Development of novel reagents for tuberculosis detection

Nqobile Angel Cebile Ngubane

A thesis submitted to the Faculty of Health Sciences, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, in fulfilment of the requirements for the degree of Doctor of Philosophy

07 June, 2013

DECLARATION

I, Nqobile Angel Cebile Ngubane declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy in the University of KwaZulu-Natal, Durban. It has not been submitted before for any degree or examination at this or any other university.

……..

……16th…….day of July 2013

DEDICATION

This work is dedicated to the memory of my father, Robert Mlungile Ngubane (1930-2000). For his love of education and life and all the support a child could ask for from a father. To my mother, Cecilia Thamsanqa Ngubane, for the depths of her love. There is no greater love.

PUBLICATIONS

Nqobile A. C. Ngubane, Lionel Gresh, Thomas R. Ioerger, James C. Sacchettini, Yanjia J. Zhang, Eric J. Rubin, Alexander Pym, and Makobetsa Khati (2012). High-throughput sequencing enhanced phage display identifies peptides that bind mycobacteria. PLoS One. ***Manuscript submitted.***

Nqobile A. C. Ngubane, Lionel Gresh, Eric J. Rubin, Alexander Pym, and Makobetsa Khati (2012). Selection of RNA aptamers against *M.tuberculosis* EsxG protein using surface plasmon resonance. Biochem Biophys Res Commun. ***Manuscript completed.***

PRESENTATIONS

Nqobile A. C. Ngubane, Lionel Gresh, Thomas R. Ioerger, Yanjia J. Zhang, Eric J. Rubin, Alexander Pym, and Makobetsa Khati .

Development of peptide aptamers for the detection of tuberculosis.

CSIR Emerging Researcher Symposium, 13 October 2011, Pretoria, South Africa (Oral).

Nqobile A. C. Ngubane, Lionel Gresh, Thomas R. Ioerger, Yanjia J. Zhang, Eric J. Rubin, Alexander Pym, and Makobetsa Khati .

Development of peptide aptamers for the detection of *M.tuberculosis*.

3rd SA TB Conference, 12-15 June 2012, Durban, South Africa (Oral).

Nqobile A. C. Ngubane, Lionel Gresh, Thomas R. Ioerger, Yanjia J. Zhang, Eric J. Rubin, Alexander Pym, and Makobetsa Khati .

Development of peptide aptamers for the detection of tuberculosis.

IBC Drug and Diagnostic Development Conference, 6-8 August 2012, San Francisco, USA (Poster).

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following people and institutions whose contribution was invaluable for the successful completion of this work.

I would like to express my deepest gratitude to my thesis supervisors; Dr Alexander Pym, Dr Makobetsa Khati and Prof Eric Rubin, for their continuous guidance and support. Special thanks to Prof Eric Rubin for giving me an opportunity to conduct research as part of his lab at Harvard University. To my fellow colleagues; the CSIR aptamer team, for their exceptional support during the past few years. Great thanks to Dr Lionel Gresh, for his critical scientific input and support throughout the project. Thanks to Dr. Noman Siddiq, for the training in the Biosafety level-3 laboratory. To the Rubin lab members, for reminding me that scientific research with all its sleepless nights is still the most exciting career one can pursue. Thanks to Dr Stoyan Stoychev for the mass spectrometry work done during this project. Thanks to Dr Thomas R. Ioerger and Dr Hamilton Gansen for their work on high throughput sequencing. Thanks to my family and friends for their understanding and always being my strength.

I would also like to thank the Colombia University, Forgarty AIDS and TB International Training and Research Programme for funding 6 months of my one year research period at Harvard University. Also thanks to the Howard Hughes Medical Institute (HHMI), for funding the rest of the 6 months research period at Harvard University.

Thanks to the Council for Scientific and Industrial Research (CSIR) as well as the Department of Science and Technology (DST) for funding this project. Additional thanks to CSIR Biosciences for funding my PhD studentship and enabling me to attend the IBC Drug and Diagnostic Development Conference in San Francisco, USA.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendices

LIST OF TABLES

Chapter 2

Chapter 3

Appendices

ABBREVIATIONS

µg	Microgram
µl	Microlitre
µM	Micromolar
ASSURED	Affordable, Sensitive, Specific, User friendly, Rapid and robust, Equipment-free, and Delivered
BSA	bovine serum albumin
bp	base pair
CFP-10	culture filtrate protein 10
CFU	colony forming unit
DNA	deoxyribonucleic Acid
DNase	deoxyribonuclease
EDTA	Ethylenediamine Tetra Acetic Acid
ELISA	Enzyme-linked Immunoabsorbent Assay
ESAT-6	early secretory antigenic target of 6 kDa
HIV	Human Immunodeficiency Virus
kDa	kiloDalton
LAM	Lipoarabinomannan
M	Molar
mg	Milligram
ml	Millilitre
mM	Millimolar
nM	Nanomolar
ng	nanogram

LB	Luria Bertani broth
NaCl	sodium chloride
PAGE	polyacrylamide gel electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PFU	Plaque forming unit
POC	Point-of-care
RNA	Ribonucleic Acid
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
SPR	Surface Plasmon Resonance
Tween-20	polyoxyethylene sorbitan monolaurate
Tween-80	polyoxyethylene sorbitan monooleate
V	volt
WT	wild-type

ETHICS

The research described in this thesis was approved by the Faculty of Medicine Ethics Committee (BREC REF: BE152/11) of the University of KwaZulu-Natal.

ABSTRACT

Tuberculosis (TB) is one of the most prevalent infectious diseases worldwide and causes high morbidity and mortality, despite the widespread availability of effective antibiotics against most strains of *Mycobacterium tuberculosis,* which is the causative agent of TB. One of the primary reasons that hinder TB control is that many cases of active disease go undetected or are discovered late. This is, in large part, due to the relative insensitivity and limited specificity, amongst other limitations, of the current TB diagnostics tests. Moreover, *M. tuberculosis* infection can be asymptomatic and latent, or cause active disease. Therefore, an ideal or effective TB diagnostic needs to distinguish between these two states. The aim of this study was to develop novel diagnostic reagents for *M. tuberculosis* using phage displayed peptides and nucleic acid aptamers with a view to discerning latent from active TB.

Using a linear (X_{12}) and constrained (CX_7C) phage display libraries, five rounds of selection (biopanning) were performed. Ten phage displayed peptides that bind to the mycobacteria surface were selected. These phage clones were identified using both random clone picking and high throughput (HTP) sequencing. A phage clone displaying the CPLHARLPC peptide was identified by HTP sequencing as the most enriched, representing 82.49% of the selected CX_7C phage population. Further characterization showed that it bound better to different mycobacteria species, including *M. tuberculosis*, than the unselected phage library. Moreover, using surface plasmon resonance (SPR) technology, the chemically synthesised

CPLHARLPC peptide was shown to bind *M. tuberculosis* H37Rv whole cell lysate and not non-mycobacteria lysates.

In addition, using the systematic evolution of ligands by exponential enrichment (SELEX) protocol and SPR technology, 2'-Fluoro-pyrimidine-RNA aptamers were selected against the mycobacteria ESX-3 secreted protein, EsxG. At least five aptamers were identified after five rounds of selection. Two of these aptamers, GH43 and GH78, not only bound EsxG with high affinities, K_D 8.04 ± 1.90 nM and K_D 78.85 ±9.40 nM respectively, but also preferentially bound EsxG better than the EsxA homologue.

Taken together, these findings suggest that a combination of phage display, SELEX and HTP sequencing can be a useful tool for the identification of specific detection reagents that can bind to mycobacteria and its associated targets. These reagents could be exploited to develop alternative molecular probes for TB diagnostics.

CHAPTER 1: *INTRODUCTION AND LITERATURE REVIEW*

1.0 Introduction

Tuberculosis, which is caused by *M. tuberculosis*, remains one of the major infectious diseases that lead to morbidity and death. Despite the widespread availability of effective antibiotics against drug sensitive *M. tuberculosis*, the World Health Organisation (WHO) estimates that in 2011 there were between 8.3 and 9.0 million incidences of TB globally of which 26% were in Africa (Figure 1.1). In the same year, there was an estimated 0.8 to 1.1 million deaths of HIV negative people globally that resulted from TB (World Health Organisation, 2012).

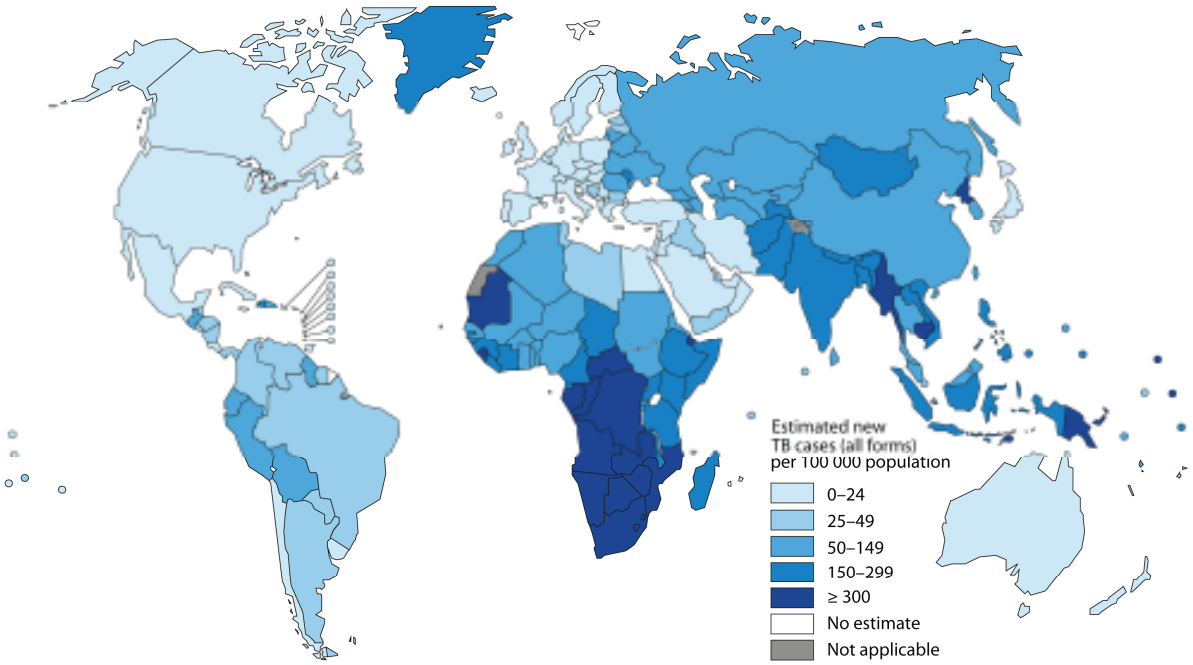


Figure 1. 1: Estimated tuberculosis (TB) incidence rates, 2011 (WHO, Global tuberculosis report, 2011)

The difficulty in controlling the spread of TB can be attributed to several reasons. These include the emergence of drug resistant *M. tuberculosis* strains (Senior, 2009), the considerable effort and expense incurred by supervising directly observed therapy (DOTS) and the socio-economic challenges that made continued supervision of treatment difficult. However, one of the primary reasons that hinder TB control is that many cases of active disease go undetected or are discovered late (Storla *et al.*, 2008, Reid & Shah, 2009). This is largely due to several limitations of current TB diagnostics such as low sensitivity and low specificity (Sohn *et al.*, 2009) and are reviewed in section 3.0 of this chapter. An additional complication in the diagnosis of TB is that *M. tuberculosis* infection can be latent or cause active disease. In an individual with an efficient cell-mediated immunity, the tuberculosis infection may be arrested permanently and subclinically. This state is called the latent TB infection (LTBI) or persistent state. However, if the initial infection cannot be controlled or the infected person becomes immunocompromised, infection may progress to the extent that it becomes clinically detectable. This state is called active TB (Dannenberg, 1994). This poses a particular challenge because it cannot be distinguished from latent TB using most of the current diagnositics. In addition to being able to distinguish latent from active TB, an ideal diagnostic should be affordable, sensitive, specific, user-friendly, rapid, equipment-free and deliverable to the end-user (ASSURED).

Generally, two broad approaches are used to develop diagnostic tests. First, host associated markers with levels or activity that are changed in response to the pathogen, can be measured. Serological antibody detection tests that measure the circulating antibody response to *M. tuberculosis* antigens are an example of this kind of diagnostic test. The second approach is to detect the *M. tuberculosis* bacilli and/or a part of the

bacilli. The main advantage of this approach is that it provides a direct evidence of bacterial presence. Morevoer, this approach is more likely to differentiate between LTBI and active disease. For this reason, this study adopted the second approach and targeted the the *M. tuberculosis* cell surface and its secreted protein EsxG for the developed detection reagents.

The literature review summarizes the current knowledge on *M. tuberculosis* with a focus on the cell wall and type VII secretion system, to give context to their use as selection targets for this project. Also included, is a review of the limitations and advantages of the current TB diagnostics. Phage display and aptamer technology are also reviewed, focusing on their application in diagnostics. Finally, the scope and the objectives of the research project are summarised followed by the structure of the thesis.

2.0 Mycobacterium tuberculosis

M.tuberculosis belongs to the *Mycobacterium* genus. There are currently more than 120 recognised species of mycobacteria (Lettieri, 2007, Tortoli, 2006). The mycobacteria bacillus has variable lengths between 1.5 and 4.0 μm depending on the period of culture, and a diameter of 0.3-0.5 μm (Wayne and Diaz, 1986). There are two groups of mycobacteria, fast growers like *M. smegmatis* and slow growers like *M. tuberculosis*. The fast growing species are characterised by their ability to form colonies on solid media in less than 7 days while the slow growing species require more than 7 days (Shinnick and Good, 1994). Generally, the slower growing species are pathogenic to humans and animals while this is not usually the case with the faster growing mycobacteria species (Shinnick and Good, 1994).

Although *M. tuberculosis* is the predominant cause of human TB, other mycobacteria species can cause human and/or animal TB. These species are grouped together within the *Mycobacterium tuberculosis* complex (MTBC). The MTBC consists of closely related organisms namely, *M. tuberculosis*, *M. africanum*, *M. bovis*, *M. bovis* BCG, *M. caprae*, *M. microti*, *M. pinnipedii*, and *M. canettii* (Brosch et al., 2002, Cousins et al., 2003, Cousins et al., 1994, Tsukamura et al., 1985). The mycobacteria characteristic of interest for this study is its complex cell envelope and the associated specialised type VII secretion (T7S) system which are described below.

2.1 The cell envelope of *M. tuberculosis*

The cell envelope of *M. tuberculosis* is a complex structure with diverse components that can serve as potential biomarkers for suitable diagnostics. It comprises mainly of two segments, the cytoplasmic inner membrane (IM) and the cell wall. The cell wall segment consists of a peptidoglycan layer (PG) that is covalently attached to arabinogalactans (AG) (Figure 1.2). Arabinogalactans are in turn attached to mycolic acids, which have long meromycolate and short α-chains and account for 30 to 40% of the cell envelope mass. These form part of the mycobacterial outer membrane (MOM). The MOM also consist of a large variety of non covalently bound extractable lipids (Daffe and Draper, 1998), some of which have longer fatty acids that complement the shorter α-chains of the mycolic acids, while others have shorter fatty acids that complement the longer chains. Distributed amongst the MOM are cell wall proteins, pore forming proteins, (Niederweis, 2003), phosphatidylinositol mannoside (PIMS), phthiocerol-containing lipids, lipomannan (LM) and lipoarabinomannan (LAM) (Brennan, 2003a).

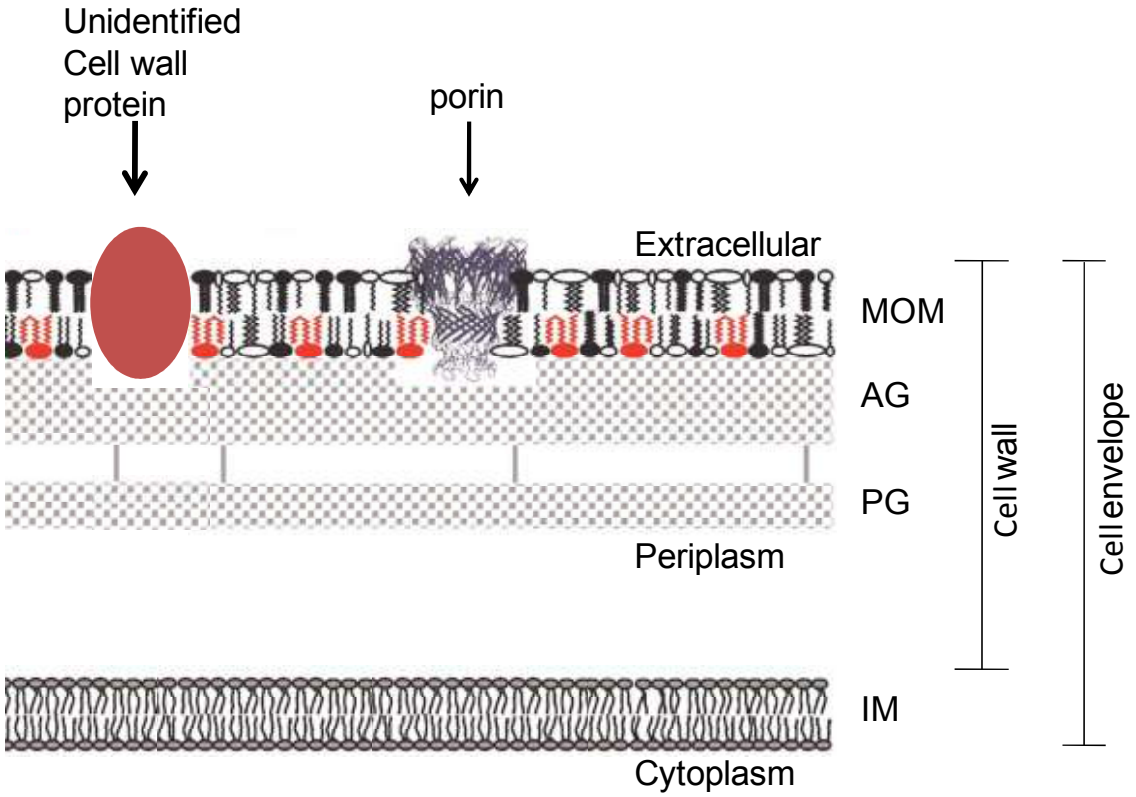


Figure 1. 2: Schematic of the mycobacteria cell wall

Schematic representation of the mycobacterial cell envelope adapted from (Cook et al., 2009). Mycolic acids which are covalently linked to the arabinogalactan (AG)–peptidoglycan (PG) co-polymer are drawn in red and inserted in the folded conformation as part of the inner leaflet of the mycobacterial outer membrane (MOM). As part of the MOM, a porin is represented in blue, an unidentified cell wall protein in red, while the extractable lipids are represented in black. The inner membrane (IM) and the periplasmic layer are also indicated on the schematic.

Indeed, diagnostics based on bacterial components are successfully used in a variety of infections, including Legionella (Dominguez et al., 1999), Cryptococcus (Frank et al., 1993) and Histoplasma (Connolly et al., 2007). Morevoer, the mycolic acids of mycobacteria which are species specific, have previously been investigated for TB diagnostics using various techniques like high-performance liquid chromatography

(Butler and Guthertz, 2001, Butler and Kilburn, 1988) and electrospray ionization-tandem mass spectrometry (Song et al., 2009). However, these tests are limited in that they are laboratory based tests and can only be used on cultured mycobacteria and not directly on clinical samples.

A test for a mycobacterial glycolipid, lipoarabinomannan (LAM), has recently been evaluated on both urine and sputum, although it is limited by the low sensitivity range of 6%-50.7% although it showed a good specificity range between 87-100% (Dheda et al., 2010, Hamasur et al., 2001, Reither et al., 2009). There are likely to be numerous other cell wall components that are shed by the bacterium and can be possibly detected in patients' samples. The primary reason that LAM has been investigated further is that, unlike many other *M. tuberculosis* glycolipids, it is antigenic and, therefore, straightforward to measure by enzyme-linked immunosorbent assay (ELISA). Since many of the potential targets on the surface of the bacterium are glycolipids and are unlikely to be antigenic, the conventional use of an antibody as a detection probe will be limited against some of the cell wall components. Hence, this research project explored alternatives to antibodies such as phage displayed peptides and nucleic acid aptamers to develop as detection reagents for the *M. tuberculosis* cell envelope components.

Mycobacteria like other intracellular pathogens secrete proteins to stimulate a supportive environment for survival and replication during infection of the host. To achieve this, mycobacteria uses at least four pathways to transport proteins across its impermeable cell wall. These are the SecA1 mediated general secretory pathway (Pugsley, 1993, Wiker and Harboe, 1992), SecA2-operated pathway (Braunstein et al., 2003), the twin-

arginine translocation system (McDonough et al., 2008, Saint-Joanis et al., 2006) and the recently named, specialised type VII secretion (T7S) system (Abdallah *et al.*, 2007).

2.2 The type VII secretion (T7S) system of *M. tuberculosis*

In the virulent strain of *M. tuberculosis* the T7S system consists of five paralogous gene clusters namely, ESX-1 (genes *Rv3866-Rv3883c*), ESX-2 (*Rv3884c-Rv3895c*), ESX-3 (*Rv0282-Rv0292*), ESX-4 (*Rv3444c-Rv3450c*) and ESX-5 (*Rv1782-Rv1798*) (Gey Van Pittius et al., 2001, Smith, 2003). Of the five ESX loci, only three have been demonstrated to encode functional secretion systems in mycobacteria. These are ESX-1 which is required for secretion of virulence factors (Ganguly et al., 2008, Lewis et al., 2003, Smith, 2003, Stanley et al., 2003), ESX-3 which is required for optimal iron acquisition under low iron environments (Serafini et al., 2009, Siegrist et al., 2009) and ESX-5 which exports Pro-Glu (PE), pro-pro-glu (PPE) and proline-glutamic acid_polymorphic guanine-cytosine-rich sequence (PPE_PGRS) proteins, that are associated with virulence in mycobacteria (Abdallah et al., 2006, Abdallah et al., 2009).

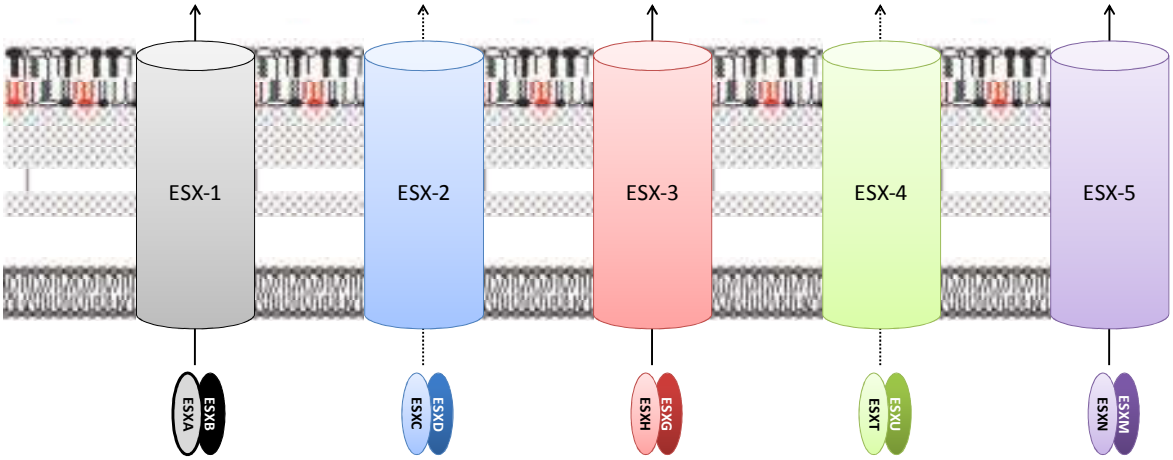


Figure 1. 3: The five ESX systems and their corresponding secreted WXG100 protein.

The Esx proteins corresponding to each T7S system gene cluster are indicated below its associated ESX secretion system illustration. Several other substrates have been demonstrated to be secreted by the ESX secretion systems and are not depicted in this fugure. Of the five ESX loci, only three (indicated by the solid arrow), ESX-1, ESX-3 and ESX-5, have been demonstrated to encode functional secretion systems in mycobacteria.

One of the extensively characterised T7S secreted proteins is the ESAT-6:CFP10, also known and the EsxA:EsxB protein complex that is secreted by the ESX-1 secretion system (Smith, 2003, Stanley et al., 2003), and is required for pathogenicity of *M. tuberculosis* (Guinn et al., 2004, Hsu et al., 2003, Lewis et al., 2003, Pym et al., 2002). Both EsxA and EsxB proteins are part of the 23 small (EsxA-W) secreted proteins in *M. tuberculosis* that belong to the WXG100 family and are encoded by the *esx* gene pairs. This protein family is characterised by their size of approximately 100 amino acids, their helical structure and a helical bend that is formed by the conserved Trp-Xaa-Gly (W-X-G) motif (Pallen, 2002). There are 11 genomic loci in the *M. tuberculosis* genome that encode for the *esx* gene pairs (Cole et al., 1998). Five of the *esx* gene pairs are flanked

by the genes coding for components of the T7S system (ESX-1-ESX-5), which is involved in the export of Esx proteins (Abdallah et al., 2007, Champion et al., 2006, Gey Van Pittius et al., 2001). This duplication of the *esx* gene clusters is also present in other mycobacteria species including *M. bovis*, *M. laprae*, the non pathogenic mycobacteria strain *M. smegmatis* (Gey Van Pittius et al., 2001). Interestingly, *M. leprae* which is believed to possess the minimal gene set required for pathogenesis has only retained the genes corresponding to the *esx* cluster 1 and cluster 3 (Cole, 2001, Gey Van Pittius et al., 2001). This suggests that these two *esx* gene clusters are essential for either mycobacteria virulence or growth. Indeed, the ESX-3 system is the only ESX system that is essential for growth in *M. tuberculosis* (Sassetti et al., 2003), while ESX-1 is essential for virulence of mycobacteria. Of relevance to this thesis is the ESX-3 secretion system and its secreted protein, EsxG, which was targeted in this study.

2.3 The mycobacteria ESX-3 secterion system

The Esx-3 secretion system has been implicated in mycobacteria growth under low iron conditions (Serafini et al., 2009, Siegrist et al., 2009). Iron, is an essential nutrient for most microorganisms. At least 40 enzymes encoded by the *M. tuberculosis* genome require iron as a cofactor (Cole et al., 1998). Iron acquisition is a challenge for intracellular pathogens as it is not freely available in the host, but instead is tightly bound to iron-binding host proteins (Weinberg, 1999). For this reason, successful intracellular pathogens have evolved complex ways to overcome the iron limited environment experienced in the host during infection. Mycobacteria are no exception, as, there is now evidence indicating that *M. tuberculosis* does in fact encounter low iron conditions during infection of macrophages (Gold et al., 2001).

One of the ways that mycobacteria has evolved to acquire iron is by the use of small-molecule compounds, called siderophores, for high affinity chelating of ferric iron (Miethke and Marahiel, 2007). The siderophores are generally secreted by the organism to compete for the available iron in their environment (De Voss et al., 1999). The genus of mycobacteria has two kinds of siderophores, mycobactin and exochelin (Gobin and Horwitz, 1996, Wong et al., 1996).

While non pathogenic mycobacteria such as *M. smegmatis*, produce both mycobactin which are molecules distinguished mainly by the presence of a phenyloxazolidine ring (Harrison et al., 2006), and exochelin which is a peptidic siderophore (Ratledge and Ewing, 1996), the pathogenic *M. tuberculosis* produces two classes of mycobactin, namely carboxymycobactin which is a more polar form and is secreted by the bacillus and mycobactin that is less polar and is cell bound (Ratledge, 2009). The synthesis of the siderophoresis is highly regulated, occurs exclusively under low iron conditions (Rodriguez et al., 2002), and is essential for growth in low iron medium. DeVoss and colleagues (2000) demonstrated this by showing that a mutant strain (mbtB) of *M. tuberculosis* which is defective in the synthesis of mycobactin is unable to replicate in low iron medium and macrophages (De Voss et al., 2000).

Iron homeostasis in *M. tuberculosis* is mainly maintained by the iron-dependent regulator (IdeR) protein. In the presence of iron, IdeR binds to the promoters and represses the transcription of *mbt* genes which are essential for siderophore synthesis (Gold et al., 2001). In addition, IdeR represses the transcription of the ESX-3 gene cluster in both *M. tuberculosis* (Rodriguez et al., 2002) and *M. smegmatis* (Maciag et al.,

2009), while low iron conditions induce the transcription of the ESX-3 cluster in *M. tuberculosis* (Rodriguez et al., 2002).

One of the known substrates of ESX-3 is the EsxG:EsxH heterodimer (Lightbody et al., 2004, Siegrist et al., 2009). A recent study by Siegriest and colleagues (2009) demonstrated that a *M. smegmatis* mutant which lacks ESX-3 (ΔESX-3) is unable to be rescued by the exogenous addition of purified siderophore, mycobactin, indicating that the mutant is unable to use iron-bound mycobactin. Serafini and colleagues (2009) reported similar findings using a *M. tuberculosis* ESX-3 conditional mutant. Serafini and colleagues (2009) further hypothesised that ESX-3 is responsible for the secretion of some unknown factor(s) which is required for the optimal uptake of iron and zinc. Since EsxG and EsxH are known to be secreted by the ESX-3 system, it raises the question if one or both of these proteins are possibly the unknown factor secreted by this secretion system for optimal iron uptake.

Taken together, I reasoned that since EsxG and EsxH are secreted by the ESX-3 system which in turn is induced under low iron conditions both of these two proteins are likely to be present in an infected host and can be investigated as a potential tuberculosis biomarker, if appropriate detection were available. If suitable reagents that can detect either EsxG or EsxH were developed, they would also be useful tools to study the role of these individual proteins in iron acquisition of *M. tuberculosis*.

3.0 Current TB diagnostics

The various available TB diagnostic tests fall in two major groups, those that detect the pathogen and/or itsassociated biomarkers and those that measure host associated biomarkers in response to the pathogen infection. Both groups of tests have advantages when applied in an appropriate environment but also posess inherent limitations. However, to date we do not have but urgently need, a TB diagnostic with high specificity and sensitivity and also offers the ease of use at point of care fascilities in resource constrained environments where TB is mostly predominant.

3.1 Direct detection of MTBC or its associated components in clinical samples

3.1.1 Sputum smear microscopy

The most widespread test for diagnosing TB is the sputum smear microscopy test. This test detects mycobacteria by staining and identification of acid fast bacteria (AFB) under a microscope. There is a conventional smear microscopy test which uses a light microscope to detect AFB on the sputum sample. Its sensitivity is up to 70–80% in some settings (Steingart et al., 2007c) but it cannot detect paucibacillary disease. The sensitivity of smear microscopy is iimproved by the use of a fluorescence microscope that uses an acid-fast fluorochrome dye to stain the mycobacteria for visualising under the microscope (Kivihya-Ndugga et al., 2003, Bell and Brown, 1962, Githui et al., 1993). While this test only requires one piece of equipment which is a microscope, it still needs trained personnel to be conducted effectively. Both of these requirements necessary for smear microscopy might to be a limitation for their adoption in resource limited settings. Furthermore, the test sensitivity is significantly lower in areas with

high HIV prevalence due to a low bacterial burden in sputum of HIV positive patients (Steingart et al., 2006).

3.1.2 Solid Culture

Culture based tests are based on detection of mycobacteria growth on specialised culture medium. The culture based methods are still the gold standard in TB diagnostics. They can be performed in liquid culture as described in 3.1.3 or in solid culture like the Lowenstein-Jensen medium. The main advantage of culture tests is that they offer a higher sensitivity than smear microscopy as they can detect as low as 10 bacilli per milliliter of sputum when compared to the minimum range of detection by smear microscopy of 5000 - 10000 bacilli per millilitre of sputum (American Thoracic Society. Committee on Revision of Diagnostic, 2000). However, the major lmitatation of this test is that it can take up to eight weeks to get the result as it is dependent on bacterial growth (Kent and Kubica, 1985).

3.1.3 Automated Liquid Culture

Liquid culture methods offer a quicker turnaround time from sample collection to diagnosis when compared to solid culture medium, without a compromise on sensitivity. For example, the previous radiometric system Bactec TB 460 (Becton-Dickinson, USA) required on average 13 days to give a positive result for TB (Abe et al., 1992). However, it had its limitations. In addition to complexity of performing this test, the Bactec TB 460 system used radioisotopic reagents which required specialised disposal methods resulting in additional costs associated with performing the test (Palacios et al.,

1999). However, there are now less complex, automated systems based on liquid culture that do not require the use of radioisotopic reagent. These include the BACTEC 9000 MB system which has an average time of 10 days to mycobacteria detection (Pfyffer et al., 1997a, Zanetti et al., 1997), the Difco ESP Culture System II which averages 13.1 days to detection of MTBC (Woods et al., 1997) and the MB/BacT system (Palacios et al., 1999) which takes an average of 13.2 days to mycobacteria detection (Nogales et al., 1999) and and can also be used for antibiotic drug susceptibility (Diaz-Infantes et al., 2000). Furthuremore, there are now culture tests like the the BBL Mycobacteria Growth Indicator Tube (MGIT) that do not require instrumentation but can be inspected visually for mycobacteria detection (Pfyffer et al., 1997b). The test requires on average 9.9 days to detect mycobacteria (Pfyffer et al., 1997b). While great improvements had been achieved in improving the time to detection using liquid cultures, these test fall short of being used at point of care and require trained personeel to be performed correctly. This limits their use in resource poor environments.

3.1.4 Nucleic acid amplification test (NAAT)

Nucleic acid amplification tests (NAAT) are based on amplification of nucleic acids from MTBC using various nucleic acid amplification techniques. There are three main NAAT, the Amplified Mycobacterium tuberculosis Direct (MTD) Test™ (Gen-Probe®, San Diego, CA), the Amplicor® Mycobacterium tuberculosis (Amplicor MTB) Test (Roche® Diagnostic Systems, Branchburg, NJ) and the GeneXpert MTB/RIF assay (Cepheid, Sunnyvale, CA). The MTB and the Amplicor MTB test have been shown to have a detection sensitivity between 83% and 100% and a specificity between 97% and 100% when compared to culture (Moore et al., 2005). The most recent of these tests, the

GeneXpert, has been shown to have a turnaround time of less than 2 hours and a detection limit of approximately 100 organisms (van Zyl-Smit et al., 2011). Thefast turnaround time to MTBC detection and the high sensitivity and specificity of these tests offers a much needed improvement in TB diagnostics. However, the cost of the equipment, the consumables and the training of personnel required by these tests limits their usability in poor environments.

3.1.5 Loop mediated isothermal amplification (LAMP) test

The LAMP test is based on the amplification of a target gene in mycobacteria without the use of a thermocycler. The LAMP test is performed in a single tube with all the reagents, and the reaction occurs in 60 min, under isothermal conditions between 60ºC - 65ºC. The results of the test can be acquired without the use of instrumentation by adding SYBR Green I to the reaction. The test has been successfully used to detect M. tuberculosis by targeting the *gyrB* gene (Iwamoto et al., 2003) and the16S rRNA (Pandey et al., 2008). At the reaction time of 60min, the LAMP test using primers targeting the *gyrB* gene showed similar sensitivity to the Amplicor MTB test for the detection of MTBC, with a detection limit of 5 to 50 copies of purified DNA (Iwamoto et al., 2003). The LAMP test offers a simpler alternative to the conventional NAAT, because it eliminates the need for expensive equipment.

3.1.6 Lipoarabinomannan (LAM) antigen detection test

The test detecting the mycobacterial glycolipid, LAM, has demonstrated that detection of TB associated components can be exploited in the development of TB diagnostics. A metanalysis review by Flores and colleagues (2011) showed that this test has a

sensitivity of 87% in sputum and 53% in urine (Flores et al., 2011). They further showed that the sensitivity of the test is higher (47%) in HIV-infected individuals when compared to HIV-uninfected (14%) individuals. While the LAM test offers a good proof-of-principle that an antigen detecting test can be performed on easily obtainable clinical samples like urine, the low sensitivity of the test remains a limitation in its use for TB diagnosis.

3.1.7 Phage amplification test

Phage amplification tests use bacteriophages to infect live *M. tuberculosis* and detect the bacilli by subsequently amplifying the phage (Wilson et al., 1997) or detecting the light that is emitted by the phages with a luciferase reporter after they have infected the live bacilli (Jacobs Jr et al., 1993). A meta-analysis study by Pai and colleagues (2005) showed that of the two studies included in their meta-analysis which performed phage assays directly on sputum; one study reported a sensitivity of 100% and a specificity of 99% (Albert et al., 2004), while the other reported a sensitivity of 86% and specificity of 73% (Butt et al., 2004, Pai et al., 2005). While the sensitivity and specificity of this test is higher than that of smear microscopy, its limitation is that it requires equipment similar to that of culture based test and can only be performed in a laboratory environment with trained personeel.

3.2 Indirect detection of MTBC infection by measuring host associated biomarkers

The major limitation shared by this group of diagnostic test is the inability to differentiate between latent and active TB infection. Nonetheless, in settings where TB

is not prevalent or an alternative diagnostic test is either not available or effective, some of these tests are used.

3.2.1 Tuberculin skin test (TST)

This test measures the patients' immune response to *M. tuberculosis* purified protein derivative (PPD). While the test has been successfully used for diagnosis of latent TB, it has limitations which include subjectivity in the administration and reading of the test results. Furthuremore, BCG vaccination increases the probability of a positive TST result in the absence of latent infection (Wang et al., 2002). The cross reactivity of this test with BCG vaccination wich is still used in countries with a high burden of TB disease, and the inability of the test to distinguish latent from active TB, limits its application in high burden countries.

3.2.2 Serological antibody detection test

The serological antibody tests are based on the detection of antibodies against *M. tuberculosis* in clinical samples. While serological tests are successfully used against other pathogens, the sensitivity and specificity results from these tests in TB diagnsostic are very variable and inconsistent (Steingart et al., 2007b, Steingart et al., 2007a). This is the primary reason that the WHO in July 2011 issued a warning that serological blood tests should not be used for the diagnosis of active TB.

3.2.3 Adenoside deaminase (ADA)

This tests measures the levels of ADA, a purine degrading enzyme implicated in mononuclear phagocyte maturation (Fischer et al., 1976, Carson and Seegmiller, 1976),

in pleural fluid (Greco et al., 2003). This test sensitivity for pleural fluid ADA of ≥50 U·L−1 was shown to be 95% and specificity 89% when compared to thoracoscopy, which is the most sensitive test for diagnosis of TB pleurisy (Diacon et al., 2003). This test offers an alternative to the diagnosis of TB pleurisy in settings where thoracoscopy is not affordable.

3.2.4 T-cell-based interferon-y release assays (IGRA's)

The gamma interferon (IFN-γ) release assays (IGRA) are based on the measurement of the IFN-γ release by foreign epitope-stimulated T cells (Tsiouris et al., 2006). An IGRA test using ESAT-6 and CFP-10 antigens was demonstrated to have a specificity of 98.1% and the sensitivity of 89.0% (Mori et al., 2004). This type of test is commercially available as QuantiFERON-TB Gold In-Tube (Cellestis Inc, Carnegie, Australia). Another commercially available IGRA is the T-SPOT.TB (Oxford Immunotec Ltd, Abingdon, UK), which measures the total number effector T cells expressing IFN-γ. The main advantage of the IGRAs is that they simultaneously respond to multiple targets, which allows for continues improvement of the specificity and sensitivity if multiple specific targets are identified and tested. The main limitation is that to perform the test, trained personnel and specialised equipment is required which limits it applicability in poor settings.

