Chapter 2: Review of Literature
2.1 Tuberculosis and Its Current Epidemic

In spite of more than 100 years of research, tuberculosis continues to be a serious threat to the global health. Tuberculosis (TB) accounts for more than one-fourth of all preventable adult deaths in the world (Global Tuberculosis Control, World Health Organization report, http://www.who.int/gtb/publications/glovrep01/, 2001). Tuberculosis is caused by the bacterium *M. tuberculosis*. Approximately 32% of world’s human population is infected with this bacterium. More than eight million new cases of TB are diagnosed each year, and almost two billion people are latently infected with the tubercle bacillus.

The symptoms of TB include low-grade fever, night sweats, fatigue, weight loss and a persistent cough. In some cases there might not be any obvious symptoms. Generally, the risk of infections is proportional to the intensity of exposure. Although skin testing reveals that one-third of the world population is positive either presently or previously for bacteria, the infection does not necessarily lead to disease, a stage termed the latent infection. The latent tuberculosis infection probably represents a balance between pathogen virulence factors and the host immune response that leads to long-term containment, but not eradication, of persistent bacilli (Manabe and Bishai, 2000).

In cases where the host immune mechanisms start waning, for example in HIV infected patients, the balance shifts in favour of the pathogen, ultimately resulting in proliferation of bacilli and setting up of active tuberculosis disease.

The registered number of new cases of TB worldwide roughly correlates with economic conditions, maintaining its tightest grip on the populations of poor and under developed countries of Africa, Asia, and Latin America. In industrialized countries, the steady drop in TB incidence began to level off in the mid-1980s and then stagnated or even began to increase. WHO estimates that eight million people get TB every year, of which 95% live in developing countries (Figure 2.1).

Re-emergence of TB as a serious public health threat worldwide has majorly been because of a significant increase in multiple-drug-resistant TB (mrdTB) as well as synergism between Human Immunodeficiency Virus (HIV) and *M. tuberculosis* infection. On an average, out of ten immuno-competent people who are infected with *M. tuberculosis* one will have a chance of falling sick in their lifetimes, but among those with HIV, one in ten per year will develop active TB, while one in two or three tuberculin test positive AIDS patients will develop active TB. In industrialized countries, these cases make up only a small minority of TB cases. However in
developing countries, the impact of HIV infection on the TB situation, especially in the 20-35 age group, is worthy of concern. While wealthy industrialized countries with good public health care systems can be expected to keep TB under control, in much of the developing world a catastrophe awaits. It is crucially important that support be given to research efforts devoted to better understanding of the host-pathogen interactions of the bacilli, improving the diagnosis of TB, creating new, highly effective anti-TB medications and developing an effective TB vaccine, shortening the amount of time required to ascertain drug sensitivities.

2.2 Mycobacterium tuberculosis

2.2.1 General History

It is believed in general that the causative agent of TB, *M. tuberculosis* invaded human life from cattle during the age of the domestication of cattle approximately 10,000-15,000 years ago. The presence of tuberculosis has been documented in the prehistoric remains of humans (4000 BC) as well as in Egyptian mummies (3000-2400 BC). In India, the oldest recorded references of TB cases are dated around 2000 BC. The name tuberculosis is derived from a Greek term *Phthisis* that means consumption. It was Hippocrates who first identified tuberculosis as a fatal disease around 460 BC and named it *phthisis*. In 1839, J.L. Schoenlein gave the disease its present name tuberculosis.

The bacillus-causing tuberculosis, *M. tuberculosis*, was described on March 24, 1982 by Robert Koch, who later in 1905 received the Nobel Prize for medicine or physiology for this discovery.

The first Tuberculosis vaccine BCG (Bacillus of Calmette and Guerin) was developed from attenuated bovine strain tuberculosis by Albert Calmette and Camille Guerin in 1906. In 1946, the antibiotic streptomycin was developed as a treatment drug for TB. However with emergence of multi-drug resistant strains since 1980, when first such case was reported, the hopes that the disease could be completely eliminated have been dashed. The resurgence of tuberculosis resulted in the declaration of a global health emergency by the World Health Organization in 1993.

Genetically, *M. tuberculosis* is a relatively young strain. The original disease may have originated 15,000 years ago. More research is necessary to determine if the
multiple strains are due to cattle-to-human transmission or due to the nomadic lifestyles of humans that brought them into closer contact with multiple races of people.

2.2.2 Taxonomic Classification

Higher order taxa: Bacteria (kingdom); Actinobacteria (class); Actinobacteridae (order); Actinomycetales (sub-order); Corynebacterineae (family); Mycobacteriaceae (genus); Mycobacterium tuberculosis complex; Mycobacterium (species) tuberculosis (sub species).

Characteristics of Tuberculosis complex organisms are:

- Obligate aerobes growing most successfully in tissues with high oxygen content, such as the lungs.
- Facultative intracellular pathogens usually infecting mononuclear phagocytes (e.g. macrophages).
- Slow-growing with a generation time of 12 to 18 hours (c.f. 20-30 minutes for Escherichia coli), a physiological characteristic that may contribute to its virulence.
- Hydrophobic with high lipid content in the cell wall. Because the cells are hydrophobic and tend to clump together, they are impermeable to the usual stains, e.g. Gram’s stain.
- Known as "acid-fast bacilli" because of their lipid-rich cell walls, which are relatively impermeable to various basic dyes unless the dyes are combined with phenol. Once stained, the cells resist decolorization with acidified organic solvents and are therefore called "acid-fast". (Other bacteria, which also contain mycolic acids, such as Nocardia, can also exhibit this feature.)

2.2.3 M. tuberculosis Complex

Members of Mycobacterium tuberculosis complex (MTBC) are the causative agents of tuberculosis in humans and animals. Notwithstanding close relationship at genetic level, they differ in epidemiology, pathogenicity, geographic range, host preference, and in importance vis-a-vis the tuberculosis disease in humans.
A flowchart listing all the members of MTBC is shown in figure 2.2. The last member of this list has still not been officially recognized on the list of bacterial names (http://www.bacterio.cict.fr). A novel member, *M. pinnipedii*, proposed by Cousins *et. al.* (2003) has been well-characterized as a species of *M. tuberculosis* complex. Genetically, all members of this complex are extremely similar in having 99.9 % similarity at the nucleotide level and identical 16S rRNA sequences (Chimara *et. al.*, 2004).

*M. bovis* is the etiological agent of TB in cows and, infrequently, in humans. Both cows and humans can serve as reservoirs. Humans can also be infected by *M. bovis* through the consumption of non-pasteurized milk. This route of transmission can lead to the development of extra-pulmonary TB, exemplified in history by bone infections that led to hunched backs.

