APPLICATION OF *Aequorea victoria* GREEN FLUORESCENT PROTEIN (GFP) FOR EXPRESSION IN *Mycobacterium bovis* BCG

by

KHAIRUL EZANI BINTI NAJAMUDIN

Thesis submitted in fulfillment of the
requirements for the degree
of Master of Biomedicine (Molecular Biology and Cell)

November 2007

This thesis is dedicated to my beloved parents, brothers, yusran and his family for their patience's and encouragements.

THE GREEN GENE
By Christopher Dinesh Raj

Oh, fluorescent one

From the time they cloned you
And the time, I first set my eyes upon you
Something drives us
To do something new with you

You have been emitted and excited
Mutated to yellow, cyan and blue
Broken and permutated

You have been driven by a promoter
Be it viral T7 or the brain's own L7
Been expressed in every cell type
Fused to the you-name-it gene at the N or C terminal

Confused, I am

The cells don't mind you
And the Ni-Ta columns will bind you
Provided you have been His-Tag

Only to make you pure

You have been glorified here
And stolen there

Still you shine like an emerald

Allow me to make a new construct out of you
Mutate and transform you
Shine blue light upon of you
And I am certain my peers will turn green
Yes, green with envy

Because my GFP is brighter

ACKNOWLEDGEMENT

All praise and thanks are due to ALLAH; the possessor of all the Excellencies; for gratuitously giving me the ingredients of success. Invoke the blessings of ALLAH on the noble Prophet Mohammad peace is upon him, who taught us to be thankful.

Over the span of time during this project was conducted, I have received assistance and/or advices from numerous people whom I would like to acknowledge at this time.

It is my great honour to thank my supervisor, Prof. Zainul Fadziruddin Zainuddin, for his support, excellent guidance and supervision throughout all the experimental work, research investigation and providing all the necessary facilities to carry out in this study. It is my great pleasure to work under your supervision which gave me a tremendous experience as a postgraduate student.

I wish to thank my Co-supervisor, Dr. Shaharum Shamsuddin, for many insightful conversations, advice and kind guidance at all times. His steadfast guidance and constant accessibility are greatly appreciated. It has been a privilege to work with you.

Thanks also to Prof. Norazmi Mohd. Nor, School of Health Sciences and Assoc. Prof. Dr. M. Ravichandran, School of Medical Sciences for supporting my experimental work. I would also like to extend my appreciation to all my friends and colleagues in the Advance Molecular Laboratory and Immunology Laboratory, USM, who provided both friendship and assistance during the course of this study.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Ab	Antibody
Amp	Ampicillin
bp	Base Pair
BCG	*M. bovis* bacilli Calmette-Guérin
°C	Degree Celsius
Da	Dalton
ddH_2O	Deionized distilled water
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
dNTPs	Deoxy nucleotides triphosphate
HIV	Human immunodeficiency virus
IPTG	Isoprophyl-β-D-thiogalactopyranoside
Kan	Kanamycin
kb	Kilobase
kDa	Kilodalton
L	Liter
M	Molar
mM	Milimolar
µl	Microlitre
µm	Micrometer
mA	Mili-ampere
min	Minute
mg	Miligram
ml	Mililitre
MW	Molecular weight
ng	Nanogram
OD	Optical density
PCR	Polymerase chain reaction
Pfu DNA polymerase	*Pyrococcus furiousus* DNA polymerase
%	Percentage
RNase	Ribonuclease
rpm	Rotations per minute
RT	Room temperature
SDS	Sodium Dodecyl Sulphate

Sec	Seconds
Taq DNA polymerase	*Thermus aquaticus* DNA polymerase
Tm	Melting temperature
TB	Tuberculosis
U	Unit
UV	Ultraviolet
V	Volt
v/v	Volume/Volume
w/v	Weight/Volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactosylpyranoside

LIST OF APPENDICES

PENGGUNAAN PROTIN BERPENDAFLUOR HIJAU (GFP) *Aequorea Victoria* UNTUK PENGEKSPRESAN DI DALAM *Mycobacterium bovis* BCG

ABSTRAK

Protin pendafluor hijau (GFP) yang berasal daripada obor-obor, *Aequorea victoria* adalah merupakan salah satu penanda yang berpotensi dan mempunyai banyak kelebihan sebagai penanda yang praktikal kerana ia tidak memerlukan sumber luaran seperti substrat dan faktor bersama untuk berpendafluor walaupun masih bergantung kepada sel perumah yang hidup. Ia juga dilaporkan pernah digunakan di dalam *M. tuberculosis*, tetapi dengan penggunaannya di dalam strain patogenik ini bagi tujuan penyaringan asai berkadar tinggi boleh menimbulkan masalah keselamatan. Oleh yang demikian, strain selamat iaitu strain vaksin *Mycobacterium bovis* BCG telah digunakan untuk mengekpreskan penanda bagi memudahkan pemantaun pemilihan asai secara berkala tanpa perlu mematikan sel bakteria. Di dalam kajian ini, gen asal *GFP* telah dibangunkan semula secara sintetik dengan menggunakan gen yang berkecenderungan kepada kodon mikobakteria dan potensinya sebagai penanda yang praktikal telah diakses di dalam strain vaksin mikobakteria, *M. bovis* BCG. Gen sintetik yang dinamakan sebagai *EzyGFP* telah berjaya dibangunkan dengan menggunakan teknik himpunan tindakbalas rantaian polimerase (PCR). Di dalam pendekatan yang kedua, versi *GFP* yang telah diubahsuai iaitu *GFPuv* telah digunakan sebagai jalan alternatif. Gen ini telah dibanyakkan secara asli di dalam *Escherichia coli* sebelum vektor perumah mikobakteria telah dibangunkan bagi tujuan memasukkan gen ini ke dalam *M. bovis* BCG. Bagi menghasilkan vektor perumah mikobakteria, gen yang mengkodkan replikasi asal mikobakteria telah dikeluarkan daripada plasmid sumber, pMNM012 dan dimasukkan ke dalam pEzyGFP, pGFPuvK dan pGFPuvK65 secara berasingan,

