THE REGULATION AND FUNCTION OF THE ESAT-6 GENE CLUSTER OPERONS OF *MYCOBACTERIUM TUBERCULOSIS*

Jeanine Botha

Dissertation presented for approval for the degree of Masters in Medical Sciences (Medical Biochemistry) at the University of Stellenbosch

Promoter: Dr. N. C. Gey van Pittius

Co-promoter: Prof. R. M. Warren

Stellenbosch

December 2006

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work, and has not, to my knowledge, previously in its entirety or in part been submitted at any university for a degree.

……………………………….. ……..……………………….

J. Botha (Student number: 12890464) Date

Summary

The ESAT-6 gene cluster regions are duplicated 5 times in the genome of *Mycobacterium tuberculosis*. ESAT-6 gene cluster region 1 is the most frequently studied region as it contains RD1 (region of difference 1). RD1 is a 9.5 Kb deletion region confirmed to be involved in mycobacterial virulence and pathogenesis, and is present in virulent *M. bovis* strains, yet absent in all attenuated *M. bovis* BCG vaccine strains. The antigens CFP-10 and ESAT-6, which both evoke strong T-cell responses in experimental animals and humans, are situated in the RD1 region, and are thought to be key antigens in mycobacterial virulence. The absence of this region from the genomes of all BCG vaccine strains, led to the conclusion that the mechanism of attenuation of *M. bovis* BCG was due to the loss of RD1. Studies have shown that this attenuation is attributed to the loss of cytolytic activity mediated by secreted ESAT-6 (and some of the genes responsible for its secretion), which in turn results in reduced tissue invasiveness.

The potent T-cell antigens ESAT-6 and CFP-10 are secreted without ordinary *sec*-dependent secretion signals. A study of the potential functions of the proteins encoded by the ESAT-6 gene clusters shows that most of these proteins have a potential to function in a protein-dependent ATP-binding cassette active transport system. It has been shown that ESAT-6 gene cluster region 1 is responsible for the secretion of the ESAT-6 and CFP-10 genes contained in this region, explaining the absence of any ordinary *sec*-dependent secretion signals in the amino acid sequences of members of this family.

In order to elucidate the regulation of expression of the ESAT-6 gene cluster region 1, shown to encode for a secretion system for ESAT-6 and CFP-10 and to be involved in virulence, an operon analysis and promoter identification experiments were carried out in this study. The analysis of the ESAT-6 gene cluster region 1 showed the existence of more than one operon in this region and three constitutively-expressed promoters driving the expression of the genes in the operons. These results provide insight into the functional relationship (regulatory and secretory mechanisms) between the genes contained within ESAT-6 gene cluster region 1.

None of the other four ESAT-6 gene cluster regions have been proven to also encode secretion systems. Preliminary studies indicated that the ESAT-6 gene cluster region 3 is expressed in its entirety as one single operon and a strong promoter involved in the expression of this region was identified. Mtb9.9A (the ESAT-6 antigen of the ESAT-6 gene cluster region 5) have also been shown to evoke strong T cell responses and to be secreted without any ordinary secretion signal. During the present study, we thus aimed to investigate the secretion of Mtb9.9A in order to determine whether it is also secreted by a dedicated secretion system encoded by ESAT-6 gene cluster region 5. The fact that region 5 was shown to be the last of the four duplications is important, as a positive result with this region would indicate whether the other four gene clusters share a similar secretion function.

ESAT-6 gene cluster regions 2, 4 and 5 were isolated in the present study to form part of subsequent ESAT-6 gene cluster region secretion studies. Mtb9.9A was cloned, expressed and purified for antibody-generation, Resulting antibodies were used in an antigen secretion analysis. The secretion analysis entailed the integration of the isolated ESAT-6 gene cluster region 5 into the genome of *M. smegmatis* and investigation of the influence of the genes (contained in region 5) on the secretion of a heterologously expressed Mtb9.9A-HA-tagged fusion protein. We therefore attempted to show whether the proteins encoded by the ESAT-6 gene cluster region 5 also function together as a mycobacterial membrane-bound complex involved in protein-dependent transport and if so, whether this transport system is responsible for the active secretion of the native ESAT-6 antigen (designated Mtb9.9A) of region 5.

Pectora roborant cultus recti

This study opens the way for the understanding of the regulation, transport- and secretion mechanisms of important T-cell antigens of the mycobacteria, thereby giving insight into and building onto our understanding of the pathogenicity of *Mycobacterium tuberculosis.* A better understanding of these mechanisms could lead to the development of efficient strategies to either terminate or enhance secretion of antigens, which in turn will have an impact on drug and vaccine design and development.

Opsomming

Die ESAT-6 geen groep gebiede word 5 keer herhaal deur die genoom van *Mycobacterium tuberculosis*. Die ESAT-6 geen groep 1 is die mees bestudeerde geen groep omdat dit die RD1 delesie geen gebied bevat. RD1 is ‘n 9.5 Kb gebied wat betrokke is by mikobakteriële virulensie and patogenie en is teenwoordig in *M. bovis* virulente stamme, maar afwesig in alle nie-virulente *M. bovis* BCG vaksien stamme. Die antigene, CFP-10 en ESAT-6, wat beide ‘n sterk immuun reaksie in laboratorium diere en mense uitlok is ingesluit in die RD1 geen gebied en is belangrike antigene wat betrokke is by mikobakteriële virulensie. Die afwesigheid van RD1 uit die genome van alle BCG vaksien stamme, het aanleiding gegee tot die gevolgtrekking dat die meganisme van *M. bovis* se ontwikkeling tot die nie-virulente *M. bovis* BCG deur die delesie van RD1 toegeskryf kan word. Studies het ook gewys dat die ontwikkeling tot nie-virulensie ‘n gevolg is van verlore sitolitiese aktiwiteit (gemedieer deur gesekreteerde ESAT-6 asook die gene verantwoordelik vir die sekresie van ESAT-6) wat weer tot gevolg gehad dat weefsel inname deur die bakterie verlaag is.

Die T-sel antigene, ESAT-6 en CFP-10, word gesekreteer sonder enige normale *sec*-afhanklike sekresie-seine. ‘n Studie oor die potensiële funksies van die proteïne, gekodeer deur die ESAT-6 geen groepe, wys dat die meeste van hierdie proteïne ‘n potensiaal het om te funksioneer in ‘n protein-afhanklike ATP bindings-kasset aktiewe transport sisteem. Daar is gewys dat die ESAT-6 geen groep 1 verantwoordelik is vir die sekresie van die ESAT-6 en CFP-10 gene in hierdie gebied, wat die afwesigheid van enige *sec*-afhanklike sekresie-seine in die aminosuur volgorde van lede van hierdie familie verduidelik.

‘n Operon analise en promoter identifikasie eksperimente is gedurende die studie uitgevoer om die regulering van die uitdrukking van die ESAT-6 geen groep 1, wat vir ‘n sekresie sisteem van ESAT-6 en CFP-10 kodeer en betrokke is by virulensie, te bepaal. Die operon analise van die ESAT-6 geen groep 1 het die bestaan van meer as een operon en drie konstitatiewe uitgedrukte promoters, wat die uitdrukking van die gene in die operons aandryf, in hierdie geen groep aangedui. Die resulte het inligting verskaf in verband met die

funksionele verhouding (regulerend en sekreterend) tussen die gene wat ingesluit is in geen groep 1.

Nie een van die ander vier ESAT-6 geen groepe is bewys om sekresie sisteme te kodeer nie. Vorige studies het aangedui dat die ESAT-6 geen groep 3 uitgedruk word as een operon waarby 'n sterk promoter betrokke is by die uitdrukking van die geen groep. Daar is bewys dat Mtb9.9A (die ESAT-6 antigeen ingesluit in die ESAT-6 geen groep 5) ook sterk T-sel reaksies ontlok en sonder enige gewone sekresie sein gesekreteer word. Gedurende hierdie studie het ons dus beoog om die sekresie van Mtb9.9A te ondersoek om sodoende te bepaal of dié antigeen ook gesekreteer word deur 'n sekresie sisteem van die ESAT-6 geen groep 5. Die feit dat geen groep 5 die laaste duplikasie van die 4 duplikasies van die ESAT-6 geen groepe is, is belangrik omdat 'n positiewe resultaat van hierdie geen groep 'n aanduiding sal wees of die ander 4 geen groepe 'n verwante sekresie funksie toon.

ESAT-6 geen groepe 2, 4 en 5 is geïsoleer om deel te vorm van toekomstige ESAT-6 geen groep sekresie studies. Mtb9.9A is gekloneer, uitgedruk en geïsoleer vir die vervaardiging van teenliggame teen dit. Die gevolglike teenliggame is gebruik in 'n antigeen sekresie analise. Die sekresie analise het die integrering van die ESAT-6 geen groep 5 in die genoom van *M. smegmatis* behels asook die invloed van die gene (ingeslote in geen groep 5) op die sekresie van die heterologiese uitgedrukte Mtb9.9A-HA fusie-proteïn. Daar dus gepoog om te wys of die proteïne gekodeer deur die ESAT-6 geen groep 5 saam funksioneer as 'n mikobakteriële membraan-gebonde kompleks, betrokke by protein-afhanklike transport. Indien bogenoemde bevestig word, sal daar ook gepoog word om te bewys dat hierdie transport-sisteem verantwoordelik is vir die sekresie van die ESAT-6 geen groep 5 ESAT-6 antigeen, Mtb9.9A.

Hierdie studie sal die weg baan om die mikobakteriële regulering, transport- en sekresie meganismes van belangrike T-sel antigene van mikobakterië (wat insig sal gee om die patogenisiteit van *Mycobacterium tuberculosis*) beter te verstaan. Dit kan lei tot die ontwikkeling van strategië om die sekresie van antigene te beëindig of te verbeter wat gevolglik 'n impak sal hê op vaksien ontwerp en ontwikkeling.

Acknowledgements

Eerstens wil ek my wonderlike, almagtige Hemelse Vader bedank vir Sy hand van bewaring oor my en hulp deur die tyd van navorsing en die skryf van my tesis. Sonder U genade en seëninge sou ek nie gekom het waar ek vandag is nie.

Tweedens, Mamma, ongelooflik baie dankie dat Mamma die beste moeder en vriendin is wat enige dogter kan begeer. Dankie vir al Mamma se liefde, gebede en ondersteuning deur hierdie tyd. Dankie dat Mamma in my bly glo het, selfs in tye waar ek amper moed opgegee het en nie meer in myself geglo het nie. Ek het nie woorde om Mamma genoeg te bedank nie. Het Mamma liewer as lief…

Dan my sussie, Wilma. Wat sê mens vir ‘n goue mens, so propvol begrip en onselfsugtige liefde. Dankie dat jy my noord-ster is, my kompas. Jy het my nie help dra nie… jy het my gedra. Ek is so lief vir jou.

Jannes, dankie dat jou liefde, begrip en ondersteuning onvoorwaardelik en onophoudelik was deur hierdie tyd. Selfs toe ek op moedsverlore vlakte was het jy my ondersteun, moed ingepraat en bly koffietjie aandra toe ek voor die rekenaar gesit het. Ek het jou baie lief.

Die res van my familie: Mariska, Johantjie, Bernardus, Celeste en Johan. Ek waardeer jul ondersteuning en moed-inpraat geweldig baie. Ek is lief vir julle en bevoorreg om julle in my lewe te hê.

Verder, my vriende en vriendinne: Michelle, Jacolien, Lizl, Gerda-Marie, Natalie, Stephanie, Ameri, Steven, Pieter, Christiaan, Roper, Pieter R. Baie dankie vir die rol wat julle in my lewe speel. Jul liefde en ondersteuning en daar-wees word opreg waardeer.

Pectora roborant cultus recti

Nico, vir jou wil ek ekstra baie spesiaal dankie sê. Dankie vir al jou geduld, begrip en ondersteuning tydens ons saamwerk en die skryf van my tesis. Ek sal jou nooit genoeg kan bedank vir alles nie. Ek waardeer jou as promotor en goeie vriend wat my geinspireer het om altyd nog ‘n trappie hoër te klim, nog hoër standaarde te stel en dat jy my geleer het hoe wonderlik die wêreld van navorsing kan wees.

Rob and Prof van Helden. Thank you for giving me this wonderful and life-changing opportunity to be part of your research-group and believing in me, but above all for your patience and giving me freedom to follow my dreams. It was truly a privilege and an enriching experience.

Don, het nie woorde om jou genoeg te bedank vir al jou hulp en bystand nie. Jy is dierbaar, geduldig en sommer net die beste. Jy was nie net ‘n kollega nie, maar is ‘n baie goeie vriend. Jy het altyd gesorg dat ‘n frons op my gesig plek gemaak het vir ‘n glimlag. Dankie!

The rest of the Medical Biochemistry Department: Thank you for your friendliness and excepting me as one of your own.

Pectora roborant cultus recti

This project was financially supported by the following institutions to which I am extremely grateful: Harry Crossley Foundation, MRC of South Africa and the University of Stellenbosch.

Table of Contents

		Page Number
1)	**Chapter 1: Introduction**	**1**
	1.1 Background	2
	1.2 Problem Statement	12
	1.3 Hypothesis	13
	1.4 Preliminary Studies	14
	1.5 Aims	15
2)	**Chapter 2: Promoter analysis of the ESAT-6 gene cluster region 1**	**16**
	2.1 Introduction	17
	2.2 Materials and Methods	19
	2.3 Results	28
	2.4 Discussion and Conclusion	37
3)	**Chapter 3: Isolation of the ESAT-6 gene cluster regions 2, 4 and 5**	**39**
	3.1 Introduction	40
	3.2 Materials and Methods	41
	3.3 Results	47
	3.4 Discussion and Conclusion	51
4)	**Chapter 4: Expression and purification of the ESAT-6 protein, Mtb9.9A**	**53**
	4.1 Introduction	54
	4.2 Materials and Methods	56
	4.3 Results	61
	4.4 Discussion and Conclusion	65
5)	**Chapter 5: Secretion analysis of the ESAT-6 gene cluster region 5**	**66**
	5.1 Introduction	67
	5.2 Materials and Methods	68
	5.3 Results	73
	5.4 Discussion and Conclusion	76
6)	**Chapter 6: Conclusion and future directions**	**78**
7)	**Addendum**	**83**
8)	**List of References**	**87**

Pectora roborant cultus recti

List of Abbreviations

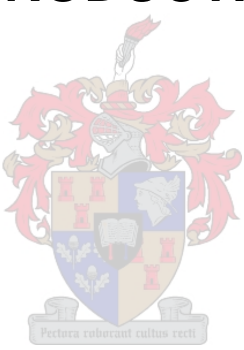
ABC	ATP-binding casette
Amp	Ampicillin
BCG	Bacille de Calmette et Guerin
BLAST	basic local alignment search tool
bp	basepair
BSA	bovine serum albumin
CF	culture filtrate
CFP-10	culture filtrate protein 10
^{0}C	degrees Celsius
CSU	California State University
DNA	deoxyribonucleic acid
E.	*Escherichiae*
EDTA	ethylenediaminotetraacetic acid
ESAT-6	6 kDa early-secreted antigenic target
G+C	guanine + cytosine
g/l	grams per litre
GST	glutathione-S-transferase
HA	hemagglutinin
HIV	human immunodeficiency virus
HRPO	horseradish peroxidase
Hyg	Hygromycin
Kan	Kanamycin
kDa	kilodalton
LB	Luria-Bertani (medium)
M.	*Mycobacterium*
µg	microgram
µM	micromolar
µl	microliter

M	molar
ng	nanogram
nm	nanometer
OADC	oleic acid/albumin/dextrose/catalase
OD	optical density
ONPG	*o*-nitrophenyl-β-D-galactoside
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
%	percent
P_{ESREG3}	ESAT-6 gene cluster region 3 promoter
PPD	purified protein derivative
RBS	ribosome binding site
RD	region of difference
RNA	ribonucleic acid
rpm	revolutions per minute
RT-PCR	reverse transcriptase polymerase chain reaction
SDS	sodium dodecyl sulphate
STE	sodium/tris/EDTA buffer
TAE	tris/acetic acid/EDTA buffer
TB	tuberculosis
TIGR	The Institute for Genomic Research

Pectora roborant cultus recti

CHAPTER ONE

INTRODUCTION



BACKGROUND

1.1 Brief history

For thousands of years people all over the world have been falling ill and dying of the respiratory/lung disease that they had known as consumption or in Greek terms, phthisis. Tuberculosis (TB), as it is known today, is a chronic infectious disease caused by the bacterium *Mycobacterium tuberculosis* (*M. tuberculosis*), which usually attack the lungs (pulmonary TB). *M. tuberculosis* can spread to and damage other parts of the body, such as the digestive and urogenital tracts, bones, joints, nervous system, lymph nodes and skin. This is called extrapulmonary TB, and is far less common. The disease is normally characterized by the development of granulomas or tubercles in infected tissues. Symptoms include the coughing up of blood, fever, weight loss and fatigue (Smith 2003).

In the initial stage of *M. tuberculosis* infection (about 3 to 8 weeks after *M. tuberculosis* contained in inhaled aerosols becomes implanted in alveoli), the bacteria are disseminated by the lymphatic circulation to regional lymph nodes in the lung. Then for the next approximately 3 months hematogenous circulation of bacteria to many organs including other parts of the lung takes place – sometimes in the form of fatal tuberculosis meningitis. Inflammation of the pleural surfaces can then occur during the next few months, but may last up to 2 years. Finally, more slowly developing extrapulmonary lesions may develop, e.g, those in bones and joints (van Crevel et al 2002 and Smith 2003)

Pectora roborant cultus recti

Tuberculosis has been identified in individuals as early as 8000, 5000 and 2400 BC, where bone and skeletal remains from the Neolithic period, the Heidelberg region and ancient Egypt respectively showed scars of tuberculosis (Herzog 1998). Similar bone deformities were identified in skeletons found throughout Italy, Denmark and the Middle East, indicating that TB was distributed all over the world at that time (Smith 2003).

The evolutionary origin of *M. tuberculosis* is uncertain. Some hypothesize that it originated from soil – from where it is believed to have initially infected cattle, after which it evolved to infect humans with the domestication of cattle 10000-25000 years ago (Smith 2003). Others believe that the two species of *M. tuberculosis* (human TB) and *M. bovis* (bovine TB) have diverged independently from a common precursor (Smida 1988). Brosch et al (2002) stated that the genome of *M. bovis* is smaller than that of *M. tuberculosis* (Gordon et al 2001), indicating that *M. bovis* is the final member of a separate lineage (represented by *M. africanum*, *M. microti* and *M. bovis*) that branched from the progenitor of *M. tuberculosis* (Brosch et al 2002).

In North America and Europe, tuberculosis increased during the 17th century and peaked in the 18th century. The reason for this occurrence was the unbearable socioeconomic conditions during the industrial revolution, which was the result of a sudden population explosion, malnutrition, poor hygienic and sanitary conditions and a lack of medical care (Murray 2004). However, during the early 19th century, a turnaround occurred when the death rate owing to tuberculosis decreased substantially, due to an improvement of socioeconomic conditions which entailed improved nutrition, living and working conditions and a step towards taking public health measures (Murray 2004).

Pectora roborant cultus recti

Different theories and misconceptions concerning the disease have been put forward by important physicians in history like Sylvius (1679), Manget (1702), Marten (1720), Brehmer (1854) and Villemin (1865). The first notion towards understanding tuberculosis came with the acknowledgment by Dr Benjamin Martenby that the disease may be caused by an airborne organism (McKinney et al 1999). However, it was not until 1882, when Robert Koch discovered a staining technique (acid-fast staining) which enabled him to see *M. tuberculosis* under the microscope, that the world started forming a better understanding of the disease. The next significant impact on the history of tuberculosis was when Wilhelm Konrad von Rontgen discovered the radiation that bears his name, known today as X-rays. A major breakthrough in the prevention of tuberculosis during the 19th century followed after Louis Pasteur developed the

technique to attenuate the virulence of a living microbe as a means of producing a vaccine. Following in his footsteps, the French bacteriologists Albert Calmette and Camille Guérin developed an attenuated form of *Mycobacterium bovis* (*M. bovis*) between 1908 and 1921. They had discovered that growth of *M. bovis* in ox bile reduced its virulence and after performing 231 serial passages of a single bacterium they showed that it had totally lost its ability to cause active and fatal tuberculosis in animals. The first *M. bovis* BCG vaccine immunization in humans was carried out in Paris in 1921(Herzog 1998, Murray 2004). Trials that have been conducted during the 1940s and 1950s in developed countries such as the UK, Denmark and North America revealed the BCG vaccine's efficiency of between 70 and 80% (Andersen 2001).

During the early 20th century, the only way to treat people was admittance to sanatoriums where patients could rest, eat nutritious food and get lots of fresh air. If no success followed with this treatment, physicians resolved to surgical treatment of pulmonary tuberculosis by means of pneumothorax (lung collapse) and thoracoplasty (rib removal; Herzog 1998). These forms of treatment ended with the discovery of a drug treatment. In 1940, in the middle of World War 2, Selman A. Waksman and his team were able to isolate an effective anti-TB antibiotic, actinomycin, but unfortunately it proved to be too toxic for use in humans and animals (Herzog 1998). However, success came in 1943, when streptomycin (purified from *Streptomyces griseus*) showed maximal inhibition of *M. tuberculosis* in test animals and in 1944 the drug was administered for the first time to a critically ill TB patient with great success. The only remaining side effect of streptomycin was the negative impact it had on hearing (Herzog 1998). There was an almost simultaneous development of another active anti-TB agent by Jörgen Lehmann in Sweden, known as para-aminosalicylic acid (PAS). This agent was designed to look like the salicylic acid that *M. tuberculosis* metabolized, but would kill the bacterium when it was broken down. The appearance of resistance against the monotherapy of either streptomycin or PAS in *M. tuberculosis* strains, made it evident that there was a need to develop new drugs and the use of a combination of antibiotics to overcome the acquisition of the resistance. Therefore, PAS

Pectora roborant cultus recti

(1949), isoniazid (1952), pyrazinamide (1954), cycloserine (1955), ethambutol (1962) and rifampin (rifampicin; 1963) were developed as anti-TB agents (Herzog 1998).

However, the success-rate of drug-treatment for TB has declined during the past fifty years. The reason for this is the non-compliance of patients taking anti-TB drugs, which causes the emergence of multi-drug resistant tuberculosis. Subsequently, these patients with multi-drug resistant tuberculosis could transmit the drug-resistant strain. The treatment of drug-resistant TB requires extensive chemotherapy (up to two years of treatment), is often more than 100 times more expensive than treatment of drug-susceptible TB, and is also more toxic to patients (WHO report 2005).

1.2 Tuberculosis today

Approximately eight million people develop active tuberculosis (TB) each year, with about 2 million people dying annually. It is also known that more than a third of the world's population has been infected with *M. tuberculosis* (WHO report 2005, Shin et al 2004, Stewart et al 2003, Mustafa 2002). Between 5 – 10% of individuals who become infected subsequently develop clinical disease (Bloom et al 1992). An individual with active pulmonary tuberculosis is able to infect on average between 10 and 15 people every year. The initial infections of most individuals are controlled by a cell-mediated immune response that inhibits the onset of disease. However, the remaining population of viable mycobacteria (with their thick cell walls providing resistance against the host's microbicidal mechanisms of macrophages) can lie dormant in the host for years. Disease is then caused by one of two possibilities: 1) reactivation by residual bacteria from the initial infection or 2) failure to control a subsequent reinfection (Selwyn et al 1989 and Stewart et al 2003).

Tuberculosis remains one of the leading killers among all infectious diseases. There is an increase in tuberculosis infection, which is caused by several factors including the growing HIV epidemic (Ravn et al 1999), the increase in multi-drug resistant *M. tuberculosis* strains (Andersen

Pectora roborant cultus recti

2001) and the waning efficacy of the currently used BCG vaccine (Brandt et al 2002, Sørensen et al 1995, Agger et al 2001, Agger et al 2002). Furthermore, being a live vaccine, BCG can be pathogenic in immunocompromised recipients (Mustafa et al 2003). It is hypothesized that the waning efficacy of the BCG vaccine can be attributed to interactions between the vaccine and environmental mycobacteria in tropical regions, the genetic variability in the populations studied, differences in BCG strains, doses and vaccination schedules and administration of the vaccine to individuals already infected (Andersen 2001). It has also been suggested that storage of the vaccine as well as inadequate strains of BCG could explain the declining protective effect of BCG (Fine et al 1995).

1.3 Mycobacterium tuberculosis

M. tuberculosis is a gram-positive, rod-shaped, slow-growing bacillus containing a complex cell wall with a very high lipid content. The presence of mycolic acids and other lipids outside the peptidoglycan layer makes mycobacteria acid-fast (basic fuchsin dye cannot be removed from the cell by acid alcohol treatment, Prescott et al 2002).

The *M. tuberculosis* complex includes *M. tuberculosis*, *M. bovis* (including *bovis* BCG), *M. africanum* and *M. microti*, the first three of which are the causitive agents of tuberculosis in humans and animals, and the last of which are pathogenic only in rodents (Brosch et al 2000).

The complete sequencing of the genome of *M. tuberculosis* laboratory strain H37Rv revealed its size of 4.40 Mb with a high G+C content of 67% and around 4000 genes (Cole et al 1998). Only approximately 40% of the genes have known functions and 16% bear a resemblance to known proteins. More than 250 of the organism's genes are responsible for lipid metabolism and it is thought that *M. tuberculosis* obtains much of its energy from host lipid degradation (Prescott et al 2002).

Although a great deal has been learned from the genome sequencing of some of the mycobacterial strains our comprehension of its pathogenicity is still limited. One of the reasons for this is the problems that scientists encounter during mycobacterial research. These include 1) the bacterium's slow growth, 2) the clumping together of cells in liquid media, 3) the fact that it has to be worked with under strict biolevel 3 conditions, 4) the genome's high G+C content which makes PCR amplification more difficult and 5) the difficulty to express recombinant mycobacterial proteins in *Escherichia coli*.

1.3 Mycobacterium tuberculosis virulence/pathogenicity

Unlike other disease-causing bacteria like *Escherichia coli*, *Shigella dysenteriae*, *Corynebacterium diphteriae* and *Vibrio cholerae*, which have classical virulence factors, the cause of virulence of *M. tuberculosis* still remains to be elucidated.