Other human pathogens belonging to the *Mycobacterium* genus include *M. avium* which causes a TB-like disease especially prevalent in AIDS patients, and *M. leprae*, the causative agent of leprosy. *M. tuberculosis*, the causative agent of tuberculosis, *M. bovis* and *M. leprae* are obligate pathogens. In addition, several other mycobacterial species, particularly members of the *M. avium-intracellulare* complex, are opportunistic pathogens, *i.e.* they occur naturally in the environment but may cause serious disease in humans occasionally, especially in immunosuppressed individuals.

*M. tuberculosis* is the etiologic agent of tuberculosis in humans. Humans are the only reservoir for the bacterium. The following three tests have conventionally been used to distinguish *M. tuberculosis* from other mycobacteria.

- **The Niacin Test:** *M. tuberculosis* produces large amounts of niacin. However, drug resistant strains do not confirm to this test.
- **The Catalase Test:** *M. tuberculosis* has a heat sensitive catalase. However, isoniazid (INH) resistant *M. tuberculosis* lack catalase and thus doest not confirm to this test.
- **The Nitrate Test:** *M. tuberculosis* can reduce nitrate. *M. bovis*, an animal pathogen which causes a disease indistinguishable from TB, can also reduce nitrate.
2.2.4 Drug Resistant Strains
According to the World Health Organization, in some countries, the strains of \textit{M. tuberculosis} that are resistant to one or several chemotherapeutic agents account for up to 41\% of total \textit{M. tuberculosis} clinical isolates (Portevin \textit{et. al.}, 2004).

In general, drug resistance in TB occurs as a result of mutations in tubercle bacillus and is not transmitted by plasmids. Once exposed to a single effective anti-TB medication, the predominant bacilli sensitive to that drug are killed; the few drug resistant mutants, likely to be present if the bacterial population is large, will multiply freely. Since it is very unlikely that a single bacillus will spontaneously mutate to resistance to more than one drug, giving multiple effective drugs simultaneously will inhibit the multiplication of these resistant mutants. It was first observed in the 1950s, that the strains defective in production of catalase were overproducing peroxidase and were found to be INH resistant. Since then, it has been shown that a mutation or deletion in the catalase gene is responsible for 10-25\% of INH resistance. The rest of the 75-90\% of INH resistance is due to mutations in an enoyl-acyl carrier protein reductase involved in mycolic acid synthesis (Heym \textit{et. al.}, 1994). The present understanding is that INH is activated by catalase and then binds to the reductase, inhibiting cell wall synthesis. If reductase production is up regulated, the enzyme could overcome the effect of INH. On the other hand, a mutation can also prevent INH binding.

Although rare, resistance to rifampicin (RIF) is also increasing because of widespread application and results in selection of mutants resistant to other components of short-course chemotherapy. RIF had long been believed to target the mycobacterial RNA polymerase and thereby kill the organism by interfering in the transcription process (Ovchinnikov \textit{et. al.}, 1981; Levin and Hatfull, 1993).

Similarly, in the case of Ethambutol (EMB), which is a first-line anti-MTB drug with a broad spectrum of activity the genetic basis for resistance to it were largely obscure. Specificity of EMB for mycobacterial species, however, indicated that its target may have been involved in the construction of the outer cell wall. Indeed, earlier studies of Takayama and colleagues demonstrated that administration of EMB led to rapid cessation of mycolic acid transfer to the cell wall and equally rapid accumulation of trehalose mono- and di-mycolates (Takayama \textit{et. al.}, 1979; Belanger \textit{et. al.}, 1996).
Resistance to pyrazinamide (PZA) a structural analog of nicotinamide has been attributed to a single point mutation that results in the substitution of His to Asp at position 57 of the \textit{pncA} gene from \textit{M. bovis} isolates (Scorpio and Zhang, 1996).

\subsection*{2.2.5 Transmission and Proliferation of \textit{M. tuberculosis}}

The primary route of transmission of \textit{M. tuberculosis} is through aerosol but it can also transmit itself through the gastrointestinal tract. The cough of people with active TB contains droplet nuclei carrying infectious organisms which can remain suspended in the air for several hours. The infection takes place in the event of these droplets reaching the alveoli of the lungs of the healthy individual by inhalation. Only 10\% of immunocompetent people infected with \textit{M. tuberculosis} develop active disease in their lifetime whereas the other 90\% do not become ill and cannot transmit the organism. Nonetheless, in some groups such as infants or the immunodeficient individuals (e.g. those with AIDS or malnutrition), the proportion who develop clinical TB is much higher (Official Statement of the American Thoracic Society and the Centers for Disease Control and Prevention, 2000).

Once the bacillus has reached the lungs, the organism is taken up by alveolar macrophages and carried to lymph nodes, from where it may spread to multiple organs. The cell mediated immunity (CMI) and hypersensitivity (DTH) develop in two to eight weeks after infection, that then leads to the characteristic reaction to the tuberculin test and, in immunocompetent individuals, containment of infection. Inflammatory immune responses eventually result in lung damage. TB treatment requires a 6 to 12 month regimen with at least two antibiotics. Failure to complete the full course of drug therapy can lead to \textit{M. tuberculosis} organisms that are resistant to one or more anti-tuberculosis drugs, severely limiting effective treatment options. This makes it absolutely essential to treat TB patients with the recommended four drug regimen of isoniazid, rifampin, pyrazinamide and ethambutol or streptomycin.

\subsection*{2.2.6 Morphology of \textit{M. tuberculosis}}

The bacteria comes in the shape of 1-4 $\mu$m long, thin rods, usually straight or slightly curved that frequently show irregular beading due to vacuoles and polyphosphate granules. It often appears singly and occasionally in threads. The bacterium is non-motile, non-spore forming, and non-capsule forming. Mycobacteria contain mycolic acids and complex long-chain fatty acids that are found otherwise only in \textit{Nocardia}.
and *Corynebacterium*. Owing to the presence of N-glycolylmuramic acid in the place of N-acetyl muramic acid in the peptidoglycan cell wall, *M. tuberculosis* is an acid-fast bacterium, which means it has the ability to retain the original stain even after washing. *M. tuberculosis* has a thick, waxy coat that makes the stain irregular.

### 2.2.7 Genome of *M. tuberculosis*

The elucidation of complete genome sequence of H37Rv, the most widely used strain of *M. tuberculosis* has broadened the horizons for the research and development of new prophylactic and therapeutic interventions against tuberculosis. *M. tuberculosis* genome is 4,411,529 base pairs long with 3,924 predicted ORFs, and a relatively high G+C content of 65.6 % (Cole *et. al.*, 1998). Genomic informations related to *M. tuberculosis* are centralized in TubercuList (http://www.pasteur.fr/Bio/TubercuList/). At 4.41 Mb, *M. tuberculosis* is one of the largest known bacterial genomes, coming in just short of *E. coli* (3.6 Mb), and a distant third to *Streptomyces coelicolor* (8 Mb). As deduced from its genome, *M. tuberculosis* has the potential to manufacture all of the machinery necessary to synthesize all of its essential vitamins, amino acids, and enzyme co-factors. Out of all the protein coding genes of *M. tuberculosis* ~40 % are considered as conserved hypotheticals whereas remaining genes have been classified into 11 broad groups based on the database comparisons of their possible functions. A key feature of mycobacterial genome is that a significant portion of the *M. tuberculosis* genome amounting to 9.3 % of total coding capacity (225 genes) *i.e.* approximately 372 kb in length is devoted to genes involved in lipid metabolism (Cole *et. al.*, 1998; Cole *et. al.*, 2001).