yang mana masing-masing mempunyai gen *GFP* tertentu. Plasmid yang terhasil kemudiannya dimasukkan ke dalam *E. coli* Top 10 dan *M. bovis* BCG sebelum pemblotan western dan pengesanan pendafluor menggunakan fluorometer dilakukan bagi membuktikan pengekpresan protin terbabit. Kajian pengekpresan secara pemblotan western di dalam *E. coli* Top 10 dan *M. bovis* BCG telah membuktikan kewujudan EzyGFP di dalam bentuk protin bakteria secara keseluruhan. Walaubagaimanapun, kehadiran pendafluor di dalam *E. coli* Top 10 dan *M. bovis* BCG rekombinan tidak dapat dikesan. Keputusan ini mencadangkan bahawa protin telah diekpreskan dalam bentuk yang tidak larut sebagai badan-badan terkandung yang mana tidak aktif kerana lipatan GFP yang salah. Di dalam pendekatan yang kedua, GFPuv telah ditemui diekpreskan di dalam bakteria supernatan secara pemblotan western di dalam *E. coli* Top 10 dan *M. bovis* BCG rekombinan. Walau bagaimanapun, isyarat pendafluor hanya dapat dikesan di dalam *E. coli* Top 10 rekombinan yang mengekspreskan GFPuv hanya setelah isyarat peptida dibuang. Oleh yang demikian, kegunaan isyarat peptida dalam perembesan GFP dapat disimpulkan boleh menimbulkan masalah dalam menghasilkan lipatan gen yang baik. Menariknya, peningkatan penghasilan GFPuv di dalam *E. coli* Top 10 rekombinan dapat dilihat apabila ekpresinya didorong oleh promoter HSP65 yang berasal daripada mikobakteria. Walaupun *M. bovis* BCG rekombinan yang mengekpreskan GFP protein secara efektif tidak berjaya ditemui di dalam kajian ini, data yang terhasil adalah berguna untuk penyelidik lain bagi menghasilkan satu sistem penyaringan yang selamat bagi ubat anti-tuberkulosis.

APPLICATION OF *Aequorea victoria* GREEN FLUORESCENT PROTEIN (GFP) FOR EXPRESSION IN *Mycobacterium bovis* BCG

ABSTRACT

A potential marker, green fluorescent protein (GFP) derived from the jellyfish *Aequorea victoria* offer many advantages as a viability reporter, as it requires no external source of substrate nor cofactors to fluoresce but is dependent on the host cell being alive. Its use has been reported in *M. tuberculosis*, but the incorporation of potentially pathogenic strains for high throughput screening would raise issues of safety. Therefore, a safe strain such as the vaccine strains *M. bovis* BCG if made to express an easily detectable viability reporter marker would be useful tool as a screening assay. In this study, a synthetic version of the wild type *GFP* gene sequence incorporating mycobacterial codon bias, was used to explore its potential as a marker of viability in the mycobacterial vaccine strain, *M. bovis* BCG. The synthetic gene designated as *EzyGFP* was successfully constructed by using assembly polymerase chain reaction (PCR). In the second approach, a modified version of *GFP*, *GFPuv*, was also tested as an alternative candidate. To propagate these genes in *Escherichia coli* and then deliver these genes into *M. bovis* BCG, mycobacterial shuttle vectors were constructed. To create mycobacterial shuttle vectors, the mycobacterial origin of replication was excised from a source plasmid, pNMN012, and ligated into plasmids designated as pEzyGFP, pGFPuvK and pGFPuvK65 respectively which contain the *GFP* gene of interest. The resultant plasmids were grown in *E. coli* Top10 and *M. bovis* BCG and expression was confirmed by Western blotting and fluorometry. Expression studies by Western blotting in *E. coli* Top10 and *M. bovis* BCG successfully detected the presence of EzyGFP in total bacterial protein (pellet). However, fluorescence in both *E. coli* and

M. bovis BCG were not detectable. These results suggest that the protein was expressed in insoluble form as inclusion bodies which are inactive due to incorrect GFP folding. In the second approach, GFPuv protein was successfully detected in bacterial supernatant by Western blotting in recombinant *E. coli* Top10 and *M. bovis* BCG. However, the fluorescence signal was only observed in the recombinant *E. coli* Top10 expressing GFPuv after removal of the mycobacterial signal peptide. Thus, the uses of the signal peptide for GFP secretion appear to be detrimental to the proper folding of the gene. Interestingly, an enhanced production of GFPuv in recombinant *E. coli* Top10 was observed when its expression was driven by the mycobacterial HSP65 promoter. Although an effective *M. bovis* BCG strain expressing the GFP protein for rapid discovery of anti-tuberculous compounds did not materialize in this study, the data generated would be useful for other researchers to develop a safe screening assay system for anti-tuberculous drugs.

CHAPTER ONE

1.0 Introduction

1.1 History of Tuberculosis

Tuberculosis (TB) is an infectious disease, which is caused by a bacterial pathogen called *Mycobacterium tuberculosis* (MTB). It has been present in human populations since more than 4 000 years ago (Donoghue *et al.*, 2004). It is evident by the discovery of a spinal fragment from an ancient Egyptian mummy that showed definite pathological signs of TB decay as early as 3000 BC (Bedeir, 2004). Around 460 BC, an ancient Greece philosopher, Hippocrates, identified the widespread disease as phthisis and it was believed to be an inherited and non-contagious disease (Madkour *et al.*, 2004b). It is also believed that diseases named as phthisis (wasting), white plague, Pott' disease (TB of the bones), lupus vulgaris (TB of the skin) and consumption (the "classic" case of lung disease) (Lerner, 1997) are all the same disease which is nowadays called tuberculosis.

In his Opera Medica of 1679, Franciscus De La Boe or Sylvius identified the actual tubercles as a consistent and characteristic change in the lungs and other areas of consumption patients which lead towards the development of abscesses and cavities (Iseman, 2000). He believed the disease to be contagious. According to Benjamin Marten's publication in 1720, TB was believed cause by specific living creatures, which only attack certain people. By speculating the nature of transmission of consumption patients, he understood that TB was a much widespread problem. Then, in 1865 an army surgeon, Jean-Antoine Villemin had elaborated a series of experiments on the inoculation of animals with tuberculous tissue and the transmission of this "creature" from animal to animal (Barnes, 2000, Osoba, 2004). On the basis of this evidence, he concluded that the disease was caused by a specific microorganism which had never been seen by him.

The specific agent described by Villemin was finally identified by Robert Koch in 1882, a German physician. In a presentation to the Berlin Physiological society, Koch announced that he had succeeded in identifying and culturing the causal agents of TB: a rod-shaped bacterium called tubercle bacillus (Kaufmann and Schaible, 2005, Barnes, 2000). This bacterium was known as *M. tuberculosis* in 1886 (Grange, 1982). Koch also described a standard principle for identification and isolation of specific infectious microorganisms, which is now known as Koch's Postulates (Barnes, 2000).

1.2 Epidemiology of Tuberculosis

TB remains a significant global public health problem representing a major cause of mortality and morbidity. According to the World Health Organization (WHO), one third of the world's population is infected, with more than 23 000 people developing active TB and close to 5 000 mortality everyday. Each year there are an estimated eight million new cases of TB and it is estimated that around 1.7 million deaths annually occur due to TB (WHO, 2002). If the current trend continues, it is estimated that over the next 20 years, there will be one billion new TB infections and 36 million people will die with one every nine seconds (TB Alliance, 2003). The predicted scenario suggests that in the next decade another 90 million people will develop TB and 30 million people will die (Glassman, 2004). In order to control TB, a comprehensive cost-effective strategy called direct observed therapy (DOTS) was designed aimed at reducing mortality, morbidity and transmission of TB. However, it is estimated that only 32% of new smear-positive tuberculosis cases worldwide were managed under this strategy in 2001 (Espinal and Raviglione, 2004). The emergence of multidrug resistant tuberculosis and co-infection with the Human Immunodeficiency Virus (HIV) has posed additional challenges to tuberculosis control. In year 2001, there were 3 813 109 (62 per 100 000 people) cases notified

worldwide, of which 50% was notified in Asia and 21% in Africa (Espinal and Raviglione, 2004).

