CHAPTER 1

INTRODUCTION

History of Tuberculosis

Molecular Epidemiology
A) Definition
B) IS6110 and DR based fingerprinting of *M. tb.*
C) Fluorescent Amplified Fragment Length Polymorphism
D) Molecular Genotyping

Assay for Drug Resistance

Molecular Basis of Multiple Drug Resistance
A) Rifampicin
B) Isoniazid
C) Streptomycin
D) Fluoroquinolones
E) Drug Efflux Pumps
The Greeks called it Phthisis (wasting away), medieval Europe knew it as Kings evil, to Hindus it was Sosha (cough) or Rajayakshma (wasting) and others as Consumption, Scrofula, Potts disease or lupus vulgaris (Reibman, 1995). Today we realize that all these were nothing but manifestations of one disease—Tuberculosis (TB) caused by pathogens belonging to the *Mycobacterium tuberculosis* complex. TB finds mention in ancient Indian texts and was first described by Hippocrates as Phthisis, meaning wasting away of the body, which remains a primary symptom of untreated TB and carries a lot of economic significance even today. TB remained an obscure disease in ancient civilizations probably due to a very low number of cases. However, there is evidence that the disease existed in the ancient world as seen in 3000-year old mummies which show classical bone anomalies associated with TB (Salo et al, 1994). The disease assumed importance and notoriety in feudal Europe where poor living conditions, poverty and nonexistent treatment caused epidemics of TB with high mortality. The disease spread to America, Asia, Africa and other pacific islands causing further mortality. Tuberculosis, inspite of being curable, remains a deadly killer worldwide. The morbidity and mortality statistics of TB make it one of the most infectious diseases known to humans. It is estimated that almost one-third of the human population has been infected by *M. tuberculosis* (Bloom and Murray, 1992). The disease, once believed to have been conquered, has resurfaced and this time with more resistant strains. *M. tuberculosis* infection has proved to be the most frequent opportunistic infection in Acquired Immune Deficiency
Syndrome (AIDS) patients and remains the single largest killer of AIDS patients. With the AIDS pandemic, the spread of *M. tuberculosis* and mortality due to it has increased dramatically (Iseman and Sabarbaro, 1992).

Historically, the causative organism is supposed to have evolved into its human host via domesticated animals, when humans started settling down and an agrarian social order came into effect (Bates and Stead, 1993). Mycobacterium probably emerged from the soil to find a niche first infesting and then infecting various mammals and birds. Based on genetic relatedness *Mycobacterium bovis*, a common animal pathogen, and *Mycobacterium tuberculosis* share a common ancestor. It is reasonable to infer that at some period of time *M. bovis* underwent subtle changes to adapt to a parasitic role in humans. The tubercle bacilli scores over *M. bovis* in that it has a natural reservoir in humans, diminished virulence potential in animals and exhibits airborne transmission.

The renaissance in the seventeenth century saw the scientist trying to understand and find cures to this dreaded scourge. Jean Antoine Villemin in 1865 demonstrated that TB was contagious and seventeen years later Robert Koch isolated and cultured *M. tuberculosis* from crushed tubercles. These discoveries were of immense importance in fighting TB. By the late eighteenth century, sanitaria for TB patients were established which helped in containing the disease. Further, the pasteurization of milk reduced the cases of *M. bovis* infection. Albert Calmette and Camille Guérin introduced the BCG, an attenuated strain of *M. tuberculosis*, as a vaccine in the year 1908, which is still
in use. Finally the discovery of antibiotics and isoniazid provided the much needed drugs to combat TB (Iseman, 1994).

Effective health programs, a battery of new drugs and a general improvement in living conditions put a brake to the TB epidemics. The number of cases of TB and the mortality rates in the early nineteenth century became so low that people believed that the battle against TB was won. Initial euphoria over the ability to cure TB led to overconfidence in people involved in the management and treatment of the disease. The tedious drug regimens were not monitored carefully and health programs were allowed to become lax. This, combined with AIDS pandemic, allowed the disease to re-emerge.

Because HIV infects, disables, and kills the cell that is central to tuberculosis immunity- the CD4\(^+\) or helper T lymphocyte- the viral epidemic led to a dramatic upsurge in tuberculosis in regions such as sub-Saharan Africa, south-east Asia and regions in South and North America. Analysis of nosocomial outbreaks of TB in some New York hospitals demonstrated clearly that the impairment of defences associated with HIV disease facilitated the transmission of multidrug-resistant tuberculosis (MDR-TB) (Pitchenik et al, 1990; Edlin et al, 1992; Coronado et al, 1993).