### 2.2.8 Cell Envelope of *M. tuberculosis*

The mycobacterial cell envelope has been found to play a fundamental role in the physiology of these bacteria, both in terms of pathogenicity and resistance to antibiotics. Infact, the sluggish growth of *M. tuberculosis* is also accredited to the presence of the tough cellwall that resists the passage of nutrients into the cell and inhibits waste products to be excreted out of the cell (Lederer *et. al.*, 1975).

The cell envelope of *M. tuberculosis* is unique both in terms of its molecular composition and the architectural arrangement of its constituents (Adam *et. al.*, 1969). A key feature of the mycobacterial envelope is its high lipid content and the variety of lipid compounds with unusual structures (Lederer *et. al.*, 1975).
A simplified diagram presenting the current understanding of cell-wall architecture is shown in figure 2.3. The specialized cell envelope of *M. tuberculosis* resembles a modified Gram positive cell wall. It is composed of two segments, upper and lower. There is an additional peptidoglycan (PG) layer present beyond the cell membrane in *M. tuberculosis* that is covalently attached to arabinogalactan (AG), which in turn is attached to the mycolic acids. This is termed the cell wall core—the mycolyl arabinogalactan–peptidoglycan (mAGP) complex. The upper segment is composed of free lipids and is embellished with a variety of cell-wall proteins, the phosphatidylinositol mannosides (PIMs), the phthiocerol-containing lipids, lipomannan (LM), and lipoarabinomannan (LAM). These free lipids, proteins, LAM, and PIMs are solubilized, and the mycolic acid–arabinogalactan–peptidoglycan complex remains as the insoluble residue during the process of cell wall extraction. It can be considered that these lipids, proteins, and lipoglycans are the signaling effector molecules in the disease process, whereas the insoluble core is essential for the viability of the cell and therefore should be addressed in the context of new drug development (Brennan, 2003). The cell envelope of *M. tuberculosis* also includes adhesins but does not contain any known toxins (Vergne and Daffe, 1998).

The high concentration of lipids in the cell wall of *M. tuberculosis* has been associated with following properties of the bacterium:

- Impermeability to stains and dyes
- Resistance to many antibiotics
- Resistance to killing by acidic and alkaline compounds
- Resistance to osmotic lysis via complement deposition
- Resistance to lethal oxidations and survival inside of macrophages

### 2.2.9 Virulence Mechanisms and Virulence Factors

Traditionally, bacterial virulence has been attributed to factors such as toxins, capsules and fimbriae etc. However, virulence of *M. tuberculosis* has been accredited to a large number of structural and physiological properties of the bacterium. Some of them are listed below:

1) Mechanisms for cell entry of *M. tuberculosis* whereby the tubercle bacillus can bind directly to mannose receptors on macrophages via the cell wall-
associated mannosylated glycolipid, LAM, or indirectly via certain complement receptors or Fc receptors.

2) *M. tuberculosis* can grow intracellularly which is an effective means of evading the immune system. In particular, antibodies and complement are ineffective. Once *M. tuberculosis* is phagocytosed, it can inhibit phagosome-lysosome fusion. The bacterium may remain in the phagosome or escape from the phagosome; in either case, finding a protected environment for growth in the macrophage.

3) *M. tuberculosis* interferes with the toxic effects of reactive oxygen intermediates produced in the process of phagocytosis by any of two mechanisms: either compounds including glycolipids, sulfatides and LAM down-regulate the oxidative cytotoxic mechanism, or Macrophage uptake via complement receptors may bypass the activation of a respiratory burst.

4) Antigen 85 complex: This complex is composed of a group of proteins secreted by *M. tuberculosis* that are known to bind fibronectin. These proteins may aid in walling off the bacteria from the immune system and may facilitate tubercle formation, although evidence of this is as yet lacking.

5) Slow generation time: Because of which immune system may not readily recognize the bacteria or may not be triggered sufficiently to eliminate them. Many other chronic disease are caused by bacteria with slow generation times, for example, slow-growing *M. leprae* causes leprosy, *Treponema pallidum* causes syphilis, and *Borrelia burgdorferi* causes Lyme disease.

6) Cord factor: The cord factor is primarily associated with virulent strains of *M. tuberculosis*. Although its exact role is unclear, it is known to be toxic to mammalian cells.

7) High lipid concentration in cell wall, as mentioned previously, accounts for impermeability and resistance to antimicrobial agents, resistance to killing by acidic and alkaline compounds in both the intracellular and extracellular environment, and resistance to osmotic lysis via complement deposition and attack by lysozyme.

2.2.10 Lipids of *M. tuberculosis*

The genus *Mycobacterium* is known for its high lipid content and elaborated cell envelop with great variety of glycolipids of rather exotic structure, constituting up to
40% of dry weight. Because pathogenic mycobacterial species are facultative intracellular parasites that infect and reside in host cells, some of them may represent potential virulent factors as they have been shown to inhibit both macrophage antimicrobial activities and lymphoproliferation. These activities may be derived, at least in part, from the modulation of cellular functions through the interactions between host membranes and surface-exposed lipids whose structures are different from those of mammalian cell membrane components.

The constitutive lipids are of two types:

a) **Covalently linked lipids**: This first class of lipid is composed of high molecular weight (up to $C_{90}$) alpha-branched, beta-hydroxylated fatty acids called mycolic acids that esterify the cell wall arabinogalactan.

b) **Loosely-associated/free lipid molecules**: The lipids of this class are a mixture of ubiquitous, type-specific and species-specific lipids that are readily extractable with organic solvents. Some of these lipids are abundant cell envelope components and are exposed on the bacterial surface. These comprise the species-specific phenolic glycolipids, glycopeptidolipids, sulfatides, and lipooligosaccharides, and the ubiquitous phosphatidylinositolmannosides. In few cases purified glycolipids have been shown to profoundly affect the physical and functional properties of biologic membranes. Through specific or non-specific interactions with host membranes mycobacterial glycolipids may contribute to the pathogenicity by modulating the activity of the infected cells. Therefore, the enzymes involved in the biosynthesis of the biologically active glycolipids represent potential drug targets.

### A. 2.3 Polyketide Synthases

#### 2.3.1 Introduction

Polyketides are a large family of structurally varied natural products that have a broad range of pharmacological properties and, together with their semi-synthetic derivatives, command a vital role in human and veterinary medicine. A large number of polyketides have been clinically approved as drugs for treating disorders such as bacterial and viral infections, cancer, cardiovascular diseases, and inflammation.
These therapeutically important agents, as the name implies, are polycarbonyl compounds synthesized from repetitive Claisen condensation reactions that link small 2- to 3-carbon acyl groups derived from coenzyme A (CoA) thioesters precursors (Figure 2.4). The process is similar in many respects to bacterial and mammalian fatty acid synthesis (Pfeifer and Khosla, 2001).