In 19th century, TB started to decline in Europe due to industrialization, better sanitation and improvement of living conditions. Nowadays, TB in Western Europe, Australia, Canada, Japan and the United States only becomes the problem of selected groups while the general populations are minimally effected (WHO, 2003b). TB infection occurs predominantly among immigrants from high prevalence settings, the homeless, intravenous drug addicts and HIV infected people. Previous studies have shown a rapid spread of TB occurring in crowded living conditions, such as in nursing homes, hospitals (Baum and Wolinsky, 1983), homeless shelters, schools, military barracks, and prisons (Bignell, 1994, Brewer *et al.*, 2001). This is because TB is a disease of poverty, which affects the poorest people in the world's poorest countries (WHO, 2001) such as India, Indonesia, Bangladesh, Vietnam, Cambodia, Thailand and Myanmar. It is estimated every six out of ten TB cases world wide live in Asia and the Pacific region (Narain and Ying-Ro, 2004).

According to the WHO Report in 2003, more than 90% of global TB cases and death occur in the developing world, where 75% of cases are in the most economically productive age group (15 – 54 years). It is estimated, an adult with TB will lose an average of three to four working months, which causes the lost of 20 – 30% of an annual household income. The widening gaps between rich and poor in various populations and the collapse of health infrastructures in countries facing severe economic crisis or civil unrest became the reasons for the increase of global TB burden. Other reasons include inadequate case detection, diagnosis and cure, and the impact of human immunodeficiency virus (HIV) pandemic in most developing countries (WHO, 2003b). TB will act in lethal synergy with HIV virus; doubling the

infection of many patients, weakening their immune systems and hastening death (Kaufmann and Schaible, 2005).

Over the past ten years in Malaysia, TB showed a steady increase in yearly notification of new cases which include infectious and other forms of TB. In the year 2000, 15 507 cases was notified and TB was announced as the second highest communicable disease in 2001 (Iyawoo, 2004). In the year 2002, the incidence rate per 100 000 population was 58.67, a slight decrease of cases compared to 2001 (CDC Malaysia, 2006). The highest disease burden state is Sabah which had 3 908 notified cases in 2001 (Dony *et al.*, 2004) followed by Wilayah Persekutuan, Sarawak (1 885 cases) (Dony *et al.*, 2004) and Pulau Pinang. Out of these numbers, 10% of the notified cases were among immigrants from high TB burden neighboring countries (Iyawoo, 2004).

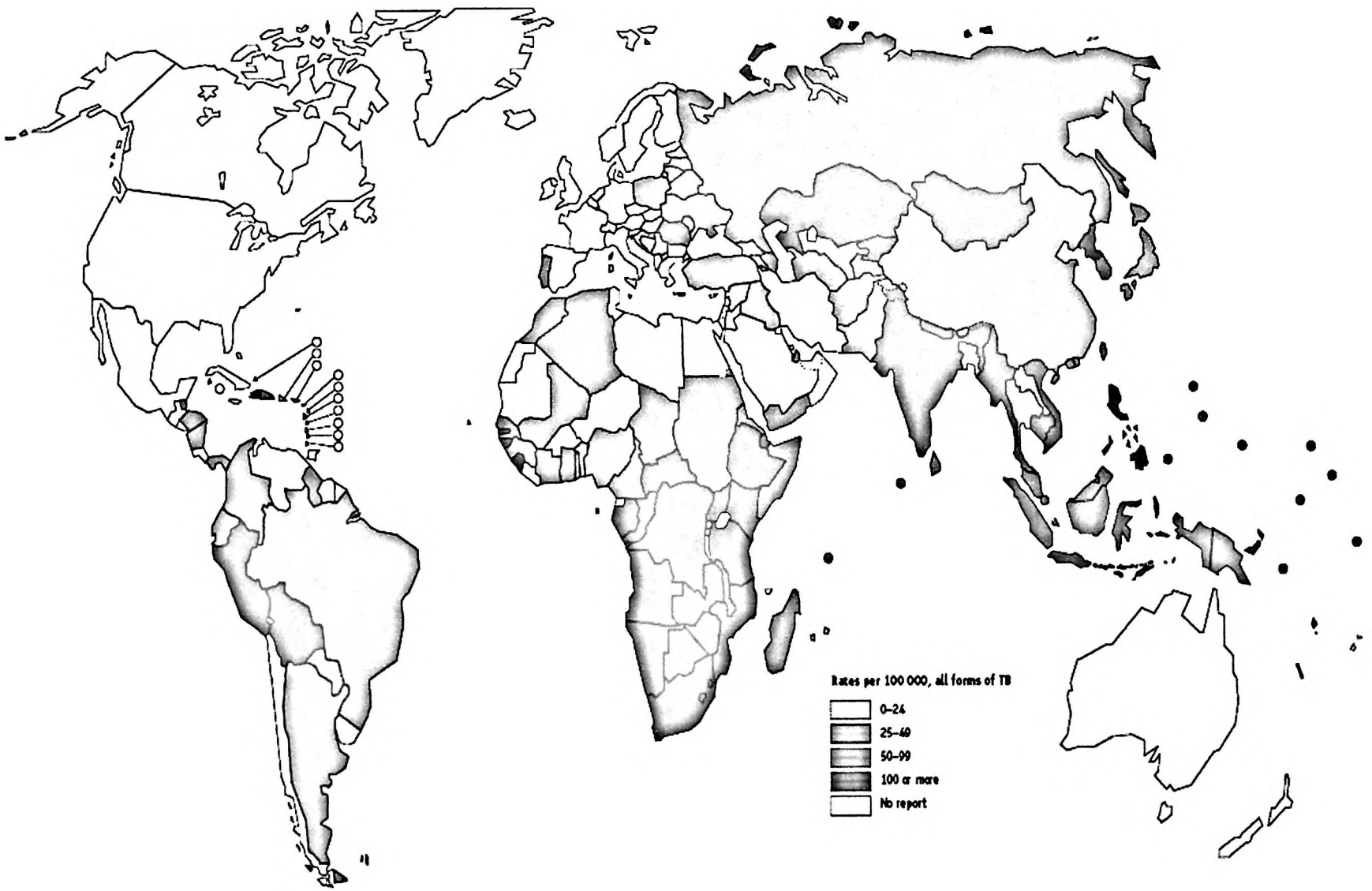

Figure 1.1 Tuberculosis notification rates by country, 2002 (Adapted from WHO, 2004)

Table 1.1: Incidence of tuberculosis cases in Malaysia from 1998 until 2005. (Adapted from CDC Malaysia, Kementerian Kesihatan Malaysia).