The vaccine strain *M. bovis* BCG has been used as a human tuberculosis vaccine for more than 80 years with more than 3 billion doses administered (Hsu et al 2003), yet we have only recently begun to understand the mechanisms that caused the attenuation of BCG. Mahairas et al (1996) compared the genome sequences of the avirulent BCG and virulent *M. bovis* and found three regions of difference (designated RD1, RD2 and RD3) present in the genome of *M. bovis* but absent in the genome of BCG. Subsequently, Behr et al (1999) identified 16 deletions (including RD1-RD3), present in the *M. tuberculosis* genome but absent in BCG. One of these deletions, RD1, is a 9455 bp region that is absent in all attenuated BCG strains. It was therefore hypothesized that RD1 was the primary cause of attenuation, which resulted in the generation of avirulent BCG from virulent *M. bovis*. It was recently confirmed that the RD1 region contributes to bacterial virulence (Hsu et al 2003, Lewis et al 2003, Pym et al 2002).

The RD1 gene products are fascinating for a variety of reasons: not only do they represent potential virulence factors, but they are also vaccine candidates. The best studied RD1 proteins are CFP-10 (culture filtrate protein 10 kDa; *esxB*) (Berthet et al 1998, Colangeli et al

2000, Dillon et al 2000) and ESAT-6 (early secreted antigen target 6 kDA; *esxA*) (Andersen et al 1995, Brandt et al 2000, Pallen 2002, Sorensen et al 1995), both of which encode T-cell antigens that elicit strong T cell responses in experimental animals and humans and lack detectable *secA* secretion signals (Colangeli et al 2000, Andersen et al 1995, Arend et al 2000, Brusasca et al 2001).

1.5 The ESAT-6 gene cluster regions

The completion of the genome sequencing of *M. tuberculosis* (Cole et al 1998) enabled the identification of five duplicated gene cluster regions, described as the ESAT-6 loci (Tekaia et al 1999, Gey van Pittius et al 2001, see Figure 1), due to the fact that the gene encoding the previously-identified ESAT-6 antigen was present within this region. One of these gene cluster regions also encompassed the previously-identified RD1 deletion region shown to be involved in virulence (see Figure 1). These gene clusters were designated region 1 (Rv3866-Rv3883c), region 2 (Rv3884c-Rv3895c), region 3 (Rv0282-Rv0292), region 4 (Rv3444c-Rv3450c) and region 5 (Rv1782-Rv1798), according to the numbering of another previously-identified gene family named the mycosins (Brown et al 1999). Apart from the mycosins (which are secreted cell-wall associated subtisilin-like serine proteases) and the members of the ESAT-6 and CFP-10 gene families, other genes included in these clusters encode for cell-wall-associated ABC-transporters and ATP-binding proteins (Gey van Pittius et al 2001, Brodin et al 2005).

ESAT-6 is an important T-cell antigen of *M. tuberculosis* and is not only present in multiple copies in association with the ESAT-6 gene clusters, but is also distributed as 6 additional single copies in the genome (Gey van Pittius et al 2001). Proteins of this family are about 100 amino acids (10 kD) and are characterized by a conserved motif Trp-X-Gly (WXG; Pallen et al 2002). ESAT-6 has been part of several vaccine studies where it is included in DNA and subunit vaccines since it elicits strong T- and B-cell responses in humans and animals (Andersen et al 1995, Colangeli et al 2000, Arend et al 2000, Brandt et al 2000, Olsen et al 2000,

Pectora roborant cultus recti

Skjøt et al 2001). It has been suggested that ESAT-6 plays a role in the lysis of alveolar epithelial cells and in escape from macrophages (Guinn et al 2004 and Hsu et al 2003).

The *esxA* and *esxB* genes (encoding ESAT-6 and CFP-10, respectively) are organised in an operon (Berthet et al 1998). Renshaw et al (2002) showed that ESAT-6 and CFP-10 form a tight 1:1 complex *in vitro*. Studies done by Gey van Pittius (PhD Thesis 2002), Pym et al (2003), Stanley et al (2003) and Guinn et al (2004) showed that ESAT-6 and CFP-10 (lacking a *sec*-dependant secretion signal) are secreted via the ESAT-6 system-1 (ESX-1), a dedicated secretion apparatus encoded by the ESAT-6 gene cluster region genes flanking *esxA* and *esxB*. The proteins implicated in this process consist of a member of the AAA-family of ATPases (Rv3868), which may perform chaperone-like functions and assist in the assembly and disassembly of the protein complexes, and a few putative membrane proteins with 1, 3 or 11 transmembrane domains (Rv3869, Rv3870, Rv3876 and Rv3877). The latter proteins as well as a protein with an ATP binding site (Rv3871), could be involved in forming a transmembrane channel for the translocation of the effector molecules (Hsu et al 2003, Pym et al 2003, Stanley et al 2003, Guinn et al 2004, Brodin et al 2004, Converse et al 2004).

Pectora roborant cultus recti

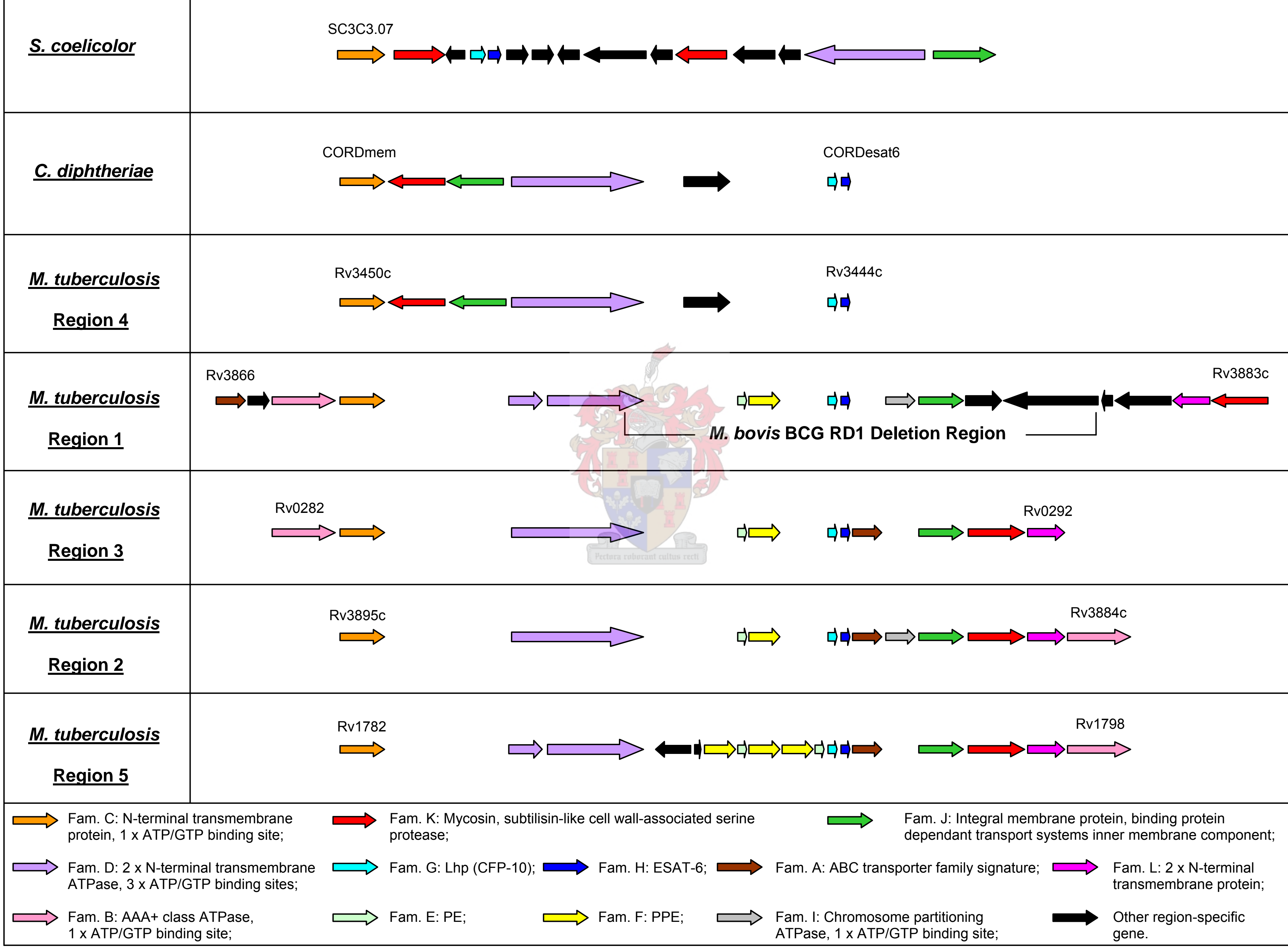
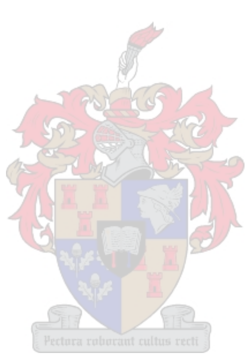

S. coelicolor
SC3C3.07
C. diphtheriae
CORDmem
CORDesat6
M. tuberculosis
Region 4
Rv3450c
Rv3444c
M. tuberculosis
Region 1
Rv3866
Rv3883c
M. bovis BCG RD1 Deletion Region
M. tuberculosis
Region 3
Rv0282
Rv0292
M. tuberculosis
Region 2
Rv3895c
Rv3884c
M. tuberculosis
Region 5
Rv1782
Rv1798
Fam. C: N-terminal transmembrane protein, 1 x ATP/GTP binding site;
Fam. K: Mycosin, subtilisin-like cell wall-associated serine protease;
Fam. J: Integral membrane protein, binding protein dependant transport systems inner membrane component;
Fam. D: 2 x N-terminal transmembrane ATPase, 3 x ATP/GTP binding sites;
Fam. G: Lhp (CFP-10);
Fam. H: ESAT-6;
Fam. A: ABC transporter family signature;
Fam. L: 2 x N-terminal transmembrane protein;
Fam. B: AAA+ class ATPase, 1 x ATP/GTP binding site;
Fam. E: PE;
Fam. F: PPE;
Fam. I: Chromosome partitioning ATPase, 1 x ATP/GTP binding site;
Other region-specific gene.

Figure 1: Schematic representation of the genomic organization of the genes present in the five ESAT-6 gene cluster regions of *M. tuberculosis* H37Rv as well as the regions in *C. diphtheriae* and *S. coelicolor*. ORF's are represented as blocked arrows showing the direction of transcription, with the different colours reflecting the specific gene family and the length of the arrow reflecting the relative lengths of the genes. Annotations of *M. tuberculosis* H37Rv genes are according to Cole et al (1998). Black arrows indicate unconserved genes present in these regions. Gaps between genes do not represent physical gaps between genes on the genome, but have been inserted to aid in indicating conservation among gene positions. Gene families were named arbitrarily according to their position in *M. tuberculosis* H37Rv region 1. The regions were named after the numbering system of Brown et al (2000) used arbitrarily for the five mycosin (subtilisin-like serine protease) genes identified from these regions. *M. tuberculosis* regions are shown in order of suggested duplication events and not by numbering (Gey van Pittius et al 2001).



PROBLEM STATEMENT

The regulation of expression of virulence-associated systems as well as secretion of virulence determinants of *M. tuberculosis* is largely unknown. The ESAT-6 gene cluster region 1 has been shown to play an important role in mycobacterial virulence and as such it is important to elucidate the regulation of expression of this system. Furthermore, it has been shown that this virulence is due to secretion of an important T-cell antigen family. The secretory mechanism/function of the ESAT-6 gene cluster region 5 has not yet been analyzed and is an important choice for secretion analyses, as it is the last duplicated ESAT-6 gene cluster region (in the phylogenetic duplication events that followed from the ancestral ESAT-6 gene cluster region 4) and thereby possesses the greatest chance of function changes and evolution. Furthermore, region 5 contains a potent T-cell antigen, Mtb9.9A, which has already been shown to be secreted by *M. tuberculosis* by an as yet unknown mechanism of secretion.

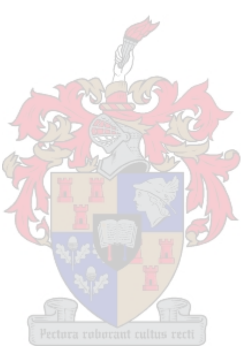
Pectora roborant cultus recti

HYPOTHESIS

We hypothesize that all five the ESAT-6 gene cluster regions are transcribed as one or more operons and are regulated by specific promoters. Furthermore, we hypothesize that the proteins encoded by the genes situated in these five ESAT-6 gene cluster regions, respectively, function together to form a dedicated multi-component secretion system for the specific secretion of the members of the ESAT-6 protein family located within the regions.

PRELIMINARY STUDIES

Preliminary studies have indicated that at least one of the ESAT-6 gene cluster regions (region 3) is expressed as a single polycistronic RNA, forming an operon structure, and is regulated by a single promoter designated $P_{ESREG3.}$ Initial results also suggest that the whole ESAT-6 gene cluster region 1 is required for the secretion of the ESAT-6 antigen, indicating that the genes included in region 1 function together to form a multi-component secretion system (Gey van Pittius PhD Thesis 2002). These results must be confirmed by studying the regulation and function of the other four ESAT-6 gene cluster regions.



AIMS

The aims of this study were:

1. to investigate the regulation of the ESAT-6 gene cluster region 1 by isolating and cloning its promoter(s) and looking at expression levels.

2. to isolate the ESAT-6 gene cluster regions 2, 4 and 5 for use in further studies.

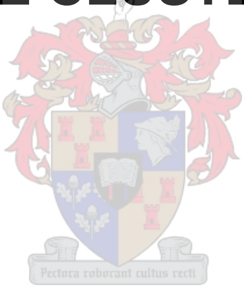
3. to isolate the ESAT-6 protein of the ESAT-6 gene cluster region 5 for use during secretion analysis studies.

4. to study the secretion of the members of the ESAT-6 antigen family by integrating the isolated ESAT-6 gene cluster region 5 into the genome of *Mycobacterium smegmatis* (*M. smegmatis)* and investigating their influence on the secretion of a heterologously expressed ESAT-6-HA fusion protein.

Pectora roborant cultus recti

CHAPTER TWO

PROMOTER ANALYSIS OF THE ESAT-6 GENE CLUSTER REGION 1



2.1 Introduction

The ESAT-6 gene cluster is situated in the genome of *M. tuberculosis* and other mycobacteria (Tekaia et al 1999 and Gey van Pittius et al 2001), as well as in the genomes of members of the Corynebacteria and Streptomyces (Gey van Pittius et al 2001). This gene cluster is duplicated five times in the genome of *M. tuberculosis* and contains members of the important T-cell antigen ESAT-6 gene family, leading to the clusters being designated the ESAT-6 loci (Tekaia et al 1999). These gene clusters are designated region 1 (Rv3866-Rv3883c), region 2 (Rv3884c-Rv3895c), region 3 (Rv0282-Rv0292), region 4 (Rv3444c-Rv3450c) and region 5 (Rv1782-Rv1798) consistent with the arbitrary numbering system used previously to classify the five mycosin (subtilisin-like serine protease) genes identified from these regions (Brown et al 2000).

Of the five duplicated ESAT-6 gene cluster regions of *M. tuberculosis*, region 1 has been the most frequently studied region. The reason for this is because it contains the RD1 (region of difference 1) deletion region (Mahairas et al 1996), which is a 9.5 Kb DNA segment present in virulent *M. bovis* strains, yet absent in all attenuated *M. bovis* BCG vaccine strains (Mahairas et al 1996 and Behr et al 1999). Ahmad et al (1999) showed that the RD1 region contains 14 open reading frames, of which RD1 antigens CFP-10 (Dillon et al 2000, Colangeli et al 2000, Berthet et al 1998) and ESAT-6 (Gey van Pittius et al 2001, Andersen et al 1995 and Sorensen et al 1995) both evoke strong T- and B-cell responses in experimental animals and humans (Colangeli et al 2000, Andersen et al 1995, Arend et al 2000 and Brusasca et al 2001). The RD1 deletion region is important due to the fact that it has recently been confirmed to be involved in mycobacterial virulence and pathogenesis (Pym et al 2002 and Lewis et al 2003). Hsu et al (2003) concluded that the mechanism of attenuation of *M. bovis* BCG (the loss of RD1) is attributed to the loss of cytolytic activity mediated by secreted ESAT-6 (and some of the genes responsible for its secretion), which in turn results in reduced tissue invasiveness.

The fact that the clusters are conserved between different bacterial species as well as the close proximity of the genes contained in the clusters indicates the possibility of one or more operons. Berthet et al (1998) identified an operon in the ESAT-6 gene cluster region 1 from which the *esxA* (encoding for ESAT-6) and *esxB* (encoding for CFP-10) are co-transcribed. More recently, it was shown that the ESAT-6 gene cluster region 3 is expressed in its entirety as one single operon and the promoter involved in the expression of ESAT-6 gene cluster region 3 was identified and designated P_{ESREG3} (Gey van Pittius PhD Thesis 2002, manuscript in preparation). The genes within the clusters encode for proteins involved in the active transport of the members of the ESAT-6 protein family (Tekaia et al 1999, Gey van Pittius et al 2001, Guinn et al 2003, Stanley et al 2003, Young et al 2003). Therefore, analysis of the operonic nature of the ESAT-6 gene cluster region 1 and possible promoters driving its expression will shed light on the functional relationship between the genes contained within the operon. This will aid in the improvement in the understanding of pathogenicity and virulence of *M. tuberculosis*.

In this study, we demonstrated the existence of 3 promoters driving the expression of ESAT-6 gene cluster region 1. We also showed that the ESAT-6 gene cluster region 1 is expressed as more than one polycistronic RNAs. These results form part of the initial steps to elucidate the regulatory and secretory mechanisms of the ESAT-6 antigen family of *M. tuberculosis*.

Pectora roborant cultus recti

2.2 Materials and Methods

2.2.1 DNA sequence analyses

Six intergenic regions were selected in order to identify potential promoters driving the expression of the genes contained within the ESAT-6 gene cluster region 1 (see Figure 2.1). These regions were selected on the basis of their sizes (at least 50 bp) and position within the clusters (upstream of both gene regions comprising region 1). Previous studies showed that the intergenic region Rv3874-75 contains a promoter that drives the expression of the *esxB*/*esxA* gene operon (Berthet et al 1999, Gey van Pittius, PhD Thesis 2002), so this region was not selected for further analyses. Intergenic regions Rv3862c-63, Rv3863-64 and Rv3864-65 were selected for this analysis because of their positions (upstream of the ESAT-6 gene cluster region 1, see Figure 1.1) and also because of their sizes, being 328 bp, 246 bp and 91 bp respectively (see Figure 1.1). Intergenic regions Rv3871-72 and Rv3875-76 were also selected because of their sizes (146 bp and 117 bp, respectively) as well as their positions lying up- (Rv3871-72) and downstream (Rv3875-76) of the *esxB*/*esxA* operon (see Figure 2.1). Finally, the intergenic region Rv3884c-83c was selected on the basis of its size (225 bp) and the fact that it lies directly downstream from region 1. The intergenic region sequences were selected with 100 base pairs on each side, to include the possibility of (or parts of) promoters lying within the adjacent gene sequences. The DNA sequences of the intergenic regions were obtained from the *M. tuberculosis* H37Rv genome sequence database, Tuberculist (http://genolist.pasteur.fr/Tuberculist/). Multiple sequence alignments of promoter regions were done with the program, ClustalW1.5 on the ClustalW WWW server (http://www.ebi.ac.uk/ClustalW/). The P_{AN} promoter sequence used as positive control in this study originates from *M. paratuberculosis* (Murry et al 1992) but is also conserved in H37Rv as shown in Figure 2.2. We did not clone any intragenic regions for analysis of potential promoters, as intragenic promoters occur infrequently in bacteria and none are known in *M. tuberculosis*.

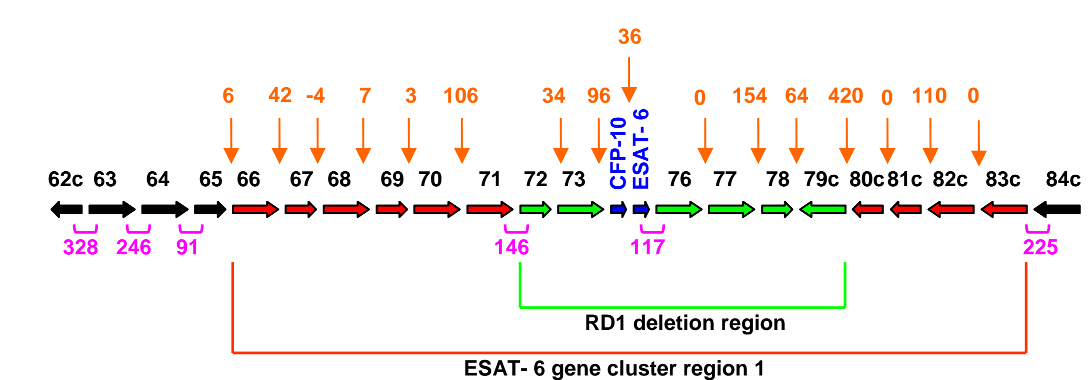


Figure 2.1 A schematic representation of the six intergenic regions (in purple) selected for the promoter study of the ESAT-6 region 1. The horizontal green arrows indicate the genes within the RD1 deletion region and the horizontal red arrows represent the genes present within the ESAT-6 gene cluster region 1. The horizontal black arrows represent the genes up- and downstream of region 1. The numbers in pink represent the amounts of basepairs of the six intergenic regions used in the promoter study. The numbers in orange (above vertical orange arrows) represent the amounts of basepairs of the rest of the intergenic regions of this gene region. The numbers above the arrows are the allocated gene number for each gene, preceded by Rv38. Therefore, this is an illustration of the gene region Rv3862c to Rv3884c.

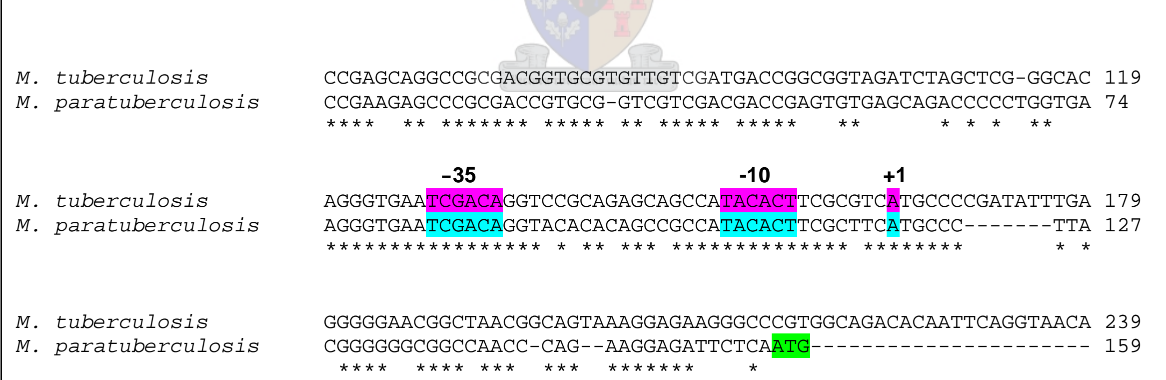


Figure 2.2 Multiple sequence alignment of the P_{AN} promoter sequence present in the genomes of *M. paratuberculosis* and *M. tuberculosis*, indicating the conservation of the promoter elements between these mycobacteria. The -35 and -10 promoter regions as well as the transcription start site +1 are highlighted in pink (*M. tuberculosis*) and in turquoise (*M. paratuberculosis*). The start codon of the ORF2 gene of *M. paratuberculosis* is highlighted in green.

2.2.2 Bacterial strains

Escherichia coli XL1-Blue was used as a host for all the subcloning steps of this study. As blue-white β-galactosidase promoter-screening should be done directly in a mycobacterial host (Timm et al 1994), the mycobacterial strain used to study promoter activity was *M. smegmatis* mc^2155 (Snapper et al 1990).

2.2.3 Media and culture conditions

E. coli was grown overnight on solid or liquid (with shaking) Luria-Bertani (LB) medium (Sambrook et al 1989). *M. smegmatis* was grown for 2 days at 37°C (with shaking, 200 rpm) in Middlebrook 7H9 broth (Difco) supplemented with filter-sterile ADC (albumin-dextrose-catalase - 0.5% BSA, 0.2% glucose, 0.015% catalase) and containing 0.05% Tween 80 (Sigma). Electro-competent cells of *E. coli* and *M. smegmatis* were prepared as described in Addendum A. Transformants of *E. coli* and *M. smegmatis* were selected by addition of Ampicillin (50 μg/ml, Roche) or Kanamycin (50 μg/ml, Roche). *E. coli* transformants were selected using blue/white screening on LB plates supplemented with X-gal (20 mg/ml, 0.032% final concentration) and IPTG (200 mg/ml, 0.0032% final concentration), unless noted otherwise. For determining β-galactosidase activity on solid media, *M. smegmatis* cells were grown on Middlebrook 7H11 agar supplemented with filter-sterile OADC (0.005% oleic acid, 0.5% BSA, 0.2 % glucose, 0.02% catalase, 0.085% NaCl) and containing 0.05% Tween 80 (Sigma). For *β*-galactosidase assays in liquid media, mycobacterial cultures (10 ml) were cultured in Kirchner's broth (0.3% Na_2HPO_4, 0.4% KH_2PO_4, 0.107% $MgSO_4.7H_2O$, 0.25% $C_6H_5Na_3O_7.2H_2O$, 2% glycerol, 0.5% apargine) instead of 7H9, due to a large amount of BSA protein in the 7H9 media, which would interfere with protein concentration determinations of the whole cell lysate. These cultures were grown to an optical density of 1 at 600 nm. The cells were centrifuged and the resulting pellets were each resuspended in 500 μl of a 1 x PBS (0.8% NaCl, 0.02% KCl, 0.02% KH_2PO_4, 0.09% Na_2HPO_4, pH 7.3) and sonicated at 4.5 setting in a Misonix cup sonicator on ice for a total of 5 minutes (15-second bursts with 30-second intervals). For RNA extraction, *M. smegmatis* was grown at 37°C for 2 days with shaking (200 rpm) in Middlebrook 7H9 broth (Difco) supplemented with filter-

sterile ADC (0.5% BSA, 0.2% glucose, 0.015% catalase) and containing 0.05% Tween 80 (Sigma).

2.2.4 Primers

All oligonucleotide primers used in this study were designed by making use of the Primer Premier program and are listed in Table 2.1. These primers were selected based on length and Tm.

