# Investigating the localisation of the ESX-3 secretion system in *Mycobacterium smegmatis*

by
Nastassja Lise Steyn

*Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Medical Sciences (Molecular Biology) in the Faculty of Medicine and Health Sciences at Stellenbosch University*

Supervisor: Prof NC Gey van Pittius
Co-supervisor: Prof RM Warren

December 2012

**Declaration**

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own and original work, that I am the authorship owner thereof (unless to the extent of explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Nastassja Lise Steyn

September 2012

**Abstract**

*Mycobacterium tuberculosis* is a pathogenic organism that infects a third of the world's population and causes approximately 2 million deaths per year. Extensive research has been done on this pathogen, however our knowledge of the mechanisms of pathogenicity remain limited. The *M. tuberculosis* genome contains five ESAT-6 gene cluster regions, ESX-1 to 5, which encode specialized type VII secretion systems. These secretion systems are known to secrete members of the ESAT-6/CFP-10 and PE/PPE protein families, some of which contribute to the pathogenicity and phagosomal escape of the pathogen. ESX-3 has been shown to be essential for *in vitro* growth and survival of *M. tuberculosis.* The expression of ESX-3 in *M. tuberculosis* is regulated by IdeR and Zur, in response to intracellular iron and zinc concentrations, respectively. Interestingly, ESX-3 is not essential for the growth and survival of the saprophytic organism *M. smegmatis.*

In this study, we aimed to identify the subcellular localisation of the individual components of the ESX-3 secretion system in the non-pathogenic, fast-growing organism *M. smegmatis*. The esx conserved component (ecc) genes from ESX-3 were expressed from the episomal expression vector pDMNI as fusion proteins with green fluorescent protein (GFP). MSMEG_0615 (eccA$_3$), MSMEG_0616 (eccB$_3$), MSMEG_0623 (eccD$_3$) and MSMEG_0626 (eccE$_3$) were successfully cloned into pDMNI and expression of fusion proteins was confirmed by Western blotting for MSMEG_0615-GFP, MSMEG_0616-GFP and MSMEG_0626-GFP in *M. smegmatis.* In the *M. smegmatis* ESX-3 knock-out (with MSMEG_0615 to MSMEG_0626 deleted) expression was confirmed for MSMEG_0615-GFP and MSMEG0626-GFP. Fluorescent microscopy determined that MSMEG_0615-GFP localised to a single mycobacterial pole in both strains. MSMEG_0616-GFP and MSMEG_0626-GFP were found to be membrane associated in *M. smegmatis*, while MSMEG_0626-GFP was found to be membrane associated in the *M. smegmatis* ESX-3 knock-out.

The unipolar localisation of MSMEG_0615-GFP suggests that the assembled ESX-3 secretion system apparatus is situated at a single pole in *M. smegmatis*. Therefore, we hypothesize that MSMEG_0615 might act as a recruiter protein that is involved in the assembly of ESX-3 at the mycobacterial pole.

## Opsomming

*Mycobacterium tuberculosis* is 'n patogene organisme wat 'n derde van die wêreld se bevolking infekteer en eis jaarliks 2 miljoen lewens deur tuberkulose. Ten spyte van uitgebreide navorsing, is daar min kennis oor die meganismes van patogenisiteit van hierdie organisme. Die *M. tuberculosis* genoom bevat vyf duplikasies van die ESAT-6 geen groep gebiede, ESX-1 tot 5, wat kodeer vir gespesialiseerde Tipe VII sekresie sisteme. Hierdie sekresie sisteme is bekend vir die sekresie van lede van die ESAT-6/CFP-10 en PE/PPE proteïen families, waarvan sommige bydra tot die patogenisieit en fagosomale ontsnapping van hierdie organisme. ESX-3 is noodsaaklik vir die *in vitro* groei en oorlewing van *M. tuberculosis.* Die uitdrukking van ESX-3 in *M. tuberculosis* word gereguleer deur IdeR en Zur in reaksie op intrasellulêre yster en sink konsentrasies, onderskeidelik. ESX-3 word nie benodig vir die groei en oorlewing van die saprofitiese organisme *M. smegmatis* nie.

Hierdie studie was gemik om die sub-sellulêre lokalisering van ESX-3 te identifiseer in die nie-patogeniese en vinnig-groeiende organisme, *M. smegmatis*. Die "esx conserved component" (ecc) gene van ESX-3 is uitgedruk vanaf die episomale uitdrukkingsvektor pDMNI as gekombineerde proteïene met die groen fluoreserende proteïen (GFP). MSMEG_0615 (eccA$_3$), MSMEG_0616 (eccB$_3$), MSMEG_0623 (eccD$_3$) en MSMEG_0626 (eccE$_3$) is suksesvol gekloneer en die uitdrukking van die gekombineerde proteïene is bevestig deur Western oordrag vir MSMEG_0615-GFP, MSMEG_0616-GFP en MSMEG_0626-GFP in *M. smegmatis.* In die *M. smegmatis* ESX-3 uitklopmutant (met MSMEG_0615 tot MSMEG_0626 uitgeslaan) is uitdrukking bevestig vir MSMEG_0615-GFP en MSMEG0626-GFP. Fluoresensie mikroskopie het bepaal dat MSMEG_0615-GFP gelokaliseer is by 'n enkele mikobakteriese pool in beide stamme. MSMEG_0616-GFP en MSMEG_0626-GFP was membraan-geassosieerd in *M. smegmatis,* terwyl en MSMEG_0626-GFP geassosieer het met die membraan in die *M. smegmatis* uitklopmutant.

MSMEG_0615 het gelokaliseer by 'n enkele pool in *M. smegmatis* en dit dui aan dat die saamgestelde ESX-3 sekresie sisteem apparaat slegs by 'n enkele pool voorkom in *M. smegmatis.* Ons hipotiseer dat MSMEG_0615 dalk mag optree as 'n werwer proteïen wat betrokke is by die samestelling van die ESX-3 sekresie sisteem by die mikrobakteriese pool.

**Contents Page**

## Acknowledgements

There are many people to thank for the work that will be presented in this thesis. Firstly, I would like to thank Mae Newton-Foot for all the help and guidance that she has given me the past two years. You were my go to person, thank you very much. I would also like to thank my supervisor, Nico and my co-supervisor, Rob for all the guidance and help that you have given me. I would like to thank the department, lab 424 and the University of Stellenbosch for allowing me to continue my studies and giving me the resources that were needed to be able to complete this project.

A special thanks to Lize Engelbrecht (CAF – Stellenbosch University) and Mohamed Jaffer (Electron microscopy unit – Cape Town University) for the help with the fluorescent microscopy and electron microscopy, respectively.

My friends have helped me both emotionally and academically over the duration of this project and I would like to thank each and every one of you: Mae, Lani, Lizaan, Louise, Anzaan, Zhuo and Keith. You guys are amazing. I would like to thank my family, I would not have been able to do this without your continued support.

Lastly, I would like to thank Christopher. You know how much you help me every day. Thank you.

## List of Abbreviations

| | |
|---|---|
| °C | degrees Celcius |
| µF | microFarad |
| µFd | microFaraday |
| µg | microgram |
| µl | microliter |
| 7H11 | BBL$^{TM}$ Seven H11 Agar Base |
| 7H9 | Difco$^{TM}$ Middlebrook 7H9 Broth |
| amp | ampicillin |
| Amp$^R$ | Ampicillin resistance |
| *aph* | kanamycin resistance gene |
| APS | ammonium persulphate |
| Arg | Arginine |
| Asp | Aspartic acid |
| ATP | Adenosine triphosphate |
| BCG | Bacille Calmette at Guérin |
| CFP-10 | culture filtrate protein 10 |
| DNA | Deoxyribonucleic acid |
| DtxR | Diphtheria toxin repressor (protein family) |
| *E. coli* | *Escherichia coli* |
| *ecc* | esx conserved components |
| EDTA | Ethylenediamine tetra acetic acid |
| ESAT-6 | early secretory antigenic target of 6 kDa |
| *esp* | ESX-1 secretion associated protein |
| ESX | ESAT-6 secretion system |
| EsxA$_{ms}$ | ESAT-6 from ESX-1 in *M. smegmatis* |
| EsxB$_{ms}$ | CFP-10 from ESX-1 in *M. smegmatis* |
| *et al.* | and others |
| Fur | Ferric uptake regulator |
| GC | Guanine and Cytosine |
| GFP | green fluorescent protein |
| GTP | Guanine triphosphate |
| His | Histidine |
| HRP | horseradish peroxidase |
| IdeR | iron dependent repressor |
| IKE | Immune killing evasion vaccine strain |
| IKEPLUS | Immune killing evasion PLUS vaccine strain |

| | |
|---|---|
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| kan | kanamycin |
| Kan$^R$ | Kanamycin resistance |
| kb | kilobases |
| KCl | potassium chloride |
| kDa | kiloDalton |
| kV | kiloVolt |
| *lacZ* | β-galactosidase gene |
| LB | Luria-Bertani broth |
| Lys | Lysine |
| *M.* | *Mycobacterium* |
| MgCl$_2$ | Magnesium Chloride |
| ml | millliliter |
| MycP | mycosin protease |
| ng | nanogram |
| NTP | nucleotide triphosphate |
| OD | optical density |
| *oriE* | Origin of replication for *E.coli* |
| *oriM* | Mycobacterial origin of replication |
| PBS | phosphate buffered saline |
| PCR | polymerase chain reaction |
| PE | proline-glutamic acid (mycobacterial protein family) |
| PPE | proline-proline-glutamic acid (mycobacterial protein family) |
| Pro | Proline |
| P$_{smyc}$ | promoter for mycobacteria |
| RBS | ribosome binding site |
| RD | Region of Difference |
| repA | replication protein A |
| repB | replication protein B |
| RNA | ribonucleic acid |
| rpm | revolutions per minute |
| rrnB | transcription terminator |
| SDS | sodium dodecyl sulphate |
| SDS-PAGE | sodium dodecyl sulphate - polyacrylamide gel electrophoresis |
| Sec | general secretion pathway |
| Ser | Serine |
| SirA | Iron regulated ABC transporter siderophore-binding protein |

| SOC | super optimal catabolate repression broth |
| T7SS | Type-VII secretion system |
| TAE | tris-acetic acid-EDTA buffer |
| Tat | Twin-arginine translocation pathway |
| TEM | Transmission Electron Microscopy |
| TEMED | N,N,N',N'-tetramethylethylenediamine |
| tet | tetracycline |
| Thr | Threonine |
| Tm | annealing temperature |
| tris | tris(hydroxymethyl)aminomethane |
| Tween-80 | polyoxyethylene sorbian monooleate |
| U | Units |
| UV | ultraviolet |
| V | Volt |
| WCL | whole cell lysate |
| WT | wild-type |
| WXG | tryptophan-X-glycine (mycobacterial protein family) |
| X | Variable amino acid |
| X-gal | 5-bromo-4-chloro-3-indolyl-$\beta$-D-galactopyranoside |
| YFP | Yellow fluorescent protein |
| ZN | Ziehl-Neelsen |
| Zur | zinc uptake regulator |
| $\beta$ | Beta |
| $\Omega$ | Ohm |

**List of Tables**

## List of Figures

# **Problem Statement**

*M. tuberculosis* is the causative agent of tuberculosis, a disease that infects approximately a third of the world's population and results in millions of deaths each year. Despite extensive research over decades the mechanism of pathogenicity still remains unclear. *M. tuberculosis* contains five ESAT-6 gene cluster (ESX) regions within its genome, each of which is believed to encode a type VII secretion system that is responsible for the secretion of proteins from within the cell. These secretion systems have been shown to be important for the virulence and survival *M. tuberculosis*. A better understanding of how and where in the mycobacterial membrane these secretion systems are assembled is needed. More knowledge on the structure, functions, regulation and substrates of these secretion systems are required to identify possible drug targets or vaccine candidates that could help curb the tuberculosis epidemic.

# CHAPTER 1

## Introduction

## 1.    Introduction

*Mycobacterium tuberculosis* is a pathogenic organism that is the causative agent of tuberculosis disease in humans. It is estimated that a third of the world's population is infected with *M. tuberculosis* and that approximately 2 million people annually succumb to this disease (WHO 2011). The use of the attenuated *M. bovis* bacilli Calmette-Guérin (BCG) and *M. microti* as vaccines has had little success in the prevention of disease in areas of high tuberculosis prevalence (Pym *et al.* 2002). These vaccine strains have a deletion known as the region of difference one (RD1) which has been shown to have caused the attenuation of these organisms (Pym *et al.* 2002, Brodin *et al.* 2002). This deletion spans part of the ESAT-6 gene cluster region 1 and includes the genes for the T-cell antigens CFP-10 (10 kDa culture filtrate protein, EsxB) and ESAT-6 (6 kDa early secreted antigenic target, EsxA) (Berthet *et al.* 1998).

*M. tuberculosis* possesses five paralogues of the ESAT-6 gene cluster, ESX-1, -2, -3, -4, and -5 (Figure 1.1). Interestingly, ESAT-6 gene clusters have also been identified in some high G+C Gram-positive Actinobacteria, such as in the closely related mycolata species which have mycolic acids in their cell wall, *Corynebacterium* and *Nocardia*.  However, multiple copies of these gene clusters can only be found within the genus *Mycobacterium*. (Gey van Pittius *et al.* 2001). ESX-4 has been identified as the ancestral region as its location as well as gene arrangement is the same as that of *Corynebacterium diphtheria* (Gey van Pittius *et al.* 2001). ESX-4 is also the smallest of the ESAT-6 gene cluster regions and contains the fewest genes (Gey van Pittius *et al.* 2001). Phylogenetic analyses have indicated that the order in which the ESX were duplicated is 4, 3, 1, 2 and 5 (Gey van Pittius, personal communication).

The ESAT-6 gene clusters regions are hypothesized to encode a specialized type VII secretion system that exports extracellular proteins across the membrane of the mycobacteria (Abdallah *et al.* 2007). These secretion systems are commonly referred to as the ESAT-6 secretion systems or ESX secretion systems.

**Figure 1.1 The five ESAT-6 gene cluster regions in *M. tuberculosis* H37Rv.**

The figure shows the five ESAT-6 gene cluster regions of *M. tuberculosis* H37Rv and their genetic organization. *ecc* is <u>e</u>sx <u>c</u>onserved <u>c</u>omponent and esp is <u>E</u>SX-1 <u>s</u>ecretion-associated <u>p</u>rotein. Source: Bitter *et al*. 2009
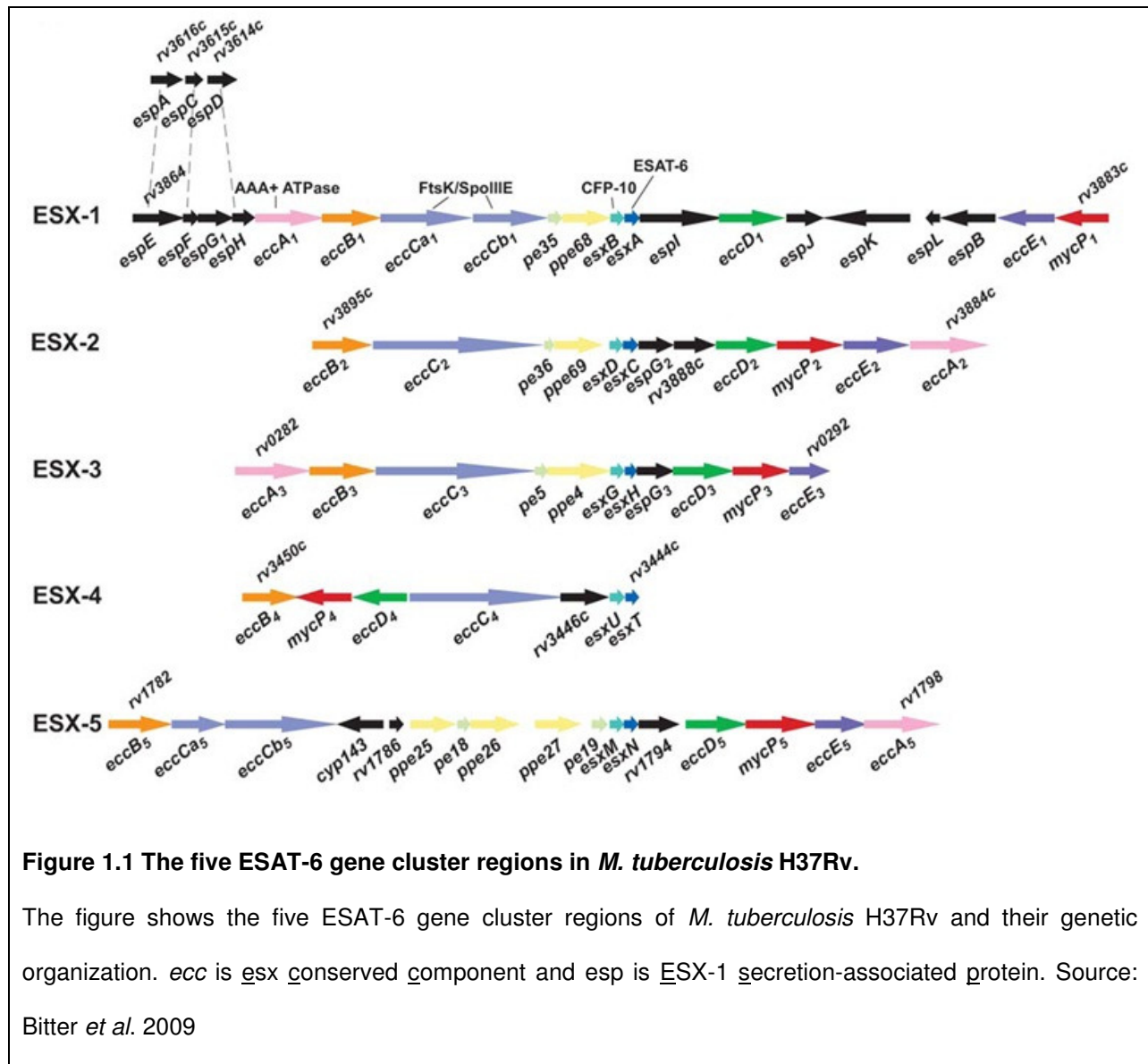
As mentioned above, *M. tuberculosis* contains five ESX secretion systems, ESX-1 to 5. ESX-1 is known to secrete the ESAT-6 and CFP-10 antigens which lack known secretion usually associated with the Tat and Sec secretion pathways (Berthet *et al.* 1998, Sorensen *et al.* 1995, Feltcher, Sullivan & Braunstein 2010). These antigens and the other genes that are situated in the RD1 deletion (in ESX-1) are required for the pathogenicity of the organism, as well as the cell to cell migration of *M. tuberculosis* and *M. marinum* (Guinn *et al.* 2004, Gao *et al.* 2004, Abdallah *et al.* 2006). ESX-3 is the only gene cluster shown to be essential for the *in vitro* growth and survival of *M. tuberculosis* (Sassetti, Boyd & Rubin 2003). It has

also been shown that the expression of ESX-3 is regulated by Iron dependent repressor (IdeR) and Zinc uptake regulator (Zur) in response to the availability of iron and zinc, respectively (Rodriguez *et al.* 2002, Maciag *et al.* 2007). The most recently evolved ESX gene cluster, ESX-5, has not been shown to be important for *in vitro* growth of *M. tuberculosis* (Sassetti, Boyd & Rubin 2003, Sassetti, Rubin 2003). It has, however, been implicated in cell death and has also been shown to be required for the export of many proline-glutamic acid motif (PE)/ proline-proline-glutamic acid motif (PPE) family proteins in *M. marinum* and *M. tuberculosis* (Abdallah *et al.* 2006, Abdallah *et al.* 2011, Abdallah *et al.* 2009, Bottai *et al.* 2012). The presence of ESX-5 is also able to distinguish between fast- and slow-growing mycobacteria (Gey van Pittius *et al.* 2006). Little is known about the functions of ESX-2 and ESX-4, but studies indicated that neither of these two secretion systems are required for in *vitro* growth or virulence (Sassetti, Boyd & Rubin 2003, Sassetti, Rubin 2003). It is believed that divergent evolution as well as differential regulation signals resulted in these ESX secretion systems having such diverse functions (Abdallah *et al.* 2007).

## 2.     The type VII secretion system components

The ESX loci contain a number of conserved genes that encode the core components of the type VII secretion system and as such these gene clusters appear to be homologues. However, some ESAT-6 gene cluster regions encode genes that are unique to them. Comparative studies of the five different ESX systems in *M. tuberculosis* identified six genes that are present in all the ESAT-6 gene cluster regions and as such are hypothesized to encode the core components that make up the type VII secretion system and its substrates (Abdallah *et al.* 2007). These genes encode the following proteins: homologues of ESAT-6 and CFP-10, a member of the FtsK/SpoIIIE family, a subtilisin-like serine protease (mycosin), an integral membrane protein with 10-11 transmembrane domains and another member of the transmembrane protein family (Gey van Pittius *et al.* 2001). Aside from these six genes, all ESX gene clusters, except for ESX-4, possess homologues of the PE and PPE genes and these proteins, together with the ESAT-6 and CFP-10 homologues are predicted to be the substrates of these secretion systems. The conserved components EccA, -B, -C, -D and -E and the subtilisin-like serine protease, MycP (Table 1.1), are all required for the secretion of the ESAT-6/CFP-10-like heterodimers as well as other ESX

**Table 1.1: The core components (apart from the secretion products EsxA/EsxB and PE/PPE) of the *M. tuberculosis* ESX secretion systems.**

| | Description | Gene annotation in ESAT-6 gene cluster | | | | |
|---|---|---|---|---|---|---|
| | | **ESX-1** | **ESX-2** | **ESX-3** | **ESX-4** | **ESX-5** |
| **EccA** | AAA+ class ATPases, CBXX/CFQX family, SpoVK, 1 x ATP/GTP-binding site | Rv3868 | Rv3884c | Rv0282 | | Rv1798 |
| **EccB** | Amino terminal transmembrane protein, possible ATP/GTP binding motif | Rv3869 | Rv3895c | Rv0283 | Rv3450c | Rv1782 |
| **EccC** | DNA segregation ATPase, ftsK chromosome partitioning protein, SpoIIIE, YukA, 3 x ATP/GFP-binding sites, 2 x amino-terminal transmembrane protein | Rv3870-Rv3871 | Rv3894c | Rv0284 | Rv3447c | Rv1783-Rv1784 |
| **EccD** | Integral inner membrane protein, binding-protein-dependent transport systems inner membrane component signature, putative transporter protein | Rv3877 | Rv3887c | Rv0290 | Rv3448 | Rv1795 |
| **EccE** | 2 x amino-terminal transmembrane protein | Rv3882c | Rv3885c | Rv0292 | | Rv1797 |
| **MycP** | Mycosin, subtilisin-like cell wall-associated serine protease | Rv3883c | Rv3886c | Rv0291 | Rv3449 | Rv1796 |

substrates (Figure 1.2). ESX-4 does not have a copy of EccA or EccE but it is believed that these genes were incorporated with the evolution of the duplications of the other ESAT-6 gene cluster regions. These conserved proteins, the ESX and the PE/PPE substrates, will further be discussed in more detail.



**Figure 1.2. The predicted location of conserved ESX gene products within the inner mycobacterial membrane.**

The figure shows the hypothesized location of conserved ESX gene products within the membrane of *M. tuberculosis* where they assemble to form the type VII secretion system. The channel present in the outer membrane represents a hypothetical pore protein, the existence of which is unconfirmed. Source: Bitter *et al.* 2009.

**2.1 ESX secretion substrates**

Our understanding of how type VII secretion systems are assembled and how they secrete target proteins is largely based on work done on ESX-1. The most well known ESX-1 secreted proteins are ESAT-6 and CFP-10, which are secreted as a stable heterodimer (Figure 1.3 A) (Renshaw *et al.* 2002). CFP-10 and ESAT-6 are members of the WXG protein family which are approximately a 100 amino acids in length each, and can typically be found in pairs (Gey van Pittius *et al.* 2001, Renshaw *et al.* 2002, Pallen 2002). The *M. tuberculosis* genome contains 23 pairs of genes for ESAT-6 and CFP-10, which are located at 11 loci and are often preceded by genes from the PE and PPE protein family (Gey van Pittius *et al.* 2001). The secretion of the ESAT-6/CFP-10 protein complex (EsxA and EsxB transcribed from the ESX-1 region) is important for the pathogenicity of some mycobacterial species. For example: the absence of

ESX-1 secretion in *M. bovis* BCG, *M. microti* and *M. tuberculosis* H37Ra results in the attenuation of these strains (Pym *et al.* 2002, Frigui *et al.* 2008). Although EsxA and EsxB are required for the pathogenicity of the organism, the mechanism of their involvement remains unclear. Another ESX gene, EsxH (Rv0288) from ESX-3, has also been shown be to a strong antigen in tuberculosis patients and BCG vaccinated individuals (Skjot *et al.* 2000). Despite the diverse functional roles for each of the five ESAT-6 gene cluster regions, it is hypothesized that the Esx protein complexes adopt similar backbone topologies to the reported EsxA·EsxB complex with unique surface properties and features that determine their functional roles. It has been determined that the core structure of EsxG·EsxH is similar to that of EsxA·EsxB but that this structure might possess a specific $Zn^{2+}$ binding site (Figure 1.3 B). Unlike the EsxA·EsxB heterodimer, there is no evidence of interaction with the surface of macrophage/monocyte-like cells. This suggests that EsxA·EsxB is involved in pathogen-host cell signaling and that EsxG·esxH is involved in iron and zinc acquisition (Ilghari *et al.* 2011). Other proteins that are secreted by the ESX-1 secretion system include EspR, EspA, EspB and EspC (Feltcher, Sullivan & Braunstein 2010).