Incidence rate of tuberculosis (all forms) in Malaysia for 1998 – 2005 (per 1000 of populations)							
1998	1999	2000	2001	2002	2003	2004	2005
63.6	65.6	67.8	62.3	58.67	52.47	54.50	36.19

Incidence of tuberculosis (all forms) in Malaysia for 1998 – 2005															
1998		1999		2000		2001		2002		2003		2004		2005	
Case	Death	Case	Death	Case	Death	Case	Death	Case	Death	Case	Death	Case	Death	Case	Death
14115	1059	14908	1191	15057	1295	14830	1326	14389	1290	13144	249	13942	310	9456	135

1.3 Mycobacteria

Mycobacterium is the only genus in the Family *Mycobacteriaceae,* which is found within the Order Actinomycetales. Etymologically, "mycobacterium" is derived from Greek word, fungus (*myces*) and small rod (*bakterion*). The use of the name fungus was derived from the tendency of these microorganisms to spread diffusely in a mold-like growth pattern in liquid medium (Iseman, 2000). Mycobacterium is often classified into the *Mycobacterium tuberculosis* complex, Non-tuberculous mycobacteria (NTM) and *Mycobacterium leprae* (Levinson and Jawetz, 2000). The *M. tuberculosis* complex consists of MTB together with *M. bovis, M. africanum* and *M. microti.* Meanwhile, NTM refer to all other species in the mycobacteria family, most of which are saprophytic but some may cause human diseases.

On the basis of their growth rates, mycobacteria can be divided into two main groups. The rapid growing bacteria, usually non-pathogenic are seen as visible colonies which are grown from diluted inoculums in less than seven days. Meanwhile, the potentially pathogenic slow-growing mycobacteria usually need several weeks before visible colonies appear (Kaufmann and Hahn, 2003). Mycobacteria are obligate aerobes, straight or slightly curved bacteria which vary from 0.3 to 0.6 by 1.0 to 10 μm in size, sometimes with branching (Iseman, 2000). They are nonmotile and nonspore forming, with high lipid content cell walls which include mycolic acids. Outside of the genus *Mycobacterium*, mycolic acids found only in *Nocardia* and *Corynebacterium* (Tortora *et al.*, 2001).

Biochemical and electron microscopic studies indicated that the MTB cell wall consist of four layers. The innermost layer is composed of a peptidoglican layer whereas the next three surface layers are composed of lipids. These lipid complexes in the cell wall consist of peptidoglicolipids (mycosides) cord factor and sulpholipids

that give shape and rigidity. Mycosides is functionally similar to the O-antigen which determines sero-agglutination and bacteriophage susceptibility of the organism (Osoba, 2004).

The mycobacteria cells are hydrophobic; therefore, they tend to clump together and are impermeable to the usual stains unless the dyes are combined with phenol. Once stained, the cells resist decolourization with acidified organic solvents and are therefore called "acid fast" (Osoba, 2004). This characteristic reflects the unusual composition of cell wall, which is responsible for the resistance of mycobacteria to environmental stresses, such as drying. In fact, these bacteria can survive for six to eight months in dried sputum and are very resistant to chemical antimicrobials used as antiseptics and disinfectant. They are also resistant to 3% of HCl and 6% of H_2SO_4, and to 4% of NaOH. However, mycobacteria are sensitive to moist heat at 60°C, disinfectants such as alcohol, glutaraldehyde and formaldehyde, and ultraviolet (UV) radiation (Tortora *et al.*, 2001).

MTB genomic DNA also contains high guanine and cytosine (GC) content, ranging from 62-70%, which affects the utility of *Escherichia coli* as a surrogate genetic host (Tortora *et al.*, 2001). The genome of MTB comprised of a single circular chromosome of approximately 4.4 million base pairs (bp) with about 4, 000 predicted open reading frames (ORF), including 56 insertion sequences and two prophages (Gordon *et al.*, 1999, Castell *et al.*, 2005). A complex regulatory network consisting of 103 repressors and activators, 13 sigma factors and 11 two-component systems appear to be involved in the regulation of gene expression (Cole *et al.*, 1998, Gordon *et al.*, 1999).

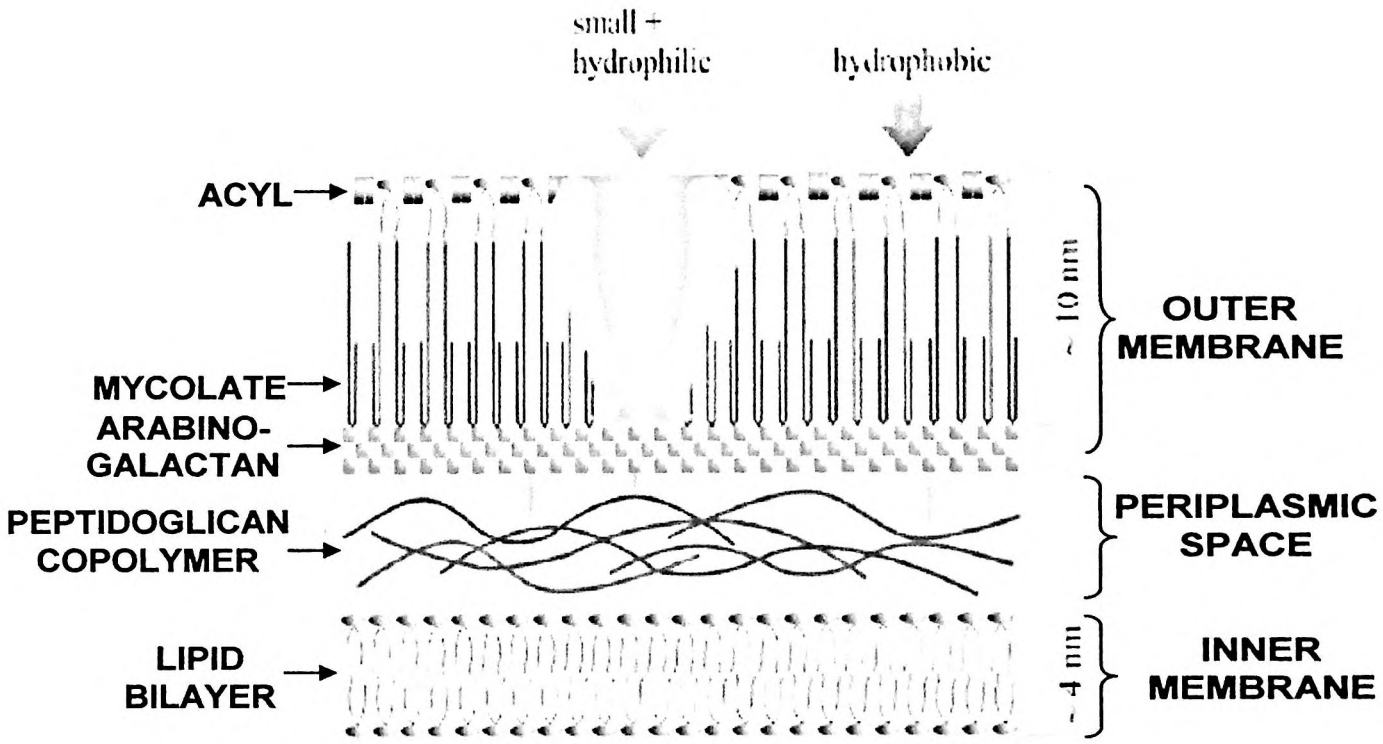

Figure 1.2: Illustration of composition of outer and inner membrane of mycobacteria cell wall. (Illustration was adapted and modified from http://www.diss.fu-berlin.de/2005/142/kap1.pdf)