Drug resistance in *M. tuberculosis* was recognized soon after streptomycin was introduced as an anti-tuberculous drug in the early 1950s. In later years, with the introduction of new drugs, resistance to them was also discovered. Not much notice was given to this resistance as it remained restricted to one drug and could be combated with other new drugs (Heym and Cole, 1996). It was soon
recognized that treatment with a battery of drugs as compared to a single drug therapy was much more effective in eliminating and preventing multi drug resistant tuberculosis (MDR TB). Simply, the combination of drugs assured that bacilli which developed resistance to one drug were killed by another. Secondly, the drugs were able to combat the different type of bacilli viz. a viz. actively dividing, dormant or resting bacterium. The mechanism of resistance to most drugs remained unknown till the last decade. After the initial discovery of five or six anti-tuberculous drugs, no better drugs have been made available against \textit{M. tuberculosis}. Most of the new drugs have proved to be too toxic for human consumption.

A recently published WHO report reviewing the global status of tuberculosis has pointed to an increasing incidence of drug resistance tuberculosis (Cohn et al, 1997). Rates of primary resistance to isoniazid, administered as a single agent, ranged from 0 to 16.9% (median rate, 4.1%); to streptomycin, 0.1%-23.5% (median, 3.5%); to rifampicin, 0-3.0% (median, 0.2%); and to ethambutol, 0-4.2% (median, 0.1%). The rates of acquired resistance to these agents, which were higher than those of primary resistance, were as follows: isoniazid, 4.0%-53.7% (median rate, 10.6%); streptomycin, 0-19.4% (median, 4.9%); rifampicin 0-14.5% (median, 2.4%); and ethambutol, 0-13.7% (median, 1.8%). The highest rates of multidrug-resistant tuberculosis have been reported in Nepal (48.0%), Gujarat, India (33.8%), New York City, (30.1%), Bolivia (15.3%), and Korea (14.5%). Further the report points to the alarming increase in the number of TB patients on the subcontinent. India in particular has been singled
Introduction

out as having the greatest burden of TB patients. This number is going to increase in the coming years unless more effective health programs like directly observed treatment, short course (DOTS) are implemented. There are only a couple of studies on the prevalence of drug resistance tuberculosis in India (Chowgule et al, 1998, Das et al, 1995). Moreover, there are no studies from India, on the epidemiology of drug resistance strains or characterizing the mutations leading to resistance in the multidrug resistant strains.

MOLECULAR EPIDEMIOLOGY

Most of our understanding of tuberculosis (TB) dynamics in populations have been derived indirectly by inference from descriptive epidemiological data, as well as from clinical trials conducted to test the efficacy of antituberculous drugs and vaccine immunization programs. Usually a small number of infected individuals end up with the disease, also the long incubation period between infection and disease makes it difficult to identify source of infection and site of transmission. The diagnosis of TB can be established definitively, only by isolation of *Mycobacterium tuberculosis* by culture from the individual suspected of having the disease. Microscopic examination of sputum can detect organisms only when large numbers are present, thus it has sensitivity of no more than 50 to 70 percent compared to culture. Because of high prevalence of tuberculosis and widespread use of Bacillus Calmette-Guérin (BCG) vaccine, a positive tuberculin skin test does not identify the source or duration of the infection. The ability to determine transmission links is thus, limited
by technology. The integration of molecular methods of strain identification with conventional epidemiology can be used to establish transmission links and to identify individual and environmental risk factors. The discovery of polymorphic DNA in *Mycobacterium tuberculosis* (*M. tuberculosis*) has led to strain differentiation becoming an important tool in the epidemiology of TB (Eisenach et al, 1988).

One new tool that has been applied to TB epidemiology is restriction fragment length polymorphism (RFLP) analysis-based DNA fingerprinting that identifies specific strain of microorganisms. This technique has found widespread use in *M. tuberculosis* characterization because of identification of repetitive DNA sequences that are present in the genome in variable numbers and locations. These include the insertion sequences (IS), direct repeats (DR), polymorphic GC-rich sequences (PGRS), major polymorphic tandem repeats (MPTR) and enterobacterial repetitive intergenic consensus sequences (ERIC). The greater polymorphism shown by these repetitive sequences offer better identification power over the earlier typing techniques such as pulse field gel electrophoresis. Another advantage of these techniques, is their standardization, which offers greater intra- and inter-laboratory comparisons of fingerprints. Few years back a new method 'Spoligotyping' was introduced for diagnosis and epidemiology of *M. tuberculosis*. It is based on the detection of various non-repetitive spacer sequences located between the direct repeats in the DR Locus. It is a PCR based method hence the results can be obtained in less time (Goguet et al, 1997). However the discriminative power of spoligotyping compared to IS6110
fingerprinting is less, and consequently, its use for epidemiological studies of *M. tuberculosis* is limited. The latest method for strain differentiation is the Fluorescent Amplified Fragment Length Polymorphism (FAFLP).