Considering the various inherent limitations in the current TB diagnostics, there is an urgent need to develop a test that can be used at point-of-care (POC). The advantage of a test that detects the pathogen or a pathogen associated marker is that it can distinguish between latent and active disease. However, the challenge in selecting a pathogen

biomarker is that they are likely to be present in low concentrations in clinical samples. To date, antibody based tests have performed poorly, showing mainly low specificity. For this reason this study investigated the use of nucleic acid aptamers, as well as phage display as detectiom probes, which have proved to be a good alternative to antibodies.

4.0 Aptamers in diagnostics

Traditionally, antibodies are used as detection probes in diagnostic applications. In the past, one of the major limitations in the use of antibodies in diagnostic was attaining adequate production scales to meet market needs. However, improvements in the antibody production processes today have enabled monoclonal antibodies to be produced indefinitely by continuous culturing of a single selected clone resulting in culture processes that can produce up to 5 g/l of antibody concentration (Birch et al., 2008, Winder and Rob, 2005). Indeed, this advance in antibody production has enabled continuous use of antibodies in various applications including diagnostics. Nonetheless, there are still limitations in the use of antibodies as diagnostic probes (Jayasena, 1999) These include the batch to batch variation of antibodies which require an immunoassay to be performed and re-optimised for each produced batch. Also, antibody-target binding properties cannot be changed on demand. In addition, their sensitivity to temperature which results is an irreversible denaturing, limits their application in diagnostics as the final environment cannot always be regulated. Antibodies also bind to a restricted number of macromolecules which are "antigenic". To address some of these limitations, this study considered aptamers as an alternative class of detection probes for diagnostics.

Aptamers are artificial nucleic acid or peptide ligands which are selected using an in vitro process and bind to a specific target molecule with high specificity and affinity. Their ability to recognize and bind their targets is similar to that of antibodies (Tasset et al., 1997, Xu and Ellington, 1996). They can be potent inhibitors of protein function (Kumar *et al.*, 1997), which also makes them applicable as ideal candidates in drug development as well as pathway elucidation in scientific research**.** For example, there is already an FDA-approved RNA aptamer, Macugen, which is an inhibitor of vascular endothelial growth factor (VEGF) and is used for the treatment of age-related macular degeneration (AMD) (Lee et al., 2005)**.** In addition, aptamers also offer the following advantages as alternatives to antibodies:

- The in vitro selection of aptamers simplifies their identification when compared to the in vivo selection of antibodies
- Unlike antibody selection, which is limited to in vivo conditions, aptamer selection conditions can be varied to support their intended use in the final diagnostic assay.
- Antibodies are difficult to generate against non-antigenic or toxic molecules, but the in vitro selection process of aptamers enables their selection against such targets.
- The chemical synthesis of aptamers versus the common cell culture methods used in the production of antibodies allows for minimal batch to batch variations.
- In addition, chemical synthesis of aptamers also allows for the attachment of functional groups which enable simple derivatisation of aptamers with other reporter molecules such as biotin and fluorescein for ease of detection or read out in diagnostic applications.

- Finally, the temperature stability and reversible denaturing ability of aptamers allows for long term storage and ease of transportation at ambient temperature which is of benefit especially in POC diagnostic assays which in the case of TB diagnostics will generally be used at primary health care facilities in remote areas.

4.1 Nucleic acid aptamers in diagnostics

Nucleic acid aptamers have been developed for diagnostics purposes. These include a thrombin-binding DNA aptamer (Tombelli et al., 2007) which has been investigated in various assays (Baldrich et al., 2004, Cai et al., 2002, Gronewold et al., 2005, Radi et al., 2006, Zhang et al., 2006) for the diagnosis of diseases associated with coagulation abnormalities. Another aptamer with diagnostic potential is the RNA aptamer that bind to S-adenosylhomocystein (SAH), which has potential for use in the diagnosis of cardiovascular disease (Gebhardt et al., 2000). At least three nucleic acid aptamers have been developed that can detect whole cell bacterial pathogens. These include aptamers that can detect *Campylobacter jejuni* (Tombelli et al., 2007), which is the most common cause of food borne infections (Altekruse et al., 1999), DNA aptamers that detect *Francisella tularensis* (Tombelli et al., 2007) which causes tularemia and DNA aptamers that bind *M. tuberculosis* (Chen et al., 2007). One DNA aptamer (NK2) selected against *M. tuberculosis* H37Rv has been shown to prolong the survival rates of mice challenged with intravenous therapy H37Rv and is being further evaluated for its therapeutic potential (Chen et al., 2012). This success in aptamer selection has led to continuous investigation of aptamers as detection probes in diagnostics.

The use of aptamers in point of care diagnostics has been explored in various platforms. Aptamers have been successfully demonstrated to be adaptable to POC detection methods previously developed for antibodies. These include the use of aptamers in different biosensor platforms (Navani and Li, 2006) like the quartz crystal protein biosensor (Liss et al., 2002) and the rapid plasmon resonance biosensor (Proske et al., 2005). However, aptamer-based biosensors are currently not user-friendly for use at point of care. The ideal POC test for TB in resource limited settings will need a colorimetric read out system, where the results can be interpreted with the naked eye. Two platforms with a colorimetric readout have been successfully explored using aptamers. The first one uses gold nanoparticles functionalized with DNA aptamers to construct sensors with colorimetric read out in response to various analytes such as adenoside, cocaine and potassium ions (Liu and Lu, 2006). In the second platform, Lui and colleagues (2006) further demonstrated a DNA aptamer based cocaine-sensing lateral flow device (Liu et al., 2006). The lateral flow was used to detect varying concentrations of cocaine in a buffer solution as well as cocaine in undiluted human blood serum (Liu et al., 2006). Both platforms can be exploited for TB POC diagnostic using aptamers.

Nucleic acid aptamers are selected against a particular target by an invitro selection process called the systematic evolution of ligands by exponentialenrichment (SELEX) (Ellington and Szostak, 1990, Tuerk and Gold, 1990). SELEX is an in vitro selection process consisting of sequential steps of selection and amplification to reduce a library of nucleic acids with randomised sequences to a minimised subset of one or more sequences that bind tightly to the target of choice (Que-Gewirth and Sullenger, 2007).

The first step in the SELEX process is the design and chemical synthesis of the random nucleic acid library. There are three things to consider during the design of the library, namely the length of the random region, type of randomization and the chemistry of the nucleic acid pool (Marshall and Ellington, 2000). Usually 10^{14} unique sequences are used for the initial round of selection (Gold et al., 1995). The second step in the SELEX process is the selection of ligand sequences that bind to the target. This is achieved by the incubation of the library with the target under appropriate temperature and buffer conditions (Gopinath, 2007). Different ratios of nucleic acid to target can be evaluated according to the level of stringency required for the experiment. However, the common practice is to start with a low ratio of nucleic acid to target, to validate whether all nucleic acid molecules bind the target (Marshall and Ellington, 2000). The third step of SELEX involves the use of various partitioning methods to separate the sequences that bind to the target from non-binding sequences. Advantages and limitations of the various partitioning methods, which include the use of nitrocellulose membrane filtration, affinity surfaces, column matrices, centrifugation, surface plasmon resonance (SPR), flow cytometry and capillary electrophoresis have been reviewed (Gopinath, 2007).

Successful use of the SPR based SELEX, which was adopted for this study, has been previously reported. These include RNA aptamers selected against the surface envelope glycoprotein (gp120) of the HIV-1 Ba-LR5 strain (Khati et al., 2003) and those selected against the human influenza virus hemagglutinin (Misono and Kumar, 2005). The SPR SELEX is based on an optical method that measures the refractive index on a sensor surface and is reviewed by (Van Der Merwe, 2000). In brief, the Biacore system is the most common SPR instrument used and has a sensor surface that forms the floor of a 20-

60 nl flow cell, with a continuous flow of running buffer. Typically, one molecule (the ligand) is immobilized onto the sensor surface and its binding partner (the analyte) is injected in a sample buffer through the flow cell, to monitor their interaction. Real time measurements are acquired as the analyte binds to the ligand. This results in an increase in the refractive index and is plotted as the response units (RUs) (Van Der Merwe, 2000). A typical SELEX protocol using and SPR instrument is illustrated (Figure 1.4).

The final step of the SELEX process involves the amplification of bound aptamers to be used as input library in the following round of SELEX. For nucleic acid aptamers the amplification of the bound sequences is achieved by PCR and in the case of RNA aptamers, this is followed by in vitro transcription to convert the DNA into RNA. Traditional SELEX usually consists of 8-12 rounds of selection (Bowser, 2005). At the end of SELEX, monoclonal binding sequences are selected from the final population and validated for binding to their target.

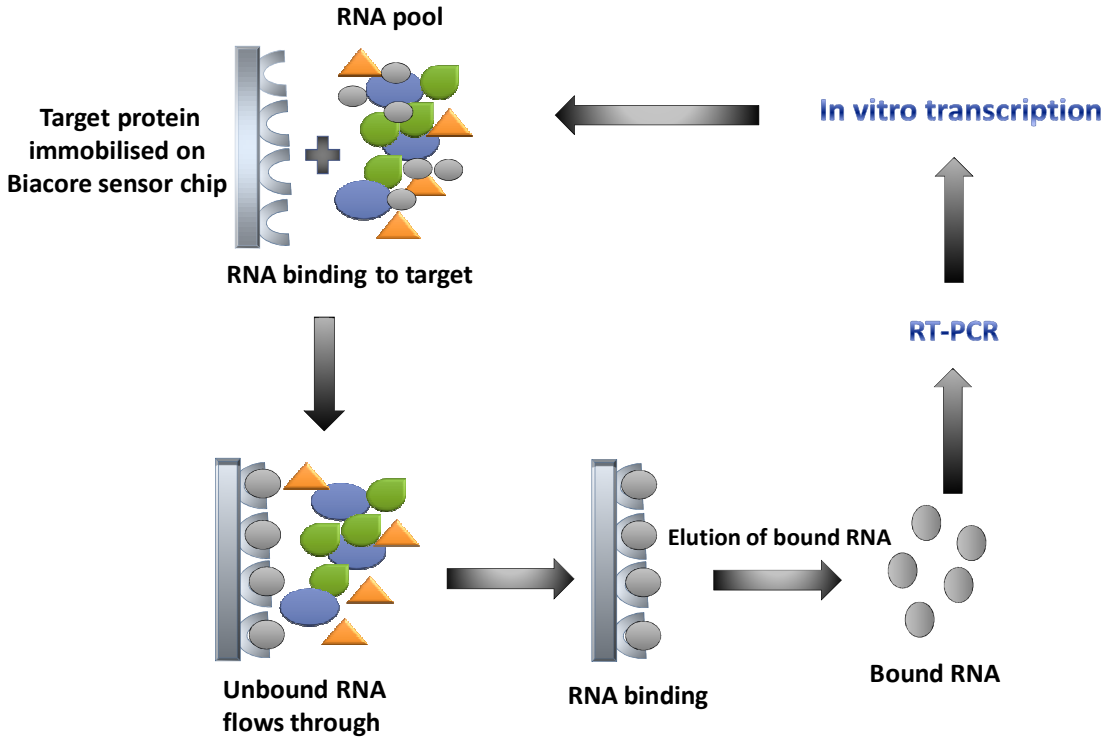


Figure 1. 4: Schematic showing a typical SELEX using SPR, Biacore.

4.2 Phage display in diagnostics

There are several technologies available for the selection of peptide reagents. These include phage display technology, the two-hybrid approach, ribosome display and the mRNA display (James, 2001). This study adopted the phage display technology approach for the selection of peptides that bind to *M. tuberculosis*. Phage display has previously been used to identify peptides with diagnostic potential. Anandakumar and colleagues used a phage library that displayed 12-mer random peptides to identify candidate peptides for the diagnosis of opportunistic infections caused by *Candida albicans* (Anandakumar et al., 2011). Larbanoix and colleagues (2008) used a phage library that displayed a 7-mer cyclic peptide on the pIII protein to isolate peptide that bind amyloid plaque for the potential diagnosis of Alzheimers (Larbanoix et al., 2010).

Stratmann and colleagues (2002) selected 12-mer peptides using random phage display against *M. paratuberculosis* with a diagnostic potential for routine screening in bulk milk (Stratmann et al., 2002).

This technology was first introduced by G. Smith in 1985 (Smith, 1985). It enables the display of a large number (up to 10^{10}) of diverse peptides, proteins and antibodies on the surface of the filamentous phage (Azzazy and Highsmith, 2002). The most commonly used phage is the M13 filamentous phage (Figure 1.5), which infects *E. coli* as its natural host (Mullen et al., 2006). M13 is rod shaped with a circular genome of 6000-8000 bases which is enclosed by an envelope that is made of five different proteins (Kehoe and Kay, 2005, Smith and Petrenko, 1997). The phage is covered by several copies of the pVIII protein, while one end contains five copies of each minor proteins pIII and pVI. There are 3-5 copies of the minor proteins pVIII and pXI on the other end of the phage. While foreign peptides have been displayed on the N-terminus of all phage coat proteins (Jespers et al., 1995, Smith and Petrenko, 1997), the most commonly used are the pIII and the pVIII. While protein and antibody phage display libraries are in use, libraries that display random peptides are the most common. Randomised oligonucleotides of a certain length are cloned as a fusion protein to one of phage coat proteins, which results in their expression as a peptide-capsid fusion proteins (Azzazy and Highsmith, 2002). The phage displayed peptides can vary between 6 and 43 amino acids in length (Burritt et al., 1995, McConnell et al., 1996). Furthermore, structurally constrained peptides have been successfully displayed and screened for target binding and are advantageous in selecting high affinity peptides (McLafferty et al., 1993).

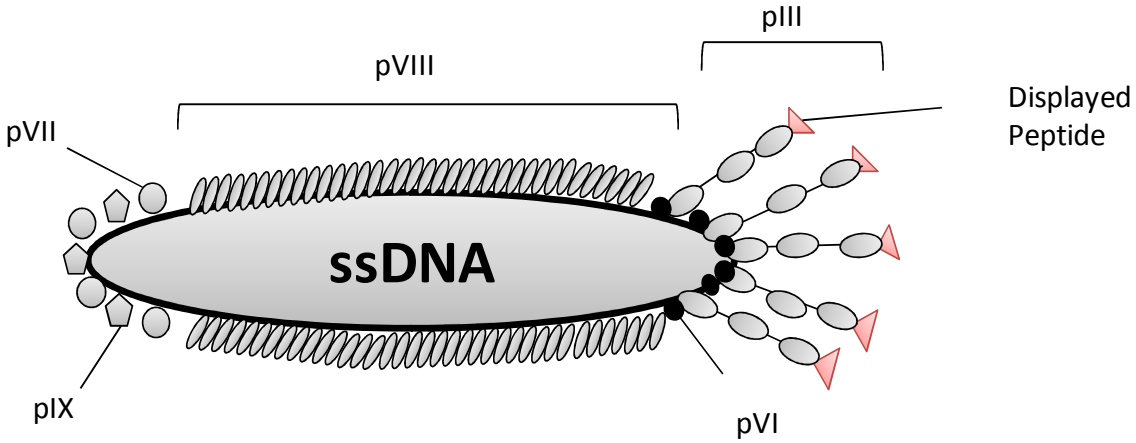


Figure 1. 5: Schematic view of the M13 filamentous phage particle displaying a peptide on the PIII protein.

Target binding phage clones are selected in an invitro process called biopanning. Biopanning involves several steps as illustrated in figure 1.6. In brief, the library is allowed to interact with the target. This is followed by several washes to remove unbound phage. Specifically bound phage will then be eluted from the ligand. This can be accomplished by using a solution of a known ligand for the target to compete with the bound phage or adding bacterial host (Smith and Petrenko, 1997). In a case where there is no known ligand for the target, a general low pH Glycine-HCl buffer can be used for non specific disruption of the binding interaction. The eluted phage is amplified by infection of *E.coli* bacteria (Smith and Petrenko, 1997). Usually three to five rounds of biopanning are necessary to select for a specific pool of phage. However, biopanning rounds should continue until a binding assay of the enriched pool indicates that the selected population is enriched (Kehoe and Kay, 2005, Menendez and Scott, 2005). The biopanning protocol can be modified at any stage for a specific target as described in chapter 2.

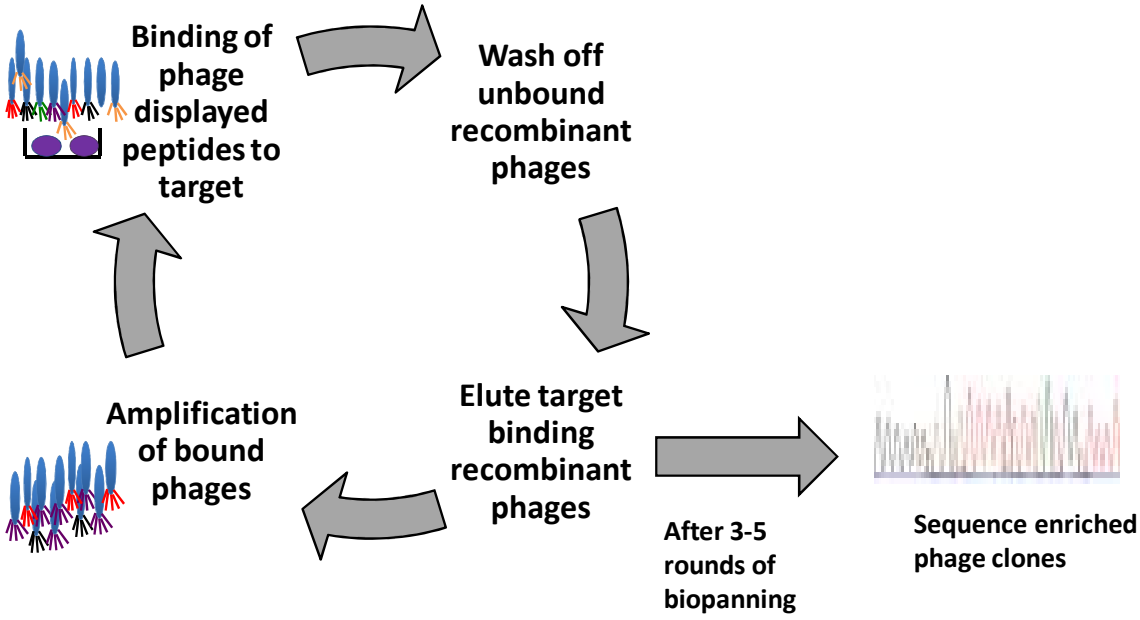


Figure 1. 6: Schematic showing a typical solid phase biopanning protocol

5.0 Aim and objectives

The aim of this thesis was to develop phage displayed peptides and nucleic acid aptamers, respectively, for detection of *M. tuberculosis*. These reagents were developed with a view to diagnosing TB.

Thus, the objectives of this study were:

- To isolate phage-displayed peptides that can detect *M. tuberculosis* bacilli
- To demonstrate that the isolated peptide reagents could bind their *M. tuberculosis* targets in different forms, namely, intact bacilli and lysed bacilli.
- To isolate and characterize RNA aptamers against *M. tuberculosis* ESX-3 secreted protein, EsxG

6.0 Structure of the thesis

In Chapter 2, the biopanning experiments against *M. tuberculosis* from two phage display libraries are described. This chapter describes the different binding properties of phage pools enriched from three different biopanning experiments, highlighting the importance of optimizing biopanning for each target. Subsequently, the phage pools resulting from the final biopanning protocolexperiment were assessed for enrichment using high-throughput sequencing (HTS) and validated by characterising their ability to bind the targeted *M. tuberculosis*. High-throughput sequencing (HTS) in addition to the traditional random clone picking method, were used to identify the highly enriched clones from the biopanning study.

In Chapter 3, the phage clones identified using both methods are assessed for enrichment and ability to bind mycobacteria. The highly enriched phage clone as identified by HTS was characterised for binding to its target using different species of mycobacteria. The displayed peptide was further synthesized. Furthurmore, the binding specificity and the nature of its mycobacteria ligand evaluated.

Chapter 4 describes the isolation of RNA aptamers with nanomolar affinities against the ESX-3 secreted protein EsxG using surface plasmon resonance.

Finally, Chapter 5 provides a general discussion and conclusion to the thesis.

CHAPTER 2

DEVELOPMENT OF PEPTIDE APTAMERS AGAINST SURFACE EXPOSED TARGETS ON M. TUBERCULOSIS, USING PHAGE DISPLAY TECHNOLOGY AND ILLUMINA SEQUENCING

Chapter Summary

Bacterial cell wall components have been previously used as infection biomarkers detectable by antibodies. However, the surface of the *Mycobacterium tuberculosi*s, the causative agent of tuberculosis, is mainly composed of glycolipids which are unlikely to be antigenic. This makes the probing of biomarkers on the surface of *M. tuberculosis* cell wall limited, using antibodies. Phage display is an approach that can find peptides that bind their targets with high affinity using the selection process called biopanning. The success of the biopanning experiment is dependent on various steps during the protocol. These include the immobilization of the target, the stringency in the removal of unbound phages through multiple washes and the strategy for elution of bound phages from the target. In this chapter, I investigated three biopanning protocols with varied strategies for immobilization of the target, stringency of the washing steps as well as elution methods. I demonstrated that the immobilization of *M. tuberculosis* intact bacilli on a surface treated for binding both hydrophilic and hydrophobic molecules performed better than both the untreated surface and in-solution biopanning against the *M. tuberculosis* bacilli. Furthermore, I found that the elution strategy using both low pH and sonication resulted in an enriched pool with higher binding signal to the targeted *M. tuberculosis* bacilli when compared to the conventional low pH elution strategy without sonication. I further confirmed the enrichment of the selection using high-throughput Illumina sequencing, showing the reduction in the diversity of the selected pools throughout the selection protocol. Taken together, this data supports the use of phage display technology to enrich for a population of peptides that bind mycobacteria, which can be further investigated as detection reagents for potential TB biomarkers.

1.0 Introduction

Pathogen-associated biomarkers, have advantages over current immunological biomarkers as they identify the presence of an organism that is independent of the state of the disease, such as active and latent TB. Thus far, the only available pathogen-associated tests that are used - mainly on sputum samples are smear microscopy, culture (Levy et al., 1989), and nucleic acid amplification tests (Davis et al., 2011, Ling et al., 2008). In the case of extrapulmonary TB, or in pediatric and immunocompromised patients, where individuals would have difficulty producing a sputum sample or the bacilliary burden is low, tests that probe for biomarkers that can be detected in samples other than sputum are critical. Currently, these include assays that detects LAM (Chan et al., 2000, Sada et al., 1992) in urine, the volatile organic compounds breath test (Fend et al., 2006, Phillips et al., 2007), and whole blood culture (Ravn et al., 2005, Wallis et al., 2001). However, all these tests have varying limitations which include low sensitivity, low specificity or poor cost-effectiveness (reviewed in detail in chapter 1). Therefore, it is critical that new biomarkers are identified to improve diagnosis of tuberculosis.

I hypothesised that numerous cell wall associated components are shed by the mycobacterium during infection. These might possibly be detected in patient samples such as sputum, serum and urine, if suitable detection reagents were available. However, it is well known that the surface of *M. tuberculosis* is mainly composed of glycolipids (Lopez-Marin, 2012) which are possibly non-antigenic. Antibodies, which are the conventional reagents used for biomarker probing or pull-down assays, are limited, because by definition they can only identify antigenic components. This is

possibly one of the reasons these potential non-antigenic biomarkers, have not yet been identified. As a result, we currently do not have reagents to probe for these potential biomarkers in patients samples. Thus, we employed phage display technology to identify peptides that can potentially bind to both antigenic and non-antigenic surface components of mycobacteria.

Panning of phage display libraries has successfully identified peptides that bind intact bacteria (Stratmann et al., 2002) and viruses (Chen et al., 1996). The technology involves the display of a random peptide sequence appended to a recombinant viral protein on the surface of a bacteriophage (Adda et al., 2002). The typical selection as previously detailed in chapter 1, involves exposure of the unselected library to the target, and removal of unbound phages. The bound phages are then eluted and amplified by infection of host bacteria under selective pressure.

The biopanning protocol can be optimized for every single target. There are invariably three main parts which can be varied. First, the immobilization strategies, that can either be surface panning or solution phase panning. Surface panning protocols are the most common, mainly because of their simplicity. Traditionally, the target is directly coated onto the plastic surface. One of the limitations of this protocol is the inaccessibility of the ligand binding site as a result of steric blocking and/or partial denaturing of the target along the surface (Biolabs, 2012). Moreover, since targets have variable surface properties, it might be necessary to experiment with differently treated surfaces to find the best surface for immobilising the specific target. Alternatively, the solution phase panning can be adopted, where the target is allowed to interact with the phage in

solution. This can be followed by affinity capture of the phage-target complexes to recover the bound phage (Cwirla et al., 1990) or by simply centrifugation of the phage-target complex (Stratmann et al., 2002). While this protocol falls short of the technical simplicity of surface immobilisation, it offers two main advantages. First, during solution panning, all the potential ligand binding sites on the target are accessible to the phage. Second, using this protocol minimizes the potential denaturing of the target that can be the result of immobilisation on a solid surface.

The second variable part of the biopanning protocol is the washing steps to remove target unrelated binders. Stringency in the washing step is usually influenced by increasing the amount of detergent in the washing buffer and/or raising the number of washing steps in the successive rounds, allowing for the removal of any target unrelated binders without compromising the binding of the target binding phages (Lunder et al., 2008).

The last variable part of the biopanning protocol is the elution strategy used to recover target binding peptides. In the case where a ligand of the target is known, competition elution is an option. However, non-specificelution using low pH buffers is commonly used (Koivunen et al., 1994, Koolpe et al., 2002, Lowman et al., 1991, Verhaert et al., 2002). Different non-specificelution protocols are reviewed by Lunder and colleagues (2008), and they demonstrated that the addition of sonication to low pH elution results in a single step isolation of high affinity peptides.

Traditionally, at the end of biopanning, random phage clones are selected and evaluated for binding to the intended target. While this method can identify enriched clones, it is not able to evaluate the enrichment levels of each individual clone at the end of selection, nor can it assess the success of the biopanning experiment by providing the global trend of enrichment. However, HTP sequencing has made possible the sequencing of millions of inserts allowing for a higher resolution of the selected pool of the displayed peptides (Dias-Neto et al., 2009, 't Hoen et al., 2011). The reduction in the sequence diversity throughout the selection rounds of biopanning together with the analysis of the enriched peptides at the end of the selection can be used to evaluate the success of the biopanning experiment.

This chapter describes the three biopanning protocols evaluated for the enrichment of phage displayed peptides that bind to intact *M. tuberculosis* (Δleu/Δpan). To evaluate the three panning protocols, we employed a library that displays random 7-mer peptides (CX_7C) at the tip of the pIII minor coat protein of the M13 phage. The displayed peptide is flanked by two cysteine residues, which are oxidized during phage assembly to a disulfide bond, resulting in a loop constrained peptide. Three rounds of panning against the targeted *M. tuberculosis* (Δleu/Δpan) were performed using all three protocols and the eluates from the final round of panning were evaluated in a binding assay against the target. Subsequently, biopanning protocol 3, which showed the best enrichment and specificity to the target after three rounds of biopanning was used to select for *M. tuberculosis* binding peptides using an additional phage library that displays random 12-mer peptides (X_{12}). Finally, HTP sequencing together with ELISA, which determined binding, were used to demonstrate the success of the selected biopanning protocol.

2.0 Materials and methods

2.1 Bacterial strains and growth conditions

The bacterial strains used in this study were *M. tuberculosis* H37Rv, *M. tuberculosis* ΔleucineD and ΔpanthothenateCD double auxotroph (Δleu/Δpan), *M. smegmatis* $mc^2$1551, *M. bovis* (BCG) and the *E. coli* ER2738 strain for phage amplification.

Mycobacteria cultures were grown on Middlebrook 7H9 media (Sigma, St. Louis, MO) supplemented with the Middlebrook oleic albumin dextrose catalase (OADC) (bovine albumin fraction V 5 g/l, dextrose 2 g/l, catalase 0.004 g/l, oleic acid 0.05 g/l, sodium chloride 0.85 g/l) (Sigma, St. Louis, MO), 0.2% glycerol and 0.05% Tween-80. *M. tuberculosis* (Δleu/Δpan) media was further supplemented with both leucine (50 µg/ml) and pantothenate (24 µg/ml) (Sampson et al., 2004). Mycobacteria cultures were grown at 37ºC with agitation. In experiments involving preserving the mycobacteria capsular layer, all mycobacteria cultures were grown in the absence of tween and agitation. For further use, all mycobacteria were centrifuged at 8000 rpm and resuspended in carbonate buffer (35 mM $NaHCO_3$, 15 mM Na_2CO_3 [pH 9.8]).

The *E. coli* ER2738 strain was grown in Luria-Bertani (LB) medium with tetracycline (20 µg/ml final concentration) for amplification and to titer phage. For the plating of the phage isopropyl β-D-thiogalactopyranoside (IPTG; 1mM final concentration) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal; 60mM final concentration) were added to LB agar plates before plating. All bacterial cultures were grown at 37°C with agitation.

2.2 Immobilization of the target mycobacteria for solid phase panning

Mycobacteria suspensions were resuspended in carbonate buffer (35 mM $NaHCO_3$, 15 mM Na_2CO_3 [pH 9.8]) and adjusted to an optical density of 1.0 at 660 nm, corresponding to approximately 10^8 bacteria per ml (Stratmann et al., 2002). Polystyrene petri dish wells (Sigma Aldrich, MO, USA), used for biopanning 1 or MaxiSorp™ surface microtiter plate wells (Nunc, Roskilde, Denmark) used for biopanning 3, were filled with 200 µl of bacterial suspension and incubated overnight at 4ºC. Wells were blocked for 1 hour at 4°C with blocking buffer (0.1M $NaCO_3$ [pH 8.6], 5mg/ml bovine serum albumin (BSA)) before incubation with the phage library.

2.3 Selection of phage displayed peptides-biopanning

Three biopanning protocols were evaluated for the selection of *M. tuberculosis* (Δleu/Δpan) binding peptides. A brief description of the three biopanning protocols is given in Table 2.1.

In Biopanning Protocol 1, for the first round of panning, 10 µl of the library (~2×10^{11} phages) was diluted to 100 µl with Tris-buffered saline supplemented with 0.1% Tween-20 (TBST) (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.1% tween) and incubated in bacteria-coated polystyrene petri dish wells for 1 h at room temperature with gentle agitation. Nonbinding phages were then discarded and the wells were washed 10 times with 200 µl of TBST. Finally, the elution of the bound phage was carried out using 100 µl the low-pH elution buffer (0.2 M Glycine-HCl [pH 2.2]). The eluent was then neutralised with 15 µl of 1 M Tris-HCl [pH 9] and amplified by infecting its bacterial host *E.coli* ER 2738. In the following rounds of biopanning an average of 2×10^{11}

plaque-forming units (PFU) of the amplified phage from the preceding selection round was used. After the three rounds of biopanning, the phage pools from the final round were assessed for binding to the targeted *M. tuberculosis* (Δleu/Δpan) and a related non-pathogenic mycobacteria species *M. smegmatis* by ELISA.

Biopanning Protocol 2 was a solution-based panning. Similar to biopanning protocol 1, in the first round 10 µl of the library (~2×10^{11}phages) was diluted to 100 µl with TBST. The phage suspension was subsequently incubated in a micro centrifuge tube with 100 µl of *M. tuberculosis* (Δleu/Δpan) culture that was suspended in PBS (0.01 M phosphate buffer, 0.0027 M potassium chloride and 0.137 M sodium chloride [pH 7.4]). The mycobateria and phage complex were collected by centrifugation (5 min, room temperature, 16,000 × g) and nonbinding phages discarded. Mycobacteria were washed ten times with 200 µl TBST, by centrifugation (5 min, room temperature, 16,000 × g). The elution of the bound phage was performed by combining three elution fractions. The first fraction was carried out using 100 µl of the low-pH elution buffer (0.2 M Glycine-HCl [pH 2.2]) and immediately neutralised with 15 µl of 1 M Tris-HCl [pH 9]. This was followed by a 200 µl TBST wash and an additionalelution with 100 µl of the low-pH elution buffer. All three elution fractions were combined after each round, amplified, tittered and 2×10^{11} PFU were used as input phage for the following round. After the three rounds of biopanning, the phage pools from the final round were assessed for binding to the targeted *M. tuberculosis* (Δleu/Δpan) strain and a related non-pathogenic mycobacteria species *M. smegmatis* by ELISA.

Biopanning Protocol 3 was a solid surface-based panning. In the first round of panning, 10 µl of the library (~2×10^{11} phages) was diluted to 100 µl with phosphate buffered saline supplemented with 0.1% Tween-20 (PBST) (0.01 M phosphate buffer, 0.0027 M potassium chloride and 0.137 M sodium chloride, pH 7.4, 0.1% Tween) and incubated in bacterial-coated MaxiSorp™ surface microtiter plate wells for 1 h at room temperature with gentle agitation. Nonbinding phages were then discarded and the washes were performed as previously described(Lunder et al., 2008). In brief, the wells were washed 25 times with 200 µl of PBST followed by four washes with 200 µl of the low pH elution buffer (0.2 M Glycine-HCl, pH 2.2). Finally, the elution of the bound phage was carried out using 100 µl of the low-pH elution buffer and sonicated in a sonicator water bath (50 kHz) for 10 minutes. The eluent was then neutralised with 15µl of 1 M Tris-HCl [pH 9]. In the following rounds of biopanning an average of 2×10^{11} PFU were used. Five rounds of biopanning were performed. The first three rounds were targeted against *M. tuberculosis* (Δleu/Δpan) and the phage eluates from round 3 were assessed for binding to the targeted *M. tuberculosis* (Δleu/Δpan) and compared to biopanning 1 and 2 phage eluates. Subsequently, round 3 was followed by a subtraction round against *M. smegmatis*. The fifth round of positive selection was again against the targeted *M. tuberculosis* (Δleu/Δpan). After the final round of biopanning, the phage eluates from round 5 were assessed for binding to the targeted *M. tuberculosis* (Δleu/Δpan) and two other mycobacteria species, namely *M. smegmatis* and BCG. In addition to binding analysis, the unselected library, round 3, round 4 and round 5 phage eluates were also subjected to HTP Illumina sequencing.

2.4 Phage Amplification

Phage eluates were amplified according to the suppliers' instructions (New England Biolabs, MA, USA). In brief, the phage eluates were used to infect an early log phase growing culture of *E. coli* ER2738 host cells. After 4.5 h of growth at 37°C, bacteria were removed by centrifugation (10 min, at 4ºC, 12,000 × g) and phages in the supernatant were precipitated by adding one-sixth volume of 20% polyethylene glycol-8000 and 2.5 M NaCl overnight at 4°C. The precipitate was resuspended in 100 µl of PBS, and amplified eluates were titered according to the suppliers (New England Biolabs, MA, USA) instruction to determine phage concentration.

2.5 Phage ELISA

MaxiSorp™ surface microtiter plate wells (Nunc, Roskilde, Denmark) were coated with 100 µl of mycobacteria suspension, with an optical density of 1.0 at 660 nm in carbonate buffer (35 mM $NaHCO_3$, 15 mM Na_2CO_3 [pH 9.8]), and incubated overnight at 4ºC. Plates were blocked overnight at 4°C with blocking buffer (0.1M $NaCO_3$ [pH 8.6], 5mg/ml bovine serum albumin (BSA)). A separate set of wells were blocked with blocking buffer without previous mycobacteria immobilization as negative controls (no target control). The required concentration of each amplified phage eluate was resuspended in PBS to the final volume of 100 µl and transferred to the coated wells. The bacteria and phage eluates were incubated for 1.5 h at room temperature. Wells were then washed six times with 200 µl of PBST. Horseradish peroxidase (HRP)-labelled mouse anti-M13 monoclonal antibody (Amersham Biosciences, Amersham, UK) was diluted in PBS (1:5,000). Two hundred microliters of the antibody was added per well and incubated for 1 h at room temperature. This was followed by washing of the wells six times with PBST. One hundred and fifty microliters of substrate solution

(2,2'-azino-di-[3-ethylbenzthiazoline sulfonate] diammonium salt) (Thermo Scientific, MA, USA) was added and incubated for 30 min at 37°C. The reaction was stopped with 100 µl of 1% SDS. Absorbance was determined using a Multiskan GO microplate reader (Thermo Scientific, MA, USA) at the wavelength of 405 nm.

2.6 Phage DNA preparation

The amplification of the displayed peptides was performed with PCR using primers spanning the variable region in the gp3 phage coat protein (Appendix table 1). Two sequencing reads were performed during Illumina sequencing. The forward primer (Illumina_R1) contained the sequencing region for Illumina read 1 sequencing (Oligonucleotide sequences © 2006-2008 Illumina, Inc.) and was the same for all the phage DNA amplifications. The reverse primers (Illumina_F1-F5) contained a homology sequence required for annealing to the Illumina sequencing flow cell, a sequencing region for Illumina read 2 sequencing and a six-nucleotide barcode to enable multiplexing (Oligonucleotide sequences © 2006-2008 Illumina, Inc.). This enabled amplicons to be directly sequenced on an Illumina Genome Analyzer II as previously described (Griffin et al., 2011). The PCR reaction mix consisted of 0.5 µM of each primer and 5U of the GoTaq® DNA polymerase mix (Promega, WI, USA) in a 100 µl final volume. Whole phage PCR (denaturation at 95°C for 1 minute, annealing at 55°C for 2 minutes and extension at 72°C for 1minute and 30 seconds) was performed as previously described (Kingsbury and Junghans, 1995). Cycles varied from 10 to 25, and the number of reaction tubes varied from 2 to 5, according to the amount of amplicon available for each sample. The amplicons were analysed on a 2% agarose gel. To validate the incorporation of the Illumina adaptor sequences, the PCR amplicons were cloned into a pGEMT-Easy cloning kit (Promega, WI, USA) according to the

manufacturer's instruction. Four to five clones were randomly selected from each pool and the insert sequenced (Inqaba Biotechnology, Pretoria, RSA).

2.7 Phage DNA high throughput sequencing using Illumina technology

DNA Amplicons resulting from the PCR (0.45–0.85 μg) were sent to the department of computer science (Texas A&M University, Texas, USA) for DNA sequencing and analysis. The DNA was sequenced using the Illumina Genome Analyzer II. The nucleotide sequence of the amplified region of the phage pIII protein (gp3) gene was reconstructed by aligning and combining the two paired-end reads. The 36 bp (X_{12} library) and 21 bp (X_7 library) variable region was extracted by trimming off constant bases and was translated into amino acid sequences of length 12 using the Illumina GA Pipeline software, which were then clustered for statistical analysis.

3.0 Results

3.1 Selection of phage displayed peptides that bind to intact *M. tuberculosis*

In order to identify phage displayed peptides that bind to intact *M. tuberculosis*, a 12-mer (X_{12}) and a constrained 7-mer (CX_7C) phage library were used for biopanning on immobilised *M. tuberculosis* (Δleu/Δpan). This non-pathogenic auxotrophic *M. tuberculosis* (Δleu/Δpan) strain was used as a model organism because it can be grownoutside a biosafety level 3 laboratory but will have a cell wall equivalent to pathogenic *M.tuberculosis*. The basic protocol for biopanning is described in chapter 1 (section 1.5.4). I initially tested three biopanning protocols using the CX_7C phage library (Table 2.1). The protocols tested varied in the strategies for immobilisation of

the target, stringency of the washing step and the elution of bound phages. Three positive rounds of selection were performed and the resulting phage eluates were evaluated for binding to the target, *M. tuberculosis* (Δleu/Δpan) and a related non-pathogenic mycobacteria species *M. smegmatis*using an ELISA assay.

3.1.1 Biopanning Protocol 1

The biopanning protocol 1, was a solid surface panning experiment which was performed according to the phage library manufactures instructions (New England Biolabs, MA, USA). In brief, the *M. tuberculosis* (Δleu/Δpan) culture was immobilized directly on a plastic surface of a sterile polystyrene petri dish. The conditions for the binding of the library to the target and the washes are detailed in section 2.3. The elution strategy for biopanning protocol 1 involved a nonspecific disruption of the binding interaction between the phages and the target using a low pH buffer, 0.2 M Glycine-HCl buffer pH 2.2.