1.4 Pathophysiology of TB

1.4.1 Tuberculosis: The Disease

Tuberculosis is defined as a state of suffering from active, progressive invasion of an organ or organs by *M. tuberculosis* characterized pathologically by formation of granulomas and by cell-mediated hypersensitivity. This disease is typically manifested by constitutional symptoms that correlate with the specific organ system affected. The main reservoir of *M. tuberculosis* is the patient with pulmonary tuberculosis. Some of the pulmonary cavities are rich with bacilli with up to 100 million bacilli in a cavity of approximately two cm in diameter (Ait-Khaled and Enarson, 2003). Pulmonary tuberculosis refers to the disease that involves the lung parenchyma while extra pulmonary tuberculosis involves organs other than the lung such as the pleura, lymph nodes, genitourinary tract, skin, joints, bones and meninges (WHO, 2003b). Extra pulmonary TB is very common among patients co-infected with HIV. In HIV patients with pulmonary TB, 60-80% will develop extra pulmonary TB in contrast to 17% of non-HIV patients with pulmonary TB.

After being infected, a series of immune response occurs and determines whether the infection will be eradicated, becomes latent or develop progressive primary tuberculosis that will result in the development of an active disease (Daley, 2004). The infection is often clinically identified by detection of a delayed-type hypersensitivity reaction to an intradermal injection of purified protein derivative (PPD) which is a mixture of mycobacterial antigens. In primary tuberculosis, the disease will result in calcified lesions at the initial site of infection called Ghon lesions or granulomatous focus (Daley, 2004). This primary focus is the site of tuberculosis-specific caseating necrosis which contains 1000 - 10 000 bacilli (Ait-Khaled and Enarson, 2003). The bacilli will lose their viability and multiply more slowly, and survive for months or years.

The weakening of cellular immune response (immunosuppression) will cause endogenous reactivation which will lead to the development of active (post primary) TB. In the absence of treatment and immune deficiency, the risk of post primary TB is estimated at about 5-10% in the year after primary infection and 5% for the remainder of the individual's life span (Ait-Khaled and Enarson, 2003). It was estimated that approximately 90% of reinfection in adult cases are due to endogenous reactivation of latent infection. Active TB can be also caused by exogenous reinfection such as those infected by human immunodeficiency virus (HIV) or in patient with diabetes mellitus, cancer of the neck and head, silicosis, leukemia, substance abuse and low body weight (CDC, 2002).

1.4.2 Symptoms of TB

Most people infected with TB often have no symptoms or may only experience a flu-like illness. However, in active disease in the secondary stage, the symptoms include cough, chest pain, fatigue, weight loss, poor appetite, swollen glands, chills, fever, and sweating at night (CDC, 2002). Tuberculosis of the lung is usually symptomatic by dry cough that eventually result in productive cough with mucoid, muco-purulent, blood-stained or with massive sputum haemoptysis (Madkour *et al.*, 2004a) and patients usually develop severe breathing problems. Meanwhile, symptoms that are related to extrapulmonary tuberculosis includes tuberculous monoarthritis, Pott's disease, genitourinary symptoms, or those related to the organs involved but may correlate with respiratory symptoms at the time of presentation.

1.4.3 Transmission of TB

Transmission of tuberculosis occurs by the airborne or aerosol route and through the gastrointestinal tract by ingestion of inoculums. Coughs (containing approximately 3500 bacilli) or sneezes (~1 million bacilli) results in active TB droplet nuclei which can remain suspended in the air for several hours (Ait-Khaled and Enarson, 2003). The infection occurs when the organism is inhaled and reaches the alveolus of the lungs. Everyone is susceptible to *M. tuberculosis* through airborne particles regardless of age, sex, and race. However, not everyone becomes infected or develops clinical symptoms of TB. There are at least three factors influencing the disease transmission; firstly, the number of viable bacilli in a patient's sputum and the concentration of mycobacteria-containing droplets in the air. Secondly, the frequency and duration of an exposed person breathing contaminated air and lastly the immune status of the exposed individual (Horsburgh, 1996). It is estimated only one in ten of immunocompetent people infected with *M. tuberculosis* will develop active disease in their lifetime (CDC, 2002).

1.4.4 Diagnosis of TB

The tuberculin test is a serological method that was used as detection and to determine the prevalence of infection in human populations. Tuberculin contains a number of polyantigenic proteins derived from metabolic products of *M. tuberculosis* bacilli (Ait-Khaled and Enarson, 2003). The injection of tuberculin will provoke delayed type hypersensitivity immune response that is demonstrated by the appearance of localized infiltration of inflammatory cells into the skin after several hours.

Identification of TB in laboratory clinical specimens conventionally relies on microscopic examination of acid fast-stained sputum sediment or other specimens

(AFB smears) (WHO, 2003a, Hanna, 2004). Other staining method includes fluorochrome and Kinyoun staining. However, only the Ziehl-Nelseen method is recommended by the International Union against TB and Lung Disease (IUATLD) and WHO (Enarson *et al.*, 2000). Although the sensitivity is not too high (50–80% of patients with pulmonary TB will have positive smear) this method is still useful as cases with positive smears contribute to higher fatality rates compared to negative smear cases (Frieden *et al.*, 2003). Microscopic diagnosis is followed by culturing of the specimens and confirmation of the identification of bacterial colonies grown by definitive biochemical tests or PCR-based assays. The sensitivity of culture is approximately 80–85% with a specificity of 98% (Strang *et al.*, 1987, Schoeman *et al.*, 1997).

There are three different types of traditional culture media available which are egg based (Lowenstein-Jensen and Ogawa), agar based (Middlebrook 7H10 and 7H11 media) and broth (Middlebrook 7H9 and 7H12 media). Traditional detection method requires two to eight weeks to obtain the primary culture and an additional several weeks for drug-susceptibility testing to be performed. Although the organism takes longer period to grow in solid culture compared to liquid culture media (six weeks compared to 7 to 21 days), both types of cultures should be used for the detection (Morgan *et al.*, 1983). This is because the solid media will allow examination of colony morphology and identification of mixed cultures, whereas the broth media enable more rapid diagnosis and drug susceptibility testing. For rapid identification of mycobacteria, molecular approaches have been applied and numerous nucleic acid based assays have been developed. These include transcription-mediated amplification, one based on strand displacement amplification, oligonucleotide ligation amplification and polymerase chain reaction which enhance sensitivities and specificities (Madkour, 2004).

