GENERATION AND CHARACTERISATION OF RNA APTAMERS AGAINST ESAT6 PROTEIN FROM MYCOBACTERIUM TUBERCULOSIS

BAKHTIAR AFFENDI BUKARI

UNIVERSITI SAINS MALAYSIA

2017

GENERATION AND CHARACTERISATION OF RNA APTAMERS AGAINST ESAT 6 PROTEIN FROM MYCOBACTERIUM TUBERCULOSIS

by

BAKHTIAR AFFENDI BUKARI

Thesis submitted in partial fulfilment of the requirements

for the degree of

Master of Science

July 2017

ACKNOWLEDGEMENT

A daunting, overwhelmingly difficult task when I first began, the production of this thesis has been a challenging stretch of my Master's programme. Without key persons along the way, I never would have been able to finish this. Not a chance. This section is dedicated to those involved.

A special mention to my supervisors, Dr Citartan Marimuthu and Prof Tang Thean Hock for teaching me the ins and outs of scientific research, helping with my research project, and perhaps more importantly in the long run, giving me an idea of what kind of researcher and graduate I should be. Thank you for agreeing to take me as your student and telling me how I could improve myself. Your experiences were absolutely invaluable to me and has helped in so many ways for me to finish this study. I am grateful to the both of you for helping me ploughing through this chapter of my life.

Members of the Infectomics Cluster, past and present, thank you ever so much for your assistance during my time at the lab. It hasn't always been smooth sailing and circumstances can be downright discouraging with regards to our respective research projects, but together I hope that we have made it easier for everyone to complete our research projects through our discussions, interactions and lab work assistance in one way or another. (Granted, there's one problem that is pretty much impossible for us to solve – AC unit leakages.)

Last but not at all least, my family - mother and siblings - for always supporting my decision to continue studying and giving all they can to help. I couldn’t have been luckier to have them in my life. I hope that I have made you all proud for me reaching this stage of education. This thesis is dedicated especially to you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND SYMBOLS

A	Adenine
APS	Ammonium persulfate
ATP	Adenosine 5‟-triphosphate
Bis	N, N‟-methylene bisacrylamide
bp	Base pair(s)
BSA	Bovine serum albumin
C	Cytosine
°C	Degrees Celsius
cDNA	Complementary DNA
C-terminal	Carboxy-terminal
CFP10	Culture Filtrate Protein
CTP	Cytidine 5‟-triphosphate
ddH_2O	Double-distilled water
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
ELAA	Enzyme-Linked Aptamer Assay

ELISA	Enzyme-Linked Immunosorbent Assay
ESAT6	Early Secretory Antigenic Target
et al.	and others
EtBr	3, 8-diamino-5-Ethyl-6-phenyl phenanthridinium Bromide
g	Gravitational acceleration
g	Gram
G	Guanine
GTP	Guanosine 5‟-triphosphate
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidise
IPTG	Isopropyl-β-D-thiogalactopyranoside
KCl	Potassium chloride
K_d	Dissociation constant
kDa	Kilodalton
LB	Luria Bertani medium
M	Molar, $\frac{Mole}{Litre}$
Mg^{2+}	Magnesium ion
Min	Minute(s)
mL	Milliliter
mM	Millimolar
Na^{+}	Sodium ion

NaCl	Sodium chloride
$NaOAc.3H_2O$	Sodium acetate trihydrate
NaOH	Sodium hydroxide
ng	Nanogram
nM	Nanomolar
NFM	Nitrocellulose filter membrane
nt	Nucleotide(s)
N-terminal	Amino-terminal
-OH	Hydroxyl
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline with Tween20
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Rotations per minute
RT	Room temperature
RT-PCR	Reverse transcription-PCR
s	Second(s)
SELEX	Systematic Evolution of Ligands via Exponential Enrichment
ssDNA	Single-stranded DNA
T	Thymine
TAE	Tris–Acetic Acid–EDTA

TBE	Tris-Boric Acid-EDTA
TEMED	N,N,N‘,N‘-Tetramethylethylenediamine
Tris	Tris-(Hydroxymethyl)-Aminomethane
tRNA	Transfer RNA
U	Units of enzymatic activity
UTP	Uridine 5‘-triphosphate
u.v.	Ultraviolet
V	Volt (s)
v/v	Volume per volume
w/v	Weight per volume
X–gal	5‘-Bromo-4‘-Chloro-3‘-Indolyl-β-D-galactoside
µg	Microgram
µL	Microliter
µM	Micromolar
2'-F	2'-fluoro
2'-NH_2	2'-amino
γ 32P	Gamma Phosphorus-32

PENJANAAN DAN PENCIRIAN APTAMER RNA TERHADAP PROTEIN ESAT6 DARIPADA MYCOBACTERIUM TUBERCULOSIS

ABSTRAK

Aptamer adalah ligan kimia terbina daripada jujukan nukleotida pendek yang berkemampuan untuk berinteraksi dengan protein pada kadar afiniti dan ketentuan yang tinggi. Aptamer dihasilkan melalui suatu process yang dinamakan Evolusi Ligan Secara Sistemik melalui Pengkayaan Eksponen *(Systematic Evolution of Ligands via Exponential Enrichment*) atau secara ringkasnya, SELEX. Oleh kerana kadar kestabilan struktur and ketentuan yang tinggi di antara aptamer dengan protein sasarannya, aptamer mempunyai potensi yang amat tinggi untuk menjadi alatan biologi molekul yang berguna. *Early Secretory Antigenic Target* (ESAT6), ialah sejenis protein dengan jisim molekul 6 kDa yang dihasilkan oleh *Mycobacterium tuberculosis* dan dipercayai mempunyai peranan yang penting dalam penyebaran dan perkembangan penyakit tuberkulosis (TB). Protein ini terletak pada lokus yang dikenali sebagai Region of Difference 1 (RD1) di dalam genom *Mycobacterium* sp. dan kehilangan lokus ini menyebabkan bakteria tersebut tidak dapat untuk menjangkiti hos. Berdasarkan bukti tersebut, ESAT6 adalah berpotensi untuk menjadi penanda biologi untuk penyakit TB. Dengan itu, penjanaan aptamer terhadap molekul ini dapat menjadi satu batu loncatan untuk penghasilan ujian diagnosis TB yang dapat menjimatkan masa, wang dan sebagai kegunaan terapeutik. Matlamat kajian ini adalah untuk menghasilkan aptamer RNA yang boleh berinteraksi dengan ESAT6 secara khusus dan afiniti yang tinggi. Sebelas kitaran SELEX telah dijalankan menggunakan koleksi helaian tunggal RNA rawak. Ketegaran interaksi di antara jujukan RNA dan