**Figure 1.3. The structures of the EsxA·EsxB and EsxG·EsxH complexes.**

A. A ribbon representation of the CFP-10·ESAT-6 complex with ESAT-6 in blue and CFP-10 in red. Source: Lightbody *et al.* 2008. B. A ribbon representation of the EsxG·EsxH complex. EsxG is shown in red and EsxH is shown in blue. Source: Ilghari *et al.* 2011. The images clearly show the two helix-turn-helix hairpin structures formed by the Esx- Esx complexes and the flexible N and C termini of both proteins within the complexes.

It is hypothesized that the first PE and PPE genes were inserted into the first duplication of ESX-4 to ESX-3 and that they were subsequently duplicated with the other ESAT-6 gene cluster regions (Gey van Pittius *et al.* 2006). The PE and PPE protein family is unique to the mycobacteria and although the precise function(s) of these proteins remain(s) unknown, they have been implicated in bacterial virulence (Brennan *et al.* 2001, Li *et al.* 2005, Ramakrishnan, Federspiel & Falkow 2000). These PE/PPE gene families have expanded extensively in the slow growing and pathogenic mycobacteria. In *M. marinum*, various PPE and PE family proteins, such as PPE41 and members of the Polymorphic GC-rich-repetitive sequence (PE_PGRS) subfamily, are secreted in an ESX-5 dependent manner (Abdallah *et al.* 2006, Abdallah *et al.* 2009). ESX-5 has also been shown to be required for the secretion of PPE41 in *M. tuberculosis* and it has been shown that the PE and PPE proteins encoded by ESX-5 might be involved in the ESX-5 secretion machinery (Bottai *et al.* 2012). The *ppe-pe* genes are flanked by *ecc* (esx conserved components) genes that are predicted to encode membrane components and ATP-binding proteins that form the secretion system (Brodin *et al.* 2004, Bitter *et al.* 2009).

**2.2 ESX Conserved Component A (eccA)**

EccA is a cytoplasmic AAA ATPase from the CbxX/CfqX subfamily of AAA+ ATPases (Luthra *et al.* 2008). The AAA+ class ATPases are believed to have a wide range of functions including protein degradation, DNA replication, recombination, transcriptional regulation, vesicular fusion, peroxisome biogenesis and the assembly of membrane complexes (Snider, Houry 2008, Ogura, Whiteheart & Wilkinson 2004). In *M. tuberculosis* and *M. marinum* EccA is an essential component of the ESX-1 secretion system, although it's exact role and functions have not been characterized. It has been found that disruption of Rv3868 (EccA$_1$ in *M. tuberculosis*) prevents the secretion of EsxA, EsxB, EspA and EspB however it does not affect the expression of these proteins (Luthra *et al.* 2008).

AAA$^+$ family proteins are known to form hexameric rings of homo- or hetero-oligomers (Ogura, Whiteheart & Wilkinson 2004). EccA$_1$ exists as multiple hexamers with a molecular mass of 380kDa. It is believed that the N-terminal domain of this protein does not possess any ATPase activity and that this compact domain is a monomer in contrast to the hexameric association that's observed with the full length protein. The C-terminal domain was found to predominantly act as a dimer in the absence of ATP; however in the

presence of ATP the protein forms higher order oligomers (Luthra *et al.* 2008). *In silico* methods were used to generate a model of the hexameric association of Rv3868 (Figure 1.4). The resulting model suggests that the binding pocket is lined with Pro-336, Glu-337, Thr-338, Lys-340 and Arg-429 and that the ATP lies in a defined area within the binding pocket in which Arg-429 is part of a neighbouring subunit and acts as a sensory arginine (Ogura, Whiteheart & Wilkinson 2004). This residue senses the presence of the nucleotide within the binding site giving rise to the associated mechanical outcomes of AAA-ATPases (Luthra *et al.* 2008).
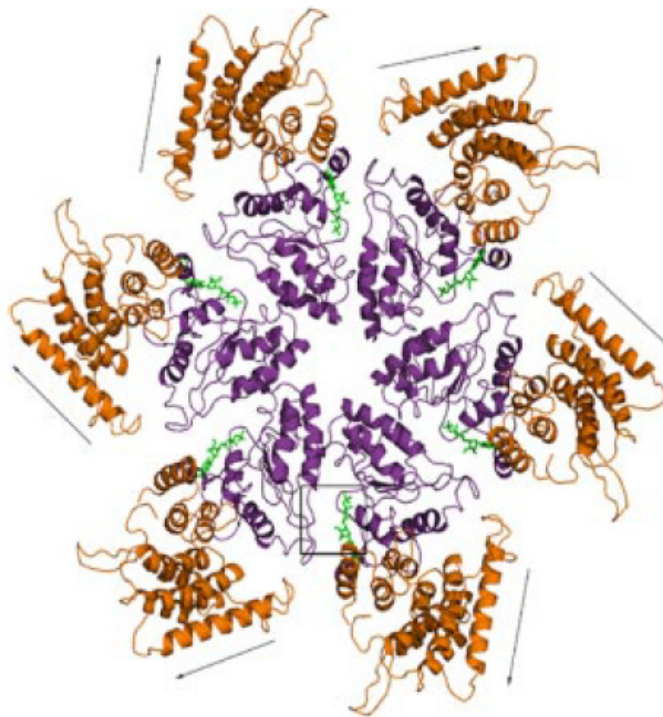


**Figure 1.4. The hexameric structure formed by Rv3868.**

Rv3868 forms a hexameric structure that possesses ATPase activity. The arrows indicate the predicted movement of the N-terminal domain. The ATP binding site is marked with a box. Source: Luthra *et al.* 2008

It has been predicted that Rv3868 might be involved in the translocation of substrates across the mycobacterial membrane. Rv3868 may interact with a gating protein to specifically modulate secretion of virulence factors. Alternatively, Rv3868 may be involved in translocation of substrates by transferring energy to co-proteins of the ESX-1 secretion system. When an ATP molecule is bound to Rv3868, the two domains of the molecule are in close proximity, however upon the release of the nucleotide there is a conformational change as to make the molecule more accessible to interactions with proteins of the ESX-1 secretion system (Luthra *et al.* 2008).

It is believed that the C-terminal region of EsxB is a signal sequence that is recognized by Rv3871 and that this sequence targets the EsxA/EsxB heterodimer for secretion by the ESX-1 system (Champion *et al.* 2006). Another protein that has a high sequence similarity to the C-terminal region of EsxB is EspC (Rv3615c). These two proteins share conserved regions in the last 25 amino acids but have divergent signal sequences that indicate different targeting sequences for these two substrates. It has been shown that the two signal sequences are both required for ESX-1 secretion and that the EspC signal sequence is not functionally equivalent to the EsxB signal sequence (DiGiuseppe Champion *et al.* 2009).  It was determined through yeast two-hybrid studies that EspC interacted with Rv3868 and that ExsB failed to interact with Rv3868. Rv3868 showed no direct interactions with EsxA/EsxB proteins as well as no stimulation of the ATPase activity in assays (Luthra *et al.* 2008). This suggests that the Rv3868/EspC interaction is specific (DiGiuseppe Champion *et al.* 2009).

EccA is thus an ATPase with a co-factor-induced "open–close" movement that allows for the secretion of the EsxA/EsxB heterodimer (Figure 1.4), possibly by providing energy to other proteins of the ESX-1 secretion machinery, such as EspE in order to secrete the EsxA/EsxB complex (Luthra *et al.* 2008, DiGiuseppe Champion *et al.* 2009).

**2.3 EccB**

EccB is a possible amino terminal transmembrane protein with a ATP/GTP binding motif. It has been established that interruption of Rv3869, the EccB from ESX-1 results in the loss of EsxA/EsxB secretion

despite normal expression of these two virulence factors (Brodin *et al.* 2006). No further information is available at this stage regarding the function of EccB.

## 2.4 EccC

In *M. tuberculosis*, EccC is an FtsK/SpoIIIE-like protein that is transcribed from one gene in ESX-4, -3 and 2 (Rv3447c, Rv0284 and Rv3894c, respectively), but from two genes in ESX-1 and ESX-5 (Rv3870-Rv3871 and Rv1783-Rv1784, respectively) (Figure 1.1). Members of the FtsK/SpoIIIE protein family have a diversity of functions that includes chromosome partitioning, conjugative transfer of plasmid DNA, and the export of virulence factors (Burton, Dubnau 2010).

It is predicted that Rv3870 is an integral membrane protein and that Rv3871 is a cytoplasmic AAA ATPase that is required for the secretion of the EsxA/EsxB heterodimer (Feltcher, Sullivan & Braunstein 2010). It has been determined that the 7 C-terminal residues of EsxB are required for the secretion of the EsxA and EsxB complex by ESX-1. A model has been proposed for the targeting of secreted substrates by ESX-1 in *M. tuberculosis*, in which the stable EsxA/EsxB heterodimer interacts with Rv3871 via the C-terminal domain of EsxB. The substrate-bound Rv3871 interacts with Rv3870, which targets the EsxA/EsxB complex for secretion (Champion *et al.* 2006).

## 2.5 EccD

EccD is a probable multi-transmembrane spanning protein with short periplasmic and cytoplasmic loops that form a channel or translocation pore in the inner membrane of the mycobacterium (Feltcher, Sullivan & Braunstein 2010). In *M. tuberculosis* it is believed that Rv3877 forms a pore or a channel that can be used for protein translocation across the inner membrane of the organism (Feltcher, Sullivan & Braunstein 2010). Disruption of Rv3870, Rv3871 and Rv3877 abolishes the secretion of EsxA and EsxB and causes the organism to become less virulent (Stanley *et al.* 2003).

## 2.6 EccE

EccE (Rv3882c in ESX-1) is a possible transmembrane protein situated in the inner membrane of *M. tuberculosis* facing the periplasm (Abdallah *et al.* 2007). Rv3882c has been shown to interact with

Rv3614c (EspD), a cytosolic protein. This interaction is essential for the secretion of EsxA and EsxB (MacGurn *et al.* 2005). No further information is available at this stage regarding the function of EccE.

## 2.7 Mycosin P

Sequence analysis of the mycosins suggests that this group of proteins are membrane-anchored serine proteases of which each protein contains a conserved catalytic triad (Asp, His, Ser), the presence of which is typical of members of the subtilisin protease family. The mycosin proteases all have N-terminal signal sequences and C-terminal transmembrane anchors which allow them to ne anchored in the membrane of the organism (Brown *et al.* 2000, Ohol *et al.* 2010). The non-pathogenic organism *M. smegmatis* possesses three mycosin proteases MycP1, 2 and 3, which were shown to be localised in the bacterial cell wall. This data supports the hypothesis that mycosins might possess signal peptides with C-terminal transmembrane regions (Brown *et al.* 2000). MycP1 is essential for ESX-1 secretion in *M. tuberculosis* and *M. smegmatis (Ohol et al. 2010).*

Microbial pathogens can utilize proteases as virulence factors for a number of functions such as activation of regulatory proteins or peptides, inactivation of host defense molecules, acquisition of nutrients, tissue damage and the processing of secreted signaling molecules that regulate gene expression (Dave *et al.* 2002).

A MycP1 knockout strain expressing a mycosin protease with a modified active site displayed increased secretion of EsxA, EsxB, EspA and EspR. It is believed that the increased secretion of ESX-1 substrates is the direct result of the loss of the protease activity of MycP1 and that MycP1 negatively regulates ESX-1 secretion in *M. smegmatis* by cleaving the activator protein, EspB, which is essential for ESX-1 secretion. It is thought that MycP1 regulates the amount of EsxA which is secreted to obtain an optimal balance between virulence and immunogenicity (Ohol *et al.* 2010), as EsxA stimulates the immune response. Thus over-stimulation of the immune response may suppress *M. tuberculosis* growth *in vivo*.

### 3.      The ESX-3 secretion system

All mycobacteria possess an ESX-3 gene cluster region. ESX-3 is essential for the *in vitro* growth and survival of *M. tuberculosis* (Sassetti, Boyd & Rubin 2003, Sassetti, Rubin 2003), however, ESX-3 was found to not be essential for growth in *M. smegmatis* (Siegrist *et al.* 2009). In *M. tuberculosis* H37Rv, the ESX-3 operon encompasses the genes *Rv0282* to *Rv0292,* while in *M. smegmatis* the operon encompasses the genes *MSMEG_0615* to *MSMEG_0626*. This region has been implicated in cation homeostasis and antigenicity in tuberculosis.

### 3.1  The role of ESX-3 in metal cation homeostasis

The availability of iron is very important for the survival of *M. tuberculosis,* which uses this element as a co-factor for enzymes that are involved in redox reactions and other essential functions(Rodriguez *et al.* 2002). Due to iron not being readily available within the host organism, mycobacteria produces mycobactins which are high affinity iron chelators, however the amount of iron in the mycobacterial cell needs to be regulated as high amounts could be lethal (Rodriguez *et al.* 2002). Mycobacteria have four potential iron-dependent regulators belonging to different families: FurA and Zur (previous annotation: FurB) from the ferric uptake regulator (Fur) family and IdeR and SirR are members of the diphtheria toxin repressor (DtxR) family (Rodriguez *et al.* 2002, Maciag *et al.* 2007).

IdeR is essential in *M. tuberculosis* and is an iron dependent DNA binding protein which interacts with specific DNA sequences called "iron-boxes" in operon regions of iron-regulated genes to control their expression (Gold *et al.* 2001, Rodriguez *et al.* 1999, Dussurget *et al.* 1999). DNA microarray studies have shown that 51 genes are regulated by IdeR and that all of these genes have sequences resembling iron boxes in their upstream regions. The identified IdeR controlled genes encode a number of proteins with diverse functions, these include: transporters (Rv0282, Rv0283, Rv0284), enzymes involved in lipid metabolism (Rv1344, Rv1345 and Rv1347) and members of the PE/PPE protein family (Rv0285, Rv0286 and Rv2123) (Rodriguez *et al.* 2002). A number of these genes are situated within the ESX-3 gene cluster, indicating that ESX-3 is regulated by IdeR in response to the amount of iron available within the cell. IdeR has been shown to be responsible for iron dependent siderophore production in *M. tuberculosis*

and is required for an efficient response to oxidative stress (Rodriguez *et al.* 2002, Gold *et al.* 2001). In *M. smegmatis* IdeR was found to repress siderophore expression when bound to iron (Dussurget, Rodriguez & Smith 1996).

Zinc is also an essential element for *M. tuberculosis* growth*,* and it acts as a co-factor for enzymes and DNA binding proteins. Zinc also serves as a structural scaffold for several other proteins (Maciag *et al.* 2007). Zur is a repressor protein and within a *zur* mutant, genes under the transcriptional control of this protein are constitutively expressed (Maciag *et al.* 2007). This repressor protein binds a target sequence in the presence of zinc. DNA microarray studies identified 24 genes that were directly regulated by Zur. Several genes under Zur dependent transcriptional control are homologous to genes involved in zinc uptake in other bacteria (Maciag *et al.* 2007). In a *M. smegmatis zur* deletion mutant, plasmids with the following promoters of *M. tuberculosis* genes were found to be over-expressed: Rv0106, Rv0280, Rv0282 and Rv2059. Rv0280 and Rv0282 are both present in ESX-3. Zur was also found to represses the expression of five ESAT-6 and CFP-10 like proteins from *M. tuberculosis*, namely *esxG*, *esxH*, *esxQ*, *esxR* and *esxS (Maciag et al. 2007).* Two of the 5 *esx* genes (*esxG* and *esxH*) regulated by Zur are located within ESX-3 and all 11 genes of ESX-3 from *M. tuberculosis* were upregulated in the *zur* deletion mutant, which suggests they are in an operon (Maciag *et al.* 2007).

It would thus appear that the promoter upstream of ESX-3 in *M. tuberculosis* is regulated by both Zur and IdeR proteins. In *M. smegmatis*, ESX-3 only responds to iron and not to zinc (Maciag *et al.* 2009). Iron and zinc dependent expression of ESX-3 along with the increased expression of ESX-3 genes under acid stress conditions support the hypothesis that ESX-3 is expressed during tuberculosis infection (Maciag *et al.* 2009).

Mycobacteria use two classes of siderophores to acquire iron: exochelin and mycobactin (Ratledge, Dover 2000). Pathogenic mycobacteria only use mycobactin and the structurally related carboxymycobactin to acquire iron (Ratledge, Dover 2000). ESX-3 has been shown to be required for iron acquisition via the mycobactin pathway and survival during infection. ESX-3 is also required for the growth of *M. bovis* BCG and *M. smegmatis* under iron limiting conditions.  It is believed that ESX-3 is required for the uptake of iron-bound mycobactin (Siegrist *et al.* 2009).

15

In a *M. tuberculosis* conditional mutant, ESX-3 has been shown to be required for growth in the absence of iron. The growth phenotype that is observed with this conditional ESX-3 mutant can be complemented by additional iron and zinc or the culture filtrate from wild-type *M. tuberculosis*. ESX-3 is thus required for the optimal uptake of iron and zinc and is thus essential for the growth of the organism in 7H9 and 7H10. ESX-3 is responsible for the secretion of unknown factors that are required for optimal iron and zinc uptake and the growth of *M. tuberculosis (Serafini et al. 2009)*.

### 3.2  ESX-3 in immunogenicity and vaccine development

The use of vaccines to combat tuberculosis infection has had little success in areas of high tuberculosis prevalence. Currently *M. bovis* BCG and *M. microti* are used as vaccines but they seem to provide little or no protection against *M. tuberculosis* infection (Pym *et al.* 2002). There is thus a desperate need for new effective vaccines to help against the fight with tuberculosis.

Recently, a new potential vaccine was created by knocking out ESX-3 in *M. smegmatis*. This vaccine was called the IKE strain due to the loss of the "immune killing evasion" that was observed in mouse models (Sweeney *et al.* 2011). Mice that were infected with the IKE strain could contain the infection with an innate immune response, indicating that ESX-3 genes mediate an evasion of innate immune responses. Complementation of the IKE strain with the *M. tuberculosis* ESX-3 genes (*rv0278 – rv0303)* resulted in a strain, IKEPLUS, which was shown to be rapidly cleared and was unable to kill immunocompetent mice. This potential vaccine candidate produced a potent and long lived adaptive immune response during infection that resulted in reduced levels of *M. tuberculosis* bacilli in IKEPLUS-immunized mice (Sweeney *et al.* 2011).

IKE and IKEPLUS strains both possess desirable properties for anti-tuberculosis vaccines. Both these mutants were found to be highly attenuated in mice and they rapidly induced a cytokine response. IKEPLUS has the added benefit of being an *M. smegmatis* recombinant strain that contains a cosmid with *M. tuberculosis* ESX-3 genes. These genes being present in a vaccine strain have the potential to serve as specific antigens which could elicit an adaptive immune response upon infection with *M. tuberculosis*. (Sweeney *et al.* 2011)

**4.        The Localisation of ESX secretion systems in Mycobacteria**

Many pathogenic bacteria possess specialized virulence associated secretion systems, for example type III or type IV secretion systems, which are specifically active at the bacterial poles. This implies that the polar localisation of virulence related protein secretion is a feature that can be associated with pathogenesis (Jaumouille *et al.* 2008). The localisation of these secretion systems to a specific compartment or compartments of the cell allows for site directed secretion of virulence factors or effecter proteins that might be important for the pathogenesis of the organism (Carlsson *et al.* 2009, Wirth *et al.* 2012).

Recently it has been shown that ESX-1 is localised at the polar region in *M. marinum* as well as in *M. smegmatis* (Carlsson *et al.* 2009). Immunofluorescent microscopy was used to identify the location of the Rv3864 homologue in *M. marinum*, MMAR_5439. Rv3864 (EspE) was shown to be an ESX-1 substrate of which a part remains surface associated upon secretion. It was determined that EspE localises at the poles (at both the old poles and the new poles), indicating that ESX-1 secretion primarily occurs at the poles in *M. marinum* (Carlsson *et al.* 2009). A KasB-deficient mutant strain that possesses a more permeable cell wall and as such allows for an anti-serum to interact with the *M. marinum* homologue of Rv3870 (EccCa$_1$) was used to determine the localisation of this protein in the mycobacterial membrane. EccCa$_1$ also localised to the poles, supporting the observation that ESX-1 localises to the bacterial poles. The majority of cells displayed a unipolar localisation, predominantly at the new bacterial pole (Carlsson *et al.* 2009). *M. marinum* forms actin tails at one of their poles upon reaching the cytosol of infected host cells (Stamm *et al.* 2003). EspE localised to the bacterial pole with actin tails, indicating that ESX-1 is active at the poles and that actin polymerization occurs primarily at the new bacterial poles. It is believed that new peptidoglycan is inserted at the new poles, making the new poles the position of active cell growth. Thus, ESX-1 localises to the region of active cell wall production. These findings implicate a functional relationship between the secretion site and active cell growth that could possibly be of general importance to gram positive bacteria, including the mycobacteria (Carlsson *et al.* 2009).

The use of immunofluorescent microscopy and a ΔkasB mutant strain to determine the subcellular localisation of ESX-1 in *M. marinum* does have its limitations, as a number of cytoplasmic components of ESX-1 cannot be visualized by making use of this method. Recently, a study was performed in which fluorescent protein fusions were used to determine the subcellular localisation of ESX-1 components in *M. smegmatis*. The ESX-1 proteins were fused to a yellow fluorescent protein with a penta-glycine linker chain to allow native folding of the N-terminal ESX-1 protein (Wirth *et al.* 2012). EspE from *M. smegmatis* (MSMEG_0055) and *M. tuberculosis* (Rv3864) were found to localize at the bacterial poles, however unlike the *M. marinum* orthologue EspE remained unsecreted in *M. smegmatis*. EccCb from *M. smegmatis* (MSMEG_0062) and *M. tuberculosis* (Rv3871) were also found to localize at the bacterial poles (Wirth *et al.* 2012). These findings indicate that the cellular proteins and signals that are necessary for the polar localization of the ESX-1 apparatus are conserved in *M. marinum*, *M. smegmatis* and *M. tuberculosis* (Wirth *et al.* 2012). The majority of cells in which the YFP-tagged proteins were expressed exhibited predominantly unipolar localisation. The presence of the fluorescently tagged proteins in a single polar focus indicated that the poles of the mycobactrial cells are functionally distinct. (Wirth *et al.* 2012)

EccCb-YFP fusions from *M. smegmatis* and *M. tuberculosis* failed to localize to the bacterial poles in a *M. smegmatis* strain that has a large ESX-1 deletion, indicating that localization of these proteins to the poles is specific and not an artifact of protein aggregation or expression (Wirth *et al.* 2012). The localization of these proteins to the polar regions might be dependent on the expression of the other proteins that are found in the ESX-1 operon. The failure of EccCb to localize to the polar regions of the cell in a *M. smegmatis* strain that possesses a deletion of EccCa showed that other proteins that are encoded from the ESX-1 region are also required for the localization of this protein to the polar regions of the cell. These findings where further supported by the observation that the expression of *M. tuberculosis* ESX-1 in a *M. smegmatis* ESX-1 knock-out mutant restored the polar localistion of EccCb$_{ms}$. This suggests that the ESX-1 apparatus in *M. tuberculosis* is recruited to the mycobacterial cell pole. The functional conservation of the ESX-1 secretion system also implies that the observed species-specific differences are due to different secreted proteins rather than differences in their secretion systems (Wirth *et al.* 2012). Unsurprisingly, *M. smegmatis* EsxA and EsxB YFP fusions did not localise to any regions within the cell.

18

This is expected because it is hypothesized that only structural components of the ESX secretion systems will localise to the polar regions of the cell (Wirth *et al.* 2012).

Transposon mutagenesis screens identified non- *esx1*-encoded proteins that were found to be essential for EsxA/EsxB secretion as well as conjugal DNA transfer in *M. smegmatis*. These proteins are encoded from a three gene operon located 7 kb upstream of ESX-1, Msmeg0044-0046. These genes are known as *saeA-C*. A SaeC-YFP fusion protein was found to unipolarly localize at the polar region of the cell. SaeA-YFP was found to have conditional polar localisation. Polar localisation of SaeA was only observed when cultured in the absence of Tween-80. It is thought that the detergent properties of Tween-80 combined with the overexpression of the fusion protein resulted in the alteration of the cell wall structure and as such it disturbed the localisation of the fusion protein. Tween-80 did not have any effect on any of the other fusion proteins.  The SaeB-YFP fusion did not show any polar localisation (Wirth *et al.* 2012).

To determine if fusion proteins localised that the same bacterial pole, SaeC-YFP was co-expressed with either $EspE_{ms}$-TdTomato or $EccCb_{ms}$TdTomato in *M. smegmatis*. In all cases where dual fluorescence was observed, these proteins co-localised to the same mycobacterial pole. The co-localisation of these proteins indicates that ESX-1 associated proteins as well as the structural components of this secretion system are transported to the same location regardless of where they are encoded (Wirth *et al.* 2012). SaeC-YFP localised to the bacterial pole even in the absence of ESX-1 core components, however $EccCb_{ms}$ was unable to localise at the bacterial pole in the absence of *saeC*. It is believed that the defect in DNA transfer as well as EsxA/EsxB secretion in *Δsae* strains are due to ESX-1 not being assembled or being assembled at a different location within the membrane. The localisation of SaeC-YFP to the polar regions independently of ESX-1 encoded proteins indicates that SaeC is an early recruit to the cell pole as well as being a recruiter for ESX-1 assembly (Wirth *et al.* 2012). Time lapse fluorescent microscopy showed that Sae-YFP localized to the old bacterial pole following cell division (Wirth *et al.* 2012). This indicates that all ESX-1 components that show polar localisation also localise at the old bacterial pole, due to these proteins co-localising to the same bacterial pole.