1.5 Treatment of TB

There was no effective treatment available against TB in human until the 1940s and the standard treatment was only rest and fresh air (Leibert and William, 2004). Therefore, the discovery of streptomycin in 1944 had evoked a golden age for the development and discovery of anti-tuberculous drug (Duncan, 2003). Streptomycin is purified from *Streptomyces griseus* which act by damaging bacterial cell membrane and inhibiting protein synthesis. A rapid succession of new anti-TB drug appeared in the following years when scientist discovered para-aminosalicyclic acid (PAS). PAS act by exerting a bacteriostatic effect on *M. tuberculosis*, which blocks the conversion of para-aminobenzoic acid into folic acid (Di Perri and Bonora, 2004). The discovery of PAS was important because with streptomycin monotheraphy, resistant mutants began to appear within a few months.

In 1951, the most potent which became the most widely used anti-TB drug, isoniazid was introduced (Leibert and William, 2004). Following isoniazid, pyrazinamide (1954), cycloserine (1955), ethambutol (1962) and rifampin (rifampicin; 1963) were introduced as anti-TB agents. Aminoglycosides such as capreomycin, viomycin, kanamycin and amikacin and then quinolones (e.g. ofloxacin and ciprofloxacin) are only used in the drug resistant situations. Each of the main anti-tuberculous drugs varies in its capacity to kill mycobacteria, to sterilize lesions and the ability to prevent the development of drug resistance (Kaufmann and Hahn, 2003). Isoniazid will act on several targets in the mycobacterial cell, mainly in the mycolic synthesis pathway; rifampin has the ability to rapidly diffuse across the hydrophobic cell envelope and disrupt protein synthesis; pyrazinamide and ethambutol are thought to act on the mycobacterial cell wall biosynthesis pathway and streptomycin acts by preventing mRNA translation.

The discovery and widespread application of the revolutionized TB treatment, contributes to the use of highly effective regimens which only require six months. Multi-drug therapy of tuberculosis was found to be the cornerstone of an effective treatment regimen for TB and also prevented the development of multi-drug resistance. However, there has been no major breakthrough in the drug development which had occurred over the last 40 years, which contributed to the emergence of multi-drug resistance (TB Alliance, 2002). Today, it is estimated that less than 50% of TB cases are diagnosed and less than 60% of these cases are cured.

1.5.1 Emergence of Multi-Drug Resistance in TB

Drug resistance is defined as a temporary or permanent capacity of an organism to remain viable or multiply in the presence of the concentration that would normally destroy or inhibit its growth. Multi-drug resistance in MTB (MDR-TB) can be defined as resistance to isoniazid and rifampicin, with or without resistance to other anti-tubercular drugs (Kaufmann and Hahn, 2003). It is the most severe form of bacterial resistance and usually occurs in chronic TB cases. The problem is man-made because wild isolates of MTB that have never been exposed to anti-TB drug virtually clinically sensitive to the normal regimens (WHO, 2004). However, there are a few isolates of MTB specifically from Madras that showed higher intrinsic levels of resistance to PAS, but does not contribute to the overall burden of resistance (WHO, 2004).

The incidence of MDR-TB varies greatly from region to region all over the world with 300 000 new cases per year according to WHO records and currently affecting 50 million patients (McKenna, 2004). The prevalence of MDR-TB was exceptionally high in almost all former Soviet Union countries surveyed, with a high proportion of isolates resistant to three or four drugs (WHO MDRTB Report, 2003).

According to a WHO MDRTB report, in 2003 there are 90 080 cases with 77 175 new cases and 12 905 previously treated cases developing resistance to the four main antituberculous drugs. Thus, the World Health Organization (WHO) and the International Union against Tuberculosis and Lung Disease (IUATLD) started the Global Project on Anti-tuberculosis Drug Resistance Surveillance in 1994, which further on found that drug resistance was ubiquitous in the world (Kaufmann and Hahn, 2003).

Prevalence of drug resistance against MTB is classified into three classes which are primary, acquired and combined. Primary resistance is the resistance among new cases and is defined as the presence of isolates of MTB in patients who denies having any prior anti-TB treatment (within at least one month). Acquired mycobacterial drug resistance is usually among patients who had had previous treatment or confessed to have been treated for TB for one month or more (WHO MDRTB Report, 2003). It refers as resistance to one or more anti-drugs, which arises during the course of treatment due to non-adherence to the recommended regimen or to incorrect prescribing. This type of resistance was firstly noted in the early introduction of streptomycin in the management of TB (Harkin and Condos, 2004). According to a WHO report in 1997, the rates are quite high in some countries with 36% higher than the average rate. Combined prevalence of drug resistance is defined as the prevalence of resistance in the surveyed population regardless of prior treatment (WHO MDRTB Report, 2003).

Since the early 1990s, several outbreaks of MDR-TB have been reported as a consequence of inconsistent and inappropriate use of anti-TB drugs. These included human error in prescription of chemotherapy treatment, management of drug supply, case management and in the delivery process of drug to patients. The mutation rates of mycobacteria are also high because genetic mutation can occur as

often as once in every million divisions. Each spontaneous genetic mutation might contribute to the resistance of an existing drug (WHO MDRTB Report, 2003). Genetic studies have showed the consequence of spontaneous mutation in genes that encode either the target of the drug, or enzymes that are involved in drug activation (Somoskovi *et al.*, 2001). In fact, tubercles can contain up to one hundred billion mycobacteria, and this increases the chance of mycobacteria developing resistance to any individual drug (Ait-Khaled and Enarson, 2003). MDR-TB resistance develops by sequential acquisition of mutation at different loci, usually due to inappropriate patient treatment (Somoskovi *et al.*, 2001).