Varying amounts of phages (2×10^5 PFU – 2×10^{11} PFU) from the final round of selection, round 3, were evaluated for binding to the targeted *M. tuberculosis* (Δleu/Δpan) and a related species of mycobacteria, *M. smegmatis*. There was no significant difference between the binding signals of the selected pool of phages to both mycobacteria species when compared to the control wells which had no prior immobilisation of target (Figure 2. 1A). This data is most likely indicative of an unsuccessful immobilization strategy employed during this protocol which was not suitable for mycobacteria resulting in the failure to immobilise mycobacteria during biopanning. To overcome this limitation, a solution phase biopanning protocol was

evaluated in protocol 2. In addition, the elution strategy used in this protocol was evaluated for efficiency before proceeding to protocol 2. To evaluate the elution protocol, three elution fractions were collected (described in 2.3, biopanning protocol 2) and each fraction was titered to monitor the presence of phage in each fraction. Phage was present in all three fractions tested for elution (Appendix figure 1). This resulted in the adoption of a modified elution strategy for the biopanning protocol 2 where all three elution fractions were combined before phage amplification and used in the subsequent selection round.

3.1.2 Biopanning Protocol 2

To address the possibility that mycobacteria might not be efficiently immobilized on the polystyrene surface of the petri dish, a solution-phase panning protocol was used. In brief, the phage library was allowed to react with the targeted mycobacteria in solution. The washing steps were performed under similar conditions to those used in protocol 1, with the exception of a low speed centrifugation step to recover the mycobacteria bacilli with the bound phages as described in section 2.3.

After three rounds of panning, a similar ELISA was used to access the level of enrichment during the selection. There was a significant difference ($p < 0.01$) between the binding signal of the selected pool of phages to both mycobacteria species when compared to the control wells which had no prior immobilisation of target (Figure 2. 1B). This result was the same when 2×10^8 PFU and 2×10^{11} PFU were used as input phages for the ELISA. However, there was no significant difference in the binding

signal of the phage pools to mycobacteria when compared to the control wells when 2×10^5 PFU were used as input for the assay.

When comparing the binding of the selected phage population between the two mycobacteria species tested, there was no significant difference to the binding of the selected phage pool to the targeted *M. tuberculosis* (Δleu/Δpan) when compared to *M. smegmatis* (Figure 2. 1B). Taken together, this data shows that the solution-phase biopanning protocol performed better in enriching for a phage population that binds to mycobacteria. However, this protocol enriched for a population that did not discriminate between the targeted *M. tuberculosis* (Δleu/Δpan) and the related *M. smegmatis*, indicating that the elution protocol might have indiscriminately allowed for enrichment of phage clones that were common to the genus of mycobacteria. To evaluate whether we couldperform a panning experiment that can enrich for *M. tuberculosis* species specific phage displayed peptides, we performed a third biopanning protocol that had been previously demonstrated (Lunder et al., 2008) to enrich for high affinity clones.

3.1.3 Biopanning Protocol 3:

The biopanning protocol 3 was a solid surface panning experiment similar to protocol 1, with the exception that the *M. tuberculosis* (Δleu/Δpan) culture was immobilized directly on a pre-treated plastic surface with a high affinity for molecules with mixed hydrophilic/hydrophobic domains, namely the MaxiSorp™ plate (Nunc, Denmark). This immobilization protocol was performed mainly to improve the immobilization efficiency of mycobacteria on the solid surface and was previously demonstrated to be

effective when immobilizing mycobacteria by Stratmann and colleagues (Stratmann et al., 2002). In addition, the washing protocol was made more stringent by increasing the number of washes using a similar tween supplemented buffer as forprotocol 1 and 2. Four additional washes were performed using a low pH buffer followed by sonication as previously described by Lunder and colleagues (2008). Three positive rounds of panning were performed against the targeted *M. tuberculosis* (Δleu/Δpan). When evaluating enrichment of target binding phages using a similar ELISA, our data showed that this protocol yielded a phage population that had a significantly higher binding signal when binding to *M. tuberculosis* (Δleu/Δpan) when compared to *M. smegmatis* or the control wells (Figure 2. 1C). This observation was true when 2×10^{10} PFU were used as input, while there was no significant binding difference when 2×10^{8} PFU or less was used. This protocol was further used for a biopanning experiments starting with the X_{12} library and similarresults were obtained (Figure 2. 1D). Taken together, our data indicated that the third biopanning protocol was more efficient in selecting peptides specific to *M. tuberculosis* (Δleu/Δpan).

Additional rounds of biopanning were performed to further enrich for *M. tuberculosis* (Δleu/Δpan) specific clones. In brief, to remove peptides binding to cell wall components common to the mycobacteria genus, a subtraction round (round 4) was performed against *M. smegmatis*. A final positive panning round (round 5) was performed to further enrich for peptides that were specific to *M. tuberculosis* (Δleu/Δpan). A schematic of the full biopanning protocol 3 is illustrated in Figure 2. 2.

Table 2. 1: Overview of Biopanning Protocols

Protocol	Biopanning 1	Biopanning 2	Biopanning 3
Target immobilisation	Untreated solid surface –polystyrene petri dish	Solution phase - microcentrifuge tube	Treated solid surface with high affinity to molecules with mixed hydrophilic/hydrophobic domains– MaxiSorp™ plate
Washing step	10 X TBST (0.1% tween)	10 X TBST (0.1% tween)	25 X PBST (0.1% tween) 4X low pH glycine-HCl pH 2.2 buffer
Elution strategy	Low pH	Low pH	Low pH and sonication
Elution step	1 X Glycine-HCl pH 2.2	Combined: 1X Glycine-HCl ph2.2 1X PBS wash 1X glycine-HCl ph2.2	1X Glycine-HCl pH 2.2 And sonication for 10 min
Number of selection rounds	3	3	3 *
The selection was considered successful when the average ELISA value of the selected pools binding to *M. tuberculosis* (Δleu/Δpan) significantly exceeded the binding to *M. smegmatis* after three rounds of selection.			

*Two additional panning rounds were performed on this protocol as detailed in figure 2.2

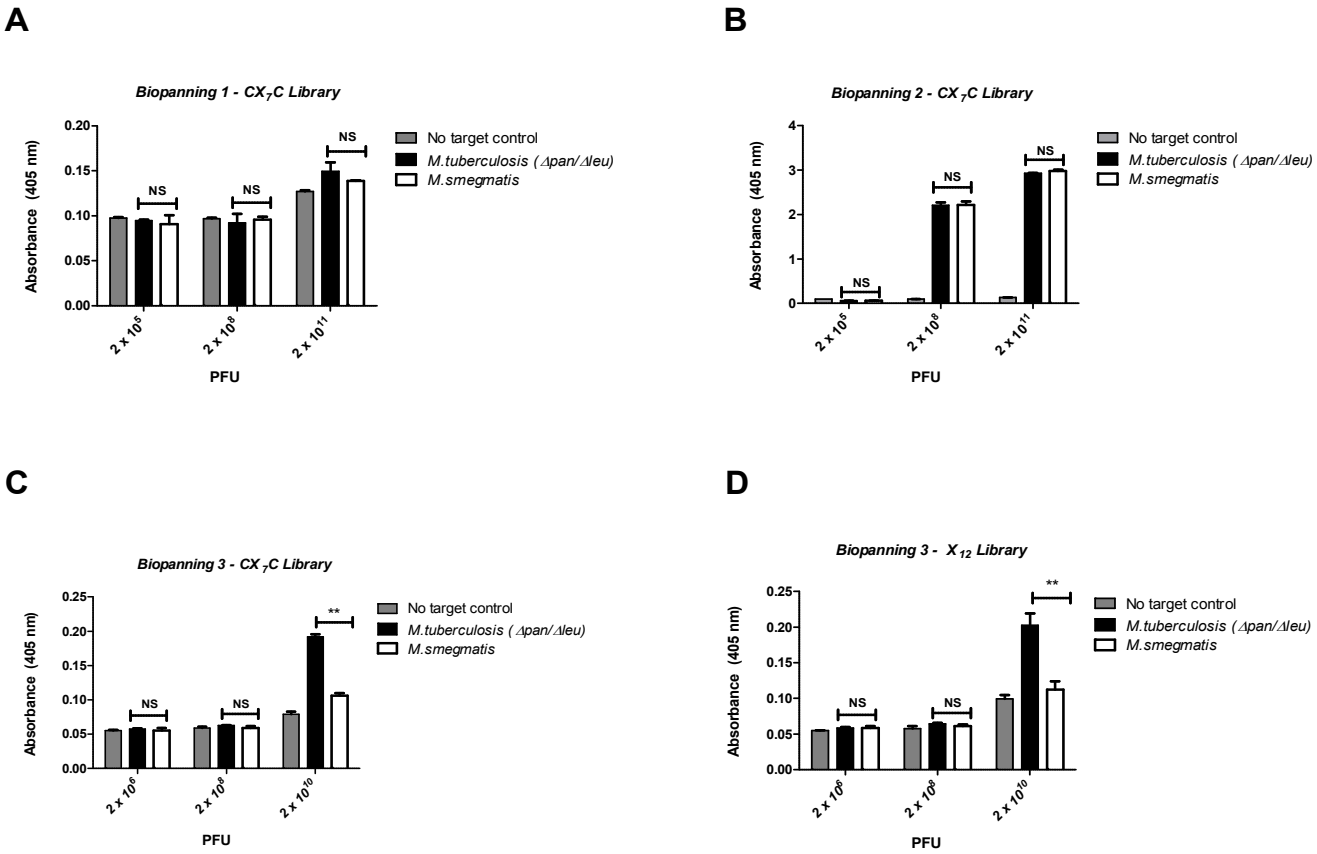


Figure 2. 1: Binding comparison to mycobacteria of phages selected by different biopanning protocols using a whole cell ELISA assay

Three biopanning experiments were performed against *M. tuberculosis* (Δleu/Δpan) and evaluated after three rounds of panning. Binding analysis to *M. tuberculosis* (Δleu/Δpan) and *M. smegmatis* of the final phage populations were performed and compared between the three biopanning protocols. A two-tailed, unpaired t-test was used to analyse significance (NS = no significance, *$p<0.05$, **$p<0.01$). Error bars represent standard deviations of the arithmetic means of optical densities at 405 nm performed in triplicates.

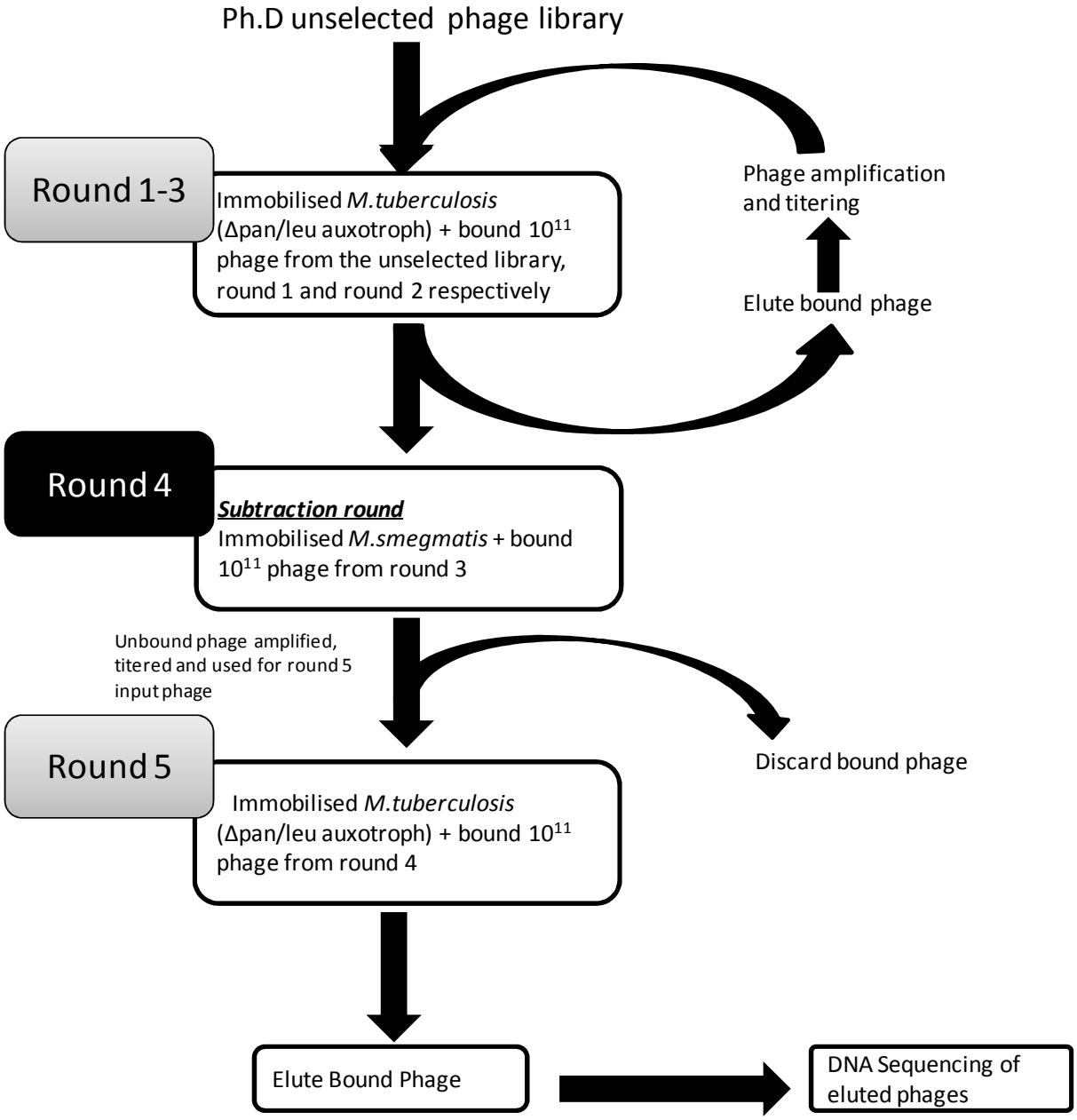


Figure 2. 2: Schematic oftreated solid phase biopanning strategy (protocol3)

Three rounds of biopanning were performed against *M. tuberculosis* (Δleu/Δpan). After the third round of panning, eluted phages were subjected to a subtraction round against *M. smegmatis* (round 4). Round 5 was performed against the target *M. tuberculosis* (Δleu/Δpan).

3.2 Characterisation of the enrichment process from the biopanning protocol 3 using high throughput Illumina sequencing

In order to evaluate the trend of enrichment during biopanning, we performed HTP sequencing on the naive phage library before selection and after three, four and five rounds of biopanning. The selected phage eluates were amplified by PCR**.** The forward primer (Illumina_F1) for the PCR was the same for all phage amplification. The primer sequence was adapted from the Illumina multiplexing PCR primer 1.01 (Illumina, 2006-2008) with an additional 33 bases homologous to gp3 region of the phage genome upstream of the random peptide insert. The reverse primer (Illumina_R1-R8) sequences were adapted from the Illumina multiplexing PCR primer 2.01 (Illumina, 2006-2008) with an additional 18 bases homologous to the conserved phage sequence and a six base multiplex index tag which was unique for each phage pool. All primer sequences are listed in the appendix, supplementary table 1.

The PCR conditions are detailed in section 2.6. Resulting amplicons were quantified and the expected size of approx. 200 bp was confirmed on a 2% agarose gel (Figure 2. 3A). To validate the incorporation of the sequences needed for multiplexed Illumina sequencing, the PCR amplicons were cloned into a pGEMT-Easy vector. Four to five clones were randomly sequenced (Inqaba Biotechnology, Pretoria, RSA) from each pool and their sequences were aligned with the corresponding primer pair and the multiplex index tag. A representative alignment for the PCR product using the multiplex index tag 1 is shown in Figure 2. 3B.

Illumina sequencing was then used to determine the unique sequences of the amplified phage inserts from each selection round. The nucleotide sequences of the amplified region of the gp3 gene, was reconstructed by aligning and combining the two paired-end reads (Texas A&M University, Texas, USA). The 36 bp variable region was extracted by trimming off constant bases and was translated into amino acid sequences using the Illumina GA Pipeline software, which were then clustered for analysis (Texas A&M University, Texas, USA).

3.2.5 Characterization of CX_7C biopanning enrichment

Approximately 1.5 million sequence reads from each selection round were acquired during HTP sequencing. This represented $1.36x10^6$ unique peptides from the unselected library (Figure 2. 4A). While this fell short of the theoretical complexity of the library, $1.23x10^9$ heptapeptides, it represented sufficient depth to measure the quantitative enrichment of relevant peptides. To confirm successful enrichment during selection, we characterised the reduction in diversity of the phage eluates in the consecutive rounds of panning. There was over two log reduction in the overall diversity of the phage eluates by the end of biopanning, in round 5 (Figure 2. 4A). To illustrate, the number of unique sequences decreased from 1,361,688 in the unselected library to 5665 after the final round of panning. This is indicative of a successful selection for a subpopulation of phage clones during the biopanning experiment.

3.2.1 Characterization of X_{12} biopanning enrichment

A similar number of sequence reads was obtained from each selection round of the X_{12} biopanning when compared to the CX_7C biopanning experiment. This represented $0.73x10^6$ unique peptides in the unselected library (Figure 2. 4B). While the theoretical complexity of this library is $4.1 x10^{15}$ dodecapeptides, the X_{12} phage library used for this study had an initial complexity in the order of 10^9, representing less than one millionth of the possible dodecapeptides. Taking this into consideration, the sequencing depth achieved by the HTP experiment represented sufficient depth to measure quantitative enrichment of relevant peptides from this library. Importantly, there was over one log reduction in the sequence diversity after three rounds of positive selection, while additional rounds of selection did not further reduce the diversity of the enriched pool (Figure 2. 4B). A similar trend was also evident for the CX_7C selection. This suggests that for both selections, additional rounds of panning, round 4 and 5, did not have a dramatic effect on enrichment.

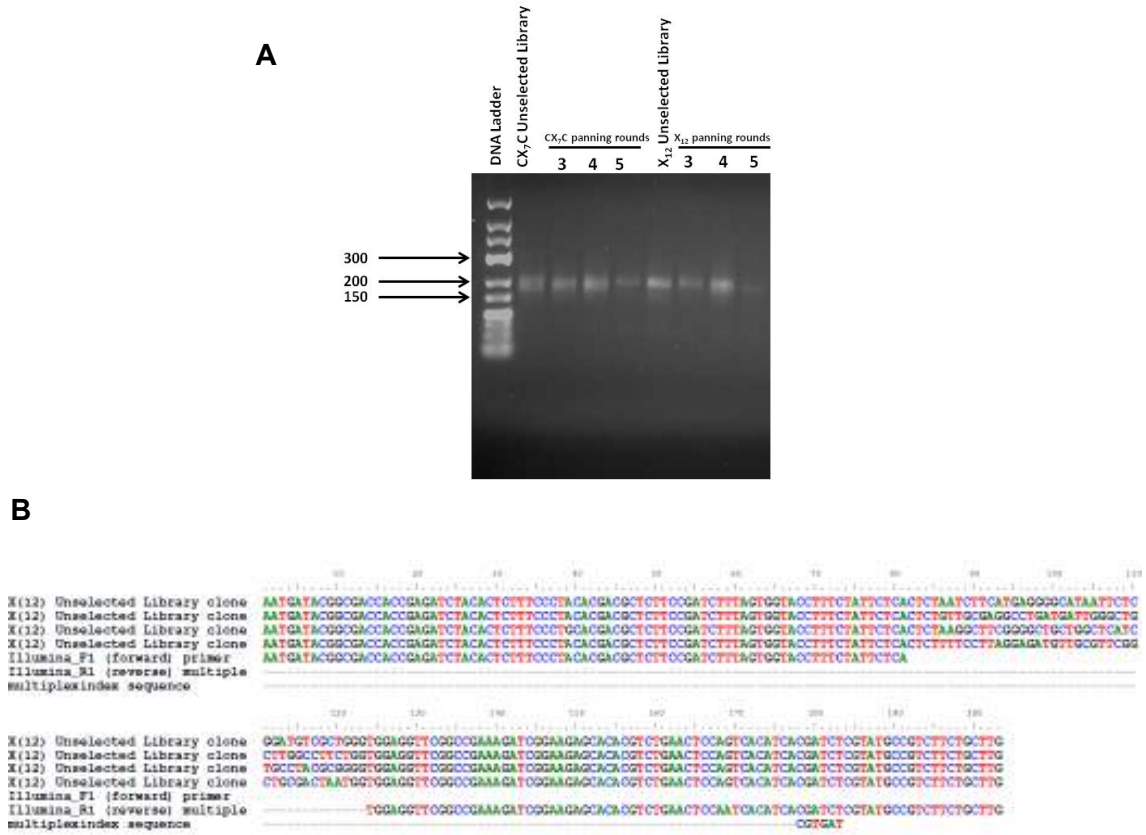


Figure 2. 3: PCR amplification of phage inserts from the CX_7C and X_{12} library pools.

(A) PCR products from phage library insertsamplified using primers containing HTP Illumina adaptor sequences and multiplexing tags and visually evaluated on a 2% agarose gel. (B) DNA alignments of sequences obtained from four clones picked from the X_{12} library and subjected to amplicon sequencing using primer sets designed for HTP sequencing. Sequence alignment of the forward and reverse primer sequences confirms the presence of the Illumina adaptor sequences, while the alignment of the unique multiplex tag was also confirmed for each round of selection.

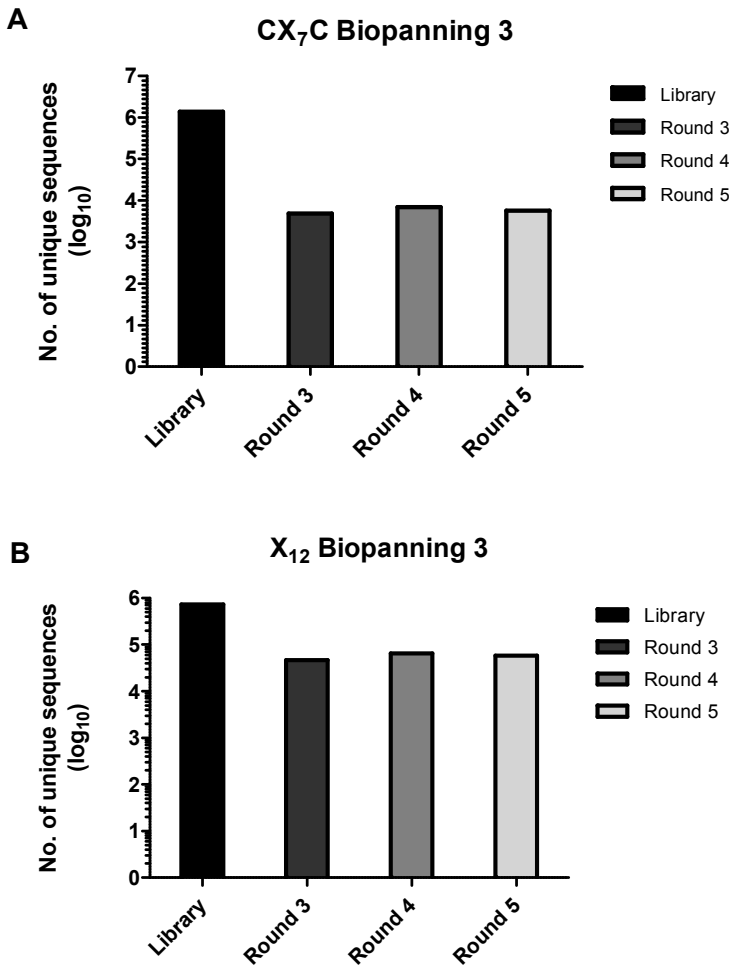


Figure 2. 4: Sequence enrichment profiles

Sequence enrichment profiles after high-throughput sequencing of the phage displayed libraries using Illumina technology. Histograms of the number of unique peptides observed in the different rounds of biopanning from the (A) CX_7C library selection and the (B) X_{12} library selection.

3.3 Characterisation of the binding of the phage populations enriched during biopanning protocol 3

To evaluate the effectiveness of the two additional rounds of biopanning, binding signals of round 3 and round 5 phage eluates were compared to each other using an ELISA assay. For the CX_7C biopanning, there was a significantly ($p < 0.01$) higher

binding signal of the round 5 phage eluates to *M. tuberculosis* (Δleu/Δpan) compared to the phage pool from round 3 (Figure 2. 5A). This indicates that the two additional rounds of biopanning resulted in the enrichment of *M. tuberculosis* (Δleu/Δpan) binding peptides. Similar results were observed for the X_{12} biopanning experiment, showing a significant ($p < 0.05$) increase in binding signal to *M. tuberculosis* (Δleu/Δpan) of phage eluates from round 5 when compared to those from round 3 (Figure 2. 5B).

To assess the effectiveness of the subtraction round, round 4 against *M. smegmatis*, binding signals of the phage eluates to *M. smegmatis* were assessed before and after the subtraction round. Phage eluates from both libraries showed a significantly ($p < 0.01$) higher binding signal to *M. smegmatis* after the subtraction round (Figure 2. 5A-B). This data indicates that the subtraction round was not successful in eliminating *M. smegmatis* binding phages. Alternatively, it is possible that the majority of the phage clones in round 5 bind to a common surface exposed component between the members of the genus mycobacteria.

To verify possible promiscuous binding of the selected phage populations to members of the *Mycobacterium* genus, we tested whether a similar pattern of increased binding was observed with BCG, which is another related species of *M. tuberculosis* (Δleu/Δpan). Indeed, for both libraries, the phage eluates from round 5 showed a significant increase ($p < 0.01$) in binding to BCG when compared to round 3 phage eluates (Figure 2. 5A-B). Taken together, this data suggests that thetwo additional rounds during biopanning protocol 3 enriched for a population of

phage clones that potentially bind promiscuously to the surface of bacteria belonging to the genus of mycobacteria.

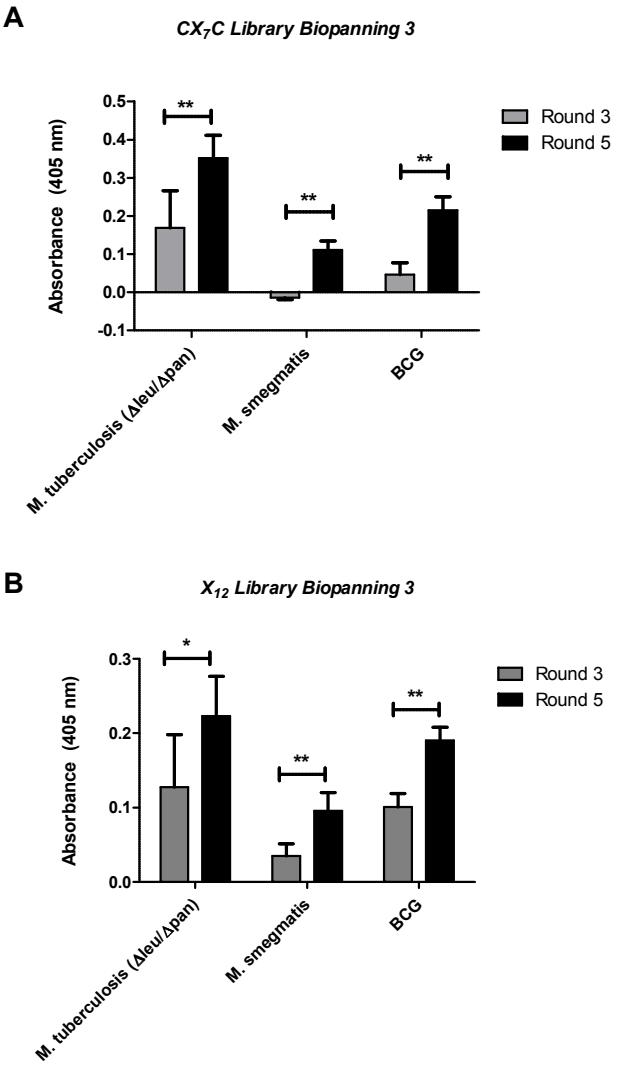


Figure 2. 5: Comparison between the binding of the phage eluates before and after the subtraction round using ELISA

Binding analysis to *M. tuberculosis* (Δleu/Δpan), *M. smegmatis* and BCG, of the round 3 and round 5 phage eluates from the (A) CX_7C and (B) X_{12} biopanning experiments. A two-tailed, unpaired t-test was used to analyse significance (*$p<0.05$, **$p<0.01$). Error bars represent standard deviations of the arithmetic means of optical densities at 405 nm measuredin triplicates, which were normalised to wells without previous mycobacteria immobilization.

3.4 Binding of the enriched phage eluates to mycobacteria cultured under different conditions

A previous finding showed that mycobacteria cultured in the absence of detergent and agitation have a thick outermost capsule-like layer, while those grown under the traditional laboratory conditions in the presence of both detergent and agitation have this layer partially or completely removed (Sani et al., 2010). To determinewhether the selected phage population binding is affected by the presence of the capsular layer in mycobacteria, the final pool of phages wastested for itsability to bind mycobacteria when it is grown under these two different culturing conditions. The phage eluate from round five of both the CX_7C and X_{12} biopanning experiments showed no significant difference in binding to *M. tuberculosis* (Δleu/Δpan) when grown under both conditions (Figure 2. 6A-B). However, for both libraries, there was a significantly higher binding signal of the final phage pool to *M. smegmatis* ($p < 0.01$) and BCG ($p < 0.05$) when grown in the absence of detergent and agitation as compared to the traditional laboratory culturing conditions that include detergent and agitation (Figure 2. 6A-B). Taken together, these results show that the selected population of phages is able to bind mycobacteria grown under both conditions. However, for both libraries the binding of the final phage eluates to *M. smegmatis* and BCG is enhanced by the presence of the capsule-like layer, while the binding to *M. tuberculosis* (Δleu/Δpan) is not affected by the presence of the capsule-like layer.

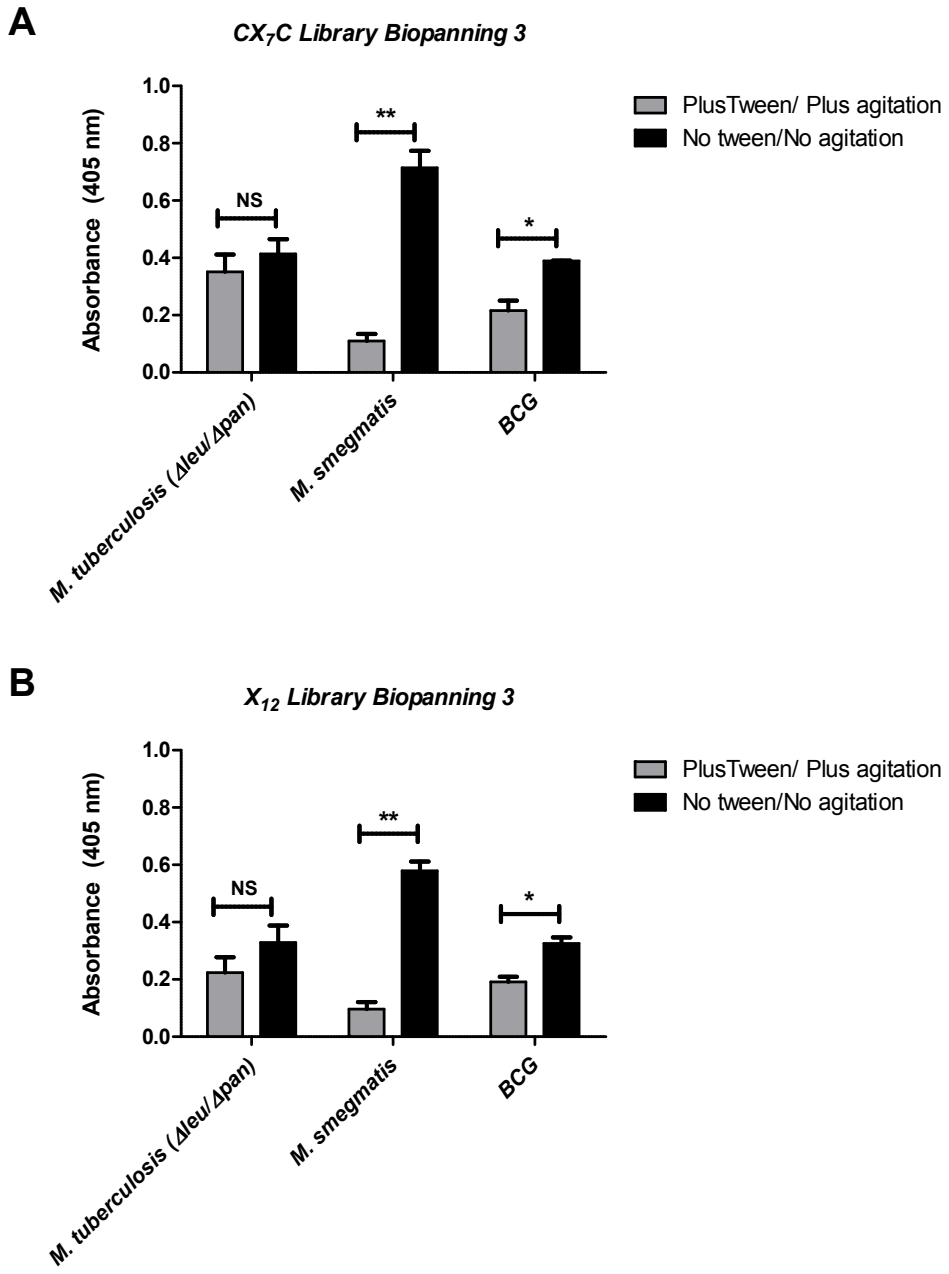


Figure 2. 6: Binding of the phage eluates from round 5 to mycobacteria grown under different culturing conditions using ELISA

Phage eluates from the final round of selection of the (A) CX_7C and (B) X_{12} biopanning experiments were amplified and 5×10^{11} PFU were used for a plate binding assay with *M. tuberculosis* (Δleu/Δpan), *M. smegmatis* and BCG as solid-phase antigens. A two-tailed, unpaired t-test was used to analyse significance (NS = no significance, **$p<0.01$, *$p<0.05$). Error bars represent standard deviations of the arithmetic means of optical densities at 405 nm performed in triplicates, which were normalised to wells without previous mycobacteria immobilization.

4.0 Discussion

Phage display technology has made it possible to search for peptides that bind to various protein (Koivunen et al., 1994, Pasqualini et al., 1995) and non protein targets, like silica (Naik et al., 2002) and apatite-based material (Segvich et al., 2009). In this study, I applied phage display technology with the aim of identifyingpeptide ligands that bind to *M. tuberculosis*. I evaluated three biopanning protocols varying mainly in immobilisation of the target, stringency in the washing steps of unbound phages as well as elution of the bound phage strategies. The data showed that immobilisation of mycobacteria on the Maxisorp™ surface, which has a high affinity to molecules with mixed hydrophilic and hydrophobic domains, performed better when compared to untreated surfaces. While the solution based panning yielded the highest binding signal of the phage eluates, the final phage population failed to discriminate binding between *M. tuberculosis* and *M. smegmatis* after three rounds of biopanning. This is likely due to the extensive elution method adopted for this panning protocol, which combined three elution fractions of the bound phage. While this elution method achieved the enrichment of mycobacteria binding phage clones, these phage clones were non-discriminating, possibly binding to surface components that are common among mycobacteria species. Taken together, thefindings indicate that biopanning protocol 3 was the best suited protocol for selecting phage displayed peptides against mycobacteria.

Traditionally, the success of a biopanning experiment is confirmed by evaluating the binding affinity of randomly picked clones to the target. While this method

usually leads to the identification of phage displayed peptides that bind to their target, it falls short in describing the global enrichment trend during biopanning as well as identifying the most enriched clones which are likely to be peptides with the best binding affinities to the target. HTP sequencing has made possible the sequencing of millions of inserts allowing for a higher resolution of the selected pool of the displayed peptides. In this study we used both a binding assay and HTP sequencing to validate the success of the selected panning protocol. The higher resolution of the selection pools enabled us to demonstrate the reduction in the number of unique peptides during biopanning, which is indicative of the enrichment of a subset of phages. The HTP sequencing results correlated with the binding assay of the selected phage population to the targeted *M. tuberculosis* (Δleu/Δpan), which showed an increased binding of the phage pools at the end of selection when compared to the phage pools at round 3. Surprisingly, the additional subtraction round against *M. smegmatis* did not result in the exclusion of *M. smegmatis* binding phage clones. As a result, theenriched phage population is likely composed of phage clones that bind to surface molecules conserved across the three mycobacterium species. This data is indicative of the invariable nature of the mycomembrane among different mycobateria species (Brennan and Crick, 2007).

Recently, Sani and colleagues (2010) demonstrated the presence of the labile capsular layer that covers the outer membrane of both pathogenic and non pathogenic mycobacteria. They further showed that the capsular layer is influenced by the culturing conditions used in the laboratory which involve detergent and agitation, resulting in the shedding of the capsule into the medium. The capsular

layer is known to be composed of polysaccharides, proteins and small amounts of lipids (Daffe and Etienne, 1999, Lemassu et al., 1996, Ortalo-Magne et al., 1995). Sani and colleagues (2010) showed that the detection of these components, namely the major capsular polysachharide (α-glucan), the glycolipid phosphatidylinositol mannosides (PIMs) as well as the EspE protein associated with the capsule, was weaker when mycobacteria is grown with agitation. Since our biopanning experiments were performed on *M. tuberculosis* (Δleu/Δpan) cultures grown under the conventional laboratory conditions, we investigated whether the presence of the capsule affected the binding of the enriched phage population. We had hypothesised that, if the phage eluates has a common binding partner amongst the different mycobacteria species, the binding trend will be similar between species. Surprisingly, binding of the phage to *M. tuberculosis* (Δleu/Δpan) was not affected by the presence of the capsule while binding to both *M. smegmatis* and BCG was enhanced. This result possibly indicates a different binding partner on the cell surface of *M. tuberculosis* (Δleu/Δpan) that is not affected by the presence of the capsular layer. Alternatively, it is possible that the non pathogenic *M. tuberculosis* (Δleu/Δpan) strain is not able to form a capsular layer under the tested conditions. Taken together, these results suggest that the enhanced binding of the phage pool to *M. smegmatis* and BCG when the capsular layer is present is likely a result of the increased presence of either the capsule associated glycans, glycolipids or proteins.

In conclusion, we demonstrated the selection of a binding population of phage clones against *M. tuberculosis* (Δleu/Δpan) using a Maxisorp™ surface based biopanning protocol with the modified washing and elution strategy. We further demonstrated that the presence of the capsular layer on mycobacteria does not

diminish the binding of the phage pools to mycobacteria. However, in the case of *M. smegmatis* and BCG the binding is enhanced by the presence of the capsular layer. Notwithstanding, this protocol sufficed to allow for the identification of highly enriched clones that bind to mycobacteria, which is discussed in chapter three.

CHAPTER 3

CHARACTERISATION OF A HIGHLY ENRICHED PHAGE CLONE IDENTIFIED AFTER BIOPANNING BY HIGHTHROUGHPUT ILLUMINA SEQUENCING

Chapter Summary

At the end of a successful biopanning experiment, the discovery of the highly enriched clones or strong binders remains a challenge. Traditionally, the random clone-picking method is used to identify phage displayed peptides that were enriched during biopanning. However, depending on the sequence diversity at the end of the selection, this method may not necessarily identify the highestenriched clones. High-throughput (HTP) sequencing has now made it possible to sequence millions of inserts enabling a more quantitative approach to the identification of highly enriched clones resulting from a biopanning experiment. In this chapter, I used both the traditional random clone picking method and high throughput sequencing, to identify the highly enriched clones during biopanning. The most highly enriched phage (clone 1), displaying the CPLHARLPC peptide, was only identified by Illumina sequencing but not by random clone picking. This phage clone is able to bind mycobacteria significantly better than the unselected phage library. Surface plasmon resonance showed that the chemically synthesised CPLHARLPC peptide binds *M. tuberculosis* H37Rv after bacterial lysis and does not bind to non-mycobacteria lysates. Furthermore, the target for the phage 1 displayed peptide was characterised to be a mycobacteria protein of about 15 kDa in size. These observations demonstrate that phage display technology combined with high-throughput sequencing is a powerful tool to identify peptides that can be used for investigating potential biomarkers for TB and other bacterial infections.