Polyketides derive their enormous diversity in structure through a number of programmed events that are dictated by the polyketide synthase (PKS) and involve the selection of starter and extender units, carbon chain length, folding, degree of reduction, and termination. Post-PKS tailoring events such as glycosylation, acylation, alkylation and oxidation further add to polyketide structural diversity. PKSs utilize a wide assortment of starter units, such as short-chain (branched) fatty acids, various acyclic and aromatic acids and amino acids, in the assembly of their products. In many cases, the nature of the primer unit provides important structural and biological features to the molecule.

Approximately two-thirds of the known bioactive polyketide natural products originate from the actinomycetes. Other major microbial sources of polyketides include the myxobacteria and filamentous fungi.

### 2.3.2 Classification

PKSs have been classified into three distinct families on the basis of protein architecture (Pfeifer and Khosla, 2001).

**A. Type I that mimics type I (vertebrate) FAS**

This subclass termed as type I PKSs closely resembles vertebrate (type I) FAS. Depending upon mode of action it can be further classified into two types:

**A.1. Iterative type I PKS:** The fungal PKS represents this class of PKSs. The proteins are multidomain _i.e_ composed of a small number of separate proteins, active sites of which are used in an iterative fashion for assembly and functional group manipulation of side chains. An example is the lovastatin LNKS, that catalyzes the formation of dihydromonacolin, the polyketide precursor for lovastatin biosynthesis (Hendrickson _et al._., 1999).

**A.2. Modular type I PKS:** This PKS, known as modular PKSs, is exemplified by the deoxyerythronolide B synthase (Hutchinson, 1999). Here, discrete multifunctional enzymes control the sequential addition of thioester units
and their subsequent modification to produce macrocyclic compounds. In effect, polyketide biosynthesis proceeds in a processive fashion on a very large megasynthase in which each active site is used once during the overall catalytic cycle. This class of PKS is equally abundant both in gram positive as well as gram negative bacteria.

B. Type II that mimics type II (bacterial and plant) FAS

Bacterial PKS systems with an architectural relationship to type II fattyacid synthases represent type II PKS category. Also acting in an iterative fashion, these multienzyme systems produce aromatic polyketide products. However, individual active sites occur as distinct polypeptides rather than as domains within a single multifunctional polypeptide. An example is the doxorubicin PKS, which synthesizes the tetracyclic skeleton of this anthracycline antibiotic (Kantola et. al., 2000).

Type I and type II PKSs are multifunctional proteins that have primarily been reported from the microbial world. Both these classes of PKSs require a 4'-phosphopantetheine arm to which the growing polyketide chain is acylated during the biosynthesis process. In contrast, the type III category of PKSs consists of a modest size of (40-47 kDa), homodimeric proteins that use CoA thioesters as substrates and synthesize polyketides by using a single catalytic center (Sankaranarayanan et. al., 2004).

C. Type III that mimics plant chalcone synthase (CHS)

The type III PKSs are related to the plant enzyme superfamily that includes chalcone and stilbene synthase (Hopwood, 1997). It is perhaps the simplest type of PKS that have been exclusively studied from the plant kingdom for several decades. The recent discovery of these plantlike PKS proteins in bacterial systems has provided a new dimension to our understanding of the importance and hence the ubiquitous distribution of these enzymes (Cole et. al., 1998).

2.3.3 FAS vs PKS

PKSs and FASs are structurally and mechanistically related proteins that synthesize products through a common strategy of repetitive decarboxylative condensation of simple acyl-CoA thioesters (Sankaranarayanan et. al., 2004). The assembly line like nature of type I PKSs draws a strong parallel with the mammalian as well as microbial type I FAS. Crucially however, the FAS modules function iteratively while
those belonging to the PKS process the carbon chain only once, before the chain is passed on to the next module. The single use of a module within a PKS means that the final polyketide molecule nearly always contains partially processed chemical functionalities and chiral carbon centres. Also, unlike fatty acids, the structures of polyketides are far more diverse due to variations in the fatty acid synthesis theme and post-PKS modifications (Pfeifer and Khosla, 2001). This structural diversity is also reflected in diversity in their biological modes of action.

2.3.4 PKSs of *M. tuberculosis*

The intense interest in studying the mycobacterial cell envelope is now further boosted by the presence of examples of every known lipid and polyketide biosynthetic system encoded in the genome, including enzymes usually associated with mammals and plants (Cole *et. al.*, 1998).

Mycobacteria, which are members of the Actinomycetales, are also primarily saprophytic soil bacteria. The genus includes two of the most dangerous pathogens known to man namely *M. tuberculosis* and the leprosy-causing *M. leprae*, (Bentley *et. al.*, 2002). Mycobacteria synthesize polyketides by several mechanisms. Interestingly, *M. tuberculosis* genome has revealed a remarkable array of PKSs. The *M. tuberculosis* chromosome contains 18 PKS gene clusters (Cole *et. al.*, 1998). In fact, there are many more of these enzymes than there are known metabolites. Thus, there may be new lipid and polyketide metabolites that are expressed only under certain conditions, such as during infection and disease. The fact that six of these clusters were found in the severely reduced *M. leprae* chromosome suggests that they serve indispensable functions in the bacteria. In mycobacteria, most of PKSs have been implicated in the biosynthesis of composite lipids, and several PKS mutants of mycobacteria show attenuation in their virulence properties. However, no polyketide product has been isolated thus far (Cole *et. al.*, 2001).

- Type I PKS of Mycobacteria

The proposed Claisen-type reaction involved in the condensation of two fatty acid chains to form mycolic acid is very similar to the one catalyzed by type I PKS during the formation of methyl branched fatty acids in *M. tuberculosis*. A systematic search for the type I PKS of *M. tuberculosis* (Cole *et. al.*, 1998) revealed the presence of nine
putative PKS enzymes that were common between *M. leprae* and *M. tuberculosis*. Both these species are known to produce various types of mycolic acids. Seven of these nine putative enzymes have been shown to be involved in the biosynthesis of the phthiocerol dimycocerosates and structurally related phenolglycolipids in *M. tuberculosis* (Azad et al., 1996; Constant et al., 2002).

In *M. ulcerans*, Type I PKS system is involved in the production of mycolactone, a toxin involved in pathogenicity. Mycolactone causes cell death via apoptosis as well as immunosuppression (Quadri et al., 1998; George et al., 1999).