ESAT6 dinaikkan secara beransur-ansur dengan mengubah jumlah protein, koleksi jujukan RNA dan pesaing dengan setiap kitaran SELEX. Koleksi RNA yang terhasil pada kitaran SELEX ke 11 dikaji selanjutnya dengan cerakin *nitrocellulose filter binding* dan afiniti terhadap ESAT6 telah dinilai. Berdasarkan keputusan cerakin *nitrocellulose filter binding*, pengkayaan koleksi jujukan RNA telah disahkan berlaku. Cerakin tersebut juga mempostulatkan bahawa terdapat calon-calon aptamer di dalam koleksi RNA kitaran SELEX ke 8 dan 11. Analisis jujukan populasi RNA pada koleksi RNA kitaran ke 8 dan 11 menunjukkan kehadiran beberapa jenis jujukan RNA yang dapat dinobatkan sebagai calon aptamer terhadap ESAT6. Kajian selanjutnya telah mengenal pasti dua spesis aptamer, dinamakan RAE6-1 dan RAE6-3, yang mempunyai afiniti terbaik terhadap ESAT6. Dengan menggunakan lengkuk regresi bukan linear, kadar penceraian, K_d, telah ditentukan sebagai 595.8 nM untuk RAE6-1, dan 561.9 nM bagi RAE6-3. Maklumat jujukan sepsis aptamer ini digunakan untuk penjangkaan struktur sekundernya. Jujukan kedua-dua aptamer ini setersunya disunting untuk memendekkan saiz tanpa mengubah struktur utama aptamer tersebut. Dengan menggunakan kaedah pemotongan langkah demi langkah, sebanyak 10 nukleotida dapat disunting keluar untuk aptamer RAE6-1, memendekkan aptamer tersebut dari 80 ke 70 nukleotida. Bagi aptamer RAE6-3, 8 nukleotida dapat disunting keluar, memendekkan saiz aptamer tersebut dari 80 ke 72 nukleotida. Secara kesimpulannya, dua aptamer RNA, RAE6-1 dan RAE6-3, telah berjaya dihasilkan melalui proses SELEX dan pencirian aptamer-aptamer tersebut juga telah dijalankan.

GENERATION AND CHARACTERISATION OF RNA APTAMERS AGAINST ESAT6 PROTEIN FROM MYCOBACTERIUM TUBERCULOSIS

ABSTRACT

Aptamers are chemical ligands made up of short nucleotides sequences that are able to bind to target proteins with high affinity and specificity. They are generated using a process called Systemic Evolution of Ligands via Exponential Enrichment (SELEX). Due to their chemical stability and high specificity against the target, aptamers have the potential to become very useful biological tools. The 6 kDa Early Secretory Antigenic Target, or ESAT6, is a secretory protein produced by *Mycobacterium tuberculosis* and is thought to be a major player in mycobacterial pathogenesis. The protein is found on a locus known as Region of Difference 1 (RD1) and the loss of this locus has been shown to render the *Mycobacterium* sp. unable to cause severe TB. Following this line of evidence, ESAT6 could potentially be a good biomarker for TB infection. As such, producing an aptamer against this protein could prove valuable in the attempt to create a novel and economical TB diagnostic assay with the potential in treating the disease. The objective of this study is to generate RNA aptamers that can bind specifically and with high affinity to ESAT6 protein. Eleven SELEX cycles were carried out using the N40-randomised RNA pool. Stringency of the binding reaction in each SELEX cycles was increased gradually by varying the amounts of protein, RNA pool and the competitor. The resulting RNA pool from the 8^{th} and 11^{th} cycle of SELEX was subjected to filter binding assay to assess its binding against ESAT-6 protein. Filter binding assay against the target protein confirmed that binding enrichment of the RNA pool has indeed occurred. Nitrocellulose Filter

Membrane Binding assay suggested the presence of potential binders in the RNA pool. Sequence analyses revealed that different species of potential candidate aptamers were present in both the 8^{th} and 11^{th} cycle pool. Further investigations with the binding assay identified aptamer species, RAE6-1 and RAE6-3 as having the best affinity towards ESAT6. Using a non-linear regression curve, the dissociation constant, K_d, of the aptamers were determined to be 595.8 and 561.9 nM for RAE6-1 and RAE6-3, respectively. The aptamers were then truncated to reduce them to the shortest possible length without compromising the main secondary structures. A total of 10 nucleotides were able to be excised from RAE6-1, shortening its length from 80 to 70 nucleotides. Meanwhile, 8 nucleotides were removed from RAE6-3, making the aptamer shorter from 80 to 72 nucleotides long. In conclusion, two RNA aptamers, RAE6-1 and RAE6-3, have been successfully generated using the SELEX method and the aptamers were characterised.