1.0 Introduction

Short peptides identified from a phage display library have been previously used in a capture PCR to detect *M. paratuberculosis* from milk of infected dairy herds (Stratmann et al., 2002). This study aimed to identify phage clones that bind specifically to *M. tuberculosis* or mycobacteria. These clones could subsequently be used to probe for its target in patients' clinical samples and ultimately identify a novel biomarker for TB diagnosis.

However, one of the challenging steps in the use of phage display technology is the identification of the most promising candidates at the end of the biopanning experiment. While the random clone picking method is commonly used for the identification of the highest enriched clone during biopanning, it falls short of providing a sequencing depth that can enable a global evaluation of the enriched phage population. However, with the availability of HTP sequencing it is now possible to sequence millions of inserts displayed by the enriched phage population, allowing for a higher resolution of the selected pool of the displayed peptides. Thus, we employed both the traditional clone picking method together with HTP sequencing to identify highly enriched phage clones during biopanning. Surprisingly, we found that the most abundant phage clone, as identified through HTP sequencing, was missed by the traditional clone picking method. The identified clone was further characterised for binding specificity to mycobacteria and the nature of its target was characterised.

2.0 Materials and Methods

The bacterial strains used, their growth conditions, immobilisation of target mycobacteria, phage amplification, phage ELISA and phage DNA preparation and sequencing are as previously described in chapter 2 unless stated otherwise.

2.1 Phage 1 DNA amplification and sequencing

The DNA sequence which encoded for the phage displayed peptide was amplified by phage colony PCR. The primer pair sequences, forward (Find_Phage1_F1) and reverse primer (Find_Phage1_R1) are listed in Appendix table 1. The PCR reaction mix consisted of 0.5 µM of each primer, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 1 × DNA polymerase buffer (Promega, WI, USA) and 5U of the GoTaq® DNA polymerase mix (Promega, WI, USA) in a 100 µl final volume. Whole phage PCR (denaturation at 95°C for 1 minute, annealing at 55°C for 2 minutes and extension at 72°C for 1minute and 30 seconds) was performed as previously described (Kingsbury and Junghans, 1995). Thirty PCR cycles were performed. The resulting amplicons were separated on a 2% agarose gel. Three insert positive clones were randomly selected and sequenced. The sequencing results were compared to the phage sequences obtained from HTP Illumina sequencing, and the confirmed clone was amplified and titered for further characterisation

2.2 Mycobacteria specificity phage ELISA for the Illumina identified enriched clones

Microtiter plate wells were coated with 100 µl of mycobacteria suspension per well, with an optical density of 1.0 at 660 nm in carbonate buffer (35 mM NaHCO3, 15 mM Na_2CO_3 [pH 9.8]), and incubated overnight at 4ºC. A separate set of wells were blocked with blocking buffer without previous mycobacteria immobilization as negative controls (no target control). Plates were blocked overnight at 4°C with 200 µl of the gelatin (0.5%)-supplemented supernatant of an *E. coli* strain ER2738 F′ culture infected with the whole phage library. Each concentration of the selected and amplified phage clone was made up to one hundred microliters in PBS andtransferred to coated wells. Plates were incubated for 1.5 h at room temperature. Wells were then washed six times with PBST. Horseradish peroxidase (HRP)-labelled mouse anti-M13 monoclonal antibody (Amersham Biosciences) was diluted in PBS (1:5,000). Two hundred microliters of the antibody was added per well and incubated for 1 h at room temperature. This was followed by washing the wells six times with PBST. One hundred and fifty microliters of substrate solution (2,2'-azino-di-[3-ethylbenzthiazoline sulfonate] diammonium salt) (Thermo Scientific) was added and incubated for 30 min at 37°C. The reaction was stopped with 100 µl of 1% SDS and absorbance recorded at the wavelength of 405 nm.

2.3 Peptide Synthesis

Both peptides (Biotin-ACPLHARLPCG and its scrambled derivative Biotin-ACHLRPPLACG) were synthesised by GL Biochem (Shanghai, China), with the

C-C disulphide bridge. The peptides were supplied as a powder with purity above 85%.

2.4 Surface plasmon resonance (SPR) biosensor analysis

A Biacore™ 3000 instrument (GE Healthcare UK Ltd, Buckinghamshire, England) was used for SPR measurements. Instrument temperature was set to 25ºC and HBS-N (10 mM Hepes and 150 mM NaCl, [pH 7.4]) was used as running buffer. 50 µg/ml (80 µl) of streptavidin, in sodium acetate buffer [pH 4.5], was immobilised by amine coupling on the CM5 sensor chip (GE Healthcare UK Ltd, Buckinghamshire, England). Immobilisation was performed at a flow rate of 10µl/min for 7 min. The biotinylated peptides were captured using the previously immobilised streptavidin. A total of 60 µl of 100 µg/ml biotinylated-peptide in PBS (sample flow cell) was loaded onto the chip at the flow rate of 10 µl/min. No streptavidin or peptide was immobilised on the negative control flow cell. Binding of whole cell lysates to the biotinylated peptide was then analysed by diluting the lysate in HBS-N buffer (10 mM Hepes and 150 mM NaCl, [pH 7.4]) and passing it over the chip at 10 µl/min. The *M. tuberculosis* H37Rv whole cell lysate binding was analysed at concentrations of 100 and 500µg/ml of total protein, while the unrelated bacteria whole cell lysates were analysed at 100µg/ml of total protein concentration.

2.5 Protease digestion of *M. tuberculosis* H37Rv lysate

M. tuberculosis H37Rv whole cell lysate (BEI Resources, VA, USA) was treated with 1 mg/ml Pronase E (Sigma, St. Louis, MO) for 2 hours. The protease

digestion reaction was inactivated by heating at 90ºC for 20 minutes. The negative control reaction was treated in a similar manner in the absence of Pronase E.

2.6 Pull-down assay

Biotinylated peptides (60 µl; 1 mg/ml) were captured on 100 µl of streptavidin conjugated beads (Resyn Biosciences, Pretoria, Republic of South Africa). Two washing buffers, PBS and the same buffer with additional 0.5 M NaCl, were used during the experiment. The beads were washed three times with the relevant washing buffer prior to incubation with the *M. tuberculosis* H37Rv lysate. Subsequently, the streptavidin beads were mixed with 500 µg total proteins of *M. tuberculosis* H37Rv whole cell lysate. The final volume was adjusted to 500 µl with PBS and the mixture was placed on a rotating shaker with gentle mixing at 4ºC for 16 hours. Following incubation, the beads were pelleted using a magnet and the supernatant removed. The beads were washed three times each with 1 ml of the respective wash buffer. After the last wash, the pulled down mixture was resuspended in 50 µl of Laemmli sample buffer (63 mM Tris HCl [pH 6.8], 10% Glycerol, 2% SDS, 2% (v/v) 2-mercaptoethanol, 0.1% (w/v) Bromophenol Blue), denatured at 95ºC for 5 min, separated by electrophoresis onto a 12% SDS–polyacrylamine minigel and stained with Coomassie staining.

2.7 SDS PAGE gel analysis

The 12% SDS/PAGE gel was used which was comprised of a stacking gel of 4% (w/v) acrylamide and a separating gel of 12% (w/v) acrylamide. Gel electrophoresis of samples was carried out at 180V for 1 hour.

2.8 Liquid chromatography-tandem mass spectrometry LC-MS/MS analysis

Protein bands of interest were in-gel trypsin digested as described in (Shevchenko et al., 2006). In short, protein bands were destained using 50mM NH_4HCO_3/50% CH_3OH followed by in-gel protein reduction (50mM DTT in 25 mM NH_4HCO_3) and alkylation (55 mM iodoacetamide in 25 mM NH_4HCO_3). Proteins were digested over night at 37°C in 5 – 50 µl of 10 ng/µl tryspin depending on the gel piece size. Peptides were extracted using 50% acetonitrile (ACN)/5% formic acid (FA) and vacuum dried.

Samples were re-suspended in 35 µl 2% ACN / 0.2% FA and analysed using a Dionex Ultimate 3000 RSLC system coupled to a QSTAR ELITE mass spectrometer. Peptides were first de-salted on an Acclaim PepMap C18 trap (75 µm x 2 cm) for 8 min at 5 µl/min using 2 % ACN/0.2 % FA, then separated on AcclainPepMap C18 RSLC column (75 µm x 15 cm, 2 µm particle size) connected to the trap column via a 10-port switching valve. Peptide elution was achieved using a flow-rate of 500 nl/min with a gradient: 4-60 % B in 30 min (B: 80% ACN/0.1 % FA). Nano-spray was achieved using a MicroIonSpray head assembled with a New Objective, PicoTip emitter. An electrospray voltage of 2.0 - 2.8 kV was applied to the emitter. The QSTAR ELITE mass spectrometer was operated in Information Dependent Acquisition (IDA) using an Exit Factor of 2.0 and Maximum Accumulation Time of 2.5 sec. MS scans were acquired from m/z 400 to m/z 1500 and the three most intense ions were automatically fragmented in Q2 collision cells using nitrogen as the collision gas. Collision energies were chosen automatically as function of the mass of the ion (m) upon the charge (z) (*m/z*).

3.0 Results

3.1 Identification and characterisation of selected clones

Ten phage plaques from the final round of phage eluates were randomly selected and sequenced from each library. In addition to the traditional random clone picking method used to identify enriched peptides, high throughput Illumina sequencing was performed as described in chapter 2 (section 2.6) and used to identify the highest enriched peptides from both libraries.

3.1.1 CX_7C randomly selected phage clones

Three unique sequences, namely phage 2, 3, and 4, were obtained from the six successfully sequenced randomly selected plaques (Table 3. 1). Two clones, phage 3 and phage 4, were represented more than once (Table 3. 1). HTP sequencing, however, described a different quantitative landscape. HTP sequencing identified phage 1 as the most abundantly enriched clone, which was not identified using random clone picking. Phage 1 represented more than 80% of the enriched phage population as early as round three of biopanning. Nonetheless, there is some degree of correlation between the peptide sequences that were identified by traditional clone picking and the top five sequences identified using high-throughput sequencing. That is, all three unique sequences identified during random clone picking were in the top five of the most abundant peptides identified by HTP. Moreover, since the most abundant peptide had a frequency of more than 80% after the first three rounds of selection, this means that with the current sequencing depth, further rounds of selection will less likely have lead to the

identification of peptides that could not be found in the current available sequencing data.

To establish whether the selected peptides were not binding to non-targeted substrates and/or other components used during the biopanning process like BSA, we used the free web tool for scanning, reporting and excluding possible target-unrelated peptides from real binding peptides. This tool is called the Scanner and Reporter Of Target-Unrelated Peptides (SAROTUP) (Huang et al., 2010). None of the four sequences selected were identified as non-specific binders. However, the SAROTUP web server indicated that the phage1 displayed peptide has been previously isolated against an antibody targeting the surface haemagluttinin protein of the influenza A H1N1virus (Zhong et al., 2011) (Table 3. 1).

3.1.2 X_{12} randomly selected phage clones

Sequencing data from the ten randomly selected clones identified six unique sequences, namely phage 5, 6, 7, 8, 9 and 10 (Table 3. 1), while sequencing of one plaque was unsuccessful. Two clones, phage 9 and phage 10, were represented more than once (Table 3. 1). There was a good degree of correlation between the peptide sequences that were identified by traditional clone picking and the top fivesequences identified using high-throughput sequencing. That is, all the randomly selected clones with the exemption of phage 6, were also identified in the top ten identified by high-throughput sequencing. In constrast to CX_7Cbiopanning which resulted in an enrichment of a single peptide to levels higher than 80% of the final phage population, for X_{12} biopanning, none of the peptides were enriched to

levels higher than 12% of the total sequenced population. This means that with the current sequencing depth, further rounds of selection could have possibly led to further enrichment of specific peptides.

The six clones identified using HTP were scanned on the SARUTOP web server for non target related binding. Similar to the CX_7C sequences, none of the six sequences selected were identified as nonspecific binders using the SAROTUP web server. However, the peptide displayed by phage clone 7 has been previously isolated against several targets (Cao et al., 2011, Carter et al., 2006, Dickerson et al., 2008, Ejima et al., 2010, Estephan et al., 2009, Hardy et al., 2007, Kim et al., 2010, Mohammadi et al., 2007, Rothenstein et al., 2012, Roy et al., 2008, Sawada et al., 2011, Tipps et al., 2010, Yang et al., 2010, Zhao et al., 2010) indicating that it is possibly a non-specific binding peptide (Table 3. 1).

Table 3. 1: Summary of selected phage clones

Phage library	Phage clone	Percentage representation in the selected population				Phage displayed peptide sequence	SAROTUP: Target Unrelated Peptide scanner for other targets
		Random picking at round 5		HTP sequencing at round 5			
		n	%	n	%		
CX_7C	1[a†]	6	0	1 655 954	82.49	CPLHARLPC	Anti-influenza A H1N1 monoclonal antibody IV.C102 and SIV sera from patients (Zhong et al., 2011)
	2[a,b]	6	16.67	1 655 954	0.81	CHYDGARAC	None found
	3[a,b]	6	33.33	1 655 954	0.92	CDHGYLPSC	None found
	4[a,b]	6	50.00	1 655 954	5.05	CFDTRSLVC	None found
X_{12}	5[a,b†]	9	11.11	1 467 822	9.17	YMPNPFTASKWK	None found
	6[a]	9	11.11	1 467 822	Not detected	YYNNPLLISRLP	None found
	7[a]	9	11.11	1 467 822	11.93	LLADTTHHRPWT	Protein tonB (Carter et al., 2006), Umbilical endothelial cells (Hardy et al., 2007), Anti-MUC1 monoclonal antibody PR81(Mohammadi et al., 2007), Semiconductor crystalline surface ZnSe (Estephan et al., 2009), Bevacizumab-treated tumor vasculature in nude mice (Cao et al., 2011), Mouse embryonic stem cells (Zhao et al.), Glycine receptor subunit alpha-1(Tipps et al., 2010), Anti-EGFR monoclonal antibody 12H23 (Yang et al., 2010), Hyperbranched poly(phenylene vinylene) (Ejima et al.), In vitro M cell co-culture system (Kim et al.), Solubilized matrices (Integra ®) (Sawada et al., 2011), Hydroxyapatite, HA(Roy et al., 2008), titania (TiO_2) (Dickerson et al., 2008), Oxygen-terminated sides of single-crystalline ZnO and Zinc-terminated sides of single-crystalline ZnO (Rothenstein et al., 2012).
	8[a]	9	11.11	1 467 822	2.31	DLHPRRPPTIHD	None found
	9[a]	9	22.22	1 467 822	1.53	SITYMHHQPQKY	None found
	10[a]	9	33.33	1 467 822	6.46	YPAPQPLVTKTS	None found

[a]Clones identified in the top ten highly enriched sequenced by HTP sequencing

[b]Clones identified through random sequencing

[†] The highest enriched clone at round 5, as identified by HTP sequencing

3.2 Binding characterisation of the randomly selected recombinant phages

Binding of the randomly selected phage clones was investigated using an immobilised *M. tuberculosis* (Δleu/Δpan) strain (Figure 3.1A). All the randomly selected clones binding signal to *M. tuberculosis* (Δleu/Δpan) was significantly higher ($p < 0.01$) than the assay negative control without phage input (Figure 3.1). While there was no significant difference in the binding signal to *M. tuberculosis* (Δleu/Δpan) amongst all the phage clones identified from the X_{12} biopanning, phage 3 and 4 from the CX_7C biopanning bound *M. tuberculosis* (Δleu/Δpan) with a significantly higher binding signal ($p<0.05$) when compared to phage 2 which was also identified from the same library.

To evaluate the specificity of the selected phages, their binding to *M. tuberculosis* (Δleu/Δpan) was compared to their binding to *M. smegmatis* and BCG (Figure 3.1 B-C). All the clones binding signal to both BCG and *M. smegmatis* was higher than the negative control. However, there was no significant difference in binding of all clones to the three species of mycobacteria tested.

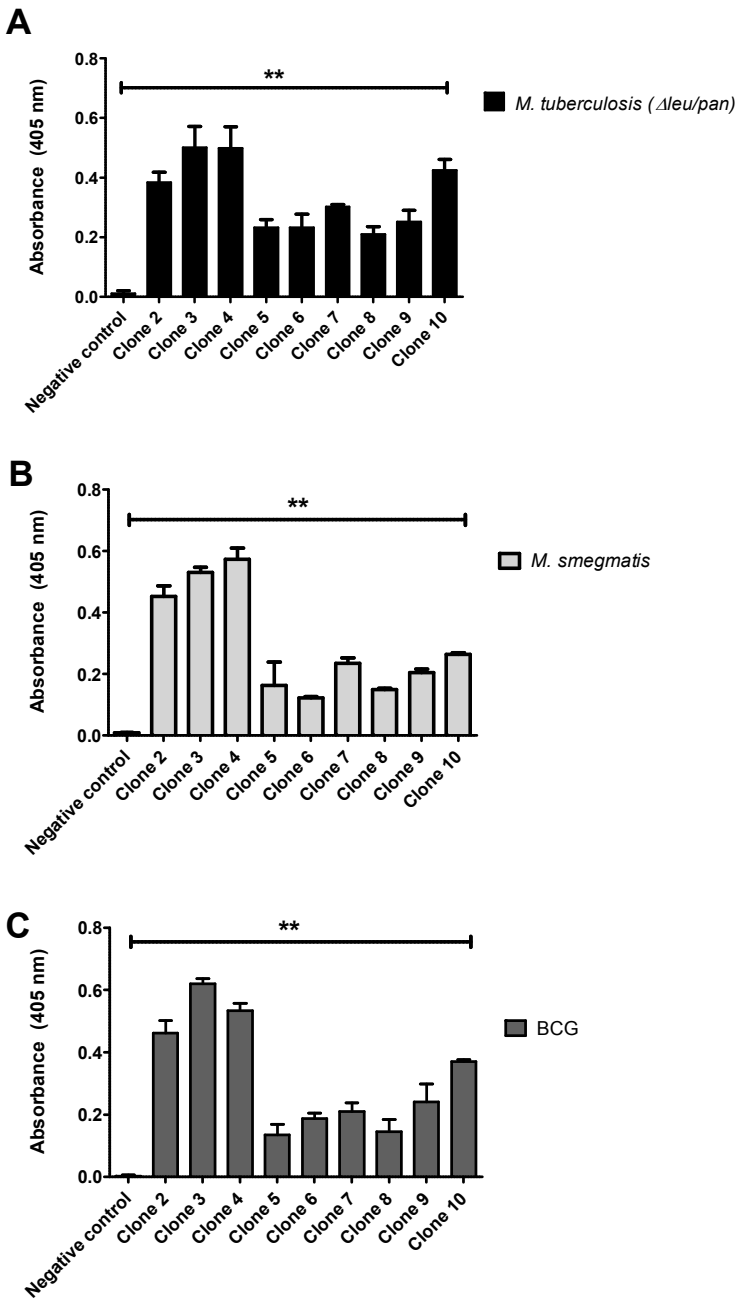


Figure 3. 1: Binding of the randomly selected phage clones to mycobacteria using an ELISA

Binding of the individual phages to (A) *M. tuberculosis* (Δleu/Δpan), (B) *M. smegmatis* and (C) BCG as compared to the assay buffer control without phage input (negative control). All values were normalised to wells without previous mycobacteria immobilization. A two-tailed, unpaired t-test was used to analyse significance (**p<0.01). Error bars represent standard deviations of the arithmetic means of optical densities at 405 nm performed in triplicates

3.3 Identification of the highest enriched phage (clone 1) from CX_7C biopanning

PCR was used to recover phage 1, which HTP sequencing identified as the highest enriched clone. A PCR primer pair was designed to specifically amplify the sequence encoding for the displayed peptide of phage 1. The forward primer (Find_Phage1_F1) sequence washomologous to the gp3 region of the phage genome upstream of the random peptide insert, while the reverse primer (Find_Phage1_R1) was homologous to 15 bases which encoded for part of the displayed peptide. Both primer sequences are listed in appendix supplementary table 1.

The monoclonal phages(2 and 3) identified by random picking, were plated on LB agar plates and the phage clones used as input phage to demonstrate the specificity of the primer pair on a PCR assay as described in section 2.1 of this chapter. There was no detectable amplification of the phage insert from both phage clones during the primer test PCR (Figure 3. 2A). The phage population from the final round of the CX_7C panning was alsoplated on LB agar plates and individual clones were used to screen for phage 1. Ten phage clones were selected from the plated, final round of enriched phages and subjected to phage 1 specific colony PCR. Six out of the ten phage clones were positive for amplification of the phage 1 random insert (Figure 3. 2B). Three positive clones were sent for sequencing and all three resulting sequences were identical to the phage 1 sequence identified using Illumina sequencing. One of the confirmed clones was amplified and titered for further characterisation.

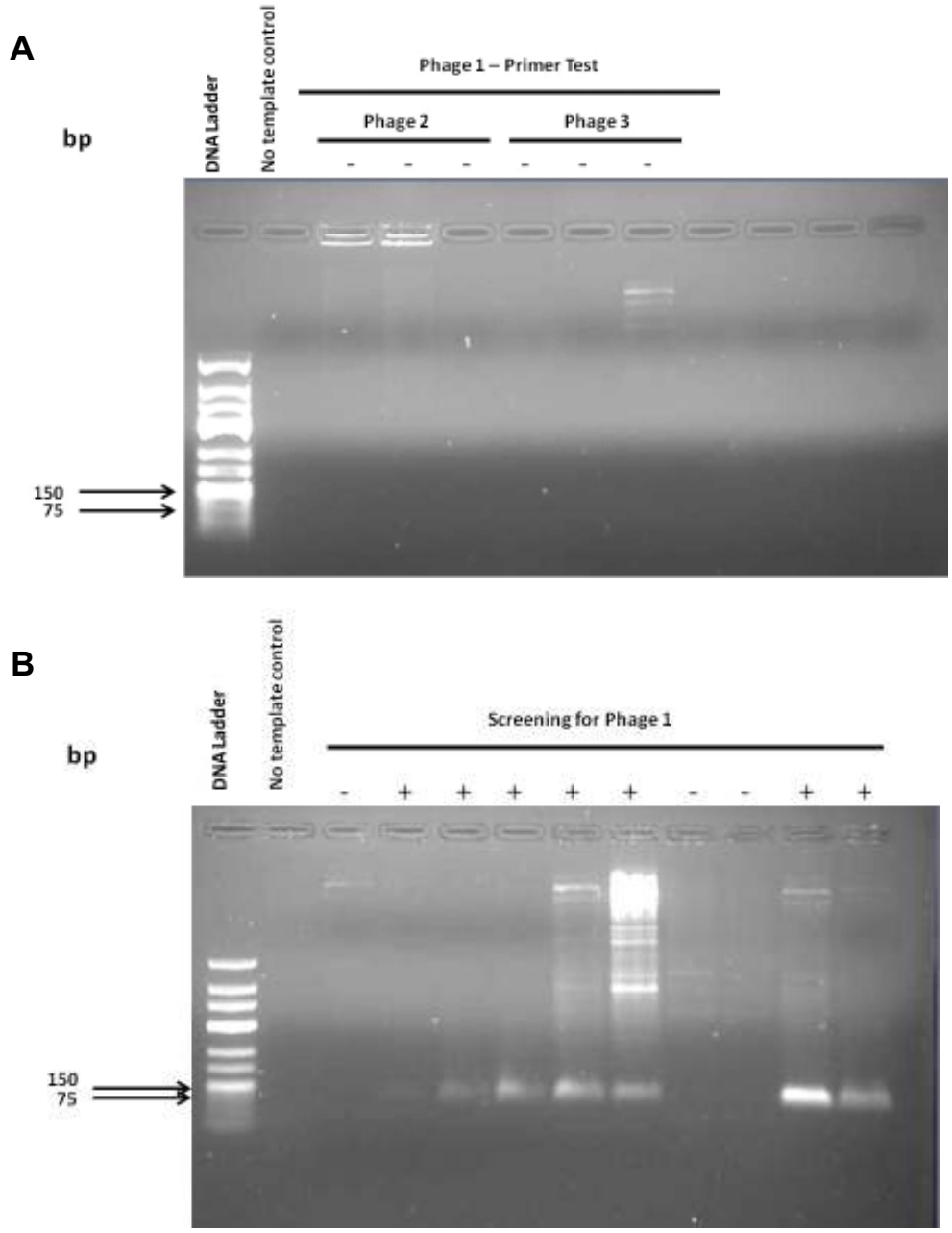


Figure 3. 2: Colony PCR for the identification of phage 1

(A) The PCR prime pair (Find_Phage1_F1/R1) was tested for specificity using randomly picked phage 2 and phage 3 as amplicons. No amplification was detected (-) for phage 1 insert. (B) Ten phage clonesfrom round 5 of CX_7C biopanning phage population were tested for positive amplification using the Find_Phage1_F1/R1PCR primer pair for amplification of phage 1 peptide insert. (+) positive for phage 1 insert.

3.4 Binding characterisation of the highly enriched, HTP sequencing identified Phage 1

The isolated highly enriched phage 1, was investigated for its ability to bind the *M. tuberculosis* (Δleu/Δpan) strain which was used during biopanning as well as binding to the pathogenic strain *M. tuberculosis* H37Rv, using ELISA. Phage 1 binding signal to both strains of mycobacteria was at least twice as high when compared to its binding to the unselected phage library which was used as a control (Figure 3. 3A-B). However, there was no significant difference in binding the pathogenic strain *M. tuberculosis* H37Rv when compared to *M. tuberculosis* (Δleu/Δpan). In addition, phage 1 was evaluated on its ability to discriminate between bacteria from the same genus of mycobacteria by comparing its binding to *M. tuberculosis* to two other mycobacteria species namely, *M. smegmatis* and *M. bovis* BCG. Our results show that phage 1 consistently bound to all the three mycobacteria species with a binding signal similar to that of *M. tuberculosis* (Δleu/Δpan) which was targeted during panning. That is, there was no significant difference in binding to all mycobacteria species tested. These results were consistent when tested using two different phage inputs of 5×10^{11} and 1×10^{12} PFU.

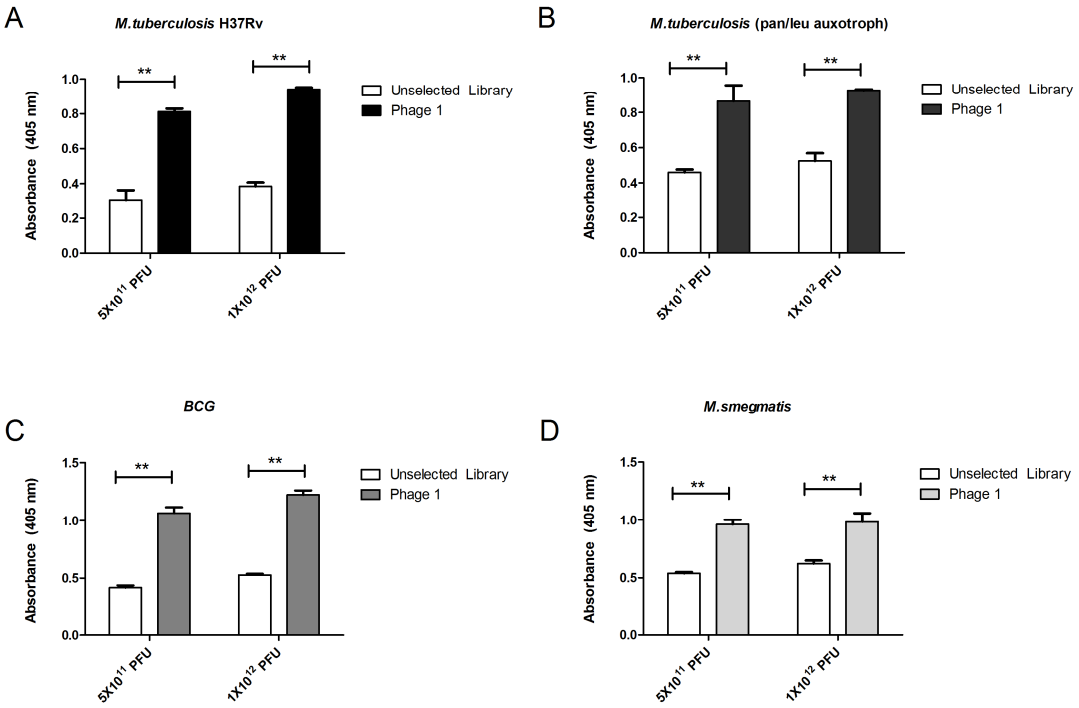


Figure 3. 3: Characterisation of the HTP identified phage (clone 1) binding to mycobacteria

Phage clone 1 (A-D) was amplified and 5×10^{11}PFU and 1×10^{12} PFU were used for a plate binding assay with (A) *M. tuberculosis* H37Rv, (B) *M. tuberculosis*(Δleu/Δpan), (C)*BCG* and (D) *M. smegmatis* as solid-phase antigens. The unselected library was used as a control. A two-tailed, unpaired t-test was used to analyse significance (**p<0.01). Error bars represent standard deviations of the arithmetic means of the normalised optical densities at 405 nm performed in triplicates.

3.5 Binding of the synthetic peptide displayed by phage 1 to *M. tuberculosis* H37Rv lysate

In order to test whether the highly enriched displayed peptide can bind its ligand when it is not displayed by the carrier phage, the peptide displayed by phage 1, CPLHARLPC (phage1-synpeptide) and a scrambled derivative CHLRPPLAC (phage1-synpeptide-Sc) were synthesised with a biotinylation modification on the N-terminus. The binding of the synthesised peptides to the whole cell lysate from *M. tuberculosis* H37Rv was evaluated using the SPR technology.

Biotinylated peptides were immobilised on a streptavidin sensor chip (Figure 3. 4A) and whole cell lysates were injected at different concentrations of total protein. When the peptide binding was measuredat 470 seconds after the start of lysate injection, there was a 7.1 fold increase in binding of the whole cell lysate to phage1-synpeptide at 100 µg/ml (Figure 3. 4B) as compared to the negative control which had no prior immobilisation of the synthesised peptides; and an even higher fold increase of 22.9 when the whole cell lysate concentration was increased to 500 µg/ml (Figure 3. 4C). The scrambled derivative had a fold increase of 3.8 compared to the negative control, at the total protein concentration of 100 µg/ml and only increased to 7.1 fold when at 500 µg/ml of the total protein concentration. This data indicates that, Phage1-synpeptide has a stronger and more specific interaction with H37Rv whole cell lysate than that of its scrambled derivative.

To further evaluate the binding specificity of the selected peptides to mycobacteria, their ability to bind to whole cell lysate from other unrelated bacteria was evaluated. The tested bacteria included both Gram negative and Gram positive

bacteria that are potential upper respiratory tract pathogens and the *E. coli* strain used for the amplification of phage during biopanning. Both the synthesised peptide and its scrambed derivative did not bind to any of the non mycobacteria lysates tested. That is, all non mycobacteria lysates tested showed similar binding to the synthesised peptide, its scrambled derivative and the negative control (no peptide) flow cell (Figure 3. 4D-G). This data suggests thatthe selected peptide specifically binds mycobacteria.

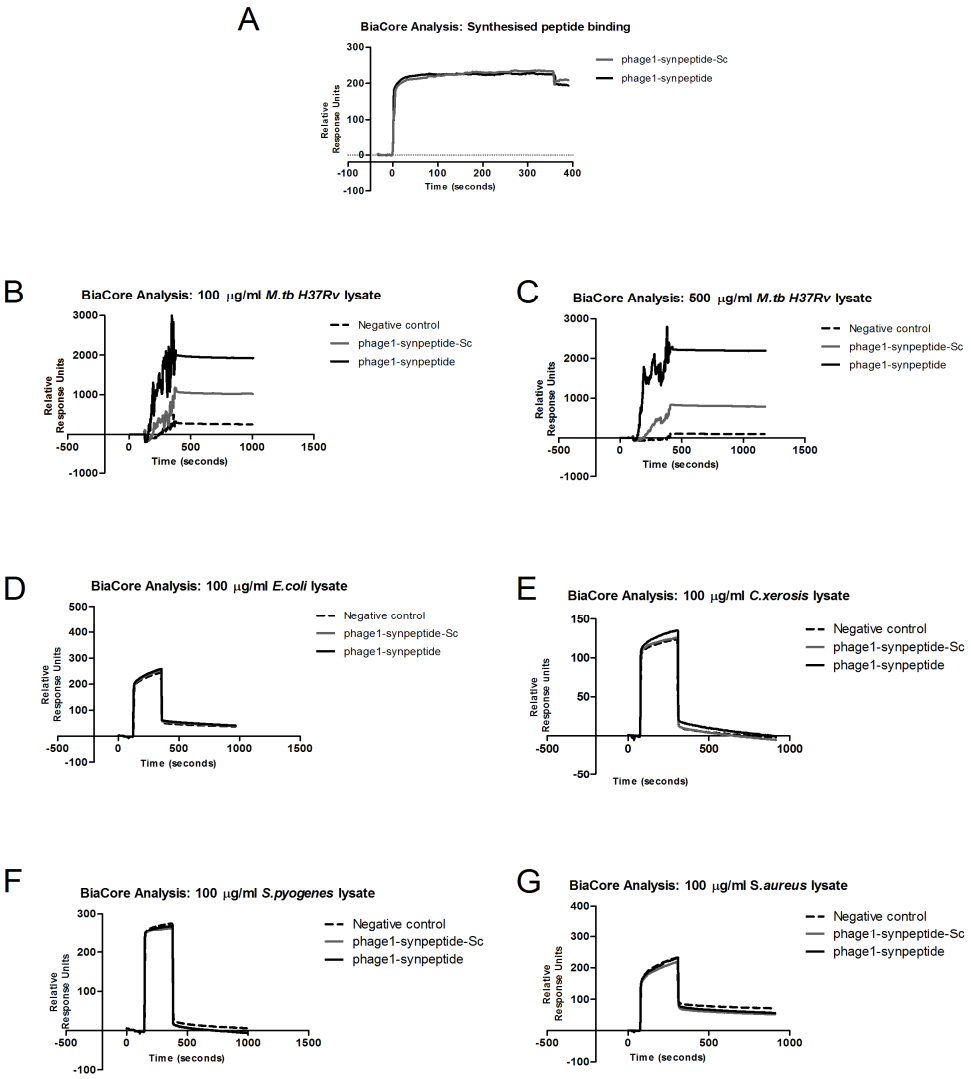


Figure 3. 4: Sensograms showing the association of bacteria whole cell lysates with immobilized synthetic peptides using SPR Biacore™ 3000

(A) Phage1 synthetic peptide (phage1-synpeptide) and scrambled peptide (phage1-synpeptide-Sc) were captured with covalently bound streptavidin on a CM5 chip. No peptide was captured on the covalently bound streptavidin for the negative control. Whole cell lysates of *M. tuberculosis* H37Rv with a total protein concentration of 100 µg/ml (B) and 500 µg/ml (C) were injected over the immobilised peptides. To evaluated specificity whole cell lysates with a total protein concentration of 100 µg/ml from unrelated bacteria, (D) *Escherichia coli* ER2738, (E) *Corynebacterium xerosis,* (F) *Streptococcus pyogenes,* (G) *Staphylococcus aureus*, were injected over the immobilised peptides. Changes in surface plasmon resonance were monitored in real time and are shown in response units.

3.6 Characterisation of the mycobacteria cell wall associated binding partner for Phage1 synthetic peptide (phage1-synpeptide)

To determine if the target of phage 1 displayed peptide on the surface of mycobacteria is a protein, we tested its binding to protease-treated *M. tuberculosis* whole cell lysate. The digestion was performed as detailed in section 2.5 of this chapter. Our results showed that the binding of phage1-synpeptide to *M. tuberculosis* whole cell lysate is abrogated after the lysate has been incubated for 2 hrs with Protenase E (Figure 3.5A-B). This data suggests that phage1-synpeptide binding partner is likely to be of a peptide or protein nature.

In order to determine which mycobacteria protein interacts with the phage 1 displayed peptide, we performed a pull down assay from *M. tuberculosis* H37Rv whole cell lysate using phage1-synpeptideas a capture peptide. Two wash buffers were evaluated, namely PBS and a more stringent wash buffer which was PBS with a highersalt concentration of 0.5M NaCl. The scrambled derivative, phage1-synpeptide-Sc, and the streptavidin beads without prior immobilisation of the phage 1 peptide were used as negative controls. A protein of approximately 15 kDa in size was pulled down from the *M. tuberculosis* whole cell lysate by phage1-synpeptide irrespective of the stringency of the wash buffer (Figure 3.6). Moreover, this peptide was not detectable on the PAGE gel using Coomassie staining when phage1-synpeptide-Sc was used a capture peptide, indicating the specific nature of the interaction between the phage 1 displayed peptide and its mycobacteria binding partner (Figure 3. 6).

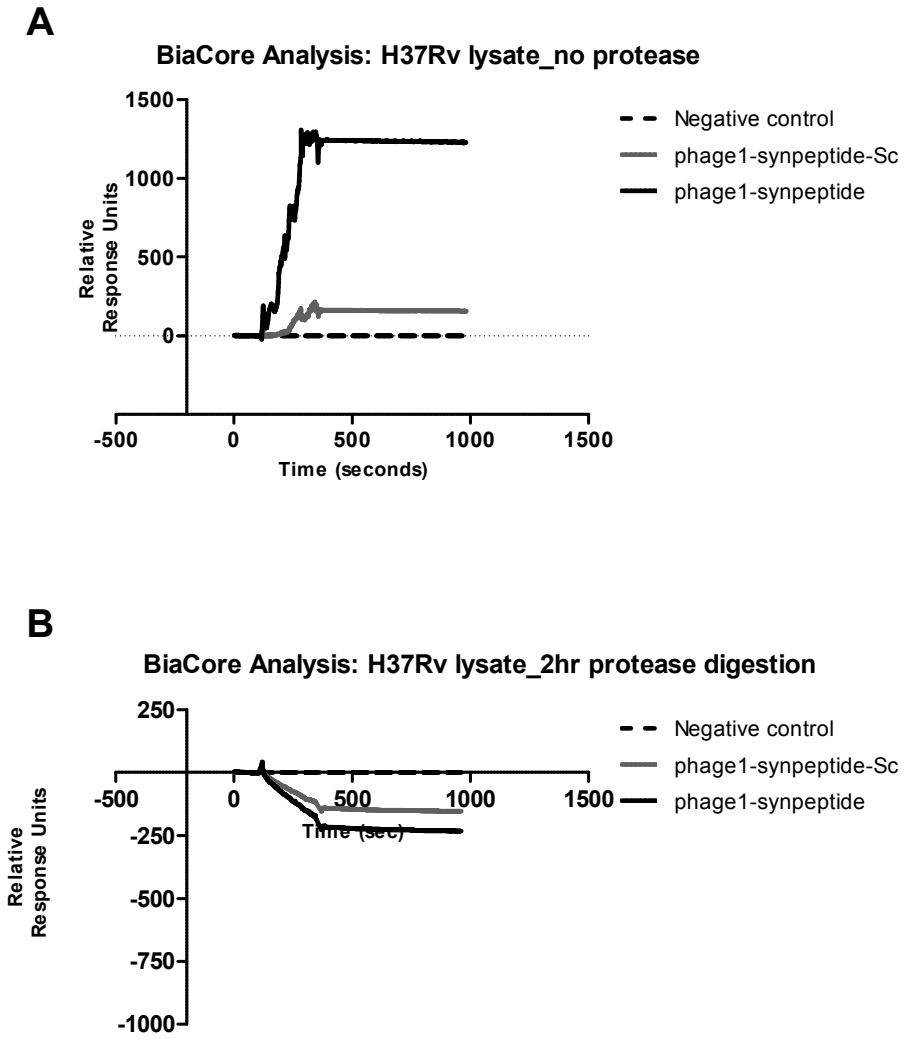


Figure 3.5: Sensograms showing the association of protease digested *M. tuberculosis* H37Rv whole cell lysatewith immobilized synthetic peptides using surface plasmon resonance Biacore™ 3000

(A) Biotinylated phage1 synthetic peptide (phage1-synpeptide) and biotinylated phage 1 scrambled peptide (phage1-synpeptide-Sc) were captured with covalently bound streptavidin on a CM5 chip**.** Protease-digested whole cell lysates of *M. tuberculosis* H37Rv with a total protein concentration of 100 µg/ml were injected over the immobilised peptides. (B) Control experiment using undigested *M. tuberculosis* H37Rv whole cell lysate. Changes in surface plasmon resonance were monitored in real time and are shown in response units.