**IS6110 and DR based fingerprinting of *Mycobacterium tuberculosis***

The initial attempts (Collins and De Lisle, 1985) of typing *M. tb.* by restriction digestion patterns were only partially successful. The use of multiple restriction enzymes, the complexity of the resulting patterns, and the limited ability to compare results from different experiments made the broad application of this approach impractical. However, by using repetitive sequences as probe it has become possible to overcome most of the difficulties in the earlier approaches. The insertion element IS6110 has been most extensively studied as the genetic marker for *M. tb.* DNA polymorphism (Mazurek et al, 1991). This 1335-bp insertion sequence (IS) element, described a decade ago (Thierry et al, 1990), is related to the IS3 family and has been found exclusively in all isolates of the *M.tb* complex (McAdam et al, 1991). There are usually 1 to 25 copies of this element but clinical isolates from Asia having no copy have also been described (van Soolingen et al, 1991). IS6110 has an imperfect 28 basepair (bp) inverted repeat at its ends, and generates a 3- to 4-bp target duplication on insertion. Sequence alignment of different copies of IS6110 from different strains showed no difference (Thierry et al, 1990) except the one from BCG (named IS987) which has a 6
nucleotide deletion (Hermans et al, 1991). There seems to be no consensus target sequence for insertion [as revealed by sequencing the chromosomal junctions (Mendiola et al, 1992)] but recent reports indicate that the insertion may not be absolutely random. Mycobacterium bovis BCG contains a single copy of the insertion element and in all strains this copy is integrated at the same site in the chromosome. The flanking region of this insertion site has multiple copies (~49) of virtually identical direct repeats each separated by 35 to 41 bp spacer DNA (Figure I). The insertion takes place in the 30th DR. It was shown (Hermans et al, 1991) that even M. tuberculosis strains have a single or tandem copies of the element at this position. It seems that the DR cluster is a specific, hot spot region for integration of insertion elements in the chromosome of M. tuberculosis complex strains. In another report (Fang and Forbes, 1997) the authors have described a 267 bp nucleotide region in the M. tuberculosis genomic sequence which can harbour the IS6110 at six different locations. They named it ipl (IS6110 preferential locus) and all insertions at this site are in the same orientation.

Another group (Mchugh and Gillespie, 1998) compared M. tuberculosis fingerprints from two collections at London and Tanzania for checking the probability of finding a band at a certain position. They calculated the frequency of occurrence of a band at a given position in the fingerprints included in the study. They report that certain bands appear at certain positions more frequently than could be attributed to chance, thus, indicating that certain regions in the M. tuberculosis genome might be actually preferred sites for insertions.
Figure 1: A schematic representation of a copy of IS6110 inserted in the DR region of Mycobacterium tuberculosis genome. The INS1 and INS2 primers are used for amplifying a region of the element which is used as the probe in RFLP. (IR-Inverted Repeat)
Inspite of these studies the IS6110 have been used and proven to be an effective probe for strain identification in various epidemiological settings.

To facilitate the intra- and inter-laboratory comparisons of fingerprints obtained using this technique, a standard methodology for typing was proposed in 1993 (van Embden) and applied world over for epidemiological purposes. In this method the clinical specimen of \textit{M. tuberculosis} is cultured for DNA extraction. The DNA is subjected to digestion with restriction endonuclease, Southern blotted and probed with (or a part of) the repetitive sequence. Three important parameters necessary for standardization are the specificity of the restriction enzyme, the DNA probe and the use of an appropriate molecular mass standard. The recommended enzyme is \textit{PvuII} that cuts the IS6110 only at one site. The probe is usually an amplified region of the IS6110 that is at the right side of the \textit{PvuII} site. This approach minimizes the number of bands generated and simplifies the analysis. A molecular standard in the range of 0.5-10 kilobase (Kb) covers most of the IS6110 hybridizing fragments.