Table 2.1 PCR and RT-PCR oligonucleotide primers used in this study

Name of primer	Primer sequence (from 5' to 3')	Length of primer	Tm [°C]	Length of product
Panf	cgcgacggtgcgtgttgt	18 bp	60	178 bp
Panr	catctgcggtgttacctgaa	18 bp	60	
386263f	ttactgtcatgtctacgttcctcca	25 bp	72	399 bp
386263r	atcctaaccgttgcacccttctact	25 bp	74	
386364f	tcgaagcactgatccgtcagatcaa	25 bp	74	272 bp
386364r	tttatgtgtttccttacgctcgccg	25 bp	74	
386465f	gacaagaaggaagacgaggaaggcg	25 bp	78	190 bp
386465r	tgggttgaaccctcctctacgttta	25 bp	74	
387172f	gttaagattatttcattgccgg	22 bp	60	149 bp
387172r	atttgcctacttcttccctcc	21 bp	62	
387576f	tcactgggatgttcgcatagggc	23 bp	72	134 bp
387576r	tgacaacctctcagagtgcgctcaa	25 bp	76	
388483f	gtttggtagcagccgcatat	20 bp	60	231 bp
388483r	ctcccagaacactccattcgtt	22 bp	66	

Pectora roborant cultus recti

2.2.5 DNA manipulations for promoter cloning

Standard protocols were followed for all DNA manipulations (Sambrook et al 1989). DNA extractions on small scale were carried out using the Wizard SV Gel & PCR Clean-up System (Promega) as well as the Wizard SV Minipreps kit (Promega). DNA extractions on a large scale

were performed using the Nucleobond AX kit (Macherey-Nagel). Enzymes used for DNA digestion and ligation were purchased from either Roche Molecular Biochemicals or Promega. These enzymes were used as recommended by the manufacturers. Restriction endonuclease-digested DNA was isolated from agarose gels and purified using the Wizard SV Gel & PCR Clean-up System. Oligonucleotide pairs (see Table 2.1) were used as primers to PCR amplify the P_{AN} promoter sequence (positive control) as well as the intergenic region sequences (see Figure 2.1) from *M. tuberculosis* H37Rv genomic DNA (received as part of NIH, NIAID Contract No. HHSN266200400091C, CSU). The PCR amplified P_{AN} and intergenic region sequences were cloned into pGem-T Easy (Promega) giving rise to the following clones: (pGemTPan, pGemT6263, pGemT6364, pGemT6465, pGemT7172, pGemT7576 and pGemT8483). The inserts were then excised by digestion with *Eco*RI and cloned into the *Eco*RI site of pBluescript II KS$^+$ (Stratagene) resulting in pBluePan, pBlue6263, pBlue6364, pBlue6465, pBlue7172, pBlue7576 and pBlue8483, respectively. Inserts were selected for correct orientation by PCR amplification with the pBluescript II KS$^+$ T7 primer and the inserts' reverse primers (Table 2.1) before they were excised by digestion with *Bam*HI and *Kpn*I restriction endonucleases. The *Bam*HI and *Kpn*I fragments were finally cloned upstream of a promoter-less *lacZ* gene into the corresponding sites of the mycobacterial-*E.coli* shuttle promoter-probe vector pJem15 (Figure 2.3, Timm et al 1994). The resulting constructs were designated pJemPan, pJem6263, pJem6364, pJem6465, pJem7172, pJem7576 and pJem8483. During the cloning of the P_{AN} promoter sequence and the intergenic regions, all pGem-T Easy, pBluescript II KS$^+$ and pJem15 constructs were verified by DNA sequencing. For all cloning steps, the constructs were electroporated into *E. coli* XL1Blue cells and transformants were selected on LB plates containing Ampicillin (pGem-T Easy and pBluescript II KS$^+$) and Kanamycin (pJem15). For determining β-galactosidase activity, the final pJem15 constructs were electroporated into *M. smegmatis* and the transformants were selected on Middlebrook 7H11 agar plates containing Kanamycin and X-gal. All essential constructs created during the cloning of the intergenic regions were verified by DNA sequencing.

Pectora roborant cultus recti

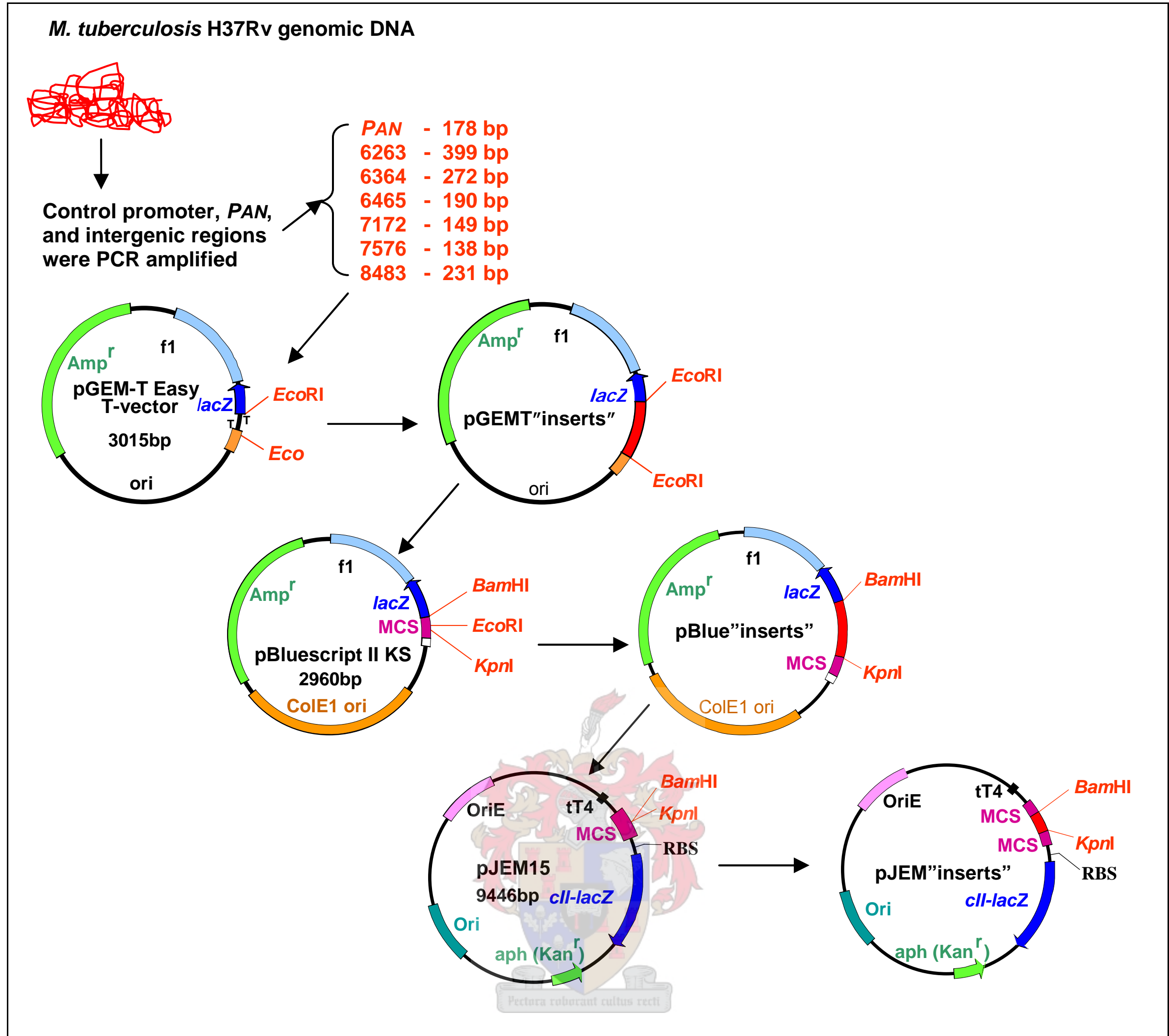


Figure 2.3 A schematic representation of the subcloning of the P_{AN} promoter and intergenic region sequences (after PCR amplification from *M. tuberculosis* H37Rv genomic DNA) through pGem-T Easy and pBluescript II KS^{+} into the promoter probe vector pJem15.

2.2.6 β-Galactosidase assays

For the detection of *β*-galactosidase activity on solid media *M. smegmatis* transformants were plated onto Middlebrook 7H11 media containing 0.001% X-gal. *β*-galactosidase assays were performed in liquid media as described by Pardee et al (1959) and Timm et al (1994). Briefly, sonicated extracts (see section 2.2.3) of *M. smegmatis* transformed with the different pJem15 constructs were assayed. Protein concentrations in sonicated extracts were calculated by means of the BioRad protein determination assay (as stipulated by the manufacturer

instructions). *β*-galactosidase assays entail the hydrolysis of the *β*-galactosidase substrate, *o*-nitrophenyl-β-D-galactoside (ONPG), where one unit of *β*-galactosidase is defined as producing 1 μMole *o*-nitrophenol per minute from ONPG at 28°C, pH 7.0. The reactions, each consisting of equal amounts of protein (10μg); 100 x Mg solution (10% 1M $MgCl_2$. 32% β-mercaptoethanol); 1 x ONPG and 0.1M Phosphate buffer (82% 0.2M $NaH_2PO_4.2H_2O$, 18% 0.2M Na_2HPO_4), were stopped by addition of 0.5ml 1M Na_2CO_3 after a 30 minute incubation at 28°C. β-galactosidase activity was spectrophotometrically calculated by measurement of *o*-nitrophenol at an optical density (OD) of 420 nm where 1μMole/ml *o*-nitrophenol has an OD_{420} of 0.0075. β-galactosidase assays were performed in triplicate. β-galactosidase units were calculated as follows: 1Unit = 200 x OD_{420}/mg of protein/minute, as was described by Pardee et al (1959), Timm et al (1994) and Berthet et al (1998).

2.2.7 RNA preparation

All steps performed during RNA isolation and purification were under strict RNase free conditions. *M. smegmatis* culture (50 ml) was grown to an OD_{600} of above 0.8. This culture was then centrifuged at 4000rpm for 10 minutes. After the removal of the supernatant, Trizol (GibcoBRL) was added to the pellet at a volume of 0.85 ml per 10 ml original culture. Cells were resuspended by pipetting and transferred to Blue FastPrep tubes containing silicone beads (Bio101). The cells in the FastPrep tubes were ribolyzed twice in a FastPrep bead-beater (Bio101) at speed setting 6.5 for 45 seconds and put on ice in-between steps. The cell mixture was incubated at room temperature for 5 minutes, after which 0.2 ml chloroform was added per 0.75 ml Trizol, and mixed by vortexing. The mixture was incubated at room temperature for 5 minutes and centrifuged at 4ºC for 15 minutes. The supernatant was removed from the beads/cell debris, transferred to a 1.5 ml microfuge tube and 0.5 ml isopropanol was added per 0.75 Trizol. After thorough mixing, the tube was incubated at room temperature for 10 minutes and centrifuged at 4ºC at 13 000 rpm for 10 minutes. The supernatant was removed and 250 μl 100% ethanol was added and mixed by pipetting. The RNA was then purified using the Qiagen RNeasy mini-kit, as described by the manufacturer. Briefly, the RNA/DNA/ethanol mixture was

added to a RNeasy mini column in a 2 ml collection tube and centrifuged at 13 000 rpm for 15 seconds. RW1 buffer (350 µl) was added to the RNeasy column and centrifuged at 13 000 rpm for 15 seconds. A DNaseI incubation mix (consisting of 10µl DNaseI stock solution added to 70 µl of RDD buffer) was then added to the column and incubated at room temperature for 15 minutes. RW1 buffer (350 µl) was added and centrifuged for 15 seconds. The column was transferred to a new collection tube and 500 µl RPE buffer added and centrifuged for 15 seconds. RW1 (500 µl) was added again and centrifuged for 2 minutes at 13 000 rpm. The RNeasy column was then placed into a 1.5 ml collection tube (also supplied by the kit). The RNA was first eluted in 30 µl of RNase-free H_2O, centrifuged for 1 minute at 13 000 rpm and then in 20 µl of RNase-free H_2O (again centrifuged at 13 000 rpm for 1 minute) for a total of 50 µl of RNA. As a final precaution to remove all possible DNA from the isolated RNA, the eluted 50 µl of RNA was put through an additional DNase process (performed as described by the manufacturer, Qiagen). Briefly, the 50 µl RNA sample was made up to 100 µl with RNase-free H_2O. RLT buffer (350 µl with added β-mercaptoethanol at a volume of 10µl per 1 ml buffer RLT) was added and mixed thoroughly. Two hundred and fifty microlitres of 100% ethanol was then added and mixed thoroughly by pipetting. The 700 µl mixture was added to an RNeasy mini column, placed in a 2 ml collection tube and centrifuged for 15 seconds at 10 000 rpm. The flow-through and collection tube were discarded and the column placed into a new 2 ml collection tube. RPE buffer (500 µl) was added to the mini column and again centrifuged for 15 seconds at 10 000 rpm and the flow-through discarded. RPE buffer (500 µl) was again added to the column and centrifuged for 2 minutes at the same speed to dry the silica-gel membrane. For the elution of RNA, the RNeasy column was transferred to a new 1.5 ml collection tube (supplied) and 30 to 50 µl of RNase-free water was added directly to the column. After centrifugation for 1 minute at 10 000 rpm, the RNA concentration was calculated by means of the Nanodrop Spectrophotometer. RNA was confirmed to be free of any DNA contamination by PCR analysis and stored at -20°C.

Pectora roborant cultus recti

2.2.8 RT-PCR analysis

RT-PCR was performed by a combination of the Roche Titan One-Tube RT-PCR System kit and the Promega HotstarTaq PCR System, as described by the manufacturers. For the Reverse Transcriptase (RT) part of the process, two separate master mixes were prepared: The first master mix contained the RNA template, dNTP's (200 μM final concentration), reverse primers of the intergenic regions (50 pmol each), DTT (10 mM final concentration), RNaseI (RNase inhibitor; 1 U/μl) and water to a final volume of 25 μl. The second master mix contained the 5X RT-PCR buffer, the Reverse Transcriptase enzyme mix, and water to a final volume of 25 μl. These two mastermixes were mixed together immediately prior to cycling. Reverse transcription was performed at 50ºC for 45 minutes. Thereafter, PCR was performed on the cDNA for 35 cycles by using the HotStarTaq PCR amplification system (Qiagen) under the prescribed manufacturer's conditions in a Perkin Elmer GeneAmp 2400 (see Table 2.2). Results of RT-PCR reactions were visualized on a 1 % TAE agarose gel (0.5g of agarose MS-8, Whitehead Scientific, in 50 ml TAE buffer).

Table 2.2 RT-PCR cycling parameters

Number of Cycles	Reaction, temperature and time duration
1 cycle	reverse transcription at 50°C for 30 minutes
1 cycle	template denaturing at 94°C for 2 minutes
1 cycle	HotstarTaq polymerase activation at 95°C for 15 minutes
35 cycles	denaturing at 94°C for 30 seconds annealing at *x*°C* for 30 seconds elongation at 72°C for 30 seconds
1 cycle	elongation at 72°C for 7 minutes

**x*°C is the optimum annealing temperature for the specific primer pair – see Table 2.1

2.3 Results

2.3.1 Promoter analysis

The aim of the promoter analysis was to identify a promoter(s) that drive(s) the expression of genes contained within the ESAT-6 gene cluster region 1. Therefore, six intergenic regions (see Figure 2.1) as well as the P_{AN} promoter sequence (positive control, Murray et al 1992) were subcloned through pGem-T Easy and pBluescriptII KS^{+} into pJem15 (promoter-probe vector, previously described by Timm et al (1994; see Figure 2.3). The P_{AN} promoter sequence and intergenic regions were subcloned into pJem15 (promoter-probe vector) by transformation of *E. coli* and selected on LB–Kanamycin plates. The constructs were subsequently electroporated into *M. smegmatis* and plated onto Middlebrook 7H11-Kanamycin-X-gal plates.

White colonies were found on the pJem15 (negative control), pJem6364 and pJem7576 plates; light blue colonies on the pJem7172 plates and dark blue colonies on the pJemPan (positive control), pJem6263, pJem6465 and pJem8483 plates (Figure 2.4). To quantitatively confirm these results, the β-galactosidase activities were determined. This was done by means of β-galactosidase assays which make use of a spectrophotometric determination of ONPG-hydrolysis (Pardee et al 1959, Timm et al 1994). The results of the β-galactosidase assays confirmed what was observed on the X-gal plates with region Rv6263 showing very strong promoter activity and Rv6465 and Rv8483 both containing promoters of medium strength (Figure 2.5). Rv7172, Rv6364 and Rv7576 showed very little to no activity, comparable to the pJem15 negative control.

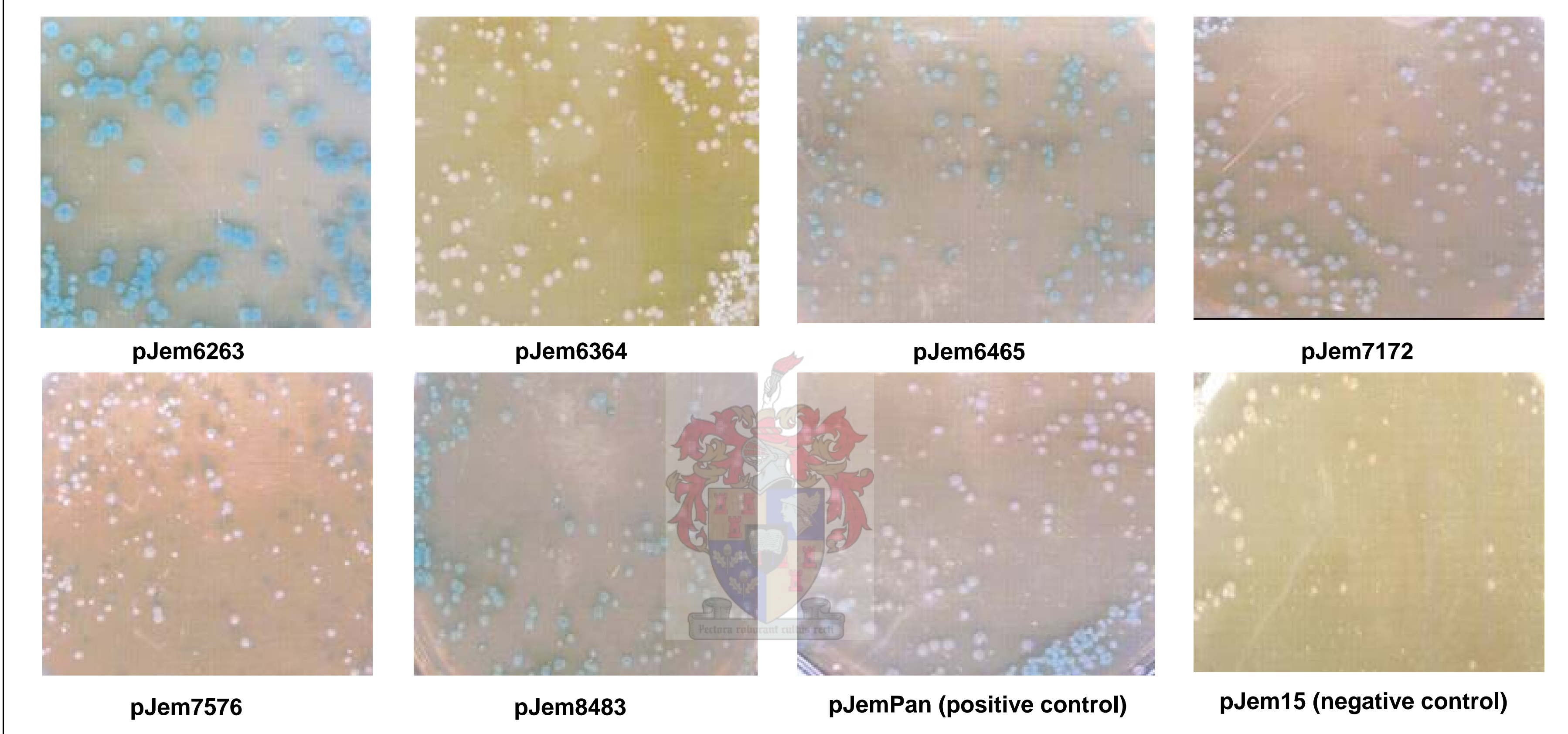


Figure 2.4 X-gal plates showing the presence of promoter activity indicated by the intensity of the blue colour of the transformed *M. smegmatis* colonies. pJemPan (containing the P_{AN} promoter sequence) serves as the positive control and therefore blue colonies are seen on the plate. pJem15 (containing no insert) serves as the negative control and thus contains white colonies. Dark blue colonies are seen with pJem6263, pJem6465 and pJem8483, indicating the presence of strong promoters.

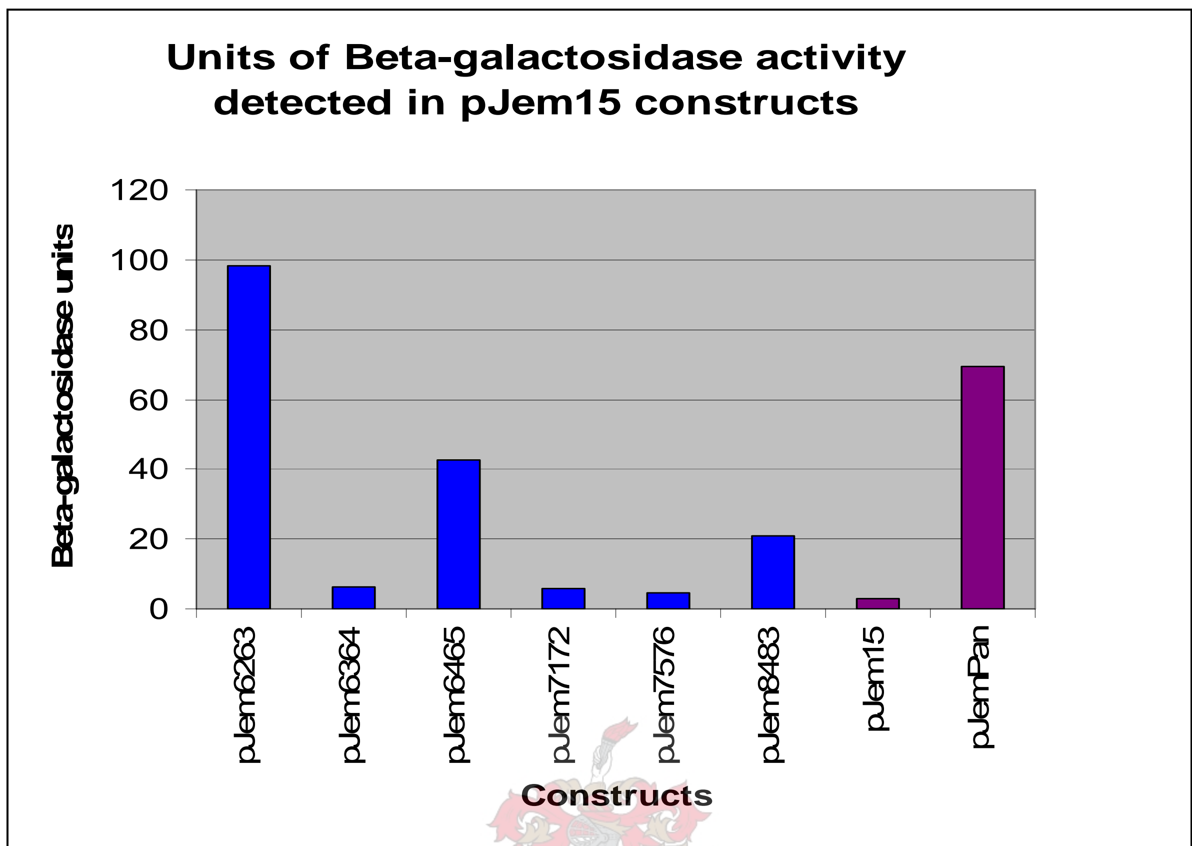


Figure 2.5 **β-Galactosidase activities of *M. smegmatis* clones containing ESAT-6 gene cluster region 1 intergenic regions.** The results are the means of three independent experiments. pJem15 represents the negative control and pJemPan the positive control.

2.3.2. Operon analysis

The results of the RT-PCR analysis performed with primers spanning the intergenic regions of the ESAT-6 gene cluster region 1 are presented in Figures 2.6 and 2.7. This shows amplification of cDNA from intergenic regions Rv3864-Rv3865 (designated IG6465), Rv3871-Rv3872 (designated IG7172), and Rv3875-Rv3876 (designated IG7576), indicating the presence of RNA species spanning these three intergenic regions. Intergenic regions Rv3862-Rv3863, Rv3863-Rv3864 and Rv3884-Rv3883c showed no products, suggesting the absence of RNA species spanning these intergenic regions. These results were expected, because region 1 starts at Rv3866 on the sense strand, and at Rv3883c on the reverse strand. Strong promoter elements were also found within intergenic regions Rv3862-Rv3863 and Rv3884-Rv3883c. The

amplification products were isolated and sequenced to confirm the identity of the respective intergenic regions. In all three cases the presence of an amplicon corresponding to the intergenic region sequence was confirmed (results not shown). It is clear from these results that the ESAT-6 gene cluster region is expressed as more than one polycistronic RNA (Figure 2.7).

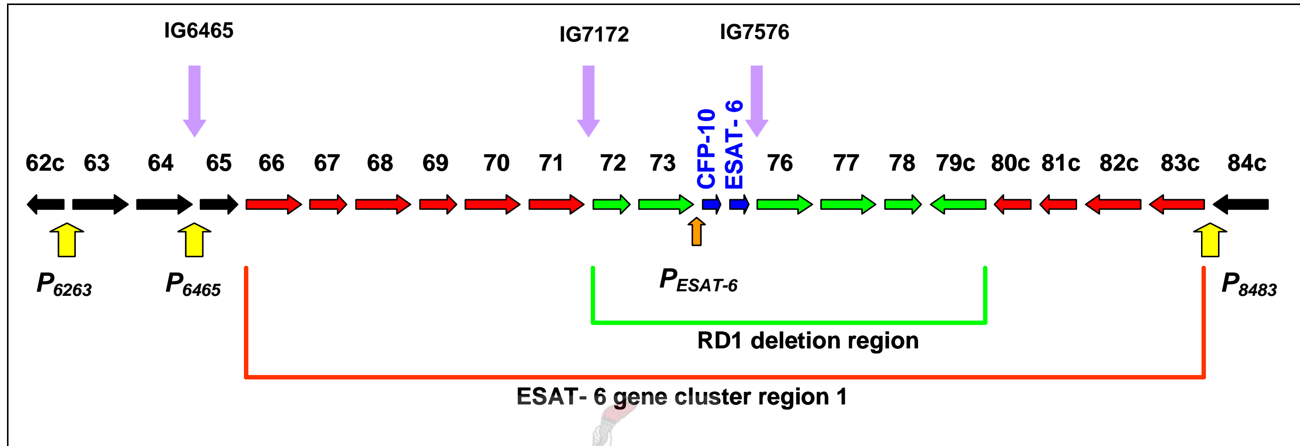


Figure 2.6 Promoter- and operon analysis. A schematic representation of the promoters (P_{6263}, P_{6465} and P_{8483}) identified in the ESAT-6 gene cluster region 1 promoter analysis (see section 2.3.1) as well as amplified RNA species (IG6465, IG7172 and IG7576). Promoters are indicated by yellow arrows. The orange arrow indicates the *esxB/esxA* operon (P_{ESAT-6}) promoter identified by Berthet et al (1998). Light purple arrows indicate the positions of the three operon amplicons IG6465, IG7172 and IG7576 (see section 2.3.2). The green arrows indicate the genes within the RD1 deletion region and the red arrows represent the genes of the ESAT-6 gene cluster region 1. The black arrows represent the genes up- and downstream of region 1.