**Iterative type I PKS**

The mas gene, a member of PDIM biosynthetic cluster in *M. bovis* and *M. tuberculosis* encodes an iterative type I PKS that produces mycocerosic acids after two to four rounds of extension of C\(_{18-20}\) fatty acids with methylmalonyl-CoA units by a type I FAS-like iterative process (Fernandes and Kolattukudy, 1997). The products of Mas action- the Mycocerosic acids, and the multimethyl-branched fatty acids are then linked as esters to a diol moiety present in the polyketides phthiocerol and phenolphthiocerol that are synthesised by an upstream modular type I PKS in members of the *M. tuberculosis* complex and *M. leprae* (Jenkin et al., 2003).

**Modular type I PKS**

A modular type I system, similar to that involved in erythromycin biosynthesis (Hopwood, 1997), is encoded by a very large operon, ppsABCDE, and functions in the production of phenolphthiocerol in *M. bovis* and *M. tuberculosis* (Kolattukudy et al., 1997). The ppsA-E gene products are responsible for the elaboration of incoming long-chain (C\(_{20-22}\)) fatty acids by successive additions of malonyl or methylmalonylCoA extender units, typical of type I PKSs like the erythromycin PKS (Cortes et al., 1990). The absence of a second type I PKS suggests that the related lipids phthiocerol A and B, phthiodiolone A and phthiotriol may all be synthesized by the same system, either from alternative primers or by differential post-synthetic modification.

Some of these PKSs have also been assumed to extend type I FAS CoA primers to produce other long-chain methyl-branched fatty acids such as mycolipenic, mycolipodienic and mycolipanolic acids or the phthioceranic and
hydroxyphthioceranic acids, or may even show functional overlaps (Kolattukudy et al., 1997). An example is the gene cluster spanning pks10, pks7, pks8 and pks9, which includes two of CHS like enzymes and two modules of an apparent type I system. The unknown metabolites produced by these enzymes are interesting because of the potent biological activities of some polyketides such as the immunosuppressor rapamycin (Hopwood, 1997; Kolattukudy et al., 1997; Cole et al., 1998).

- **Type III PKS of Mycobacteria**

The analysis of mycobacterial genome sequences has revealed three type III pks genes in *M. tuberculosis*. These pks genes have shown 40-45% similarity with both plant and bacterial type III CHS-like sequences. Two of the type III genes, pks10 and pks11 are clustered in an unusual organization with several other type I pks genes that encompass 18 kb of genomic DNA. Pks 18 protein is not flanked by any other pks-related genes. Genome analysis suggests that all the three type III pks genes are present as pseudogenes in *M. leprae*, whereas they were completely conserved in *M. bovis* BCG. *M. bovis* has therefore been used as a model to test functional expression of these genes. These novel type III PKSs from *M. tuberculosis* are involved in the biosynthesis of long-chain α-pyrones (Saxena et al., 2003) Subtractive hybridization revealed a type I Pks locus specific to *M. ulcerans* (Jenkin et al., 2003).

### 2.4 Phthiocerol Dimycocerosate

Intercalated within the lipid environment provided by the mycolic acids of the mAGP complex are the lipids that have intrigued researchers for over five decades: the phthiocerol/phenolphthiocerol dimycocerosate (DIM/PDIM), cord factor/dimycolytrehalose, the sulfolipids (SLs), the PIMs, etc. Knowledge of their roles in "signaling" events, in pathogenesis, and in the immune response is now emerging. Two structurally related families of lipids, phthiocerol diesters and phenolic glycolipids, have retained a special attention. These mycobacterial cell envelope constituents are produced by a small number of slow growing species, in particular the pathogenic species *M. leprae, M. ulcerans, M. marinum* and members of the *M. tuberculosis* complex.
2.4.1 PDIM
The tubercle bacillus is referred as a "ball of wax", mainly because of high content of a lipid called PDIM/DIM in the cell wall of the bacterium. PDIM/DIM was discovered by Rudolf Anderson at Yale in the 1940s. It is highly apolar. Phthiocerols diesters are composed of a mixture of long chain diols that are esterified by multimethyl-branched fatty acids. Depending on the asymmetric centers bearing the methyl branches (D or L series), the fatty acids are called mycocerosic or phthioceranic acid, respectively, and the corresponding complex lipids are named dimycocerosates of phthiocerol (DIMs) or diphthioceranates of phthiocerol (DIPs). Structurally, DIM/PDIM molecule is similar to other phenolic glycolipids (PGLs) in that they all consist of a lipid core or diphthioceranates of phthiocerol. However, DIM/PDIM structure differs in that it is not terminated by glycosylated aromatic nucleus as is the case for PGLs (Figure 2.5). Recent research has clearly shown that DIMs are involved in mycobacterial pathogenicity, by the demonstration that \textit{M. tuberculosis} mutants deficient in the biosynthesis or transport of these molecules display attenuated multiplication in guinea pigs and mice (Perez \textit{et. al.}, 2004).

Recently, PDIM molecules have been found to be associated with several obligate and opportunistic mycobacterial pathogens; these include members of the \textit{M. tuberculosis} complex (\textit{M. tuberculosis}, \textit{M. bovis}, \textit{M. microti}), \textit{M. leprae}, \textit{M. kansasii}, \textit{M. marinum}, \textit{M. haemophilum}, and \textit{M. ulcerans} (Brennan, 2003). In \textit{M. leprae}, PGL have been shown to be species-specific serological markers. The only non-pathogenic spp. in which PDIM has been found are \textit{M. gastri} and \textit{M. tuberculosis} strain \textit{Canetti}. Owing to its significance in pathogenesis of \textit{M. tuberculosis}, these cell envelope-associated lipids, phthiocerol dimycocerosates (PDIMs), have lately received significant attention (Camacho \textit{et. al.}, 1999; Cox \textit{et. al.}, 1999).

2.4.2 Role of PDIM
Although PDIM is supposed to play a structural role in providing a stable base for the insertion of other lipids as well as act as a fluidity modifier whose function could be to modulate cell wall viscosity, its biologic activities listed below are those demonstrated or convincingly presumed:
Review of Literature

- Antigenic in leprosy and in tuberculosis (Torgal-Garcia *et. al.*, 1988).
- Non specific Inhibition of lymphoproliferative response irrespective of the donor (Mehra *et. al.*, 1984; Prasad *et. al.*, 1987). PGL-I was able to selectively inhibit the concanavalin-A-induced proliferation of peripheral blood mononuclear cells from patients with lepromatous leprosy but not from those with tuberculoid form of the disease, nor from those of healthy persons.
- Scavenging of oxygen radicals (Chan *et. al.*, 1989; Vachula *et. al.*, 1989).
- Suppression of monocyte oxidative responses (Vachula *et. al.*, 1990).
- Activation of complement for the phagocytosis of *M. leprae* by human monocytes (Ramanathan *et. al.*, 1990; Schlesinger and Horwitz, 1991). Loss of PGL was found to correlate with an increase in the release of the pro-inflammatory cytokines tumour-necrosis factor-a and interleukins 6 and 12 in-vitro. Furthermore, the overproduction of PGL by *M. tuberculosis* or the addition of purified PGL to monocyte-derived macrophages was found to inhibit the release of these pro-inflammatory mediators in a dose-dependent manner.

