. Chem.* **272:**19731-19737.

Zhou, T., S. Radaev, B.P. Rosen and D.L. Gatti. (2000). Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump. *The EMBO Journal.* **19:**4838-4845.

Zhou, T., S. Radaev, B.P. Rosen and D.L. Gatti. (2001). Conformational Changes in Four Regions of the *Escherichia coli* ArsA ATPase Link ATP Hydrolysis to Ion Translocation. *J. Biol. Chem*. **276:**30414-30422.