CHAPTER 1: INTRODUCTION

1.1 Aptamer

Aptamers are short, single stranded sequences of DNA or RNA with high specificity and affinity to their targets. The name itself is a portmanteau, derived from the Latin word '*aptus'*, meaning 'to fit' and the Greek word '*meros'*, meaning 'part' (Ellington & Szostak, 1990), alluding to the stereochemical nature of these molecules which grants them binding capabilities to their targets. The term itself was first used by Ellington and Szostak in their seminal paper in 1990. The specificity and affinity of aptamers towards their target molecules meant that they are sometimes also called chemical or synthetic antibodies.

Aptamers form secondary structures as a result of intramolecular base pairing, which allows them to interact with the molecular structures and functional groups of their target proteins (Kim & Gu, 2013). The resulting conformation will interact with its target via shape complementarity, hydrogen bonding, electrostatic attraction and stacking (Hermann & Patel, 2000; Patel, 1997).

Aptamers have tremendous potential to be very useful biological tools and should be developed and investigated further for uses in diagnostic and therapeutic fields.

1.1.1 A Brief History of Aptamer & SELEX

Two research groups from America, Tuerk and Gold, 1990, and Ellington and Szostak, 1990, independently developed an *in vitro* method to select RNA molecule against specific target proteins. Interestingly, and as typically happens in science, both of these groups did not set out to outright create a synthetic affinity reagent. In fact, in Tuerk and Gold's case, the study that lead to the development of this methodology is an extension of their previous research on bacteriophage T4 DNA polymerase interactions with its own mRNA (Tuerk & Gold, 1990).

Ellington and Szostak on the other hand was intrigued by the idea of active binding sites that could have spontaneously arisen from a collection of random sequence of nucleotides and what that would mean for theories on the origin and evolution of life (Ellington & Szostak, 1990). Nonetheless, both groups are now widely regarded as pioneers in the field of aptamers and responsible for coming up with the process of developing and selecting aptamers from a randomised pool of nucleotides.

This method, dubbed Systematic Evolution of Ligand via Exponential Enrichment or SELEX by Tuerk and Gold, involves single stranded DNA or RNA library being incubated with a target protein, followed by partitioning to remove the unbound from the target-bound sequences (Tuerk & Gold, 1990). A simplified representation of the whole process is provided in Figure 1.1. The bound sequences are collected and amplified using polymerase chain reaction (PCR) to increase their

population within the sequence library. In the case of DNA aptamer generation, the double stranded DNA (dsDNA) has to be converted into single stranded DNA (ssDNA). RNA aptamer generation on the other hand involves an *in vitro* transcription step to convert the dsDNA into RNA. The cycle of incubation, partitioning and amplification is repeated for a few more times until the pool of RNA or DNA is populated with high specificity and affinity aptamer candidates (Gopinath, 2011).

Competing reagents and counter selection steps would typically be included in the SELEX assay with additional cycles to induce a higher selective pressure for the isolation of high specificity and affinity aptamer. As the term “evolution” in the name SELEX suggests, sequences that can bind to the target protein will continue to proliferate throughout the generations within the pool while the ones that do not will be removed. Following this iterative cycle of incubation with target protein, separation of bound from unbound species, and sequence amplification over the course of the experiment, a nucleic acid pool containing enriched potential aptamer sequences that are able to strongly bind to the intended target protein will be obtained.