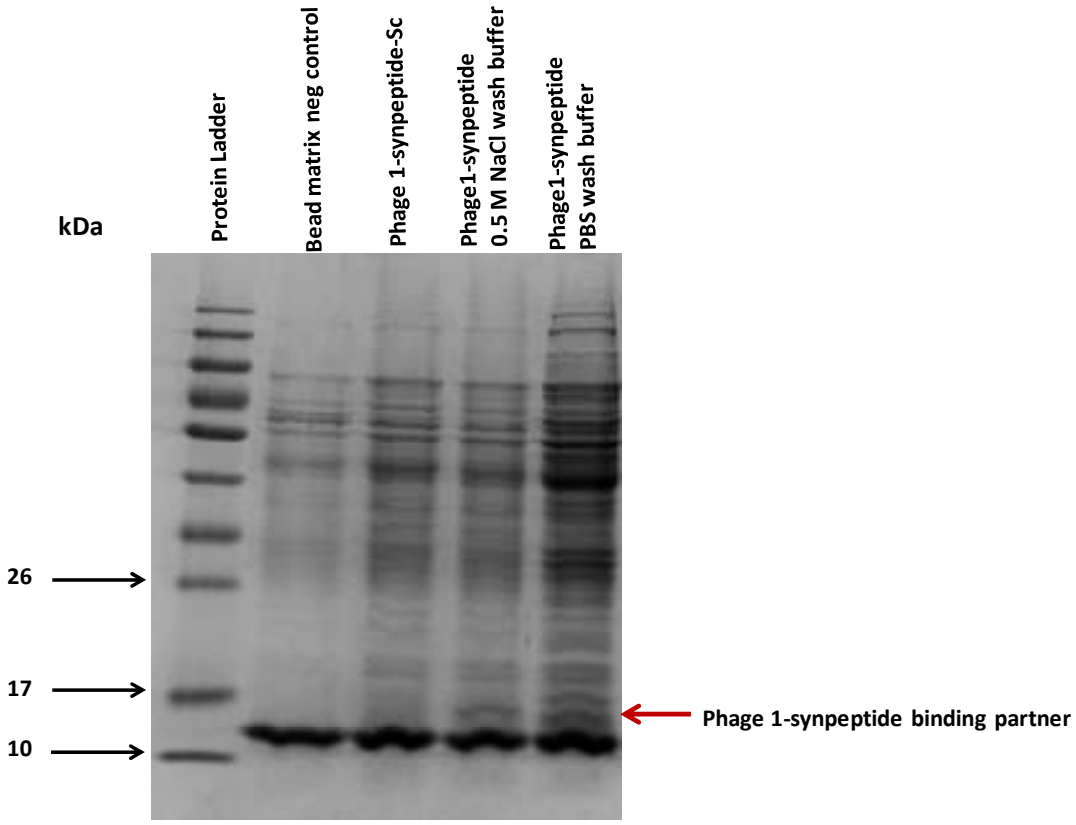


Figure 3. 6: SDS–PAGE of proteins pulled down from *M. tuberculosis* H37Rv whole cell lysate incubated with Streptavidin MagResyn™ magnetic beads and various peptides

Coomassie stained 12% SDS–polyacrylamide electrophoresis gel of proteins obtained from pull down experiments. Lane 1: molecular weight protein ladder. Lane 2: proteins pulled down from *M. tuberculosis* H37Rv whole cell lysate by the streptavidin beads in the absence of phage1-synpeptide. Lanes 3: proteins pulled down from *M. tuberculosis* H37Rv whole cell lysate by the streptavidin beads that had been pre-incubated with 0.5mg of bio-phage1-synpeptide-Sc. Lane 4: proteins pulled down from *M. tuberculosis* H37Rv whole cell lysate by the streptavidin beads that had been pre-incubated with 0.5 mg of bio-phage1-synpeptide. The red arrow indicates the possible bio-phage1-synpeptide binding protein from *M. tuberculosis* H37Rv.

To identify the mycobacteria protein that associates with the phage 1 displayed peptide, we utilised mass-spectrometry as described in materials and methods section2.8. The gel band corresponding to 15kDa on the PAGE gel was cut from each pull down fraction and analysed using mass-spectrometry. Figure 3.6 A, describes the numbers of proteins identified from the corresponding band of each pull down fraction. All proteins that were present in the beads only negative control were excluded as matrix non specific binders before any further analysis of the data (Figure 3. 6A). Furthermore, before the final analysis of the data, proteins identified from the scrambled derivative pull down also were excluded as non specific binders.

This resulted in two groups of possible mycobacteria proteins that associate with peptide1-synpeptide protein. Twenty eight were identified from the pull down using PBS as a wash buffer and 19 using the more stringent wash buffer of 0.5 M NaCl. The two groups had five proteins in common (Figure 3. 6B) which are listed in Table 3.2. Since these five proteins are common between the two experiments, they are the most likely candidates of the phage 1 displayed peptide target. Three of the proteins are hypothetical proteins which have not been previously characterised. Two of the proteins, identified as probable ribosomal proteins, have been characterised by experimentation or inferred by homology, to be associated with either the plasma membrane or the cell wall in mycobacteria (Table 3.2).

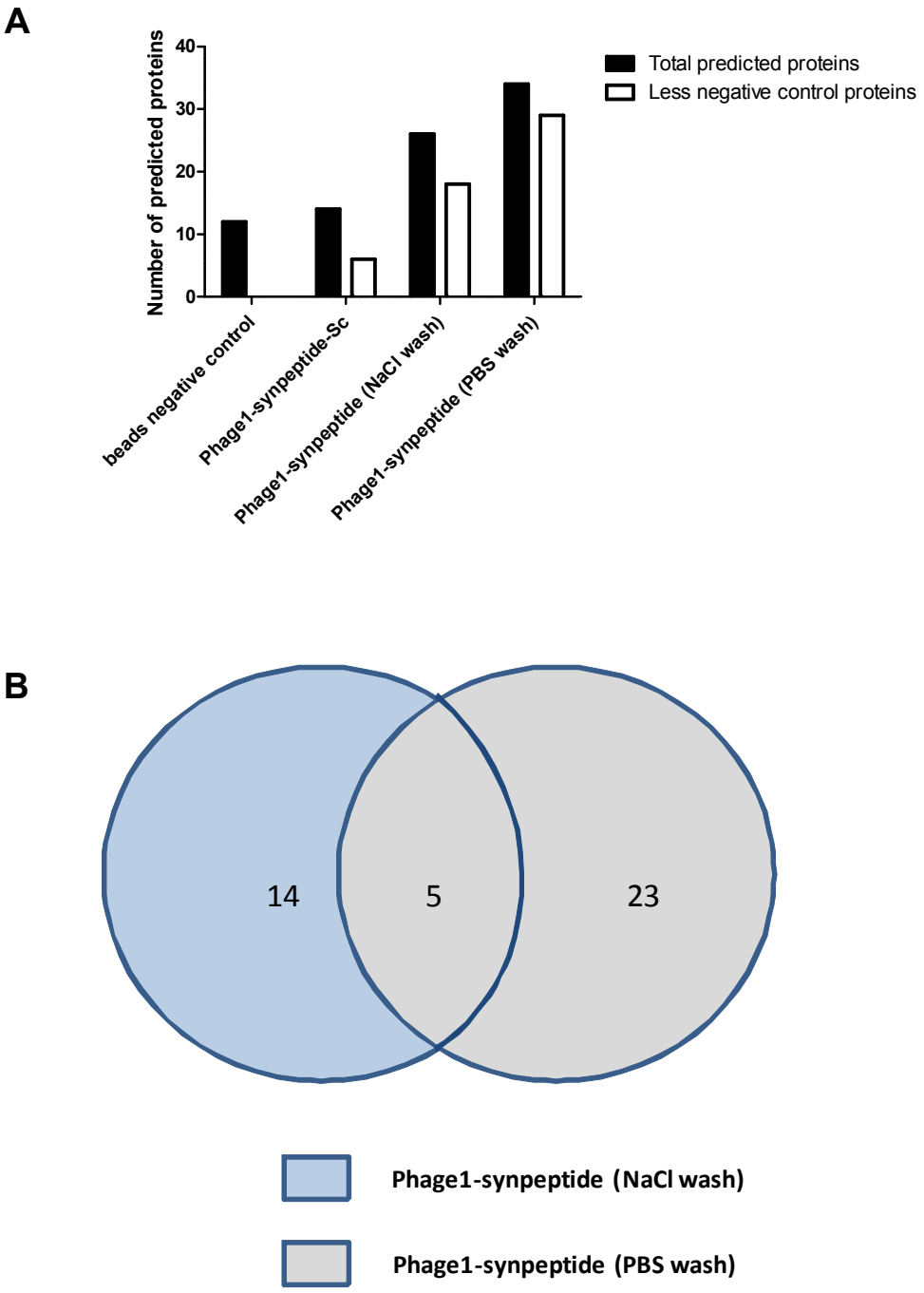


Figure 3. 7: A profile of proteins identified by mass spectrometry

(A) Total predicted proteins before and after the subtraction of the nonspecific proteins from the bead matrix negative control and (B) Venn diagram illustrating the distinct and overlapping protein compositions of the phage1-synpeptide affinity pull down assays performed using two different wash buffers.

Table 3. 2: LC-MS/MS overlapping predicted proteins

			Peptide 1 (NaCl wash)		Peptide 1 (PBS wash)	
Protein Name	Localization	Accession number	Unique peptides	Coverage (percentage)	Unique peptides	Coverage (percentage)
hypothetical protein Rv2901c	uncharacterised	P65051	4	55.5	4	73.3
probable ribosomal protein S11	plasma membrane (inferred from homology)	O06326	4	36	3	38.9
Hypothetical protein	uncharacterised	Q8VKL3_MYCTU	2	23.6	3	34
probable ribosomal protein L11 rplK	cell wall (Mawuenyega et al. 2005)	P66056	2	24.7	2	29.6
Hypothetical protein	uncharacterised	Q7D7S7_MYCTU	3	40.4	2	14.9

4.0 Discussion

In this chapter, we used the traditional clone picking method and high-throughput sequencing to identify the highest enriched clones during biopanning. Both HTP sequencing and random clone picking revealed multiple peptides that were enriched during the selection. However, only HTP sequencing enabled quantitative measures of enrichment, which allowed us to compare and rank our multiple hits. Our results showed that only one phage clone, phage1, resulting from the biopanning of the CX_7C library was significantly enriched during the selection. Interestingly, the traditional random clone picking method failed to identify this clone, despite the fact that HTP sequencing showed that this clone represented over 80% of the sequenced population after the first three rounds of selection. This result is in agreement with earlier findings showing that when comparing HTP sequencing to the traditional clone picking method, HTP sequencing accelerates the discovery of binders. This result demonstrates the advantage offered by the higher

resolution of the selection pools enabled by HTP sequencing, which will not be possible using random clone picking, since this method is limited by the number of clones that can be analysed.

When evaluating the binding of phage 1 to mycobacteria using an ELISA, there was no significant difference in the binding of this phage clone to the three mycobacterium speciestested. This suggests that phage 1 binds to a surface molecule conserved across all mycobacteria species. However, there are reagent similarities, like the plastic polystyrene microtiter wells, between the biopanning experiments and an ELISA. It is possible that these similarities could contribute falsely to apparent binding as previously discussed by others. This necessitates the use of a different method that does not include these materials, to further validate intended target binding. Indeed, when the peptide displayed by phage 1 was evaluated on surface plasmon resonance technology, the results showed that this peptide is able to associate with the *M. tuberculosis* H37Rv whole cell lysate while its scrambled counterpart exhibited minimal binding. The diminished binding of the scrambled peptide shows that the phage1-synthetic peptide sequence is important for its specific interaction with mycobacteria. Furthermore, this peptide showed no binding to unrelated bacteria that were tested which is indicative of specificity to mycobacteria.

To further confirm target related enrichment of the phage 1 displayed peptide, the peptide sequence was searched on the SARUTUP web server which has a database for non target related peptide sequences resulting from phage display (Huang et al.,

2010). This peptide was not identified as a non target binding peptide. However, a phage displaying the same peptide, CPLHARLPC, has recently been isolated and characterised as binding to the IV.C102 H1N1 monoclonal antibody and the swine-origin influenza virus A sera (Zhong et al., 2011). This monoclonal antibody has previously been demonstrated to bind to the type A H1N1 influenza strain epitopic peptide localised in residues 207-225 of the hemagglutinin HA1 subunit (Kiselar and Downard, 1999). However, there is no sequence similarity between the IV.C102 monoclonal antibody epitope (AIYHTENAYVSVVSSHYNR) on the hemagglutinin protein and the peptide displayed by phage 1. Luchesse and colleagues (2009) further characterised that peptide AIYHTENA is the minimal determinant epitopic region required for IV.C102 binding (Lucchese et al., 2009), which only has two amino acids in common, histidine and alanine, to the phage 1 displayed peptide. Nevertheless, the phage 1 displayed peptide contains five hydrophobic amino acids (out of nine), and the minimal epitopic peptide of the IV.C102 antibody also includes four hydrophobic amino acids (out of the eight). This may suggest that the interaction of this peptide with either the IV.C102 H1N1 monoclonal antibody or the mycobacteria cell wall associated target is via hydrophobic interactions.

The surface of mycobacteria has multiple possible binding partners for the phage 1 displayed peptide. These potential ligands vary in their nature, ranging from cell wall proteins, glycans and free lipids (Brennan, 2003b). In this work, we have demonstrated that the highly enriched phage 1 peptide binds to amycobacterial protein of approximately 15 kDa in size. However, the identification of the target protein requires further validation. Nonetheless, our mass spectrometry

experiments leads to five potential targets, of which three are hypothetical and have not been characterised to date. However, there is limited literature on two of these potential targets, namely the probable ribosoml proteins S11 and L11 RplK. These are predicted or known to be associated with either the cell membrane or the cell wall.

There isreported evidence of a large number of ribosomal proteins that are membrane associated in *M. tuberculosis*. Sinha and colleagues (2004) reported 28 ribosomal proteins that are membrane associated (Sinha et al., 2005), while in the preceding year Gu and colleagues (2003) also reported a number of membrane associated ribosomal proteins in *M. tuberculosis* (Gu et al., 2003). The localisation of the S11 protein is inferred from homology to similar proteins. While the 15 kDa L11 rplK protein (Festa et al., 2010) has been shown using 2DLC-MS/MS to be localised in the cell wall of *M. tuberculosis* (Mawuenyega et al., 2005). The size of these proteins is in agreement with our SDS PAGE size estimate of the potential phage 1 peptide binding partner, while their localisation agrees with the hypothesised localisation of the mycobacteria partner based on the selection target, which was surface exposed target. Taken together, these two proteins makes for promising phage 1 targets for further validation studies.

In conclusion, these findings show that phage display combined with HTP sequencing identified a specific peptide that bind to a mycobacteria-associated ligand. The peptide CPLHARLPC, which binds to a mycobacteria surface exposed protein, could be further developed as detection probe for TB infection.

CHAPTER 4

SELECTION OF RNA APTAMERS AGAINST *M. TUBERCULOSIS* EsxG PROTEIN USING SURFACE PLASMON RESONANCE

Chapter Summary

EsxG is one of the secreted substrates of the ESX-3 secretion system found in mycobacteria. ESX-3 has been hypothesised to be responsible for the secretion of some still unrecognized factor required for the optimal uptake of iron (Siegrist et al., 2009) and zinc (Serafini et al., 2009). There is evidence that *M. tuberculosis* experiences low iron conditions during macrophage infection (Gold et al., 2001)and that EsxG expression is repressed by the iron-dependent repressor (IdeR) when iron is available in both *M. tuberculosis* (Rodriguez et al., 2002) and *M. smegmatis* (Maciag et al., 2009). Taken together, we reasoned that EsxG might possibly be expressed in the infected host and can be investigated as a potential tuberculosis biomarker if an appropriate probe such, as an aptamer, was available to detect the protein. Aptamers are nucleic acid molecular probes, which have been explored as alternative recognition molecules to antibodies. In addition, some aptamers are able to inhibit the function of their targets similar to inhibitory antibodies (Bock et al., 1992, Hirao et al., 2000, Rhie et al., 2003, Tasset et al., 1997, Xu and Ellington, 1996). This chapter describes the isolation of 2'-fluoro-pyrimidine substituted RNA aptamers against the *M. tuberculosis* ESX-3 secreted protein, EsxG. After five rounds of aptamer selection using surface plasmon resonance at least five RNA aptamers were selected from an initial combinatorial library with a diversity of 10^{14}. Two high affinity monoclonal aptamers, GH 43 and GH 78, were found to preferentially bind EsxG over its homologous EsxA protein, which could be used for the development of novel diagnostics or techniques to investigate the role of EsxG in infection.

1.0 Introduction

Iron is an essential element for many organisms in a variety of metabolic and regulatory pathways. Since the innate immune system has evolved various mechanisms to restrict iron availability to microbial invaders, intracellular pathogens like *M. tuberculosis* developed countermeasures to obtain iron from their host (Ganz, 2009). There is now evidence that indicates *M. tuberculosis* does in fact encounter low iron conditions during infection of macrophages (Gold et al., 2001).

The mycobacteria siderophore mediated iron acquisition pathway is described in chapter 1 (section 1.4). Of importance is that the ESX-3 secretion system,which is one of five type 7 secretion systems (T7S) identified in mycobacteria, has been shown to be essential for growth in low iron conditions (Serafini et al., 2009, Siegrist et al., 2009). In brief, Siegrist and colleagues (2009) convincingly demonstrated that a *M.smegmatis* mutant which lacks ESX-3 (ΔESX-3) cannot be rescued by the exogenous addition of purified siderophore, mycobactin, indicating that the mutant is unable to use iron-bound mycobactin. In addition, Serafini and colleagues (2009) used a *M.tuberculosis* conditional mutant to demonstrate that ESX-3 is responsible for the secretion of some unknown factor(s) which is required for the optimal uptake of iron and zinc.

Until recently there was no high-throughput method to screen for possible T7S substrates, since the T7S had no known signal peptide for secretion. Daleke and colleagues (2012) have only recently identified the YxxD/E motif as a signal

peptide for the T7S, which can now be used to identify a number of potential unknown T7S substrates (Daleke et al., 2012). To date, there is evidence to show that the heterodimer EsxG:EsxH is secreted by ESX-3(Lightbody et al., 2004, Siegrist et al., 2009). Since the ESX-3 system is critical in low iron environmentsit is probable that EsxG or EsxH are expressed in an infected host and can be investigated as a potential tuberculosis biomarker.

Traditionally, antibodies are used as recognition probes for biomarker discovery. However antibodies have their limitations. Their synthesis requires animals or cell culture and they are limited by definition to immunogenic targets. In the recent past, aptamers have been investigated as alternative recognition molecules to antibodies. They are simpler to produce by means of chemical synthesis and are not limited to immunogenic targets. They have the ability to recognize, bind and in some cases inhibit function of their targets similar to that of antibodies (Bock et al., 1992, Hirao et al., 2000, Rhie et al., 2003, Tasset et al., 1997, Xu and Ellington, 1996).

This chapter describes the selection of RNA aptamers against EsxG using surface plasmon resonance technology.

2.0 Materials and Methods

2.1 Expression of EsxG protein

EsxG from *M. tuberculosis* H37Rv was cloned into the bacterial vector pET41A (gift of Professor Eric Rubin, Harvard Medical School, Boston, MA). The vector was used to transform *Escherichia coli* BL21 (DE3) cells and recombinant clones were selected on Kanamycin (50 µg/ml). Expression studies were performed as follows; a single colony was used to inoculate 5 ml LB media (containing Kanamycin 50 µg/m), followed by incubation at 37ºC overnight. The pre-inoculum (2 % (v/v) of the overnight culture) was then used to inoculate LB medium, the culture was incubated at 37 ºC until the OD_{600nm} was between 0.5 and 0.6. The expression was induced by addition of IPTG to a final concentration of 0.5mM, then incubation at 30ºC for 4 h. The cells were harvested by centrifugation using the Sorvall RC-5B super speed centrifuge at 10000 RPM for 20 min at 4ºC.

2.2 Purification of EsxG protein

The cell pellet was resuspended in 20mM Tris buffer [pH 7.5], containing 0.5M NaCl, and lysed using sonication for non denaturing purification. The cells were also lysed using B-PER (in phosphate buffer, 50mM [pH 7.5]), bacterial protein extraction reagent (Thermo Fisher Scientific, Rockford, Illinois) for denaturing conditions. Prior to purification the supernatant was ultracentrifuged using the Sorvall RC-5B super speed centrifuge at 20 000 RPM for 30 min at 4ºC. The recombinant protein was purified using immobilized metal affinity chromatography (IMAC) packed with Protino® Ni-TED resin (Machery-Nagel, Düren, Germany).

Bound tagged EsxG was eluted with 250 mM imidazole in 20 mM phosphate buffer containing 0.5 M NaCl. Eluted fractions were buffer exchanged to 20mM Tris buffer [pH 7.5] using the Vivaspin™ sample concentrators (GE Healthcare Lifesciences, Buckinghamshire, United Kingdom) with a molecular weight cut off at 10 kD. This size exclusion method was also used to remove proteins that co-purified with EsxG. The BCA protein assay kit (Thermo Scientific, Illinois, U.S.A) was used to determine the concentration of the purified protein. The purity of protein was assessed in a 18% Tris-tricine SDS/PAGE gel stained with Coomassie blue. EsxG expression was confirmed using mass spectrometry. Identification was on an algorithm *Thorough* search in Protein Pilot. An identification confidence of 95 % was selected during searches.

2.3 SDS/PAGE and Western blotting

The 18% Tris-tricine SDS/PAGE gel was used as described in (Schagger and von Jagow, 1987). Gel electrophoresis of samples was carried out at 200V for 2 hours. Gels comprised a stacking gel of 5% (w/v) acrylamide and a separating gel of 18% (w/v) acrylamide. Proteins separated by SDS/PAGE were transferred to Hybond-P PVDF transfer membrane (GE Healthcare Lifesciences, Buckinghamshire, United Kingdom) with a Bio-Rad Transblot semi-dry transfer cell at 12V for 2 hrs. To confirm the presence of the histidine tagged EsxG protein the protein bands were detected using a 1:10 000 dilution of the anti-histidine HRP-conjugated antibody (Qiagen, Hilden, Germany). A pre-stained ladder was used for the Western blot and the detection was accomplished with a colorimetric HRP substrate, TMB (3,3′,5,5′-tetramethylbenzidine) (Thermo Fisher Scientific, Rockford, Illinois).

2.4 Mass spectrometry analysis, LC-MS/MS

The mass-spectrometry protocol is described in chapter 3 section 2.8.

2.5 Preparation of the RNA library

The DNA library contained 50 nucleotides of random sequences flanked by 44 and 40 residues on the 5' and 3', respectively, giving a full sequence of 134 nucleotides, and was synthesized by Integrated DNA Technologies (Coralville, Iowa). The 5'-fixed region was 5'-AAT TAA CCC TCA CTA AAG GGA ACT GTT GTG AGT CTC ATG TCG AA-3'. The 3' fixed region that contained the T7 promoter for *in vitro* transcription (underlined and bolded) was 5'-**TAA TAC GAC TCA CTA TAG GG** AGA CAA GAC TAG ACG CTC AA-3'.

2.6 Amplification of the DNA Library

The DNA library was amplified by PCR under mutagenic conditions. The reaction mix consisted of 170 pmol of the single stranded DNA library, 1.1 μM of the T3 primer (T3SELEX5') and 3.3 μM of the T7 primer (T7SELEX3'), 150 μM of each dNTP, 10 mM of MgCl, and 5U of the GoTaq® DNA polymerase mix (Promega) in a 100μl final volume. The PCR conditions were as follows; template denaturing was performed at 93ºC for 30 seconds, annealing was at 55ºC for 30 seconds and elongation at 72ºC for 1 min with the final extension at 72ºC for 8 min. Cycles varied from 1 to 5, according to the maximum amount of amplicon obtained before amplification occurred. The resulting PCR product was cleaned up using the Wizard® SV Gel and PCR clean-up system (Promega, Wisconsin, USA).

2.7 In vitro transcription

The resulting dsDNA was converted to an RNA library using the T7 RNA polymerase. In the transcription reaction mixture, CTP and UTP were replaced with2' fluoropyrimidine triphosphates, 2'-F-CTP and 2'-F-UTP (TriLink, USA) to produce ribonuclease resistant RNA. For a 100 µl transcription, 225 pmol ofthe DNA template was used with 1.5 mM (final concentration) each of the 2' fluoropyrimidine triphosphates (TriLink, USA), 0.5 mM of unmodified purine nucleotides (Fermentas, USA), a transcription buffer (40 mM TrisCl [pH 7.9], 6 mM $MgCl_2$, 10mM DTT, 2 mM spermidine) and 200 units of T7 RNA polymerase (New England BioLabs, Massachusetts, USA) were used. The reaction mixture was incubated at 37^{o}C, overnight. Subsequently, the template DNA was removed by DNase I digestion for one hour. Free nucleotides were removed from the RNA transcript using a Sephadex-G50 column. Proteins were removed from the RNA using phenol/chloroform extraction as described in (material and methods2.8) and the RNA was subsequently precipitated using ethanol as described in (material and methods 2.9).

2.8 Phenol/chloroform extraction of RNA

An equal volume of Tris-buffered [pH 5] phenol/chloroform/isoamyl alcohol (v/v 25:24:1) was added to the RNA followed by vigorous mixing using a vortex and centrifugation at 14 000 rpm for 5 min. The aqueous (upper) phase containing RNA was transferred to a clean Eppendorf tube. An equal volume of chloroform:isoamylalcohol (v/v) was added to remove any trace of phenol followed by vigorous mixing using a vortex and centrifugation at 14 000 rpm for 5 min. The

RNA contained in the aqueous (upper) phase was transferred to a clean Eppendorf tube and precipitated with ethanol.

2.9 Ethanol precipitation of RNA

The RNA was precipitated by addition of 0.1 vol. of 3 M sodium acetate [pH 5.2], and 2.5 vol. of absolute ethanol. After 2 hours at -80^{o}C, the RNA was pelleted by centrifugation at 14 000 rpm (4^{o}C) for 30 min. The pellet was washed once with 3 vol. of 70% ethanol to remove residual salt, air-dried and resuspended in distilled water. The transcribed RNA poolintegrity was visually assessed on a 7 M urea gel, and the RNA quantified using a NanoDrop spectrophotometer (Thermo Scientific, MA, USA).

2.10 In vitro selection of RNA aptamers against EsxG protein using the NTA system on the BIAcore™ 3000

The SELEX process was performed using surface plasmon resonance on the BIAcore™ 3000 instrument. The HBS-P (0.01 M HEPES [pH 7.4], 0.15 M NaCl, 0.005% Surfactant P20) continuous flow buffer (GE Healthcare Lifesciences, Ohio, USA) was used in all the biacore experiments on the nitrilotriacetic acid (NTA) chips. All samples which were injected onto the sensor chip were spun at 14 000 rpm for 1 min to dislodge air-bubbles. Briefly, the recombinant EsxG was immobilized on the NTA sensor chip (GE Healthcare Life Sciences, Ohio, USA) via nickel chelation. The NTA chip was briefly washed with 20 µl of regeneration solution (HBS-P supplemented with 0.35M EDTA, adjusted to [pH 8.3]). The surface of the NTA sensor chip wasfirst activated with Ni^{2+} ions using 30 µl of the

nickel solution (500 µM $NiCl_2$ in HBS-P, 50 µM EDTA) at a flow rate of 20 µl/min, to saturate the chip. Following surface activation, 30 µl of 200 nM of EsxG (His_6) was injected at 2 µl/min. The transcribed RNA pool (representing approximately 10^{14} unique molecules) was denatured at 95ºC for 4 min and refolded at room temperature in HMCKN buffer (10 mM Hepes [pH 7.4], 1 mM $MgCl_2$, 1 mM $CaCl_2$, 2.7 mMKCl, 150 mM NaCl) for an additional 4 min. The RNA pool was injected for 2 min into the flow cell at a flow rate of 20 µl/min. After each selection cycle the bound RNA was eluted from the NTA using theelution buffer (7 M urea, 5 mM EDTA). The recovered RNA was then converted to cDNA by reverse transcription using the RevertAid M-MuLV Reverse Transcriptase (Fermentas, USA) according to manufacturer's protocol. The cDNA was amplified by PCR as described in (2.6) with the exception of the $MgCl_2$ which was reduced to 1.5 mM. The number of cycles was determined for each round of selection to avoidoveramplification.

After five rounds of selection, the enriched RNA pool was reverse transcribed; the resulting cDNA amplified byPCR and cloned into the TA cloning vector using the pGemT-Easy cloning kit (Promega, Wisconsin, USA) as per manufactures instruction. The pGemT-Easy ligation mix was transformed into DH5α chemically competent cells (Invitrogen, Carlsbad, CA, USA) as described in 2.11. The individual clones were screened for insert and the insert sequence identified by DNA sequencing and screened for the binding using the biacore.

2.11 Transformation of bacterial competent cells

DH5α chemically competent cells (Invitrogen, Carlsbad, CA, USA) were thawed on ice, and 2 µl of the pGemT-Easy ligation reaction was directly added into 50 µl

competent cells and mixed gently. The transformation mix was incubated on ice for 30 min., followed by heat shock for 30 seconds in the 42°C water bath and immediately placed back on ice. 250 µl of SOC medium (Invitrogen, Carlsbad, CA, USA) equilibrated at room temperature was added to the transformation mix, and the vial was shaken (225 RPM) at 37°C for at least 1 hour. Different volumes (10-50 µl) of transformed bacteria were plated on LB agar plates containing 50 µg/ml ampicillin and incubated for 16 hours at 37°C.

2.12 Colony PCR for validation of monoclonal aptamer clones

The T3SELEX5' and T7SELEX3' primer pair (Appendix Table 1) was used for screening insert positive clones for colony PCR. A single colony from the plated transformed bacteria was used for each PCR reaction. The reaction mix consisted of 0.5 µM of the T3SELEX5 primerand 0.5 µM of the T7SELEX3' primer,80 µM of each dNTP, 1.5 mM of MgCl, and 1U of the GoTaq® DNA polymerase mix (Promega) in a 100 µl final volume. The PCR conditions were as follows, template denaturing was performed at 95ºC for 10 min, annealing was at 55ºC for 30 seconds and elongation at 72ºC for 30 seconds with the final extension at 72ºC for 8 min. Twenty five cycles of PCR were performed. The resulting PCR product was evaluated for the presence of a 150 bp insert which represented thepresence of the insert. Insert positive clones were subsequently sequenced to obtain monoclonal aptamer sequences.

2.13 Binding assay using the nickel chelation on theBIAcore™

Screening of monoclonal RNA aptamers was performed using the modified NTA immobilization protocol as described in (Willard and Siderovski, 2006). Briefly, the BIAcore™ 3000 biosensor instrument (GE Healthcare Lifesciences, Ohio, USA) flow rate was set at 5 µl/min. Nickel was bound to the NTA chip using the nickel solution (500 µM $NiCl_2$ in HBS-P, 50 µM EDTA). The carboxymethylated dextran (CMD) surface was then activated using 70 µl of *N*-hydroxysuccinimide (NHS)/1-ethyl-3(3-dimethylaminopropyl)-carbodimide hydrochloride (EDC). EsxG (30 µl, 200nM) in pH 7.4 running buffer, was then captured at a low flow rate (2 µl/min). This was followed by the blocking of the uncoupled amino reactive sites on the CMD surface by injecting 70 µl of 1 M ethanolamine-HCl[pH 8.5]. Nickel and non-covalently coupled EsxG were subsequently stripped from the surface of the sensor chip using HBS-P buffer containing 350mM EDTA.

2.14 Binding assay using the amine coupling on the BIAcore™

A standard amine coupling protocol was followed according to manufactures instructions (Biacore, GE Healthcare Lifesciences, Ohio, USA). The BIAcore™ 3000 biosensor instrument (GE Healthcare Lifesciences, Ohio, USA) flow rate was set at 10 µl/min and the carboxymethyl groups on the chip were activated by injecting 50 µl of 0.5 M N-hydroxysuccinimide (NHS), and 0.2 M N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC). The EsxG protein was diluted into 10 mM sodium acetate, pH 4.5 to a final concentration of 2 µM and was injected onto the activated surface, resulting in amine coupling of the protein to the activated surface. The remaining activated carboxymethyl groups were blocked

by injecting 70 µl of 1 M ethanolamine-HCl, pH 8.0, which also helped elute non-covalently bound ligand.

2.15 Binding kinetics analysis using the amine coupling on the BIAcore™

Affinity measurements were performed at room temperature (25^{o}C), which was the same temperature used during selection. 1000–1200 RU of EsxG was covalently immobilized on the chip as described in (material and methods 2.14). The concentration of the monoclonal RNA aptamer was measured using the nanodrop, and dilution series of the aptamer was prepared (5 nM– 500 nM). Each concentration of the aptamer (20 µl, starting with the lowest) was injected at 10 µl/min and allowed to dissociate over 600 seconds. To ensure that there was no residual RNA left on the surface before the next concentration injection, the ligand was regenerated by injecting 10 µl of freshly prepared 100 mM NaOH to remove the RNA that was still bound. Data was analysed using the BIAevaluation 3.0 (Biacore, GE Healthcare UK Ltd, Buckinghamshire, England) and GraphPad Prism 5.0 (GraphPad software Inc., USA) software and the affinity constant was calculated from the k_{on} and k_{off}.

2.16 Iron Starvation Experiments

The iron starvation experiments were performed by Marte Singsås Dragset at Harvard University. The strains and methods used for these experiments have been previouly described (Siegrist et al., 2009). Briefly, *M. smegmatis* was cultured in 7H9 medium 0.05% Tween-80 and 10% albumin-dextrose-catalase (ADC) in the presence of kanamycin and hygromycin as needed. The culture was then

subcultured once in chelated Sauton's medium (consisting of 60 ml of glycerol, 0.5 g of KH2PO4, 2.2 g of citric acid monohydrate, 4 g of asparagine, and 0.5% Tween-80). After adjustment of the pH to 7.4, the medium was stirred 1–2 days at room temperature with 10 g of Chelex 100 resin (Sigma). The medium was filtered, and 1 g of MgSO4•7H2O was added as asterile solution), and then diluted 1:1,000 in chelated Sauton's medium without antibiotics containing either aptamer or HMCK aptamer refolding buffer. The growth curves for the cultures was monitored over 72 hours and recorded as absorbance readings at 600nm.

3.0 Results

3.1 Purification of EsxG protein

His-tagged EsxG was expressed in soluble form in *E. coli* cells and subsequently purified using IMAC. The concentration of the purified protein was about 300 µg/ml and the total yield was 0.6 mg per litre of culture as quantified by the BCA assay. Disruption of the *E.coli* pellet was performed under native conditions using sonication as well as denaturing conditions using the B-PER reagent (Pierce Protein Research Products, Rockford, Illinois) which contained detergent. The purification procedure involved a single pass over the Ni-TED column, resulting in the capture of most of the expressed EsxG by the column as indicated by the very faint band corresponding to EsxG in the lane showing the flow through (Figure 4. 1A). However, several contaminating bands were consistently observed along with the eluted EsxG protein (Figure 4. 1A). Co-purified protein bands were removed by size exclusion. Using this approach we were able to purify pure fractions of the expressed protein.

The molecular mass of the purified EsxG was estimated at 11 kDa when compared to the protein standards (Biorad, CA, USA) on 12% SDS-PAGE gel (Figure 4. 1B). A similar band migration pattern was observed whether the protein was purified under native or denaturing conditions using the B-PER reagent (Pierce Protein Research Products, Rockford, Illinois). The presence of EsxG protein was validated by Western blot using the horseradish peroxidase conjugated anti-histidine (antihistidine-HRP) antibody (Figure 4. 1C).

In addition, mass spectrometry-based peptide sequencing was used to confirm protein identity. The results from the gel digestion of the protein bands and an in solution digest of the purified fraction identified EsxG (His_6) with 100% and 89.3% sequence coverage, respectively. Taken together, these results indicate that EsxG was successfully purified using nickel chromatography together with size exclusion, under these conditions.

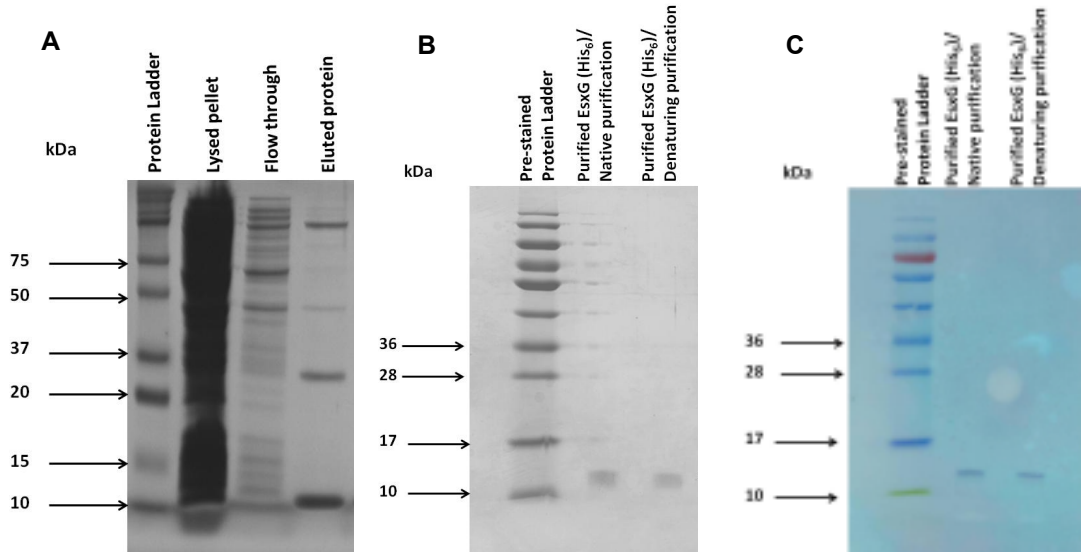


Figure 4. 1: Purification and SDS-PAGE analysis of purified EsxG protein

EsxG was purified on a Ni-TED column. (A)Samples from various steps of the purification procedure were analyzed by 12% SDS-PAGE gel, and the bands were visualized by Coomassie protein staining. Lane 1 showing the lysed pellet after sonication; lane 2 showing the flow through material from Ni-TED columns; lane 3 shows the protein eluted with 50 mM imidazole (B) Coomassie stained 12% SDS-PAGE gel of purified EsxG after size exclusion purification. (C) Western-blot of EsxG using the horseradish peroxidase conjugated anti-histidineantibody. A prestained ladder was used for the Western blot and the detection was accomplished with a colorimetric substrate, TMB.