Despite the high degree of IS6110-associated restriction fragment length polymorphism (RFLP) among \textit{M. tuberculosis} strains, no IS6110 transposition has been observed during prolonged \textit{in vitro} culturing or \textit{in vivo} growth (Niemann et al, 1999, Boer et al, 1999) suggesting that the frequency of transposition is relatively low. This makes the IS6110 an excellent genetic marker to trace individual \textit{M. tuberculosis} strains during microepidemics, nosocomial infections and the clonal dissemination of multiple drug-resistant strains. This fact is

Another marker used in fingerprinting of \textit{M. tuberculosis} is the 36-bp Direct Repeat (DR) sequence of which multiple copies are present at the hot spot of IS6110 integration and these copies are usually interspersed by non repetitive sequences of 36 to 41 bp (Hermans et al, 1991). The DR probe can be used as a secondary molecular marker to support or rule out clustering of strains with fewer than five copies of IS6110. DR has proved especially useful in strains having none, one or low copy numbers of IS6110 (van Soolingen et al, 1993). This marker has been used in many studies and is routinely used for characterizing strains with less than five IS6110 copies (Sola et al, 1997, Torrea et al, 1995, Yang et al, 1994, van Soolingen, 1993).

**Fluorescent Amplified Fragment Length Polymorphism**

Fluorescent amplified fragment length polymorphism (FAFLP) analysis provides a means of examining DNA segments distributed over the entire genome of an organism. This information can be used to dissect the routes of infection, study evolutionary genetics and identify novel genes involved in resistance and virulence of pathogens in an epidemiological setting. It is a modification of Amplified Fragment Length Polymorphism (AFLP) technique, commonly used to type infectious agents (Lindstedt et al, 2000, Aarts et al, 1999). This technique offers more discriminatory powers than other fingerprinting
techniques. The information generated by this technique can be digitized offering quicker dissemination and inter laboratory comparisons. Other advantages of this technique are the requirement of less amount of starting DNA and the range of markers generated which can be used for diagnosis or may possibly reflect the resistance and virulence loci in the genome of M.tb. strains.

The technique is based on the digestion of known/unknown DNA with two restriction enzymes (REs). Generally, one enzyme is a frequent cutter while the other is a rare cutter. The digested fragments are ligated to adaptors (with RE recognition sequences identical to the ones used in digestion). Primers based on the adaptor sequences are then used for PCR amplification. Sequences to which both adaptors are attached are exponentially amplified while fragments with one attached adaptor amplify linearly and are thus in negligible quantity. A second optional amplification can be carried out by primers, which extend into the unknown sequence by attaching one to three bases at the 3' end of the primers. This leads to selective amplification of sequences having bases complementary to the ones attached to the primer(s). The primer (forward or reverse) is tagged with fluorescent dyes, which when excited, emit light of a particular wavelength, which forms the basis of visualizing the amplified DNA. These signals represent a particular length of the DNA fragment amplified and form the basis of typing. The strains with the identical set of amplified products represent a single strain. This technique offers the advantage of typing the strains based on the polymorphism present
over the whole genome and thus has added phylogenetic value (Goulding et al, 2000).

**Molecular Genotyping**

According to Sreevatsan et al (1997), the *M. tuberculosis* strains can be classified into three genetic groups based on two polymorphisms that occur with high frequency in the genes encoding catalase-peroxidase (*katG*) and the A subunit of DNA gyrase (*gyrA*). Group 1 has the *katG* codon 463 CTG (Leu) and *gyrA* codon 95 ACC (Thr); group 2 has the *katG* 463 CGG (Arg) and the *gyrA* codon 95 ACC (Thr), and group 3 has the *katG* codon 463 CGG (Arg) and the *gyrA* codon 95 AGC (Ser). In terms of evolutionary hierarchy, Group 1 represents the older isolates followed by Group 2 and Group 3.

**ASSAY FOR DRUG RESISTANCE**

The three conventional methods still in use for susceptibility testing are Absolute concentration method, Proportion method, and Bactec radiometric analysis. The proportion method is a more popular and widely used technique to determine susceptibility of *M. tuberculosis*. In this method various dilutions of inoculum are grown on control medium to give colonies which are counted. Similarly, the colonies grown on drug containing media are also counted and compared to the number of organisms in the inoculum. Thus the proportion of organisms resistant to the drug can be measured and expressed as a percentage of the total mycobacterial population (Cannetti et al, 1969). Another method in use for susceptibility testing is the
mycobacteria growth indicator tubes (MGIT) method (a modification of the Bactec method). MGITs contain enriched Middlebrook 7H9 broth with an oxygen-sensitive fluorescent sensor to indicate microbial growth. Drugs are added to the broth at critical concentrations for susceptibility testing.