The unique subcellular localisation of this specialized secretion system should come as no surprise as macromolecular complexes are commonly found at a specific location within the cell. Examples of these include complexes that are involved in cell division and DNA replication (Burton, Dubnau 2010). As mentioned before, many pathogenic bacteria localise their specialized secretion systems at the bacterial poles. The mechanism whereby the protein recognizes a polar region remains unknown. It has been suggested that a number of proteins recognize the polar regions of the cell by membrane curvature (Lenarcic *et al.* 2009, Huang, Ramamurthi 2010, Ramamurthi 2010), although it appears that membrane curvature is not the answer for mycobacteria, because the majority of cells only have fluorescence at one of the poles, predominantly the old pole (Wirth *et al.* 2012). It is hypothesized that because cell wall synthesis occurs at this old polar region that the cell wall growth process or a component of it may contribute to the localisation of ESX-1. The localisation of ESX-1 to the old pole where active cell wall synthesis occurs, indicates that it might be involved in the generation and modification of the cell wall (Wirth *et al.* 2012). The studies which investigated the localisation of ESX-1 in *M. marinum* and *M. smegmatis* both indicated that ESX-1 localises to the pole where active cell wall synthesis occurs. However, these two studies had contradicting results as to whether ESX-1 localises at the old or new mycobacterial pole. Recent studies have shown that active cell wall synthesis predominantly occurs at the old mycobacterial pole (Aldridge *et al.* 2012, Flardh 2010). It might therefore be assumed that ESX-1 localises to the old pole. However, it is possible that cell wall growth occurs at different poles in these different species or that the methods used to determine the localisation of ESX-1 in *M. marinum* was insufficient to determine at which mycobacterial pole this secretion system localises as no time-lapse microscopy was used. The inactivation of ESX-5 by a disruption in $EccD_5$ (Rv1795) was found to affect the cell wall stability. This indicates that without the secretion of EsxA secretion substrates from ESX-1 and ESX-5, the mycobacterial cell wall becomes more permeable and that these secretion systems might be involved in the generation of the cell wall (Bottai *et al.* 2012).

## 5. Conclusion

Tuberculosis is an infectious disease that is caused by the pathogenic organism, *M. tuberculosis*. This disease causing mycobacterium possesses five ESX gene cluster regions that are associated with pathogenicity. It is hypothesized that these ESX gene cluster regions encode a type VII secretion system that exports or is involved in the export of antigenic proteins across the mycobacterial membrane. ESX-3 has been shown to be essential for the growth and survival of *M. tuberculosis*, however it has been determined that ESX-3 is not required for survival in *M. smegmatis*. In this thesis we investigate the subcellular localisation of the esx conserved component (ecc) proteins from ESX-3 in *M. smegmatis*. Knowing the subcellular localisation of these essential proteins will increase our knowledge of this important secretion system and might help to develop more effective medical approaches that will deliver drugs to the appropriate location as to selectively interfere with the protein function.

## 6. Aim and Objectives

Aim:

To investigate the cellular location of components of the *M. smegmatis* ESX-3 secretion system in *M. smegmatis.*

Objectives:

1. To select a number of genes of interest located in the ESX-3 secretion system of *M. smegmatis*, which could be used as surrogate markers for the determination of the localization of the assembled ESX-3 secretion apparatus.
2. PCR amplification of the genes of interest.
3. Cloning of the genes of interest into the expression vector, pDMNI.
4. Expression of the selected ESX-3 proteins as fusion proteins with green fluorescent protein (GFP), in wild type *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out.
5. Identification of the localisation of the selected ESX-3 fusion proteins in wild type *M. smegmatis* and in the *M. smegmatis* knock-out strain.

# CHAPTER 2

## Materials and Methods

## 2.1 Bacterial strains

*Escherichia coli* XL-1 Blue was used for all cloning procedures. *M. smegmatis* mc$^2$155 (Snapper *et al.* 1990) and a *M. smegmatis* mc$^2$155 ESX-3 knock-out strain (kind gift from Mae Newton-Foot, Stellenbosch University) was used to express and to determine the localisation of  the esx conserved components (ecc) from ESX-3 in *M. smegmatis*. The *M. smegmatis* ESX-3 knock-out strain is an unmarked mutant with at deletion spanning from MSMEG_0615 to MSMEG_0626.

## 2.2 Media and culture conditions

Culture media compositions are described in Table 2.1. All antibiotics and supplements are described in Table 2.2.

*E. coli* was cultured in liquid Luria-Bertani (LB) broth with shaking or on LB agar plates at 37°C overnight. Liquid or solid media that was used to culture *E. coli* contained either ampicillin (50µg/ml) or kanamycin (50 µg/ml) as required. LB agar underlined with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and isopropyl-β-D-thiogalactopyranoside (IPTG) when required.

*M. smegmatis* strains were grown in Difco$^{TM}$ Middlebrook 7H9 Broth with shaking and on BBL$^{TM}$ Seven H11 Base plates at 37°C for 2-3 days. Media contained kanamycin (25 µg/ml) where appropriate.

Bacterial strains used in this study were stored in 40% glycerol at -80°C.

## 2.3 Ziehl-Neelsen staining

Cultures of *M. smegmatis* were screened for contamination using Ziehl-Neelsen staining (ZN staining) as described by Kent and Kubica (1985). Briefly, cultures of *M. smegmatis* were heat fixed to microscope slides. Heat fixed cultures were flooded with ZN Carbol Fuschin and heated with a flame until steaming, followed by an incubation period of 5 minutes. Slides were rinsed with water and decolourised with 5% acid-alcohol solution for 2 minutes, followed by a wash step with water. Slides were counterstained using Methylene Blue for 1-2 minutes, rinsed with water and allowed to air dry. Reading of slides was done with a light microscope under a 100X oil immersion lens. Uncontaminated *M. smegmatis* cultures appeared as pink acid fast bacilli and contaminated cultures would contain blue coloured organisms.

**Table 2.1: Media composition.**

| Medium | Composition | Source |
|---|---|---|
| Luria-Bertani broth (LB) | 0.5% (w/v) Sodium chloride<br>1% (w/v) Tryptone<br>0.5% (w/v) Yeast extract | Sigma-Aldrich<br>Merck<br>Merck |
| Luria-Bertani agar (LB agar) | 1.2% (w/v) Bacteriological agar<br>0.5% (w/v) Sodium chloride<br>1% (w/v) Tryptone<br>0.5% (w/v) Yeast extract | Merck<br>Sigma-Aldrich<br>Merck<br>Merck |
| Difco$^{TM}$ Middlebrook 7H9 | 0.47% (w/v) Middlebrook 7H9<br>supplemented with 0.5% (v/v) glycerol, 0.5% (v/v) glucose, 0.2% (v/v) Tween-80 after autoclaving | Becton Dickson (BD) |
| BBL$^{TM}$ Seven H11 Base | 1.9% (w/v) BBL$^{TM}$ Seven H11 Base, supplemented with 0.5% glycerol, 0.5% (v/v) glucose, 0.2% (v/v) Tween-80 after autoclaving | Becton Dickson (BD) |
| SOC | 2% (v/v) Glucose<br>1% (v/v) Magnesium chloride<br>0.25% (w/v) Potasium chloride<br>0.05% (w/v) Sodium Chloride<br>2% (w/v) Tryptone<br>0.5% (w/v) Yeast Extract | Merck<br>Sigma-Aldrich<br>Merck<br>Sigma-Aldrich<br>Merck<br>Merck |

**Table 2.2: Antibiotics and supplements.**

| Antibiotic/Supplement | Stock concentration | Solvent | Sterilisation | Storage | Supplier | Working concentration *E. coli* | *M. smegmatis* |
|---|---|---|---|---|---|---|---|
| Ampicillin (amp) | 50 mg/ml | ddH$_2$O | Filtered | -20°C | Roche | 50 µg/ml | |
| Kanamycin (kan) | 50 mg/ml | 0.9% NaCl solution | Filtered | 4°C | Sigma-Aldrich | 50 µg/ml | 25 µg/ml |
| Tetracycline (tet) | 5 mg/ml | Ethanol | | -20°C | Sigma-Aldrich | 50 µg/ml | |
| 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) | 20 mg/ml | N,N'-dimethyl formamide | | -20°C | Roche | 100µl per 90mm plate | |
| Isopropyl-β-D-thiogalactopyranoside (IPTG) | 0.1 M | ddH$_2$O | Filtered | -20°C | Biosolve | 100µl per 90mm plate | |
| Potassium Chloride (KCl) | 1M | ddH$_2$O | Autoclaved, 121°C, 100kPa, 20 min | Room temp | Merck | | |
| Magnesium Chloride (MgCl$_2$) | 1M | ddH$_2$O | Filtered | Room temp | Sigma-Aldrich | | |
| Glucose | 1M | ddH$_2$O | Filtered | -20°C | Merck | | |
| Glucose | 50% (v/v) | ddH$_2$O | Filtered | 4°C | Merck | | 0.5% v/v |
| Glycerol | 50% (v/v) | ddH$_2$O | Filtered | 4°C | Merck | | 0.5% v/v |
| Polyoxyethylene-sorbitan monooleate (Tween-80) | 20% (v/v) | ddH$_2$O | Filtered | 4°C | Sigma-Aldrich | | 0.2% v/v |

## 2.4 Electrocompetent cells

*2.4.1 Electrocompetent E. coli*

Electrocompetent *E.coli* was prepared by streaking out *E. coli* XL-1 Blue onto LB agar plates containing tetracycline (50 µg/ml); plates were incubated at 37°C overnight. A single colony was picked and inoculated into 50 ml LB containing tetracycline (50 µg/ml) and incubated at 37°C overnight with shaking. Overnight cultures were inoculated into 2L of fresh antibiotic-containing medium at a temperature of 37°C (dilution 1:100) and grown to $OD_{600}$ = 0.7 at 37°C with shaking. Cultures were placed on ice and all further steps were done under ice cold conditions (4°C). Cultures were centrifuged at 4000xg for 10 minutes at 4°C. The cell pellet was washed in an equal volume to the culture volume of 10% glycerol (that was stored at 4°C), followed by centrifugation at 4000xg for 10 minutes at 4°C. The wash step was repeated and the cells were pooled in a 50 ml falcon tube, followed by centrifugation at 4000xg for 10 minutes at 4°C. Cells were resuspended in 4 ml of 10% glycerol. Aliquots of 100 µl were prepared followed by freezing in liquid nitrogen and storage at -80°C.

*2.4.2 Electrocompetent M. smegmatis*

Electrocompetent *M. smegmatis* was prepared by inoculating *M. smegmatis* into Difco^TM Middlebrook 7H9 and grown to $OD_{600}$ = 0.5, at 37°C with shaking. Cultures were incubated on ice for 1 hour and all further steps were done under ice cold conditions. Cells were pelleted by centrifugation at 3220xg for 10 minutes at 4°C, followed by resuspension in 10% glycerol (stored at 4°C) equal to culture volume. The resuspended culture was centrifuged at 3220xg at 4°C for 10 minutes to pellet the cells. The wash step was repeated and cells were resuspended in ice cold 10% glycerol using a volume of 2 ml per 100 ml of culture. Electrocompetent *M. smegmatis* was freshly prepared before use.

**2.5 DNA manipulation**

*2.5.1 Primers*

The sequence information used for the design of primers was obtained from the JVCI CMR *Mycobacterium smegmatis* MC2 Genome Page (http://cmr.jcvi.org/cgi-bin/CMR/GenomePage.cgi?org=gms). Primers were designed using Primer Premier 5.0 (PREMIER Biosoft International) and were ordered from Integrated DNA Technologies (IDT, through Whitehead Scientific). For primer sequences refer to Table 2.3.

Cloning primers were designed to amplify the genomic region of the genes of interest, MSMEG_0615, MSMEG_0616, MSMEG_0617, MSMEG_0623 and MSMEG_0626 from *M. smegmatis* mc$^2$155 (Figure 2.1). Primers were designed to amplify the genes of interest in frame. A mycobacterial ribosome binding site (RBS) as well as a restriction enzyme cutting site for *Eco*RI were incorporated into the forward cloning primers with the exception of one primer. Due to *Eco*RI cutting within the amplified MSMEG_0615 gene, MSMEG_0615 forward primer was designed to have the restriction enzyme cutting site *Mfe*I. *Eco*RI and *Mfe*I are isoschizomers because they generate the same overhangs upon digesting DNA. The RBS enables the ribosome to bind to the RNA after transcription enabling translation of the protein.  Forward cloning primers were HPLC purified due to their length that ranged from 53 to 56 bases. Reverse cloning primers were designed to have a *Hind*III restriction enzyme cutting site. The restriction enzyme cut sites allow for easy cloning from the pGEM-T Easy vector into the pDMNI vector where the genes of interest will be expressed as fusion proteins with GFP.

**Table 2.3: Primer sequences.**

| Primer name | Sequence (5' – 3') | Primer length | Tm* | Restriction site |
|---|---|---|---|---|
| MSMEG_0615 f | CAATTG<u>TGACACTTGAGGAGGAGAGCAAAATTCTC</u>ATGGGAAGTGACACGCTTGC | 55 | 60 | *Mfe*I |
| MSMEG_0615 r | AAGCTTTTGCCGGCACCGACAGTC | 24 | 60 | *Hind*III |
| MSMEG_0616 f | GAATTC<u>TGACACTTGAGGAGGAGAGCAAAATTCTC</u>ATGACCGGCCCCGTCAAC | 53 | 60 | *Eco*RI |
| MSMEG_0616 r | AAGCTTTTCGGGAGGCCTCCATACG | 25 | 60 | *Hind*III |
| MSMEG_0617 f | GAATTC<u>TGACACTTGAGGAGGAGAGCAAAATTCTC</u>ATGAGCCGGCTCATCTTCGAG | 56 | 64 | *Eco*RI |
| MSMEG_0617 r | AAGCTTTTCGGTATTCCCCTCCTCGG | 26 | 64 | *Hind*III |
| MSMEG_0623 f | GAATTC<u>TGACACTTGAGGAGGAGAGCAAAATTCTC</u>GTGCTGGCCGCCGGAGAC | 53 | 64 | *Eco*RI |
| MSMEG_0623 r | AAGCTTTCCTGTCGAGCACGAGGCTG | 26 | 64 | *Hind*III |
| MSMEG_0626 f | GAATTC<u>TGACACTTGAGGAGGAGAGCAAAATTCTC</u>ATGACCGCCCGGATAGCG | 53 | 60 | *Eco*RI |
| MSMEG_0626 r | AAGCTTTCGCCGGATGACCCGCTT | 24 | 60 | *Hind*III |
| MSMEG0617seq800 | TGGCACGACCCCACCAT | 17 | 56 | |
| MSMEG0617seq1600 | GCCGACATCTTCCGCCACT | 19 | 62 | |
| MSMEG0617seq2400 | GCAAGTTGATCGCCACCAT | 19 | 58 | |
| MSMEG0617seq3200 | AGCGACATCGCGGTCATCAA | 20 | 62 | |
| pDMNIseq f | GGGTGGGCGAGTTTGTCCTG | 20 | 66 | |
| pDMNIseq r | CCAGTAGTGCAAATAAATTTAAGGG | 25 | 68 | |
| T7 | TAATACGACTCACTATAGGG | 20 | 56 | |
| SP6 | ATTTAGGTGACACTATAG | 18 | 48 | |

*Tm was determined by: (G + C x 4) + (A + T x 2), where A, T, G and C are the number of nucleotides of the type within the primer sequence.

Underlined - Mycobacterial ribosome binding sites within the primers are indicated by underlined regions.
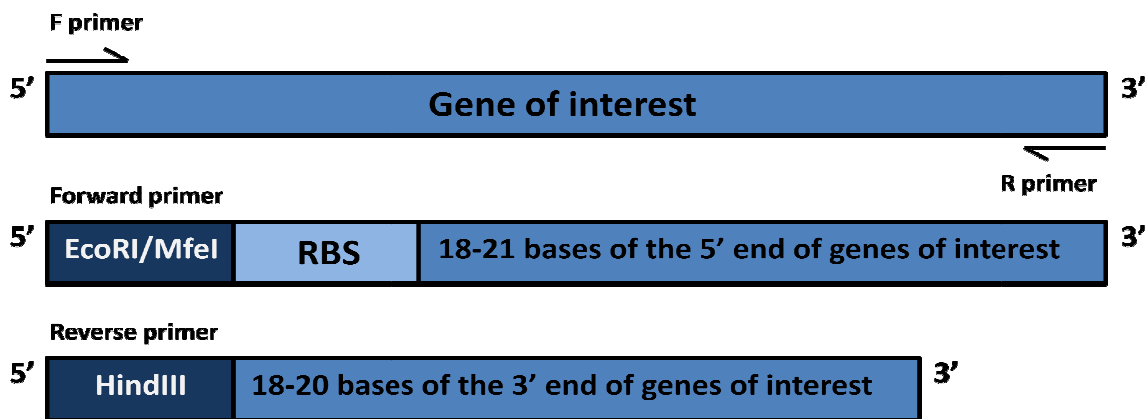
**Figure 2.1. Cloning primer design for *MSMEG_0615*, *MSMEG_0616*, *MSMEG_0617*, *MSMEG_0623* and *MSMEG_0626*.**

Cloning primers were used to amplify each of the genes of interest. The forward cloning primers were designed to incorporate an *Eco*RI/*Mfe*I cut site for cloning as well as a ribosome binding site. Reverse cloning primers were designed to incorporate a *Hind*III cut site for cloning.
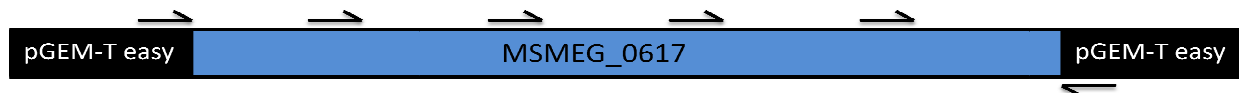


**Figure 2.2. Sequencing primers for *MSMEG_0617*.**

Sequencing primers were designed to sequence the MSMEG_0617 gene to verify that cloning was successful. (1.) MSMEG_0617seq1 sequenced from 700 bp. (2.) MSMEG_0617seq2 were designed to sequence from 1500 bp. (3.) MSMEG_0617seq3 sequenced from 2300 bp. (4.) MSMEG_0617seq4 sequenced from 3100 bp in MSMEG_0617.

29

Constructs were sequenced with sequencing primers for pGem-T Easy, namely T7 and Sp6. Internal sequencing primers were designed for MSMEG_0617 due to the length of the insert. These primers were designed to sequence the sections of the insert and to have overlapping regions for better resolution (Figure 2.2). For the sequencing of pDMNI constructs, the sequencing primers pDMNIseq f and pDMNIseq r were designed (Figure 2.3). These primers were used to confirm the presence of the insert in the vector as well as the integrity of the insert. The pDMNIseq f primer was designed to amplify from a region adjacent to the insert and the pDMNIseq r primer amplified from within the GFP sequence. Refer to Table 2.3 for primer sequences.



**Figure 2.3. Sequencing primers for the final pDMNI expression vector.**

The primers pDMNIseq f and r were designed to amplify from within the pDMNI vector cross the insert. These primers were used to check the presence and the sequence integrity of the insert within the final expression vector.
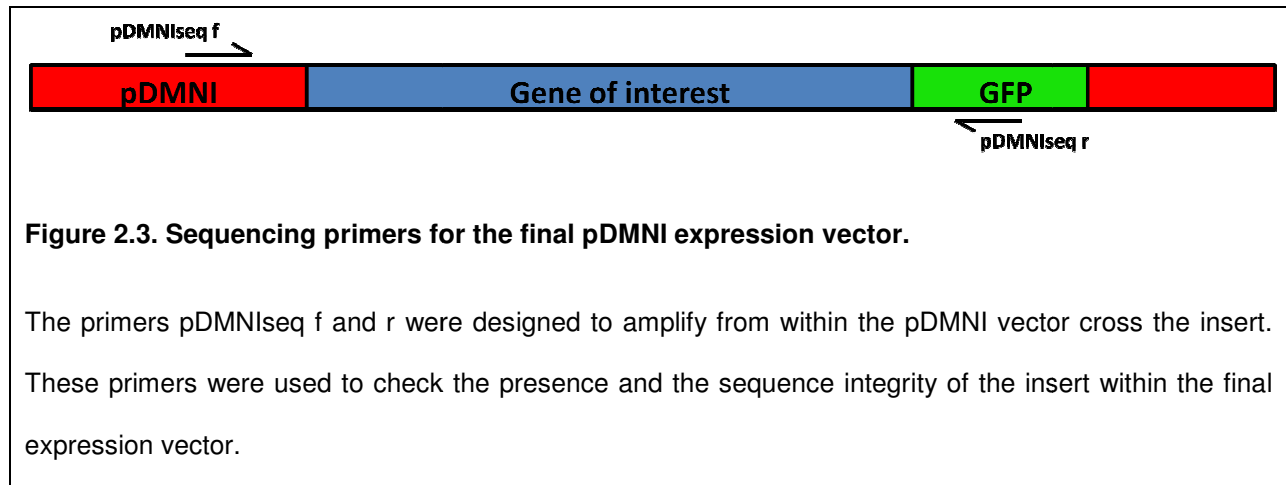
*2.5.2 Cloning vectors*

All cloning vectors are described in Table 2.4.

PCR amplified genes *MSMEG_0616*, *MSMEG_0617*, *MSMEG_0623* and *MSMEG_0626* were ligated into the commercial vector, pGEM-T Easy (Promega) and subsequently cloned into the episomal expression vector pDMNI to form fusion products with GFP. PCR-amplified MSMEG_0615 was ligated into CloneJet and subsequently cloned into pDMNI (Figure 2.4).

**Table 2.4: Cloning vectors.**

| Vector | Description | Size (bp) | Source/Reference |
|--------|-------------|-----------|------------------|
| pGEM-T easy | *E. coli* cloning T-vector, Amp$^R$, *lacZ*, *ori*E | 3015 | Promega |
| CloneJet | *E. coli* blunt or sticky end cloning vector, Amp$^R$, *eco417R*, multiple cloning site (MCS), rep (pMB1) | 2974 | Fermentas |
| pDMNI | *M. smegmatis* expression vector, Kan$^H$, GFP, *ori*E, *ori*M | 5090 | Kind gift from Ros Chapman, UCT |



**Figure 2.4: The episomal vector pDMNI.**

pDMNI is an episomal vector that were used for the overexpression of green fluorescent protein (GFP) fusion proteins. This vector contains an *aph* gene for kanamycin resistance as well as an OriE, truncated cap, repA and repB and a OriM. This vector also contains the mycobacterial promoter Psmyc as well as the GFP gene and an rrnB terminator.

*2.5.3 Polymerase Chain Reaction (PCR)*

Polymerase chain reaction was used to amplify the genes of interest from *M. smegmatis* mc$^2$155 genomic DNA. *M. smegmatis* extracted DNA was a kind gift from N. C. Gey van Pittius, and was prepared as previously described (Sambrook, Fritsch & Maniatis 1989). Genes *MSMEG_0616*, *MSMEG_0623* and *MSMEG_0626* were amplified using FastStart Taq DNA Polymerase (Roche) with the GC-rich buffer, as described by the manufacturer. Amplification conditions are described in Table 2.5. Phusion High-Fidelity DNA Polymerase (Fermentas) was used to PCR amplify the genes *MSMEG_0615* and *MSMEG_0617*. PCR master mix reactions for Phusion High Fidelity DNA polymerase reactions are described in Table 2.6. Amplification conditions for Phusion High-Fidelity DNA Polymerase are described in Table 2.7.

**Table 2.5: PCR amplification conditions for FastStart Taq DNA polymerase.**

| Temperature (°C) | Time | Reaction | Number of Cycles |
|---|---|---|---|
| 95 | 5 min | Activation of FastStart Taq | 1 |
| 95 | 30 s | Template denaturation | |
| Tm* | 30 s | Annealing of primers | 35 |
| 72 | 1 min/kb | Elongation | |
| 72 | 10 min | Elongation | 1 |
| 4 | ∞ | Storing of sample | 1 |

*Tm was determined by: (G + C x 4) + (A + T x 2), where A, T, G and C are the number of nucleotides of the type within the primer sequence.

**Table 2.6: Master mix for Phusion High-Fidelity DNA Polymerase reactions.**

| Component | 50 µl reaction | n* |
|---|---|---|
| Water | 22.5 µl | n x 22.5 µl |
| 5X Phusion HF Buffer | 10 µl | n x 10 µl |
| GC rich solution (Faststart) | 10 µl | n x 10 µl |
| dNTPs | 4 µl | n x 4 µl |
| Phusion DNA Polymerase | 0.5 µl | n x 0.5 µl |
| F primer | 0.5 µl | n x 0.5 µl |
| R primer | 0.5 µl | n x0.5 µl |
| DNA | 2 µl | n x 2 µl |

*n is the number of samples to be PCR amplified including a negative control

**Table 2.7: PCR amplification conditions for Phusion High-Fidelity DNA polymerase.**

| Temperature (°C) | Time | Reaction | Number of Cycles |
|---|---|---|---|
| 98 | 30 s | Activation of Phusion High-Fidelity DNA Polymerase | 1 |
| 98 | 10 s | Template denaturation | |
| Tm* | 30 s | Annealing of primers | 35 |
| 72 | 1 min/kb | Elongation | |
| 72 | 10 min | Elongation | 1 |
| 4 | ∞ | Storing of sample | 1 |

*Tm was determined by: (G + C x 4) + (A + T x 2), where A, T, G and C are the number of nucleotides of the type within the primer sequence.

*2.5.4 Agarose Gel Electrophorsis*

PCR products and restriction enzyme digests were electrophoretically fractionated at 100V on a 1.0 % agarose gel in TAE (40 mM tris, 20 mM acetic acid, 1 mM ethylenediamine tetra acetic acid (EDTA)) with ethidium bromide. Agarose gels were visualized under UV light.

GeneRuler[TM] 100bp plus DNA ladder and GeneRuler[TM] 1kb plus DNA ladder (Fermentas) were used as appropriate to determine band sizes. Where appropriate, bands of the correct size were excised from the agarose gel and purified using the Wizard[®] SV Gel and PCR Cleanup Kit (Promega).