2.4.3 PDIM-Virulence Factor of *M. tuberculosis*

Using signature-tagged mutagenesis, Camacho and coworkers and Cox and his colleagues (Camacho *et. al.*, 1999; Cox *et. al.*, 1999) independently reported the finding that some *M. tuberculosis* mutants that are defective in the production of a lipid called PDIM are severely attenuated, with many exhibiting accentuated colony morphology change. They also show much higher cell-wall permeability, all fitting in with previous findings of Brennan and Goren (Brennan and Goren, 1977) that a PDIM-less variant of *M. tuberculosis* H37Rv showed attenuated growth in the guinea pig. Their findings strongly implied the role of PDIM and similar molecules as important virulence determinants. Consequently, the cluster of genes responsible for PDIM synthesis was termed as the ‘virulence gene cluster’ of *M. tuberculosis*. Recent advances in mycobacterial genetic techniques and the availability of the *M. tuberculosis* genome (Cole *et. al.*, 1998) have provided the opportunity to test directly the virulence of PDIM-deficient mutants. For example, mutation of *mmpL7* (Cox *et. al.*, 1999), the gene product of which may be involved in PDIM transport, and mutation of genes involved in PDIM synthesis, including *msl7* and *pks10* (Sirakova...
et. al., 2003) and pks7 (Rousseau et. al., 2003), yield strains with attenuated growth in the murine model. Over the years, the vital role played by many of the gene components of this virulence cluster in the pathogenesis of \textit{M. tuberculosis} has been firmly established (Azad et. al., 1996; Azad et. al., 1997; Fitzmaurice and Kolattukudy 1998; Sirakova et. al., 2002).

### 2.4.4 PDIM Biosynthesis Cluster

The publishing of the sequence of the entire \textit{M. tuberculosis} genome has initiated greater interest towards understanding the mechanism of drug resistance as well as the pathogenesis of the bacterium.

The biosynthesis of \textit{M. tuberculosis} virulence determinant PDIM has just begun to be elucidated and has been shown to involve a synthetic pathway in which a precursor, a C22-C24 fatty acid is elongated by three malonyl-CoA and two methylmalonyl-CoA units to yield a lipid core (Brennan, 2003).

The studies carried out by Camacho and his group as well as Cox and coworkers (Camacho et. al., 2001; Cox et. al., 1999) have neatly addressed the role of PDIM in disease pathogenesis. Using signature-tagged transposon mutagenesis, Cox and coworkers constructed a library of signature-tagged transposon mutants of this bacterium. The attenuated mutants allowed the identifications of 13 loci important for the multiplication of the bacterium within the lungs of mice. By means of these as well as other biochemical and immunological studies spanned over the years, it has now been shown that these thirteen different genes that are clustered on a 50 kb fragment of \textit{M. tuberculosis} genome are the ones that are involved directly/indirectly in PDIM biosynthesis. Seven out of these 13 genes have been shown to be involved in the biosynthesis of compounds of the cell envelope: phthiocerol, phenolphthiocerol and mycoside B (Azad et. al., 1996; Azad et. al., 1997; Fitzmaurice and Kolattukudy, 1998). Fascinatingly, these molecules are restricted to eight species of mycobacteria, namely, \textit{M. tuberculosis}, \textit{M. bovis}, \textit{M. africanum}, \textit{M. marinum}, \textit{M. ulcerans}, \textit{M. leprae}, \textit{M. gastri} and \textit{M. kansasii}. Seven of these are pathogenic. The vital role played by many of the gene components of this virulence cluster in the pathogenesis of \textit{M. tuberculosis} has now been firmly established (Rainwater and Kolattukudy, 1985; Azad et. al., 1997; Fernandes and Kolattukudy, 1998; Fitzmaurice and Kolattukudy 1998).
The PDIM biosynthesis cluster is characteristic of a type-I modular PKS, composed of five ORFs of phenolphthiocerol/phthiocerol PKS (ppsA-E PKS), the single ORF of mycocerosic acid synthase or mas, two genes for two acylCoA synthases or Acoas (fadD26 and fadD28), single gene for a type II thioesterase (tesA) and three genes for enzymes of ABC transporter family drrA-C and one gene for mmpL7, a transporter protein of RND permease superfamily that are thought to be involved in PDIM transport. The location of constituent genes in relation to each other within the PDIM biosynthetic gene cluster is shown in (Figure 2.6A).

Based on modular logic of PKSs, the biosynthetic machinery of PDIM can be visualized by dividing the complete process in five parts: 1). Activation and transfer of fatty acids onto the PpsA protein, 2). Synthesis of phthiocerol by modular PKS proteins, 3). Synthesis and release of mycocerosic acids, 4). Transesterification of mycoserosic acids onto the phthiocerol, 5). Transportation and localization of PDIM to cell wall.

The postulated pathways and organization of the gene cluster for the biosynthesis of phthiocerol and phenolphthiocerol are shown in figures 2.6A and 2.6B.

Step 1: Synthesis of phthiocerol by modular type I PpsA-E PKS proteins
Earlier work by Kolattukudy and his colleagues (Kolattukudy et. al., 1997) suggested that PpsA-E genes of the cluster were responsible for phthiocerol synthesis through a modular PKS mechanism. An FAS I-like system synthesizes a straight-chain fatty acyl group attached to its enzyme. Module I (PpsA) contains an acyl transferase domain, a ketoacyl synthase domain and a keto reductase domain, allowing introduction of the hydroxyl group. PpsB functions in identical manner. However, PpsD and E use methylmalonyl CoA, allowing for the introduction of the characteristic methyl branches of phthiocerol (Figures 2.6A and 2.6B).

Step 2: Synthesis and release of mycocerosic acids
Mycocerosic acid synthase encoding gene, located in the midst of PDIM operon is transcribed in the direction opposite to that of ppsA to E genes. It has been shown to encode an iterative PKS enzyme that acts like FAS I (Rainwater and Kolattukudy, 1985; Azad et. al., 1997; Fernandes and Kolattukudy, 1998; Fitzmaurice and Kolattukudy 1998). After four rounds of extension of a C₁₈ fatty acid using methylmalonyl CoA, MAS produces long-chain, multiply methylated branched fatty acid,
called mycocerosic acid. MAS have been functionally reconstituted \textit{in vitro}. MAS is a 464 kDa dimeric protein and contains six catalytic active sites. It is physiologically significant that the pps gene cluster occurs immediately upstream of mas, which encodes the multifunctional enzyme mycocerosic acid synthase (MAS), as their products phthiocerol and mycocerosic acid esterify to form the very abundant cell-wall-associated molecule PDIM (Rousseau \textit{et. al.}, 2003).