3.2 Amplification and transcription of the DNA library

The first step in the SELEX experiments was to use a pool of variant sequences from which RNA ligands of relatively high affinity for EsxG could be selected. For this experiment we chose a similar DNA library previously used by Khatiand colleagues (2003) to isolate high affinity aptamers against the surface glycoprotein of HIVgp120.

A 134-mer DNA template randomized at 50 nucleotide positions (Figure 4. 2A) was amplified by mutagenic PCR with varying number of cycles to achieve sufficient template without over amplification of the PCR product (Figure 4. 2B). Four PCR cycles resulted in the cleanest single band of the expected size to be used for transcription (Figure 4. 2B). Fewer cycles did not yield enough template, while, additional cycles resulted in over amplification of the PCR product as indicated by the presence of higher molecular weight bands (Figure 4. 2B).

The resulting dsDNA template was purified using the membrane based system, Wizard® SV Gel and PCR Clean-Up System kit (Promega, Wisconsin, USA) and transcribed in vitro using 2'-fluoro-pyrimidines as described in (2.7). The transcription product was digested using DNase I to remove the template DNA. Free nucleotides were removed from the resulting RNA using a Sephadex-G50 column prior to the removal of proteins using phenol/chloroform. Subsequently, the RNA was precipitated using ethanol and the quality confirmed on a denaturing urea gel (Figure 4. 2C).

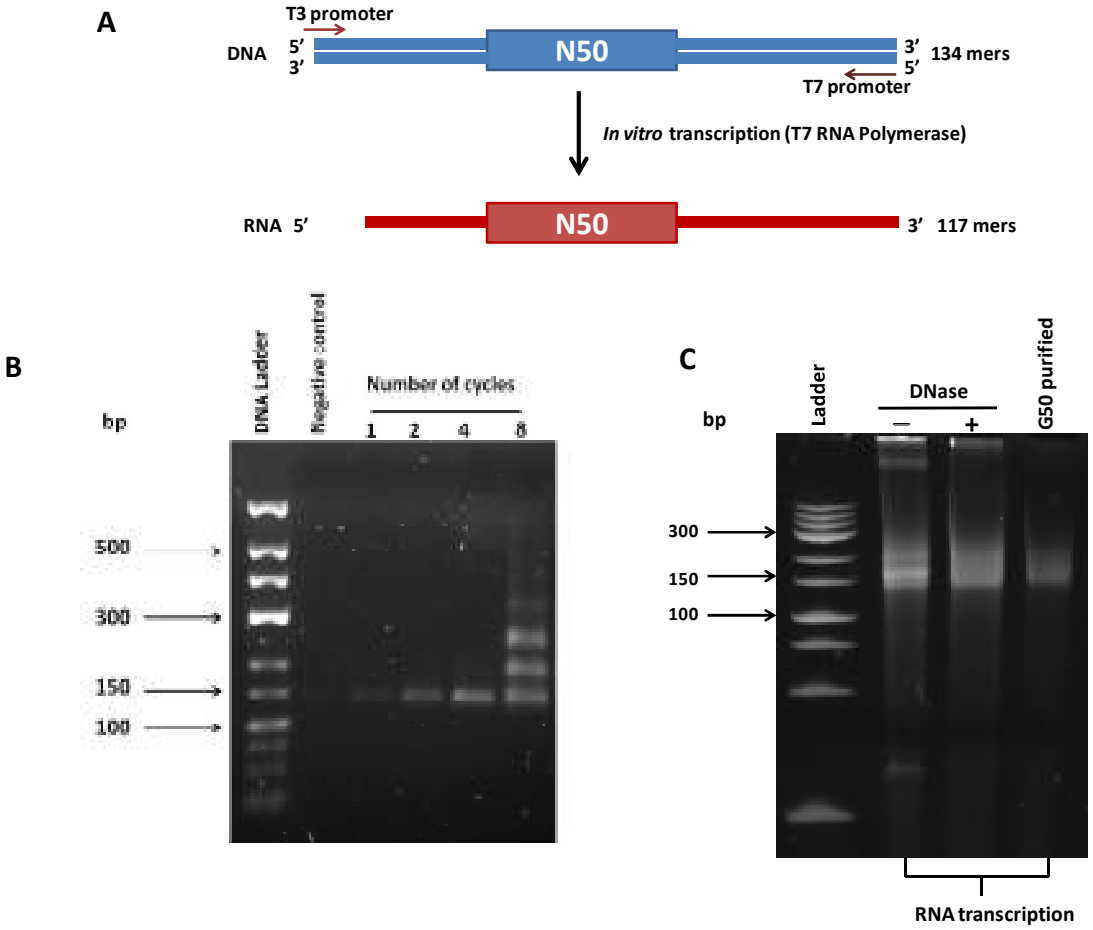


Figure 4. 2: PCR amplification and transcription of the variant template

(A) The double–stranded DNA template was transcribed in vitro using the T7 RNA polymerase to give a 117 nucleotide 2'F–RNA molecule. (B) Agarose gel analysis of the PCR products, using the T3SELEX5' and T7SELEX3' primer pair, yielded by varying numbers of cycles, showing the cleanest band of interest produced after 4 cycles. (C) Urea gel analysis of the RNA transcription, showing the RNA transcript before and after DNaseI digestion and the cleanest RNA transcript after desalting using a Sephadex-G50 column and ethanol precipitation.

3.3 Selection of RNA aptamers using the BIAcore™

For this selection the target protein was immobilized on the NTA chip (Figure 4. 3A) as described in (2.13). We reasoned that the NTA system will be more suitable than the CM5 chip because it allowed for a similar orientation of the recombinant

protein for every round of selection, standardizing the SELEX surface. This advantage in orientation is achieved because the NTA system will immobilise the protein ligand via the histidine tag, meaning all proteins on the surface of the sensor chip will be in the same orientation. In contrast, coupling on the CM5 chip results in a random orientation of the protein on the sensor chip surface because the active succinamide esters on the sensor surface will form covalent bonds with the reactive primary amines on the lysine residues and/or N-terminus of the protein ligand.

During the first selection cycle, approximately 10^{14} RNA sequences (20 μg) were injected into the flow cell for binding to the recombinant EsxG as described in (2.10). The bound RNA was eluted off the immobilized EsxG using7M urea buffer to provide one fraction (4μl) of RNA. The eluted RNA was used as a template to regenerate the binding RNA species via reverse transcription, PCR, and in vitro transcription reaction.

During the subsequent rounds of selection, the amount of starting RNA for each round was decreased to 10 μg. This was done to reduce the competition among sequences for binding to the limited amount of EsxG on the chip surface, while progressively enriching for the selected species of RNA.

Prior to the final round of selection, round five, the RNA was pre-cleared by passing it over a nickel charged flow cell before selection on EsxG. This was done to reduce the selection of aptamers that might bind to the chip matrix. After five rounds of selection the amount of target bound species increased when compared to

the unselected library, as indicated by the increase in response units of the selected biacore selection (Figure 4. 3B). This indicated that the anti-EsxG aptamers were enriched in the round 5 polyclonal pool of RNA.

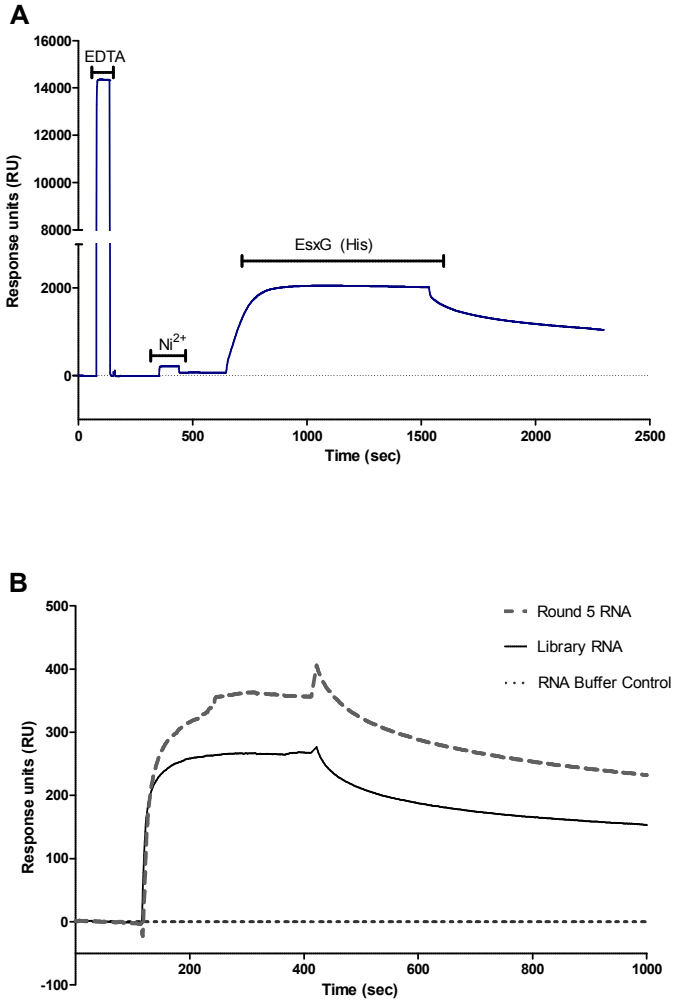


Figure 4. 3: Isolation of aptamers that bind to EsxG

(A) Nickel was bound to an NTA sensorchip by the injection of nickel solution (500µM $NiCl_2$ in HBS-P, 50µM EDTA). EsxG (200 nM) in eluent buffer was captured on the biosensor surface by injection at a low flow rate (30 µl, 2µl/min) before binding of polyclonal RNA pools. (B) After five rounds of selection, the polyclonal pool of 2'F-RNA was analysed for binding to covalently bound EsxG on a CM5 chip. The binding of the final round of selection was compared to the initial library before selection. The binding curve was derived after subtraction of measurements from the RNA buffer only control.

3.4 Cloning and sequencing of monoclonal aptamers

After five rounds of selection the resulting PCR product was cloned into pGEM-T Easy vector using TA cloning. Mutagenic PCR which was used during selection was abandoned by reducing the concentration of $MgCl_2$from 7.5 mM to 1.5 mM during cloning and subsequent screening procedures. This allowed for the high fidelity amplification of the template and hence conserved the nucleotide sequence of the EsxG-binding species. The TA cloned aptamers were used for bacterial transformation as described (2.11) and colony PCR, using the T3SELEX5' and T7SELEX3' primer pair (Appendix Table 1) was performed to screen bacterial transformants for inserts. Figure 4. 4 shows a representative agarose gel of colony PCR products used to evaluate the presence of the monoclonal aptamer insert. A total of 100 clones were screened, of which 79 clones were insert positive.

All the 79 insert clones were sequenced of which sequencing was successful for 57 of the clones. Aptamer clones were arbitrarily numbered, with the prefix G, denoting the EsxG target protein. Of the 57 successful sequences, three clones (G23, G49 and G68) had a random region of less than 50 nucleotides, while one clone (G26) had three additional nucleotides inserted in the random region resulting in the random region of 53 bases. Five clones had identical sequences (G31 and G34; G36 and G71; G43 and G94; G70 and G76; G78 and G98), and the rest had no detectable similarity to each other (Figure 4. 5).

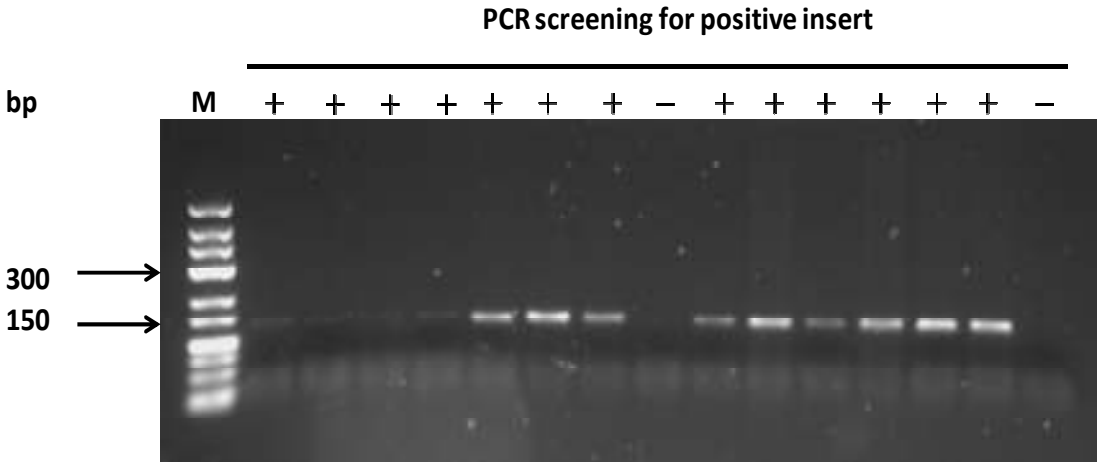


Figure 4. 4: PCR screening for insert positive clones

Agarose gel (2%) representing the PCR screening for insert–positive clones that contain monoclonal aptamers. Positive clones (+) are indicated by the presence of the expected150 base pair double stranded DNA. The clones without insert are presented by the minus sign (-). One hundred clones were screened, of which 79 clones were insert positive.

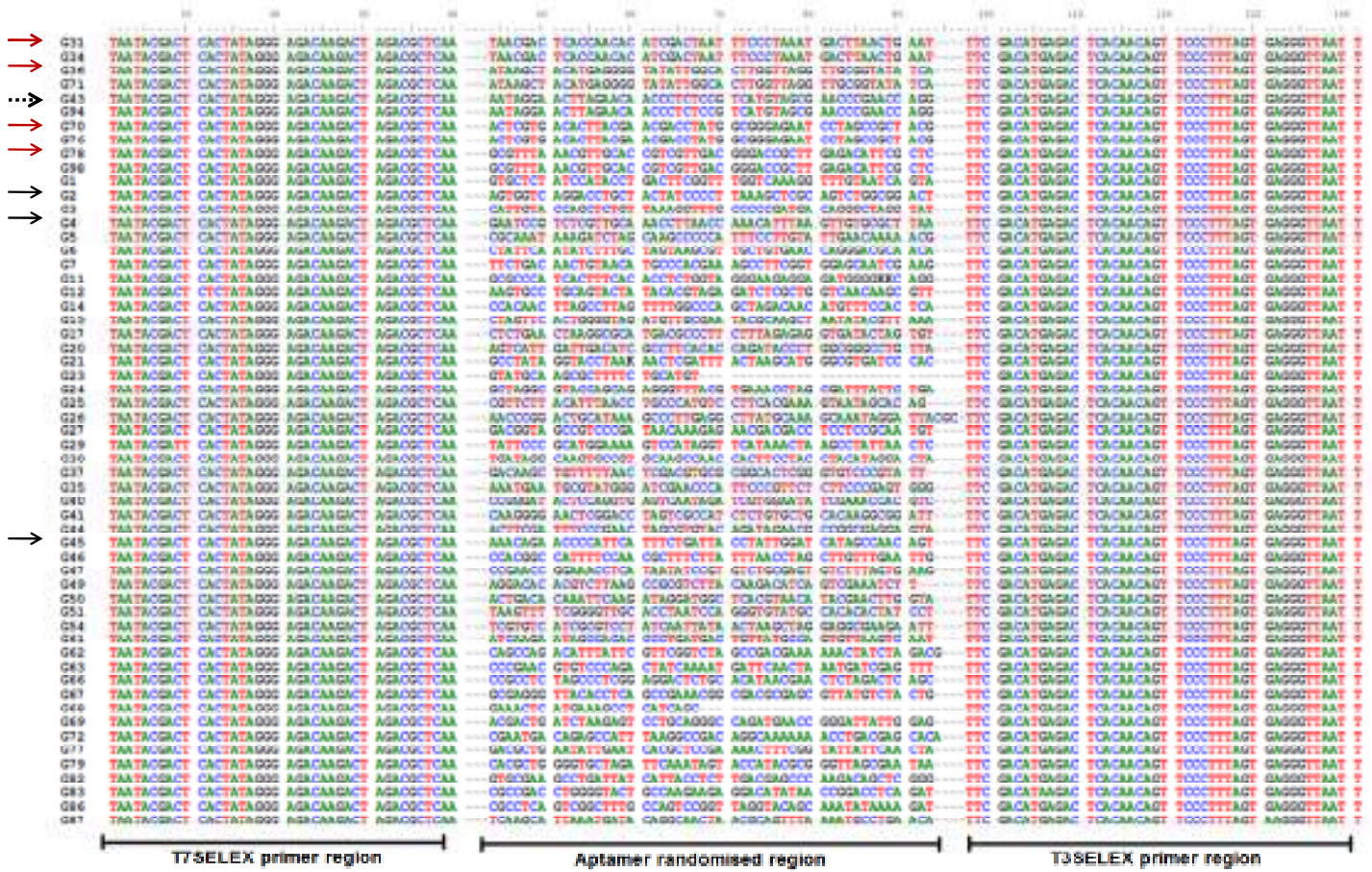


Figure 4. 5: Sequences of the monoclonal aptamers from the final round of selection.

Monoclonal aptamerswere randomly sequenced from the final round of selection. The sequences that were represented more than once are indicated by a red arrow, while the sequences that showed binding higher than the control during screening on the NTA system are indicated by a black arrow.G43 which was present in both groups is indicated by dotted arrow

3.5 Screening of monoclonal aptamers binding to the EsxG

Individual aptamer clones were screened against recombinant EsxG. The initial screening of the monoclonal aptamers was performed on the NTA chip to maintain a similar orientation of the recombinant protein to that usedduring SELEX. However, the intrinsic deficiency of the NTA system is the slow and continuous dissociation of the immobilized protein. To overcome this limitation a modified protocol for the immobilization of protein on NTA sensor chips using His_6 tag capture followed by free amine coupling was used as described by Willard and Siderovski (2006). Figure 4. 6A depicts a sensogram displaying this protocol.

Only four of the screened individual aptamer clones, namely G2, G4, G43 and G45, had a resonance signal higher than the buffer control after subtraction of the immobilization matrix flow cell without prior EsxG mobilization (Figure 4. 6B). Of these aptamers, G43 was identified twice during sequencing. However, the response units obtained using this screening method was generally low, with noaptamers having a response higher than 50 response units (RU).

Our rational was, therefore, to repeat the screening using amine coupling chemistry on the CM5 chip which results in a more stable immobilization of the protein on the biosensor surface. However, the CM5 is limited in the number of regenerations that can be performed. For this reason, we tested only clones that were identified more than once during sequencing as well as the three additional clones that gave a resonance signal higher than the buffer control during screening on the NTA chip. Interestingly, the results indicated that four of the five pairs that occurred more than

once during sequencing bind EsxG when amine coupled to the CM5 chip (Figure 4. 7A-E). However, out of the four clones that showed a binding response higher than that of the buffer control during NTA screening; only G43 which was also duplicated during sequencing, bound EsxG when amine coupled to the sensor chip surface (Figure 4. 7E-H).

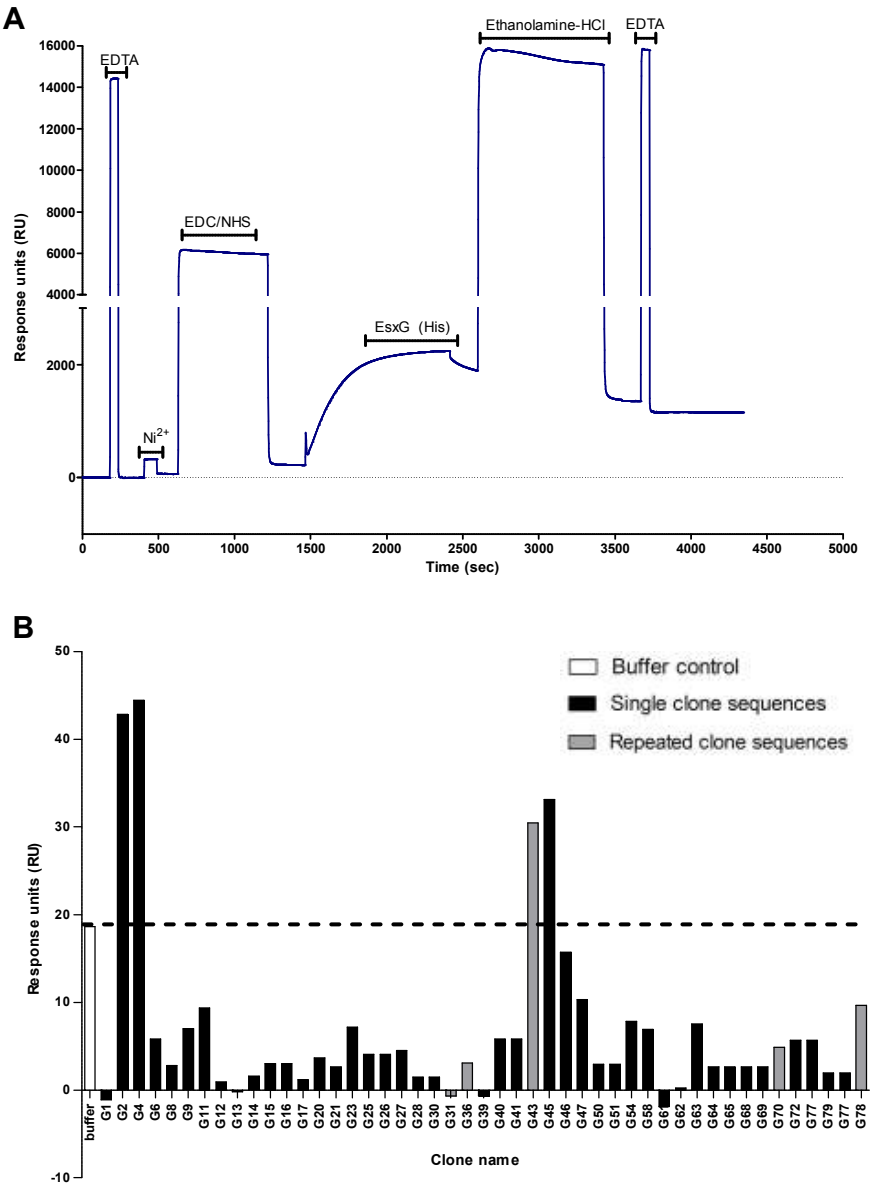


Figure 4. 6: Screening of the isolated clones on the NTA system

(A) Nickel was bound to an NTA sensorchip by the injection of nickel solution (500µM $NiCl_2$ in HBS-P, 50µM EDTA). EsxG (200 nM) in running buffer was captured on the biosensor surface by injection at a low flow rate (30 µl, 2µl/min) before binding of monoclonal RNA (B) After five rounds of selection, the monoclonal 2'F-RNAs were analysed for binding to covalently bound EsxGon a NTA chip. The clones that were presented more than twice on sequencing are labelled grey, while the buffer control is labelled white. The response units corresponding to the binding signal were derived after subtraction of measurements from the control flow cell.

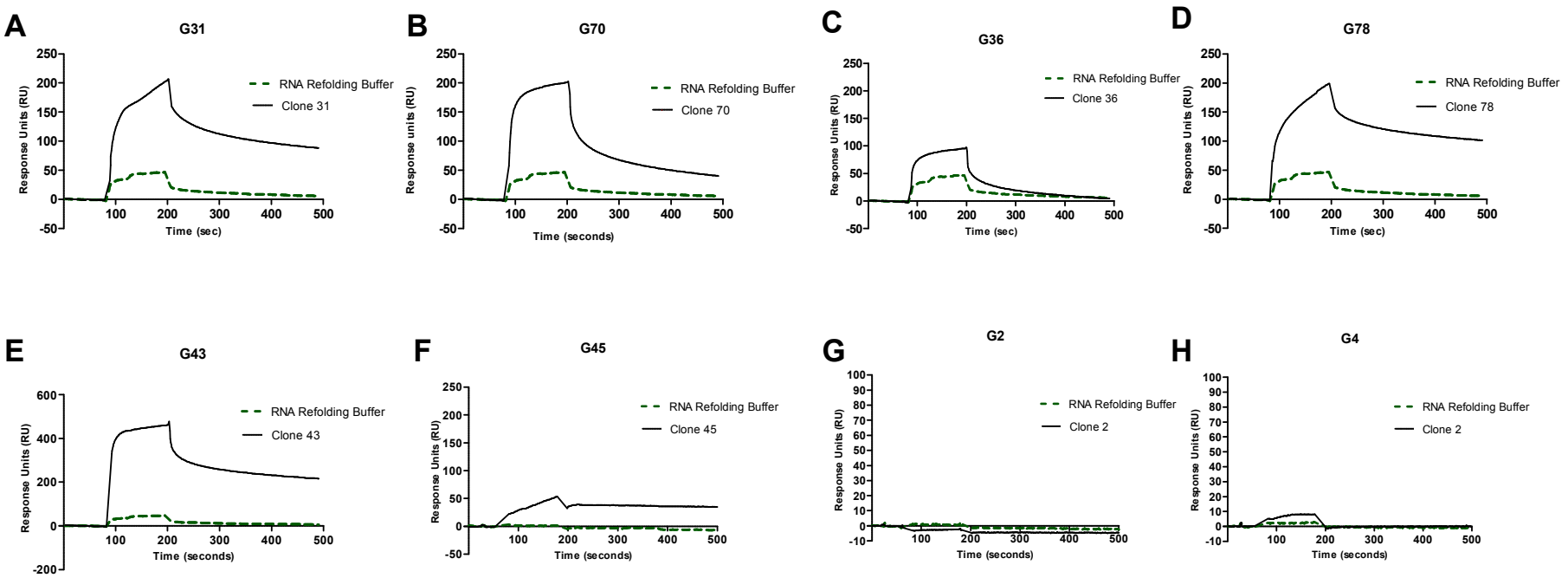


Figure 4. 7: Screening for binding of selected aptamers on the CM5 chip

Surface plasmon resonance sensorgram for EsxG interactions with (A-E) representative aptamers of the pairs that were duplicated during sequencing and (E-H) aptamers that showed abinding response higher than that of the buffer control during NTA screening. The thick black lines represent the aptamer binding to EsxG, and the dotted lines the buffer control.

3.6 Characterisation of aptamers to EsxG

Two aptamers, G43 and G78, with the highest response units during screening (Figure 4. 7D-E) were selected for further characterization.

3.6.1 G42 and G78 secondary structure prediction

With stranded RNA aptamers, intermolecular base pairing between complimentary base pairs results in the secondary structure of RNA. Depending on the environmental conditions, a given RNA sequence may exist in multiple structures in solution. We used the mfold web server (Zuker, 2003) to predict the minimum free energy (MFE) structures for each of the selected aptamers. Aptamer G43 was predicted to have a single MFE structure (Figure 4. 8A) with a Gibbs free energy of -27.07 kcal/mol. In contrast, aptamer G78 was predicted to have nine MFE structures (Appendix figure 2-A-H) with the lowest energy difference of 9.44kcal/mol. Since the energy difference between these structures is small, it is predicted that these structures simultaneously exist in solution.

The structures with the lowest MFEfor each selected aptamer are shown in Figure 4. 8A-B. There were no apparentsimilarities between the MFE structures of aptamer G43 and aptamerG78, as predicted by mfold.

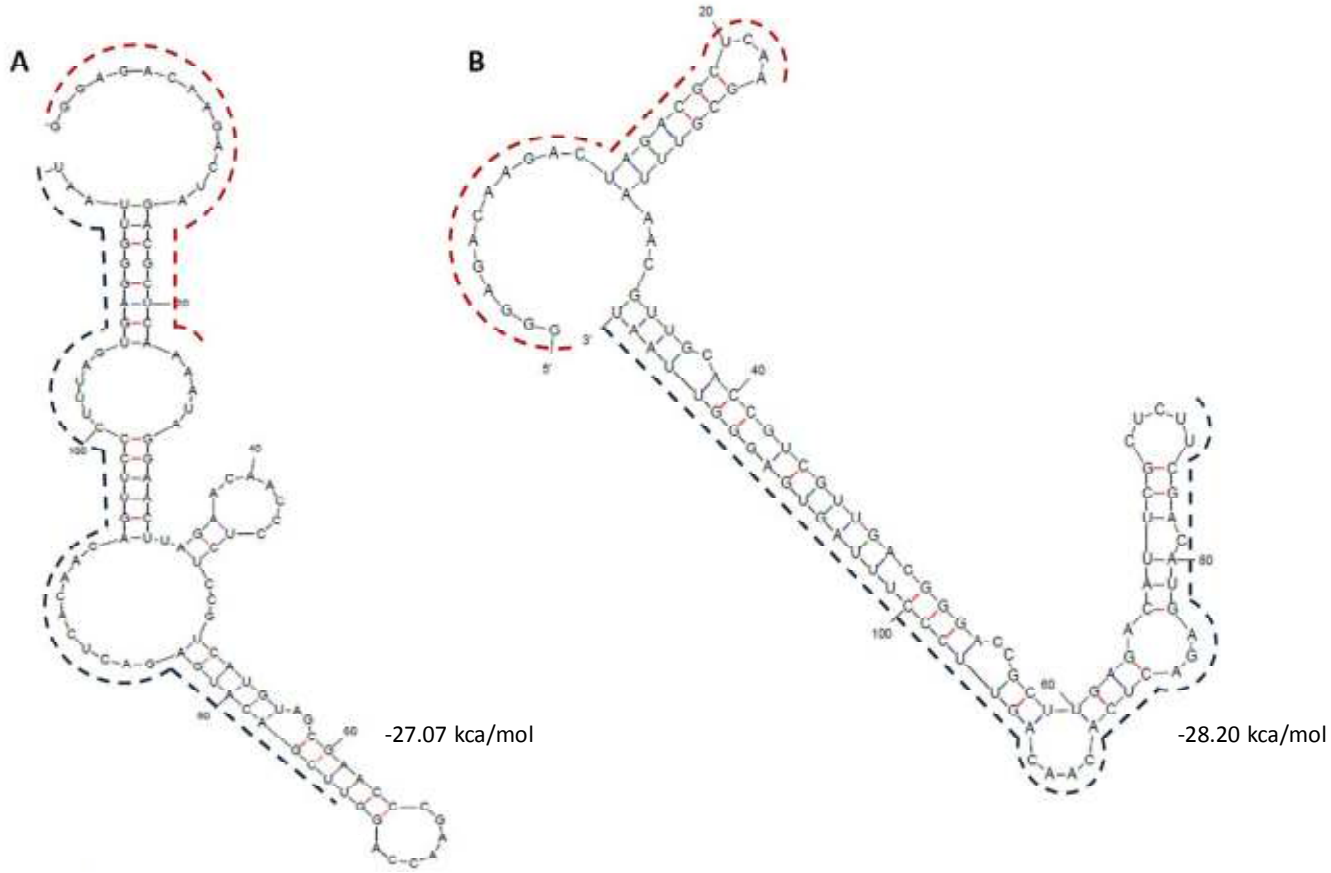


Figure 4. 8: Structures of anti-EsxG aptamers G43 and G78

The structures and free energies of aptamer (A) G43 (B) G78 were predicted by mfold. Fixed primer sequences corresponding to the T3SELEX5' and T7SELEX3' primer regions are marked by red and navy dotted line, respectively.

3.6.2 Binding kinetics of aptamers to EsxG

Binding kinetics to EsxG of the two selected aptamers was determined using SPR. EsxG was immobilized on the CM5 biacore chip using amine coupling as described in section 2.14. The concentration of the aptamer was measured using the NanoDrop spectrophotometer, and a dilution series (5nM – 500 nM) was prepared in HMCKN refolding buffer. Each concentration of the analyte was injected at 10µl/min and allowed to dissociate over 600 seconds. The affinity constant was calculated based on a Langmuir one on one curve fit using the BIAevalaution 3.0. Representative curves of the kinetic analysis obtained from the biacore are shown in Figure 4. 9A-B. The chi^2values for all the kineticanalysis were between 4.2 and 9.8. Typically the chi^2 less than 10 is indicative of good fitting to the model used, since our chi^2 values are less than 10 this means that the Langmuirone model used for our aptamer kinetics study adequately describes our data. The dissociation constants of the G43 and G78 aptamers were found to be about 8.04 ±1.90 and 78.85 ± 9.40 nM, respectively. Overall, aptamer G43 bound EsxG with about 7 times higher affinity than aptamer G78.

3.6.3 Binding specificity of aptamers to EsxG

To determine if the aptamers G43 and G78 are able to discriminate their binding to EsxG from its closely related protein, EsxA, I characterized binding using the biacore. The ligand, EsxG or EsxA was immobilized in a similar manner as during kinetics studies as described in (2.14) and 500 nM of each aptamer was injected over the immobilized EsxG. A two-tailed, unpaired t-test was used to analyse significance in the difference between binding of the aptamers to EsxG and EsxA. Response units measurements taken 100 seconds after aptamer injection has ended,

showed that both aptamers had significantly higher response units G43 ($p < 0.05$), G78 ($p < 0.01$) when binding to 2 µM of immobilized EsxG as compared to their binding to the immobilized, equimolar concentration of EsxA (Figure 4. 9C).

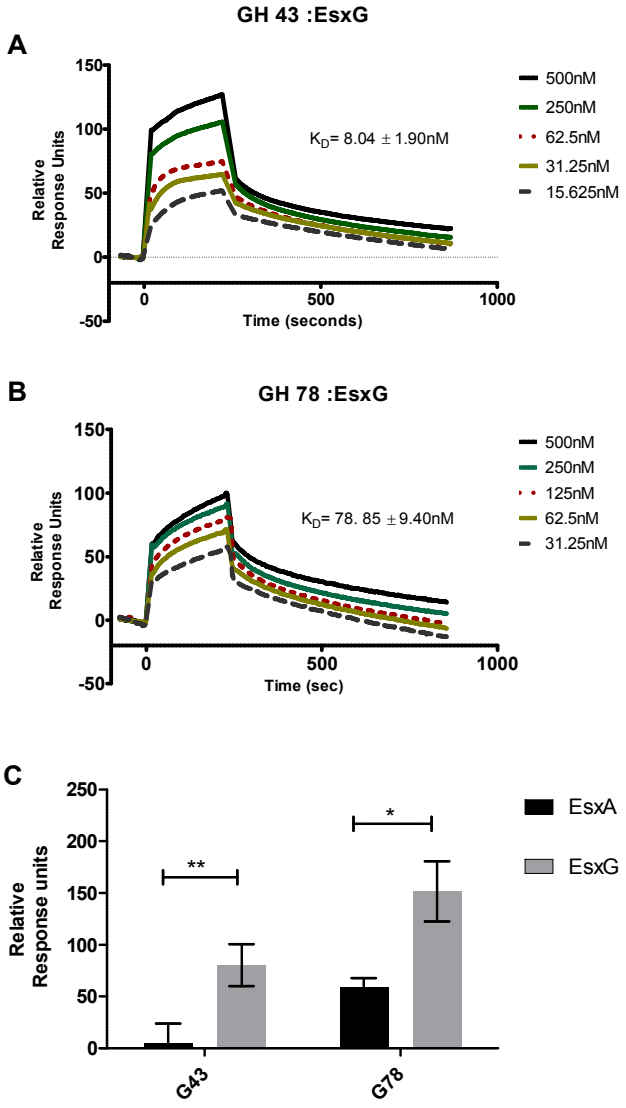


Figure 4. 9: Binding kinetics and specificity of EsxGaptamers G43 and G78

Equilibrium kinetics of RNA aptamers (A) G43 and (B) G78 binding to immobilized EsxG using the BIAcore™3000. Binding kinetic parameters were obtained using various concentrations of aptamer between (0-500 nM). The graphsare a representative and standard deviations were calculated from three experiments. (C) RNA-aptamers discriminate amongst closely related proteins, EsxG and EsxA by biacore. Relative response units were taken 100 seconds post

injection of 500 nM G43 and G78 binding to immobilized EsxG. A two-tailed, unpaired t-test was used to analyse significance (**p<0.01, *p<0.05).

3.6.4 G43 and G78 does not retard mycobacteria growth under low iron conditions

Certain aptamers are known to inhibit function of their target proteins. To exploit this aptamer feature, we used the EsxG aptamers G43 and G78 to investigate whether interfering with EsxG would interfere with the optimal uptake of iron. Siegrist and colleagues (2009) have previously shown that the *M. smegmatis fxbA*- mutant (which has an insertion in the formyltransferase required for exochelin synthesis) shows a marked growth defect under conditions of iron starvation. They further demonstrated that the double mutant strain *fxbA*-ΔESX-3 fails to grow in low iron medium.

In order to test whether EsxG contributed to optimal uptake of iron, we added 100 nM and 500 nM of the selected aptamers to a *M. smegmatis fxbA*- culture growing in low iron medium and compared it to wild-type. We hypothesised that, if the aptamers inhibit function of EsxG and EsxG is indeed a soluble factor that is secreted by the ESX-3 secretion sytem for optimal iron uptake, the addition of the aptamers to the *fxbA*- mutant will have an additive increase in the growth inhibition of this culture in low iron culture medium.

Our results were in agreement to the previous findings published by Siegrist and colleagues (2009). That is, the *fxbA*- mutants shows a marked growth defect, while

the *fxbA*$^{-}$ ΔESX-3 mutant fails to grow in low iron culture medium. However, the addition of aptamers G43 or G78 to *fxbA*- mutant culture did not show an additive growth reduction under low iron conditions (Figure 4. 10A-B). The initial growth delay observed from the cultures with the aptamers is most likely an effect of the aptamer refolding buffer (HMCKN) as a similar trend is observed from the *fxbA*$^{-}$ HMCKN control in the absence of the aptamer. Similar experiments using aptamer G31 and G70 (Appendix figure 3) showed similar results.

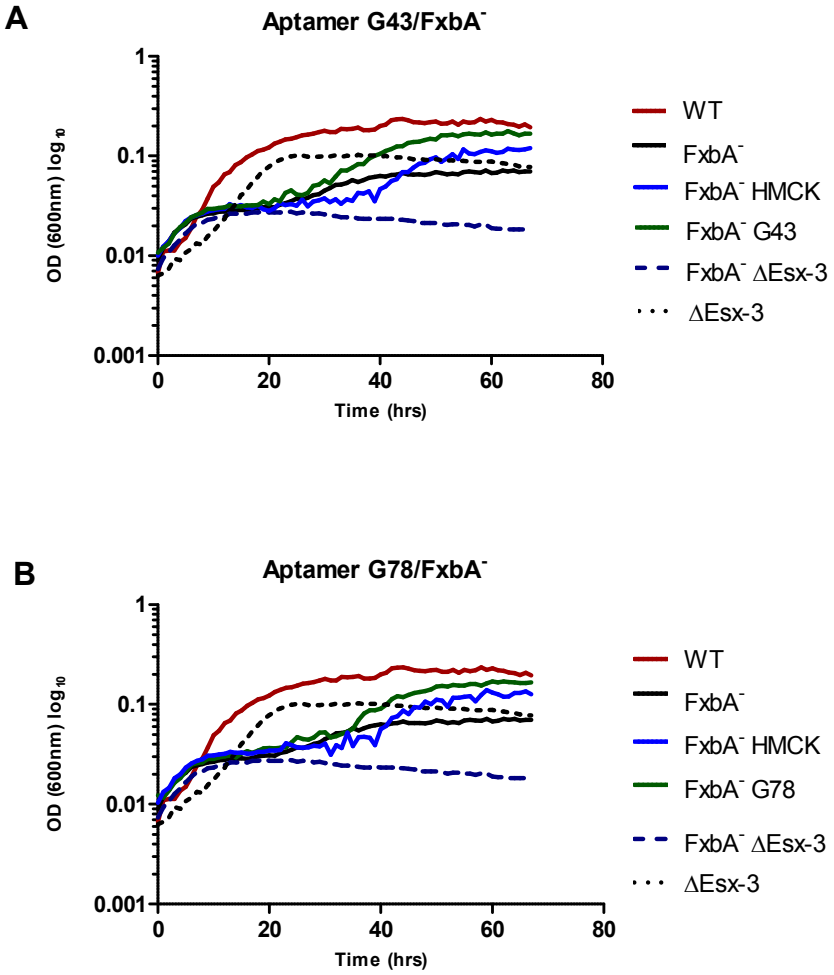


Figure 4. 10: Effect of G43 and G78 on mycobacterialgrowth under low iron conditions.