*2.5.5 Adenylation of PCR products*

The MSMEG_0617 PCR product that was amplified with Phusion High-Fidelity DNA polymerase was polyadenylated to allow for T-A cloning into pGEM-T Easy. The blunt end PCR products were purified from the reaction mixture using the Wizard[®] SV Gel and PCR Cleanup Kit (Promega). Hotstart Taq DNA polymerase (Qiagen) was used to add a single adenosine to the PCR products. For sample preparation mix please refer to Table 2.8. Adenylation took place in a single PCR cycle at 72°C for 25 minutes. Thereafter, the adenylated PCR product was purified using the using the Wizard[®] SV Gel and PCR Cleanup Kit (Promega).

**Table 2.8: Sample preparation mix for adenylation of blunt end PCR products.**

|  | 1 X 50 µl |
| --- | --- |
| 10 X buffer | 5 µl |
| $MgCl_2$ | 4 µl |
| dATP | 2 µl |
| Hotstart Taq | 0.3 µl |
| Purified Phusion PCR product | 25 µl |
| $H_2O$ | 13.7 µl |

*2.5.6 Ligations*

Amplified PCR products were ligated into the pGEM-T Easy vector (Promega) according to the manufacturer's instructions. Briefly, 1 µl of pGEM-T Easy vector, 5 µl of 2X ligation buffer, 1 µl of $T_4$ DNA ligase and 3 µl of purified PCR product were incubated at 4°C overnight.

Inserts were ligated into pDMNI using $T_4$ DNA ligase (Promega) as per the manufacturer's instructions. Inserts were ligated to 100 ng of vector at an insert:vector molar concentration of 3:1 using 3 U of $T_4$ DNA ligase. Ligation reactions were incubated at 4°C, overnight.

*2.5.7 Plasmid DNA Transformations*

Electrocompetent *E. coli* was electrotransformed using 2 µl of cloning vector or ligation reaction at 2.5 kV, 25 µF, 125 µFd and 200Ω using the MicroPulser® electroporator (Bio-rad). After transformation 1 ml of SOC (2% glucose (v/v), 2% tryptone (m/v), 0.5% yeast extract (m/v), 1% $MgCl_2$ (v/v), 0.05% NaCl (m/v), 0.25% KCl (v/v), pH adjusted to 7 with 10M NaOH) was added to the cells followed by an incubation period of 1 hour at 37°C with shaking. The transformation was plated out on LB agar supplemented with the appropriate antibiotic and incubated at 37°C overnight.

Freshly made electrocompetent *M. smegmatis* cells were electroporated with 1 µg of final pDMNI construct at 2.5 kV, 25 µF, 125 µFd and a 1000Ω using a Gene Puler® electroporator (Bio-rad). After transformation 1 ml of Difco™ Middlebrook 7H9 was added, followed by an incubation of 3 hours at 37°C

with shaking. The transformation was plated out on BBL<sup>TM</sup> Seven H11 Base plates supplemented with kanamycin and incubated at 37°C for 2-3 days.

*2.5.8 Restriction Enzyme Digestion*

Plasmid DNA was digested with restriction endonucleases *Eco*RI (Roche), *Hind*III (Roche) and *Mfe*I (NEB), as per manufacturer's instructions and in the optimal buffers. All restriction enzyme digestions were performed in a volume of 20 µl, containing 1 ug of vector DNA and 10 units of enzyme. Restriction enzyme digests were incubated at 37°C for 2-5 hours.

*2.5.9 Dephosphorylation of pDMNI*

pDMNI digested with *Eco*RI and *Hind*III was dephosphorylated with rAPid Alkaline Phosphatase (Roche) as per the manufacturer's instructions. Briefly, 1 µg of digested vector DNA, 2 µl of rAPid Alkaline Phosphatase buffer (10 X), 1 µl enzyme and nuclease free water up to 20µl was incubated at 37°C for 30 minutes, followed by the inactivation of the enzyme at 75°C for 2 minutes.

**2.6 Generation of pDMNI constructs with genes of interest**

*2.6.1 Generation of pDMNI0616, pDMNI0623 and pDMNI0626*

The DNA extracted from *M. smegmatis* mc$^2$ 155 was used as the PCR template to amplify the genes of interest, *MSMEG_0616*, *MSMEG_0623* and *MSMEG_0626* using FastStart Taq DNA polymerase (Roche). For primer sequences and amplification conditions refer to Tables 2.3 and 2.4. Amplified genes of interest were separated by agarose gel electrophoresis and the PCR amplified genes of interest were purified using Wizard® SV Gel and PCR Cleanup Kit (Promega) followed by their subsequent ligation into the commercial T-vector, pGEM-T Easy. pGEM-T Easy ligation reactions were transformed into *E. coli* and transformed cells were cultured on LB agar containing ampicillin (50 µg/ml), X-gal and IPTG. For pGEM-T Easy transformations white colonies were picked making use of blue/white selection and PCR screened with the primer sets MSMEG_0615 f and r, MSMEG_0623 f and r and MSMEG_0626 f and r to confirm the presence of the insert. Confirmed transformations were cultured in 10 ml LB liquid media supplemented with ampicillin (50µg/ml) at 37°C, overnight with shaking. Plasmids were extracted using

the Wizard® *Plus* SV Miniprep Plasmid Purification Kit. Purified plasmids were quantified using the Nanodrop® ND1000 Spectrophotometer (Thermo Fisher Scientific) and sequencing was done to verify the integrity of the insert. These constructs were named pGemT0616, pGemT0623 and pGemT0626, respectively.
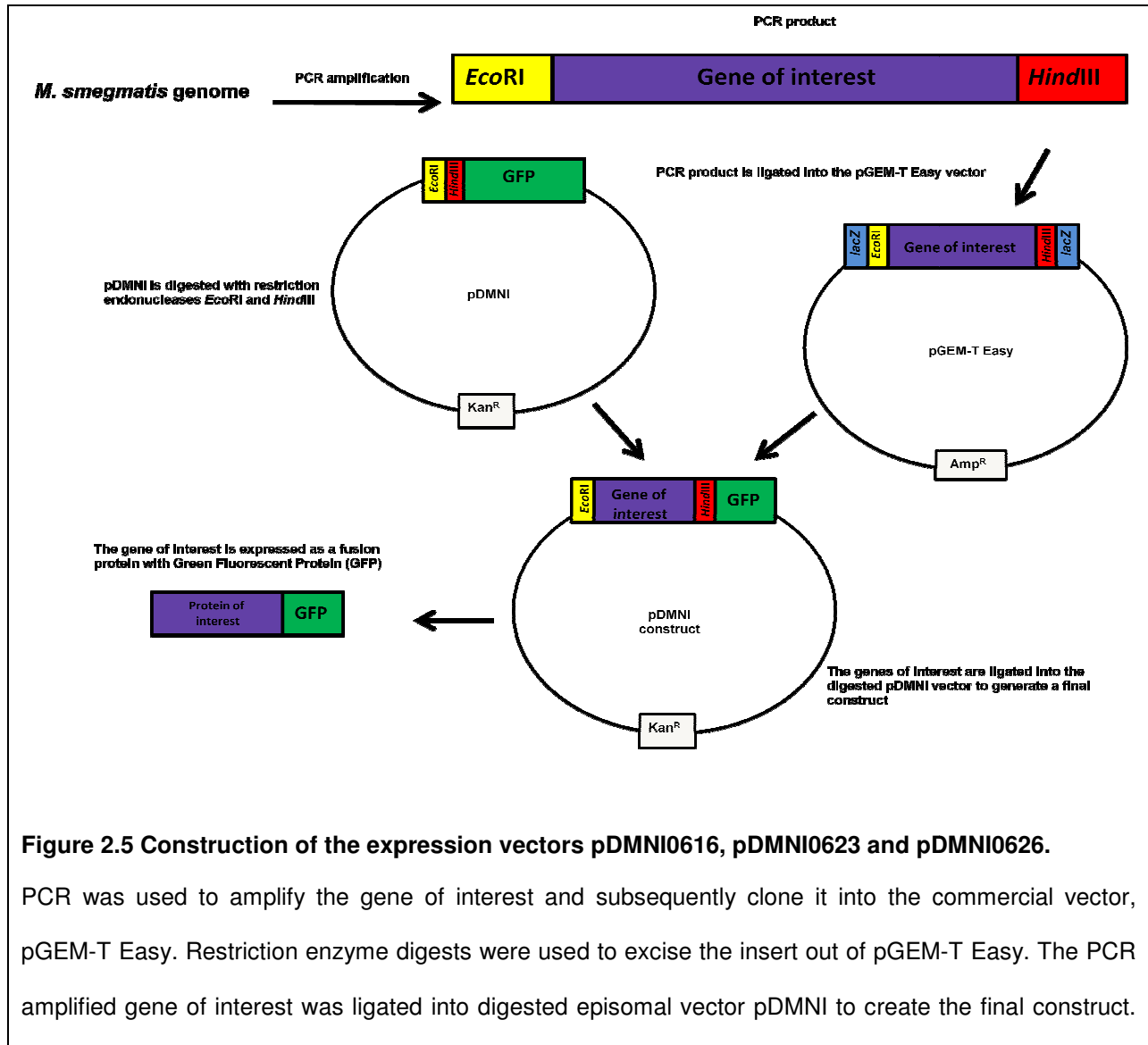


**Figure 2.5 Construction of the expression vectors pDMNI0616, pDMNI0623 and pDMNI0626.**
PCR was used to amplify the gene of interest and subsequently clone it into the commercial vector, pGEM-T Easy. Restriction enzyme digests were used to excise the insert out of pGEM-T Easy. The PCR amplified gene of interest was ligated into digested episomal vector pDMNI to create the final construct.

pGEM-T Easy constructs and the episomal vector, pDMNI were digested with the restriction endonuclease *Eco*RI followed by purification with the Wizard® SV Gel and PCR Cleanup Kit (Promega). Purified digests were subsequently digested with the restriction endonuclease *Hind*III. Restriction digests were separated on a 1% agarose gel and bands with the appropriate size were excised and purified using the Wizard® SV Gel and PCR Cleanup Kit (Promega). Purified inserts were quantified using the Nanodrop® ND 1000 Spectrophotometer (Thermo Fisher Scientific).

Digested pDMNI vector was dephosphorylated with rAPid Alkaline Phosphatase (Roche). Purified inserts were ligated into the episomal vector pDMNI to create the final construct. pDMNI-insert ligations were transformed into *E. coli* and cultured on LB agar containing kanamycin (50 µg/ml) at 37°C, overnight. Colonies were PCR amplified to confirm the presence of the insert followed by the extraction of the final constructs with the use of Wizard® *Plus* SV Miniprep Plasmid Purification Kit from 10 ml *E. coli* cultures. Sequencing was done to verify the integrity of the insert in the final construct. These constructs were named pDMNI0616, pDMNI0623 and pDMNI0626 (Figure 2.5 and 2.6), respectively.
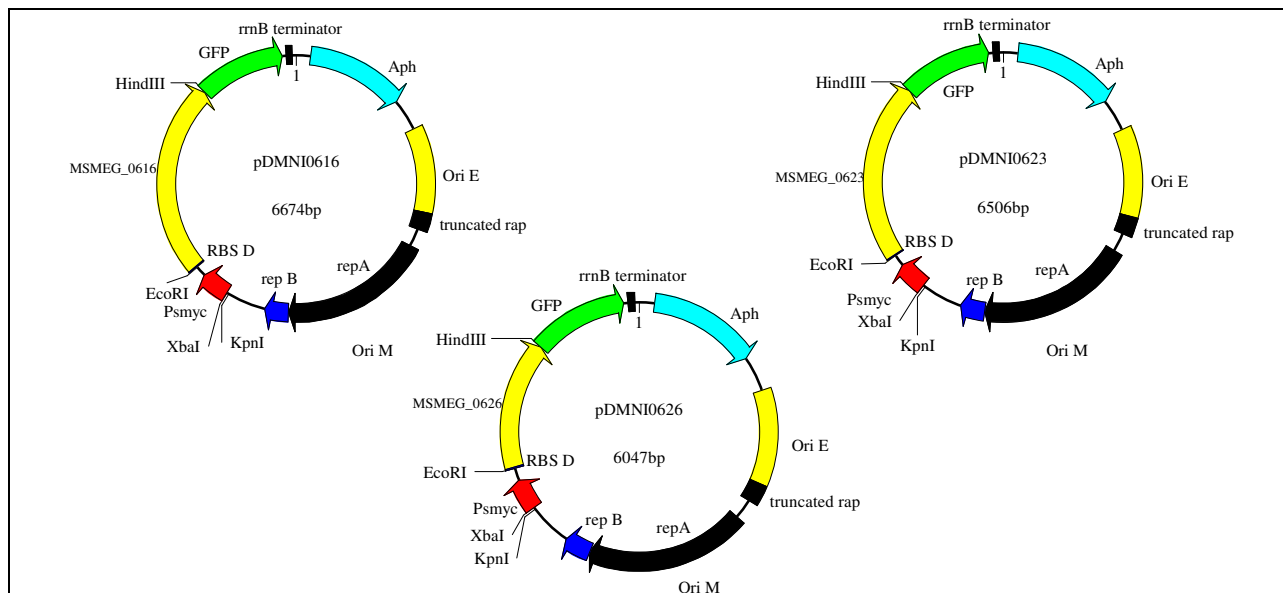


**Figure 2.6 The final constructs pDMNI0616, pDMNI0623 and pDMNI0626.**

The vector maps of the final constructs pDMNI0616, pDMNI0623 and pDMNI0626 shows the genes of interest having been cloned into the episomal vector pDMNI. These genes of interest will be expressed as fusion proteins with GFP.

*2.6.2 Generation of pDMNI0615*

*M. smegmatis* mc$^2$ 155 was used as the PCR template to amplify the gene of interest, *MSMEG_0615* with Phusion High-Fidelity DNA Polymerase. For primer sequences, sample preparation and amplification conditions refer to Tables 2.3, 2.5 and 2.6, respectively. The amplified product was separated by agarose gel electrophoresis on a 1% agarose gel, excised and purified using Wizard$^®$ SV Gel and PCR Cleanup Kit (Promega) followed by its subsequent ligation into the commercial CloneJET vector.

The CloneJET ligation reaction was transformed into *E. coli* and subsequently cultured on LB agar plates containing ampicillin (50 µg/ml). Colonies were picked from the CloneJET transformation plates and PCR screened to confirm the presence of the insert. The plasmid was extracted using the Wizard® *Plus* SV Miniprep Plasmid Purification Kit. The purified plasmid was quantified using the Nanodrop® ND1000 Spectrophotometer (Thermo Fisher Scientific) and sequencing was done to verify the integrity of the insert. This construct was named CJ0615.

The insert was digested out of CJ0615 with the restriction endonucleases *Mfe*I and *Hind*III. The restriction digest was separated on a 1% agarose gel and the band with the appropriate size was excised and purified using the Wizard® SV Gel and PCR Cleanup Kit (Promega). The purified insert was quantified using the Nanodrop® ND 1000 Spectrophotometer (Thermo Fisher Scientific). The insert was subsequently ligated into the *Eco*RI and *Hind*III digested and dephosphorilated pDMNI episomal vector. This pDMNI-insert ligation was transformed into *E. coli* and cultured on LB agar containing (50 µg/ml) at 37℃, overnight. Colonies were PCR amplified to confirm the presence of the insert followed by the extraction of the final construct with the use of Wizard® *Plus* SV Miniprep Plasmid Purification Kit from 10 ml *E. coli* cultures. Sequencing was done to verify the integrity of the insert in the final construct. This final construct was named pDMNI0615 (Figure 2.7).
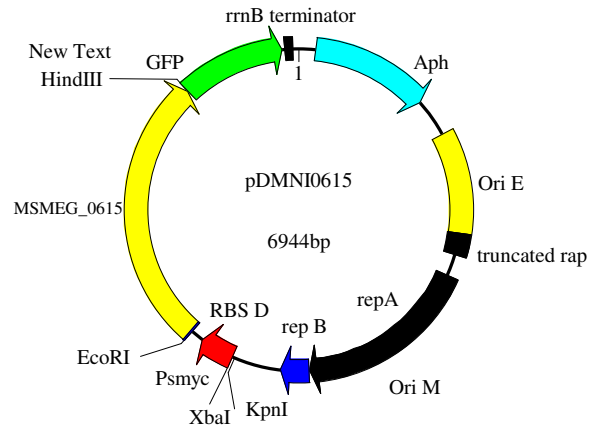
**Figure 2.7 The final construct pDMNI0615.**

The vector map of the final construct pDMNI0615 shows the gene of interest having been cloned into the episomal vector pDMNI. *MSMEG_0615* will be expressed as a fusion protein with GFP.

*2.6.3 Generation of pGemT0617*

*M. smegmatis* mc$^2$ 155 was used as the PCR template to amplify the gene of interest, *MSMEG_0617* with Phusion High-Fidelity DNA Polymerase. For primer sequences, sample preparation and amplification conditions refer to Tables 2.3, 2.5 and 2.6, respectively. The amplification product was separated by agarose gel electrophoresis on a 1% agarose gel. The PCR amplified MSMEG_0617 was purified using Wizard® SV Gel and PCR Cleanup Kit (Promega). The purified PCR product was adenylated to allow for T-A cloning with pGEM-T Easy followed by purification with the Wizard® SV Gel and PCR Cleanup Kit (Promega) and quantification using the Nanodrop® ND1000 Spectrophotometer (Thermo Fisher Scientific). The insert was subsequently ligated into the commercial T-vector, pGEM-T Easy.

The pGEM-T Easy ligation reaction was transformed into *E. coli* and cultured on LB agar containing ampicillin (50 µg/ml), X-gal and IPTG. White colonies were selected making use of blue/white selection and PCR screened to confirm the presence of the insert. The confirmed transformations were cultured in 10 ml LB liquid media containing ampicillin (50µg/ml) at 37℃, overnight with shaking. Plasmids were extracted using the Wizard® *Plus* SV Miniprep Plasmid Purification Kit. The purified plasmids were quantified using the Nanodrop® ND1000 Spectrophotometer (Thermo Fisher Scientific) and sequencing

was done with the primer set Sp6 and T7 as well as the primers MSMEG0617seq800, MSMEG0617seq1600, MSMEG0617seq2400 and MSMEG0617seq3200 to verify the integrity of the insert. This construct was named pGemT0617.

## 2.7 *In silico* transmembrane structure analysis of the ESX-3 components of *M. smegmatis*

*In silico* analyses were used to predict the structure of the proteins encoded by the genes of interest located in ESX-3. Protein sequence data was obtained from the JVCI CMR *Mycobacterium smegmatis* MC3 Genome Page (http://cmr.jcvi.org/cgi-bin/CMR/GenomePage.cgi?org=gms) and subjected to the TMHMM Server v2.0 transmembrane prediction analysis (TMHMM, http://www.cbs.dtu.dk/services/TMHMM/) to determine the transmembrane structure of the ESX-3 components. Proteins sequence data was subsequently subjected to the SignalP server (SignalP 4.0 Server, http://www.cbs.dtu.dk/services/SignalP/) to discriminate between signal peptides and transmembrane regions (Petersen *et al.* 2011).

## 2.8 Expression of GFP tagged proteins in *M. smegmatis*

### 2.8.1 Transformation of the expression vectors into M. smegmatis

Electrocompetent wild type and ESX-3 knock-out strains of *M. smegmatis* were transformed with 1 μg of the final constructs, pDMNI, pDMNI0615, pDMNI0616, pDMNI0623 and pDMNI0626. These transformations were cultured on Middlebrook 7H11 plates containing kanamycin (25 μg/ml) at 37°C for 3 to 4 days. Transformations were confirmed by PCR.

### 2.8.2 Whole cell lysate preparation

Whole cell lysates were prepared from *M. smegmatis* and *M. smegmatis* ESX-3 knockout cultures that contained the episomal vectors pDMNI, pDMNI0615, pDMNI0616, pDMNI0623 and pDMNI0626. Cultures of 50 ml were grown to an $OD_{600}$ = 0.6 – 0.8. Cultures were spun down in 50 ml conical tubes at 3220xg for 10 minutes at 4°C. The cell pellet was resuspended in 500 μl ice cold PBS containing Protease Inhibitor Cocktail set III (Calbiochem) and transferred to a 2 ml ribolyser tube with glass beads. The cells were lysed by ribolysing in a BIO 101/Servant FastPrep FP120 4 times for 20 seconds, with 1 minute

40

intervals on ice. Samples were thereafter incubated for 5 minutes on ice and centrifuged in 1.5 ml safe-lock tubes (eppendorf) for 5 minutes at 16 873xg at 4°C. The supernatant, containing the whole cell lysate, was removed and stored at -80°C.

### 2.8.3 Protein concentration determination

Protein concentrations of whole cell lysates were determined spectrophotometrically (Ultrospec 4051, LKB Biochrom) using the Bio-Rad *RC DC* Protein Assay according to manufacturer's instructions. The Quick Start™ Bovine Serum Albumin Standard Set (2 mg/ml) was used to generate a standard curve ranging from 0 mg/ml to 2 mg/ml. A standard curve was plotted according to the BSA absorbance readings at 595 nm and whole cell lysate concentrations were determined from the standard curve.

### 2.8.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

All reagents and buffers used for SDS-PAGE are described in Table 2.9 and 2.10. Whole cell lysates prepared from *M. smegmatis* and *M. smegmatis* ESX-3 knock-out strains containing the constructs pDMNI, pDMNI0615, pDMNI0616, pDMNI0623 and pDMNI0626 were separated on 12% SDS-PAGE gels.

### 2.8.4.1 Sample Preperation

Five microliter of 2X SDS sample buffer (refer to Table 2.10) was added to 1µg of protein sample. The sample was mixed and incubated at 100°C for 5 minutes.

### 2.8.4.2 Gel preparation

The gel apparatus was assembled according to the manufacturer's instructions (Bio-Rad).

The separating gel was prepared according to Table 2.11 and poured into the apparatus. A layer of Isopropanol was added onto the surface of the separating gel to remove bubbles and level the surface of the gel. The separating gel was allowed 20 minutes to set and the isopropanol was poured off. The stacking gel was prepared according to Table 2.12 and poured over separating gel until the apparatus was full. The comb was inserted and the gel was allowed 20 minutes to set.

**Table 2.9: SDS-PAGE Reagents.**

| Reagent | Concentration | Diluent | Storage | Source |
|---|---|---|---|---|
| 10% SDS | 10% (m/v) | Water | Room temperature | Sigma-Aldrich |
| 20 % SDS | | | | |
| EDTA | 0.5 M | Water | Room temperature | Fluka |
| Bromophenol Blue | 0.75% (m/v) | water | Room temperature | Fluka |
| Ammonium persulphate (APS) | 15% (m/v) | Water | 4°C | Sigma-Aldrich |
| Acrylamide/bisacrylamide | 30:0.8% (m/v) | | 4°C, dark container | Bio-Rad |
| TEMED | N,N,N',N'-Tetramethylethylenediamine | | Room temperature, dark container | Sigma-Aldrich |
| β-Mercaptoethanol | | | Room temperature, dark container | |
| Glycine | | | Room temperature | Merck |
| Tris | tris(hydroxymethyl)aminomethane | | Room temperature | Merck |
| Glycerol | | | Room temperature | Merck |

**Table 2.10: SDS-PAGE Buffers.**

| Buffer | Component | Volume/Weight/ Concentration | Diluent | Storage | pH |
|---|---|---|---|---|---|
| Separating buffer | Tris | 1.5 M | Water | 4°C | pH 8.8 (adjust with HCl) |
| Stacking buffer | Tris | 1.0 M | Water | 4°C | pH 6.8 (adjust with HCl) |
| 2X Reducing SDS sample buffer | Stacking buffer | 3.4ml | | 4°C or -20°C | |
| | Glycerol | 2 ml | | | |
| | SDS (20%) | 3 ml | | | |
| | Bromophenol Blue | 500 µl | | | |
| | EDTA | 200 µl | | | |
| | β-Mercaptoethanol | 1 ml | | | |
| Running buffer (2L) | Tris | 6 g | Water | 4°C or Room temperature | |
| | Glycine | 28.8 g | | | |
| | SDS | 20 ml | | | |

**Table 2.11: Separating gel composition.**

| Component | 12% Separating Gel (2 x 0.75 mm mini-slab gels) |
|---|---|
| Acrylamide/bisacrylamide | 4.0 ml |
| 1.5 M Tris-HCl | 2.5 ml |
| 10% SDS | 0.1 ml |
| Water | 3.35 ml |
| TEMED | 10 µl |
| APS | 50 µl |

**Table 2.12: Stacking gel composition.**

| Component | 3% Stacking Gel (4 x 0.75 mm mini-slab gels) |
|---|---|
| Acrylamide/bisacrylamide | 1.3 ml |
| 1.0 M Tris-HCl | 1.25 ml |
| 10% SDS | 0.1 ml |
| Water | 7.4 ml |
| TEMED | 20 µl |
| APS | 50 µl |

*2.8.4.3 Electrophoresis*

Gels were placed in the running chamber and running buffer was poured into the upper and lower chambers of the apparatus to cover the top and the bottom of the gel. Combs were removed and PageRuler[TM] Prestained protein Ladder (Fermentas) was loaded as a molecular weight marker. Samples were loaded into the wells and the proteins were separated by running the gel at 120V until the dye front reached the end of the gel.

*2.8.4.4 Coomassie blue staining*

SDS-PAGE gels were soaked in Coomassie blue stain for 2 hours with shaking at room temperature and destained overnight with destaining solution. The Coomassie blue staining and destaining solutions used are described in Table 2.13.

43

**Table 2.13: Staining and Destaining solution for SDS-PAGE.**

| Solution | Component | Volume/Weight |
|---|---|---|
| Coomassie blue staining solution | Coomassie blue | 2.5 g |
| | Methanol | 450 ml |
| | Water | 450 ml |
| | Acetic Acid | 100 ml |
| Destaining solution | Glycerol | 100 ml |
| | Glacial Acetic Acid | 375 ml |
| | Methanol | 250 ml |
| | Water | 4275 ml |

*2.8.5 Western Blotting*

Buffers used for Western Blotting are described in Table 2.14. A duplicate SDS-PAGE gel was subjected to Western blotting using a mouse anti-GFP antibody (Invitrogen) and a goat anti-mouse horse radish peroxidase (HRP) conjugated antibody (Invitrogen) to confirm the expression of the fusion proteins.