\textbf{Step 3: Activation and transfer of fatty acids onto PpsA protein and Transesterification of mycoserosic acids onto the phthiocerol.}

Traditionally, the biosynthetic machinery for fatty acids and polyketides has been studied independently, but recent evidence suggests that these enzyme systems may coordinate in many organisms to generate hybrid molecules of fatty acids and polyketides (Minnikin \textit{et. al.}, 2002). Various studies with Mycobacterial FadD proteins show that there are at least two classes of FadD-like proteins, that uses two different chemical means to activate fatty acids. FadD proteins present adjacent to PKS/NRPS cluster, such as FadD26, FadD28, FadD30, FadD32 and FadD29 activate metabolic carboxylates and synthesizes acyl-adenylates, whereas other FadD proteins activate them as acyl-CoA. Initial cell-free reconstitution studies of several FadD and PKS proteins have provided direct evidence for the mechanism of priming long chain fatty acids onto PKS proteins (Trivedi \textit{et. al.}, 2005).

The gene fadD26 of PDIM biosynthesis locus encodes a putative acyl-CoA synthase (Acoas). As shown Cox and co-workers (Cox \textit{et. al.}, 1999) in case of of PDIM production-defective mutants that contained a transposon insertion in the promoter region of ppsA-E operon, the expression of this ORF might be coupled to those of the downstream genes, ppsA to ppsE, which were shown to encode subunits of a PKS producing phthiocerol and phenolphthiocerol (Azad \textit{et. al.}, 1997). Because of this ‘expression coupling’, it is likely that FadD26 is linked either to the synthesis or to the transfer of these long chain diols. Recently, Fitzmaurice and Kolattukudy (Fitzmaurice and Kolattukudy, 1997; Fitzmaurice and Kolattukudy, 1998) demonstrated that an enzyme, FadD28, exhibiting more than 55 % identity to FadD26, is specifically involved in the transfer of mycoserosic acids onto phthiocerol and phenolphthiocerol. These diols are not only esterified by mycoserosic acid but also by other lipids, and FadD28 does not seem to be required for this process.
Therefore, it is hypothesized that FadD26 might be involved in the transfer of these other substrates onto phthiocerol and phenolphthiocerol.

Also, the signature-tagged mutant that had an insertion in fadD28 (one of 36 homologues of the *E. coli* fadD gene identified in the *M. tuberculosis* genome) failed to survive in lungs of mice (Cox *et. al.*, 1999). FadD28 however, is probably involved in acyl transfer of mycocerosic acid instead of fatty acid catabolism, as a mutation of the fadD28 homologue in BCG specifically inhibit mycocerosic acid synthesis and thus PDIM synthesis (Fitzmaurice and Kolattukudy, 1997).

**Step 4: Transportation and localization of PDIM to cell wall**

The French group of Camacho and coworkers (1999), in parallel to Cox and coworkers (1999), isolated mutants with insertions within drrC, mmpL 7, upstream from fadD26 and within fad26. The results from both laboratories were similar, although no PDIM production or production of a related lipid was seen in the upstream fadD26 genotype, *i.e.* Pps mutant or the fadD28 mutant. The mmpL7 mutant, like the Pps and fadD28 mutants, is specifically defective for *in vivo* growth as its doubling time in liquid culture is identical to that of the parent Erdman strain. Subsequently, it was found that the mmpL7 mutant could synthesize PDIM, but it could not excrete DIM into the medium.

In effect, one could conclude that mmpL7 and drrC mutants were defective in DIM export/excretion. The two mutated ORFs of PDIM biosynthetic cluster, drrC and mmpL7, both putatively encode proteins exhibiting similarities to transporter molecules. So far, their substrates remain unknown. However, the phthiocerol and phenolphthiocerol esters were found in the cell envelope. Presumably, they are then either synthesized in this compartment or translocated afterwards. In both cases, transport systems might be required. Moreover, the enzyme MAS has been suggested to be associated with the cell wall, although it lacks a typical signal sequence, suggesting that a secretion system other than the general secretory pathway is involved.

**The DrrA-C operon**

The drr operon encoding polypeptides similar to ATP Binding Cassette (ABC) transporters is present in a 50 kb virulence fragment of *M. tuberculosis* genome that also contains the PDIM biosynthesis locus. The genes encoding ABC transporters
occupy 2.5% of the genome of *M. tuberculosis*. This drug pump belongs to a larger family of ABC transporters that transport a variety of substrates—from small organic and inorganic compounds and cholesterol, to large toxins and proteins. The ABC transporters all contain a transmembrane domain and an ATP binding domain, and operate as homo- or hetero-dimers. Signature-tagged mutagenesis of drr operon led to strong growth inhibition of *M. tuberculosis* in the lungs of intravenously infected mice, as well as export of PDIM to the cell surface (Camacho *et. al.*, 2001). DrrA and DrrB behave as a functional doxorubicin efflux pump (Choudhuri *et. al.*, 2002).

**MMLP7 (Mycobacterial Membrane Protein Large)**

The genome of *M. tuberculosis* contains two genes that encode RND (resistance, nodulation, and cell division) proteins designated as MmpL (Mycobacterial membrane protein Large). These proteins are characterized by the presence of 12 transmembrane domains and two extracytoplasmic loops and have been reported in the genomes of organisms from all major kingdoms of life. In Gram-negative bacteria, these proteins facilitate the transport of a large variety of drugs, heavy metals, aliphatic and aromatic solvents, bile salts, fatty acids, detergents and dyes. The colocalization of some of the mmpL genes with genes involved in polyketide biosynthesis (pks genes) and genes involved in lipid metabolism (papA and fadD) suggests a similar role of these proteins in complex lipid transport in *M. tuberculosis* (Barry, 2001; Domenech *et. al.*, 2004). Indeed, the MmpL7 protein has been shown to be involved in the transport of PDIM.

**2.4.5 Cox Model**

Based on their experimental observations pertaining to the *M. tuberculosis* mutants, Cox and coworkers (Cox *et. al.*, 1999) have proposed a model for the synthesis and export of PDIM, delineating roles for the above-mentioned gene products of the virulence cluster. The model, shown in figure 2.7, proposes that the ppsA-E gene products are responsible for the elaboration of incoming long-chain (C<sub>20</sub>-C<sub>22</sub>) fatty acids by successive additions of malonyl or methylmalonylCoA extender units, typical of type I PKSs like the erythromycin PKS (Cortes *et. al.*, 1990; Donadio *et. al.*, 1991). The mas gene product on the other hand extends C<sub>18</sub>-C<sub>20</sub> fatty acids, also by addition of methylmalonylCoA units but by an iterative process (Fernandes and Kolattukudy, 1997). The products of mas action are then linked as esters to a diol.
moiety present in the polyketides that are synthesised by the pps PKS. The two acoas are involved in the release of the polyketide products from pps and mas proteins, with one of the two acoas namely, fadD28, may also somehow helping to bring together, perhaps even condense the two polyketides to form PDIM. PDIM is then transported by DrrA-C and mmpL7 from the cytoplasm to its eventual resting place, the outer cell wall. The Cox model is based on clear experimental observations made by them and by Camacho co-workers concerning cessation of PDIM biosynthesis in fadD26, fadD28, mmpL7 and drrC insertion mutants (Camacho et. al., 1999; Cox et. al., 1999; Camacho et. al., 2001). It is further supported by studies by Kolattukudy and co-workers that tangentially implicate fadD28 in off-loading of the polyketide from mas (Fitzmaurice and Kolattukudy, 1998).