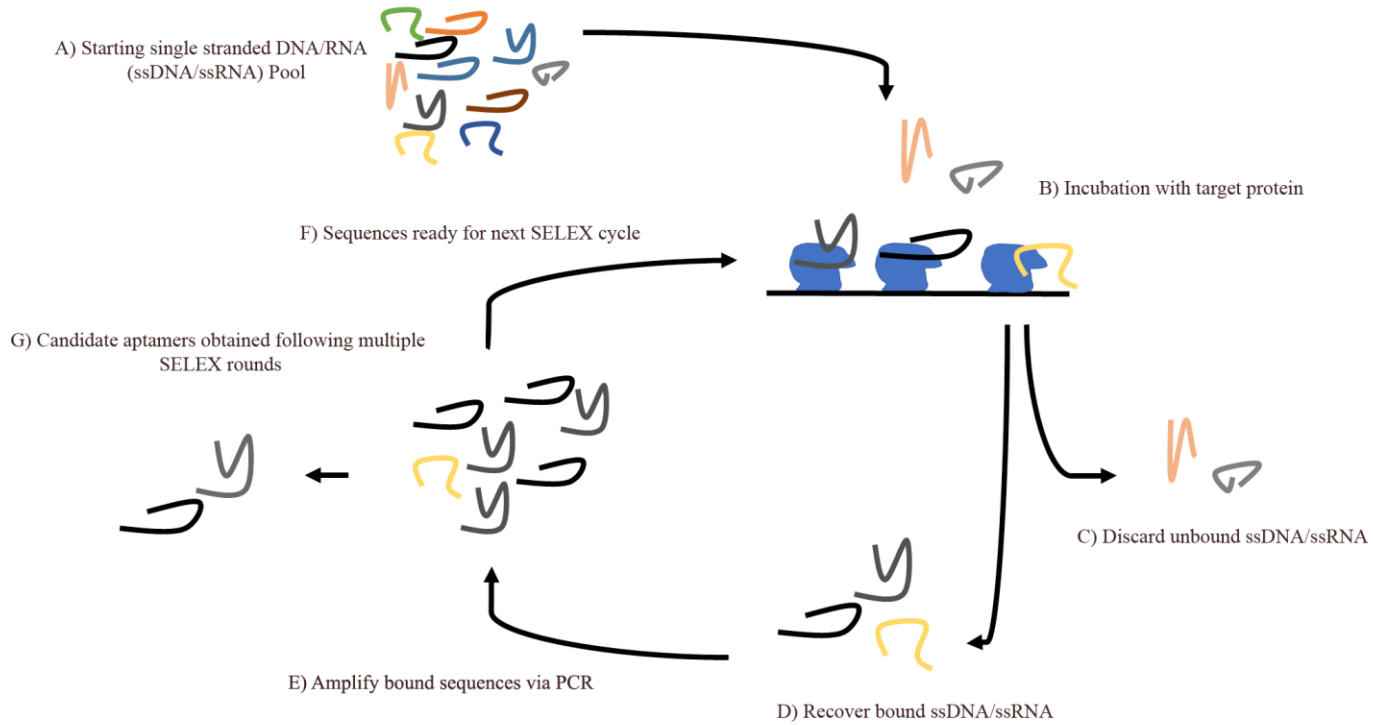


Figure 1.1 - A simplified SELEX diagram. Initial single stranded DNA/RNA pool (A) is incubated with protein of interest (B) and a partitioning step is performed to separate and remove unbound sequences from the population (C). The bound sequences will be recovered (D) and subsequently amplified using PCR. RT-PCR is necessary prior to PCR for RNA sequences (E). The pool will be converted back to either ssDNA or ssRNA before starting a new SELEX cycle (F). After 8-15 cycles, candidate aptamers can be identified for further analyses (G).

1.1.2 Advantages of Aptamers

While fulfilling the same niche, aptamers possess a number of advantages over the more established antibodies. Since they are chemically synthesised, aptamers are much more economical to produce compared to antibodies (Low, Hill, & Peccia, 2009; Sun & Zu, 2015). A recent study in 2015 estimated that an aptamer based flow cytometric assay is 1000 times cheaper compared to its antibody counterpart (Sun, Tan, & Zu, 2015). Moreover, as aptamers can be synthesised without a biological host, their development can be done faster and with no batch to batch variation problems as typically seen in polyclonal antibody productions (Jayasena, 1999; Marx, 2013; Wiberg et al., 2006). When an aptamer has been identified, the sequence can be reproduced consistently and with high fidelity each time. The fact that they do not need to be raised in a biological host also means that aptamers for toxic molecules can also be developed in the lab (Rozenblum, Lopez, Vitullo, & Radrizzani, 2016).

Compared to antibodies, aptamers are much smaller in size. A 100 bp RNA aptamer for instance is about 32 kDa which is much lighter compared to an immunoglobulin monomer, estimated at about 150 kDa (Ma & O'Kennedy, 2015; Sun et al., 2014). This size difference permits aptamers to penetrate tissue and cells much more easily to bind to their target molecules, suggesting that it could be a more sensitive reporter molecule (Jayasena, 1999). This would also mean that aptamers developed for therapeutic purposes have a much high clearance rate from the host or patients system (Hicke & Stephens, 2000).

Aptamers are also more thermodynamically stable compared to antibodies (Huang, Xi, & He, 2015; Ospina-Villa, Zamorano-Carrillo, Castañón-Sánchez, Ramírez-Moreno, & Marchat, 2016). When exposed to high temperature, they undergo reversible denaturation and are able to form their original structures upon renaturation (Jayasena, 1999; SantaLucia & Hicks, 2004).

Moreover, aptamers can be tailor-made to suit the intended downstream usage planned for them. For instance, an aptamer intended to be used in a diagnostic test can be selected at room temperature (RT) while one that is developed to be used for therapeutic purposes can be selected at 37^{o}C. Besides temperature, other parameters such as medium conditions can be integrated into the SELEX method as well. This is in contrast to antibodies which are traditionally raised in a biological host with a body temperature of 37^{o}C, which is later adopted to work in varying lab conditions (Jayasena, 1999; Stoltenburg, Reinemann, & Strehlitz, 2007).

Aptamers are less immunogenic compared to antibodies (Bouchard, Hutabarat, & Thompson, 2010; Cload, McCauley, Keefe, Healy, & Wilson, 2006; Mori, Oguro, Ohtsu, & Nakamura, 2004). This is a desirable feature to have in diagnostic assays and especially therapeutic purposes as unintended interactions with other immunoreactive molecules could lead to the trigger of immune system and inadvertent adverse physiological side effects.

Aptamers can also be easily modified through various chemical processes that could imbue them with various useful properties, increasing their utility in research, diagnostics and therapeutic applications. For instance, substitution at the 2' position of ribose on the nucleic acid backbone from hydroxyl (-OH) functional groups to amino ($-NH_2$), fluoro (-F) or methoxy ($-OCH_3$) groups confers the aptamer with an improved nuclease resistance (Zhang, 2015). Locked nucleic acid (LNA) is another type of modification that gives aptamers increased nuclear resistance and thermal stability. In LNA, a methyl bridge is introduced between the 2' oxygen and 4' carbon of the ribose group (Koshkin et al., 1998). With this modification, the aptamer will no longer be a substrate for nucleases and will help improve survivability and stability in nuclease-rich environments (Schmidt et al., 2004; Zhang, 2015). SOMAmers, or Slow Off-rate Modified Aptamers, are a new type of aptamers developed by the research group of Larry Gold, an aptamer pioneer. SOMAmers, incorporate modified nucleotides with side chains that confer more protein-like properties to the aptamer (Gold et al., 2010; Rohloff et al., 2014). The side chains on the nucleotides used in the selection process are thought to be helpful in developing aptamers with sub-nanomolar dissociation constants, give them useful properties such as nuclease resistance, increased stability, improve binding ability, attachment to beads or surfaces, and as signalling purposes (Gupta et al., 2011; Rohloff et al., 2014; Zhang, 2015).