*2.8.5.1 Protein transfer to nitrocellulose membrane*

Transfer buffer was used to presoak 2 Western blot pads, 2 Whatmann 3M sheets and 1 HybondC-extra nitrocellulose or Hybond-P PVDF membrane in a large tray. Upon the completion of electrophoresis, the blotting apparatus was stacked on the black side of the sandwich in transfer buffer: Western blot pad, Whatmann sheet, SDS-PAGE gel, nitrocellulose or PVDF membrane, Whatmann sheet and Westernblot blot pad. The closed sandwich was placed in the Westernblot apparatus with an ice pack and a magnetic stirrer bar. The apparatus was filled with ice cold transfer buffer and placed on a magnetic stirrer in the cold room (4°C). A 100V voltage was applied to the apparatus for 60 minutes during which the buffer becomes milky with bubbles when current is flowing. Thereafter the membrane was blocked with 10% blocking buffer overnight at 4°C.

**Table 2.14: Western Blotting buffers.**

| Buffer | Composition | Diluent | pH | Storage |
|---|---|---|---|---|
| Blotting/ Transfer buffer | 25 mM Tris<br>192 mM Glycine<br>20% Methanol | Water | pH should be 8.3, do not adjust | 4°C |
| TBS-T Wash Buffer | 20 mM Tris<br>137 mM NaCl<br>Tween-80 | Water | | Room temperature |
| Blocking Buffer | 10% m/v fat free milk powder<br>1% m/v Bovine serum albumin<br>10 mM Azide | TBS-T | | Room temperature |

*2.8.5.2 Western detection*

The blocked membrane was rinsed twice with TBS-T buffer and washed three times for 10 minutes by shaking in TBS-T buffer. The blocked membrane was incubated with 20 ml the primary antibody, a mouse anti-GFP antibody (Invitrogen) (1:500 dilution with TBS-T buffer) for 1 hour at room temperature, with shaking. The membrane was rinsed twice with TBS-T buffer and washed three times for 10 minutes with shaking in TBS-T buffer.  The membrane was subsequently incubated with 20ml secondary antibody, a goat anti-mouse HRP conjugated antibody (Invitrogen) (1:2000 dilution with TBS-T buffer) for 1 hour at room temperature with shaking.  The membrane was rinsed twice with TBS-T buffer, followed by washing with TBS-T buffer for 10 minutes by shaking at room temperature. The membrane was placed in a sealable plastic bag.

The membrane was incubated with ECL detection fluid (ECL Plus Western Blotting Detection System, Amersham[TM]) in a dark room according to manufacturer's instructions. Once the detection fluid was drained from the sealable plastic bag, the membrane was placed in an X-ray cassette. Signal was detected for 10-30 seconds with autoradiographic film.

**2.9 Localisation of ESX-3 components in *M. smegmatis* and *M. smegmatis* ESX-3 knock-out**

*2.9.1 Fluorescent microscopy*

Fluorescent microscopy was done at the Imaging Unit of the Central Analytical Facility at Stellenbosch University (http://academic.sun.ac.za/saf/services/services6_b.html).

Fluorescent microscopy was used to determine the sub-cellular localization of the ESX-3 components expressed from the final constructs in wild type and ESX-3 knock-out strains of *M. smegmatis*. Cells were scored as having 1 fluorescent focus, multiple foci, membrane associated fluorescence, diffuse fluorescence or no fluorescence. Images were acquired on an Olympus Cell^R system attached to an IX81 inverted fluorescence microscope equipped with an F-view-II cooled CCD camera (Soft Imaging Systems). A Xenon-Arc burner (Olympus Biosystems GMBH) as was used as a light source. Images were acquired using a 470 nm excitation filter and emission was collected with a UBG triple-bandpass emission filter cube (Chroma).

Wild type and ESX-3 knock-out strains of *M. smegmatis* were grown to an $OD_{600}$=0.7. Ten microliters of cultures were placed in Lab-tek Chambered #1.0 Borosilicate Coverglass systems (eight chamber) (Amersham). An Olympus UPlanSApo 100x/1.4 oil objective was used to acquire images and Cell^R imaging software was used to process the images and to subtract the background. A light intensity of 32,29% was used to excite the GFP.

*2.9.2 Transmission electron microscopy*

Transmission electron microscopy was performed at the Electron Microscopy Unit at the University of Cape Town. All sample buffers are described in Table 2.15.

*M. smegmatis* wild type and ESX-3 knock-out cultures transformed with the final constructs pDMNI0615, pDMNI0616 and pDMNI0626 were analyzed using transmission electron microscopy (TEM). Cultures were grown to an $OD_{600}$=0.7 and spun down at 3220xg for ten minutes at 4°C. Cells were washed with 1 X PBS, pH7.4 and resuspended in 1 ml 1 X PBS. The cell pellet was spun down at 1377xg for 10 minutes and the supernatant removed. The cell pellets were resuspended in a volume of 20 times greater than

46

that of the pellet with freshly prepared TEM fixative. Cells were fixed for 2 hours at 4°C followed by centrifugation at 1377xg for 10 minutes. The fixed cells were resuspended in 1 ml 1X PBS and stored overnight at 4°C. Fixed cells were centrifuged at 1377xg for 2 minutes and washed with water, the wash step was repeated. The cell pellets were embedded in 2.5% low melting agarose followed by series of dehydration steps for 10 min with different concentrations of ethanol (30%. 50%, 70%, 80%, 90% and 100%). Cells were incubated in a 1:1 LR white resin and 100% ethanol solution for 1 hour before being embedded in LR white resin. Samples were sectioned at the Electron Microscope unit at the University of Cape Town on a Leica EM UC7 ultramicrotome (Leica Microsystems). Samples were cut into 120 nm thin sections with a diamond knife. Sections were collected on formvar coated nickel grids and were allowed to dry overnight.  The nickel grids were incubated on a 0.1M Glycine solution for 20 minutes to quench free aldehyde groups and were washed twice with TBS-Tween for 2 minutes. Sections were incubated for 30 minutes with blocking buffer followed by 2 hour incubation with the primary antibody (1:50 dilution), a mouse anti-GFP antibody, at room temperature. The sections were washed six times for 2 minutes with TBS-Tween and subsequently incubated with the secondary antibody (1:50 dilution), an anti-mouse IgG (whole molecule) gold conjugated antibody, for 1 hour at room temperature. Sections were washed six times for two minutes with TBS-Tween. Post-fixation of sections occurred on droplets of 2% gluteraldehyde in 0.1M phosphate buffer for 10 minutes followed by six wash steps of 2 minutes with TBS-Tween. Sections were rinsed with distilled water and subsequently stained with uranyl acetate for 10 minutes, followed by staining with lead citrate for 5 minutes before being viewed on a FEI Tecnai TF20 transmission electron microscope (FEI) fitted with a 4kX4k Gatan CCD camera. Images were collected and saved in dm3 format before being converted to jpg images.

**Table 2.15 Transmission electron microscopy buffers.**

| Buffer | Component | Volume/Weight/Concentration | Diluent | Storage | pH |
|---|---|---|---|---|---|
| TEM Fixative | 25 % Gluteraldehyde | 200 µl | water | Room temperature | pH 7.6 |
| | 4 % Formaldehyde | 4 ml | | | |
| | 1 x PBS | 1 ml | | | |
| 0.1 M Glycine solution | Glycine | 0.75 g | Water | Room temperature | pH 7.6 |
| | PBS | 0.1 M | | | |
| TBS-Tween (1 L) | Tris | 6.1 g | Water | Room temperature | pH 7.6 |
| | NaCl | 9 g | | | |
| | Tween 20 | 0.5 ml | | | |
| Blocking Buffer | BSA | 1 g | | Room temperature | pH 7.6 |
| | Cold Fish Gelatin | 0.1 ml | | | |
| | 0.05 M TBS | 100 ml | | | |
| Dilution Buffer | BSA | 0.1 g | | Room temperature | |
| | 0.05 M TBS | 100 ml | | | |
| | Tween 20 | 30 µl | | | |
| 2% Gluteraldehyde PBS | 25 % Gluteraldehyde | 4 ml | | 4 °C | |
| | 0.1 M PBS | 46 ml | | | |

# CHAPTER 3

## Results

**3.1. Protein structure analysis of ESX-3 components in *M. smegmatis***

The transmembrane structure of the proteins of interest, MSMEG_0615, MSMEG_0616, MSMEG_0617, MSMEG_0623 and MSMEG_0626 from ESX-3 in *M. smegmatis* was determined *in silico*, using the TMHMM Server v2.0 transmembrane prediction server.

*3.1.1 MSMEG_0615 (EccA)*

MSMEG_0615 is a conserved protein of 608 amino acids and is predicted to be an AAA+ ATPase that shows homology to the CBXX/CFQX and SpoVK family proteins. The results of the *in silico* analysis predicted that MSMEG_0615 is a cytoplasmic ATPase with no transmembrane domains that are involved in the secretion of ESX-3 substrates (Figure 3.1). No signal sequence was identified in the MSMEG_0615 protein (Figure 3.2).



**Figure 3.1 MSMEG_0615 is predicted to be a cytoplasmic protein.**

The TMHMM transmembrane prediction of MSMEG_0615 indicates that this protein does not possess any transmembrane regions, confirming that MSMEG_0615 is a cytoplasmic protein.
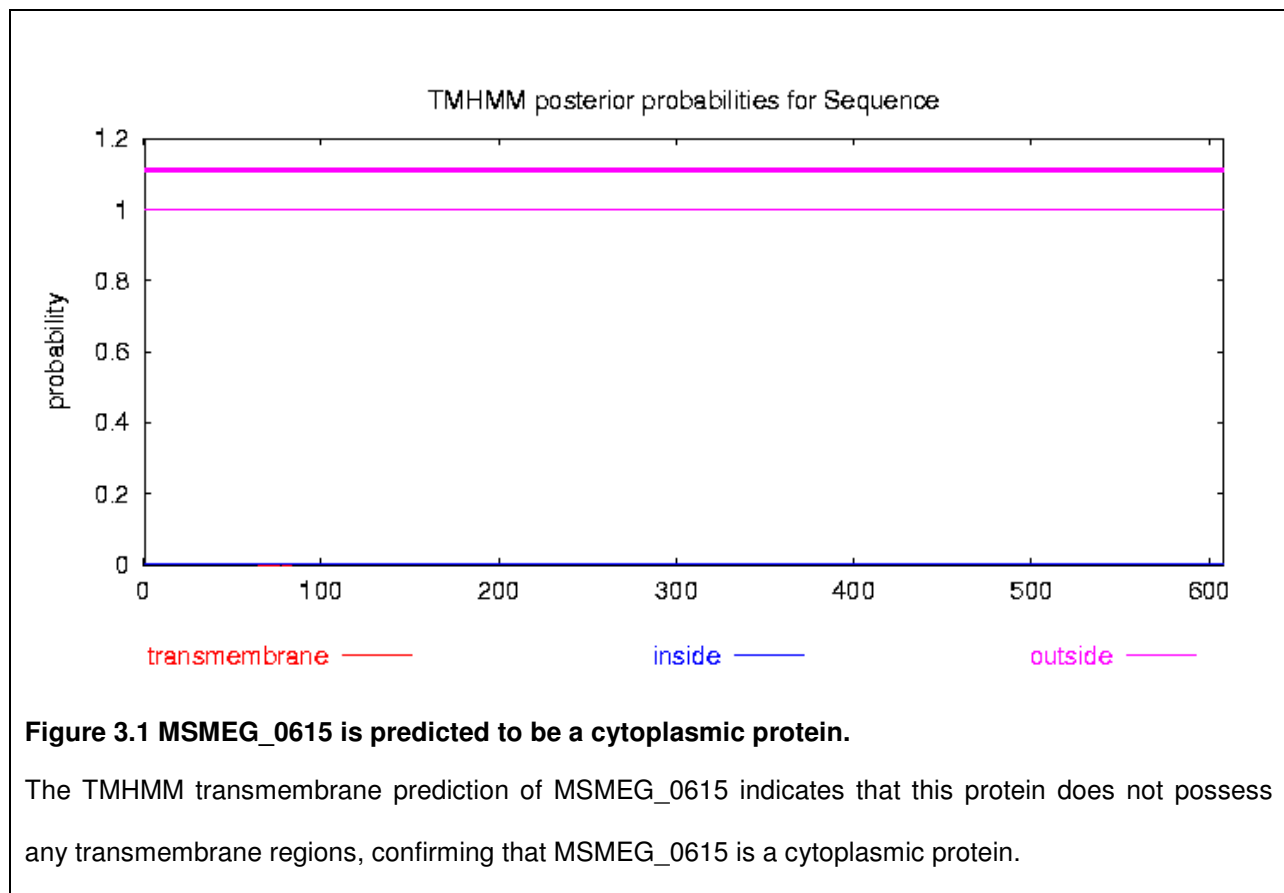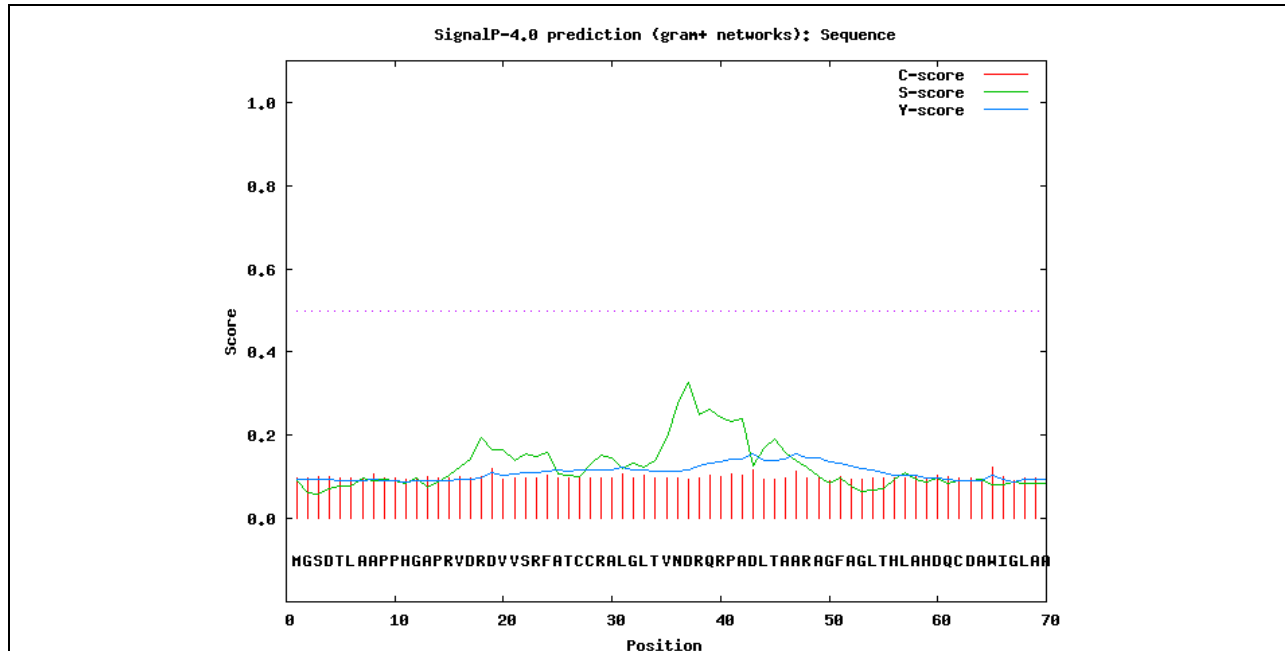
**Figure 3.2 SignalP analysis of MSMEG_0615 protein.**

The SignalP signal sequence prediction of MSMEG_0615 did not identify any signal sequences in the first 70 amino acids of MSMEG_0615.

*3.1.2 MSMEG_0616 (EccB)*

MSMEG_0616 is a conserved hypothetical protein of a 518 amino acids. The results of the *in silico* analysis predict that MSMEG_0616 contains a transmembrane region at the N-terminal domain of the protein. It is also predicted that the C-terminal domain of the MSMEG_0616 is located on the outside of the mycobacterial plasma membrane (Figure 3.3). No signal sequence was identified in the MSMEG_0616 protein confirming that this protein does have a transmembrane region (Figure 3.4).

51
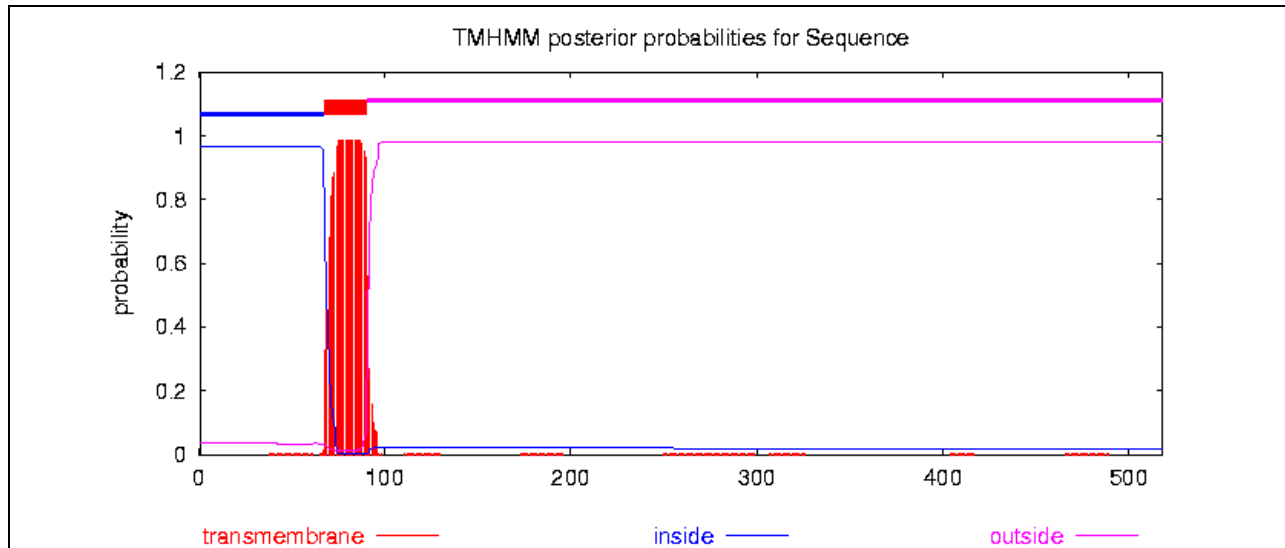
**Figure 3.3 MSMEG_0616 is an N-terminal membrane-associated protein.**

The TMHMM transmembrane prediction of MSMEG_0616 indicates that this protein is an N-terminal

transmembrane protein.



**Figure 3.4 SignalP analysis of MSMEG_0616 protein.**

The SignalP signal sequence prediction of MSMEG_0616 indicates that the first 120 amino acids of

MSMEG_0616 does not possess a signal sequence, confirming that MSMEG_0616 has a

transmembrane region.

*3.1.3 MSMEG_0617 (EccC)*

MSMEG_0617 is a conserved protein of 1325 amino acids and is predicted to be an FtsK/SpoIIIE family protein with 3 ATP/GTP binding sites. FtsK/SpoIIIE family proteins are predicted to be DNA translocases that are involved in the translocation of chromosomal DNA during conjugative DNA transfer and septum formation (Crozat, Grainge 2010, Marquis *et al.* 2008). MSMEG_0617 might also be involved in providing energy for the translocation of ESX-3 substrates due to the 3 ATP/GTP binding sites. TMHMM also predicts that the C-terminal region of MSMEG_0617 is on the outside of the mycobacterial membrane (Figure 3.5). No signal sequence was identified in the first 200 amino acids of MSMEG_0617 protein, confirming that this protein does have a transmembrane region (Figure 3.6).



**Figure 3.5 MSMEG_0617 is predicted to be an N-terminal transmembrane protein.**

The TMHMM transmembrane prediction of MSMEG_0617 indicates that this protein is an N-terminal transmembrane protein with a possible signal sequence.
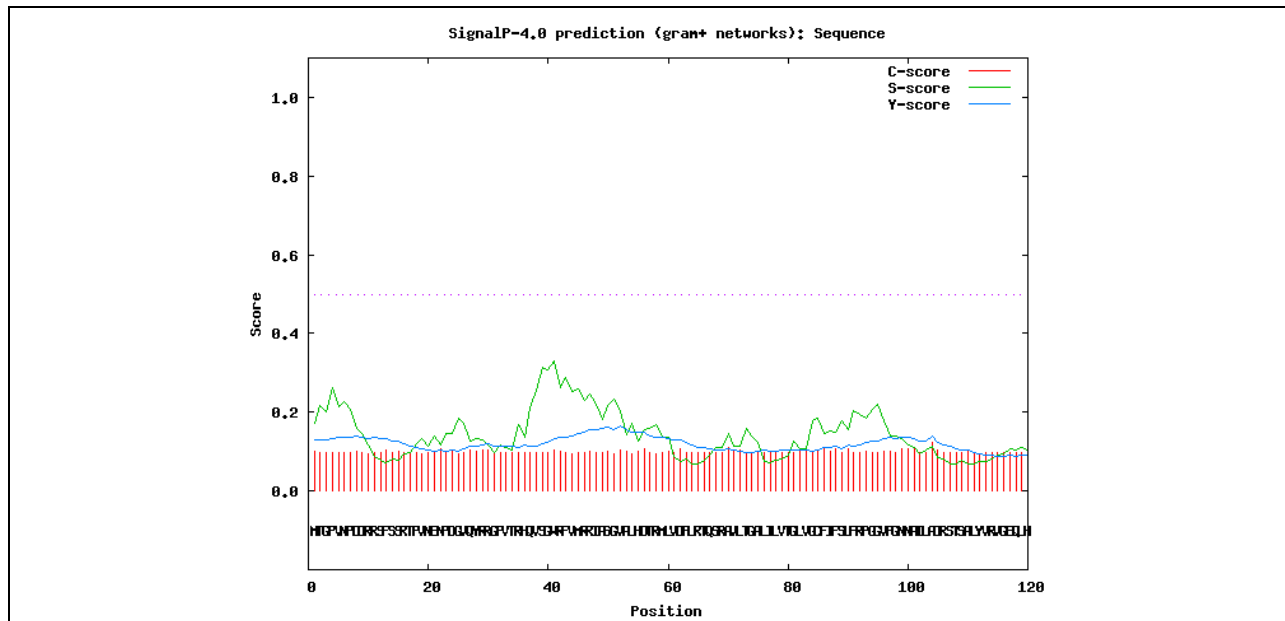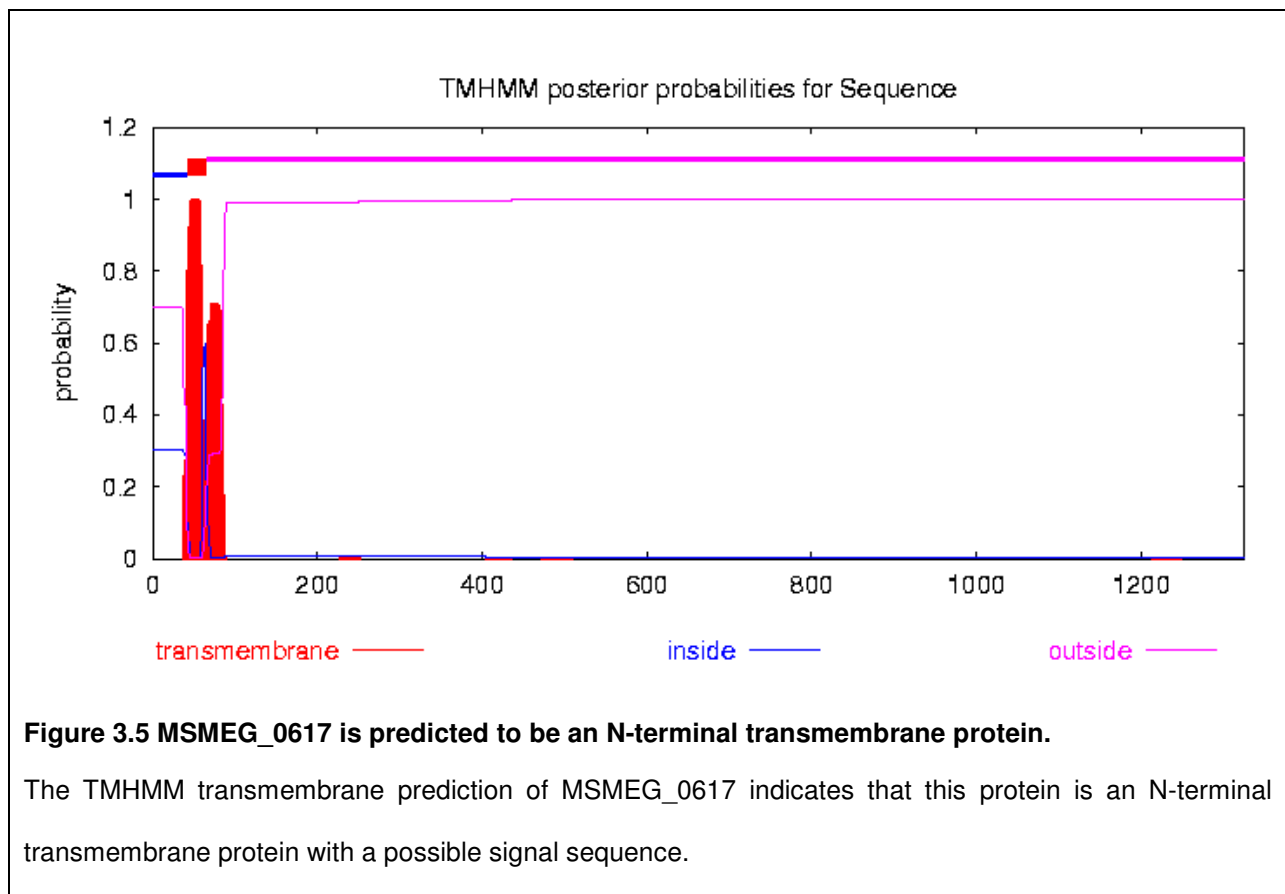
**Figure 3.6 SignalP analysis of MSMEG_0617 protein.**

The SignalP signal sequence prediction of MSMEG_0616 indicates that the first 200 amino acids of MSMEG_0616 does not possess a signal sequence and that the transmembrane region identified is a true transmembrane region.