2.5 An Overview of Bacterial two-hybrid System

The coordinated interaction of proteins with other biomolecules is a fundamental regulatory principle in cell biology. The analysis of protein interactions is primary to our understanding of the molecular basis of human diseases and has become central to many aspects in biomedical research. There is a need therefore to identify molecules that can specifically disrupt such a complex, e.g. for the targeted inhibition of particular biochemical pathways or the development of therapeutic agents.

Among the available methods for detecting protein-protein interactions, the two-hybrid systems are the most popular genetic method (Figure 2.8). In general, two fusion proteins called the bait and target are coexpressed in the cells. When the bait (usually a protein of interest) and its fused complement interact with the target and its fused complement, a biochemical pathway is activated. In contrast to biochemical methods that require purified proteins or antibodies, genetic two-hybrid systems allow interactions between known protein pairs to be quickly detected or allow screening of uncharacterized proteins from an appropriate library that specifically interact with a given protein of interest.

2.5.1 BacterioMatch™ Bacterial two-hybrid System

Stratagene’s BacterioMatch™ two-hybrid system is a fast, simple, and efficient method for detecting protein-protein interactions in vivo. As compared to conventional yeast two-hybrid systems, it offers the ability to screen larger libraries
for harder-to-find binding partners. In addition, the use of *E. coli* makes the system significantly faster than yeast systems.

2.5.2 Advantages of the BacterioMatch™ two-hybrid system Over Traditional Yeast two-hybrid Systems

The BacterioMatch™ two-hybrid system offers the ability to screen larger libraries for more difficult-to-find binding partners, due to the 1000-fold higher *E. coli* transformation efficiency. All screening is done in bacteria; transformation and plating is easy and colonies grow by the next day. Two-hybrid screening in bacteria also reduces the chance that the host harbors a eukaryotic homolog of one of the interacting protein partners, which may help to reduce false positives or potentially toxic effects.

2.5.3 The Workings of the BacterioMatch™ two-hybrid System

The BacterioMatch two-hybrid system is based on transcriptional activation as shown in figure 2.9. A protein of interest (the bait) is fused to the full-length bacteriophage λcl protein (λcl), containing the amino-terminal DNA-binding domain and the carboxyl-terminal dimerization domain. The corresponding target protein is fused to the N-terminal domain of the α-subunit of RNA polymerase (RNAPα). The bait is tethered to the λ operator sequence upstream of the reporter promoter through the DNA-binding domain of λcl. If the bait and target interact, they recruit and stabilize the binding of RNA polymerase close to the promoter and activate the transcription of the ampicillin-resistant reporter gene in the BacterioMatch two-hybrid reporter strain. The β-galactosidase reporter gene provides an additional mechanism to validate putative protein-protein interactions. The second reporter also helps in identification of false positives. The β-galactosidase gene expression is detected on X-gal plate supplemented with β-galactosidase inhibitor, phenyl ethyl β-D-thiogalactoside, which eliminates the background blue color formed in the reporter strain due to the basal transcription from the test promoter. Under such assay conditions, false positives do not produce enough β-galactosidase activity to form blue colonies, whereas strains bearing specific bait and target interactions turn blue (Dove *et. al.*, 1997; Dove and Hochschild, 1998).
2.5.4 The Sensitivity of the BacterioMatch™ two-hybrid System

The system can detect interactions between pairs of protein domains with equilibrium dissociation constants in the high nanomolar range. To date, a positive interaction has been detected using the dimerization domain of the yeast transcriptional activator Gal4 and a domain derived from the mutant form of the Gal11 protein, Gal11p. This protein pair has an equilibrium dissociation constant of approximately $10^{-7}$ M.
Figure 2.1 Map released by WHO (2005) depicting TB incidence rates.
Figure 2.2 Flow chart showing the organization of *Mycobacterium tuberculosis* complex (MTBC).
Figure 2.3 Pictorial representation of the Cell wall of *M. tuberculosis*.
Figure 2.4 Schematic representation of the decarboxylative condensation reaction in polyketide biosynthesis. The electrophile is attached to the ketosynthase (KS), whereas the nucleophile is attached to an ACP (adapted from Pfeifer and Khosla, 2001).
Figure 2.5 Model structures of PDIM and PGL.
The position of a variable number of carbohydrate residues (S) attached to the phenyl moiety are shown (adapted from Reed et al., 2004).
Figure 2.6 Organization of PDIM biosynthesis gene cluster and postulated pathways for the biosynthesis of phthiocerol and phenolphthiocerol.

A. The virulence gene cluster of *M. tuberculosis* that is responsible for the production and export of PDIM molecule (adapted from Rainwater and Kolattukudy, 1985).

B. The gene and domain organization of the enzymes postulated to catalyze the various biosynthetic steps of PDIM. *E* represents enzymes to which the intermediates may be bound during biosynthesis. KS, ketoacyl synthase; AT, acyl transferase; KR, keto reductase; DH, dehydratase; ER, enoyl reductase; ACP, acyl carrier protein.
Figure 2.7 A model representing the workings of the virulence cluster as proposed by Cox and co-workers (1999).
<table>
<thead>
<tr>
<th>Methods</th>
<th>E. coli</th>
<th>Mammalian</th>
<th>Yeast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial ToxR hybrid System</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Yeast two-hybrid system</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Mammalian two-hybrid system</td>
<td>No</td>
<td>Yes</td>
<td>*</td>
</tr>
<tr>
<td>Sos-recruitment system</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Split-ubiquitin system</td>
<td>No</td>
<td>Yes</td>
<td>?</td>
</tr>
<tr>
<td>Bacterial two-hybrid system (RNA polymerase)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Bacterial DNA loop formation system</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Bacterial hybrid repressor system</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Protein-splicing system</td>
<td>Yes</td>
<td>Yes</td>
<td>*</td>
</tr>
<tr>
<td>β-Galactosidase complementation system</td>
<td>Yes</td>
<td>Yes</td>
<td>*</td>
</tr>
<tr>
<td>DHFR complementation system</td>
<td>Yes</td>
<td>Yes</td>
<td>*</td>
</tr>
<tr>
<td>Adenylyl-cyclase reconstitution system</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

* Method applicable in principle to the listed cells. DHFR, dihydrofolate reductase.

**Figure 2.8** Analysis techniques for detecting *in vivo* protein-protein interactions.
Figure 2.9 Schematic diagram of Bacterial two-hybrid system on RNA polymerase recruitment.

αN, N-terminal domain of RNA polymerase α subunit; X and Y, interacting polyfused to λ cl and α N; β, β', σ, subunits of RNA polymerase.