There have been many studies demonstrating the high applicability of aptamers-based diagnosis in actual diagnostic conditions, no doubt due to the many advantages conferred by aptamers (LaVan, McGuire, & Langer, 2003; Liu, Mazumdar, & Lu, 2006; Mok & Li, 2008). These findings are certainly welcome and

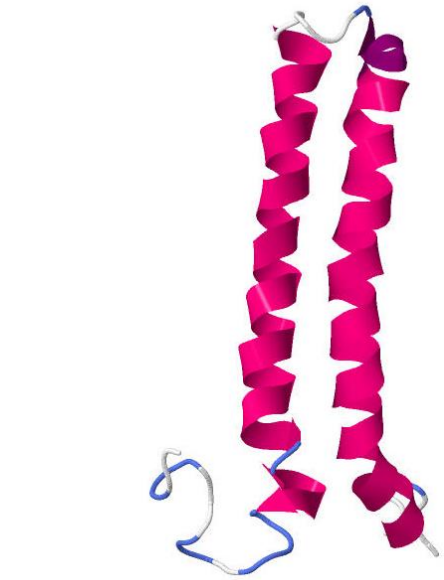
support the effort to promote and advertise the untapped potentials of aptamer in the many fields of biology.

1.2 Tuberculosis and *Mycobacterium tuberculosis*

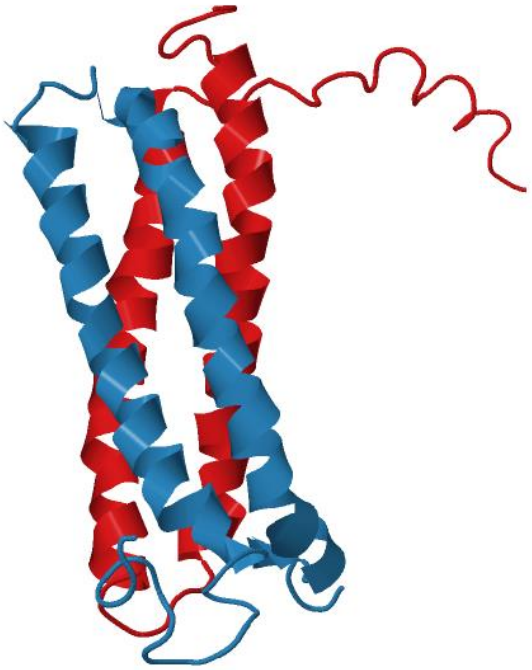
Tuberculosis (TB) is one the most devastating infectious disease known to man. The World Health Organisation (WHO) estimated that 10.4 million people worldwide in 2015 suffered from the disease caused by *Mycobacterium tuberculosis* (*Mtb*). Closer to home, the Global WHO report also estimated that Malaysia saw 27,000 incidences of TB in the year 2015 (WHO, 2016). The ease and increasing movements of people in and out of national borders within the South-East Asian region especially from high tuberculosis burden countries such as Thailand and Vietnam has been suggested by Dr Chong Chee Keong, the Malaysian Ministry of Health (MOH) director of disease control division, as one of the factors in the surge of TB cases over the last 20 years (Ng Benedict, 2014). In the effort to manage the disease and patients suffering from it, the MOH has announced that a National Tuberculosis Strategic Plan for 2016-2020 is in the process of being prepared and implemented (Roslan, 2016).

1.2.1 ESAT6

M. tuberculosis has the ability to bypass hosts' immune system using complex strategies. One such tactic is by releasing a protein called Early Secretory Antigenic Target 6 (ESAT6) into its environment in the course of infection. ESAT6 is a 6 kDa protein made up of 95 amino acid residues forming two alpha helices motifs connected by a loop (Figure 1.2) (Poulsen, Panjikar, Holton, Wilmanns, & Song, 2014; Renshaw et al., 2005).



(a)



(b)

Figure 1.2 – Structural representation of ESAT6. (a) ESAT6 structure is comprised of 2 alpha helices connected by a loop. (b) ESAT6-CFP10 heterodimer. Protein models generated using Phyre[2] protein folding prediction tool (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) and RCSB Protein Data Bank (http://www.rcsb.org/pdb/ngl/ngl.do?pdbid=1WA8).

It forms a 1:1 heterodimer complex with a neighbouring gene product called CFP10 (culture filtrate protein, 10 kDa) thus giving them structural stability when released into the environment, and is released via the ESX1 type VII secretion system (Abdallah et al., 2007; Lightbody et al., 2008; Renshaw et al., 2005). This is supported by the fact that both ESAT6 as well as CFP10 are part of the WXG100 protein superfamily, noted for its predisposition to form either homo- or heterodimeric complexes from mono-cistronic gene or bi-cistronic genes, respectively (Pallen, 2002; Poulsen et al., 2014).

Previous studies have shown that ESAT6 is a potent T-cell antigen, strongly eliciting the production of interferon-gamma (IFN-γ) as an immune response (Cardoso et al., 2002; Marei et al., 2005; Simeone, Bottai, & Brosch, 2009; van Pinxteren, Ravn, Agger, Pollock, & Andersen, 2000). The secreted protein has been also shown to confer protection to the invading pathogens against host immune system and even allowed them to spread the infection even further (Peng & Sun, 2016).

1.2.2 Role of ESAT6 in Pathogenesis

M. tuberculosis survives in its host's environment through sophisticated strategies involving ESAT6. Studies have shown that the bacterium relies on the secretion of ESAT6 to escape the phagosomal membrane of macrophage and dendritic cells and enter into the cytosol (van der Wel et al., 2007). Other studies have suggested that secreted ESAT6 promotes apoptosis of the cell when the bacterium is internalised by macrophages (Derrick & Morris, 2007; Welin, Eklund, Stendahl, & Lerm, 2011).

While it has been noted that ESAT6 forms a tight 1:1 with CFP10, the acidic environment of phagosomal membrane triggers the detachment of ESAT6 from the ESAT6-CFP10 heterodimer, leading to the disruption of lipid bilayer by ESAT6 and pore formation on the membrane, allowing the mycobacterial cells to escape cytolysis (de Jonge et al., 2007).