The progress made in molecular genetics of mycobacteria has resulted in newer molecular methods for detection of drug resistance in clinical isolates. These include Polymerase Chain Reaction-Single Strand Conformational Polymorphism (PCR-SSCP) (Heym et al, 1995, Telenti et al, 1993a), automated DNA sequencing (Telenti et al, 1993b) and a biological assay using luciferase reporter phages (Jacobs et al, 1993). While the former two methods involve screening of mutational hot spots in the genes encoding drug targets, the latter is based on the viability of the mycobacterium in the presence of the drug(s).

The principle of PCR-SSCP is based on the fact that the two denatured strands of the DNA (in this case PCR-amplified) molecule adopt stable intra-molecular conformations that may differ from the wild type upon mutation(s). This causes a change in the electrophoretic mobility of the strands. The other way to get an unambiguous result is to sequence such hot spots within the genes. This is possible using the automated sequencing systems available. The luciferase assay involves the infection of mycobacteria with phage carrying the firefly luciferase gene. In the presence of ATP, found only in viable organisms, luciferase produces light from its substrate luciferin. If mycobacteria are treated with drug(s) prior to infection with the
reporter phage, light is only produced by viable or drug-resistant mycobacteria (Jacobs et al, 1993).

**MOLECULAR BASIS OF MULTIPLE DRUG RESISTANCE**

The mycobacterial strains which are resistant to at least two of the frontline drugs rifampicin and isoniazid are considered to be multi-drug resistant (MDR). Rifampicin and isoniazid form the backbone of the short course therapy for treating tuberculosis. These drugs are the most effective, and resistance to them would necessitate using drugs which are more toxic coupled with a longer treatment period (Bloom and Murray, 1992).

The acquisition of resistance by the bacterium is a random event, and the chance of acquiring multidrug resistance is $10^{-14}$ (Harkin and Harris, 1995). These strains have been demonstrated all over the world. The frequency of resistance to multiple drugs varies geographically and acquired (secondary) resistance is more common than primary resistance. Delineation of the molecular mechanisms of antimicrobial agent resistance will and has resulted in the development of several strategies for early detection and treatment of drug resistant strains.

Resistance to drugs primarily arises due to mutations in genes coding for drug target proteins (Ramaswamy and Musser, 1998). However, this is not the exclusive mechanism employed by the pathogen to evade killing by the drug. All mycobacteria, especially *M.tb.*, have a complex cell wall which blocks drug entry into the cell and thus prevents the drug from accessing it's cytoplasmic targets. A more
recent mechanism which is well documented in other systems is the presence of membrane proteins which act as drug efflux pumps (Lewis, 1994). The *M. tb.* genome has nearly 20 such pumps (Cole et al, 1998). However, none of these have been conclusively shown to operate in a clinical setting, to give rise to drug resistance. Another mechanism for acquiring resistance that has been documented in other systems, but not yet in *M.tb.*, is drug alteration.

**Rifampicin**

Resistance to anti-tubercular drugs is attributed to the specific change(s) in nucleotide sequences within the cellular targets of these drugs. The probabilities with which resistance to anti-tubercular drugs is acquired varies and is the least (10^{-8}) for rifampicin (Shimao, 1987). Resistance to rifampicin is a relatively rare event and leads to selection of mutants that are already resistant to other components of short-course chemotherapy. Therefore, rifampicin resistance is often regarded as an excellent surrogate marker for MDR-TB (Siddiqi et al, 1998; CDC, 1993). The association of the RNA polymerase β (rpo β) subunit gene with resistance to rifampicin has been documented previously and subsequent reports from various groups have confirmed this association in clinical isolates of *M.tb* (Musser, 1995; Kapur et al, 1994; Williams et al, 1994).