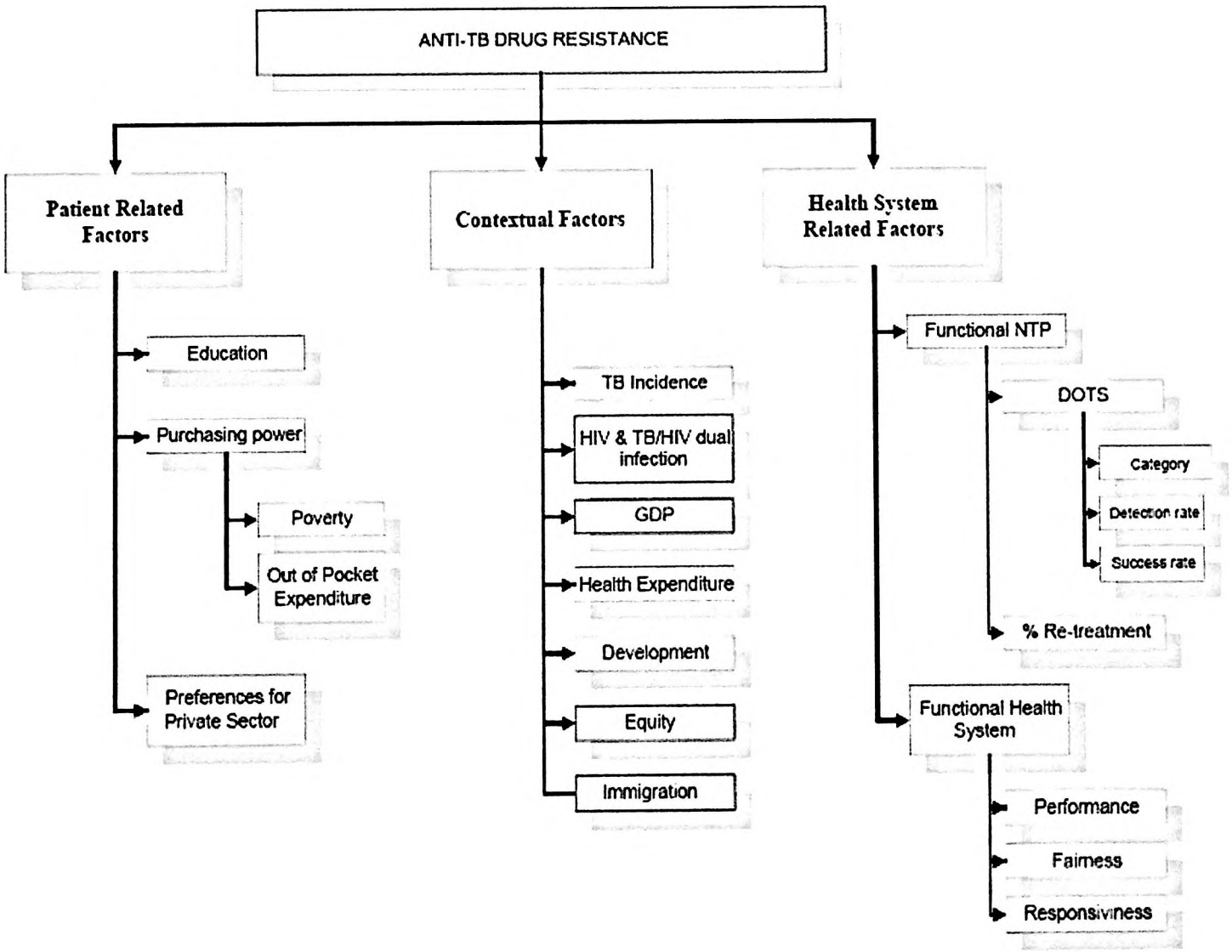

Figure 1.3: Conceptual model of ecological factors affecting TB Drug resistance in the populations.

(Chart was adapted and modified from WHO MDRTB Report, 2003).

1.5.2 Drug Development in TB

The Global Alliance for TB Drug Development has urged to speed development of new drugs in the pipeline with a broader spectrum to treat tuberculosis due to the increasing pattern of multi-drug resistance. The progress in TB drug development has been impeded by the belief that there was little need for new agents, the high cost of drug development and that global market was insufficient to enable a return on investment (TB Alliance, 2000). This is because it is estimated that the development of new drugs takes about ten years with costs of up to US$900 million (Wagner, 2003). This perception had led to the situation that most TB drugs are already decades old. Moreover, no new anti-TB drug has been introduced for over 40 years and worsening the phenomenon is that curing of MDR-TB would take one to two years. Furthermore, many of the current anti-mycobacterial drugs are prodrugs, which require *in-vivo* activation before they can exert their killing (BarryIII *et al.*, 2000, Somoskovi *et al.*, 2001). The requirement for activation may provide greater opportunity for the organism to develop resistance. In addition to serious side effects found in many patients, the lengthy treatment period often leads to the occurrence of non-compliance, which fuels the emergence of MDR strains. In fact, the only available TB vaccine, *M. bovis* BCG, has also been clinically proven not totally effective and consistent in certain places (Sanjay, 1999). This has driven scientists to develop new vaccines and anti-TB drugs.

In addition to the lack of an economic drive for the development of new TB drugs, the development is also impeded by the slow growth and highly infectious nature of this organism. This phenomenon resulted in the screening of compounds for anti-mycobacterial activity by conventional microbiological methods to be extremely slow and inefficient. To develop a new pipeline for TB drug, a safe and reliable assay must be developed for high or at least higher throughput screening.

1.6 Drug Susceptibility Testing for Anti-TB Properties

Drug susceptibility testing is performed to determine a minimal concentration of drugs that are sufficient to inhibit or kill mycobacteria. It can be defined as a process for determining the concentration or biological activity of a substance by measuring the effect on the organism compared to the standard. The new guidelines by the National Committee of Clinical Laboratory Standards (NCCLS), currently recommend repeating susceptibility tests at least every three month (Pfyffer, 2003). This is because drug susceptibility testing is important for monitoring the possible development of drug resistance in patients. Drug susceptibility test is divided into direct and indirect methods. A direct method is defined as a method in which specimens are inoculated directly into the drug containing medium before the susceptibility test is performed (Mitchison, 2005). An indirect method is performed on bacteria that have first been successfully isolated from the specimen. This will be further discussed in the next section.

One of the key techniques required in the screening of anti-TB drug compounds is the sensitivity assay. The normal method for sensitivity testing on agar or egg based slants are obviously not suitable for this work. Other methods for sensitivity testing are available such as the Epsilometer (E-) test (Dhandayuthapani *et al.*, 1995), and Mycobacterium Growth Indicator Tube (Bardarov *et al.*, 2003), but these do not lend themselves to high throughput screening and are relatively expensive. Newer systems such as BACTEC 460 and BACTEC 9000MB system are much more expensive (Collins *et al.*, 1998). These systems are relatively three times more expensive compared to the microplate Alamar blue assay (which cost $0.03 per well) and become unsuitable for large scale screening (Collins and Franzblau, 1997). In order to overcome such problem, the use of fluorescent recombinant mycobacteria expressing the firefly luciferase gene (Arain *et al.*, 1996, Bardarov *et al.*, 2003) and β-

galactosidase (Srivastava R. *et al.*, 1997), and Green Fluorescent Protein (Srivastava *et al.*, 1998, Changsen *et al.*, 2003) have been developed for high throughput-screening formats for anti-tuberculosis drug.

1.6.1 Conventional Screening Methods for Anti-TB

The conventional susceptibility testing of antimycobacterial properties is based on the growth of bacteria on a solid medium containing a fixed amount of drug. It relies on indirect methods consisting of resistance ratio, ratio proportion and absolute concentration methods. Absolute concentration method is done by inoculating a standardized suspension of mycobacteria standardized suspension onto medium containing a known concentration of antimicrobial agents (Hanna, 2004). The result can be obtained within two weeks for preliminary study and four weeks for a complete result. Resistance ratio method is similar to absolute concentration method. However, the resistance is expressed as the ratio of minimum inhibitory concentration (MIC) of the test strains divided by the MIC of the control organism used in the test (WHO, 2003a). Resistance is expressed as the percentage of colonies, which grow on critical concentrations of the test substances (0.2 mg/l for isoniazid, 2 mg/l for ethambutol, 40 mg/l for rifampicin and 4 mg/l for dihydrostreptomycin sulfate) (WHO, 2003a).