Pectora roborant cultus recti

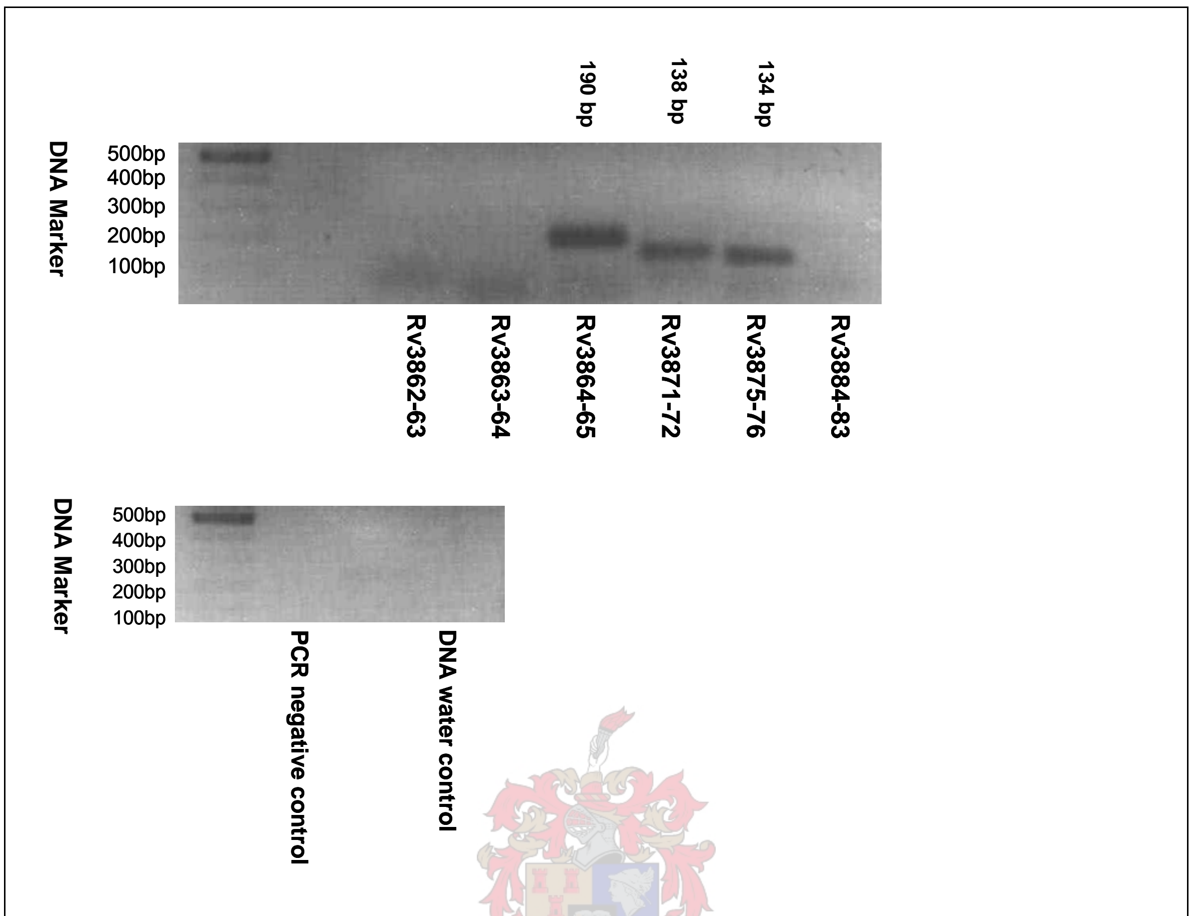


Figure 2.7 Operon analysis through identification of polycistronic RNA. RT-PCR results of intergenic regions of the ESAT-6 gene cluster region 1 indicating the presence of multiple polycistronic species in this region. RT-PCR was done on *M. tuberculosis* RNA with primers spanning the intergenic regions of the ESAT-6 gene cluster region 1. Negative controls for DNA contamination, and DNA PCR amplification were included and amplicons were separated on a 1.5% agarose gel.

2.3.3 In silico promoter analyses

In silico promoter analyses were conducted in order to identify potential promoter subelements in the intergenic regions (P_{6263}, P_{6465} and P_{8483}) showing promoter activity. The results of these analyses are illustrated in Figure 2.8. The subelements of the previously identified ESAT-6 gene cluster region promoters, P_{AN} (Murray et al 1992), P_{ESREG3} (Gey van Pittius PhD Thesis 2002) and P_{ESAT-6} (Berthet et al 1998) are also presented in this figure.

P_{AN} promoter (Rv3892c-Rv3891c): 178 bp

cgcgacggtgcgtgttgtcgatgaccggcggtagatctagctcgggcacagggtgaatcgacaggtccgca
gagcagccatacacttcgcgtcatgccccgatatttgaggggaacggctaacggcagtaaaggagaagggc
ccgtggcagacacaattcaggtaacaccgcagatg

P_{ESREG3} promoter (Rv0281-Rv0282): 281 bp

gttcgcccgcaacaccctggtcagtgccgcccgcgtctgacggcgcaccgttcgcgctgccggcaccccgg
gctccataatgaaaatcatgttcagtaagctacactctgcatatcgggctaccaacgaaatggagtatcggt
catgatcttgccagccgtgcctaaaagcttggccgcagggccgagtcgattggtcgcggtcgcctcgacagt
tagcttatgcaatgctaacttcggggcaaagttcaggcggatcggccgatggcgggcgtaggtgaa

P_{ESAT-6} promoter (Rv3873-Rv3874): 241 bp

gcaggagcgtgaagaagacgacgaggacgactgggacgaagaggacgactggtgagctcccgtaatgacaa
cagacttcccggccacccgggccggaagacttgccaacatttggcgaggaaggtaaagagagaaagtagtc
cagcatggcagagatgaagaccgatgccgctaccctcgcgcaggaggcaggtaatttcgagcggatctccgg
cgacctgaaaacccagatcgaccagg

P_{6263} promoter (Rv3862c-Rv3863): 399bp

ttactgtcatgtctacgttcctccagaaagcgttgcaggttgtagcctctgccgcgaaagcgtatcgcatt
aaccatagcgatgcaacagtttcctcctctgcctgcctagcggtgctgcggctccggttcggcgagctccga
gctctagtgcgcgcaccgccgagtaccagggcatagatcctgttaatcagctgtgtatctggcctcgccggc
gcgtatccgaccccttcgggcagatcttccaggaaaagtgttctgacatgcgacagttcaggtgtgaagtga
actgtagcggcagttcggtttggctaggaaactatttccatagcggccgtcgcgtcgctagatccaaaatg
tagcgaagtcatagcagtagaagggtgcaacggttaggat

P_{6465} promoter (Rv3864-Rv3865): 190bp

Gacaagaaggaagacgaggaaggcgtggccgaggcggagcgtgcacccatcgacgctggcaccgcggccag
ccaacgggggcaggaggggaccgtcctttgatcggacaccgagtcgccagcaggtctgtgccatagcgagtc
gaagccatagcgagtagaaagttaaacgtagaggagggttcaaccca

P_{8483} promoter (Rv3884c-Rv3883c): 231bp

gtttggtagcagccgcatatcggctagtgcggtagcaaaaccgttgagtcccggcgtggtccaggtacagc
gcctacgcgccctggtcggcgcggcggaccagcccaggcgtgacagcgttgtttgcccagccgttacgtcta
aaatgcacacaggtccgtcaagtggcccaaggtagcaacgcagctcaatgaatcgcaatgaatctcaacga
atggagtgttctgggag

Figure 2.8 Previously identified promoters, P_{AN} (Murray et al 1992), **P_{ESREG3}** (Gey van Pittius PhD Thesis 2002) **and $P_{ESAT\text{-}6}$** (Berthet et al 1998) **as well as the promoter elements identified in this study, P_{6263}, P_{6465} and P_{8483}.** Primers used for cloning are blocked. C-terminal sequences of upstream genes are shown in purple letters with stopcodon highlighted in purple. Start codons are highlighted in blue and N-terminal sequence of genes are also shown in blue letters. Known or predicted promoter elements are highlighted as follows: -35 in light blue, -10 in green, transcriptional start sites in grey and ribosome binding sites in yellow.

The hypothesized promoter elements were predicted in accordance with the *E. coli* consensus promoter sequence. Figure 2.8 thus indicates the putative -35 and -10 motifs as well as the putative transcriptional start sites (+1) and ribosomal binding sites. No homology could be seen between the putative promoter elements of this study (P_{6263}, P_{6465} and P_{8483}) and previously identified promoters, P_{AN} (Murray et al 1992) and P_{ESREG3} (Gey van Pittius PhD Thesis 2002).

Pectora roborant cultus recti

Multiple sequence alignments of the predicted promoter elements (P_{6263}, P_{6465} and P_{8483}) between *M. tuberculosis*, *M. marinum* and *M. smegmatis* are illustrated in Figure 2.9. The sequence of *M. marinum* was chosen for the alignment as it is a slow-growing, virulent mycobacterium closely-related to *M. tuberculosis*. The fast-growing, avirulent *M. smegmatis* was chosen as it is one of the most distantly related mycobacteria to *M. tuberculosis*. The genomes of both these species have been sequenced (*M. marinum* - http://www.sanger.ac.uk/Projects/M_marinum/ and *M. smegmatis* - http://cmr.tigr.org/tigr-scripts/CMR/GenomePage.cgi?database=gms).

Rv3862c-Rv3863

```
M.tuberculosis     TAACCATAGCGATGCAACAGTTTCCTCCTCTGCCTGCCTAGCGGTGCTGCGGCTCCGGTT 60
M.marinum          TAACCGTCGCAGTGCAACAGTTCCCGATGCAGGGCCGCACACCCTCCTGCTACAAC--TT 58
M.smegatis         ---CCGTAGCGATGCAACAGTTT--GATTCGAAGCGGCAGCGAGTGCTGCTACGGAATTT 55
                      ** * **  **********       *       *      * ****  *     **

M.tuberculosis     CGGCGAGCTCCGAGCTCTAGTGCGCGCACCGCCGAGTACCAGGGCATAGATCCTGTTAAT 120
M.marinum          CGCTGCGCGATGGGCCATGATAAGCATCTAACC--------TGGCATAACCAGCGGTAAT 110
M.smegatis         TG------------CGTTGGTAA-----CAAC-----------ACATGACCACCTGCGAA 87
                    *            *  *  *          *            ***           *

M.tuberculosis     CAGCTGTGTATC---TGGCCTCGCCGGCGCGTATCCGACCCCTTCGGGCAGATCTTCCAG 177
M.marinum          TAGCTGTCAATCGGCTGGGTGACACGGCGCTCCGCCAACCCCGACGCACAC-CCTTCCGG 169
M.smegatis         AAGCCAGCACTC-----GATAAAA--GCGCACGTCAGAGACCTTCACATAG--CGCACGG 138
                    ***      **      *        ****    *  *  **  *    *   *   * *

M.tuberculosis     G--AAAAGTGTTCTGACATGCGACAGTTCAGGTGTGAAGTGAACTGTAGCGGCAGTTCGG 235
M.marinum          ACAGAAACTGTTCCACTGCGCAATGCTTCGGG-GTAGGATCAAACGTAGCGGTGGCCCGG 228
M.smegatis         A--GGTTCCGTAGCGTCTC-CGTGGTTTCGCACGCGACAATTA-TGTAGCACTGGTGCGG 194
                            **         *     ***    *        *  *****    *  ***

M.tuberculosis     TTTGGCTAGGAAACTATTTCCATAGCGGGCCGTCG--CGTCGCTAGATCCAAAATGTAGC 293
M.marinum          TTTAACCAGACAACGGTCGCTTCGGAACCCCGCTGGATGTCGTTCAA---GAAACGTAGC 285
M.smegatis         TTTG-CTGGAACGTGACCAAAACCGAGGCGTAGCG------------------------- 228
                   ***  *  *               *         *

M.tuberculosis     GAAGTCATAGCAGTAGAAGGGTGCAACGGTTAGG 327
M.marinum          GAAGGCATACCAATGGAAGGGTGCAACGGTTAGG 319
M.smegatis         ----------------------------------
```

Rv3864-Rv3865

```
M.tuberculosis     --------TCGGACACCGAGTCGCCAGCAGGTCTGTGCCATAGCGAGTCGAAGCC----A 48
M.marinum          TCCGGGCAACCGGCGGCGGCTCACTGGCCTACCAAAGCCATAGCGAGCCAAACCC----A 56
M.smegatis         -------GCTACGCACCGTGGCGGCGGCTACATTAGACCACGGTATTCCCGGGTCTGCGG 53
                             *  **   *    **         ***  *     *     *

M.tuberculosis     TAGCGAGTAGAAAGTTAAACGTAGAGGAGGGTTCAACCC------------- 87
M.marinum          TAGCGAGTAGAAAGTTAGACGTAGAGGAAAGGTCTACCCCC----------- 97
M.smegatis         CCGCGCGCGGTGCGTGAGGATCCAAGGGAATGTCGGGGAATCGAGGGGATTC 105
                     *** *  *   ** *       ***     **
```

Rv3884c-Rv3883c

```
M.tuberculosis      -----------------------------------------------CAGCCGCATATCG 13
M.marinum           CGGGTCGCGCCGGTTGGCGAGGGTAAGCCCGGGCGCGTGGAATGGCATGTGTTCGTGCAG 60
M.smegatis          ------------------------------------------------AATTTCTGATCG 12
                                                                      *     *

M.tuberculosis      GCTAGTGCGGTAGCAAAACCGTTGAGTCCCGGCGTG------GTCCAGGTACAGCGCCTA 67
M.marinum           CTTTGAGCGACGCCCGGGCCGGCGCGGCACGGCGTGCG---GAGCCAGTTGCCGACGCCG 117
M.smegatis          CCTGCGGCCGGGATTTTCCGGGCAGGTAAAAAAGCGTCATCGGCGTGATCGTGATGCCCA 72
                      *   **          * *    *       * *                     *

M.tuberculosis      CGCG--CCCTGGTCGGCGCGG-CGGACCAGC-CCAGGCGTGACAGCG-TTGTTTGCCCAG 122
M.marinum           CTTGGTCACAGTTCAGGATGGACAGGCTGGCACCGGGTTGAATAGCG-TCGTTTACGCGC 176
M.smegatis          TAGACTTCCAACGTGCGAAGTGCCCGCCGGCTCGGCATGAGAAGATGCTTGTCTGCGCAG 132
                            *          *  *   *  ** *        *    * * ** * * *

M.tuberculosis      ---CCGTTACGTCTAAAATGCACACAGGTCCGTCA--AGTGGCCCAAGGTAGCAACGCAG 177
M.marinum           ---CCGTTACGTCTAAAATGCAGGCAGGACTGACACGAGCGTGCCAGAGCTTCACTGCGG 233
M.smegatis          ACACCGATCTGTTTGCTGGGGGGCTCCGGCCCGCGGCGGTCGGATCAATTCTTGTTGGGG 192
                       *** *  ** *     *       * *   *    *                 *  *

M.tuberculosis      CTCAATGAATCGCAATGAATCTCAACGAATGGAGTGTTCTGGGA 221
M.marinum           CTAGA--GAACGGAGTGATTCCGGG------------------- 256
M.smegatis          GAGTGTGAGGCGTGCAGCGCGTAGCCGTC--------------- 221
                              **    *
```

Figure 2.9 Multiple sequence alignment of the ESAT-6 gene cluster region 1 putative P_{6263}, P_{6465} and P_{8483} promoter regions from different mycobacteria. Known or predicted promoter elements are highlighted as follows: -35 in light blue, -10 in green, transcriptional start sites in grey and ribosome binding sites in yellow.

As seen in Figure 2.9, no large conserved regions can be seen when upstream (to genes Rv3863, Rv3864 and Rv3883c) intergenic regions of *M. tuberculosis*, *M. marinum* and *M. smegmatis* were aligned. This indicates that the predicted promoter elements (P_{6263}, P_{6465} and P_{8483}) are not conserved in species outside of the *M. tuberculosis* complex. This may indicate *M. tuberculosis* complex-specific regulation of these genes. Predicted promoter elements are indicated as follows: -35 motifs in light blue, -10 motifs in green, transcriptional start sites in grey and ribosome binding sites in yellow. The hypothesized promoter elements were predicted in accordance with the *E. coli* consensus promoter sequence.

2.4 Discussion and Conclusion

The functions and involvement in pathogenicity of most of the mycobacterial antigens identified to date remains largely unknown. Elucidation of the expression, regulation and intracellular functions of these antigens will contribute to the understanding of the processes involved in causing disease. Therefore, a comprehensive analysis needs to be executed in order to decipher the regulation of expression of the genes encoding these antigens. The ESAT-6 gene cluster region 1 plays an important role in the study of the pathogenicity of *M. tuberculosis*, as it contains the 9.5 kb region of difference, RD1 (Flint et al 2004, Hsu et al 2003, Pym et al 2003), that contains genes encoding proteins that form part of a novel secretion system as well as genes that encode for important antigens, CFP-10 (Dillon et al 2000, Colangeli et al 2000, Berthet et al 1998) and ESAT-6 (Gey van Pittius et al 2001, Andersen et al 1995 and Sorensen et al 1995) involved in the pathogenicity of the organism. During this study the regulation of the ESAT-6 gene cluster region 1, encoding the potent T-cell antigens ESAT-6 and CFP-10 and their secretion system, was investigated by analysis of its promoters and polycistronic nature.

Three promoters have been previously identified in the ESAT-6 gene clusters, the P_{AN} promoter situated in the ESAT-6 gene cluster region 2 (Murray et al 1992), the *esat-6* promoter from region 1 (Berthet et al 1998) and the P_{ESREG3} promoter involved in the expression of the ESAT-6 gene cluster region 3 (Gey van Pittius PhD Thesis 2002). To further elucidate the regulation of expression of the rest of the ESAT-6 gene clusters, we identified promoters involved in the expression of ESAT-6 gene cluster region 1.

The results of the promoter-probe analysis conducted during this study showed the existence, and characterized the activity of 3 promoters up- and downstream of region 1 (Figure 2.6). The 3 promoters, P_{6263}, P_{6465} and P_{8483}, were identified by β-galactosidase assays. These promoters are expressed strongly in *M. smegmatis* (Figure 2.5).

P_{6263} probably drives the expression of Rv3863 and Rv3864, whereas P_{6465} drives the expression of the gene region Rv3865 to Rv3878 (See Figure 2.6). P_{8483} most likely promotes the expression of the gene region Rv3883c to Rv3879c on the opposite strand (See Figure 2.6). To confirm that P_{6263} drives only the expression of Rv3863 and Rv3864 (and not the rest of the gene region up to Rv3878), the presence of a terminator between Rv3864 and Rv3865 has to be confirmed (see Figure 2.6). However, no terminator could be identified (data not shown). It is notoriously difficult to identify terminators in *M. tuberculosis*, and it is hypothesized that the absence of identifiable terminator sequences may be due to the G+C richness of the genome of *M. tuberculosis* (65.9% G+C, Cole et al 1998).

The results of the operon analysis conducted during this study indicated that the ESAT-6 gene cluster region 1 is expressed as more than one polycistronic RNA species (see Figure 2.7). *In silico* analyses also highlighted hypothesized promoter subelements with high probability. These results shed light on the regulatory mechanisms of the ESAT-6 antigen family of *M. tuberculosis*.

Pectora roborant cultus recti

CHAPTER THREE

ISOLATION OF THE ESAT-6 GENE CLUSTER REGIONS 2, 4 AND 5

3.1 Introduction

The genes encoding the CFP-10 and ESAT-6 proteins lie within a cluster of 12 other genes which encompasses the RD1 deletion region. These genes have been duplicated five times in the genome of *M. tuberculosis* and the duplicated gene clusters have been previously described as the ESAT-6 loci in an analysis of the *M. tuberculosis* proteome by Tekaia et al (1999). These gene clusters are designated region 1 (Rv3866-Rv3883c), region 2 (Rv3884c-Rv3895c), region 3 (Rv0282-Rv0292), region 4 (Rv3444c-Rv3450c) and region 5 (Rv1782-Rv1798) according to the numbering of the mycosin genes situated in these regions (Brown et al 1999). An examination of the genes in the clusters reveals that each of the clusters also contains (in addition to a copy of *esxA* and *esxB*, encoding ESAT-6 and CFP-10, respectively), genes encoding putative ABC transporters (integral inner-membrane proteins), ATP-binding proteins, subtilisin-like membrane-anchored cell wall-associated serine proteases (the mycosins – Brown et al 2000), and other N-terminal membrane-associated proteins (Tekaia et al 1999).

All the mycobacterial genomes that have currently been sequenced contain multiple copies of the ESAT-6 gene cluster regions. One of these regions is also contained in bacteria other than mycobacteria. A detailed description of the phylogeny and distribution of the ESAT-6 gene cluster regions is found in Gey van Pittius et al (2001).

The ESAT-6 gene cluster regions are important chromosomal domains and insight into the mechanisms of regulation and secretion of potent T-cell antigens of *M. tuberculosis* is required. In this study, we have isolated and cloned ESAT-6 gene cluster regions 2, 4 and 5 for use in subsequent secretion and antigen regulation analyses.

3.2 Materials and Methods

3.2.1 Bacterial strains

E.coli XL1Blue was used as the host for the propagation of the *M. tuberculosis* H37Rv cosmid library.

3.2.2 Media and culture conditions

E. coli was grown on solid Luria-Bertani (LB) medium (Sambrook et al 1989). Ampicillin (50µg/ml, Roche) and/or Hygromycin (50µg/ml, Roche) was added to the LB medium when antibiotic selection was required.

3.2.3 Cosmid library

The cosmid library of *M. tuberculosis* (Bange et al 1999) was a kind gift from F.-C. Bange (Medizinische Hochschule, Hannover, Germany). It was constructed by cloning the chromosomal DNA *Sau*3A fragments of *M. tuberculosis* H37Rv (approximately 40Kb) into the *Xba*I and *Bcl*I sites of the pYUB412.2 cosmid, which contains two selectable marker genes allowing for selection on either Ampicillin- or Hygromycin-containing media. This cosmid vector contains an integrase gene and integrates stably into the *attP* sites of mycobacteria only (Bange et al 1989; Figure 3.1(A))

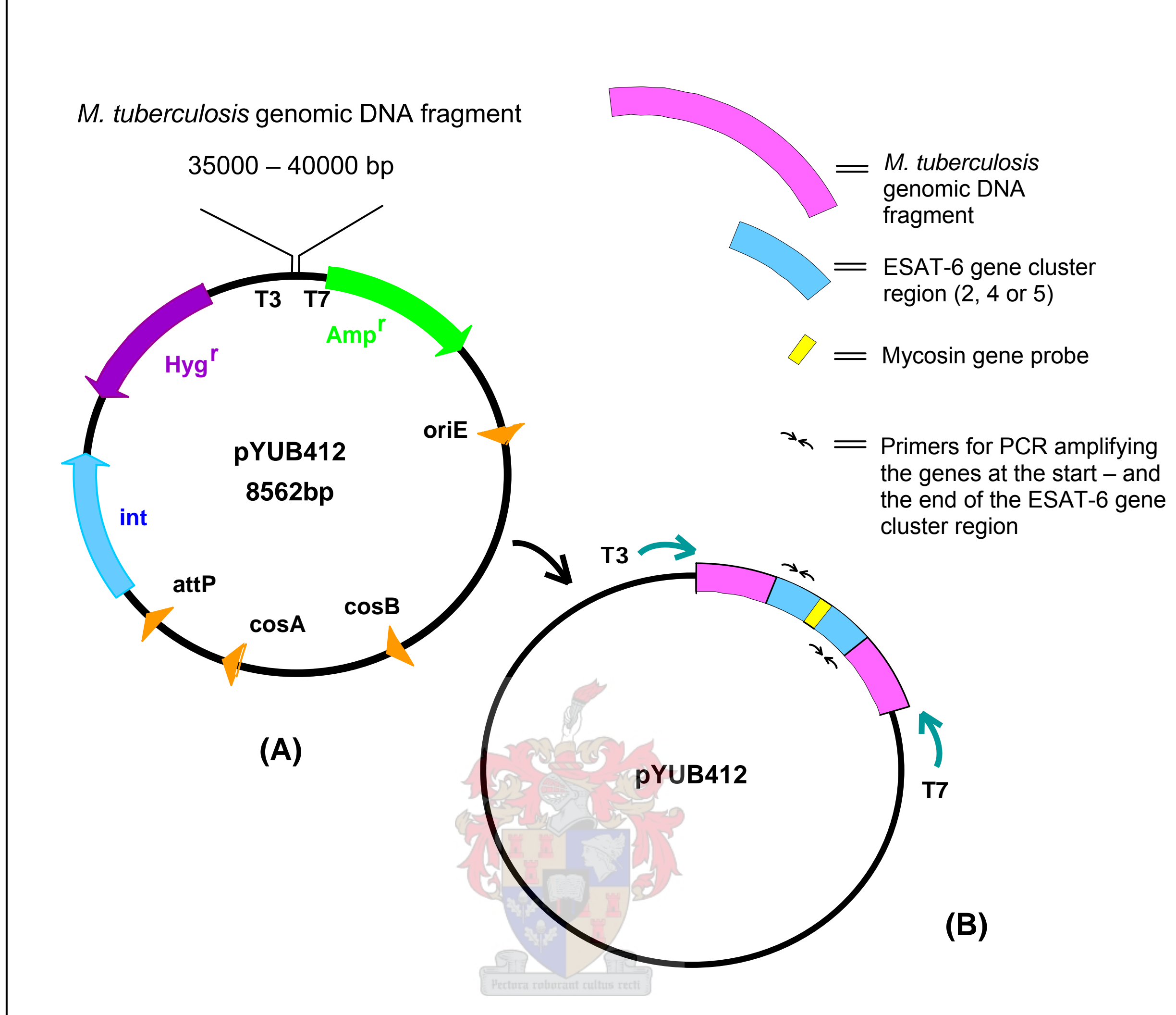


Figure 3.1 (A) A schematic representation of the *M. tuberculosis* cosmid library clones in pYUB412.2. (B) illustrates the pYUB412 construct containing the integrated ESAT-6 gene cluster 2, 4 or 5 (blue) with the Mycosin gene probe (yellow). The primers used for sequencing, T3 and T7, are also indicated.