The growth of wild-type (WT),exochelininsertional biosynthesis mutant (*fxbA*$^-$), ΔESX-3 mutant, a double mutant (*fxbA*$^-$ΔESX-3) and RNA refolding buffer control (*fxbA*$^-$ HMCK) in low iron Sauton's medium. The aptamer experimental growth curves were performed on the *fxbA*$^-$background with aptamer G43(A) G78 (B) respectively.in low iron Sauton's medium.

4.0 Discussion

Antibody based diagnostic methods are still considered the standard in clinical diagnostics. However, aptamers have been explored as alternative recognition molecules for diagnostics as in the case for the detection of thrombin in thrombin related diseases in human blood (Centi et al., 2007, Heyduk and Heyduk, 2005, Xiao et al., 2005). One of their main advantages is in their ease of synthesis that does not require animals or cell lines, making them cheaper to produce when compared to antibodies (Osborne et al., 1997).

In this study, we performed asurface plasmon resonance based SELEX, to select RNA aptamers that bind to the ESX-3 secreted protein, EsxG. Our data shows that after five rounds of selection, we selected at least four RNA sequences that bind to EsxG (Figure 4.7). This result is similar to that of an earlier study that identified RNA aptamers against the gp120 protein of HIV from the same random library after only five rounds of selection on the biacore (Khati et al., 2003). Two of the aptamers selected in this study, G43 and G78, were analysed for equilibrium kinetics and the K_D was found to be in the nanomolar concentration range. Aptamer G43 had a K_D of 8.04 ± 1.90 nM while G78 had a K_D of 78.85 ± 9.40 nM.

M. tuberculosis has five homologous T7S system (Gey Van Pittius et al., 2001, Tekaia et al., 1999). ESX-1 is the most studied of these systems and is implicated in virulence of *M. tuberculosis* (Lewis et al., 2003, Pym et al., 2002). The ESX-1 secretion system secretes the EsxA:EsxB heterodimer which is homologous to the

EsxG:EsxH dimer which is secreted by the ESX-3 system. To investigate whether the selected aptamers can discriminate amongst closely related proteins, we evaluated the binding of these aptamers to EsxA. Our data showed that both aptamers bind to EsxG significantly more than to its homologue EsxA. Taken together, this data demonstrate that aptamers G43 and G78 bind their target with affinities in the lower nanomolar concentration ranges and will likely show minimal cross reactivity with other homologous proteins secreted by other *M. tuberculosis* T7S system. Thus these will make suitable recognition probes to investigate whether EsxG can be used as a biomarker for tuberculosis.

It is known that aptamers can bind proteins and in some cases inhibit function. For example, aptamers that bind to haemagglutinin on the influenza virus also inhibit haemagglutinin-mediated membrane fusion (Gopinath et al., 2006) and the aptamers raised against the HIV gp120 protein inhibit HIV entry (Khati et al., 2003). However, the functions of EsxG or EsxG:EsxH are unknown, as a result there is no functional assay to ascertain whether the selected aptamer inhibit their function. However as an initial screen to determine if the selected aptamers interfere with protein function we used an indirect system to determine if the aptamers bind to EsxG of *M. smegmatis* and interfere with the ESX-3 function of this organism. The *fxbA*- mutant (which has an insertion in the formyltransferase required for exochelin synthesis) of *M.smegmatis* was used as a background strain for these experiments, since it shows a marked growth defect under low iron conditions (Siegrist et al., 2009). Furthermore, a *M. smegmatis fxbA*-mutant, that also lacks ESX-3 (*fxbA*$^-$ΔESX-3) fails to grow in low iron medium. We evaluated the effect caused by the addition of the selected aptamers on the growth of the *fxbA*-

mutant in low iron medium. Our findings show that both aptamer G43 and G78 had no additive growth defect on the *fxbA*- mutant. This was a preliminary experiment to illustrate how aptamers might be used to probe the function of specific *M. tuberculosis* proteins or systems. It may have given a negative result because the aptamers were selected against the *M.tuberculosis* EsxG which is about 72% similar to that of *M. smegmatis*. Alternatively the aptamers may bind to EsxG without loss of function or finally that EsxG is not required for iron uptake. Additional experiments will be required to elucidate this.

In conclusion, using surface plasmon resonance based SELEX, we have isolated strongly binding RNA aptamers after only five rounds of selection. Our results also indicate that the selected RNA aptamers are able to distinguish between closely related proteins of *M. tuberculosis* namely, EsxA and EsxG. The availability of these aptamers allows for possibly investigating whether EsxG can be used as a biomarker for tuberculosis diagnostics. I n the case where a functional assay is available to elucidate a function of protein, the ability of the aptamers to discriminate amongst homologous proteins can be exploited to elucidate the function of closely related proteins in mycobacteria.

CHAPTER 5

GENERAL DISCUSSION AND CONCLUSION

SUMMARY OF FINDINGS

The primary aim of this study was to develop reagents that can detect *M. tuberculosis* with a view toward exploiting them as alternative molecular probes in the development of TB diagnostic tests. I used two approaches: the selection of phage displayed peptides against the cell surface associated targets of mycobacteria and the selection of nucleic acid reagents, aptamers, against the *M. tuberculosis* secreted protein.

Using two libraries of phage display technology (X_{12} and CX_7C) and choosing random clones, I identified nine distinct phage displayed peptides that bind to *M. tuberculosis* with varying binding intensities. Although choosing random clones is traditionally used to identify peptides enriched during biopanning, I found that the most enriched phage clone (phage 1, displaying CPLHARLPC peptide) that resulted from the panning of CX_7C library was not identified using this method. Instead this peptide was only discovered using HTP Illumina sequencing, which revealed that after three rounds of selection more than 80% of the enriched population, was made up of phages displaying the CPLHARLPC peptide. However, this observation is limited by the small sample number, ten phage clones, used during random selection. It is possible, that if a bigger sample size was selected when choosing random clones, this clone might have been found. However, to my knowledge there is no study that has investigated the minimum sample size for evaluating enrichment using phage display technology. In fact, considering the phage display library sizes, which are in the order of 10^9, the variability of the targets and selection conditions used during biopanning, HTP

sequencing is becoming a more appropriate method for evaluating enrichment (Buller et al., 2010, Dias-Neto et al., 2009, Mannocci et al., 2008).

During further characterisation, the phage clone displaying peptide CPLHARLPC was shown to have on average a 2.5 fold increase in binding to *M. tuberculosis* when compared to the phage library before selection. However, this phage clone also binds other species of mycobacteria, namely *M. smegmatis* and BCG (Figure 3.3). This observation was similar for all phages selected during this study (Figure 3.1). This is likely due to the similiar component structures that make up the mycobacteria cell wall and are conserved across species. However, there are components on the surface of *M. tuberculosis* envelope that confer species specificity, like the mycolic acids (Shui et al., 2012) and some families of lipids such as phthiocerol diesters and phenolic glycolipids (Daffe and Draper, 1998, Rastogi and N, 1991). However, it is likely that the experimental conditions used during biopanning for this study did not favour the selection of phage clones that bind to these components. Although the enriched peptide indiscriminately binds intact mycobacteria and whole cell lysates (WCL), I found that it does not bind to lysates from other bacterial species (Figure 3.4). Using an SPR binding assay, I showed that the association of the synthesised CPLHARLPC peptide to the *M. tuberculosis* WCL is abrogated after protease digestion, indicating that the *M. tuberculosis* target for phage 1 binding is likely to be a protein (Figure 3.5). This data was collaborated by a pull-down assay which indicated that this phage binds a *M. tuberculosis* protein of approximately 15 kDa (Figure 3.6). Using mass spectrometry we further identified five potential *M. tuberculosis* proteins as possible targets for phage 1. These are listed in table 3.2. While three of these

proteins are hypothetical proteins and there is no literature available for their localisation, one of the proteins, probable ribosomal protein S11, is hypothesised by homology to be membrane associated, while the other, probable ribosomal protein L11 rplK, has been experimentally validated to be membrane associated (Mawuenyega et al., 2005). If these are truly ribosome-associated proteins, however, they would likely be found on the inner aspect of the cell membrane and would not be accessible to endogenous peptides. Other proteins on this list have been found to be associated with the cell wall but their topology is unknown. Further testingwill be required to confirm the target for phage 1 on *M. tuberculosis*.

Using SELEX and SPR technology, I isolated five nucleic acid aptamers that bind EsxG, a protein secreted by the ESX-3 secretion system of *M. tuberculosis* (Lightbody et al., 2004, Siegrist et al., 2009). Two of the aptamers, GH 43 and GH 78, were further evaluated for binding and showed binding affinity constants (K_D) in the lower nanomolar ranges, 8.04 ± 1.90 nM and 78.85 ± 9.4 nM respectively. Furthermore, since aptamers are known to inhibit function of their ligands (Gopinath et al., 2006), we tested whether EsxG aptamerswould inhibit the function of EsxG to investigate whether EsxG is the ESX-3 secreted substrate that confers the growth defect associated with the ΔESX-3 mutant under low iron conditions. However, the absence of the growth defect when these aptamers were added indicates that either EsxG alone is not responsible for this phenotype or that while these aptamers bind EsxG they do not affect its function associated with *M. tuberculosis* iron acquisition. Further experiments that exploit these aptamers will be required to differentiate between these hypotheses.

IMPLICATIONS, FUTURE DIRECTION OF THE SUDY AND CONCLUSION

Continuous improvements have been achieved in the development of TB diagnosis tests. Indeed, new, automated and more sensitive tests like the GeneXpert that can also detect drug resistance TB are now in use (Evans, 2011, Vadwai et al., 2011, Van Rie et al., 2010, van Zyl-Smit et al., 2011). However, despite these improvements, simple and inexpensive POC tests are still unavailable. One of the challenges that hinder the rapid development of TB tests is that currently there is no biomarker, host or pathogen associated, that offers 100% specificity and sensitivity for diagnosing TB. An ideal marker should also be able to differentiate between active and latent TB infection. Moreover, discovery of biomarkers is only the first stage in the development of a diagnostic test. This is usually followed by extended periods of research to develop the technology or detection reagents to use as a diagnostic test platform.

The approach adopted in this study demonstrates a shorter path to the development of a TB test. I found that specific reagents that target the cell envelope can be developed using phage display technology. The reagents I identified had genus- but not species-level specificity. Other biopanning conditions might favour selection of species-specific phage displayed peptides. These could include panning experiments targeting *M. tuberculosis* cultures grown in the absence of detergent and agitation, two *in vitro* manipulations that do not exist in clinical samples. While the selected phage population from this study showed similar binding of the phage pool to *M. tuberculosis* grown under both conditions, the

enhanced binding to BCG and *M. smegmatis* indicates species-differences in abundance of the target under these culturing conditions. These differences can be exploited by performing biopanning on cultured mycobacteria grown under different conditions. This approach can potentially identify a group of phage-display probes that can be used together to develop a more specific diagnostic test for TB. In addition, while western blot assays to validate the biotinylated CPLHARLPC peptide were unsuccessful, other techniques, such as 2-D gel electrophoresis to further separate and differentiate proteins of similar molecular weight combined with protein identification using mass spectrometry might prove more successful. Furthermore, evaluation of the selected probes on clinical samples will be required to assess its potential for use in diagnostics.

Previously, the low abundance of these Esx proteins in culture caused a concern in their application in diagnostics. However, whilst I was doing my work results were obtained for EsxA, which is homologous to EsxG, which demonstrated the protein isa potential candidate in the diagnosis of TB. Rotherham and colleagues in our lab showed that DNA aptamers raised against the homologous ESXA protein are able to detect infection in sputum samples with a sensitivity of 100% and a specificity of 68.75% (Rotherham et al., 2012). Furthermore, in the preceding year, Hong and colleagues (2011) also demonstrated the presence of the same antigen, EsxA, in urine samples (Hong et al., 2011). Taken together, this indicates that, in spite of the low abundance of these proteins in culture, they can be detected in clinical samples from patients. However, further clinical studies on patients samples are required to validate the potential application of these reagents in TB diagnostics.

Long term, these reagents can be developed into molecular probes that can be used in clinical diagnostics. Phage displayed peptides can be used either as intact phage or as peptides coupled to other support material like nanoparticles. Numerous aptamer-based POC technologies that use platforms like biosensors (Navani and Li, 2006), gold nano particles (Liu and Lu, 2006), and lateral flow device (similar to home pregnancy testing) (Liu et al., 2006) could make use of detection reagents to enable the development of a TB diagnostic test that meats the ASSURED criteria. Thus, this approach offers a clear path from discovering detection reagents to a clinically useful test.

REFERENCES

'T HOEN, P. A. C., JIRKA, S. M. G., TEN BROEKE, B. R., SCHULTES, E. A., AGUILERA, B. A., PANG, K. H., HEEMSKERK, H., AARTSMA-RUS, A., VAN OMMEN, G. J. & DEN DUNNEN, J. T. (2011) Phage display screening without repetitious selection rounds. *Analytical Biochemistry,* 421**,** 622.

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. *Nat Rev Microbiol,* 5**,** 883-91.

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. *Mol Microbiol,* 62**,** 667-79.

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. *Mol Microbiol,* 73**,** 329-40.

ABE, C., HOSOJIMA, S., FUKASAWA, Y., KAZUMI, Y., TAKAHASHI, M., HIRANO, K. & MORI, T. (1992) Comparison of MB-Check, BACTEC, and egg-based media for recovery of mycobacteria. *Journal of Clinical Microbiology,* 30**,** 878-881.

ADDA, C. G., ANDERS, R. F., TILLEY, L. & FOLEY, M. (2002) Random sequence libraries displayed on phage: identification of biologically important molecules. *Comb Chem High Throughput Screen,* 5**,** 1-14.

ALBERT, H., TROLLIP, A., SEAMAN, T. & MOLE, R. J. (2004) Simple, phage-based (FASTPlaque) technology to determine rifampicin resistance of Mycobacterium tuberculosis directly from sputum. *The International Journal of Tuberculosis and Lung Disease,* 8**,** 1114-1119.

ALTEKRUSE, S. F., STERN, N. J., FIELDS, P. I. & SWERDLOW, D. L. (1999) Campylobacter jejuni--an emerging foodborne pathogen. *Emerg Infect Dis,* 5**,** 28-35.

AMERICAN THORACIC SOCIETY. COMMITTEE ON REVISION OF DIAGNOSTIC, S. (2000) *Diagnostic standards and classification of tuberculosis in adults and children*.

ANANDAKUMAR, S., BOOSI, K. N., BUGATHA, H., PADMANABHAN, B. & SADHALE, P. P. (2011) Phage displayed short peptides against cells of Candida albicans demonstrate presence of species, morphology and region specific carbohydrate epitopes. *PLoS One,* 6**,** e16868.

AZZAZY, H. M. & HIGHSMITH, W. E., JR. (2002) Phage display technology: clinical applications and recent innovations. *Clin Biochem,* 35**,** 425-45.

BALDRICH, E., RESTREPO, A. & O'SULLIVAN, C. K. (2004) Aptasensor development: elucidation of critical parameters for optimal aptamer performance. *Anal Chem,* 76**,** 7053-63.

BELL, W. J. & BROWN, P. P. (1962) Fluorescence microscopy in the laboratory diagnosis and assessment of pulmonary tuberculosis. *The Central African journal of medicine,* 8**,** 4.

BIOLABS, N. E. (2012) Ph.D. Phage Display Libraries Instruction Manual. IN 1.1, V. (Ed.) *New england Biolabs Inc.*

BIRCH, J. R., MAINWARING, D. O. & RACHER, A. J. (2008) Use of the Glutamine Synthetase (GS) Expression System for the Rapid Development of Highly Productive Mammalian Cell Processes. *Modern Biopharmaceuticals.* Wiley-VCH Verlag GmbH.

BOCK, L. C., GRIFFIN, L. C., LATHAM, J. A., VERMAAS, E. H. & TOOLE, J. J. (1992) Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature,* 355**,** 564-6.

BOWSER, M. T. (2005) SELEX: just another separation? *Analyst,* 130**,** 128-30.

BRAUNSTEIN, M., ESPINOSA, B. J., CHAN, J., BELISLE, J. T. & JACOBS, W. R., JR. (2003) SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. *Mol Microbiol,* 48**,** 453-64.

BRENNAN, P. J. (2003a) Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis,* 83**,** 91.

BRENNAN, P. J. (2003b) Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis (Edinb),* 83**,** 91-7.

BRENNAN, P. J. & CRICK, D. C. (2007) The cell-wall core of Mycobacterium tuberculosis in the context of drug discovery. *Curr Top Med Chem,* 7**,** 475-88.

BROSCH, R., GORDON, S. V., MARMIESSE, M., BRODIN, P., BUCHRIESER, C., EIGLMEIER, K., GARNIER, T., GUTIERREZ, C., HEWINSON, G., KREMER, K., PARSONS, L. M., PYM, A. S., SAMPER, S., VAN SOOLINGEN, D. & COLE, S. T. (2002) A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Natl Acad Sci U S A,* 99**,** 3684-9.

BULLER, F., STEINER, M., SCHEUERMANN, J., MANNOCCI, L., NISSEN, I., KOHLER, M., BEISEL, C. & NERI, D. (2010) High-throughput sequencing for the identification of binding molecules from DNA-encoded chemical libraries. *Bioorg Med Chem Lett,* 20**,** 4188-92.

BURRITT, J. B., QUINN, M. T., JUTILA, M. A., BOND, C. W. & JESAITIS, A. J. (1995) Topological mapping of neutrophil cytochrome b epitopes with phage-display libraries. *J Biol Chem,* 270**,** 16974-80.

BUTLER, W. R. & GUTHERTZ, L. S. (2001) Mycolic acid analysis by high-performance liquid chromatography for identification of Mycobacterium species. *Clin Microbiol Rev,* 14**,** 704-26, table of contents.

BUTLER, W. R. & KILBURN, J. O. (1988) Identification of major slowly growing pathogenic mycobacteria and Mycobacterium gordonae by high-performance liquid chromatography of their mycolic acids. *J Clin Microbiol,* 26**,** 50-3.

BUTT, T., AHMAD, R. N., AFZAL, R. K., MAHMOOD, A. & ANWAR, M. (2004) Rapid detection of rifampicin susceptibility of Mycobacterium tuberculosis in sputum specimens by mycobacteriophage assay. *JOURNAL-PAKISTAN MEDICAL ASSOCIATION,* 54**,** 379-381.

CAI, H., XU, Y., ZHU, N., HE, P. & FANG, Y. (2002) An electrochemical DNA hybridization detection assay based on a silver nanoparticle label. *Analyst,* 127**,** 803-8.

CAO, Q., LIU, S., NIU, G., CHEN, K., YAN, Y., LIU, Z. & CHEN, X. (2011) Phage display peptide probes for imaging early response to bevacizumab treatment. *Amino Acids,* 41**,** 1103-12.

CARSON, D. A. & SEEGMILLER, J. E. (1976) Effect of adenosine deaminase inhibition upon human lymphocyte blastogenesis. *Journal of Clinical Investigation,* 57**,** 274.

CARTER, D. M., MIOUSSE, I. R., GAGNON, J. N., MARTINEZ, E., CLEMENTS, A., LEE, J., HANCOCK, M. A., GAGNON, H., PAWELEK, P. D. & COULTON, J. W. (2006) Interactions between TonB from Escherichia coli and the periplasmic protein FhuD. *J Biol Chem,* 281**,** 35413-24.

CENTI, S., TOMBELLI, S., MINUNNI, M. & MASCINI, M. (2007) Aptamer-based detection of plasma proteins by an electrochemical assay coupled to magnetic beads. *Anal Chem,* 79**,** 1466-73.

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,* 313**,** 1632-6.

CHAN, E. D., REVES, R., BELISLE, J. T., BRENNAN, P. J. & HAHN, W. E. (2000) Diagnosis of tuberculosis by a visually detectable immunoassay for lipoarabinomannan. *Am J Respir Crit Care Med,* 161**,** 1713-9.

CHEN, F., ZHANG, X., ZHOU, J., LIU, S. & LIU, J. (2012) Aptamer inhibits Mycobacterium tuberculosis (H37Rv) invasion of macrophage. *Molecular Biology Reports,* 39**,** 2157-2162.

CHEN, F., ZHOU, J., LUO, F., MOHAMMED, A. B. & ZHANG, X. L. (2007) Aptamer from whole-bacterium SELEX as new therapeutic reagent against virulent Mycobacterium tuberculosis. *Biochem Biophys Res Commun,* 357**,** 743-8.

CHEN, Y. C., DELBROOK, K., DEALWIS, C., MIMMS, L., MUSHAHWAR, I. K. & MANDECKI, W. (1996) Discontinuous epitopes of hepatitis B surface antigen derived from a filamentous phage peptide library. *Proc Natl Acad Sci U S A,* 93**,** 1997-2001.

COLE, S. T., BROSCH, R., PARKHILL, J., GARNIER, T., CHURCHER, C., HARRIS, D., GORDON, S. V., EIGLMEIER, K., GAS, S., BARRY, C. E., 3RD, TEKAIA, F., BADCOCK, K., BASHAM, D., BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R., DEVLIN, K., FELTWELL, T., GENTLES, S., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K., KROGH, A., MCLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., OSBORNE, J., QUAIL, M. A., RAJANDREAM, M. A., ROGERS, J., RUTTER, S., SEEGER, K., SKELTON, J., SQUARES, R., SQUARES, S., SULSTON, J. E., TAYLOR, K., WHITEHEAD, S. & BARRELL, B. G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature,* 393**,** 537-44.

COLE, S. T. E. K. P. J. J. K. D. T. N. R. W. P. R. H. N. G. T. C. C. H. D. (2001) Massive gene decay in the leprosy bacillus. *Nature,* 409**,** 1007.

CONNOLLY, P. A., DURKIN, M. M., LEMONTE, A. M., HACKETT, E. J. & WHEAT, L. J. (2007) Detection of histoplasma antigen by a quantitative enzyme immunoassay. *Clin Vaccine Immunol,* 14**,** 1587-91.

COOK, G. M., BERNEY, M., GEBHARD, S., HEINEMANN, M., COX, R. A., DANILCHANKA, O. & NIEDERWEIS, M. (2009) Physiology of mycobacteria. *Adv Microb Physiol,* 55**,** 81-182, 318-9.

COUSINS, D. V., BASTIDA, R., CATALDI, A., QUSE, V., REDROBE, S., DOW, S., DUIGNAN, P., MURRAY, A., DUPONT, C., AHMED, N., COLLINS, D. M., BUTLER, W. R., DAWSON, D., RODRIGUEZ, D., LOUREIRO, J., ROMANO, M. I., ALITO, A., ZUMARRAGA, M. & BERNARDELLI, A. (2003) Tuberculosis in seals caused by a novel member of the Mycobacterium tuberculosis complex: Mycobacterium pinnipedii sp. nov. *Int J Syst Evol Microbiol,* 53**,** 1305-14.

COUSINS, D. V., PEET, R. L., GAYNOR, W. T., WILLIAMS, S. N. & GOW, B. L. (1994) Tuberculosis in imported hyrax (Procavia capensis) caused by an unusual variant belonging to the Mycobacterium tuberculosis complex. *Vet Microbiol,* 42**,** 135-45.

CWIRLA, S. E., PETERS, E. A., BARRETT, R. W. & DOWER, W. J. (1990) Peptides on phage: a vast library of peptides for identifying ligands. *Proc Natl Acad Sci U S A,* 87**,** 6378-82.

DAFFE, M. & DRAPER, P. (1998) The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol,* 39**,** 131-203.

DAFFE, M. & ETIENNE, G. (1999) The capsule of Mycobacterium tuberculosis and its implications for pathogenicity. *Tuber Lung Dis,* 79**,** 153-69.

DALEKE, M. H., UMMELS, R., BAWONO, P., HERINGA, J., VANDENBROUCKE-GRAULS, C. M., LUIRINK, J. & BITTER, W. (2012) General secretion signal for the mycobacterial type VII secretion pathway. *Proc Natl Acad Sci U S A,* 109**,** 11342-7.

DANNENBERG, A. M., JR. (1994) Roles of cytotoxic delayed-type hypersensitivity and macrophage-activating cell-mediated immunity in the pathogenesis of tuberculosis. *Immunobiology,* 191**,** 461-73.

DAVIS, J. L., HUANG, L., WORODRIA, W., MASUR, H., CATTAMANCHI, A., HUBER, C., MILLER, C., CONVILLE, P. S., MURRAY, P. & KOVACS, J. A. (2011) Nucleic acid amplification tests for diagnosis of smear-negative TB in a high HIV-prevalence setting: a prospective cohort study. *PLoS One,* 6**,** e16321.

DE VOSS, J. J., RUTTER, K., SCHROEDER, B. G. & BARRY, C. E., 3RD (1999) Iron acquisition and metabolism by mycobacteria. *J Bacteriol,* 181**,** 4443-51.

DE VOSS, J. J., RUTTER, K., SCHROEDER, B. G., SU, H., ZHU, Y. & BARRY, C. E., 3RD (2000) The salicylate-derived mycobactin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages. *Proc Natl Acad Sci U S A,* 97**,** 1252-7.

DHEDA, K., DAVIDS, V., LENDERS, L., ROBERTS, T., MELDAU, R., LING, D., BRUNET, L., VAN ZYL SMIT, R., PETER, J., GREEN, C., BADRI, M., SECHI, L., SHARMA, S., HOELSCHER, M., DAWSON, R., WHITELAW, A., BLACKBURN, J., PAI, M. & ZUMLA, A. (2010) Clinical utility of a commercial LAM-ELISA assay for TB diagnosis in HIV-infected patients using urine and sputum samples. *PLoS One,* 5**,** e9848.

DIACON, A. H., VAN DE WAL, B. W., WYSER, C., SMEDEMA, J. P., BEZUIDENHOUT, J., BOLLIGER, C. T. & WALZL, G. (2003) Diagnostic tools in tuberculous pleurisy: a direct comparative study. *European Respiratory Journal,* 22**,** 589-591.

DIAS-NETO, E., NUNES, D. N., GIORDANO, R. J., SUN, J., BOTZ, G. H., YANG, K., SETUBAL, J. C., PASQUALINI, R. & ARAP, W. (2009) Next-generation phage display: integrating and comparing available molecular tools to enable cost-effective high-throughput analysis. *PLoS One,* 4**,** e8338.

DIAZ-INFANTES, M. S., RUIZ-SERRANO, M. J., MARTINEZ-SANCHEZ, L., ORTEGA, A. & BOUZA, E. (2000) Evaluation of the MB/BacT mycobacterium detection system for susceptibility testing of Mycobacterium tuberculosis. *J Clin Microbiol,* 38**,** 1988-9.

DICKERSON, M. B., JONES, S. E., CAI, Y., AHMAD, G., NAIK, R. R., KRÃ¶GER, N. & SANDHAGE, K. H. (2008) Identification and Design of Peptides for the Rapid, High-Yield Formation of Nanoparticulate TiO2 from Aqueous Solutions at Room Temperature. *Chemistry of Materials,* 20**,** 1578.

DOMINGUEZ, J., GALI, N., MATAS, L., PEDROSO, P., HERNANDEZ, A., PADILLA, E. & AUSINA, V. (1999) Evaluation of a rapid immunochromatographic assay for the detection of Legionella antigen in urine samples. *Eur J Clin Microbiol Infect Dis,* 18**,** 896-8.

EJIMA, H., MATSUNO, H. & SERIZAWA, T. (2010) Biological identification of peptides that specifically bind to poly(phenylene vinylene) surfaces: recognition of the branched or linear structure of the conjugated polymer. *Langmuir,* 26**,** 17278-85.

ELLINGTON, A. D. & SZOSTAK, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature,* 346**,** 818-22.

ESTEPHAN, E., LARROQUE, C., BEC, N., MARTINEAU, P., CUISINIER, F. J., CLOITRE, T. & GERGELY, C. (2009) Selection and mass spectrometry characterization of peptides targeting semiconductor surfaces. *Biotechnol Bioeng,* 104**,** 1121-31.

EVANS, C. A. (2011) GeneXpert--a game-changer for tuberculosis control? *PLoS Med,* 8**,** e1001064.

FEND, R., KOLK, A. H., BESSANT, C., BUIJTELS, P., KLATSER, P. R. & WOODMAN, A. C. (2006) Prospects for clinical application of electronic-nose technology to early detection of Mycobacterium tuberculosis in culture and sputum. *J Clin Microbiol,* 44**,** 2039-45.

FESTA, R. A., MCALLISTER, F., PEARCE, M. J., MINTSERIS, J., BURNS, K. E., GYGI, S. P. & DARWIN, K. H. (2010) Prokaryotic ubiquitin-like protein (Pup) proteome of Mycobacterium tuberculosis [corrected]. *PLoS One,* 5**,** e8589.

FISCHER, D., VAN DER WEYDEN, M. B., SNYDERMAN, R. & KELLEY, W. N. (1976) A role for adenosine deaminase in human monocyte maturation. *Journal of Clinical Investigation,* 58**,** 399.

FLORES, L. L., STEINGART, K. R., DENDUKURI, N., SCHILLER, I., MINION, J., PAI, M., RAMSAY, A., HENRY, M. & LAAL, S. (2011) Systematic review and meta-analysis of antigen detection tests for the diagnosis of tuberculosis. *Clin Vaccine Immunol,* 18**,** 1616-27.

FRANK, U. K., NISHIMURA, S. L., LI, N. C., SUGAI, K., YAJKO, D. M., HADLEY, W. K. & NG, V. L. (1993) Evaluation of an enzyme immunoassay for detection of cryptococcal capsular polysaccharide antigen in serum and cerebrospinal fluid. *J Clin Microbiol,* 31**,** 97-101.

GANGULY, N., SIDDIQUI, I. & SHARMA, P. (2008) Role of M. tuberculosis RD-1 region encoded secretory proteins in protective response and virulence. *Tuberculosis (Edinb),* 88**,** 510-7.

GANZ, T. (2009) Iron in innate immunity: starve the invaders. *Curr Opin Immunol,* 21**,** 63-7.

GEBHARDT, K., SHOKRAEI, A., BABAIE, E. & LINDQVIST, B. H. (2000) RNA aptamers to S-adenosylhomocysteine: kinetic properties, divalent cation dependency, and comparison with anti-S-adenosylhomocysteine antibody. *Biochemistry,* 39**,** 7255-65.

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 Biol,* 2**,** RESEARCH0044.

GITHUI, W., KITUI, F., JUMA, E. S., OBWANA, D. O., MWAI, J. & KWAMANGA, D. (1993) A comparative study on the reliability of the fluorescence microscopy and Ziehl-Neelsen method in the diagnosis of pulmonary tuberculosis. *East African Medical Journal,* 70**,** 263.

GOBIN, J. & HORWITZ, M. A. (1996) Exochelins of Mycobacterium tuberculosis remove iron from human iron-binding proteins and donate iron to mycobactins in the M. tuberculosis cell wall. *J Exp Med,* 183**,** 1527-32.

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. *Mol Microbiol,* 42**,** 851-65.

GOLD, L., POLISKY, B., UHLENBECK, O. & YARUS, M. (1995) Diversity of oligonucleotide functions. *Annu Rev Biochem,* 64**,** 763-97.

GOPINATH, S. C. (2007) Methods developed for SELEX. *Anal Bioanal Chem,* 387**,** 171-82.

GOPINATH, S. C., BALASUNDARESAN, D., AKITOMI, J. & MIZUNO, H. (2006) An RNA aptamer that discriminates bovine factor IX from human factor IX. *J Biochem,* 140**,** 667-76.

GRECO, S., GIRARDI, E., MASCIANGELO, R., CAPOCCETTA, G. B. & SALTINI, C. (2003) Adenosine deaminase and interferon gamma measurements for the diagnosis of tuberculous pleurisy: a meta-analysis. *The International Journal of Tuberculosis and Lung Disease,* 7**,** 777.

GRIFFIN, J. E., GAWRONSKI, J. D., DEJESUS, M. A., IOERGER, T. R., AKERLEY, B. J. & SASSETTI, C. M. (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog,* 7**,** e1002251.

GRONEWOLD, T. M., GLASS, S., QUANDT, E. & FAMULOK, M. (2005) Monitoring complex formation in the blood-coagulation cascade using aptamer-coated SAW sensors. *Biosens Bioelectron,* 20**,** 2044-52.

GU, S., CHEN, J., DOBOS, K. M., BRADBURY, E. M., BELISLE, J. T. & CHEN, X. (2003) Comprehensive Proteomic Profiling of the Membrane Constituents of a Mycobacterium tuberculosis Strain. *Molecular & Cellular Proteomics,* 2**,** 1284-1296.

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. *Mol Microbiol,* 51**,** 359-70.

HAMASUR, B., BRUCHFELD, J., HAILE, M., PAWLOWSKI, A., BJORVATN, B., KÃ¤LLENIUS, G. & SVENSON, S. B. (2001) Rapid diagnosis of tuberculosis by detection of mycobacterial lipoarabinomannan in urine. *Journal of microbiological methods,* 45**,** 41-52.

HARDY, B., RAITER, A., WEISS, C., KAPLAN, B., TENENBAUM, A. & BATTLER, A. (2007) Angiogenesis induced by novel peptides selected from a phage display library by screening human vascular endothelial cells under different physiological conditions. *Peptides,* 28**,** 691-701.

HARRISON, A. J., YU, M., GARDENBORG, T., MIDDLEDITCH, M., RAMSAY, R. J., BAKER, E. N. & LOTT, J. S. (2006) The structure of MbtI from Mycobacterium tuberculosis, the first enzyme in the biosynthesis of the siderophore mycobactin, reveals it to be a salicylate synthase. *J Bacteriol,* 188**,** 6081-91.

HEYDUK, E. & HEYDUK, T. (2005) Nucleic acid-based fluorescence sensors for detecting proteins. *Anal Chem,* 77**,** 1147-56.

HIRAO, I., MADIN, K., ENDO, Y., YOKOYAMA, S. & ELLINGTON, A. D. (2000) RNA aptamers that bind to and inhibit the ribosome-inactivating protein, pepocin. *J Biol Chem,* 275**,** 4943-8.

HONG, S. C., LEE, J., SHIN, H.-C., KIM, C.-M., PARK, J. Y., KOH, K., KIM, H.-J., CHANG, C. L. & LEE, J. (2011) Clinical immunosensing of tuberculosis CFP-10 in patient urine by surface plasmon resonance spectroscopy. *Sensors and Actuators B: Chemical,* 160**,** 1434.

HSU, T., HINGLEY-WILSON, S. M., CHEN, B., CHEN, M., DAI, A. Z., MORIN, P. M., MARKS, C. B., PADIYAR, J., GOULDING, C., GINGERY, M., EISENBERG, D., RUSSELL, R. G., DERRICK, S. C., COLLINS, F. M., MORRIS, S. L., KING, C. H. & JACOBS, W. R., JR. (2003) The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci U S A,* 100**,** 12420-5.

HUANG, J., RU, B., LI, S., LIN, H. & GUO, F. B. (2010) SAROTUP: scanner and reporter of target-unrelated peptides. *J Biomed Biotechnol,* 2010**,** 101932.

ILLUMINA, I. (2006-2008) Multiplexed sequencing with the illumina genome analyser.

IWAMOTO, T., SONOBE, T. & HAYASHI, K. (2003) Loop-Mediated Isothermal Amplification for Direct Detection of Mycobacterium tuberculosis Complex, M. avium, and M. intracellulare in Sputum Samples. *Journal of Clinical Microbiology,* 41**,** 2616-2622.

JACOBS JR, W. R., BARLETTA, R. G., UDANI, R., CHAN, J., KALKUT, G., SOSNE, G., KIESER, T., SARKIS, G. J. & HATFULL, G. F. (1993) Rapid Assessment of Drug Susceptibilities of< i> Mycobacterium tuberculosis</i> by Means of Luciferase Reporter Phages.

JAMES, W. (2001) Nucleic acid and polypeptide aptamers: a powerful approach to ligand discovery. *Curr Opin Pharmacol,* 1**,** 540-6.

JAYASENA, S. D. (1999) Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clin Chem,* 45**,** 1628-50.

JESPERS, L. S., MESSENS, J. H., DE KEYSER, A., EECKHOUT, D., VAN DEN BRANDE, I., GANSEMANS, Y. G., LAUWEREYS, M. J., VLASUK, G. P. & STANSSENS, P. E. (1995) Surface expression and ligand-based selection of cDNAs fused to filamentous phage gene VI. *Biotechnology (N Y),* 13**,** 378-82.

KEHOE, J. W. & KAY, B. K. (2005) Filamentous phage display in the new millennium. *Chem Rev,* 105**,** 4056-72.

KENT, P. T. & KUBICA, G. P. (1985) *Public health mycobacteriology: a guide for the level III laboratory*, US Department of Health and Human Services, Public Health Service, Centers for Disease Control.

KHATI, M., SCHUMAN, M., IBRAHIM, J., SATTENTAU, Q., GORDON, S. & JAMES, W. (2003) Neutralization of infectivity of diverse R5 clinical isolates of human immunodeficiency virus type 1 by gp120-binding 2'F-RNA aptamers. *J Virol,* 77**,** 12692-8.

KIM, S. H., SEO, K. W., KIM, J., LEE, K. Y. & JANG, Y. S. (2010) The M cell-targeting ligand promotes antigen delivery and induces antigen-specific immune responses in mucosal vaccination. *J Immunol,* 185**,** 5787-95.

KINGSBURY, G. A. & JUNGHANS, R. P. (1995) Screening of phage display immunoglobulin libraries by anti-M13 ELISA and whole phage PCR. *Nucleic Acids Res,* 23**,** 2563-4.

KISELAR, J. G. & DOWNARD, K. M. (1999) Antigenic surveillance of the influenza virus by mass spectrometry. *Biochemistry,* 38**,** 14185-91.

KIVIHYA-NDUGGA, L. E. A., VAN CLEEFF, M. R. A., GITHUI, W. A., NGANGA, L. W., KIBUGA, D. K., ODHIAMBO, J. A. & KLATSER, P. R. (2003) A comprehensive comparison of Ziehi-Neelsen and fluorescence microscopy for the diagnosis of tuberculosis in a resource-poor urban setting. *International Journal of Tuberculosis and Lung Disease,* 7**,** 1163.

KOIVUNEN, E., WANG, B. & RUOSLAHTI, E. (1994) Isolation of a highly specific ligand for the alpha 5 beta 1 integrin from a phage display library. *J Cell Biol,* 124**,** 373-80.

KOOLPE, M., DAIL, M. & PASQUALE, E. B. (2002) An ephrin mimetic peptide that selectively targets the EphA2 receptor. *J Biol Chem,* 277**,** 46974-9.

KUMAR, P. K., MACHIDA, K., URVIL, P. T., KAKIUCHI, N., VISHNUVARDHAN, D., SHIMOTOHNO, K., TAIRA, K. & NISHIKAWA, S. (1997) Isolation of RNA aptamers specific to the NS3 protein of hepatitis C virus from a pool of completely random RNA. *Virology,* 237**,** 270-82.

LARBANOIX, L., BURTEA, C., LAURENT, S., VAN LEUVEN, F., TOUBEAU, G., VANDER ELST, L. & MULLER, R. N. (2010) Potential amyloid plaque-specific peptides for the diagnosis of Alzheimer's disease. *Neurobiol Aging,* 31**,** 1679-89.

LEE, J. H., CANNY, M. D., DE ERKENEZ, A., KRILLEKE, D., NG, Y. S., SHIMA, D. T., PARDI, A. & JUCKER, F. (2005) A therapeutic aptamer inhibits angiogenesis by specifically targeting the heparin binding domain of VEGF165. *Proc Natl Acad Sci U S A,* 102**,** 18902-7.