Rifampicin is a lipophilic ansamycin and its efficacy as an anti-tubercular drug lies in its ability to diffuse across the hydrophobic cell envelope. The ansa designation connotes an aromatic center that is bridged on both the ends by an aliphatic chain. The conformational
relationship between the aromatic nucleus and the aliphatic chains is very important for microbiologic activity, probably because of the interaction of the drug with its target. It is a potent inhibitor of DNA-dependent RNA polymerase. Rifampicin is the most potent of the antituberculous agents in converting positive sputum cultures to negative. This characteristic of "sterilization" has been attributed to rifampicin's ability to affect dormant organisms. It is hypothesized that these dormant organisms become metabolically active intermittently. This window of time is sufficient for rifampicin to be effective while the same is not true for other drugs. RNA polymerase is a multisubunit oligomer consisting of a core enzyme having four polypeptide chains (α2ββ'). The holoenzyme has an additional subunit σ, that allows promoter recognition for initiation of transcription. The subunits αββ'σ are coded by the *rpo* A, *rpo* B, *rpo* C and *rpo* D genes respectively (Burgess et al, 1987). Rifampicin binds to the β subunit involved in the initiation and elongation of transcription. The *rpo* β gene of *M. tuberculosis* has been cloned and sequenced and is similar to the corresponding gene of *E. coli* (Figure II). Most of the rifampicin resistant phenotypes of *M. tuberculosis* are due to point mutations within a short region of 27 codons mapped between amino acids 511 to 533, a region also known to confer resistance in *E. coli*. Most of these point mutations are mis-sense mutations resulting in single amino acid substitutions. A small percentage of mutations are insertions and deletions. The two most frequent substitutions are His526Tyr and Ser531Leu. Additional mutations at positions 511, 512, 513, 516 and 533 have been mapped.
Figure II: The reported mutations in the rpoB gene of rifampicin resistant isolates of Mycobacterium tuberculosis. The codons 508 to 532 represent the hot spot region. The middle panel shows the wild type sequence. Top panel has the nucleotide substitutions and the bottom panel depicts the corresponding amino acid changes.
Introduction

A study from Japan (Taniguchi et al, 1996) established for the first time a relationship of these mutations to the level of resistance demonstrated by the strain. Isolates with mutations in codons 513, 526 and 531 had high levels of drug resistance (MIC ≥ 50 μg/ml). In contrast, amino acid substitutions located at position 514, 521 or 533 resulted in low-level resistance (MIC ≤ 12.5 μg/ml). These results have since been reinforced by other studies (Moghazeh et al, 1996; Ohno et al, 1996; Bodmer et al 1995).

Preliminary studies with the E. coli β subunit suggest that mutations in these regions lead to loss of binding sites for rifampicin (Tavormina et al, 1996). Most of these mutations seem to confer no growth disadvantage to M. tuberculosis unlike E. coli, where growth is severely affected.

Isoniazid

Isonicotinic acid hydrazide (Isoniazid or INH) was first synthesized in 1912 but used as an anti-tuberculous drug only after 30 years. It is the generally prescribed synthetic drug for active infection and prophylaxis. The low cost, bioavailability, excellent intracellular penetration, and narrow spectrum of action makes it an almost ideal antimicrobial agent. However, the high frequency of spontaneous development of resistance of M. tuberculosis to this drug (1 in 10^6 to 10^7) precludes its use for a single-drug regimen. While the drug has been in use since 1952 as an anti-tubercular, its mechanism of action is still not fully understood (Figure III). Resistance to isoniazid is known to arise due to mutations in the katG (catalase-peroxidase) (Musser, 1995; Saint-Joanis, 1999) inhA (enoyl-ACP reductase)
(Morris et al., 1995; Banerjee et al., 1994), \textit{ahpC-oxyR} (alkylhydroperoxidase) (Collins and Wilson, 1996; Zhang et al., 1996) and \textit{kasA} (\(\beta\)-ketoacyl-ACP synthase) gene loci (Mdluli et al., 1998). According to one of the hypotheses isoniazid acts as the prodrug which is converted to an active form (isonicotinic acid or aldehyde-bearing groups or free radicals) by the \textit{katG}-encoded catalase-peroxidase enzyme in \textit{M. tuberculosis} and is toxic to these cells (Basso et al., 1998, Heym et al., 1993). This enzyme is bifunctional with regards to its catalytic activity. In the catalase mode it converts \(2\text{H}_2\text{O}_2\) to \(2\text{H}_2\text{O}\) and \(\text{O}_2\), whereas in its peroxidase mode it accepts electrons from a variety of organic electron donors and uses them to reduce \(\text{H}_2\text{O}_2\) to water. It has been shown to play a protective role against noxious oxidative agents, like \(\text{H}_2\text{O}_2\), that accumulate during oxidative respiration (Shoeb et al., 1985). It has also been shown to be important in growth and survival of the mycobacterium within the macrophages (Li et al., 1998). Isoniazid resistance has been attributed to deletions, insertions and point mutations in this gene of the mycobacterium. Most of the mutations identified are of missense type (Musser et al., 1996; Heym et al., 1995, Rouse et al., 1995). These mutations were localized to the active site or the heme-binding site of the catalase-peroxidase enzyme. The most common mutation encountered is the replacement of Arg463 to Leu.
FIGURE III: The mechanism of action of isoniazid. The prodrug is converted to active form in the cytoplasm by the Kat G protein. The active ingredient then attacks enoyl reductase enzyme which is involved in elongation/branching of mycolic acid chains. The possibility of other targets cannot be ruled out.
In *M. bovis* BCG there is a strong correlation between this mutation and resistance but in *M.tb* there seems to be no direct relation as isoniazid susceptible strains also have this polymorphism (Sreevatsan et al, 1997). This mutation is in the carboxyl terminal of the KatG protein and is probably involved in the substrate binding. Other common mutations are Ser315Thr, Asn138Ser/His, Trp328Gly/Leu/Cys and Thr275Pro (Figure IV).