3.2.4 Primers

The oligonucleotide primers used in this experiment were designed by making use of the Primer Premier program and are listed in Table 3.1. These primers were selected based on length and Tm.

Table 3.1 A list of the PCR oligonucleotide primers

Name of primer	Primer sequence (from 5' to 3')	Length of primer	Tm [°C]	Length of product
REGION 2	**PCR confirmation of cosmid isolation**			
3895cf	cctgcgtgtcgtgggagaa	19 bp	62	414 bp
3895cr	cgagggtgggcacagtatcg	20 bp	66	
3884cf	ggacccaagacgaacgagc	19 bp	62	305 bp
3884cr	gcgtagcggtgtccaatct	19 bp	60	
REGION 4	**PCR confirmation of cosmid isolation**			
3450cf	ggtccactgctcggtattccag	22 bp	64	103 bp
3450cr	ttgatccgccattgtcgctat	21 bp	62	
3444cf	catcgaatactccgttcgtcag	22 bp	66	203 bp
3444cr	gaaccgcgtttcccaagtc	19 bp	60	
REGION 5	**PCR confirmation of cosmid isolation**			
1782f	tagtgcgggcaacaccaag	19 bp	60	363 bp
1782r	caccatagcgattcggacgt	20 bp	62	
1798f	ggcgattaccagcgtagattc	21 bp	64	1242 bp
1798r	gacggcaacaacacttgagac	21 bp	64	
MYCOSIN 2	**Probe for cosmid isolation and PCR confirmation**			
P2fRNA	attacgtcctgtcggtag	18 bp	54	174 bp
P2rRNA	gcatattcttctcgcctg	18 bp	54	
MYCOSIN 4	**Probe for cosmid isolation and PCR confirmation**			
Prot4f	aagaacgccgtcatcgtg	18 bp	56	484 bp
Prot4r	gaatcgagtcgctgctga	18 bp	56	
MYCOSIN 5	**Probe for cosmid isolation and PCR confirmation**			
Prot5f	gtgctcgtaatgtcatcg	18 bp	54	658 bp
Prot5r	catatcggcaccatatcg	18 bp	54	

3.2.5 Probes

Mycosins 2, 4 and 5 (Brown et al 1998) are situated in the ESAT-6 gene cluster regions 2, 4 and 5 respectively. These 3 protease genes were subsequently used as probes for the detection of chromosomal domains containing the ESAT-6 gene cluster regions 2, 4 and 5. Mycosin 2 (3643 basepairs) was excised from the construct pUC-2a (Brown et al 1998) by restriction digestion with *Eco*RI and *Hind*III. pUC-2a was constructed previously by cloning the mycosin 2 gene into pUC18 using the same restriction enzymes. Specific domains of mycosins 4 (484 basepairs) and 5 (658 basepairs) were PCR amplified with forward and reverse primers (prot 4f and 4r and prot 5f and 5r as listed in Table 3.1) using *M. tuberculosis* H37Rv genomic DNA as the template for PCR amplification.

3.2.6 Radioactive labeling of probes

To create radioactive probes, each of the 3 genes were labeled using the Prime-It PmT Random Primer Labeling kit (Stratagene). DNA (10 µl containing approximately 200 ng DNA complementary to either mycosins 2, 4 or 5) was added to a single-use reaction tube provided in the kit. Pure, sterile water (32 µl) was added and the mixture vortexed and incubated at 100 ºC for 5 minutes. During this 5 minute incubation period, the mixture was vortexed a number of times. The mixture was then allowed to cool to room temperature. Thereafter, 5 µl ^{32}P-dCTP was added as well as 2.5 µl Random Primed Labeling enzyme. The 50 µl reaction mixture was mixed gently and then incubated at 37 ºC for one hour. After the incubation period, 2 µl of STOP solution (supplied by the kit) was added to the tube. The reaction mixture was purified through a column containing Sephadex G50M (Sigma) beads resuspended in STE (10mM Tris HCl, 100mM NaCl, 1mM EDTA, pH 8.0) to ensure the removal of salts and unincorporated nucleotides (these are a cause of background on X-ray films). Before adding the radioactive probe mixture, the column was equilibrated by adding 50 µl of STE solution and centrifuging it for 3 minutes at moderate speed. This was repeated twice and the probe was added to the column, and centrifuged again at moderate speed for 3 minutes. After all the impurities and unincorporated nucleotides had bound to the column beads, 50µl of the STE solution was added and the

Pectora roborant cultus recti

centrifugation step repeated. The purified radio-active labeled DNA (mycosin 2, 4 and 5) was eluted in the tube.

3.2.7 Isolation of cosmids containing selected genomic regions

Colony blotting (Sambrook et al 1989) was used for the isolation of the 3 cosmids (Figure 3.2) containing the ESAT-6 gene cluster regions 2, 4 and 5. Dilutions were made of the *M. tuberculosis* H37Rv cosmid library culture stocks (Bange et al 1999) in order to obtain ±200 colonies per LB-plate. It was found that a 3000x dilution resulted in ±200 colonies per plate. Nine LB-agar plates were plated with *E. coli* containing the cosmid library, for each region. Colony blotting was performed by placing pre-wetted membrane filters onto the colonies of each plate and incubating it for one minute. These membranes were then subjected to subsequent steps of denaturing, neutralization and washing using denaturing-, neutralization- and prewashing solutions (see Solutions and Buffers in Addendum A; Sambrook et al 1989). Membranes were then baked for two hours at 80ºC in a vacuum oven. Thereafter, the membranes were extensively washed and incubated overnight in prehybridization solution at 42 ºC (see Solutions and Buffers in the Addendum). Following prehybridization, the membranes were hybridized overnight at 42ºC in hybridization solution containing either of the radioactively labeled probes (see section 3.2.6). After hybridization and subsequent washing in washing solution (2x SSC, 0.1% SDS) twice for 30 minutes each at room temperature and twice for 30 minutes each at 50 ºC, the positive clones were visualized by autoradiography. The corresponding colonies on the plates were subsequently picked. Positive clones were first confirmed by PCR amplification with primers specific for the gene of the probe and secondly with primers specific for the start genes and end genes of the ESAT-6 gene cluster regions (see Figures 3.1 and 3.3). As a final confirmation, the isolated cosmids were sequenced on an automated sequencer using the T7 and T3 sequencing primers to determine the flanking sequences of the insert and confirm that the insert includes the correct gene cluster region (2, 4 and 5). See Figure 3.4.

Pectora roborant cultus recti

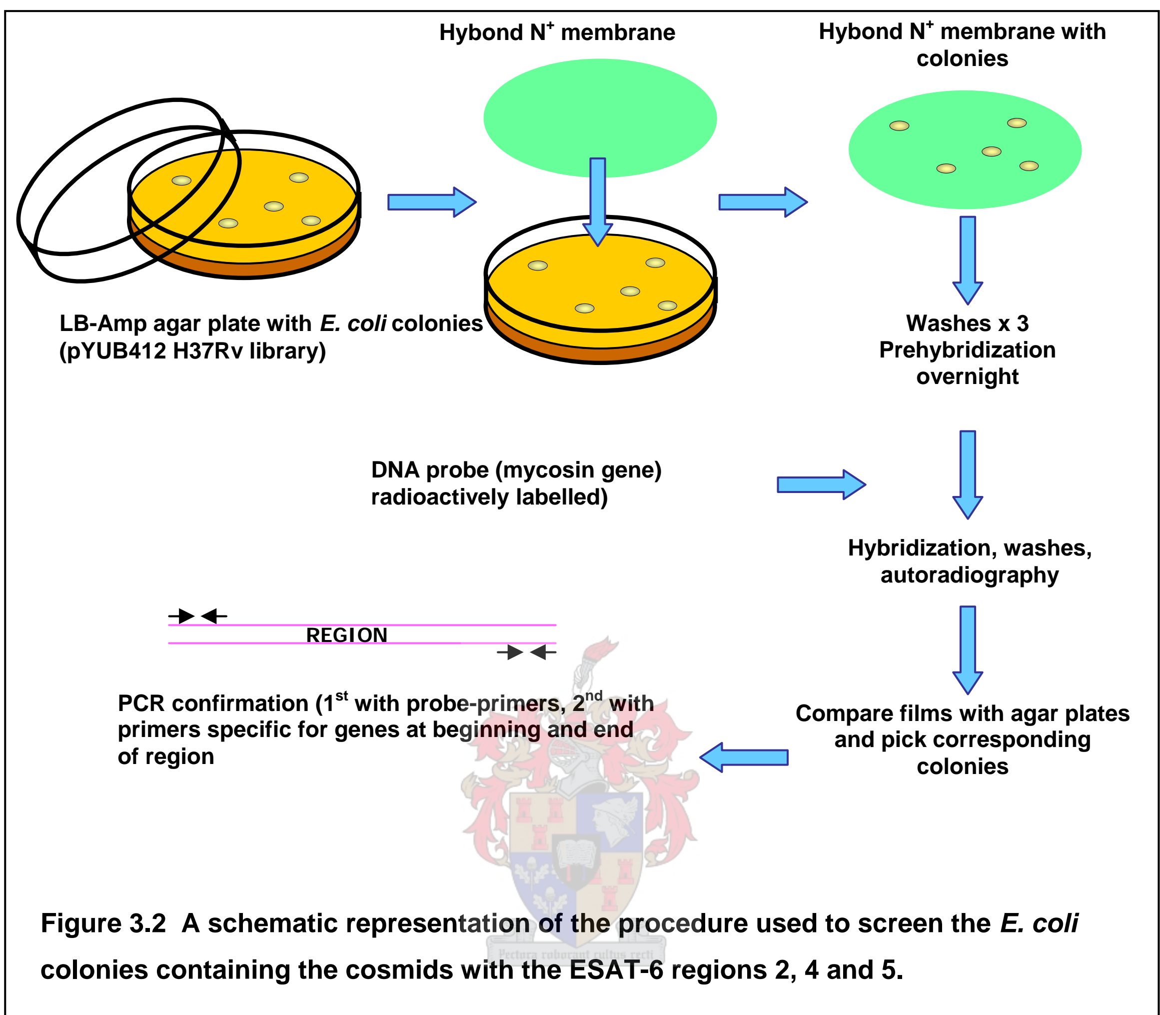


Figure 3.2 A schematic representation of the procedure used to screen the *E. coli* colonies containing the cosmids with the ESAT-6 regions 2, 4 and 5.

3.2.8 DNA sequence analyses

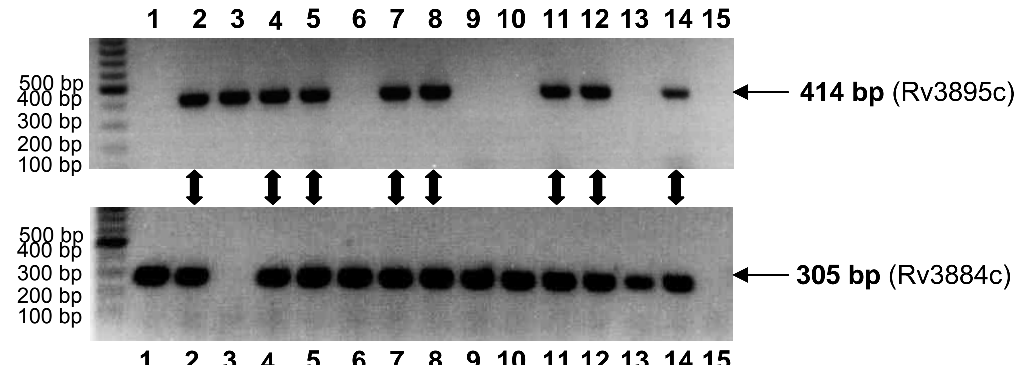
To determine which of the positive clones in this study contained regions 2, 4 and 5 in their entirety, the DNA sequence of these clones were BLASTed against the *M. tuberculosis* H37RV genomic DNA sequence on the http://www.genolist.pasteur.fr/Tuberculist/ website.

3.3 Results

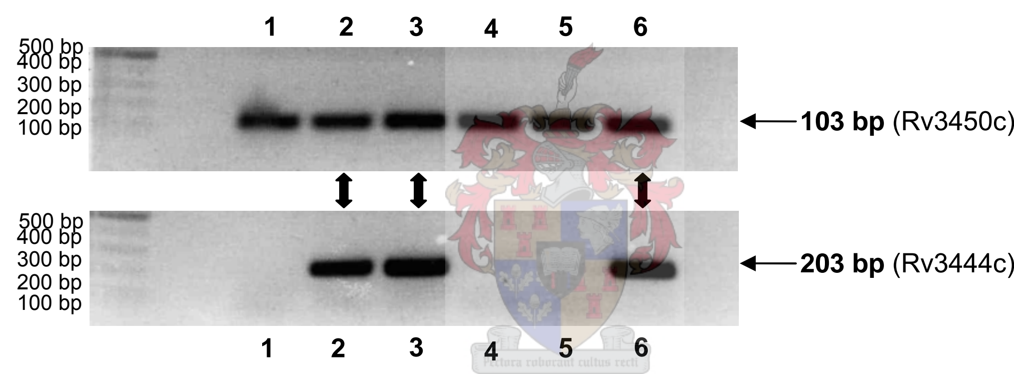
3.3.1 Isolation of cosmids containing selected genomic regions

pYUB412.2 cosmids containing *M. tuberculosis* H37Rv genomic DNA encoding for ESAT-6 gene cluster regions 2, 4 and 5, were isolated by means of colony blotting (Figure 3.2). These isolated gene cluster regions were confirmed by PCR analysis of the first and last genes of the regions (Figure 3.3 (A)-(C)). Sequencing of the flanking regions with T3 and T7 primers was performed as a final confirmation of the presence of region 2, 4 and 5 within these cosmids. The sequences (9 positive clones for region 2 and 3 positive clones for region 4 and 5 each) were BLASTed against the *M. tuberculosis* H37Rv genomic DNA sequence in Tuberculist (see DNA sequence analysis 3.2.1). The sequencing results confirmed that the clones contained the respective regions 2, 4 and 5. From each batch of positive clones per region, the smallest cosmid-fragment was chosen for use in further analyses. Figure 3.4 shows the 3 isolated cosmids (Insert A, B and C) containing the ESAT-6 gene cluster regions 2, 4 and 5 respectively.

Pectora roborant cultus recti



B: ESAT-6 gene cluster region 4 positive clones



C: ESAT-6 gene cluster region 5 positive clones

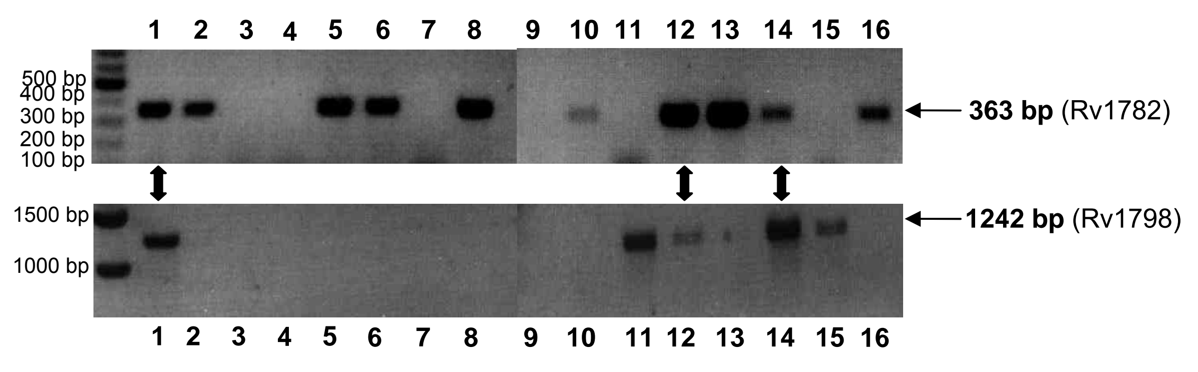


Figure 3.3 PCR analysis of the isolated cosmid with inserts containing the ESAT-6 gene cluster regions 2, 4 and 5, represented by A, B and C respectively. A: designated clones A2, A4, A5, A7, A8, A11, A12 and A14 all contain the PCR products of the first (Rv3895c) and last (Rv3884c) genes of region 2. B: designated clones B2, B3 and B6 contain the PCR products of the first (Rv3450c)- and last (Rv3444c) genes of region 4. C: designated clones C1, C12 and C14 contain the PCR products of the first (Rv1782) and the last (Rv1798) genes of region 5. These results indicate that all of the mentioned clones contain the entire regions 2, 4 and 5. Samples were run on 1% agarose gels and all PCR negative controls were negative (not shown).

ESAT-6 gene cluster region 2

Insert A12: 30805 Kb

ESAT-6 gene cluster region 4

Insert B3: 34488 Kb

ESAT-6 gene cluster region 5

Insert C1: 30805 Kb

Pectora roborant cultus recti

Figure 3.4 A schematic representation of the isolated cosmids containing the ESAT-6 gene cluster regions 2, 4 and 5. The annotated ORF numbers and position in the whole genome sequence of *M. tuberculosis* H37Rv is indicated as determined by T3 and T7 sequencing. The ESAT-6 gene cluster region 2, 4 and 5 are indicated by arrows. The sizes of the three smallest isolated cosmids are: Insert A12 – 30805 Kb, Insert B3 – 34488 Kb and Insert C1 – 34204 Kb.

3.4 Discussion and Conclusion

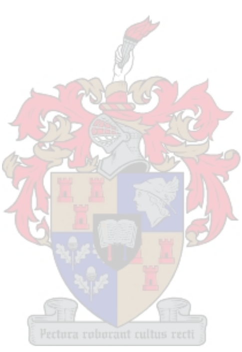
The five duplicated ESAT-6 gene cluster regions are described in detail in Gey van Pittius et al (2001). The virulence, diagnostic and potential vaccine qualities of the genes encoded by these regions have formed part of many studies (Tekaia et al 1999, Gey van Pittius et al 2001, Alderson et al 2000, Hsu et al 2003, Brodin et al 2004).

The completion of the genome sequencing of *M. tuberculosis* (Cole et al 1998) enabled the identification of five duplicated gene cluster regions, described as the ESAT-6 loci (Tekaia et al 1999, Gey van Pittius et al 2001). These gene clusters are designated region 1 (Rv3866-Rv3883c), region 2 (Rv3884c-Rv3895c), region 3 (Rv0282-Rv0292), region 4 (Rv3444c-Rv3450c) and region 5 (Rv1782-Rv1798) consistent with the arbitrary numbering system used previously to classify the five mycosin (subtilisin-like serine protease) genes identified from these regions (Brown et al 2000). Included in these clusters are genes encoding cellwall-associated ABC-transporters, ATP-binding proteins, secreted cell-wall associated subtisilin-like serine proteases (named the mycosins) and members of the CFP-10 (*esxB*) and ESAT-6 (*esxA*) gene families (encoding T-cell antigens lacking a detectable *secA* secretion signal; Gey van Pittius et al 2001).

The genomes of numerous other strains and species belonging to the genus *Mycobacterium* (as well as two species belonging to closely-related genera) contain orthologs of the ESAT-6 gene cluster regions of *M. tuberculosis* H37Rv. These include *M. tuberculosis* CDC551, *M. tuberculosis* 210, *M. bovis*, *M. bovis* BCG, *M. leprae*, *M. avium*, *M. paratuberculosis*, *M. smegmatis*, *Corynebacterium diphtheriae* and *Streptomyces coelicolor* (Gey van Pittius et al 2001, Pallen et al 2002).

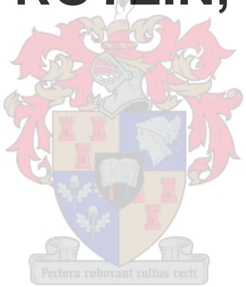
In this study, we have isolated ESAT-6 gene cluster regions 2, 4 and 5 by means of colony blotting. The isolated ESAT-6 gene cluster region 5 will form part of a secretion analysis

where the secretion of the ESAT-6 contained within this region, designated Mtb9.9A by Alderson et al (2000), is under investigation. The isolated ESAT-6 gene cluster regions 2 and 4 will also be important for future studies with regards to regulation and secretion.



CHAPTER FOUR

EXPRESSION AND PURIFICATION OF THE ESAT-6 PROTEIN, MTB9.9A



4.1 Introduction

The only available anti-tuberculosis vaccine, the live attenuated strain of *M. bovis* bacillus Calmette-Guérin (BCG), shows a variation in efficacy from one population to the other. Colditz et al (1994) have shown (by means of meta-analysis of 14 prospective trials and 12 case-control studies) that the BCG vaccine reduces the risk of tuberculosis by only 50%. Following this and the current alarming increase in the incidence of tuberculosis all over the world, there is a critical need for research in finding an improved method(s) of immunoprophylaxis against tuberculosis.

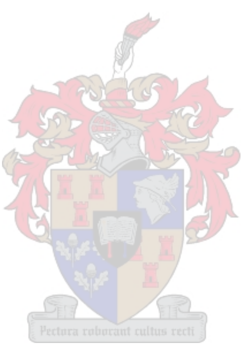
Proteins actively secreted by *M. tuberculosis* have been shown to be important targets for the immune system during the early phase of an infection (Okkels et al 2004). Such antigens are therefore candidates to be included in a future vaccine which could either be in the form of a live vaccine (e.g. a genetically modified BCG) or of a non-living subunit vaccine.

Alderson et al (2000) showed that 83% of PPD$^+$ (testing positive for purified protein derivative) donors gave a significant response to *M. tuberculosis* rMtb9.9A, indicating that it was an efficiently secreted protein. The gene encoding this Mtb9.9A protein (Rv1793 or *esxN*) is a member of the ESAT-6 gene family and is situated in the ESAT-6 gene cluster region 5. Alderson et al (2000) have shown that the quantity of Mtb9.9A in isolated culture filtrate proteins was sufficient to stimulate T cell responses. They have also indicated that although the Mtb9.9A antigen was only 10 kDa, it was found to contain at least five distinct T-cell epitopes further emphasizing its important role in the immunological response after initial *M. tuberculosis* infection. As this secreted protein is a member of the ESAT-6 protein family, and the gene encoding this protein is found within the ESAT-6 gene cluster region 5, we hypothesize that Mtb9.9A is secreted by the ESAT-6 gene cluster region 5-encoding transporter apparatus.

The objective of this study was to clone, express and purify the Mtb9.9A antigen of *M. tuberculosis* as a GST-fusion protein for antibody preparation. This was done with the aim of

using the antibodies for Western blotting to detect secretion of the ESAT-6 gene from the ESAT-6 gene cluster region 5 (see Chapter 5). This is important for the understanding the mechanism of mycobacterial protein transport and secretion mechanisms, as well as the secretion of important T-cell antigens of mycobacteria.

The results of this study could lead to the development of efficient strategies to influence, terminate or enhance the secretion of important mycobacterial antigens. In turn this might have an influence on the immunogenicity of the recombinant pathogen, which may ultimately have an impact on vaccine design and development.



4.2 Materials and Methods

4.2.1 DNA and protein sequence analyses

The DNA and protein sequence of Mtb9.9A of *M. tuberculosis* H37Rv was obtained from the Pasteur Institute's *M. tuberculosis* genome database website Tuberculist (http://genolist.pasteur.fr/Tuberculist).

4.2.2 Bacterial strains

E. coli XL1-Blue was the host for plasmid propagation as well as for the expression of protein during this experiment. Cells were made electrocompetent as described in the Addendum.

4.2.3 Media and culture conditions

For general cloning, *E. coli* transformants were grown on solid or in liquid Luria-Bertani (LB) medium as described by Sambrook et al (1989). For the expression and solubility tests as well as the purification of the Mtb9.9A, *E. coli* was grown in liquid Rich Broth (LB medium supplemented with 0.1% glucose) containing Ampicillin (50 µg/ml; Roche).

Pectora roborant cultus recti

4.2.4 Primers

The oligonucleotide primer-pair used in this experiment was designed by making use of the Primer Premier program and is listed in Table 4.1. These primers were selected based on length and Tm.

Table 4.1 A list of the PCR oligonucleotide primers

Name of primer	Primer sequence (from 5' to 3')	Length of primer	Tm [°C]	Length of product
Mtb9.9af	ggatccacgattaattaccagttcg	25 bp	62	285 bp
Mtb9.9ar	ggatccgctggagccgacg	19 bp	46	

4.2.5 DNA manipulations for protein purification

All DNA manipulations were performed as described by Sambrook et al (1989). *mtb9.9a* was PCR amplified from *M. tuberculosis* H37Rv genomic DNA using the primers listed in Table 4.1. The 5' and 3' primers were both designed to include *Bam*HI sites (See Table 4.1). The purified *mtb9.9a* amplicon was then cloned into the Glutathione S-transferase (GST) fusion protein vector, pGex-2T (Pharmacia), generating pGexMtb9.9A (Figure 4.1) and transformed into *E. coli* XL-1 Blue cells (See Addendum).