Using *Mycobacterium marinum* as a model, it was reported that further release of ESAT6-CFP10 into the extracellular space allows for the interaction between the heterodimer molecule and neighbouring epithelial cells, which then induces inflammatory matrix metalloprotease 9 (MMP-9) response in the latter (Volkman et al., 2010). This results in the recruitment of even more macrophages and granuloma expansion to the area thus allowing the invading mycobacteria to infect more macrophages (Boggaram, Gottipati, Wang, & Samten, 2013; Volkman et al., 2010).

In a recent study, ESAT6 also has been shown to have the ability to affect macrophage ability to present antigens. It was suggested that the antigenic protein interacts with Beta-2-Microglubulin (β2M), a component of the major histocompatibility complex (MHC) class I, inside the endoplasmic reticulum of the host cell. This prevents the β2M from interacting with its associated protein, Human Leukocyte Antigen-A (HLA-A), leading to partial inhibition of surface expression of β2M, thus reducing antigen presentation (Sreejit et al., 2014).

1.2.3 Region of Difference 1

ESAT6 resides on a locus spanning about 9.5 kilobases in the *M. tuberculosis* genome known as the Region of Difference 1 (RD1) which has been implicated with mycobacterial pathogenicity. This locus contains the codes for proteins thought to be molecular machineries required for the ESX1 secretory system which includes ESAT6, CFP10, PPE and PE (Figure 1.3) (M. A. Behr, 1999; Marcel A Behr & Sherman, 2007; Hsu et al., 2003).

This very region was found to be absent in the Bacille Calmette-Guérin (BCG) strain of *Mycobacterium bovis* which is being used as the TB vaccine (M. A. Behr, 1999; Brodin, Rosenkrands, Andersen, Cole, & Brosch, 2004; Frota et al., 2004). The absence of this locus rendered them relatively harmless to humans. Within the *M. tuberculosis* complex, pathogenic species such as *M. africanum*, and *M. bovis* has been found to harbour the RD1 locus while the non-pathogenic species like *M. microti* does not (Marcel A Behr & Sherman, 2007). Outside of the *M. tuberculosis* complex, *M. leprae*, which is responsible for leprosy, and *M. marinum*, which causes tuberculosis-like diseases in fish, carries homologous sequences of the *M. tuberculosis* ESAT6 in their genomes as well (Marcel A Behr & Sherman, 2007; Bekmurzayeva, Sypabekova, & Kanayeva, 2013; Volkman et al., 2010).

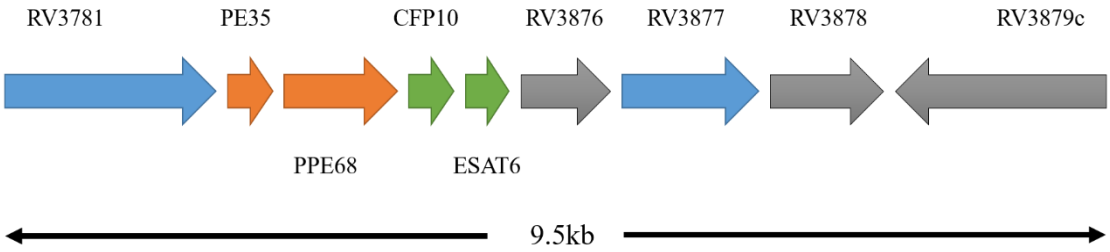


Figure 1.3- The RD1 locus where ESAT6 gene is located. The group of genes within this locus are thought to be involved in the ESX1 secretory pathway, which has a larger contribution towards mycobacterial pathogenicity. (Diagram adapted from Brodin et al., 2004)

Further studies on RD1 have shown that mycobacterial pathogenicity can indeed be affected by the presence or absence of the locus. Gao et al. in 2004 for instance used transposon mutagenesis to generate *M. marinum* strains with its RD1 disrupted from transposon insertion, resulting in reduced cytotoxicity and cell spreading capabilities (Gao et al., 2004). Similarly, deletion of the RD1 locus from virulent *M. tuberculosis* strain H37Rv rendered the strain attenuated, much like the BCG strain (Lewis et al., 2003).

Conversely, when RD1 is introduced into non-pathogenic *Mycobacterium* strains such as *M. bovis* BCG and *M. microti*, they exhibited enhanced growth and increased granuloma formation in infected cells and tissues (Pym, Brodin, Brosch, Huerre, & Cole, 2002). A similar experiment done by Guinn et al. showed that the introduction of functional copies of genes from the RD1 locus into attenuated mutant *M. tuberculosis* strains were able to restore their pathogenicity (Guinn et al., 2004).

1.3 TB Diagnosis

Based on the research findings mentioned above, it is apparent that ESAT6, contained within the RD1 locus, plays a central role in mycobacterial infection. This makes it a promising candidate for a TB diagnostic biomarker.

1.3.1 TB Culture as a Standard Approach to Diagnosis

The Centers for Disease Control and Prevention (CDC) of the United States of America considers the traditional mycobacterial sputum culture as the gold standard for TB diagnosis (CDC, 2013a). While being more sensitive and definitive as a method of TB detection, it is fraught with difficulties (Konstantinos, 2010). *M. tuberculosis* is notorious for having a long doubling time and takes 2 to 6 weeks to grow (An Wang et al., 2014). In addition to that, Wilson et al. in 2006 reported that only 56% of culture of samples taken from 54 confirmed and possible TB cases came back positive (Wilson, Nachega, Morroni, Chaisson, & Maartens, 2006). Moreover, diagnosis by culture is a technically demanding approach which also necessitates a strict biosafety practice in the lab to prevent accidental infection to laboratory personnel (Bekmurzayeva et al., 2013).