*3.1.4 MSMEG_0623 (EccD)*

MSMEG_0623 is a conserved integral membrane protein of 462 amino acids. This protein is predicted to have 10-11 transmembrane regions and is hypothesized to make up a membrane pore or channel through which ESX-3 substrates are transported to the outside of the mycobacterial membrane (Feltcher, Sullivan & Braunstein 2010). The TMHMM prediction shows that this protein has 10 transmembrane regions and that the C-terminal region of MSMEG_0623 is on the inside of the mycobacterial membrane (Figure 3.7), however it is possible that there are 11 transmembrane regions which will result in the C-terminal region of MSMEG_0623 being on the outside of the mycobacterial membrane. No signal sequence was identified in the first 200 amino acids of the MSMEG_0623 protein, confirming that the transmembrane regions identified in the first 200 amino acids are not signal sequences but transmembrane regions (Figure 3.8).
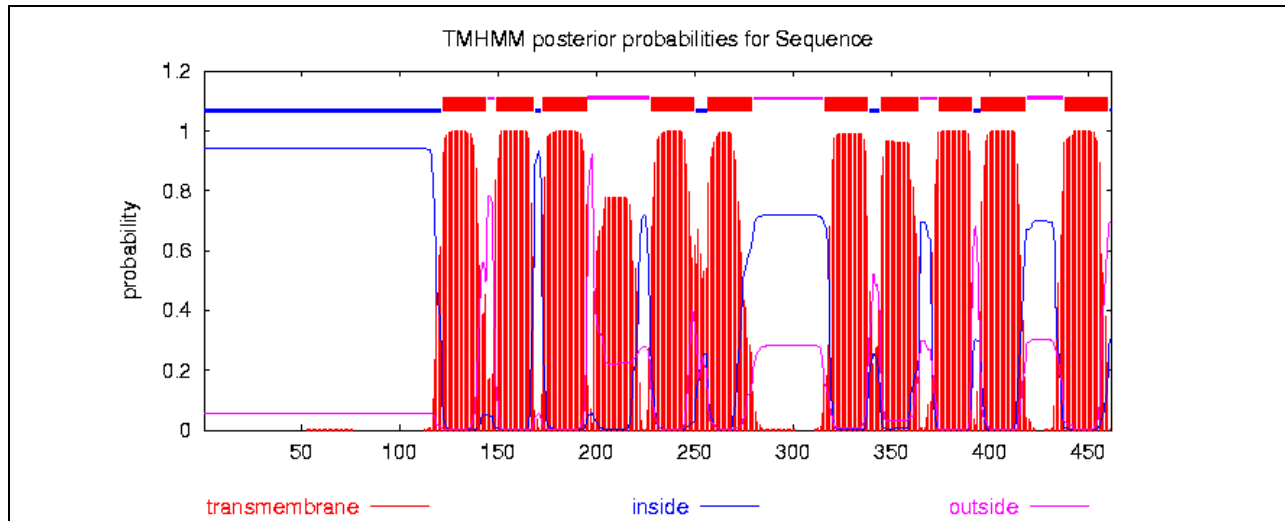
**Figure 3.7 MSMEG_0623 is predicted to be an integral membrane protein with 10-11 transmembrane regions.**

The TMHMM transmembrane prediction of MSMEG_0623 indicates that this protein is an integral membrane protein with 10-11 transmembrane regions.
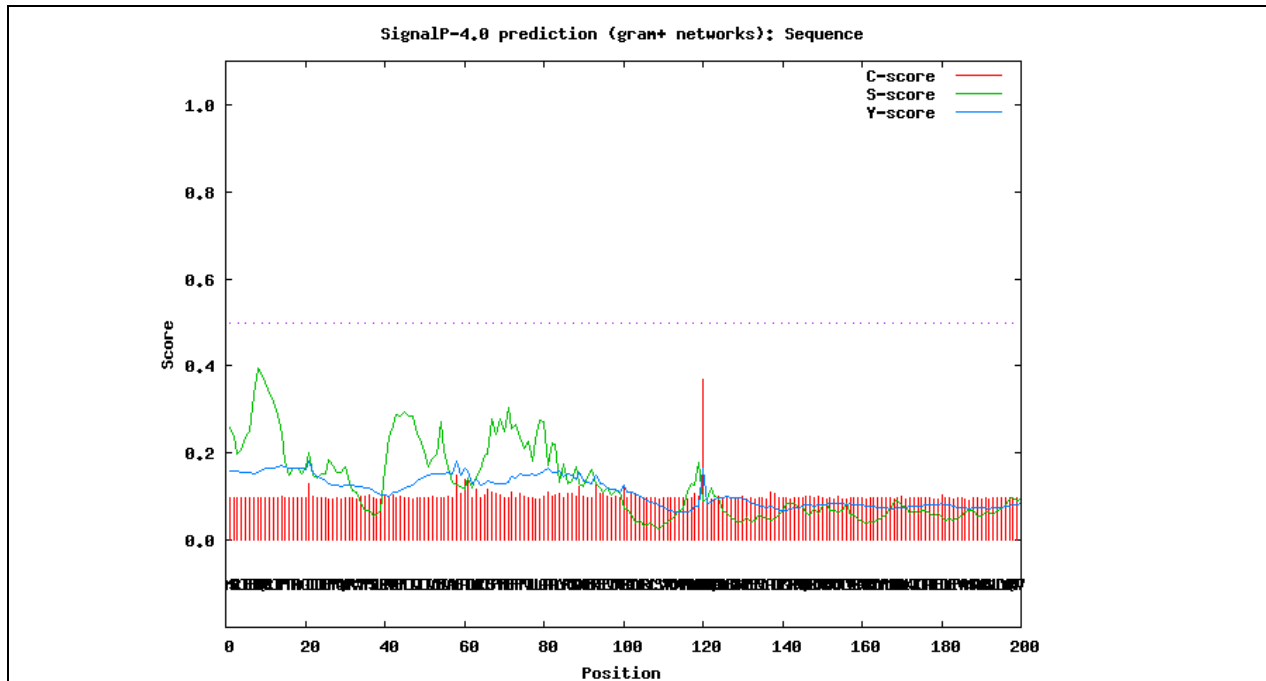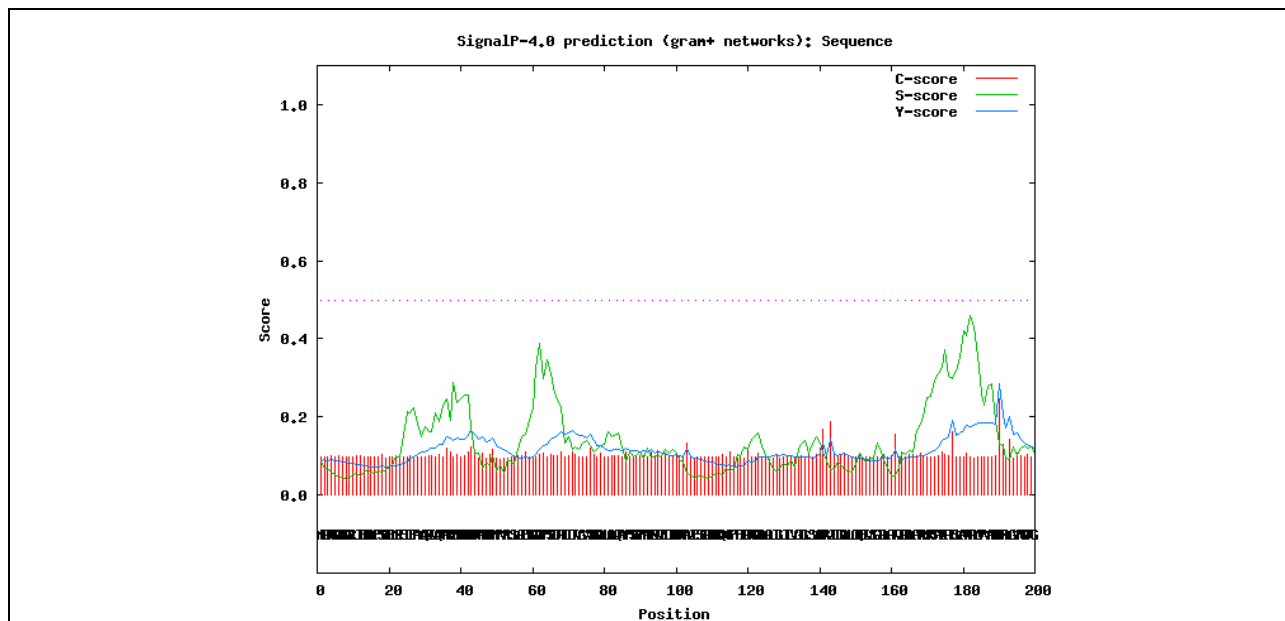


**Figure 3.8 SignalP analysis of MSMEG_0623 protein.**

The SignalP signal sequence prediction of MSMEG_0623 indicates that the first 200 amino acids of MSMEG_0623 do not possess a signal sequence confirming that the transmembrane regions identified by TMHMM are not signal sequences.

*3.1.5 MSMEG_0626 (EccE)*

MSMEG_0626 is a conserved hypothetical protein of a 309 amino acids. It is predicted that MSMEG_0626 has two N-terminal transmembrane regions and that the C-terminal of the protein is at the outside of the mycobacterial membrane (Figure 3.9). A signal sequence was identified in the first 20 amino acids of the MSMEG_0626 protein indicating that MSMEG_0626 has only one transmembrane region (Figure 3.10).



**Figure 3.9 MSMEG_0626 is predicted to have two N-terminal transmembrane regions.**
The TMHMM transmembrane prediction of MSMEG_0626 indicates that this protein has two N-terminal transmembrane regions.

**Figure 3.10 SignalP analysis of the MSMEG_0626 protein.**

The SignalP signal sequence prediction of MSMEG_0626 identified a signal sequence in the first 20 amino acids, indicating that this protein has one transmembrane region.

**3.2 Generation of pDMNI constructs**

*3.2.1 Generation of pDMNI0616, pDMNI0623 and pDMNI0626*

The genes of interest, *MSMEG_0616*, *MSMEG_0623* and *MSMEG0626*, were PCR amplified from *M. smegmatis* mc$^2$155 genomic DNA using designed primer sets. PCR products were ligated into pGEM-T Easy and transformed into electrocompetent *E. coli*. Transformants were selected on LB agar plates containing ampicillin, X-gal and IPTG and white colonies were PCR screened to confirm the presence of the inserts (Figure 3.11). These constructs were named pGemT0616, pGemT0623 and pGemT0626, respectively. DNA sequencing confirmed the integrity of the cloned genes.

57

**Figure 3.11 Cloning of MSMEG_0616, MSMEG_0623 and MSMEG_0626 into pGEM-T Easy.**

PCR using the primer sets MSMEG_0616, MSMEG_0623 and MSMEG_0626 was done to confirm the presence of the inserts in pGEM-T Easy. Lane 1 contains GeneRuler™ 1 kb Plus DNA Ladder. Lanes 2, 4 and 6 contains the PCR products amplified from the constructs pGemT0616 (1596 bp), pGemT0623 (1428 bp) and pGemT0626 (969 bp), respectively. Lane 3, 5 and 7 contains PCR reactions with no DNA and the appropriate primers, as negative controls.
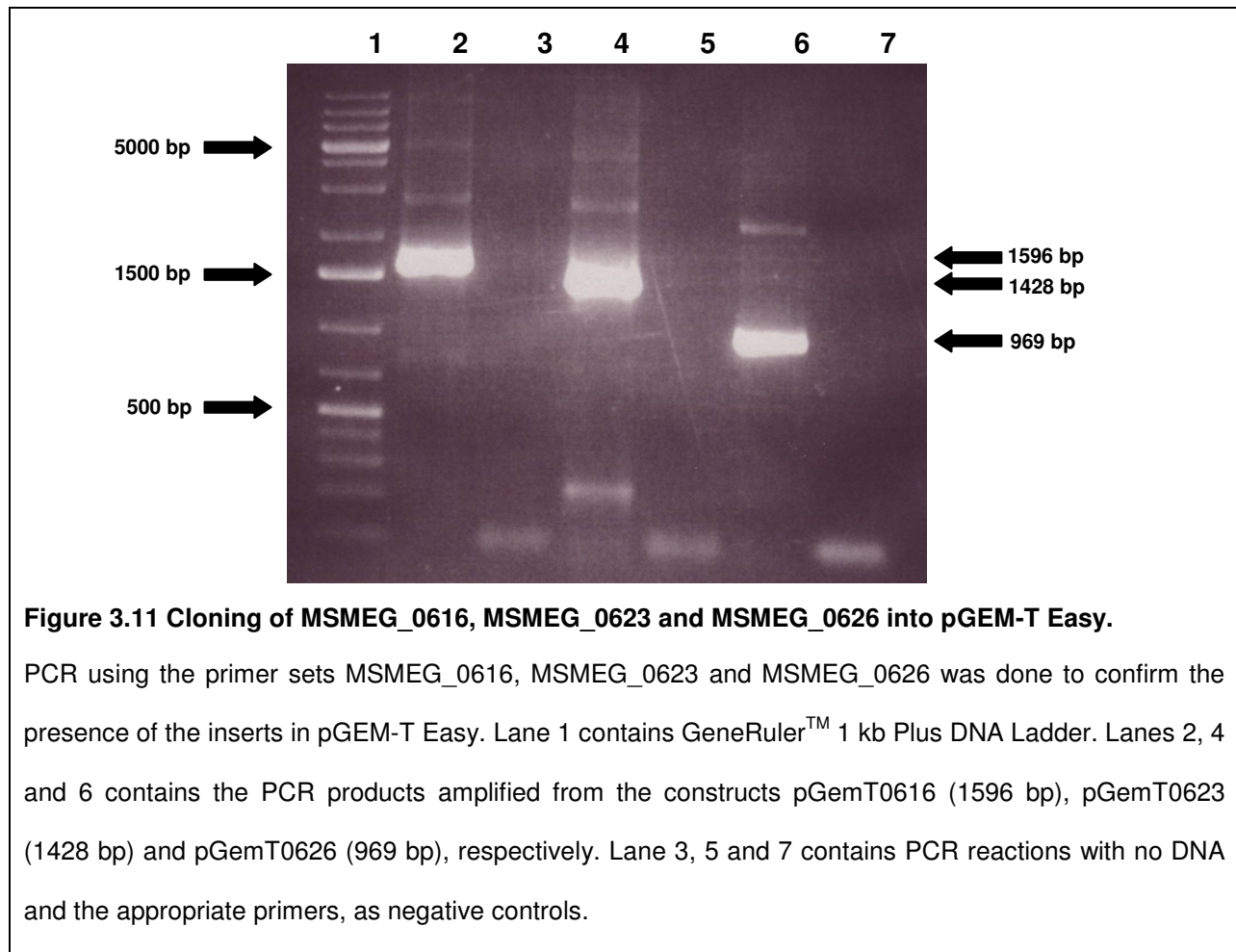
The genes of interest were excised from the pGEM-T Easy constructs using restriction enzyme digestion and subsequently ligated into the episomal vector pDMNI. Electrocompetent *E. coli* were transformed with the pDMNI ligations and transformants were selected on LB agar plates containing kanamycin. Colonies were PCR screened with the primers pDMNIseq f and pDMNIseq r to confirm the presence of the pDMNI final construct (Figure 3.12). The resulting vectors were named pDMNI0616, pDMNI0623 and pDMNI0626, respectively. DNA sequencing confirmed the integrity of the cloned genes.

**Figure 3.12 Cloning of MSMEG_0616, MSMEG_0623 and MSMEG_0626 into pDMNI.**

PCR using the primer set pDMNIseq was done to confirm the presence of the genes of interest in the episomal vector pDMNI. Lane 1 contains GeneRuler™ 1 kb Plus DNA Ladder. Lane 2, 3 and 4 contains the PCR amplicons from the final vectors pDMNI0616 (1596 bp), pDMNI0623 (1428 bp) and pDMNI0626 (969 bp), respectively. Lane 5 contains the PCR amplicon from pDMNI (336 bp) as the positive control and lane 6 contains a PCR containing no DNA as the negative control.

*3.2.2 Generation of pDMNI0615*

*MSMEG_0615* was PCR amplified from *M. smegmatis* mc$^2$155 genomic DNA using the primer set MSMEG_0615. PCR amplified MSMEG_0615 was ligated into CloneJET and transformed into electrocompetent *E. coli*. The transformation was plated out onto LB agar containing ampicillin. Colonies were picked from CloneJET transformation plates and PCR screened with the primers MSMEG_0615 f and MSMEG_0615 r to confirm the presence of the insert in CloneJet (Figure 3.13). This construct was named CJ0615.
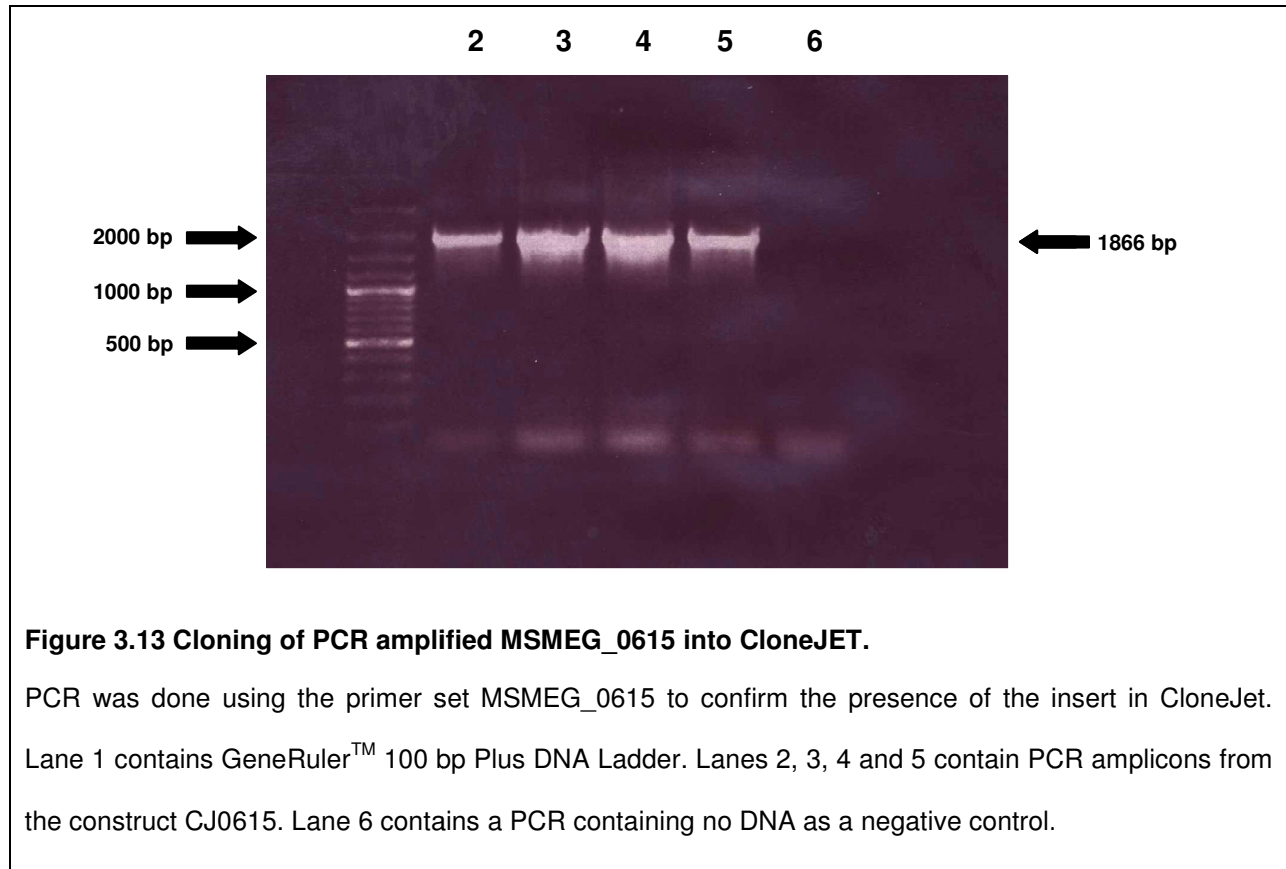
**Figure 3.13 Cloning of PCR amplified MSMEG_0615 into CloneJET.**

PCR was done using the primer set MSMEG_0615 to confirm the presence of the insert in CloneJet. Lane 1 contains GeneRuler[TM] 100 bp Plus DNA Ladder. Lanes 2, 3, 4 and 5 contain PCR amplicons from the construct CJ0615. Lane 6 contains a PCR containing no DNA as a negative control.

MSMEG_0615 was excised out of CloneJET with the restriction enzymes *Mfe*I and *Hind*III and subsequently cloned into pDMNI digested with *Eco*RI and *Hind*III. Electrocompetent *E. coli* was transformed with the pDMNI0615 ligation and cultured on LB agar plates containing kanamycin. Colonies were picked and PCR screened with the primers pDMNIseq f and MSMEG_0615 r to confirm the presence of the insert in pDMNI. This construct was named pDMNI0615 (Figure 3.14). DNA sequencing confirmed the integrity of the cloned gene.

**Figure 3.14 Cloning of PCR amplified MSMEG_0615 into pDMNI.**

PCR was used to confirm the presence of MSMEG_0615 in the episomal vector pDMNI using the primers pDMNIseq f and MSMEG_0615 r. Lane 1 contains GeneRuler$^{TM}$ 100 bp Plus DNA Ladder. Lane 2 contains the PCR amplicon from the final construct pDMNI0615. Lane 3 contains an amplification reaction with no DNA as the negative PCR control.
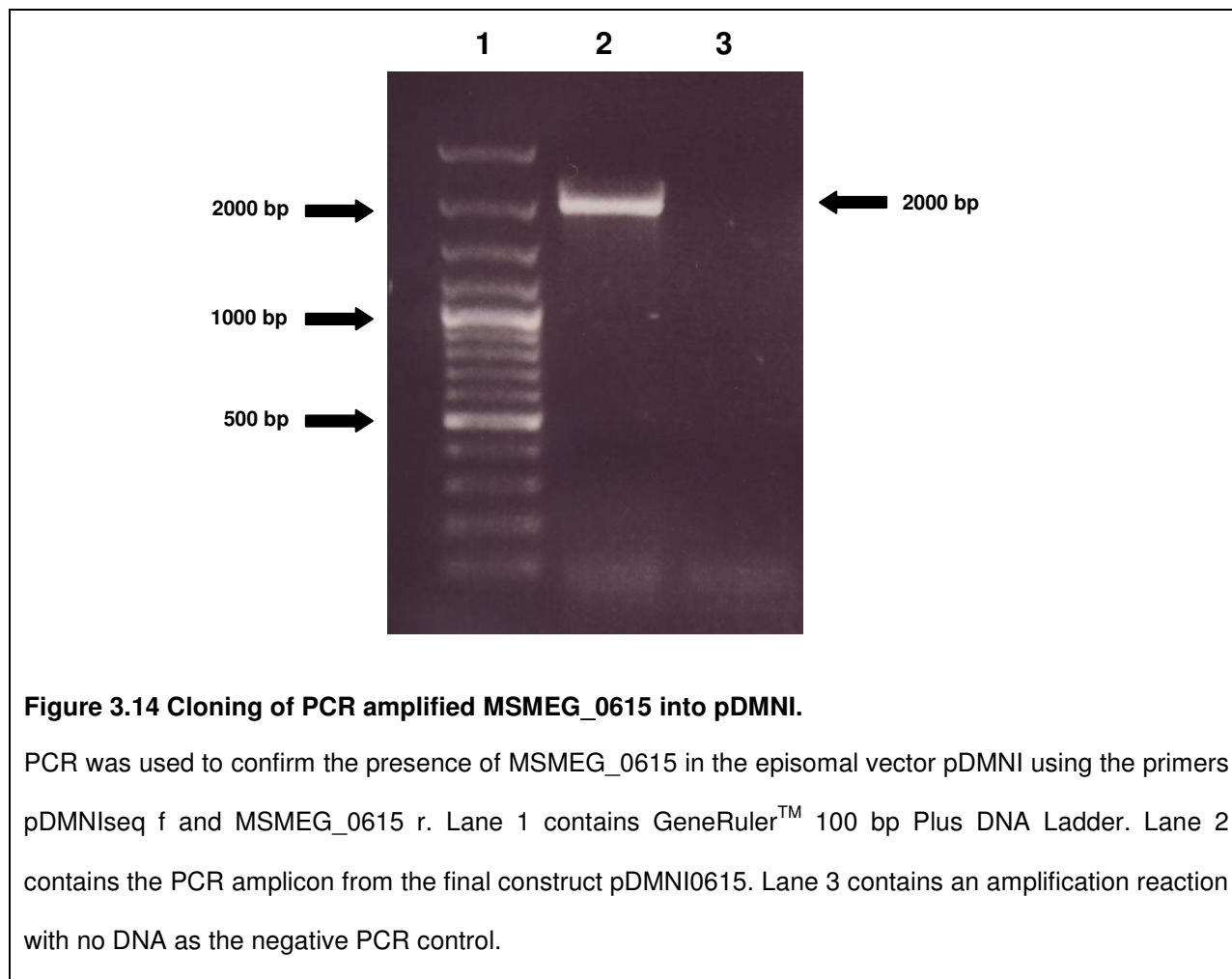
*3.2.3 Generation of pGemT0617*

*MSMEG_0617* was PCR amplified from *M. smegmatis* mc$^2$155 genomic DNA using the primer set MSMEG_0617. PCR products were adenylated and subsequently ligated into the pGEM-T Easy vector. Electrocompetent *E. coli* was transformed with the pGEM-T Easy construct and cultured on LB agar plates containing ampicillin, X-gal and IPTG. White colonies were selected and PCR screened to confirm the presence of the insert (Figure 3.15). Sequencing was performed to verify the sequence integrity of the insert, however the repeated occurance of single nucleotide polymorphisms (SNPs) in the insert was noted. For this reason, cloning of this gene was not continued.