Another hypothesis takes into account the loss of acid-fastness of *M. tuberculosis* treated with INH. This was attributed to inhibition of mycolic acid synthesis. The enoyl-ACP reductase, coded by *inhA* gene, is supposedly the target (Banerjee et al, 1994). This enzyme requires NADH as a cofactor; the mutant enzyme has been shown to have a lower affinity for NADH and cannot be saturated at NADH concentrations existing within *M. tuberculosis* (Johnsson and Schultz, 1994). The common missense mutation found in this gene was lle16Thr (Musser et al, 1996; Kapur et al, 1995). The putative regulatory region (ribosome binding site) of this gene has been also shown to be mutated in certain isoniazid resistant strains (Figure V) (Musser et al, 1996). The mechanism by which the mutations in this region causes resistance to isoniazid is not clear, however an overexpression of this protein could lead to the titrating out of the drug (Mdluli et al, 1996).

The mutations in *kasA* gene have been associated with isoniazid resistance. The G269S and G312S mutations were common at this locus (Mdluli et al, 1998).
Figure IV: A schematic representation of the reported mutations in the KatG protein. The numbers represent the codons where the polymorphism occurred and different arrows indicate the position and type of mutation.
Figure V: A schematic representation of the mutations identified in the putative ribosomal binding site of the InhA gene. The inhA locus is composed of two contiguous open reading frames (designated mabA and inhA) which are separated by a 21-bp non-coding sequence.
Streptomycin

Streptomycin is a broad spectrum antibiotic of the aminoglycoside family. It was the first drug that was shown to be effective in combating tuberculosis. Streptomycin was originally isolated from two strains of an actinomycetes species by Waksman and coworkers in 1943. But soon after its clinical use, mutant strains of M.tb resistant to streptomycin were discovered. Some other aminoglycosides such as kanamycin, amikacin, and closely related polypeptide, capreomycin serve as second line agents in the treatment of TB. They offer no advantage over streptomycin in the treatment of TB except against some drug-resistant organisms and they have similar or worse toxicity profiles.

Streptomycin is a protein synthesis inhibitor and acts at mRNA translation, misreading of genetic code and wrong proof reading by the ribosome. Resistance to streptomycin involves mutations in the 30S ribosome unit (Moazed and Noller, 1987). The rslp gene encoding the S12 protein carries point mutations that result in the conversion of critical lysine residues, K43 to arginine or threonine and K88 to arginine (Figure VI). About 70% of the resistant isolates carry the former substitution (Bottger et al, 1994, Heym et al, 1994). These regions are known to be required for binding to the S12 protein (Bottger, 1994). The substitution at codon 88 is less frequent (Meier et al, 1994). Mutations in the rrs locus (encoding 16s rRNA) have been mapped to two regions, the 530 loop and the 915 region (Figure VII) (Finken et al, 1993). The position 513 had A→C transversion and position 516 had C→T transition.
Figure VI: Schematic representation of mutations located in the \( rpsL \) gene encoding the S12 protein, associated with streptomycin resistance.
Figure VII: The secondary structure of the 16S rRNA of *Mycobacterium tuberculosis*. The mutations in the 530 loop and 904 regions are marked by arrow heads, and these lead to streptomycin resistance.
In another study with strains exhibiting high level of streptomycin resistance (>6μg/ml), there were mutations at positions 491, 512 (C→T) and 904 (A→G) (Meier et al, 1994). In yet another study mutations at the position 903 (C→G, C→A) were attributed to resistance to streptomycin (Honore and Cole, 1994). In summary almost two thirds of the resistant isolates sequenced so far had demonstrable mutations in the above two loci. This indicates that there are more mechanisms conferring streptomycin resistance and these need to be discovered.