Pectora roborant cultus recti

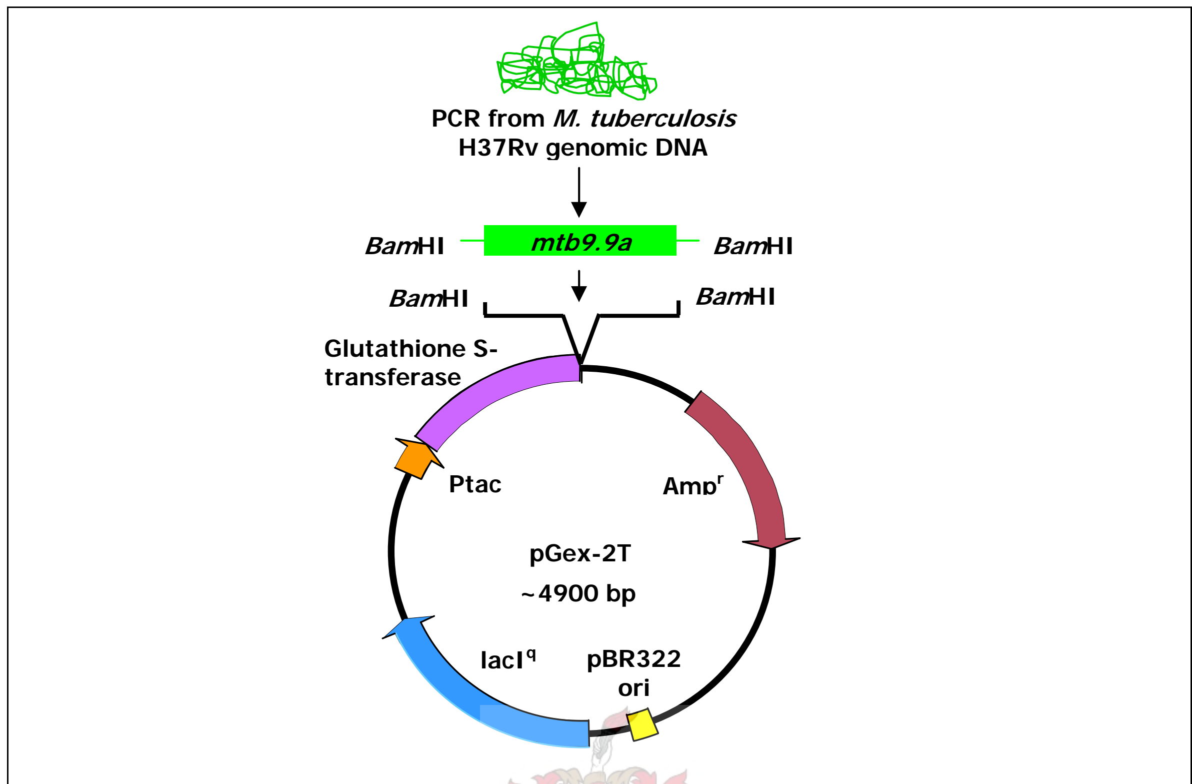


Figure 4.1 A schematic representation of the cloning of the gene, *mtb9.9a*, encoding Mtb9.9A into the fusion vector pGex-2T.

4.2.6 Expression test

For the expression test, recombinant *E. coli* transformed with pGex-2T (as control) and pGexMtb9.9A were each cultured in 20 ml LB media supplemented with glucose (see media and culture conditions 4.2.3). Cells were grown to an OD_{600} of 0.6 (approximately 6 hours), after which 1ml of each culture was centrifuged and the pellets diluted into 100 µl of a 2x reducing SDS buffer and stored in the -20ºC freezer. These samples of pGex-2T and pGexMtb9.9A represent the uninduced culture. Protein expression was induced in the remaining cultures (19 ml) by adding 60 µl 0.1M IPTG. From the time of induction, 2, 3, 4 and 24 hour samples were taken as described above, except that instead of taking 1 ml samples, 0.5 ml samples were taken. The uninduced and 2, 3, 4 and 24 hour induced samples were run on a SDS-PAGE gel and Western blotting was performed thereafter as described by Sambrook et al (1989).

4.2.7 Solubility test

The solubility test was done essentially the same as described above for the expression test. The uninduced and induced bacterial samples (3h and 24h) of pGex-2T and pGexMtb9.9A were resuspended in ice cold, sterile 1x PBS solution and sonicated in a Misonix Cup Sonicator (4.5 setting) for a total of 2 minutes (15 second bursts with 30 second intervals), before the addition of 100 µl 10% Triton X-100 to each 1 ml sonicated sample. The samples were then mixed gently on a turning-wheel for 30 minutes at 4°C and centrifuged at 13 000 rpm for 20 minutes at 4°C. The supernatants were removed into new Eppendorf tubes. One hundred microlitres of a 2x SDS-reducing buffer was added to 100 µl of each supernatant, the pellets gently resuspended and stored at -20°C. Two hundred microlitres of the remainder of the supernatants were transferred to clean Eppendorf tubes (the rest of the supernatants frozen at -20°C) and 50 µl 50% GST agarose bead slurry was added to each before they were gently mixed on a turning-wheel for 5 minutes at room temperature. 1 ml 1x sterile PBS was added to the supernatant-bead samples, vortexed briefly and centrifuged for 5 seconds to sediment the beads. The GST agarose beads bind to GST and would thus bind the soluble GST-Mtb9.9A fusion protein (expressed from pGexMtb9.9A) as well as native GST protein (expressed by pGex-2T). The supernatants were removed completely using a Hamilton syringe. The PBS-washing step was repeated twice for a total of 3 times. Finally, 100 µl 2x SDS reducing buffer was added to each bead pellet and the sample frozen at -20°C. To optimize the solubility of Mtb9.9A, experimental conditions were varied, for example the final IPTG concentration and incubation temperature after induction of protein expression were decreased, respectively, to 0.2 mM and 25°C. The cultures were also induced at OD_{600} of 0.7-0.8, harvested after 2 hours of induction, and the incubator's shaking speed was increased to 220 rpm. The solubility of the protein was confirmed by SDS-PAGE and Western blotting (Sambrook et al 1989).

4.2.8 Protein purification

Four litres of *E. coli* transformed with pGexMtb9.9A was cultured in Rich Broth (LB medium supplemented with 1% glucose) containing Ampicillin, for the purification of GST-

Mtb9.9A fusion protein. Induction of protein expression, as well as the culture conditions and harvesting were performed in the same manner as described above for the solubility test. After centrifugation at 4°C in GSA tubes (7000 rpm for 10 min), the pellets were resuspended in 25 ml ice-cold PBS (50 µl PBS per 1 ml culture) and pooled (about 280 ml). A 10mg/ml lysozyme stock solution (final concentration 10%) and protease inhibitors (2 ml 200 mM IAA at 1:100, 2 ml 100 mM PMSF at 1 mM, 0.8 ml 10 mg/ml APMSF at 0.04 mg/ml, 6 ml Protease cocktail at 1:33, 40 ml 0.5 M EDTA at 100 mM, 20 ml 1 mg/ml Trypcin at 0.1 mg/ml, 0.04 ml 10 mg/ml Aprotinin at 0.002 mg/ml and 0.1 ml 1 mg/ml Leupeptin at 0.0005 mg/ml) were added to the sample. This cocktail was sonicated in a Misonix Cup Sonicator (4.5 setting) for a total of 2 minutes (15 second bursts with 30 second intervals) before adding Triton-X 100 at a concentration of 1:10 for solubilization of proteins. The sample was gently mixed on a turning wheel for 30 minutes at 4°C and then divided into 8 Falcon tubes. The GST purification kit (Amersham) was used to purify GST-Mtb9.9A fusion protein through binding to a GST Sepharose beads column. This was done by firstly preparing a Glutathione Sepharose 4B slurry as described in the GST purification kit. The sample was centrifuged at 8000 rpm for 10 min at 4°C, after which the supernatants were pooled (± 300 ml) and divided into 6 Falcon tubes (± 50 ml per tube). 1 ml Glutathione 50% slurry was added to each 50 ml supernatant and gently mixed at room temperature on a turning wheel for 30 minutes. The supernatants, containing beads with bound GST-Mtb9.9A protein, were added to disposable columns (included in the GST purification kit). Each column was washed with 45 ml 1 x PBS for a total of 3 times. The GST-Mtb9.9A protein was eluted with Glutathione Elution Buffer (supplied by the kit) for a total of 5 times in 1.5 ml. The five 1.5 ml eluates containing purified GST-Mtb9.9A were dialysed in 1 x PBS buffer (to remove salts), after which OD_{595} readings were taken for the calculation of protein concentration (BioRad Protein Assay) and the eluate with the highest concentration (1.74 mg/ml) was sent for antibody production to Prof Dirk Bellstedt at the Department of Biochemistry, University of Stellenbosch (Bellstedt et al 1987). The purification of the GST-Mtb9.9A protein was confirmed by SDS-PAGE and Western blotting by using rabbit anti-GST antibodies (primary antibody; 1/1000 dilution; a kind gift from Dr Kevin Dennehy) and HRPO-conjugated goat anti-rabbit antibody (secondary antibody; 1/10 000; Caltag Laboratories).

Pectora roborant cultus recti

4.3 Results

4.3.1 Expression test

As observed in both SDS-PAGE and Western Blotting analyses results, high levels of GST (~26kDa) and GST-Mtb9.9A fusion protein (~35 kDa) were found to be expressed by *E. coli* transformed with pGex-2T and pGexMtb9.9A, respectively, between 2 and 24 hours after induction with IPTG (see Figure 4.2). No expression was seen in the uninduced samples. The highest levels of expression appear to occur between 2-3 hours after induction, making this the optimum time of expression.

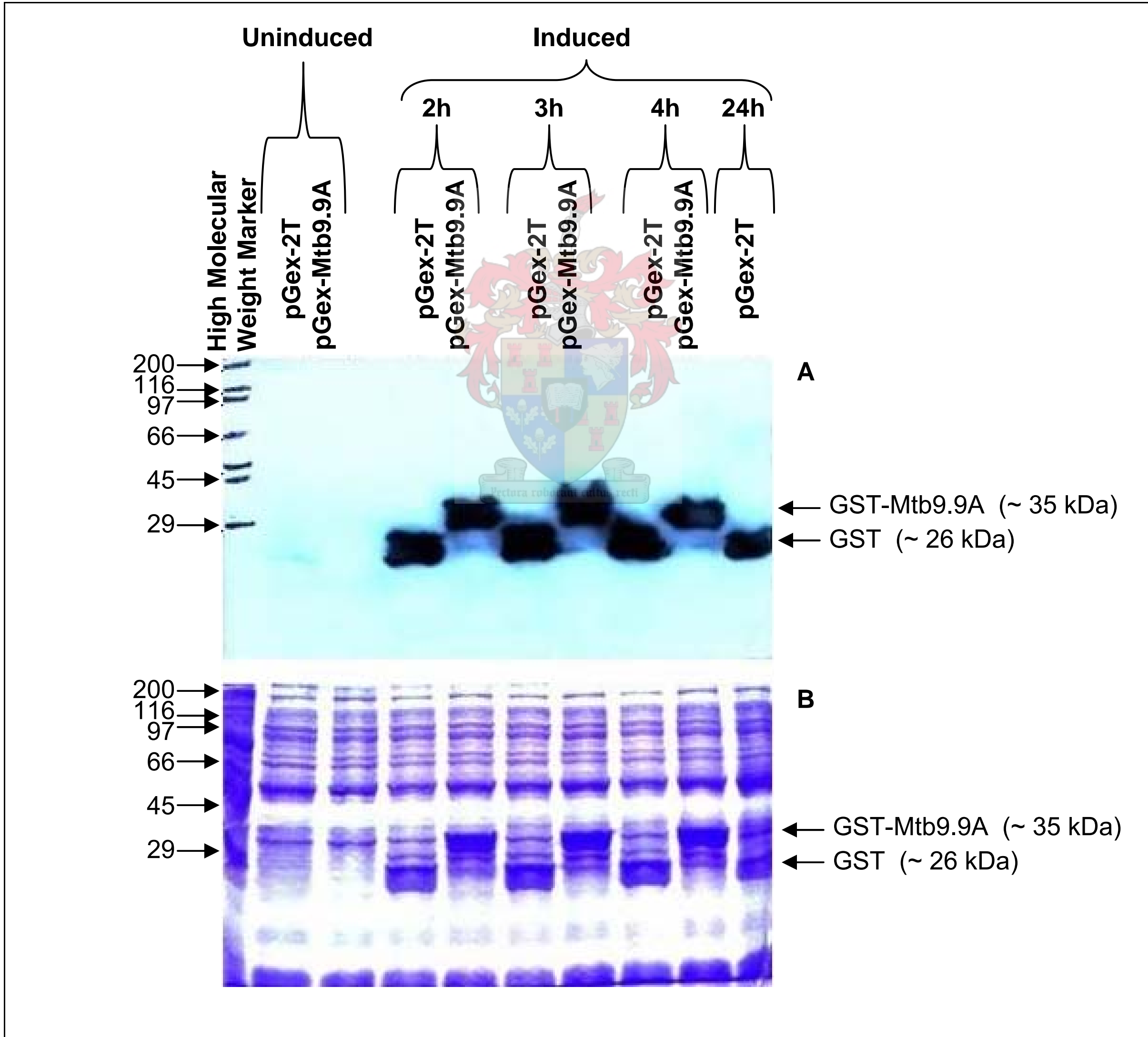
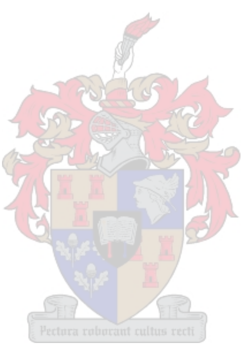


Figure 4.2 The expression of GST-Mtb9.9A at 2 – 24h. (A) represents the Western blot and (B) the corresponding SDS-PAGE gel.

4.3.2 Solubility test

In order to elucidate whether the expressed GST-Mtb9.9A fusion protein was soluble, a solubility test was performed. Optimized cultures were harvested after 2 hours of IPTG induction (optimal expression time) and cells were disrupted, after which the presence of fusion protein in the supernatant was determined by binding to GST agarose beads. The SDS-PAGE and Western Blotting results for the native GST protein show clearly that relatively high levels of GST protein was present in the pellet fraction and adequate soluble amounts of the protein could be found in the supernatant (able to bind to the agarose beads, see Figure 4.3). The overall expression levels of the GST-Mtb9.9A fusion protein was lower than that of the native GST protein, but the results were comparable in that the majority of the fusion protein could be identified in the soluble fraction of the supernatant.



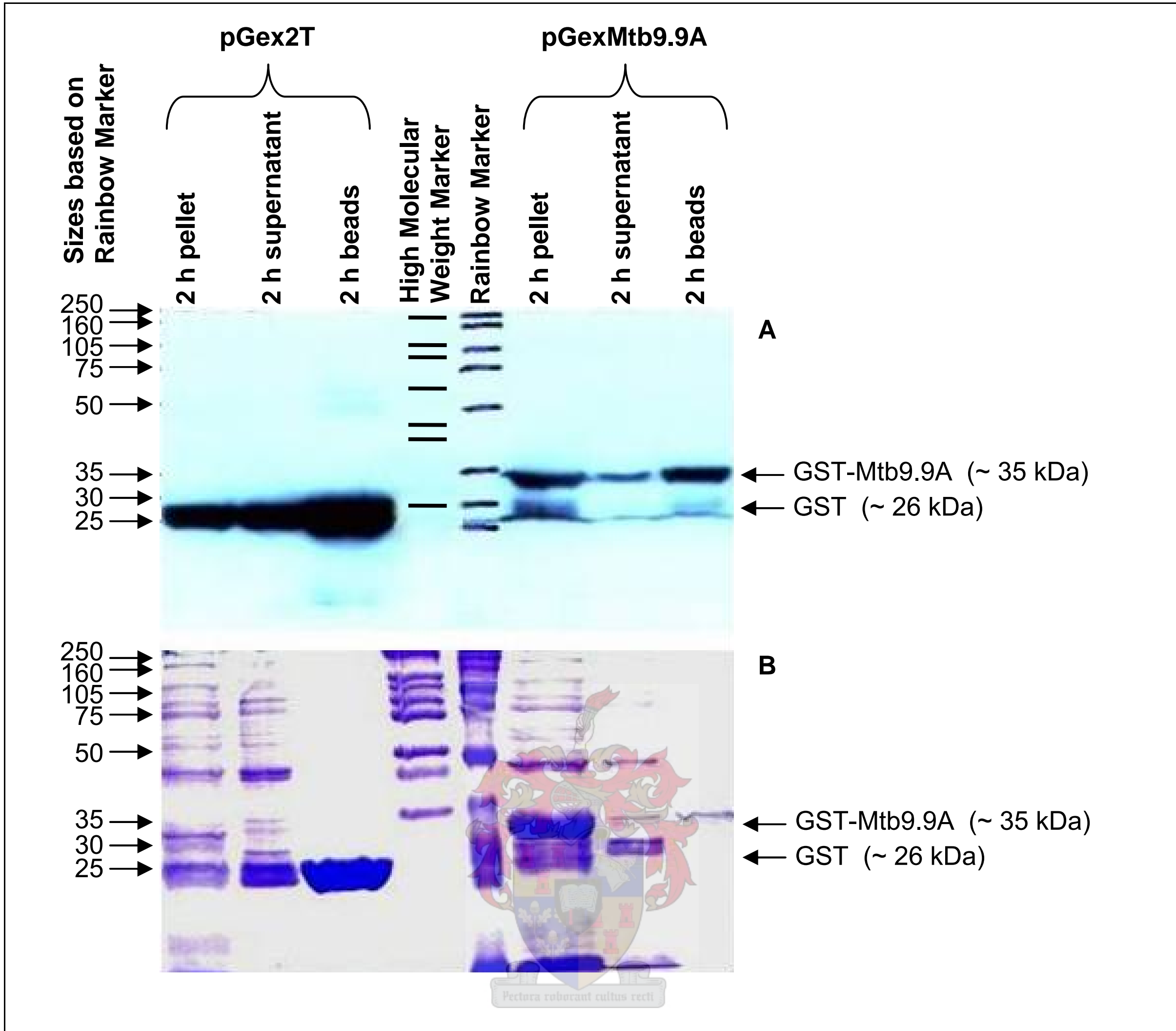


Figure 4.3 The solubility of GST-Mtb9.9A at 2h after induction of protein expression. (A) represents the Western blot and (B) the corresponding SDS-PAGE gel.

4.3.3 Purification of Mtb9.9A

GST-Mtb9.9A fusion protein was purified using a GST Sepharose 4B bead column. Five eluates were obtained and tested on SDS-PAGE and Western Blotting in order to determine purity and levels of purified protein in the eluates. The results of these analyses are presented in Figure 4.4, and show that the majority of the fusion protein (~35 kDa) was eluted in Eluate 1, after which the amount of eluted protein became progressively less. Although protease inhibitors were included in the purification, there appears to be some degradation of the purified fusion protein,

as a band was also observed at around 26 kDa in the Eluate 1 and 2 lanes, which would indicate the presence of only the GST part of the fusion protein. Eluate 1 was used to successfully generate antibodies to GST-Mtb9.9A fusion protein.

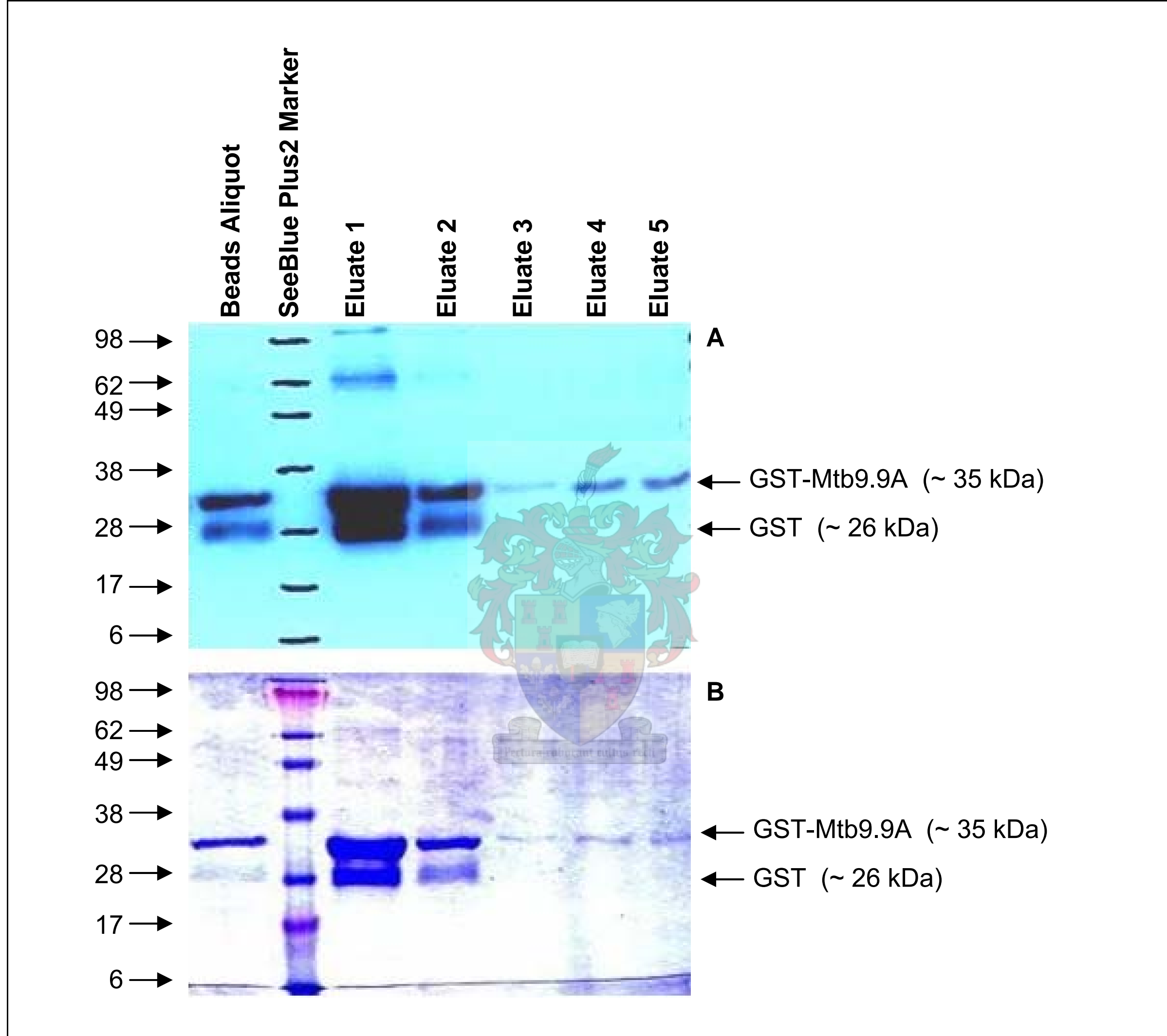


Figure 4.4 The five eluates of GST-Mtb9.9A collected after elution with glutathione elution buffer. (A) represents the Western blot and (B) the corresponding SDS-PAGE gel. The lanes containing Eluates 1 and 2 have bands representing the GST (~26kDa) protein. This indicates some degradation of the GST-Mtb9.9A fusion protein.

4.4 Discussion and Conclusion

The gene encoding the ESAT-6 low-molecular-mass protein is copied 11 times throughout the genome of *M. tuberculosis*, where 5 of the 11 copies form part of the five ESAT-6 gene cluster regions (Gey van Pittius et al 2001). ESAT-6 is a potent T-cell antigen that plays an important role during the initial stages of infection (Okkels et al 2004). Mtb9.9A is the member of the ESAT-6 antigen family from the ESAT-6 gene cluster region 5 and was first characterised by Alderson et al (2000).

In this study we have cloned, expressed and purified the important *M. tuberculosis* T-cell antigen Mtb9.9A (Alderson et al 2000). The purified GST-Mtb9.9A fusion protein was sent to Prof Dirk Bellstedt (at the University of Stellenbosch) for antibody generation in rabbits. Bellstedt et al (1987) make use of a method where the protein (to which the antibody is raised) is adsorbed to acid-treated, naked bacteria which serve as immune-carriers. The method is an improvement to previous standard procedures as only a very small amount of protein is needed for antibody production.

Pectora roborant cultus recti

The antibodies generated against Mtb9.9A formed part of a secretion analysis study to prove that the ESAT-6 gene cluster region 5 is necessary for the secretion of the Mtb9.9A antigen. This is important for understanding mycobacterial transport and secretion mechanisms, as well as the secretion of key T-cell antigens of mycobacteria, which could lead to the development of efficient strategies to interfere with antigen secretion, thereby influencing the immunogenicity of the pathogens which may ultimately have an impact on vaccine design and development.

CHAPTER FIVE

SECRETION ANALYSIS OF THE ESAT-6 GENE CLUSTER REGION 5

5.1 Introduction

The structure and organization of the ESAT-6 gene clusters have been described in detail previously (Gey van Pittius et al 2001). An analysis of the potential functions of the proteins encoded by these clusters shows that most of these proteins have a potential to function in a protein-dependent ATP-binding cassette active transport system.

It has been shown previously that ESAT-6 gene cluster region 1 is responsible for the secretion of the ESAT-6 protein present within this region, explaining the absence of any ordinary *sec*-dependent secretion signals in the amino acid sequence of this protein (Brodin et al 2004, Converse et al 2004, Gey van Pittius et al 2001, Guinn et al 2004, Hsu et al 2003, Pym et al 2003, Stanley et al 2003). During the current study, we attempted to elucidate whether the proteins encoded by the ESAT-6 gene cluster region 5 (Gey van Pittius et al 2001) also function together as a mycobacterial membrane-bound complex involved in protein-dependent transport. We aimed to investigate whether, if this is the case, this transport system is responsible for the active secretion of the ESAT-6 protein family member (designated Mtb9.9A by Alderson et al 2000) present within ESAT-6 gene cluster region 5.

5.2 Materials and Methods

5.2.1 DNA and protein sequence analyses

The DNA and protein sequence information of *M. tuberculosis* H37Rv and *M. smegmatis* mc^2155 was obtained from publicly available genome sequence databases at the Pasteur Institute (http://genolist.pasteur.fr/Tuberculist) and The Institute for Genomic Research (TIGR) (http://www.tigr.org/cgi-bin/BlastSearch/blast.cgi?organism=m_smegmatis), respectively, as well as from previously-published sequence analyses (Tekaia et al 1999, Gey van Pittius et al 2001). Sequence alignments were done using the DNAman program (version 4.1) from Lynnon Biosoft.

5.2.2 Bacterial strains

The *Escherichia coli* XL1-Blue strain was used as the host for the propagation of all cosmids and plasmids. *M. smegmatis* mc^2155 (Snapper et al 1990) was used for, cosmid integration, heterologous gene expression and protein secretion analyses (see sections 5.2.5 - 5.2.7).

5.2.3 Media and culture conditions

E. coli was grown on solid medium or in liquid Luria-Bertani (LB) medium (Sambrook et al 1989). *M. smegmatis* was grown on Middlebrook 7H11 agar supplemented with filter-sterile OADC (0.005% oleic acid, 0.5% BSA, 0.2% glucose, 0.02% catalase, 0.085% NaCl) and containing 0.05% Tween 80 (Sigma) that was used for selection of transformants on solid media. Different recombinant *M. smegmatis* clones were grown at 37°C in 200 ml Kirchner's broth (3 g/l Na_2HPO_4, 4 g/l KH_2PO_4, 1.07 g $MgSO_4.7H_2O$, 2.5 g/l Tri-sodium citrate, 20% glycerol, 5 g/l asparagine) to an optical density of 0.3 at 600 nm (early log phase). Kirchner's medium was used to acquire culture filtrate proteins in a protein-free medium that will create a secretion analysis environment free of background. For antibiotic selection Hygromycin (50 μg/ml, Roche), Kanamycin (50 μg/ml, Roche) and/or Ampicillin (50 μg/ml, Roche) were added to bacterial

cultures. For the selection of cultures containing both the integrating cosmid and plasmid the mycobacteria were grown in both Hygromycin and Kanamycin.