**Figure 3.15 Cloning of PCR amplified MSMEG_0617 into pGEM-T Easy.**

Colony PCR usint the primer set MSMEG_0617 was done to confirm the presence of the insert in pGEM-T Easy in 4 colonies. Lane 1 contains GeneRuler™ 1 kb Plus DNA Ladder. Lane 2 contains a PCR amplicon of an insert of the incorrect size and lane 3 contains no product due to an empty vector, these samples were discarded. Lanes 4 and 5 contain the PCR amplicons of the constructs pGemT0617. Lane 6 contains no DNA as a negative PCR control.

### 3.3 Expression of GFP tagged proteins in *M. smegmatis* and *M. smegmatis* ESX-3 knock-out

Electrocompetent *M. smegmatis* and *M. smegmatis* ESX-3 knock-out strains were transformed with final constructs pDMNI, pDMNI0615, pDMNI0616, pDMNI0623 and pDMNI0626. Colonies were picked from 7H11 plates containing kanamycin and PCR screened to confirm the presence of the final constructs. Whole cell lysates were prepared from *M. smegmatis* and *M. smegmatis* ESX-3 knock-out strains that were PCR positive. Whole cell lysates were separated by SDS-PAGE based on size, in duplicate (one gel was stained with coomassie blue, and the other subjected to Western blotting).

The coomassie blue stained SDS-PAGE gels of whole cell lysates from each clone showed a wide range of band sizes. No visible degradation of proteins was observed (not shown).

**Figure 3.16 Detection of GFP-fusion proteins in *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out by Western blotting.**

Western blotting analysis was performed to confirm the expression of GFP-tagged proteins in *M. smegmatis* wild type (A) and ESX-3 knock-out transformants (B). A: Lane 1 contains PageRuler[TM] Prestained Protein Ladder. Lane 2 contains the whole cell lysate for *M. smegmatis*. Lanes 3, 4, 5 and 7 shows expression of GFP (26.9 kDa), MSMEG_0615-GFP (93.2 kDa), MSMEG_0616-GFP (80.5 kDa) and MSMEG_0626-GFP (60.1 kDa) fusion proteins in *M. smegmatis* ESX-3 knock-out. Lane 6 shows no expression of MSMEG_0623-GFP fusion proteins in this strain.

B: Lane 1 contains PageRuler[TM] Prestained Protein Ladder. Lane 2 contains the whole cell lysate for *M. smegmatis* ESX-3 knock-out. Lanes 3, 4 and 7 shows expression of GFP (26.9 kDa), MSMEG_0615-GFP (93.2 kDa) and MSMEG_0626-GFP (60.1 kDa) fusion proteins in *M. smegmatis* ESX-3 knock-out. Lanes 5 and 6 shows no expression of MSMEG_0616-GFP and MSMEG_0623-GFP fusion proteins in this strain.

Western blotting was done with a mouse anti-GFP primary antibody and a HRP-conjugated goat anti-mouse secondary antibody to detect the expression of the GFP fusion proteins (Figure 3.16). Western blotting confirmed the expression of MSMEG_0615-GFP and MSMEG_0626-GFP fusion proteins in wild type *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out strain. MSMEG_0616-GFP was found to be expressed in *M. smegmatis*, but no expression could be confirmed in the *M. smegmatis* ESX-3 knock-out strain. No GFP fusion protein was detected in *M. smegmatis* or in the *M. smegmatis* ESX-3 knock-out strain that contained the vector pDMNI0623.

## 3.4 The localisation of GFP-tagged ESX-3 proteins in *M. smegmatis* and *M. smegmatis* ESX-3 knock-out

Fluorescent microscopy and transmission electron microscopy were used to determine the localisation of GFP-tagged proteins in *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out.

For the fluorescent microscopy cells were subjected to 1000x magnification and fluorescence was determined with a GFP filter (470 nm). Images were normalized for brightness and contrast. Cells were scored as having 1 fluorescent focus, multiple foci, membrane associated fluorescence, diffuse fluorescence or no fluorescence.

Immunoelectron microscopy was done on thin sections of *M. smegmatis* wild type and *M. smegmatis* cultures transformed with the episomal vectors pDMNI0615, pDMNI0616 and pDMNI0626 embedded in LR white resin. Cells were viewed with a transmission electron microscope at a magnification of 25 000x.

### 3.4.1 M. smegmatis and M. smegmatis ESX-3 knock-out without a GFP expression vector

Fluorescent microscopy determined that wild type *M. smegmatis* and *M. smegmatis* ESX-3 knock-out strains did not exhibit endogenous fluorescence (Figure 3.17 and 3.18). These strains were used as negative controls for fluorescence. In addition, immunoelectron microscopy of *M. smegmatis* determined that there was no specific binding of the primary and secondary antibodies within this organism (Figure 3.19). This strain was used as a negative control for the transmission electron microscopy.

### 3.4.2 pDMNI in M. smegmatis and M. smegmatis ESX-3 knockout

Fluorescent microscopy determined that GFP is distributed throughout the cells in *M. smegmatis* and *M. smegmatis* ESX-3 knockout strains that were transformed with the episomal expression vector pDMNI (Figure 3.17 and 3.18). These strains were used as positive controls for fluorescence.

### 3.4.3 MSMEG_0615 (EccA) in M. smegmatis and M. smegmatis ESX-3 knockout

Fluorescent microscopy determined that the MSMEG_0615-GFP fusion protein showed polar localisation in *M. smegmatis* and the *M. smegmatis* ESX-3 knock-out (Figure 3.17 and 3.18). In *M. smegmatis* the majority of the cells, 95% (64/67) exhibited a unipolar localisation for the MSMEG_0615-GFP fusion protein. It was not determined what percentage of the *M. smegmatis* ESX-3 knock-out cells exhibited unipolar localisation of MSMEG_0615-GFP. Immunoelectron microscopy of *M. smegmatis* transformed with the episomal expression vector pDMNI0615 showed that the cytoplasmic protein $EccA_3$ exhibits membrane association (Figure 3.19).

### 3.4.4 MSMEG_0616 (EccB) is membrane associated in M. smegmatis and M. smegmatis ESX-3 knock-out

Fluorescent microscopy suggested that MSMEG_0616-GFP fusion protein was membrane associated in *M. smegmatis* (Figure 3.17). Due to the lack of expression of the MSMEG_0616-GFP fusion protein in the *M. smegmatis* ESX-3 knock-out, no fluorescence was observed (Figure 3.18). Additional immunoelectron microscopy of *M. smegmatis* expressing the fusion protein MSMEG_0616-GFP indicated that the $EccB_3$ protein localises to the mycobacterial cell wall (Figure 3.19).

**3.4.1.5 MSMEG_0623 in *M. smegmatis* and *M. smegmatis* ESX-3 knock-out**

Due to the lack of expression of the MSMEG_0623-GFP fusion protein, no fluorescence was exhibited in *M. smegmatis* or in the *M smegmatis* ESX-3 knock-out strain (Figure 3.17 and 3.18).

**3.4.1.6 MSMEG_0626 (EccE) is membrane associated in *M. smegmatis* and *M. smegmatis* ESX-3 knock-out**

Fluorescent microscopy suggested that the MSMEG_0626-GFP fusion protein was membrane associated in *M. smegmatis* and the *M. smegmatis* ESX-3 knock-out strain  (Figure 3.17 and 3.18). Immunoelectron microscopy of *M. smegmatis* expressing the fusion protein MSMEG_0626-GFP indicated that the $EccE_3$ protein localises to the mycobacterial cell wall (Figure 3.19).
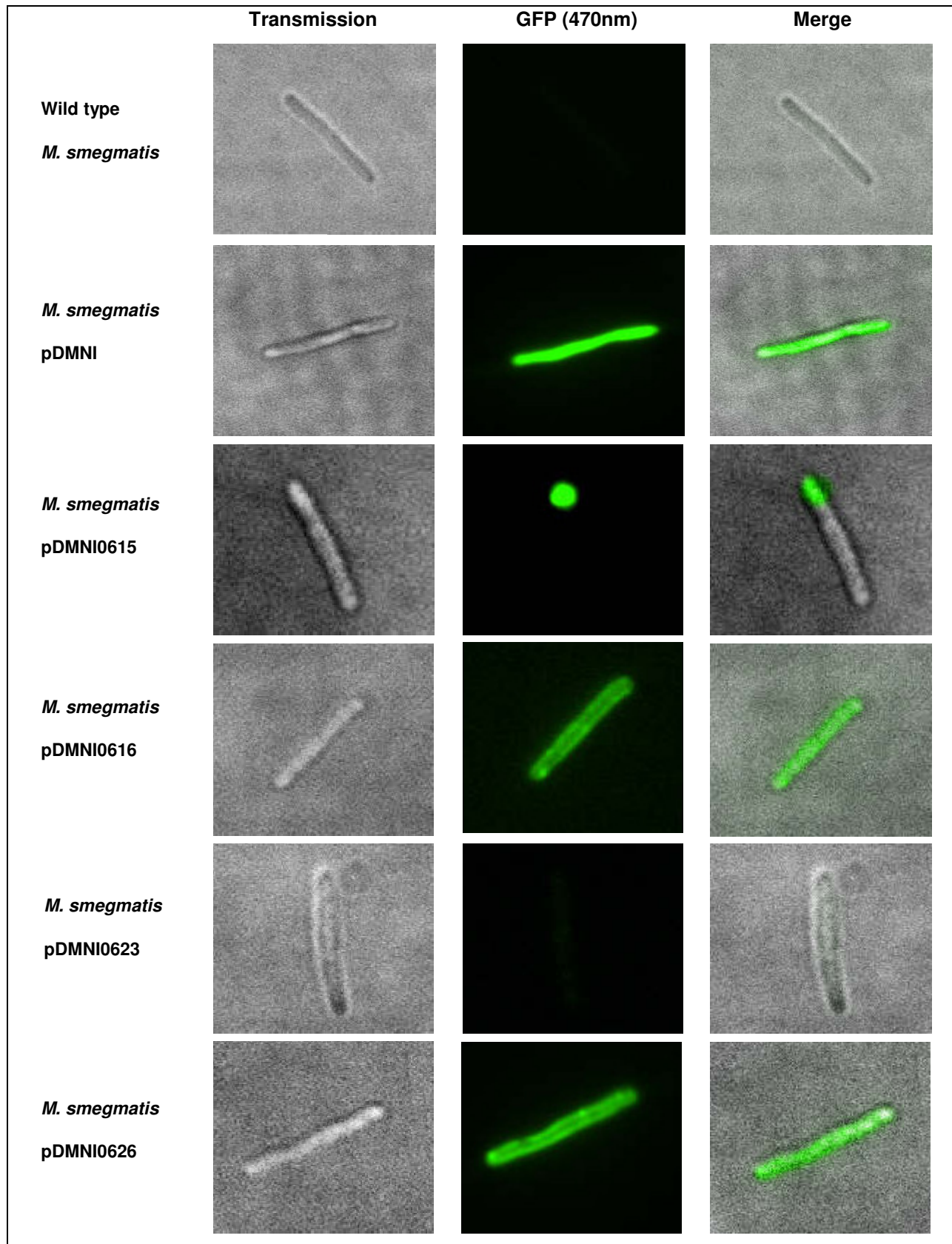
**Figure 3.17 Fluorescent microscopy of wild type *M. smegmatis* transformants.**

Representative transmission (left panel), GFP (middle) and merged (right) images of *M. smegmatis* transformants (not all images are shown). Wild type *M. smegmatis* show no fluorescence. *M. smegmatis* that contains the episomal expression vector pDMNI exhibits diffuse fluorescence throughout the cell. *M. smegmatis* expressing the MSMEG_0615-GFP fusion protein shows unipolar localisation in the organism. *M. smegmatis* cells expressing the fusion proteins MSMEG_0616-GFP and MSMEG_0626-GFP showed that these proteins localized to the mycobacterial membrane. No fluorescence was observed for *M. smegmatis* cells that contain the vector pDMNI0623.
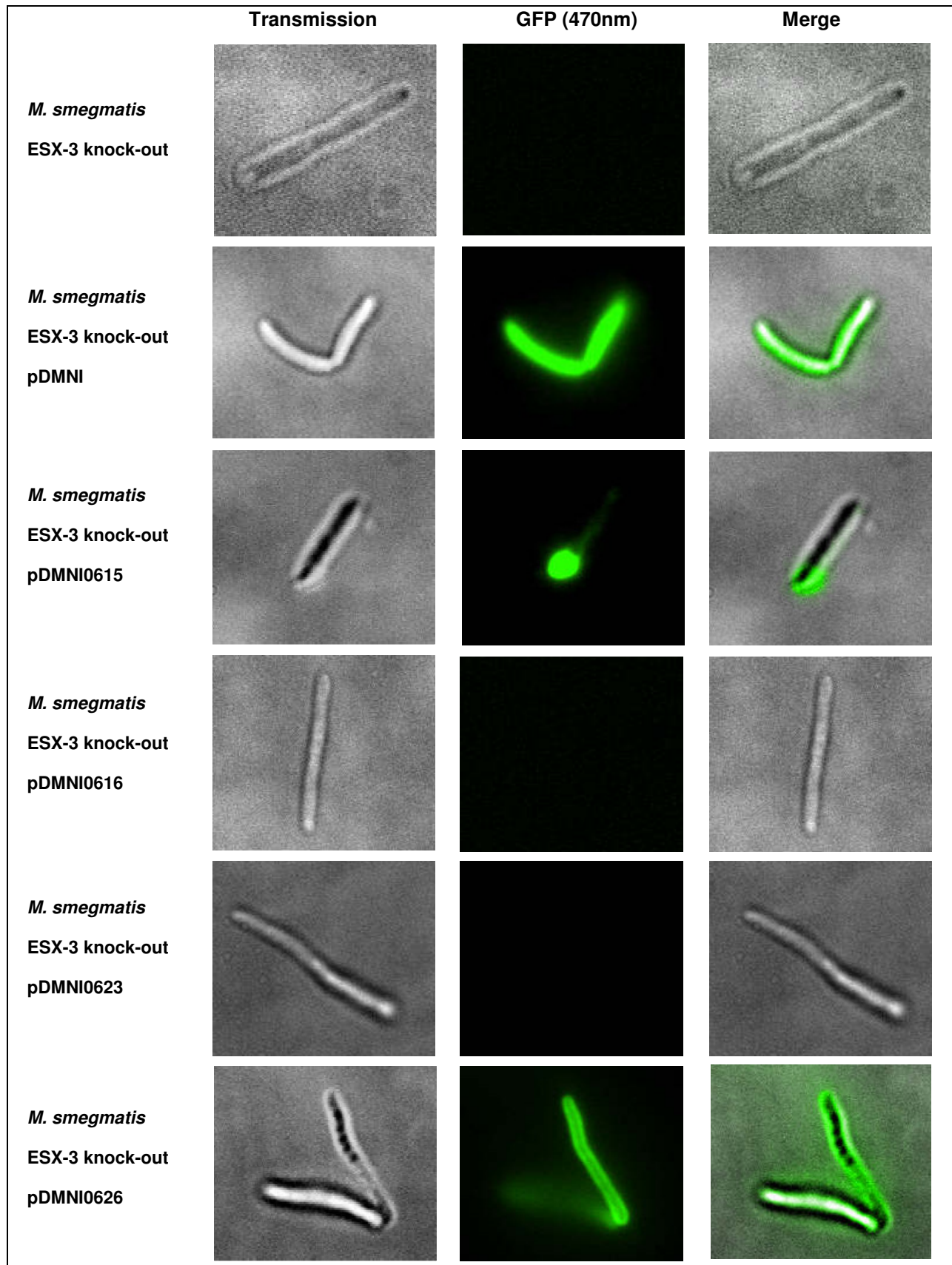
| | Transmission | GFP (470nm) | Merge |
|---|---|---|---|
| *M. smegmatis* ESX-3 knock-out | | | |
| *M. smegmatis* ESX-3 knock-out pDMNI | | | |
| *M. smegmatis* ESX-3 knock-out pDMNI0615 | | | |
| *M. smegmatis* ESX-3 knock-out pDMNI0616 | | | |
| *M. smegmatis* ESX-3 knock-out pDMNI0623 | | | |
| *M. smegmatis* ESX-3 knock-out pDMNI0626 | | | |

**Figure 3.18 Fluorescent microscopy of *M. smegmatis* ESX-3 knock-out transformants.**

Representative transmission (left panel), GFP (middle) and merged (right) images of *M. smegmatis* ESX-3 knock-out transformants. The *M. smegmatis* ESX-3 knock-out strain shows no fluorescence. The *M. smegmatis* ESX-3 knock-out that contains the episomal expression vector pDMNI shows diffuse fluorescence throughout the cell. *M. smegmatis* ESX-3 knock-out expressing the MSMEG_0615-GFP fusion protein exhibits unipolar localisation in the organism. *M. smegmatis* ESX-3 knock-out cells expressing the fusion protein MSMEG_0626-GFP showed that this protein localized to the mycobacterial membrane. No fluorescence was observed for *M. smegmatis* ESX-3 knock-out cells that contain the vectors pDMNI0616 and pDMNI0623.
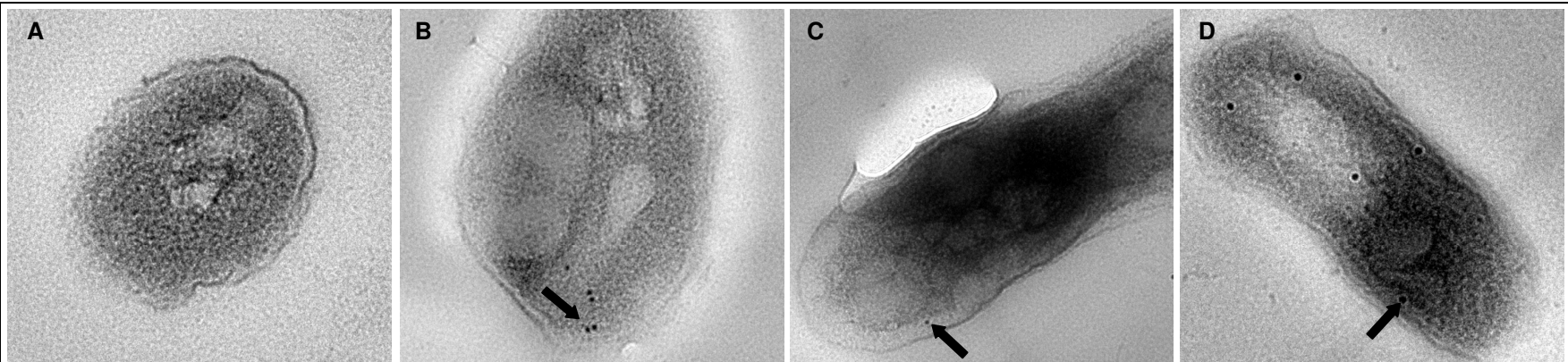
**Figure 3.19 Transmission electron microscopy of *M. smegmatis* transformants.**

Immunolabelling was done on thin sections of wild type *M.* smegmatis (A) and *M. smegmatis* pDMNI0615 (B), pDMNI0616 (C) and pDMNI0626 (D) transformants embedded in LR white resin. Binding of primary and secondary antibodies to GFP tagged proteins indicated with arrows. A: Wild type *M. smegmatis* shows no specific binding of antibodies and is used as a negative control. B: *M. smegmatis* pDMNI0615 transformant exhibits specific binding of antibodies, showing localisation of the cytoplasmic protein $EccA_3$ to the mycobacterial membrane. C: *M. smegmatis* pDMNI0616 exhibits specific binding of antibodies, showing localisation of $EccB_3$ at the mycobacterial membrane. D: *M. smegmatis* pDMNI0626 transformant exhibits specific binding of the antibodies, showing membrane localisation of $EccE_3$.

# CHAPTER 4

## Discussion and Conclusion

In this study we aimed to identify the sub-cellular location of the conserved components of the ESX-3 secretion system in *M. smegmatis* by determining where the proteins from this secretion system localise within the mycobacterial cell. We also investigated whether the location of these proteins was retained in the *M. smegmatis* ESX-3 knock-out strain.

We selected the native esx conserved component genes from ESX-3 to determine the location of this secretion system in *M. smegmatis*. The genes of interest were cloned into the episomal expression vector pDMNI and expressed as GFP fusion proteins in both the wild type *M. smegmatis* strain and the *M. smegmatis* ESX-3 knock-out strain. Western blot analysis confirmed the expression of these fusion proteins, while fluorescent microscopy and transmission electron microscopy were used to determine the location of these proteins within the mycobacterial cell. This study has provided novel insight as to how and where the ESX-3 secretion system is assembled in the mycobacterial membrane.

### 4.1 EccA encoded by *MSMEG_0615*

*MSMEG_0615* (EccA$_3$) was successfully expressed as a fusion protein in *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out strain. Fluorescent microscopy determined that this protein localized to the bacterial poles in both strains. The EccA$_3$-GFP fusion protein exhibited unipolar localisation in both strains, indicating that the localisation is specific to a single polar region in the cell. Similarly, a previous study showed that ESX-1 associated proteins specifically localised to the old mycobacterial pole following cell division in *M. smegmatis* (Wirth *et al.* 2012). ESX-1 proteins were also found to be localised to a specific mycobacterial pole in *M. marinum* (Carlsson *et al.* 2009). It is believed that EccA is a hexameric AAA+ ATPase cytoplasmic protein from the CbxX/CfqX subfamily (Luthra *et al.* 2008) and previous studies have shown that some AAA+ ATPases are involved in the assembly of membrane complexes (Snider, Houry 2008). *In silico* analysis did not identify any transmembrane regions in the *MSMEG_0615* gene, supporting the hypothesis that EccA$_3$ is a cytoplasmic protein. Immunoelectron microscopy of *M. smegmatis* cells expressing the EccA$_3$-GFP fusion protein showed that this cytoplasmic protein exhibits association with the mycobacterial membrane. We demonstrate that this protein contains the translocation signals by the fact that EccA$_3$ was translocated to the polar region in the absence of the other ESX-3 conserved components, as was shown in the *M. smegmatis* ESX-3 knock-out strain.

EccA possesses an ATP/GTP-binding site and could provide energy for the assembly of the secretion system at the polar region of mycobacterial membrane. $EccA_1$ has been found to be essential for the cellular translocation and secretion of ESX-1 substrates in *M. tuberculosis* (Luthra *et al.* 2008). However, the impact of deleting the *eccA* gene on the functionality of the remaining operon was not determined.

This is the first study to demonstrate the cellular location of EccA. Previous studies have primarily focused on $EccCa_1$ and $EccCb_1$ from the ESX-1 gene cluster and have shown that these proteins localise at the mycobacterial cell pole (Wirth *et al.* 2012). It was also established that $EccCb_1$ did not localise to the mycobacterial pole in the absence of $EccCa_1$, strengthening the hypothesis that these genes are co-transcribed and co-translated from an ESX operon. We can thus hypothesize that EccA acts as a recruiter protein of the ESX secretion system to the mycobacterial pole where it is then inserted. Homologues of *eccA* are also present in ESX-1, ESX-2 and ESX-5 gene clusters but their cellular location remains to be determined. The ancestral region, ESX-4, does not possess an *eccA* homologue.

### 4.2 EccB encoded by *MSMEG_0616*

The fusion protein $EccB_3$-GFP was expressed in wild type *M. smegmatis*, however no expression occurred in the *M. smegmatis* ESX-3 knock-out strain. Fluorescent microscopy determined that $EccB_3$ localised to the mycobacterial membrane in *M. smegmatis*. Immunoelectron microscopy of *M. smegmatis* expressing the $EccB_3$-GFP fusion protein confirmed the membrane association. *In silico* analysis identified a tranmembrane region in the amino acid sequence of $EccB_3$, supporting the hypothesis that EccB is a transmembrane protein. The absence of any polar localisation suggests that EccB may either be assembled into the ESX-3 secretion complex followed by diffusion into the cell membrane or that over-expression of this protein might result in non-specific localisation.  Non-specific localisation may result from insufficient recruiter protein encoded by the wild type ESX-3 or the lack of co-translation of the ESX-3 secretion system components.

The reason for the absence of expression of EccB in the ESX-3 knock-out mutant is unknown but may be explained by the silencing of the psmyc promoter (upstream of the *MSMEG_0616* gene) in the plasmid pDMNI0616. For example: a previous study has found that the hsp60 promoter causes plasmid instability

in *M. bovis* BCG. It was determined that various deletions occurred in the hsp60 promoter region during or directly after transformation (Al-Zarouni, Dale 2002).

## 4.3 EccD encoded by *MSMEG_0623*

Western blot analysis and fluorescent microscopy failed to detect expression of MSMEG_0623 as a fusion protein in *M. smegmatis* or in the *M. smegmatis* ESX-3 knock-out strain. *In silico* analysis identified 10-11 transmembrane regions in the $EccD_3$ protein, supporting the hypothesis that EccD is a multi-transmembrane spanning protein that forms a channel or a translocation pore in the mycobacterial membrane (Feltcher, Sullivan & Braunstein 2010).The toxic effect of overexpression of this fusion protein may result in the inactivation of the psmyc promoter resulting in no protein being expressed.

## 4.4 EccE encoded by *MSMEG_0626*

MSMEG_0626 was expressed as a GFP fusion protein in wild type *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out strain. Fluorescent microscopy indicated that this fusion protein was membrane associated, as confirmed by immunoelectron microscopy of *M. smegmatis*. *In silico* analysis identified two transmembrane regions in the $EccE_3$ protein of which one was shown to be a signal sequence. This supports the hypothesis that EccE  homologues in *M. tuberculosis* are transmembrane proteins. The membrane location of this protein, in the absence of the conserved components of ESX-3 (in the *M. smegmatis* ESX-3 knock-out strain), supports the notion of non-specific localisation.