### Fluoroquinolones

These are synthetic derivatives of nalidixic acid. The target of fluoroquinolones' action is the bacterial DNA gyrase, an ATP-dependent type II DNA topoisomerase that catalyses the negative supercoiling of DNA. This enzyme is made up of four units (α2β2) that are encoded by the *gyrA* and *gyrB* genes, respectively. Fluoroquinolones bind to the gyrase and inhibit the supercoiling of DNA (Leysen et al, 1989). The eukaryotic homologue of gyrase, topoisomerase II, is orders of magnitude less sensitive than gyrase to the quinolones; making the quinolones clinically useful antibacterials (Honk Kong chest service, 1992; Jain et al, 1996). The *gyrA* and *gyrB* genes of *M. tuberculosis* have been cloned and mutations in the quinolone-binding site have been mapped (Takiff et al, 1994; Cambau et al, 1994). In subsequent studies the hot spot region for mutations in the gyrase A gene was identified (Figure VIII).
Figure VIII: A schematic representation of the reported mutations in gyrA fluoroquinolone resistant strains of *Mycobacterium tuberculosis*. The codons 87 to 96 form the hot spot region for mutations at this locus. The middle panel has the wild type sequence. The top panel shows the nucleotide mutations and the bottom panel depicts the corresponding amino acid changes.
This extends from codon 81 to 95 and polymorphism at codons' 87, 90, 91, 94 and 95 are common (Alangaden et al, 1995). The mutations Ala90Val and Asp94Asn, His, Gly or Tyr have been implicated in conferring high degree resistance to ciprofloxacin (Xu et al, 1996). The codon 495 in gyrB gene has been shown to be polymorphic and contributes to weak resistance to quinolones (Kocagoz et al, 1996). Almost 80% of fluoroquinolone resistant M. tuberculosis have been shown to have mutations in gyrA or gyrB genes. The rest 20% of the resistant strains likely have decreased drug permeability or active efflux pumps.

**Drug Efflux Pumps**

A major contributor to drug resistance in human pathogens and cancer cells are the drug efflux pumps. Reports on presence of such pumps are available in bacterial pathogens to human cancers. Mycobacterial species are no exception and there are few reports on presence of such pumps in *Mycobacterium smegmatis* (Takiff et al, 1996; Liu et al, 1996; Banerjee et al, 1998), *Mycobacterium fortuitum* (Ainsa et al, 1998) and *Mycobacterium tuberculosis* (Silva et al, 2001; Doran et al, 1997). In the published sequence of *M. tb* strain H37Rv there are twenty such putative efflux proteins. There is no direct evidence of the involvement of any of these proteins to drug resistance in *M.tb*, however, examples in other bacterial systems prompt us to believe that *M. tb* will probably have more than one drug efflux pump. This hypothesis is supported by the fact that many researches have been unable to relate the level of resistance in MDR strains to the
mutations in the target genes. In few cases there are no mutations in the target genes that can account for the resistant phenotype.

Most bacterial drug resistant pumps are from the Major Facilitator family of efflux pumps (Lewis, 1994; Griffith et al, 1992). These are membrane translocases and use a proton motive force as a source of energy. Few examples of this family are TetA, B and C proteins which mediate tetracycline export in gram negative bacteria. NorA a quinolone resistance protein of *Staphylococcus aureus* and the multidrug resistance protein (Bmr) from *Bacillus subtilis*. Another family of transporters involved in drug efflux is the ABC (ATP binding cassette) transporters. These kinds of transporters contribute to resistance in eukaryotes and an interesting example of this family is P-glycoprotein (Griffith et al, 1992).

Few of the efflux protein genes of *M. tb*. that have been characterized are *drrABC* (ABC-transporter), *efpA* and *emrE* (resistance to ethidium bromide). Recently, it was shown that Rv1258c gene of *M.tb.* encodes a protein that is similar to Tap protein in *M. fortuitum* and confers weak resistance to tetracycline (Ainsa et al, 1998). In another study the *Mycobacterium bovis* P55 gene, located downstream from the gene that encodes the immunogenic lipoprotein P27, was characterized (Silva et al, 2001). The gene was identical to the open reading frame of the Rv1410c gene in the genome of *Mycobacterium tuberculosis* H37Rv, annotated as a probable drug efflux protein. Whether the above two proteins actually play any role in drug resistance in clinical scenario is yet to be determined.