5.2.4 Primers

The oligonucleotide primers used in this experiment were designed by making use of the Primer Premier program for the confirmation of cloned inserts and are listed in Table 5.1. These primers were selected based on length and Tm.

Table 5.1 PCR oligonucleotide primers used in this study

Name of primer	Primer sequence (from 5' to 3')	Length of primer	Tm [°C]	Length of product
Mtb9.9af	ggatccacgattaattaccagttcg	25 bp	62	285 bp
Mtb9.9ar	ggatccgctggagccgacg	19 bp	46	
ESAT-6f	gagcagcagtggaattccg	20 bp	62	271 bp
ESAT-6r	tcccagtgacgttgccttc	19 bp	60	
TB10.4f	ggatccatgtcgcaaatcatgtacaa	26 bp	74	300 bp
TB10.4r	ggatccgccgccccatttggcgg	23 bp	80	

Pectora roborant cultus recti

5.2.5 DNA manipulations for secretion analysis

All DNA manipulations were performed essentially as described by Sambrook et al (1989). See Figure 5.1 and Table 5.2 for the details of the ESAT-6 region 5 cosmid constructs and plasmid constructs used in this experiment. The episomally-replicating plasmid pSD21 was a gift from S. Daugelat (Max-Planck-Institut für Infektionsbiologie, Berlin, Germany). pSD21 contains a C-terminal HA epitope that had been codon optimized for mycobacteria (Daugelat et al 2003). pSDMtb9.9A was constructed by the insertion of the *mtb9.9a* gene in frame into the plasmid pSD21 (see Figure 5.1). Other constructs used during the secretion analysis were pMB154 (constructed by insertion of ESAT-6 into pSD21; constructed by M. Braunstein and a gift from S. Daugelat) and pSDTB10.4 (constructed by insertion of TB10.4 into pSD21 behind the

hsp60 promoter), a gift from Dr N.C. Gey van Pittius (see Figure 5.1). These were used to analyze specificity, i.e. to determine whether the ESAT-6 region 5 will only be responsible for the secretion of its own ESAT-6 protein (Mtb9.9A) and not that of another ESAT-6 gene cluster region.

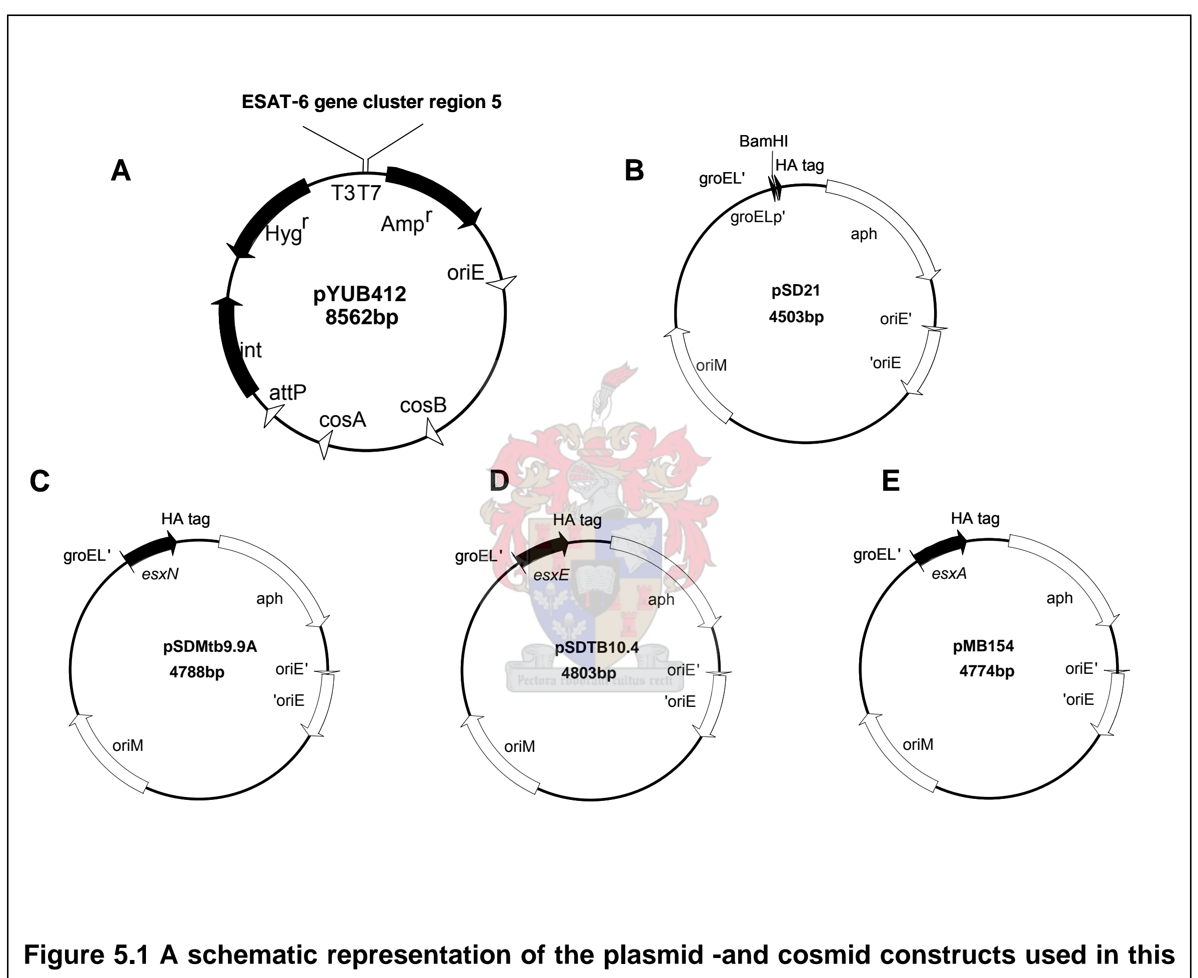


Figure 5.1 A schematic representation of the plasmid -and cosmid constructs used in this study.

Table 5.2 Plasmids and cosmid constructs used in this study

Plasmids and cosmids	Characteristics	Source of reference
pYUB412	Single copy integrating *E. coli*-mycobacterial shuttle vector cosmid library containing *M. tuberculosis* genomic DNA fragment inserts of approximately 40 000 bp, Hyg^r and Amp^r	F-C Bange (Bange et al 1999)
pSD21	*E. coli*-mycobacterial shuttle expression cloning vector for expression of C-terminal HA-tagged recombinant proteins, Kan^r	S Daugelat
pSDMtb9.9A	*E. coli*-mycobacterial shuttle expression vector for expression of the C-terminally HA-tagged *M. tuberculosis* ESAT-6 gene cluster region 5-specific ESAT-6 protein, Rv1793 (Mtb9.9A), Kan^r	This study
Cosmid containing ESAT-6 gene cluster region 5	Single copy integrating cosmid isolated from *M. tuberculosis* H37Rv genomic DNA cosmid library. Contains DNA fragment encompassing the complete ESAT-6 gene cluster region 5, Hyg^r and Amp^r	This study
pMB154	*E. coli*-mycobacterial shuttle expression vector for expression of the C-terminally HA-tagged *M. tuberculosis* ESAT-6 gene cluster region 1-specific ESAT-6 protein, Rv3875, Kan^r	S Daugelat (constructed by M. Braunstein)
pSDTB10.4	*E. coli*-mycobacterial shuttle expression vector for expression of the C-terminally HA-tagged *M. tuberculosis* ESAT-6 gene cluster region 3-specific ESAT-6 protein, Rv0288 (TB10.4), Kan^r	NC Gey van Pittius

5.2.6 Transformation of M.smegmatis

Transformation *of M. smegmatis* with the plasmid and cosmid constructs was done using electroporation, as described previously (Jacobs et al 1991). pSDMtb9.9A was transformed into wild-type *M. smegmatis.* Additionally, pSDMtb9.9A, pMB154 and pSDTB10.4 were transformed in *M. smegmatis* containing the already integrated ESAT-6 region 5 (see Chapter 2). To confirm that the *M. smegmatis* cells with the integrated region 5 were successfully transformed with pSDMtb9.9A, pMB154 and pSDTB10.4, PCR amplification of the above mentioned cultures were carried out with primers listed in Table 5.1 as described by Sambrook et al (1989).

5.2.7 Protein secretion analysis

For the protein secretion analysis, 200 ml mycobacterial culture cells was grown to an OD_{600} of no more than 0.3 (early log phase) in Kirchner's Broth to make sure that the cultures do not contain lysed cells. The cultured cells were then centrifuged at 3000 g for 10 minutes. The culture supernatants containing culture medium and secreted proteins were removed by

aspiration and the pellets were resuspended in 1 ml of a 1 x PBS and sonicated in a Misonix cup sonicator (4.5 setting) on ice for a total of 5 minutes (15 second bursts with 30 second intervals). The sonicated pellets were then centrifuged at 4 ºC at 3000 rpm for 10 minutes. The resulting supernatants containing the whole cell lysate proteins and the supernatants from the culture filtrates were filter-sterilized by serial filtration through 1.0 μM, 0.45 μM and 0.22 μM filters. The culture filtrate proteins were concentrated using the 3 kDa cut-off Centriprep 3 centrifuge concentration system (Centricon).

The protein concentrations of the concentrated culture filtrates (CF) and whole cell lysates (WCL) were determined spectrophotometrically by means of the BioRad protein concentration determination assay (as described by the manufacturer). A Tris/Glycine/SDS-PAGE gel (16% resolving gel, 10% spacer gel and 4% stack gel; Schagger and von Jagow 1987) was used to separate the low molecular weight proteins. 10 μg of each sample was loaded into each well. *M. tuberculosis* H37Rv culture filtrate proteins (a gift from C. Pheiffer, University of Stellenbosch, Tygerberg, South Africa) was used as a positive control of native ESAT-6 expression and secretion. During the Western blotting analyses rabbit anti-Mtb9.9A polyclonal antibodies (raised against purified Mtb9.9A protein, see chapter 4) were used at a dilution of 1/25 to detect the presence of secreted *M. tuberculosis* Mtb9.9A protein. Mouse anti-HA monoclonal antibodies (HA.11, Clone 16B12, Covance) were used at a dilution of 1/2000 to detect HA-tagged recombinantly-expressed Mtb9.9A protein in the culture filtrate sample of *M. smegmatis* containing the integrated ESAT-6 region 5 and co-transformed pSDMtb9.9A. Horse radish peroxidase (HRPO)-conjugated goat anti-rabbit antibodies (Caltag Laboratories) were used at a dilution of 1/10 000 as secondary antibody in the Western blotting analysis.

5.3 Results

5.3.1 DNA sequence analysis for confirmation of in-frame cloning

Figures 5.2 and 5.3 represent the sequence analysis and the translation, respectively, of Mtb9.9A-HA cloned into pSD21. As seen on in these figures, Mtb9.9A is not only present (Figure 5.2), but also cloned in-frame with GST-HA (Figure 5.3) into the expression vector pSD21.

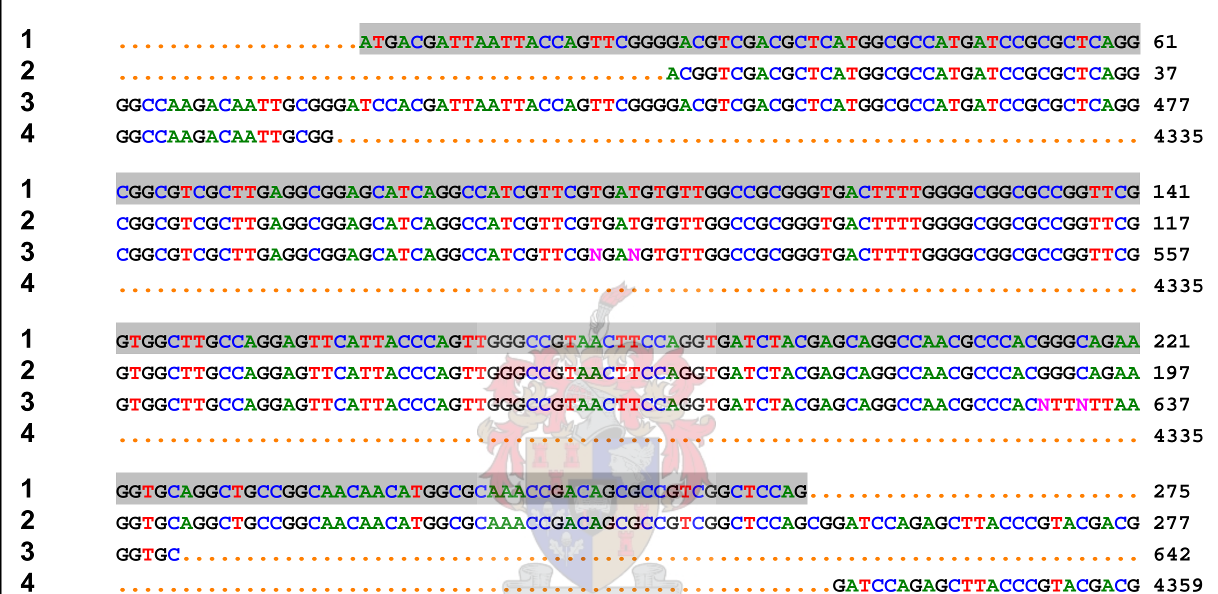


Fig 5.2 Multiple sequence alignment of *Bam*HI-Mtb9.9A-*Bam*HI fragment cloned into the pSD21 expression vector, confirming the presence of this fragment in the *Bam*HI restriction site of pSD21. Row 1 represents the Mtb9.9A insert and is highlighted in grey. Lane 2 represents the nucleotide sequencing result using the forward oligonucleotide of Mtb9.9A and lane 3 the result using the reverse oligonucleotide. Lane 4 represents the nucleotide sequence of native pSD21.

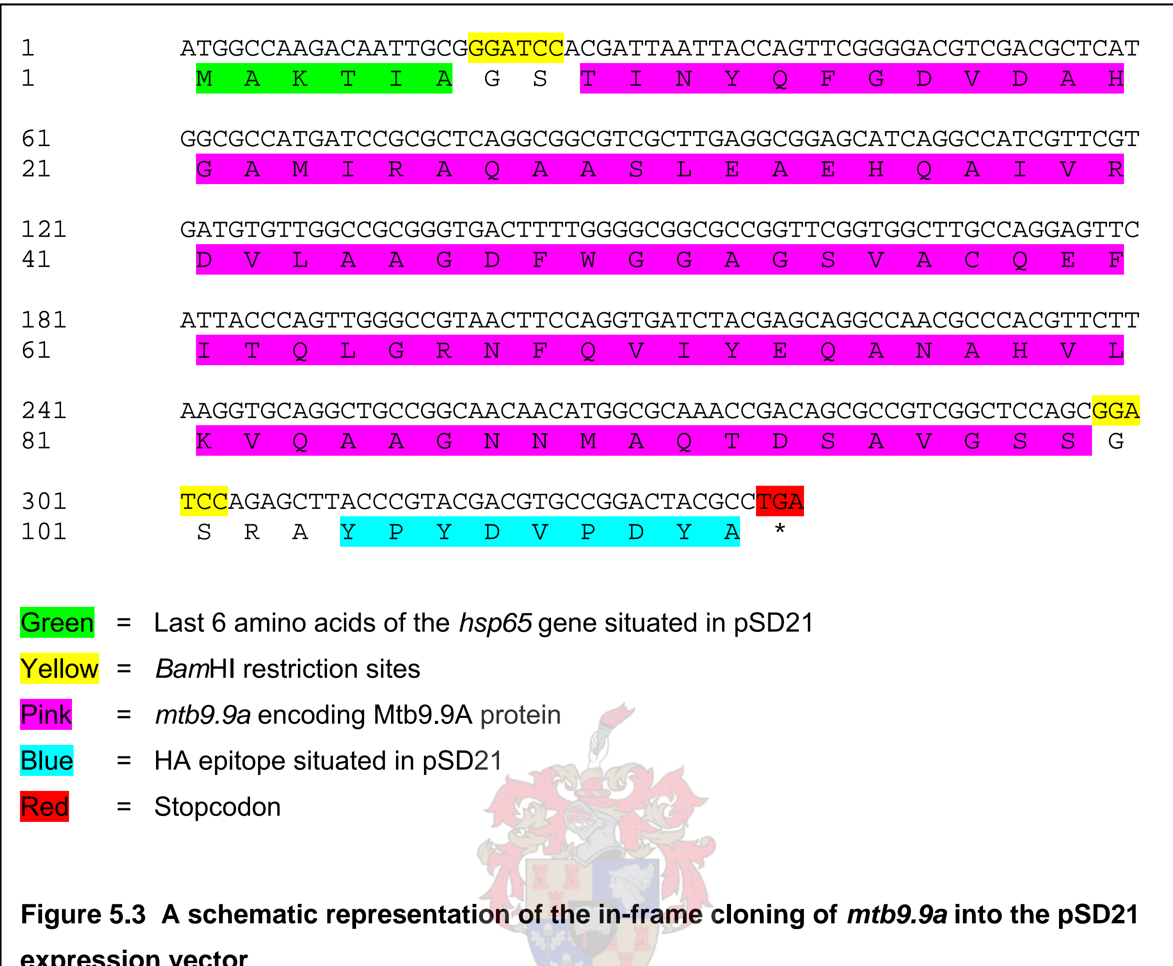


Figure 5.3 A schematic representation of the in-frame cloning of *mtb9.9a* into the pSD21 expression vector.

5.3.2 PCR amplification to confirm insert presence

Figure 5.4 is an illustration of the presence of Mtb9.9A, TB10.4 and ESAT-6 in the cultures that form part of the secretion analysis. *M. smegmatis* transformed with pSDMtb9.9A as well as *M. smegmatis* (with integrated ESAT-6 gene cluster region 5) transformed with pSDMtb9.9A , pMB154 and pSDTB10.4 were subjected to PCR analysis for insert confirmation. As seen on the illustration, Mtb9.9A, ESAT-6 and TB10.4 were present in these cultures.

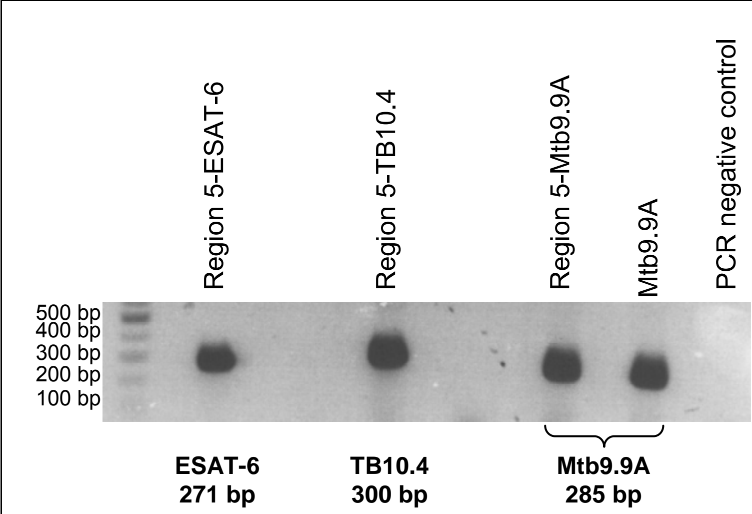


Figure 5.4 PCR analysis to confirm double transformations. PCR was performed on recombinant double transformed *M. smegmatis* containing both integrated ESAT-6 gene cluster region 5 and pMB154, pSDTB10.4 and pSDMtb9.9A, respectively. The second last lane is the PCR confirmation of *M. smegmatis* transformed with pSDMtb9.9A only. PCR was performed using primers as outlined in Table 5.1. Results showed the presence of all three genes within the recombinant *M. smegmatis* as well as within wild-type *M. smegmatis*, indicating the presence of the plasmids as were transformed.

Pectora roborant cultus recti

5.3.3 Secretion analyses

The results of the secretion analyses conducted during this study were inconclusive as no expression of Mtb9.9A could be detected in the samples tested (data not shown). This remained the case after repeated stripping and re-probing the Hybond P membrane, repeated SDS-PAGE and Western blot analyses and several changes in conditions and reagents. PCR- and sequence analyses proved that the Mtb9.9A insert was cloned in-frame of the pSD21 expression vector (see section 5.3.1), which excluded the possibility of experimental error with regards to cloning. After exhausting all options, we hypothesize that the absence of Mtb9.9A-HA protein in the tested samples is due to a lack of expression from the mycobacterial expression vector pSD21.

5.4 Discussion and Conclusion

Protein secretion forms a very important part of the virulence and pathogenicity of *M. tuberculosis.* The mycobacterium's potent T-cell antigens ESAT-6 and CFP-10 is translocated across the cell membrane during initial infection and thus plays a crucial role in the development of disease. These antigens are contained in the RD1 region (deleted in *M. bovis* BCG) and have been the subject of a number of studies focusing on diagnosis of *M. tuberculosis* infection, the search for efficient vaccine candidates and virulence (Brandt et al 2000, Wards et al 2000, Brodin et al 2004).

A large amount of research is now focusing on novel secretion mechanism(s) of *M. tuberculosis* as the ESAT-6 and CFP-10 antigens are secreted without the ordinary *Sec* secretion signals (Sorensen et al 1995 and Van Pinxteren et al 2000). This has led to the hypothesis of an alternative secretion/translocation system for *M. tuberculosis*. Studies now suggest that the whole ESAT-6 gene cluster region 1 is required for the secretion of the ESAT-6 antigen, indicating that the genes included in region 1 function together to form a multi-component secretion system (Brodin et al 2004, Converse et al 2004, Gey van Pittius PhD Thesis, Guinn et al 2004, Hsu et al 2003, Pym et al 2003, Stanley et al 2003). The ESAT-6 gene cluster secretion system of ESAT-6 gene cluster region 1 has recently been named the Snm secretion pathway (Stanley et al 2003). These authors have shown the involvement of Rv3870 (snm1), Rv3871 (snm2) and Rv3877 (snm3) in the system, and Pym et al 2003 showed that additional to the above named genes, Rv3876, is also essential for the secretion of the ESAT-6 antigen family.

M. smegmatis is a suitable host for secretion analyses as it contains the unique mycobacterial cell wall structure similar to *M. tuberculosis*. *M. smegmatis* is a fast-growing avirulent mycobacterium that is easy to transform with high efficiency which makes it very useful to study *M. tuberculosis* protein expression/secretion. *M. smegmatis*' suitability as a host for secretion analysis of the ESAT-6 gene cluster region 5 is extended by the fact that it only

contains ESAT-6 gene cluster regions 1, 3 and 4, and thus does not contain any copies of the ESAT-6 gene cluster region 5 (Rv1782-Rv1798; Gey van Pittius et al 2001).

During this study we aimed to prove that the entire ESAT-6 gene cluster region 5 is required for the secretion of its native ESAT-6, Mtb9.9A, by making use of the fast-growing avirulent *M. smegmatis* as a model organism. This study aimed to prove that, similar to what was observed for ESAT-6 gene cluster region 1, the genes included in region 5 also function together to form a multi-component secretion system. Unfortunately, the results of this study were inconclusive, as no expression of Mtb9.9A could be detected in all the samples tested. All experiments were repeated several times and cloning errors were excluded as the Mtb9.9A gene was cloned in-frame with the HA-epitope and the pSD21 expression vector. The absence of expression may be due to a number of factors, including regulation of expression within the pSD21 vector. We hypothesize that the *hsp65* promoter contained within pSD21 is very strong and possibly causes the transformed bacteria to shut down the transcription/translation machinery of the vector. Previous expression problems have subsequently been reported with the pSD21 expression vector in several laboratories (Dr NC Gey van Pittius, personal communication). Future studies with regards to secretion analysis should thus include using different expression vectors or using different techniques such as for example gene knockouts.

Pectora roborant cultus recti

CHAPTER SIX

CONCLUSION AND FUTURE DIRECTIONS

The distribution of the five duplicated ESAT-6 gene cluster regions in the genome of *M. tuberculosis* H37Rv has been described in detail by Gey van Pittius et al (2001). The gene clusters regions are designated region 1 (Rv3866-Rv3883c), region 2 (Rv3884c-Rv3895c), region 3 (Rv0282-Rv0292), region 4 (Rv3444c-Rv3450c) and region 5 (Rv1782-Rv1798) according to the numbering system used to classify the five mycosin (subtilisin-like serine protease) genes identified from these regions (Brown et al 2000).

One of the ESAT-6 gene cluster regions, region 1 (Rv3866-Rv3883c) contains the RD1 (region of difference 1; 9505 bp) first identified by Mahairas et al (1996). RD1 is hypothesized to be the primary deletion that occurred during the serial passage of *M. bovis* by Calmette and Guérin between 1908 and 1921 and also the possible reason of attenuation of *M. bovis* to *M. bovis BCG* (Brosch et al 2000 and Behr et al 1999). More recently, a study has also determined that deletion of RD1 from *M. tuberculosis* results in attenuation of virulence (Lewis et al 2003). This has led to a number of studies not only to determine the mechanism of attenuation of RD1 but also to elucidate the functions of the genes that comprise this region.

Pectora roborant cultus recti

RD1 encompasses the genes Rv3871 to Rv3879c (Cole et al 1998), which include the genes for the 6 kDa early-secreted antigenic target ESAT-6 (*esxA*) and L45 homologous protein CFP-10 (MTSA; *esxB*; Andersen et al 1995 and Berthet et al 1998). The genes *esxB* and *esxA* encode potent T-cell antigens that are secreted but lack detectable secretion signals (Sørensen et al 1995). These genes are also organised in an operon (Berthet et al 1998) and they form a tight 1:1 protein complex (Renshaw et al 2002 and Lightbody et al 2004).

During this study we looked at the ESAT-6 gene cluster region 1 by focusing on its regulation. This was a twofold analysis, firstly looking at its operonic nature and secondly identifying the promoters driving the expression of the cluster. Previous preliminary studies have indicated that the ESAT-6 gene cluster region 3 is expressed as a single polycistronic RNA, forming an operon structure, and is regulated by a single promoter designated P_{ESREG3} (Gey van

Pittius PhD Thesis 2002). The results of this study for ESAT-6 gene cluster region 1 indicated the existence of three promoters: P_{6263}, P_{6465} and P_{8483} and confirmed that region 1 is expressed as more than one polycistronic RNA and thus consists of multiple operons. These results are important for the clarification of the regulatory and secretory mechanisms of the ESAT-6 antigen family of *M. tuberculosis*.

Future studies could include analysis of the regulatory machinery of the rest of the ESAT-6 gene cluster regions, i.e. regions 2, 4 and 5. Once the promoters driving the expression of all of these regions are elucidated, the different promoter elements of all five ESAT-6 gene cluster regions can be compared under different conditions to further clarify the regulatory mechanisms and functions of the five gene clusters and the role these regions play in mycobacterial virulence.