1.3.2 Other Common TB Screening Tests

The Mantoux tuberculin skin test (TST) allows for detection of a patient's previous mycobacterial infection by intracutaneous introduction of purified TB protein derivative (PPD) into his or her arm. A skin reaction will be observed within 2 days to indicate previous exposure to TB (CDC, 2013b; Kandi, 2015). The TST has the advantage of being a relatively simpler diagnostic test and has been used for more than a century ever since Felix Mendel introduced it in 1908 (Bergmann, 2014). However, it does not have the ability to distinguish latent from active TB and any previous

exposures to TB, including through BCG vaccinations and non-tuberculosis causing mycobacteria (NTM), will give a false positive result (Cohn et al., 2000).

Another common test, known as the Acid-Fast Bacilli (AFB) smear test, could provide a presumptive result within a few hours. However, its sensitivity is highly lacking as it would need sputum samples to have a bacterial load of between 5,000 to 10,000 bacilli per millilitre. This is in contrast to bacteria culture method which would only require 10 bacilli per millilitre to grow (Lawn et al., 2013; Parsons et al., 2011).

1.3.3 ESAT6-Based TB Diagnostic Assay

Many studies on ESAT6-focused TB diagnostic assay were done with the protein of interest being the antigen that then stimulates the production of interferon-γ (Munk, Arend, Brock, Ottenhoff, & Andersen, 2001; Pollock & Andersen, 1997; Ravn et al., 2005; van Pinxteren et al., 2000). In other words, the diagnostic assays were constructed to detect the IFN-γ, and not ESAT6 itself.

For direct detection of TB via the ESAT6 protein, a group comprising of researchers from the Zhejiang University School of Medicine, Fudan University School of Medicine and Nanjing Medical University reported to have successfully developed a new anti-ESAT6 monoclonal antibody (mAb). The group used the mAb in a sandwich Enzyme Linked Immunoabsorbent assay (ELISA) diagnostic assay for the detection of *M. tuberculosis* in sputum samples (Feng et al., 2011; Leng et al.,

2014). The anti-ESAT6 mAb were reported to have 95.4% and 100% sensitivity and specificity respectively. However, there are several disadvantages associated with antibodies as mentioned in section 1.1.2. This prompts the need to develop alternative molecular recognition element against ESAT6.

Research on using ESAT6 aptamer to screen for tuberculosis were few and far between. To date, only one group from Wuhan University School of Basic Medical Sciences have been working on the development of ESAT6 aptamer as a diagnostic tool (Tang et al., 2014). Another group from University of Pretoria, South Africa, have attempted to select an aptamer against ESAT6 but ultimately was only successful in getting aptamers against CFP10, the binding partner of ESAT6 (Rotherham, Maserumule, Dheda, Theron, & Khati, 2012). Their work on ESAT6 aptamer selection, use in diagnostic assays and their respective assay sensitivity and specificity will be discussed further in Chapter 4.

The lack of studies into using ESAT6 as a biomarker for the detection of TB means that there are still plenty of new grounds to break and hopefully this particular research project will be a precursor to further development in diagnostics and therapeutics.

1.4 Objective

As a class of molecules that has a cutting edge over the antibodies in many regards, generation of an aptamer against ESAT-6 could lead to the development of an aptamer-based diagnostic assay of TB. Prior to its exploration as a diagnostic element, characterisation of the aptamer is necessary, which can provide insights into its binding affinity and structural elements. Thus the objectives of this study is:

1. To generate RNA aptamer specific against ESAT-6 protein.

2. To perform experimental and bioinformatic-based structural characterisation of the aptamer isolated against ESAT-6.
    a) Estimation of the binding affinity of the isolated aptamer against ESAT6
    b) Secondary structure determination of the aptamer
    c) Truncation of the aptamer

CHAPTER 2: MATERIALS AND METHODS

2.1 Designing Degenerate DNA library and Primers

The aptamer pool and the corresponding primer pair used in this study were synthesised by Integrated DNA Technology (IDT, Iowa, USA). Each sequence in the pool were 107 nucleotides in length and consisted of two constant primer binding sites at either ends with a 40 nucleotides randomised region between them. The DNA library has the following sequence;

5'-GGGGGAATTC**TAATACGACTCACTATA**GGGAGGACGATGCGG-(N40)-CAGACGACTCGCTGAGGATCCGAGA-3'

It is constructed using conventional solid phase phosphoramidite oligonucleotide synthesis. Each position of N can be occupied by either A, T, C or G nucleotide with a coupling efficiency of 25% each. The N40 Forward and Reverse primers' sequences are underlined. The bolded nucleotides in the forward primer represents the promoter sequence for the T7 Polymerase enzyme. The incorporation of the T7 promoter sequence allows for the synthesis of single stranded RNA from double stranded DNA molecules after each amplification round in the SELEX cycles. ESAT-6 protein was purchased from Abcam (Bristol, United Kingdom).

2.2 PCR Amplification

One µM of the N40 single stranded DNA oligonucleotide was subjected to PCR using BioRad MyCyler Thermal Cycler in a 100 µL of the reaction mixture comprising 1X PCR buffer (10 mM Tris-HCl [pH 8.8], 50 mM KCl, 0.08% [v/v] Nonidet P40), 0.2 mM dNTP (Biotools, Madrid, Spain), 0.6 µM N40 forward and reverse primers, double distilled water (ddH_20), 1.5 mM $MgCl_2$ and 5 U *Taq* Polymerase (Thermo Fisher Scientific, Massachusetts, USA). The PCR protocol used is detailed as follows; initial denaturation at 95^oC for 60 s, followed by 8-16 cycles of denaturation at 95^oC for 30 s, annealing at 55^oC for 30 s, extension at 72^oC for 30 s and final extension at 72^oC for 15 min. PCR cycles were set at a minimum of 8 cycles and additional cycles were added if PCR product observed on gel electrophoresis. The number of PCR cycles were kept low where possible to minimise the accumulation of PCR artefacts and reduce the chances of *Taq* polymerase misincorporation of nucleotides The resulting PCR product was analysed by gel electrophoresis on a 4% agarose gel in 1X TAE buffer at 90 V for 20 min. The gel was later visualisation with Gel Doc™ XR+ Gel Documentation System (BioRad, California, USA).