LEMASSU, A., ORTALO-MAGNE, A., BARDOU, F., SILVE, G., LANEELLE, M. A. & DAFFE, M. (1996) Extracellular and surface-exposed polysaccharides of non-tuberculous mycobacteria. *Microbiology,* 142 (Pt 6)**,** 1513-20.

LETTIERI, C. (2007) Nontuberculous mycobacteria: update on diagnosis and treatment. *Chest*.

LEVY, H., FELDMAN, C., SACHO, H., VAN DER MEULEN, H., KALLENBACH, J. & KOORNHOF, H. (1989) A reevaluation of sputum microscopy and culture in the diagnosis of pulmonary tuberculosis. *Chest,* 95**,** 1193-7.

LEWIS, K. N., LIAO, R., GUINN, K. M., HICKEY, M. J., SMITH, S., BEHR, M. A. & SHERMAN, D. R. (2003) Deletion of RD1 from Mycobacterium tuberculosis mimics bacille Calmette-Guerin attenuation. *J Infect Dis,* 187**,** 117-23.

LIGHTBODY, K. L., RENSHAW, P. S., COLLINS, M. L., WRIGHT, R. L., HUNT, D. M., GORDON, S. V., HEWINSON, R. G., BUXTON, R. S., WILLIAMSON, R. A. & CARR, M. D. (2004) Characterisation of complex formation between members of the Mycobacterium tuberculosis complex CFP-10/ESAT-6 protein family: towards an understanding of the rules governing complex formation and thereby functional flexibility. *FEMS Microbiol Lett,* 238**,** 255-62.

LING, D. I., FLORES, L. L., RILEY, L. W. & PAI, M. (2008) Commercial nucleic-acid amplification tests for diagnosis of pulmonary tuberculosis in respiratory specimens: meta-analysis and meta-regression. *PLoS One,* 3**,** e1536.

LISS, M., PETERSEN, B., WOLF, H. & PROHASKA, E. (2002) An aptamer-based quartz crystal protein biosensor. *Anal Chem,* 74**,** 4488-95.

LIU, J. & LU, Y. (2006) Smart Nanomaterials Responsive to Multiple Chemical Stimuli with Controllable Cooperativity. *Advanced Materials,* 18**,** 1667.

LIU, J., MAZUMDAR, D. & LU, Y. (2006) A simple and sensitive "dipstick" test in serum based on lateral flow separation of aptamer-linked nanostructures. *Angew Chem Int Ed Engl,* 45**,** 7955-9.

LOPEZ-MARIN, L. M. (2012) Nonprotein structures from mycobacteria: emerging actors for tuberculosis control. *Clin Dev Immunol,* 2012**,** 917860.

LOWMAN, H. B., BASS, S. H., SIMPSON, N. & WELLS, J. A. (1991) Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry,* 30**,** 10832-8.

LUCCHESE, G., STUFANO, A. & KANDUC, D. (2009) Proteome-guided search for influenza A B-cell epitopes. *FEMS Immunol Med Microbiol,* 57**,** 88-92.

LUNDER, M., BRATKOVIC, T., URLEB, U., KREFT, S. & STRUKELJ, B. (2008) Ultrasound in phage display: a new approach to nonspecific elution. *Biotechniques,* 44**,** 893-900.

MACIAG, A., PIAZZA, A., RICCARDI, G. & MILANO, A. (2009) Transcriptional analysis of ESAT-6 cluster 3 in Mycobacterium smegmatis. *BMC Microbiol,* 9**,** 48.

MANNOCCI, L., ZHANG, Y., SCHEUERMANN, J., LEIMBACHER, M., DE BELLIS, G., RIZZI, E., DUMELIN, C., MELKKO, S. & NERI, D. (2008) High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries. *Proc Natl Acad Sci U S A,* 105**,** 17670-5.

MARSHALL, K. A. & ELLINGTON, A. D. (2000) In vitro selection of RNA aptamers. *Methods Enzymol,* 318**,** 193-214.

MAWUENYEGA, K. G., FORST, C. V., DOBOS, K. M., BELISLE, J. T., CHEN, J., BRADBURY, E. M., BRADBURY, A. R. & CHEN, X. (2005)

Mycobacterium tuberculosis functional network analysis by global subcellular protein profiling. *Mol Biol Cell,* 16**,** 396-404.

MCCONNELL, S. J., UVEGES, A. J., FOWLKES, D. M. & SPINELLA, D. G. (1996) Construction and screening of M13 phage libraries displaying long random peptides. *Mol Divers,* 1**,** 165-76.

MCDONOUGH, J. A., MCCANN, J. R., TEKIPPE, E. M., SILVERMAN, J. S., RIGEL, N. W. & BRAUNSTEIN, M. (2008) Identification of functional Tat signal sequences in Mycobacterium tuberculosis proteins. *J Bacteriol,* 190**,** 6428-38.

MCLAFFERTY, M. A., KENT, R. B., LADNER, R. C. & MARKLAND, W. (1993) M13 bacteriophage displaying disulfide-constrained microproteins. *Gene,* 128**,** 29-36.

MENENDEZ, A. & SCOTT, J. K. (2005) The nature of target-unrelated peptides recovered in the screening of phage-displayed random peptide libraries with antibodies. *Anal Biochem,* 336**,** 145-57.

MIETHKE, M. & MARAHIEL, M. A. (2007) Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev,* 71**,** 413-51.

MISONO, T. S. & KUMAR, P. K. (2005) Selection of RNA aptamers against human influenza virus hemagglutinin using surface plasmon resonance. *Anal Biochem,* 342**,** 312-7.

MOHAMMADI, M., RASAEE, M. J., RAJABIBAZL, M., PAKNEJAD, M., ZARE, M. & MOHAMMADZADEH, S. (2007) Epitope mapping of PR81 anti-MUC1 monoclonal antibody following PEPSCAN and phage display techniques. *Hybridoma (Larchmt),* 26**,** 223-30.

MOORE, D. F., GUZMAN, J. A. & MIKHAIL, L. T. (2005) Reduction in turnaround time for laboratory diagnosis of pulmonary tuberculosis by routine use of a nucleic acid amplification test. *Diagnostic Microbiology and Infectious Disease,* 52**,** 247.

MORI, T., SAKATANI, M., YAMAGISHI, F., TAKASHIMA, T., KAWABE, Y., NAGAO, K., SHIGETO, E., HARADA, N., MITARAI, S., OKADA, M., SUZUKI, K., INOUE, Y., TSUYUGUCHI, K., SASAKI, Y., MAZUREK, G. H. & TSUYUGUCHI, I. (2004) Specific Detection of Tuberculosis Infection: An Interferon-Î³â€“based Assay Using New Antigens. *American Journal of Respiratory and Critical Care Medicine,* 170**,** 59-64.

MULLEN, L. M., NAIR, S. P., WARD, J. M., RYCROFT, A. N. & HENDERSON, B. (2006) Phage display in the study of infectious diseases. *Trends Microbiol,* 14**,** 141-7.

NAIK, R. R., BROTT, L. L., CLARSON, S. J. & STONE, M. O. (2002) Silica-precipitating peptides isolated from a combinatorial phage display peptide library. *J Nanosci Nanotechnol,* 2**,** 95-100.

NAVANI, N. K. & LI, Y. (2006) Nucleic acid aptamers and enzymes as sensors. *Curr Opin Chem Biol,* 10**,** 272-81.

NIEDERWEIS, M. (2003) Mycobacterial porins--new channel proteins in unique outer membranes. *Mol Microbiol,* 49**,** 1167-77.

NOGALES, C., BERNAL, S. & CHAVEZ, M. (1999) Comparison of the MB/BacT and BACTEC 460 TB systems. *J Clin Microbiol,* 37**,** 3432-3.

ORTALO-MAGNE, A., DUPONT, M. A., LEMASSU, A., ANDERSEN, A. B., GOUNON, P. & DAFFE, M. (1995) Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiology,* 141 (Pt 7)**,** 1609-20.

OSBORNE, S. E., MATSUMURA, I. & ELLINGTON, A. D. (1997) Aptamers as therapeutic and diagnostic reagents: problems and prospects. *Curr Opin Chem Biol,* 1**,** 5-9.

PAI, M., KALANTRI, S., PASCOPELLA, L., RILEY, L. W. & REINGOLD, A. L. (2005) Bacteriophage-based assays for the rapid detection of rifampicin resistance in Mycobacterium tuberculosis: a meta-analysis. *Journal of Infection,* 51**,** 175-187.

PALACIOS, J. J., FERRO, J., RUIZ PALMA, N., GARCÃA, J. M., VILLAR, H., RODRÃGUEZ, J., MACIAS, M. D. & PRENDES, P. (1999) Fully automated liquid culture system compared with Lowenstein-Jensen solid medium for rapid recovery of mycobacteria from clinical samples. *European Journal of Clinical Microbiology & Infectious Diseases,* 18**,** 265-273.

PALLEN, M. J. (2002) The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system? *Trends Microbiol,* 10**,** 209-12.

PANDEY, B. D., POUDEL, A., YODA, T., TAMARU, A., ODA, N., FUKUSHIMA, Y., LEKHAK, B., RISAL, B., ACHARYA, B., SAPKOTA, B., NAKAJIMA, C., TANIGUCHI, T., PHETSUKSIRI, B. & SUZUKI, Y. (2008) Development of an in-house loop-mediated isothermal amplification (LAMP) assay for detection of Mycobacterium tuberculosis and evaluation in sputum samples of Nepalese patients. *Journal of Medical Microbiology,* 57**,** 439-443.

PASQUALINI, R., KOIVUNEN, E. & RUOSLAHTI, E. (1995) A peptide isolated from phage display libraries is a structural and functional mimic of an RGD-binding site on integrins. *J Cell Biol,* 130**,** 1189-96.

PFYFFER, G. E., CIESLAK, C., WELSCHER, H.-M., KISSLING, P. & RÃ¼SCH-GERDES, S. (1997a) Rapid detection of mycobacteria in clinical specimens by using the automated BACTEC 9000 MB system and comparison with radiometric and solid-culture systems. *Journal of clinical microbiology,* 35**,** 2229-2234.

PFYFFER, G. E., WELSCHER, H. M., KISSLING, P., CIESLAK, C., CASAL, M. J., GUTIERREZ, J. & RÃ¼SCH-GERDES, S. (1997b) Comparison of the Mycobacteria Growth Indicator Tube (MGIT) with radiometric and solid culture for recovery of acid-fast bacilli. *Journal of Clinical Microbiology,* 35**,** 364-8.

PHILLIPS, M., CATANEO, R. N., CONDOS, R., RING ERICKSON, G. A., GREENBERG, J., LA BOMBARDI, V., MUNAWAR, M. I. & TIETJE, O. (2007) Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis (Edinb),* 87**,** 44-52.

PROSKE, D., BLANK, M., BUHMANN, R. & RESCH, A. (2005) Aptamers--basic research, drug development, and clinical applications. *Appl Microbiol Biotechnol,* 69**,** 367-74.

PUGSLEY, A. P. (1993) The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev,* 57**,** 50-108.

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. *Mol Microbiol,* 46**,** 709-17.

QUE-GEWIRTH, N. S. & SULLENGER, B. A. (2007) Gene therapy progress and prospects: RNA aptamers. *Gene Ther,* 14**,** 283-91.

RADI, A. E., ACERO SANCHEZ, J. L., BALDRICH, E. & O'SULLIVAN, C. K. (2006) Reagentless, reusable, ultrasensitive electrochemical molecular beacon aptasensor. *J Am Chem Soc,* 128**,** 117-24.

RASTOGI & N (1991) *Recent observations concerning structure and function relationships in the mycobacterial cell envelope: elaboration of a model in terms of mycobacterial pathogenicity, virulence and drug-resistance,* Issy-les-Moulineaux, FRANCE, Elsevier Masson.

RATLEDGE, C. (2009) Iron Metabolism. *Mycobacteria.* Blackwell Publishing Ltd.

RATLEDGE, C. & EWING, M. (1996) The occurrence of carboxymycobactin, the siderophore of pathogenic mycobacteria, as a second extracellular siderophore in Mycobacterium smegmatis. *Microbiology,* 142 (Pt 8)**,** 2207-12.

RAVN, P., MUNK, M. E., ANDERSEN, A. B., LUNDGREN, B., LUNDGREN, J. D., NIELSEN, L. N., KOK-JENSEN, A., ANDERSEN, P. & WELDINGH, K. (2005) Prospective evaluation of a whole-blood test using Mycobacterium tuberculosis-specific antigens ESAT-6 and CFP-10 for diagnosis of active tuberculosis. *Clin Diagn Lab Immunol,* 12**,** 491-6.

REITHER, K., SAATHOFF, E., JUNG, J., MINJA, L. T., KROIDL, I., SAAD, E., HUGGETT, J. F., NTINGINYA, E. N., MAGANGA, L., MABOKO, L. & HOELSCHER, M. (2009) Low sensitivity of a urine LAM-ELISA in the diagnosis of pulmonary tuberculosis. *BMC Infect Dis,* 9**,** 141.

RHIE, A., KIRBY, L., SAYER, N., WELLESLEY, R., DISTERER, P., SYLVESTER, I., GILL, A., HOPE, J., JAMES, W. & TAHIRI-ALAOUI, A. (2003) Characterization of 2'-fluoro-RNA aptamers that bind preferentially to disease-associated conformations of prion protein and inhibit conversion. *J Biol Chem,* 278**,** 39697-705.

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. *Infect Immun,* 70**,** 3371-81.

ROTHENSTEIN, D., CLAASEN, B., OMIECIENSKI, B., LAMMEL, P. & BILL, J. (2012) Isolation of ZnO-binding 12-mer peptides and determination of their binding epitopes by NMR spectroscopy. *J Am Chem Soc,* 134**,** 12547-56.

ROTHERHAM, L. S., MASERUMULE, C., DHEDA, K., THERON, J. & KHATI, M. (2012) Selection and application of ssDNA aptamers to detect active TB from sputum samples. *PLoS One,* 7**,** e46862.

ROY, M. D., STANLEY, S. K., AMIS, E. J. & BECKER, M. L. (2008) Identification of a Highly Specific Hydroxyapatite-binding Peptide using Phage Display. *Advanced Materials,* 20**,** 1830.

SADA, E., AGUILAR, D., TORRES, M. & HERRERA, T. (1992) Detection of lipoarabinomannan as a diagnostic test for tuberculosis. *J Clin Microbiol,* 30**,** 2415-8.

SAINT-JOANIS, B., DEMANGEL, C., JACKSON, M., BRODIN, P., MARSOLLIER, L., BOSHOFF, H. & COLE, S. T. (2006) Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of Mycobacterium tuberculosis, increases beta-lactam susceptibility and virulence. *J Bacteriol,* 188**,** 6669-79.

SAMPSON, S. L., DASCHER, C. C., SAMBANDAMURTHY, V. K., RUSSELL, R. G., JACOBS, W. R., JR., BLOOM, B. R. & HONDALUS, M. K. (2004) Protection elicited by a double leucine and pantothenate auxotroph of Mycobacterium tuberculosis in guinea pigs. *Infect Immun,* 72**,** 3031-7.

SANI, M., HOUBEN, E. N., GEURTSEN, J., PIERSON, J., DE PUNDER, K., VAN ZON, M., WEVER, B., PIERSMA, S. R., JIMENEZ, C. R., DAFFE, M., APPELMELK, B. J., BITTER, W., VAN DER WEL, N. & PETERS, P. J. (2010) Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS Pathog,* 6**,** e1000794.

SASSETTI, C. M., BOYD, D. H. & RUBIN, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology,* 48**,** 77.

SAWADA, R., PETERSON, C. Y., GONZALEZ, A. M., POTENZA, B. M., MUELLER, B., COIMBRA, R., ELICEIRI, B. P. & BAIRD, A. (2011) A phage-targeting strategy for the design of spatiotemporal drug delivery from grafted matrices. *Fibrogenesis Tissue Repair,* 4**,** 7.

SCHAGGER, H. & VON JAGOW, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem,* 166**,** 368-79.

SEGVICH, S., BISWAS, S., BECKER, U. & KOHN, D. H. (2009) Identification of peptides with targeted adhesion to bone-like mineral via phage display and computational modeling. *Cells Tissues Organs,* 189**,** 245-51.

SENIOR, K. (2009) Relentless spread if extensively drug-resistant tuberculosis. *The Lancet Infectious Disease,* 9**,** 403-403.

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. *J Bacteriol,* 191**,** 6340-4.

SHEVCHENKO, A., TOMAS, H., HAVLIS, J., OLSEN, J. V. & MANN, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc,* 1**,** 2856-60.

SHINNICK, T. M. & GOOD, R. C. (1994) Mycobacterial taxonomy. *European Journal of Clinical Microbiology and Infectious Diseases,* 13**,** 884-901.

SHUI, G., BENDT, A. K., JAPPAR, I. A., LIM, H. M., LANEELLE, M., HERVE, M., VIA, L. E., CHUA, G. H., BRATSCHI, M. W., ZAINUL RAHIM, S. Z., MICHELLE, A. L., HWANG, S. H., LEE, J. S., EUM, S. Y., KWAK, H. K., DAFFE, M., DARTOIS, V., MICHEL, G., BARRY, C. E., 3RD & WENK, M. R. (2012) Mycolic acids as diagnostic markers for tuberculosis case detection in humans and drug efficacy in mice. *EMBO Mol Med,* 4**,** 27-37.

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. *Proc Natl Acad Sci U S A,* 106**,** 18792-7.

SINHA, S., KOSALAI, K., ARORA, S., NAMANE, A., SHARMA, P., GAIKWAD, A. N., BRODIN, P. & COLE, S. T. (2005) Immunogenic membrane-associated proteins of Mycobacterium tuberculosis revealed by proteomics. *Microbiology,* 151**,** 2411-2419.

SMITH, G. P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science,* 228**,** 1315-7.

SMITH, G. P. & PETRENKO, V. A. (1997) Phage Display. *Chem Rev,* 97**,** 391-410.

SMITH, I. (2003) Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev,* 16**,** 463-96.

SOHN, H., MINION, J., ALBERT, H., DHEDA, K. & PAI, M. (2009) TB diagnostic tests: how do we figure out their costs? *Expert Rev Anti Infect Ther,* 7**,** 723-33.

SONG, S. H., PARK, K. U., LEE, J. H., KIM, E. C., KIM, J. Q. & SONG, J. (2009) Electrospray ionization-tandem mass spectrometry analysis of the mycolic acid profiles for the identification of common clinical isolates of mycobacterial species. *Journal of Microbiological Methods,* 77**,** 165.

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. *Proc Natl Acad Sci U S A,* 100**,** 13001-6.

STEINGART, K. R., HENRY, M., LAAL, S., HOPEWELL, P. C., RAMSAY, A., MENZIES, D., CUNNINGHAM, J., WELDINGH, K. & PAI, M. (2007a) Commercial serological antibody detection tests for the diagnosis of pulmonary tuberculosis: a systematic review. *PLoS Med,* 4**,** e202.

STEINGART, K. R., HENRY, M., LAAL, S., HOPEWELL, P. C., RAMSAY, A., MENZIES, D., CUNNINGHAM, J., WELDINGH, K. & PAI, M. (2007b) A systematic review of commercial serological antibody detection tests for the diagnosis of extrapulmonary tuberculosis. *Thorax,* 62**,** 911-8.

STEINGART, K. R., NG, V., HENRY, M., HOPEWELL, P. C., RAMSAY, A., CUNNINGHAM, J., URBANCZIK, R., PERKINS, M. D., AZIZ, M. A. & PAI, M. (2006) Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis,* 6**,** 664-74.

STEINGART, K. R., RAMSAY, A. & PAI, M. (2007c) Commercial serological tests for the diagnosis of tuberculosis: do they work? *Future Microbiol,* 2**,** 355-9.

STRATMANN, J., STROMMENGER, B., STEVENSON, K. & GERLACH, G. F. (2002) Development of a peptide-mediated capture PCR for detection of Mycobacterium avium subsp. paratuberculosis in milk. *J Clin Microbiol,* 40**,** 4244-50.

TASSET, D. M., KUBIK, M. F. & STEINER, W. (1997) Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J Mol Biol,* 272**,** 688-98.

TEKAIA, F., GORDON, S. V., GARNIER, T., BROSCH, R., BARRELL, B. G. & COLE, S. T. (1999) Analysis of the proteome of Mycobacterium tuberculosis in silico. *Tuber Lung Dis,* 79**,** 329-42.

TIPPS, M. E., LAWSHE, J. E., ELLINGTON, A. D. & MIHIC, S. J. (2010) Identification of novel specific allosteric modulators of the glycine receptor using phage display. *J Biol Chem,* 285**,** 22840-5.

TOMBELLI, S., MINUNNI, M. & MASCINI, M. (2007) Aptamers-based assays for diagnostics, environmental and food analysis. *Biomolecular Engineering,* 24**,** 191.

TORTOLI, E. (2006) The new mycobacteria: an update. *FEMS Immunol Med Microbiol,* 48**,** 159-78.

TSIOURIS, S. J., COETZEE, D., TORO, P. L., AUSTIN, J., STEIN, Z. & EL-SADR, W. (2006) Sensitivity Analysis and Potential Uses of a Novel Gamma Interferon Release Assay for Diagnosis of Tuberculosis. *Journal of Clinical Microbiology,* 44**,** 2844-2850.

TSUKAMURA, M., MIZUNO, S. & TOYAMA, H. (1985) Taxonomic studies on the Mycobacterium tuberculosis series. *Microbiol Immunol,* 29**,** 285-99.

TUERK, C. & GOLD, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science,* 249**,** 505-10.

VADWAI, V., BOEHME, C., NABETA, P., SHETTY, A., ALLAND, D. & RODRIGUES, C. (2011) Xpert MTB/RIF: a new pillar in diagnosis of extrapulmonary tuberculosis? *J Clin Microbiol,* 49**,** 2540-5.

VAN DER MERWE, P. A. (2000) Surface plasmon resonance. *Proteinâ€"Ligand Interactions: A Practical Approach***,** 3-50.

VAN RIE, A., PAGE-SHIPP, L., SCOTT, L., SANNE, I. & STEVENS, W. (2010) Xpert((R)) MTB/RIF for point-of-care diagnosis of TB in high-HIV burden, resource-limited countries: hype or hope? *Expert Rev Mol Diagn,* 10**,** 937-46.

VAN ZYL-SMIT, R. N., BINDER, A., MELDAU, R., MISHRA, H., SEMPLE, P. L., THERON, G., PETER, J., WHITELAW, A., SHARMA, S. K., WARREN, R., BATEMAN, E. D. & DHEDA, K. (2011) Comparison of quantitative techniques including Xpert MTB/RIF to evaluate mycobacterial burden. *PLoS One,* 6**,** e28815.

VERHAERT, R. M., BEEKWILDER, J., OLSTHOORN, R., VAN DUIN, J. & QUAX, W. J. (2002) Phage display selects for amylases with improved low pH starch-binding. *J Biotechnol,* 96**,** 103-18.

WALLIS, R. S., PALACI, M., VINHAS, S., HISE, A. G., RIBEIRO, F. C., LANDEN, K., CHEON, S. H., SONG, H. Y., PHILLIPS, M., DIETZE, R. & ELLNER, J. J. (2001) A whole blood bactericidal assay for tuberculosis. *J Infect Dis,* 183**,** 1300-3.

WANG, L., TURNER, M. O., ELWOOD, R. K., SCHULZER, M. & FITZGERALD, J. M. (2002) A meta-analysis of the effect of Bacille Calmette Guerin vaccination on tuberculin skin test measurements. *Thorax,* 57**,** 804-809.

WAYNE, L. G. & DIAZ, G. A. (1986) A double staining method for differentiating between two classes of mycobacterial catalase in polyacrylamide electrophoresis gels. *Anal Biochem,* 157**,** 89-92.

WEINBERG, E. D. (1999) Iron loading and disease surveillance. *Emerg Infect Dis,* 5**,** 346-52.

WIKER, H. G. & HARBOE, M. (1992) The antigen 85 complex: a major secretion product of Mycobacterium tuberculosis. *Microbiol Rev,* 56**,** 648-61.

WILLARD, F. S. & SIDEROVSKI, D. P. (2006) Covalent immobilization of histidine-tagged proteins for surface plasmon resonance. *Anal Biochem,* 353**,** 147-9.

WILSON, S. M., AL-SUWAIDI, Z., MCNERNEY, R., PORTER, J. & DROBNIEWSKI, F. (1997) Evaluation of a new rapid bacteriophage-based method for the drug susceptibility testing of Mycobacterium tuberculosis. *Nature medicine,* 3**,** 465-468.

WINDER & ROB (2005) *Cell culture changes gear,* Bushey, ROYAUME-UNI, Ten Alps Creative.

WONG, D. K., GOBIN, J., HORWITZ, M. A. & GIBSON, B. W. (1996) Characterization of exochelins of Mycobacterium avium: evidence for saturated and unsaturated and for acid and ester forms. *J Bacteriol,* 178**,** 6394-8.

WOODS, G. L., FISH, G., PLAUNT, M. & MURPHY, T. (1997) Clinical evaluation of difco ESP culture system II for growth and detection of mycobacteria. *Journal of Clinical Microbiology,* 35**,** 121-4.

WORLD HEALTH ORGANISATION, A. S. T. D. (2012) Global Tuberculosis Control. *WHO Report*.

XIAO, Y., LUBIN, A. A., HEEGER, A. J. & PLAXCO, K. W. (2005) Label-free electronic detection of thrombin in blood serum by using an aptamer-based sensor. *Angew Chem Int Ed Engl,* 44**,** 5456-9.

XU, W. & ELLINGTON, A. D. (1996) Anti-peptide aptamers recognize amino acid sequence and bind a protein epitope. *Proc Natl Acad Sci U S A,* 93**,** 7475-80.

YANG, L., JIANG, H., SHI, B., WANG, H., LI, J., WANG, H., YAO, M. & LI, Z. (2010) Identification and characterization of Ch806 mimotopes. *Cancer Immunol Immunother,* 59**,** 1481-7.

ZANETTI, S., ARDITO, F., SECHI, L., SANGUINETTI, M., MOLICOTTI, P., DELOGU, G., PINNA, M. P., NACCI, A. & FADDA, G. (1997) Evaluation of a nonradiometric system (BACTEC 9000 MB) for detection of mycobacteria in human clinical samples. *Journal of clinical microbiology,* 35**,** 2072-2075.

ZHANG, H., WANG, Z., LI, X. F. & LE, X. C. (2006) Ultrasensitive detection of proteins by amplification of affinity aptamers. *Angew Chem Int Ed Engl,* 45**,** 1576-80.

ZHAO, S., ZHAO, W. & MA, L. (2010) Novel peptide ligands that bind specifically to mouse embryonic stem cells. *Peptides,* 31**,** 2027-34.

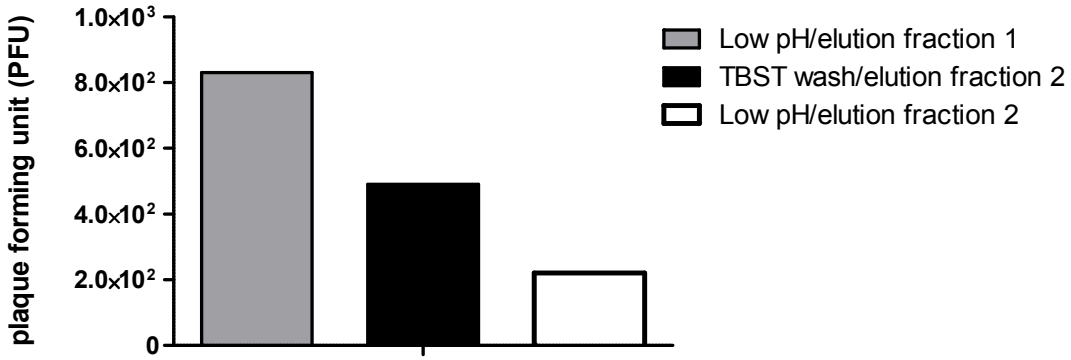
ZHONG, Y., CAI, J., ZHANG, C., XING, X., QIN, E., HE, J., MAO, P., CHENG, J., LIU, K., XU, D. & SONG, H. (2011) Mimotopes selected with neutralizing antibodies against multiple subtypes of influenza A. *Virol J,* 8**,** 542.

ZUKER, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res,* 31**,** 3406-15.

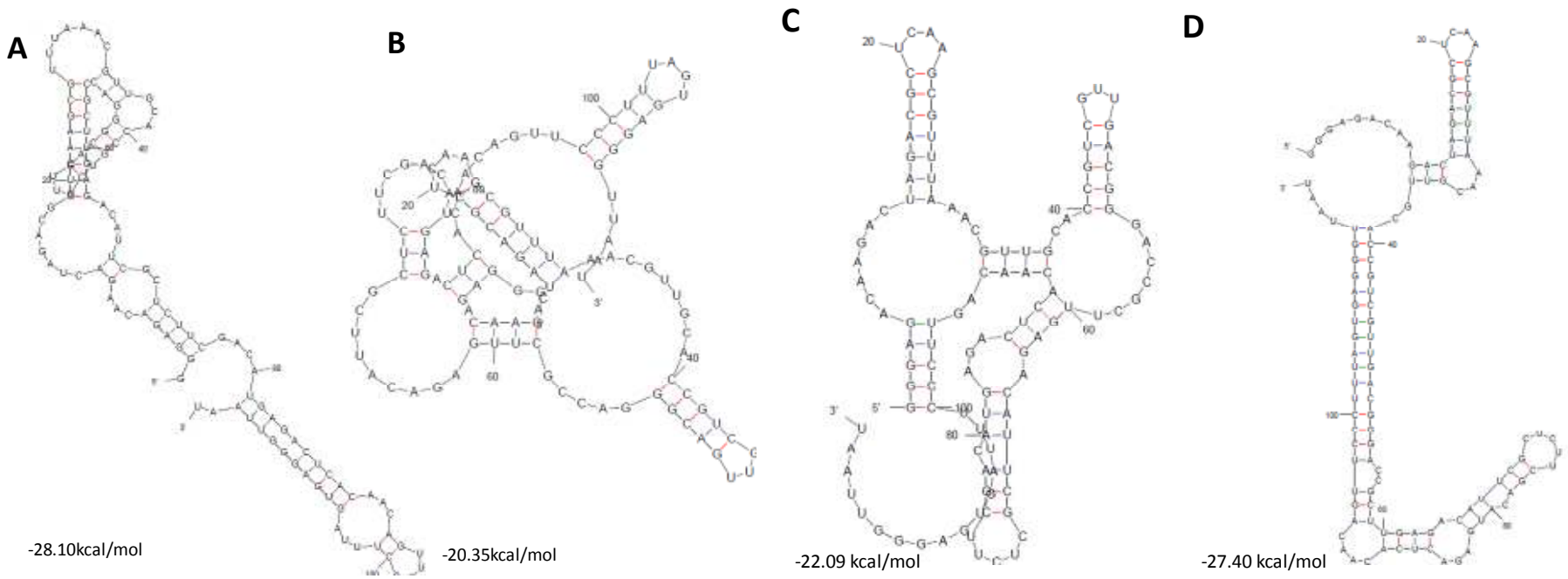
APPENDICES

Appendix Table 1: List of primer sequences

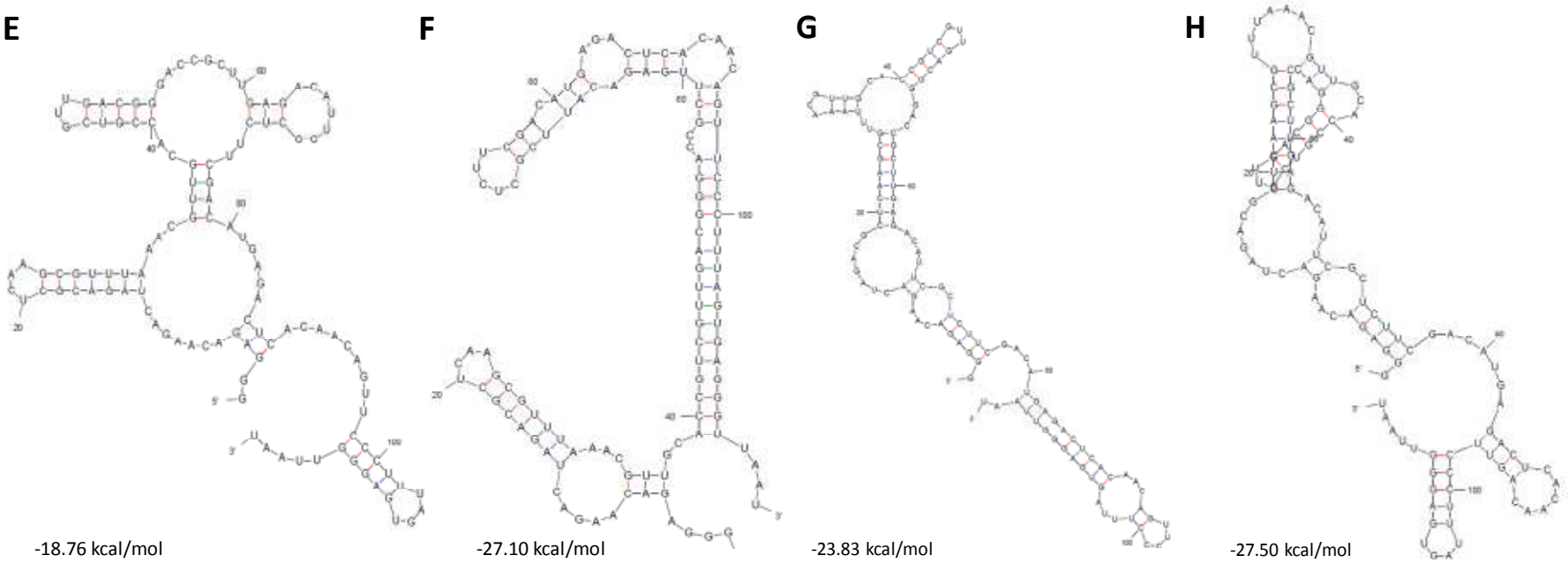
Primer Name	Primer Sequence
-96M13gIII	5´-CCCTCATAGTTAGCGTAA CG-3´
Illumina_F1	5´-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC GATCTTTAGTGGTACCTTTCTATTCTCACTCT-3´
Illumina_R1	5´-CAAGCAGAAGACGGCATACGAGAT**CGTGAT**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
Illumina_R2	5´-CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
Illumina_R3	5´-CAAGCAGAAGACGGCATACGAGAT**GCCTAA**GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
Illumina_R4	5´-CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
Illumina_R5	5´-CAAGCAGAAGACGGCATACGAGAT**ATTGGC**GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
Illumina_R6	5´-CAAGCAGAAGACGGCATACGAGAT**GATCTG**GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
Illumina_R7	5´-CAAGCAGAAGACGGCATACGAGAT**TCAAGT**GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
Illumina_R8	5´-CAAGCAGAAGACGGCATACGAGAT**CTGATC**GTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT TTCGGCCGAACCTCCACC-3´
pUC/M13 forward primer	5´-CGCCAGGGTTTTCCCAGTCACGAC-3´
pUC/M13 reverse primer	5´-TCACACAGGAAACAGCTATGAC-3´
T3SELEX	5´-AATTAACCCTCACTAAAGGGAACTGTTGTGAGTCTCATGTCGAA-3´
T7SELEX	5´-TAATACGACTCACTATAGGGAGACAAGACTAGACGCTCAA-3´



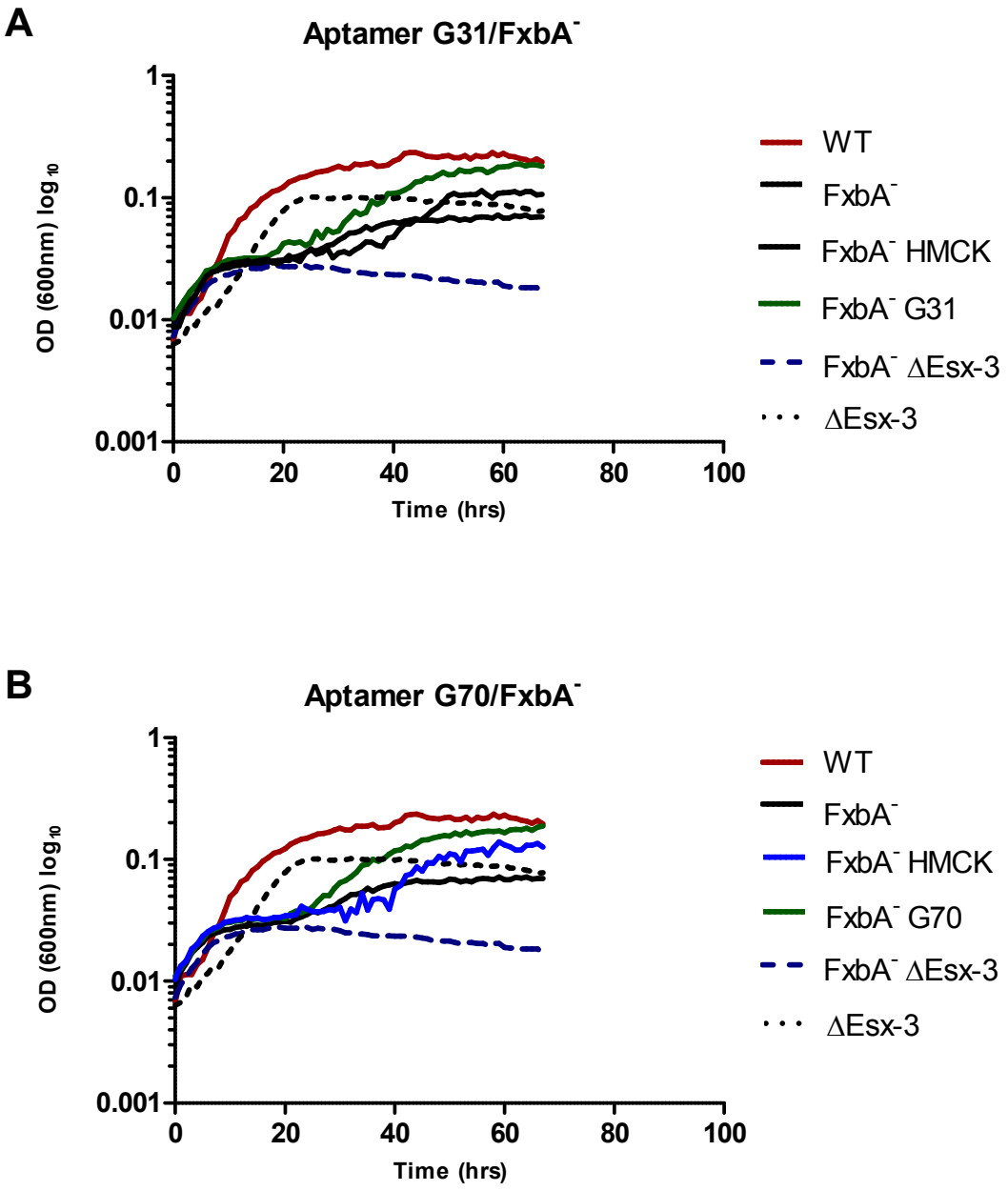
Appendix figure 1: **Three elution fractions rom biopanning protocol 2 round 1.**



Appendix figure 2: The structures and free energies of aptamer G 78 as predicted by mfold.



Appendix figure 2: The structures and free energies of aptamer G 78 as predicted by mfold.(continued)



Appendix Figure 3: G43 and G78 do not inhibit growth under low iron conditions.

Growth of (A-B) wild-type (WT),exochelin insertional biosynthesis mutant (fxbA), ΔESX-3 mutant, a double mutant (fxbA$^-$ Δesx-3) and RNA refolding buffer control (fxbA$^-$ HMCK) in low iron Sauton's medium. The aptamer experimental growth curves were performed on the fxbA$^-$background (A) (fxbA$^-$G31) and (B) (fxbA$^-$ G70) respectively, in low iron Sauton's medium.