Recent studies conducted by Guinn et al 2003, Hsu et al 2003, Stanley et al 2003, Brodin et al 2004, Flint et al 2004, Okkels et al 2004, Brodin et al 2005 and Converse et al 2005 indicate the existence of a secretion pathway (existing in the RD1 region of the ESAT-6 gene cluster region 1) where Rv3870 (Snm1), Rv3871 (Snm2), Rv3873, Rv3876 (Snm3) and Rv3877 (Snm4) all contribute to the secretion of the ESAT-6 antigen. Rv3870 and Rv3871 are hypothesized to work together to form a membrane-bound ATPase (Guinn et al 2003). Hsu et al 2003 showed by means of Western analysis of whole cell extracts and culture filtrate proteins that mutants of Rv3871 and Rv3876/77 are able to synthesize but not secrete ESAT-6 in culture filtrate. Stanley et al 2003 described the Snm pathway where ATPases Rv3870 (Snm1) and Rv3871 (Snm2) and the 12-transmembrane domain protein, Rv3877, encode components of a secretion system that work together to export ESAT-6 and CFP-10. Furthermore, it was concluded that the Snm secretion system is a major determinant of *M. tuberculosis* virulence, for mutants lacking either components or substrates are profoundly attenuated in a mouse model of infection. This reduced virulence displayed by the snm mutant cells during *in vivo* infection is hypothesized to be due to their inability to limit macrophage responses. It is also hypothesized that the Snm system functions to inhibit initial macrophage responses to infection *in vivo* and is also necessary for

normal growth kinetics during the early stages of infection with *M. tuberculosis*. Interestingly, with regards to Rv3873 and the role it plays in the secretion of ESAT-6 and CFP-10, Okkels et al 2004 demonstrated that Rv0288, encoded by *esx* cluster 3, also interacts with Rv3873 on protein level.

Following the confirmation of the secretion system of the ESAT-6 gene cluster region 1, we aimed to investigate whether the ESAT-6 gene cluster region 5 has similar secretion system functions. Phylogenetic relationships between the five duplicated ESAT-6 gene cluster regions in *M. tuberculosis* showed the following duplication events: Region 4 (ancestral region) > Region 1 > Region 3 > Region 2 > Region 5. This was done by analysing *esxA/esxB* duplication as well as separate *esxA* and *esxB* duplication (Gey van Pittius et al 2001) throughout the genome of *M. tuberculosis*. Region 5, being the last duplicate of the ESAT-6 gene cluster regions, is therefore an important region to investigate, as duplicated regions in between region 1 and 5 (regions 2 and 3) are likely to show the same secretory functions/mechanisms. Furthermore, Alderson et al (2002) has already shown that the region 5 ESAT-6 antigen, Mtb9.9A, is being secreted. Elucidating the manner by which this happened as well as comparing the secretory mechanism of the different ESAT-6 gene cluster regions will be important for understanding antigenic secretory mechanisms of *M. tuberculosis*.

Pectora roborant cultus recti

As a first step in elucidating the secretory mechanisms of the ESAT-6 antigen family, we focused on the ESAT-6 gene cluster region 5 and the region 5 ESAT-6 antigen, designated Mtb9.9A (*esxN*) by Alderson et al (2000). Alderson and coworkers (2000) showed that 83% of PPD$^+$ donors gave a significant response to rMtb9.9A, indicating that it was an efficiently secreted protein. The gene (Rv1793 or *esxN*) encoding this Mtb9.9 family member is situated in the ESAT-6 gene cluster region 5 and is thus most probably secreted by the ESAT-6 gene cluster region 5 transporter apparatus. During this study we purified and isolated the Mtb9.9A protein successfully as a GST fusion protein and generated antibodies against it. This was done to generate a tool for further secretion analyses. Due to the problems encountered with expression of Mtb9.9A in *M. smegmatis*, we were unable to conclude whether the ESAT-6 gene cluster

region 5 is responsible for the secretion of the Mtb9.9A antigen. The experimental difficulties encountered during the last part of this study may potentially be circumvented in future studies by using different vector systems or other techniques to study the secretion system. Such future directions could entail specific gene knockouts of components of the ESAT-6 gene cluster regions in order to analyse which genes associate or function together for the secretion of the ESAT-6 antigen family. Furthermore, each member of the ESAT-6 antigen family could be purified and used in immunology studies and subsequently in different combinations to form part of vaccine studies.

In conclusion, this study has made an important contribution towards the understanding of mycobacterial antigen secretion systems. We focused on the isolation, regulation and secretion of the ESAT-6 gene cluster regions 1 and 5 and the genes contained within these regions. We have identified three promoters which are involved in the regulation of the pathogenically-important ESAT-6 gene cluster region 1, and have proven that the region consists of multiple operons. We have cloned, expressed and purified another potent and important mycobacterial antigen from the ESAT-6 gene family, namely Mtb9.9A, and generated antibodies against it. We have also isolated the ESAT-6 gene cluster regions 2, 4 and 5, and used region 5 in preliminary secretion analysis studies. This study opens the way to future investigations that will form the basis for improving the understanding of the pathogenicity of *M. tuberculosis*.

Pectora roborant cultus recti

Addendum

Preparation of electro-competent cells

E. coli XL-1 Blue

Electrocompetent XL-1 Blue *E.coli* cells from a frozen glycerol stock culture were streaked onto an LB-agar plate containing tetracylcline (10µl/ml of a 5 mg/ml stock in ethanol). This was incubated overnight at 37ºC. A single colony was picked and inoculated into a Universal tube with 15 ml of LB-media containing tetracycline (5 mg/ml) for a starter culture. This was incubated overnight at 37ºC with shaking. AT the same time, 4 flasks with 250 ml LB-media containing tetracycline, were pre-warmed overnight at 37ºC with shaking. The 4 x 250 ml LB-media contained in flasks were inoculated with 1:100 of the starter culture – therefore, 2.5 ml of starter culture for each flask with 250 ml LB-media containing tetracycline. These were incubated 3-4 hours at 37ºC, with shaking, until the OD_{600} was between 0.6 and 0.8. The flasks were put on ice and all subsequent steps were performed at 4ºC. The cells were harvested by centrifugation in a GSA rotor at 5000 rpm for 10 minutes. The supernatant was removed and the pellets were carefully re-suspended in a volume of ice-cold 10% glycerol equal to the original culture volume. The bacterial cells were re-pelleted as described above and the supernatant was removed. The pellets were re-suspended again in ice-cold 10% glycerol equal to the original culture volume. Cells were again centrifuged under conditions described above. The supernatant was carefully removed and cells were re-suspended in the remaining glycerol in the centrifuge tubes. Cells were pooled and centrifuged at 7000 rpm for 10 minutes in a SS34 rotor. The supernatant was again carefully removed and bacterial cells were re-suspended in 10% glycerol using a volume of 2 ml/l of the initial culture. This was divided into aliquots of 100 µl in 1.5 ml microcentrifuge tubes and flash-frozen in liquid nitrogen. These were stored at -80ºC.

M. smegmatis mc^{2}155

An inoculate from a freeze culture containing electrocompetent *M. smegmatis* mc^{2}155 was inoculated into 10 ml Middlebrook 7H9 broth (Difco) supplemented with filter-sterile ADC (albumin-dextrose-catalase- 0.5% BSA, 0.2% glucose, 0.015% catalase). This served as the starter culture and was allowed to grow for 2 days at 37°C while shaking. A 1:200 inoculum from the starter culture was inoculated into 100 ml Middlebrook 7H9 broth (Difco) supplemented with filter-sterile ADC (albumin-dextrose-catalase- 0.5% BSA, 0.2% glucose, 0.015% catalase) and containing 0.05% Tween 80 (Sigma). This culture was allowed to grow at 37°C (while shaking) until a spectrophotometric reading gave an OD_{600} reading of ± 0.5. The culture was put on ice for about an hour and thereafter divided into two 50 ml cultures. These were then centrifuged at 4°C for 5 minutes at 5000 rpm. The resulting pellets were washed in 30 ml ice-cold, sterile 10% glycerol. The centrifuge and washing steps were repeated 4 times. The pellets were pooled and resuspended in 2 ml ice-cold, sterile 10% glycerol. This was kept on ice until electro-transformation.

Pectora roborant cultus recti

Electro-transformation

E. coli XL-1 Blue

The ligation reaction, competent cells and sterile 2 mm electroporation cuvettes were all pre-chilled on ice. 2 µl of ligation reaction was added to 50 µl of thawed competent cells, gently mixed and left on ice for 1 minute. The ligation-competent cell mix was then added to the pre-chilled cuvette and the cuvette and electroporation chamber slide were dried carefully. The Gene Pulser was preset to a voltage of 2.5 kV, a capacitance of 25 µF and a resistance of 200 ohm. The cuvette was placed into the slide, pushed into the chamber and pulsed once. The most favourable time constant expected is 3.9. After the electroporation, 1 ml of freshly prepared SOC was added to the contents in the cuvette and mixed. The transformed cells were then transferred

to a sterile 1.5 ml sterile microcentrifuge tube and incubated at 37ºC for one hour with shaking (200 rpm). One-hundred to two-hundred microlitres of cells were plated onto X-gal plates (LB-agar; autoclaved with added 0.2M IPTG, 0.02 M X-gal, 50 mM Ampicillin: in the case of pGem-T Easy and pBluescript II KS+ or 50mM Kanamycin: in the case of pJem15). These plates were incubated overnight at 37ºC

M. smegmatis mc^2155

The ligation reaction, competent cells and sterile 2 mm electroporation cuvettes were all pre-chilled on ice. 1 µl of ligation reaction was added to 200 µl of electro-competent cells, gently mixed and left on ice for 1 minute. The ligation-competent cell mix was then added to the pre-chilled cuvette and the cuvette and electroporation chamber slide were dried carefully. The Gene Pulser was preset to a voltage of 2.5 kV, a capacitance of 125 µF and a resistance of 1000 ohm. The cuvette was placed into the slide, pushed into the chamber and pulsed once. The most favourable time constant expected is between 13 and 18. After the electroporation, 1 ml Middlebrook 7H9 broth (Difco) supplemented with filter-sterile ADC (albumin-dextrose-catalase-0.5% BSA, 0.2% glucose, 0.015% catalase) was added to the contents in the cuvette and mixed. The transformed cell-mixtures were then transferred to a sterile 2 ml sterile microcentrifuge tube and incubated at 37ºC for 3 hours with shaking (200 rpm). One-hundred to two-hundred microlitres of cells were plated onto Middlebrook 7H11 agar supplemented with filter-sterile OADC (0.005% oleic acid, 0.5% BSA, 0.2% glucose, 0.02% catalase, 0.085% NaCl) also containing the necessary antibiotic(s). These plates with cells were incubated overnight at 37ºC and left for about 2 days until colonies were visible.

Pectora roborant cultus recti

Solutions and Buffers

Colony blotting

1. Denaturing solution
   - 1.5 M NaCl, 0.5 M NaOH
2. 2 x SSC
3. 10% SDS
4. Prewashing solution
   - 5 x SSC, 0.5% SDS, 1 mM EDTA (pH 8.0)
5. Washing buffer
   - 2 x SSC, 0.1% SDS
6. Neutralizing solution
   - 1.5 M NaCl, 0.5 M Tris-HCl (pH 7.2), 0.001 M Na_2EDTA
7. Prehybridization solution
   - 500 ml Formamide, 250 ml 20x SSPE, 20 ml 10% SDS, 20 ml Herring Sperm DNA (10 mg/ml), 50 ml 100x Denhardt's Solution

Pectora roborant cultus recti

List of References

A

Agger EM and Andersen P (2002) A novel TB vaccine; towards a strategy bsed on our understanding of BCG failure. *Vaccine* 21: 7-14.

Ahmad S, Amoudy HA, Thole JE, Young DB and Mustafa AS (1999) Identification of a novel protein antigen encoded by a *Mycobacterium tuberculosis*-specific RD1 region gene. *Scandinavian Journal of Immunology* 49: 515-522.

Alderson MR, Bement T, Day CH, Zhu L, Molesh D, Skeiky YA, Coler R, Lewinsohn DM, Reed SG and Dillon DC (2000) Expression cloning of an immunodominant family of *Mycobacterium tuberculosis* antigens using human $CD4^{+}$ T cells. *Jounal of Experimental Medicine* 191: 551-560.

Andersen P, Andersen AB, Sorensen AL and Nagai S (1995) Recall of long-lived immunity to *Mycobacterium tuberculosis* infection in mice. *The Journal of Immunology* 154: 3359-3372.

Andersen P (2001) TB vaccines: progress and problems. *TRENDS in Immunology* 22: 160-168.

Arend SM, Geluk A, van Meijgaarden KE, van Dissel JT, Theisen M, Andersen P and Ottenhoff TH (2000) Antigenic equivalence of human T-cell responses to *Mycobacterium tuberculosis*-specific RD1-encoded protein antigens ESAT-6 and culture filtrate protein 10 and to mixtures of synthetic peptides. *Infection and Immunity* 68: 3314-3321.

B

Bange FC, Collins FM and Jacobs WR Jr (1999) Survival of mice infected with *Mycobacterium smegmatis* containing large DNA fragments from *Mycobacterium tuberculosis*. *Tubercle and Lung Disease* 79: 171-180.

Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284: 1520-1523.

Bellstedt DU, Human PA, Rowland GF and Van der Merwe KJ (1987) Acid-treated, naked bacteria as immune carriers for protein antigens. *Journal of Immunological Methods* 98: 249-255.

Berthet FX, Rasmussen PB, Rosenkrands I, Andersen P and Gicquel B (1998) A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low- molecular-mass culture filtrate protein (CFP-10). *Microbiology* 144: 3195-3203.

Bloom BR and Murray CJ (1992) Tuberculosis: commentary on a re-emergent killer. *Science* 257: 1055-1064.

Brandt L, Elhay M, Rosenkrands I, Lindblad EB, Andersen P (2000) ESAT-6 subunit vaccination against Mycobacterium tuberculosis. *Infection and Immunity* 68: 791-795.

Brandt L, Feino CJ, Weinreich OA, Chilima B, Hirsch P, Appelberg R, Andersen P (2002) Failure of the *Mycobacterium bovis* BCG vaccine: some species of environmental mycobacteria block multiplication of BCG and induction of protective immunity to tuberculosis. *Infection and Immunity* 70: 672-678.

Pectora roborant cultus recti

Brodin P, Rosenkrands I, Andersen P, Cole ST and Brosch R (2004) ESAT-6 proteins: protective antigens and virulence factors? *TRENDS in Microbiology* 12 (nr 11): 500-508.

Brodin P, de Jonge MI, Majlessi L, Leclerc C, Nilges M, Cole ST and Brosch R (2005) Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of *Mycobacterium tuberculosis*, reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. *Journal of Biological Chemistry* 280(40): 33953–33959.

Brosch R, Gordon SV, Buchrieser C, Pym AS, Garnier T, Cole ST (2000) Comparative genomics uncovers large tandem chromosomal duplications in *Mycobacterium bovis* BCG Pasteur. *Yeast* 17: 111-123.

Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser P, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, Van Soolingen D and Cole ST (2002) A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *PNAS* 99 (no 6): 3684-3689.

Brown GD, Dave JA, Gey van Pittius NC, Stevens L, Ehlers MR, Beyers AD (2000) The mycosins of *Mycobacterium tuberculosis* H37Rv: a family of subtilisin-like serine proteases. *Gene* 254: 147-155.

Brusasca PN, Colangeli R, Lyashchenko KP, Zhao X, Vogelstein M, Spencer JS, McMurray DN and Gennaro ML (2001) Immunological characterization of antigens encoded by the RD1 region of the *Mycobacterium tuberculosis* genome. *Scandinavian Journal of Immunology* 54: 448-452.

Pectora roborant cultus recti

C

Colangeli R, Spencer JS, Bifani P, Williams A, Lyashchenko K, Keen MA, Hill PJ, Belisle J, and Gennaro ML (2000) MTSA-10, the product of the Rv3874 gene of *Mycobacterium tuberculosis*, elicits tuberculosis-specific, delayed-type hypersensitivity in guinea pigs. *Infection and Immunity* 68: 990-993.

Colditz GA, Brewer TF, Berkey CS, Wilson ME, Burdick E, Fineberg HV and Mosteller F (1994) Efficacy of BCG vaccine in the prevention of tuberculosis. Meta- analysis of the published literature. *JAMA* 271, 698-702.

Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, III, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K and Barrell BG (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393: 537-544.

Converse SE and Cox JS (2005) A Protein Secretion Pathway Critical for *Mycobacterium tuberculosis* virulence is conserved and functional in *Mycobacterium smegmatis. Journal of Bacteriology* 187(4): 1238–1245

D

Dillon DC, Alderson MR, Day CH, Bement T, Campos-Neto A, Skeiky YA, Vedvick T, Badaro R, Reed SG and Houghton R (2000) Molecular and immunological characterization of *Mycobacterium tuberculosis* CFP-1, an immunodiagnostic antigen missing in *Mycobacterium bovis* BCG. *Journal of Clinical Microbiology* 38: 3285-3290.

Pectora roborant cultus recti

Daugelat S, Kowall J, Mattow J, Bumann D, Winter R, Hurwitz R and Kaufmann SH (2003) The RD1 proteins of *Mycobacterium tuberculosis*: expression in Mycobacterium smegmatis and biochemical characterization. *Microbes and Infection* 5(12): 1082-1095.

F

Fine PEM (1995) Variation in protection by BCG: implications of and for heterologous immunity. *The Lancet* 346: 1339-1345.

Flint JL, Kowalski JC, Karnati PK and Derbyshire KM (2004) The RD1 virulence locus of *Mycobacterium tuberculosis* regulates DNA transfer in *Mycobacterium smegmatis*. *PNAS* 101 (no 34): 12598-12603.

G

Gey van Pittius NC, Gamieldien J, Hide W, Brown GD, Siezen RJ, Beyers AD (2001) The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. *Genome Biology* 2 (no 10): RESEARCH0044.

Gey van Pittius NC (2002) PhD Thesis: The Mycosins, a family of secreted subtilisin-like serine proteases associated with the immunologically-important ESAT-6 gene clusters of *Mycobacterium tuberculosis.*

Gordon SV, Eiglmeier K, Garnier T, Brosch R, Parkhill J, Barrell B, Cole ST and Hewinson RG (2001) Genomics of *Mycobacterium bovis*. *Tuberculosis (Edinb.)* 81: 157-163.

Guinn KM, Hickey MJ, Mathur SK, Zakel KL, Grotzke JE, Lewinsohn DM, Smith S and Sherman DR (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Molecular Microbiology* 51(2): 359-370.

Pectora roborant cultus recti

H

Herzog H (1998) History of Tuberculosis. *Respiration* 65: 5-15.

Hsu T, Hingley-Wilson SM, Chen B, Chen M, Dai AZ, Morin PM, Marks CB, Padiyar J, Goulding C, Gingery M, Eisenberg D, Russell RG, Derrick SC, Collins FM, Morris SL, King CH and Jacobs WR Jr (2003) The primary mechanism of attenuation of bacillus Calmette-Guérin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *PNAS* 100: 12420-12425.

J

Jacobs WR Jr, Kalpana GV, Cirillo JD, Pascopella L, Snapper SB, Udani RA, Jones W, Barletta RG and Bloom BR (1991) Genetic systems for mycobacteria. *Methods in Enzymology* 204, 537-555.

L

Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA and Sherman DR (2003) Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacilli Calmette-Guerin attenuation. *Journal of Infectious Diseases* 187: 117-123.

Lightbody KL, Renshaw PS, Collins ML, Wright RL, Hunt DM, Gordon SV, Hewinson RG, Buxton RS, Williamson RA and Carr MD (2004) Characterisation of complex formation between members of the *Mycobacterium tuberculosis* complex CFP-10/ESAT-6 protein family: towards an understanding of the rules governing complex formation and thereby functional flexibility. *FEMS Microbiology Letters* 238: 255-262.

Pectora roborant cultus recti

M

Mahairas GG, Sabo PJ, Hickey MJ, Singh DC and Stover CK (1996) Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *Journal of Bacteriology* 178: 1274-1282.

McKinney JD, Jacobs WR Jr and Bloom BR (1998) Persisting problems in tuberculosis. In 'Emerging Infections'. pp. 51-146. (Academic Press)

Murray A, Winter N, Lagranderie M, Hill DF, Rauzier J, Timm J, Leclerc C, Moriarty KM, Gheorghiu M, Gicquel B (1992) Expression of *Escherichia coli* beta-galactosidase in *Mycobacterium bovis* BCG using an expression system isolated from *Mycobacterium paratuberculosis* which induced humoral and cellular immune responses. *Molecular Microbioly* 6: 3331-3342.

Murray JF (2004) A century of Tuberculosis. *American Journal of Respiratory and Critical Care Medicine* 169: 1181-1186.

Mustafa AS, Cockle PJ, Shaban F, Hewinosn RG and Vordermeier HM (2002) Immunogicity of *Mycobacterium tuberculosis* RD1 region gene products in infected cattle. *Clinical and Experimental Immunology* 130: 37-42.

Mustafa AS, Shaban FA, Al-Attiyah R, Abal AT, El-Shamy AM, Andersen P and Oftung F (2003) Human Th1 cell lines recognize the *Mycobacterium tuberculosis* ESAT-6 antigen and its peptides in association with frequently expressed HLA Class II molecules. *Scandinavian Journal of Immunology* 57:125-134.

Pectora roborant cultus recti

O

Okkels LM and Andersen P (2004) Protein-protein interactions of proteins from the ESAT-6 family of *Mycobacterium tuberculosis. Journal of Bacteriology* 186(8): 2487-2491.

Olsen AW, Hansen PR, Holm A, Andersen P (2000) Efficient protection against *Mycobacterium tuberculosis* by vaccination with a single subdominant epitope from the ESAT-6 antigen. *European Journal of Immunology* 30: 1724-1732.

P

Pallen MJ (2002) The ESAT-6/WXG100 superfamily – and a new Gram-positive secretion system? *TRENDS in Microbiology* 10(5): 209-212.

Pardee AB, Jacob F and Monod J (1959) The genetic control and cytoplasmic expression of “inducibility” in the synthesis of β-galactosidase by *E. coli*. *Journal of Molecular Biology* 1: 165-178.

Prescott LM, Harley JP and Klein DA (2002) Microbiology 5th edition. *McGraw-Hill Companies Inc, New York*

Pym AS, Brodin P, Brosch R, Huerre M and Cole ST (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Molecular Microbiology* 46: 709-717

Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, Williams A, Griffiths KE, Marchal G, Leclerc C and Cole ST (2003) Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nature Medicine* 9(5): 533-539.

Pectora roborant cultus recti

R

Ravn P, Demissie A, Eguale T, Wondwosson H, Lein D, Amoudy HA, Mustafa AS, Jensen AK, Holm A, Rosenkrands I, Oftung F, Olobo J, von Reyn F, Andersen P (1999) Human T cell responses to the ESAT-6 antigen from *Mycobacterium tuberculosis*. *Journal of Infectious Diseases* 179: 637-645.

Renshaw PS, Panagiotidou P, Whelan A, Gordon SV, Hewinson RG, Williamson RA and Carr MD (2002) Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6·CFP-10 complex. *Journal of Biological Chemistry* 277(24): 21598-21603.

S

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (2nd edition) New York: Cold Spring Harbour.

Schagger H and von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* 166: 368-379.

Selwyn PA, Hartel D, Lewis VA, Schoenbaum EE, Vermund SH, Klein RS, Walter AT and Friedland GH (1989) A prospective study of the risk of tuberculosis among intravenous drug users with human immunodeficiency virus infection. *New England Journal of Medicine* 320: 545-550.

Pectora roborant cultus recti

Shin S, Furin J, Bayona J, Mate K, Kim JY and Farmer P (2004) Community-based treatment of multidrug-resistant tuberculosis in Lima, Peru: 7 years of experience. *Social Science & Medicine* 59:1529-1539.

Skjøt RLV, Agger EM and Andersen P (2001) Antigen discovery and tuberculosis vaccine development in the post-genomic era. *Scandinavian Journal of Infectious Diseases* 33: 643-647.

Smida J, Kazda J and Stackebrandt E (1988) Molecular-genetic evidence for the relationship of *Mycobacterium leprae* to slow-growing pathogenic mycobacteria. *Internaitonal Journal of Leprosy and Other Mycobacterial Diseases* 56(3): 449-454.

Smith I (2003) *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clinical Microbiology Reviews* 16(3): 463-496.

Snapper SB, Melton RE, Mustafa S, Kieser T and Jacobs WR Jr (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Molecular Microbiology* 4: 1911-1919.

Sørensen AL, Nagai S, Houen G, Andersen P, Andersen AB (1995) Purification and Characterization of A Low-molecular-mass T-Cell antigen secreted by *Mycobacterium tuberculosis*. *Infection and Immunity* 63: 1710-1717.

Stanley SA, Raghavan S, Hwang WW and Cox JS (2003) Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. *PNAS* 100(22): 13001-13006.

Stewart GR, Robertson BD and Young DB (2003) Tuberculosis: a problem with persistence. *Nature Reviews Microbiology* 1:97-105.

Pectora roborant cultus recti

T

Tekaia F, Gordon SV, Garnier T, Brosch R, Barrell BG, Cole ST (1999) Analysis of the proteome of *Mycobacterium tuberculosis in silico*. *Tubercle and Lung Disease* 79: 329-342.

Timm J, Lim EM, Gicquel B (1994) *Escherichia coli*-mycobacteria shuttle vectors for operon and gene fusions to lacZ: the pJEM series. *Journal of Bacteriology* 176: 6749-6753.

V

Van Crevel R, Ottenfhoff THM and Van der Meer JWM (2002) Innate Immunity to *Mycobacterium tuberculosis*. *Clinical Microbiology Reviews* 15(2): 294-309.

Van Pinxteren LA, Ravn P, Agger EM, Pollock J and Andersen P (2000) Diagnosis of tuberculosis based on the two specific antigens ESAT-6 and CFP10. *Clinical and Diagnostic Laboratory Immunology* 7: 155-160.

Pectora roborant cultus recti

W

WHO Fact Sheet N 104 (2005) http://www.who.int/mediacentre/factsheets/fs104/en/

Wards BJ, de Lisle GW and Collins DM (2000) An *esat6* knockout mutant of *Mycobacterium bovis* produced by homologous recombination will contribute to the development of a live tuberculosis vaccine. *Tubercle and Lung Disease:* 80, 185-189.

Y

Young DB (2003) Building a better tuberculosis vaccine. *Nature Medicine*: Advance online publication.