The proportion method is basically based on macroscopic comparison of the number of colonies on egg or agar-based drug-containing media. This method requires at least three weeks of incubation period due to the slow growth of mycobacteria colonies. However, this has remained as the gold standard method for a long time. There are several problems inherent in these assays, including lengthy incubation period, contamination problems, drug instability, slow growth-rate of mycobacteria, the highly pathogenic nature of the organism and the laborious

procedure (Parish and Stoker, 1998a). Therefore, there is an urgent need to develop a more rapid susceptibility testing method.

1.6.2 High Throughput Screening Assay for Anti-TB

High throughput screening is defined as a process for rapid assessment of activity in samples, usually derived from a combinatorial library or other compound collection. It allows drug screening test in a short period of time through a combination of modern robotic, data processing and control software, liquid handling devices, and sensitive detectors (http://en.wikipedia.org/wiki/High-throughput_screening). HTS system can be achieved by running parallels assays in 96 wells or more wells using automated machines. In the pharmaceutical industry, drug screenings are done up to 10,000 compounds per day and ultra-high-throughput assays (uHTS) are capable of screening up to 100,000 compounds per day (Barry III, 2003). Advances in miniaturization and automation have made the possibility of automated screening for anti-TB compounds. However, antitubercular screening have not been at the forefront of HTS or uHTS assays due to reasons already discussed above. Furthermore HTS or uHTS have tended to focus on specific targets usually produced by recombinant methods rather than whole cells. Thus the ability of compounds to cross bacterial cell walls (particularly important in mycobacteria) to reach targets within the cell will need to be tested later in whole cell assays which increases costs and processing time.

A high throughput rapid radiometric culture method using whole cells commercially known as BACTEC 460 TB system represent significant savings in time, with result generally being available within seven days (Arain *et al.*, 1996). However, the test uses vials containing a radioactive broth medium, whose unit price is substantial, and is associated with high costs for disposal of radioactive liquids. In

order to avoid radionuclide waste, the manufacturer has developed an alternative model that uses a novel fluorescent marker. The automated non-radiometric culture system, known as Mycobacterial Growth Indicator Tube (MGIT) only requires a median of 6.4 to 9.5 days in an automated format (Bardarov *et al.*, 2003) for clinical testing of drug resistance. However, it was not approved for this indication and the technology is still too expensive (Vanitha and Paramasivan, 2004) and thus unsuitable for processing large numbers of samples on a high-throughput basis for screening for anti-TB compounds.

A colorimetric drug-susceptibility testing method, known as the microplate Alamar blue assay (MABA) has been developed using a redox indicator without the needed for expensive instruments (Collins and Franzblau, 1997). However, this method requires a greater level of technical expertise to provide reliable result (Bastian *et al.*, 2001). In order to overcome this problem, the reporter gene assay techniques have been used as a high throughput screening assay.

1.6.3 Reporter Gene Screening Assay for Anti-TB

Reporter gene assay is defined as an analytical method, which allows detection and quantification of particular constituent in a mixture by tagging the constituent with a reporter gene. Reporter gene assays can also be used for the identification of promoters and enhancers and to study the correlations between their activities that are expressed from specific reporter genes. Recently, fluorescent recombinant mycobacteria expressing the firefly luciferase and β-galactosidase reporter genes have been reported for use in HTS formats. However, one drawback of the luciferase system is that the substrate, luciferin, has to be added at the required time points to induce fluorescent (Cooksey *et al.*, 1993). Whereas, β-galactosidase system need O-nitrophenyl-β-D-galactosidase (ONPG) as a substrate

(Srivastava *et al.*, 1997). Thus, this detection method required the disruption of cells, which therefore do not allow a single culture to be followed through from the beginning until the end. Cells are no longer viable to allow repeated measurements of the cultures in the same well. Repeated measurement will allow a dynamic picture of the bacterial response to various concentrations of the test material such as antimicrobial agents (Collins *et al.*, 1998).

These constraints obviously add to the complexity of the assay, which therefore require some kind of machinery to automatically add the substrate if it is to be used in HTS. In order to overcome this problem, other methods that allow the cells to remain alive would be far more useful. Furthermore if the bacteria remain alive during the assay, MBCs of the test can also be determined from the same wells by determining the numbers of CFU on solid media.

1.7 Green Fluorescent Protein

The intrinsically fluorescing Green Fluorescent Protein (GFP) (Chalfie *et al.*, 1994) has been shown to be useful in the development of recombinant Mycobacteria for screening assays (Arain *et al.*, 1996). In addition to using tubes, assays based on fluorescing recombinant mycobacteria can be performed in the microplate format for HTS. The intrinsically fluorescing nature of GFP allows the cell to grow whilst multiple measurements can be made and thus allowing for real time kinetic monitoring.

GFP was first discovered in 1961 by Osamu Shimomura *et al.* (Tsein, 1998). The protein is produced by a jelly fish, *Aequorea victoria*, which was found in the cold Pacific Northwest. GFP produces an intense and stable non-catalytically green fluorescence by absorbing blue light maximally at 395 nm and emitting green light with a peak at 509 nm (Chalfie *et al.*, 1994). The native GFP protein consists of 238

amino acids in length (27kDa). GFP folds into a unique, compact structure known as β-can that is strongly resistant to chemical denaturation and found to be stable in most cells (Corish and Tyler-Smith, 1999). The β-can structure encloses an α-helix containing the chromophore. This structure is required to provide a proper environment for the chromophore to fluoresce as shown by the fact that nascent GFP do not fluoresce (Clontech., 2001). GFP fluorescence occurs without any cofactors and this property allows GFP fluorescence to be utilized as a reporter in non-native organism. The fluorescence is derived from posttranslational cyclization of a serine-tyrosine-glycine tripeptide of the GFP protein, followed by dehydrogenation of the tyrosine. It is generated by sequential activation of two photo proteins, aequorin and GFP. Upon calcium binding, aequorin emits blue light which in turn excites GFP to fluoresce. GFP produces green fluorescence which requires only the presence of oxygen to maturate with no external compounds needed (Andersen *et al.*, 1998).

GFP has become popular reporter system for use in both prokaryotes and eukaryotes. In eukaryotes, it has been used in numerous applications, including transcriptional protein fusions to study protein targeting (Cowley and Av-Gay, 2001). In prokaryotes, it has been used primarily as a reporter for promoter activity by creating transcriptional fusions in a wide range of bacterial species, including *Escherichia coli*, *Brucella suis, Salmonella typhimurium*, and *Mycobacterium spp.* (Dhandayuthapani *et al.*, 1995, Valdivia *et al.*, 1996). GFP generally does not interfere with the growth of the host and is thus an excellent choice for non-disruptive studies of bacterial communities or other systems, which require live cells to be studied at the single cell level. In recent years, GFP has been established as a convenient reporter of cell viability due to its broad host range, cytoplasmic location and low toxicity allowing continuous production during replication, easy imaging and quantification (Collins *et al.*, 1998).