## 4.5 Limitations of this study

Determining the location of a protein within a cell by generating a fusion protein with GFP has a number of limitations. The addition of a 26.9 kDa tag onto the selected proteins of interest may interfere with the protein's function, its interactions with other proteins within the cell or its localisation. To reduce the steric interference that can be caused by the addition of such a large tag, a linker chain (SAGSAG) was used to link the protein of interest to GFP. These fusion proteins might also produce little or no fluorescence, they might have inconsistent localisation or the overexpression of the fusion proteins might be toxic to the cell (Wirth *et al.* 2012).

To determine at which pole the protein localizes following cell division, cells have to be imaged in a flow-cell microfluidic chamber that constrains the mycobacteria to a single focal plane and supplies cells with fresh media throughout the experiment (Wirth *et al.* 2012, Aldridge *et al.* 2012). Due to the absence of appropriate equipment, fluorescent time-lapse microscopy could not be used to determine to which pole the EccA-GFP fusion protein localizes in *M. smegmatis*.

To confirm that $EccB_3$ and $EccE_3$ are indeed non-specifically localised, co-localisation studies with $EccA_3$ tagged with a different fluorescent protein (e.g. YFP) could be done in future studies.

### 4.6 Conclusion

Macromolecular complexes which are involved in cell division and DNA replication have been shown to localise to specific regions within the cell. There is increasing evidence which shows that protein secretion systems also localise to specific regions within the cell (Burton, Dubnau 2010, Wirth *et al.* 2012). Type III and Type IV secretion systems are active at bacterial poles in Shigella and Agrobacterium, respectively (Jaumouille *et al.* 2008, Judd, Kumar & Das 2005). These findings indicate that pathogenic bacteria localise their virulence related secretion systems at the polar regions of the cell, it however remains unknown whether a specific location for virulence related secretion system is required for the pathogenicity of the organism.  In the mycobacterial species only ESX-1 has been shown to be localised at the bacterial pole, specifically the old bacterial pole where cell wall growth occurs (Carlsson *et al.* 2009, Wirth *et al.* 2012). In this study, we have shown that EccA from ESX-3 in *M. smegmatis* is also polarly localised. ESX-3 has been implicated in iron acquisition and it is unknown why this secretion system would localise to only one region within the mycobacterial cell. EccA might assemble the ESX-3 secretion system at the mycobacterial pole before diffusing into the rest of the membrane, explaining the non-specific localisation that was observed with EccB and EccE. It is possible that the remaining secretion systems (ESX-5 and ESX-2) might also localise to a single mycobacterial pole, most probably the old pole. The ancestral region, ESX-4, which does not possess an *eccA* gene, could possibly not have a specific location in the mycobacterial membrane if EccA is a recruiter protein.

# <u>CHAPTER 5</u>

## Future Directions

This study has demonstrated that EccA from ESX-3 is localised at a specific mycobacterial pole in *M. smegmatis*. Determining the location of the selected proteins of interest with this non-invasive method does have its limitations. Alternative methods with fewer limitations might be to create knock-ins with GFP. The linker chain and the GFP gene could be cloned into the genome of *M. smegmatis* on the 3' end of the gene of interest using homologous recombination (Gordhan, Parish 2001). By generating a knock-in strain the cell will be expressing the gene of interest at native levels from its original position in the genome, reducing the chance for toxicity issues.

In future studies it may be possible to show that the secretion of ESX-3 substrates is also localised to a mycobacterial pole to provide further evidence that the ESX-3 secretion system is present and function at a single pole in *M. smegmatis*. Known substrates EsxG and EsxH from ESX-3 in *M. smegmatis* could be tagged with a fluorescent protein and fluorescent time-lapse microscopy could be used to determine if fluorescence occurs at the mycobacterial pole. By using a different fluorescent protein, it could be shown that secretion occurs at the same pole where the ESX-3 components localise.

**5.1 EccA**

To identify the pole at which EccA localizes following cell division, time-lapse microscopy should to be used in combination with a microfluidic chamber that constrains the mycobacterial cells to a single focal plane and supplies cells with fresh media throughout the experiment (Wirth *et al.* 2012, Aldridge *et al.* 2012).

The hypothesis that EccA might act as a recruiter protein, could be tested *in vivo* pull down assays to identify proteins which interact with EccA. Proteins identified by this method could be tagged with a fluorescent protein (other than GFP) to determine if these protein do require EccA to localise to the mycobacterial membrane by making use of a MSMEG_0615 knock-out. The tagged proteins can also be used to determine if they co-localise with MSMEG_0615-GFP in *M. smegmatis*.

**5.2 EccB and EccE**

The non-specific membrane localisation exhibited by EccB-GFP and EccE-GFP could be attributed to insufficient recruiter protein being available to specifically localise the overexpressed fusion proteins. This could be rectified by using the tetracycline repressor system (Ehrt *et al.* 2005). This expression system allows for the induction or repression of gene expression in slow- and fast-growing mycobacterial species and can thus be used to create conditional mutants. Tetracycline can cross the biological membrane by diffusion, inducing expression of genes that have been cloned downstream of the conditional promoter. By not overexpressing the fusion proteins in the cell at all times, there might be enough recruiter protein available to specifically localise the fusion proteins, EccB-GFP and EccE-GFP at the mycobacterial membrane.

**5.3 EccC**

The failed cloning of MSMEG_0617 into pGEM-T Easy can be approached differently. Multiple primer sets can be used to PCR amplify sections of this gene which can be ligated together to form a single insertion fragment. This single insertion fragment can be ligated into the episomal expression vector pDMNI to generate the final construct pDMNI0617, which will subsequently express the fusion protein EccC-GFP in *M. smegmatis*. Fluorescent microscopy will be performed to identify the sub-cellular localisation the EccC-GFP in *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out.

In the case of no expression or toxicity issues which may result from with the overexpression of such a large protein in *M. smegmatis*, smaller sections of the MSMEG_0617 gene can also be PCR amplified and ligated into pDMNI. The expression of smaller sections of this protein as fusions with GFP may result in less toxicity and may prevent the shutdown of transcription or translation of this gene fragment fusion. A conditional expression mutant can also be generated making use of the tetracycline repressor system. By not overexpressing the fusion proteins in the cell at all times, transcription and translation of the EccC-GFP fusion protein may remain unaffected.

**5.4 EccD**

There was no expression of EccD-GFP in *M. smegmatis* or in the *M. smegmatis* ESX-3 knock-out strain. By PCR amplifying the gene of interest fused with the GFP gene from within the pDMNI0623 vector, this insert can be cloned into the tetracycline repressor system. The conditional expressing mutant generated can thus be expressed in the presence of tetracycline, limiting the amount EccD-GFP constitutively produced in the cell. Fluorescent microscopy and transmission electron microscopy can subsequently be used to identify the sub-cellular localisation of EccD in *M. smegmatis* and in the *M. smegmatis* ESX-3 knock-out strain.

Determining the specific localisation of the other ESX-3 proteins might strengthen the hypothesis that this secretion system is localised at the mycobacterial pole in *M. smegmatis*. Identifying proteins that interacts with EccA, which has been shown to be unipolarly localised, is of great importance for the understanding of how this secretion system localises itself to one compartment within the cell. The proteins identified in these pull down experiments might be important for future drug and vaccine design that might help to eradicate the tuberculosis epidemic.

# CHAPTER 6

**List of References**

Abdallah, A.M., Bestebroer, J., Savage, N.D., de Punder, K., van Zon, M., Wilson, L., Korbee, C.J., van der Sar, A.M., Ottenhoff, T.H., van der Wel, N.N., Bitter, W. & Peters, P.J. 2011, "Mycobacterial secretion systems ESX-1 and ESX-5 play distinct roles in host cell death and inflammasome activation", *Journal of immunology (Baltimore, Md.: 1950),* vol. 187, no. 9, pp. 4744-4753.

Abdallah, A.M., Gey van Pittius, N.C., Champion, P.A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C.M., Appelmelk, B.J. & Bitter, W. 2007, "Type VII secretion--mycobacteria show the way", *Nature reviews.Microbiology,* vol. 5, no. 11, pp. 883-891.

Abdallah, A.M., Verboom, T., Hannes, F., Safi, M., Strong, M., Eisenberg, D., Musters, R.J., Vandenbroucke-Grauls, C.M., Appelmelk, B.J., Luirink, J. & Bitter, W. 2006, "A specific secretion system mediates PPE41 transport in pathogenic mycobacteria", *Molecular microbiology,* vol. 62, no. 3, pp. 667-679.

Abdallah, A.M., Verboom, T., Weerdenburg, E.M., Gey van Pittius, N.C., Mahasha, P.W., Jimenez, C., Parra, M., Cadieux, N., Brennan, M.J., Appelmelk, B.J. & Bitter, W. 2009, "PPE and PE_PGRS proteins of Mycobacterium marinum are transported via the type VII secretion system ESX-5", *Molecular microbiology,* vol. 73, no. 3, pp. 329-340.

Aldridge, B.B., Fernandez-Suarez, M., Heller, D., Ambravaneswaran, V., Irimia, D., Toner, M. & Fortune, S.M. 2012, "Asymmetry and aging of mycobacterial cells lead to variable growth and antibiotic susceptibility", *Science (New York, N.Y.),* vol. 335, no. 6064, pp. 100-104.

Al-Zarouni, M. & Dale, J.W. 2002, "Expression of foreign genes in Mycobacterium bovis BCG strains using different promoters reveals instability of the hsp60 promoter for expression of foreign genes in Mycobacterium bovis BCG strains", *Tuberculosis (Edinburgh, Scotland),* vol. 82, no. 6, pp. 283-291.

Berthet, F.X., Rasmussen, P.B., Rosenkrands, I., Andersen, P. & Gicquel, B. 1998, "A Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10)", *Microbiology (Reading, England),* vol. 144 ( Pt 11), no. Pt 11, pp. 3195-3203.

Bitter, W., Houben, E.N., Bottai, D., Brodin, P., Brown, E.J., Cox, J.S., Derbyshire, K., Fortune, S.M., Gao, L.Y., Liu, J., Gey van Pittius, N.C., Pym, A.S., Rubin, E.J., Sherman, D.R., Cole, S.T. & Brosch, R. 2009, "Systematic genetic nomenclature for type VII secretion systems", *PLoS pathogens,* vol. 5, no. 10, pp. e1000507.

Bottai, D., Di Luca, M., Majlessi, L., Frigui, W., Simeone, R., Sayes, F., Bitter, W., Brennan, M.J., Leclerc, C., Batoni, G., Campa, M., Brosch, R. & Esin, S. 2012, "Disruption of the ESX-5 system of Mycobacterium tuberculosis causes loss of PPE protein secretion, reduction of cell wall integrity and strong attenuation", *Molecular microbiology,* vol. 83, no. 6, pp. 1195-1209.

Brennan, M.J., Delogu, G., Chen, Y., Bardarov, S., Kriakov, J., Alavi, M. & Jacobs, W.R.,Jr 2001, "Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells", *Infection and immunity,* vol. 69, no. 12, pp. 7326-7333.

Brodin, P., Eiglmeier, K., Marmiesse, M., Billault, A., Garnier, T., Niemann, S., Cole, S.T. & Brosch, R. 2002, "Bacterial artificial chromosome-based comparative genomic analysis identifies Mycobacterium microti as a natural ESAT-6 deletion mutant", *Infection and immunity,* vol. 70, no. 10, pp. 5568-5578.

Brodin, P., Majlessi, L., Marsollier, L., de Jonge, M.I., Bottai, D., Demangel, C., Hinds, J., Neyrolles, O., Butcher, P.D., Leclerc, C., Cole, S.T. & Brosch, R. 2006, "Dissection of ESAT-6 system 1 of Mycobacterium tuberculosis and impact on immunogenicity and virulence", *Infection and immunity,* vol. 74, no. 1, pp. 88-98.

Brodin, P., Rosenkrands, I., Andersen, P., Cole, S.T. & Brosch, R. 2004, "ESAT-6 proteins: protective antigens and virulence factors?", *Trends in microbiology,* vol. 12, no. 11, pp. 500-508.

Brown, G.D., Dave, J.A., Gey van Pittius, N.C., Stevens, L., Ehlers, M.R. & Beyers, A.D. 2000, "The mycosins of Mycobacterium tuberculosis H37Rv: a family of subtilisin-like serine proteases", *Gene,* vol. 254, no. 1-2, pp. 147-155.

Burton, B. & Dubnau, D. 2010, "Membrane-associated DNA transport machines", *Cold Spring Harbor perspectives in biology,* vol. 2, no. 7, pp. a000406.

Carlsson, F., Joshi, S.A., Rangell, L. & Brown, E.J. 2009, "Polar localization of virulence-related Esx-1 secretion in mycobacteria", *PLoS pathogens,* vol. 5, no. 1, pp. e1000285.

Champion, P.A., Stanley, S.A., Champion, M.M., Brown, E.J. & Cox, J.S. 2006, "C-terminal signal sequence promotes virulence factor secretion in Mycobacterium tuberculosis", *Science (New York, N.Y.),* vol. 313, no. 5793, pp. 1632-1636.

Crozat, E. & Grainge, I. 2010, "FtsK DNA translocase: the fast motor that knows where it's going", *Chembiochem : a European journal of chemical biology,* vol. 11, no. 16, pp. 2232-2243.

Dave, J.A., Gey van Pittius, N.C., Beyers, A.D., Ehlers, M.R. & Brown, G.D. 2002, "Mycosin-1, a subtilisin-like serine protease of Mycobacterium tuberculosis, is cell wall-associated and expressed during infection of macrophages", *BMC microbiology,* vol. 2, pp. 30.

DiGiuseppe Champion, P.A., Champion, M.M., Manzanillo, P. & Cox, J.S. 2009, "ESX-1 secreted virulence factors are recognized by multiple cytosolic AAA ATPases in pathogenic mycobacteria", *Molecular microbiology,* vol. 73, no. 5, pp. 950-962.

Dussurget, O., Rodriguez, M. & Smith, I. 1996, "An ideR mutant of Mycobacterium smegmatis has derepressed siderophore production and an altered oxidative-stress response", *Molecular microbiology,* vol. 22, no. 3, pp. 535-544.

Dussurget, O., Timm, J., Gomez, M., Gold, B., Yu, S., Sabol, S.Z., Holmes, R.K., Jacobs, W.R.,Jr & Smith, I. 1999, "Transcriptional control of the iron-responsive fxbA gene by the mycobacterial regulator IdeR", *Journal of Bacteriology,* vol. 181, no. 11, pp. 3402-3408.

Ehrt, S., Guo, X.V., Hickey, C.M., Ryou, M., Monteleone, M., Riley, L.W. & Schnappinger, D. 2005, "Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor", *Nucleic acids research,* vol. 33, no. 2, pp. e21.

Feltcher, M.E., Sullivan, J.T. & Braunstein, M. 2010, "Protein export systems of Mycobacterium tuberculosis: novel targets for drug development?", *Future microbiology,* vol. 5, no. 10, pp. 1581-1597.

Flardh, K. 2010, "Cell polarity and the control of apical growth in Streptomyces", *Current opinion in microbiology,* vol. 13, no. 6, pp. 758-765.

Frigui, W., Bottai, D., Majlessi, L., Monot, M., Josselin, E., Brodin, P., Garnier, T., Gicquel, B., Martin, C., Leclerc, C., Cole, S.T. & Brosch, R. 2008, "Control of M. tuberculosis ESAT-6 secretion and specific T cell recognition by PhoP", *PLoS pathogens,* vol. 4, no. 2, pp. e33.

Gao, L.Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J.N. & Brown, E.J. 2004, "A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion", *Molecular microbiology,* vol. 53, no. 6, pp. 1677-1693.

Gey van Pittius, N.C., Gamieldien, J., Hide, W., Brown, G.D., Siezen, R.J. & Beyers, A.D. 2001, "The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C Gram-positive bacteria", *Genome biology,* vol. 2, no. 10, pp. RESEARCH0044.

Gey van Pittius, N.C., Sampson, S.L., Lee, H., Kim, Y., van Helden, P.D. & Warren, R.M. 2006, "Evolution and expansion of the Mycobacterium tuberculosis PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions", *BMC evolutionary biology,* vol. 6, pp. 95.

Gold, B., Rodriguez, G.M., Marras, S.A., Pentecost, M. & Smith, I. 2001, "The Mycobacterium tuberculosis IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages", *Molecular microbiology,* vol. 42, no. 3, pp. 851-865.

Gordhan, B.G. & Parish, T. 2001, "Gene Replacement using Pretreated DNA" in *Mycobacterium Tuberculosis Protocols (Methods in Molecular Medicine)*, ed. Tanya Parish, Neil G. Stoker, Humana Press Inc., Totowa, New Jersey, pp. 77-92.

Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., Smith, S. & Sherman, D.R. 2004, "Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis", *Molecular microbiology,* vol. 51, no. 2, pp. 359-370.

Huang, K.C. & Ramamurthi, K.S. 2010, "Macromolecules that prefer their membranes curvy", *Molecular microbiology,* vol. 76, no. 4, pp. 822-832.

Ilghari, D., Lightbody, K.L., Veverka, V., Waters, L.C., Muskett, F.W., Renshaw, P.S. & Carr, M.D. 2011, "Solution structure of the Mycobacterium tuberculosis EsxG.EsxH complex: functional implications and comparisons with other M. tuberculosis Esx family complexes", *The Journal of biological chemistry,* vol. 286, no. 34, pp. 29993-30002.

Jaumouille, V., Francetic, O., Sansonetti, P.J. & Tran Van Nhieu, G. 2008, "Cytoplasmic targeting of IpaC to the bacterial pole directs polar type III secretion in Shigella", *The EMBO journal,* vol. 27, no. 2, pp. 447-457.

Judd, P.K., Kumar, R.B. & Das, A. 2005, "Spatial location and requirements for the assembly of the Agrobacterium tumefaciens type IV secretion apparatus", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 102, no. 32, pp. 11498-11503.

Lenarcic, R., Halbedel, S., Visser, L., Shaw, M., Wu, L.J., Errington, J., Marenduzzo, D. & Hamoen, L.W. 2009, "Localisation of DivIVA by targeting to negatively curved membranes", *The EMBO journal,* vol. 28, no. 15, pp. 2272-2282.

Li, Y., Miltner, E., Wu, M., Petrofsky, M. & Bermudez, L.E. 2005, "A Mycobacterium avium PPE gene is associated with the ability of the bacterium to grow in macrophages and virulence in mice", *Cellular microbiology,* vol. 7, no. 4, pp. 539-548.

Luthra, A., Mahmood, A., Arora, A. & Ramachandran, R. 2008, "Characterization of Rv3868, an essential hypothetical protein of the ESX-1 secretion system in Mycobacterium tuberculosis", *The Journal of biological chemistry,* vol. 283, no. 52, pp. 36532-36541.

MacGurn, J.A., Raghavan, S., Stanley, S.A. & Cox, J.S. 2005, "A non-RD1 gene cluster is required for Snm secretion in Mycobacterium tuberculosis", *Molecular microbiology,* vol. 57, no. 6, pp. 1653-1663.

Maciag, A., Dainese, E., Rodriguez, G.M., Milano, A., Provvedi, R., Pasca, M.R., Smith, I., Palu, G., Riccardi, G. & Manganelli, R. 2007, "Global analysis of the Mycobacterium tuberculosis Zur (FurB) regulon", *Journal of Bacteriology,* vol. 189, no. 3, pp. 730-740.

Maciag, A., Piazza, A., Riccardi, G. & Milano, A. 2009, "Transcriptional analysis of ESAT-6 cluster 3 in Mycobacterium smegmatis", *BMC microbiology,* vol. 9, pp. 48.

Marquis, K.A., Burton, B.M., Nollmann, M., Ptacin, J.L., Bustamante, C., Ben-Yehuda, S. & Rudner, D.Z. 2008, "SpoIIIE strips proteins off the DNA during chromosome translocation", *Genes & development,* vol. 22, no. 13, pp. 1786-1795.

Ogura, T., Whiteheart, S.W. & Wilkinson, A.J. 2004, "Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases", *Journal of structural biology,* vol. 146, no. 1-2, pp. 106-112.

Ohol, Y.M., Goetz, D.H., Chan, K., Shiloh, M.U., Craik, C.S. & Cox, J.S. 2010, "Mycobacterium tuberculosis MycP1 protease plays a dual role in regulation of ESX-1 secretion and virulence", *Cell host & microbe,* vol. 7, no. 3, pp. 210-220.

Pallen, M.J. 2002, "The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system?", *Trends in microbiology,* vol. 10, no. 5, pp. 209-212.

Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. 2011, "SignalP 4.0: discriminating signal peptides from transmembrane regions", *Nature methods,* vol. 8, no. 10, pp. 785-786.

Pym, A.S., Brodin, P., Brosch, R., Huerre, M. & Cole, S.T. 2002, "Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti", *Molecular microbiology,* vol. 46, no. 3, pp. 709-717.

Ramakrishnan, L., Federspiel, N.A. & Falkow, S. 2000, "Granuloma-specific expression of Mycobacterium virulence proteins from the glycine-rich PE-PGRS family", *Science (New York, N.Y.),* vol. 288, no. 5470, pp. 1436-1439.

Ramamurthi, K.S. 2010, "Protein localization by recognition of membrane curvature", *Current opinion in microbiology,* vol. 13, no. 6, pp. 753-757.

Ratledge, C. & Dover, L.G. 2000, "Iron metabolism in pathogenic bacteria", *Annual Review of Microbiology,* vol. 54, pp. 881-941.

Renshaw, P.S., Panagiotidou, P., Whelan, A., Gordon, S.V., Hewinson, R.G., Williamson, R.A. & Carr, M.D. 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. Implications for pathogenesis and virulence", *The Journal of biological chemistry,* vol. 277, no. 24, pp. 21598-21603.

Rodriguez, G.M., Gold, B., Gomez, M., Dussurget, O. & Smith, I. 1999, "Identification and characterization of two divergently transcribed iron regulated genes in Mycobacterium tuberculosis", *Tubercle and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease,* vol. 79, no. 5, pp. 287-298.

Rodriguez, G.M., Voskuil, M.I., Gold, B., Schoolnik, G.K. & Smith, I. 2002, "ideR, An essential gene in mycobacterium tuberculosis: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response", *Infection and immunity,* vol. 70, no. 7, pp. 3371-3381.

Sambrook, J.J., Fritsch, E.F. & Maniatis, T. 1989, *Molecular Cloning,* .

Sassetti, C.M., Boyd, D.H. & Rubin, E.J. 2003, "Genes required for mycobacterial growth defined by high density mutagenesis", *Molecular microbiology,* vol. 48, no. 1, pp. 77-84.

Sassetti, C.M. & Rubin, E.J. 2003, "Genetic requirements for mycobacterial survival during infection", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 100, no. 22, pp. 12989-12994.

Serafini, A., Boldrin, F., Palu, G. & Manganelli, R. 2009, "Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc", *Journal of Bacteriology,* vol. 191, no. 20, pp. 6340-6344.

Siegrist, M.S., Unnikrishnan, M., McConnell, M.J., Borowsky, M., Cheng, T.Y., Siddiqi, N., Fortune, S.M., Moody, D.B. & Rubin, E.J. 2009, "Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 106, no. 44, pp. 18792-18797.

Skjot, R.L., Oettinger, T., Rosenkrands, I., Ravn, P., Brock, I., Jacobsen, S. & Andersen, P. 2000, "Comparative evaluation of low-molecular-mass proteins from Mycobacterium tuberculosis identifies members of the ESAT-6 family as immunodominant T-cell antigens", *Infection and immunity,* vol. 68, no. 1, pp. 214-220.

Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T. & Jacobs, W.R.,Jr 1990, "Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis", *Molecular microbiology,* vol. 4, no. 11, pp. 1911-1919.

Snider, J. & Houry, W.A. 2008, "AAA+ proteins: diversity in function, similarity in structure", *Biochemical Society transactions,* vol. 36, no. Pt 1, pp. 72-77.

Sorensen, A.L., Nagai, S., Houen, G., Andersen, P. & Andersen, A.B. 1995, "Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis", *Infection and immunity,* vol. 63, no. 5, pp. 1710-1717.

Stamm, L.M., Morisaki, J.H., Gao, L.Y., Jeng, R.L., McDonald, K.L., Roth, R., Takeshita, S., Heuser, J., Welch, M.D. & Brown, E.J. 2003, "Mycobacterium marinum escapes from phagosomes and is propelled by actin-based motility", *The Journal of experimental medicine,* vol. 198, no. 9, pp. 1361-1368.

Stanley, S.A., Raghavan, S., Hwang, W.W. & Cox, J.S. 2003, "Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 100, no. 22, pp. 13001-13006.

Sweeney, K.A., Dao, D.N., Goldberg, M.F., Hsu, T., Venkataswamy, M.M., Henao-Tamayo, M., Ordway, D., Sellers, R.S., Jain, P., Chen, B., Chen, M., Kim, J., Lukose, R., Chan, J., Orme, I.M., Porcelli, S.A. & Jacobs, W.R.,Jr 2011, "A recombinant Mycobacterium smegmatis induces potent bactericidal immunity against Mycobacterium tuberculosis", *Nature medicine,* vol. 17, no. 10, pp. 1261-1268.

WHO 2011, *, WHO | Global tuberculosis control - epidemiology, strategy, financing* . Available: http://www.who.int/tb/publications/global_report/2009/en/index.html [2011, 4/12/2011].

Wirth, S.E., Krywy, J.A., Aldridge, B.B., Fortune, S.M., Fernandez-Suarez, M., Gray, T.A. & Derbyshire, K.M. 2012, "Polar assembly and scaffolding proteins of the virulence-associated ESX-1 secretory apparatus in mycobacteria", *Molecular microbiology,* vol. 83, no. 3, pp. 654-664.