2.3 Ethanol Precipitation

PCR products obtained in each SELEX cycles were subjected to ethanol precipitation before the start of the *in vitro* transcription reaction. The PCR product was added to 300 µl ddH_20, 40 µL 3 M pH 5.2 sodium acetate (NaOAc), and 1 mL absolute ethanol solution (Merck KGaA, Darmstadt, Germany). This mixture was vortexed vigorously

and incubated for 30 min at -80ºC. Afterwards, the mixture was centrifuged at 13,000 rpm, 4ºC for 15 min using Sorvall™ ST 16R centrifuge (Thermo Fisher Scientific, Massachusetts, USA). The supernatant was discarded and 1 mL of 70% ethanol was added into the tube before being centrifuged at 13,000 rpm, 4ºC for 2 min. The supernatant was discarded once more and the pellet was air-dried in a vacuum concentrator (Eppendorf, Hamburg, Germany) for 5 min. The pellet was later resuspended in 8 µL ddH_2O.

2.4 *In Vitro* Transcription

In vitro transcription was carried out using Ampliscribe™ T7-Flash™ Transcription kit (Epicentre, Wisconsin, USA). The reaction mixture was set up to contain 4 µL of the precipitated PCR product, 1X AmpliScribe T7 reaction buffer (40 mM Tris-HCl [pH 7.5], 6 mM MgCl2, 10 mM NaCl, 2 mM spermidine), 7.5 mM each of ATP, CTP, UTP, GTP, 10 mM of dithiothreitol (DTT), 20 U of AmpliScribeTM T7 Polymerase Flash Enzyme Solution (10 U/µL) and 20 U of RiboGuard RNase Inhibitor (40 U/µL). The reaction was incubated at 37ºC for 16 hours. Following that, 2 U of DNase I was added and the reaction was incubated for a further 20 min to remove any template DNA in the mixture. The reaction was later stopped with the addition of 2X RNA dye and it was heated at 95ºC for 2 min followed by snap-cooling on ice. The *in vitro* transcription reaction mixture was later run on a denaturing urea-PAGE and was purified via rapid crush-and-soak method.

2.5 Denaturing Urea-PAGE

The 10% 7 M denaturing urea PAGE gel was prepared with 1X TBE, 10% Acrylamide/Bis-acrylamide solution (BioRad, California, USA), 7 M urea (Merck KGaA, Darmstadt, Germany), ddH_2O, 0.1% TEMED (BioRad, California, USA) and 1% APS in a BioRad Mini Protean Tetra Cell system (BioRad, California, USA). Electrophoresis was carried out using BioRad Mini-PROTEAN 3 Electrophoresis Cell (Bio-Rad Laboratories, Hercules, USA) in 1X TBE buffer at 140 V for 60 min.

2.6 Rapid Crush-and-Soak-based RNA Purification

The *in vitro* transcribed RNA was purified by rapid crush and soak method which was adapted from the technique outlined by Citartan et al. (Citartan, Tan, & Tang, 2012). Upon the completion of the electrophoresis and the removal of the glass plates, the gel was then wrapped with saran wrap and placed on a silica coated plate. The location of the RNA on the gel was detected by UV shadowing. The gel was shone with a hand held short wave (254 nm) UVGL-58 UV lamp (UVP, California, USA) in a dark room and the image produced was marked. The marked area of the gel was cut and transferred to a new centrifuge tube for it to be crushed with a pipette tip. Four hundred microliters of ddH_20 was added to it and the tube was then incubated at 50^oC for 30 min. Afterwards, 1 mL of absolute ethanol was added into the tube. The mixture was shaken vigorously and centrifuged at 15,000 rpm for 2 min. The supernatant was transferred to a new tube and added with 40 µl sodium acetate. The tube was incubated for 30 min at -80^oC. The subsequent ethanol precipitation steps followed as described

in section 2.3. The pellet was then resuspended in ddH_20 and the resulting RNA concentration was determined using a NanoPhotometer P300 (Implen GmbH, Munich, Germany). Following quantification, an appropriate amount of RNA was used for each cycle of SELEX.

2.7 SELEX Cycles

In this study, each SELEX cycle carried out consisted of an iterative process of incubation of the nucleic acid pool and the target (2.7.1), partitioning to separate the target-bound and unbound nucleic acid molecules (2.7.2), elution of the target-bound nucleic acid molecules (2.7.3), amplification (2.7.4), counter selection (2.7.5) and sequence analysis (2.8)

2.7.1 Incubation of Nucleic Acid Pool and Target

For the 1st cycle of SELEX, the reaction mixture was set up to contain 9.5 µM of the initial nucleic acid pool, 20 µM of the ESAT6 protein, 100 µM yeast tRNA (Invitrogen Corporation, Carlsbad, USA) in 1X SELEX binding buffer (10 mM HEPES-KOH [pH 7.4], 150 mM NaCl) (GE Healthcare Life Sciences, Buckinghamshire, UK). Prior to the addition of the target protein and yeast tRNA, the mixture was heated at 95^{o}C followed by cooling at room temperature for 5 min. After protein and yeast tRNA has been added, the reaction was incubated at room temperature for 15 min.