CHAPTER 2: PREMISE
Mycobacterium tuberculosis, the causative organism of tuberculosis (TB) is regarded as the most successful human pathogen in history. It is believed to have infected around one third of the world’s population and costs 2 million lives annually (Dye et al., 1999; Yeu and Leung, 2008). Until 50 years ago, there were no medicines available to cure tuberculosis. The Bacillus Calmette-Guerin (BCG) vaccine introduced in 1922 is a live, attenuated strain of Mycobacterium bovis that revolutionized the area of preventive measures against TB. However, the true efficacy of BCG vaccine is still unknown and under controversy. Streptomycin was the first effective drug developed against TB in 1940s. Later more potent drugs like isoniazid and rifampicin were introduced for the treatment of TB. However, presently, strains that are resistant to a single drug have been documented in almost every country surveyed; also, strains of Mycobacterium tuberculosis resistant to all major anti-TB drugs have emerged. A particularly dangerous form of drug-resistant TB is multidrug-resistant TB (MDR-TB), which is defined as the disease caused by TB bacilli resistant to at least isoniazid and rifampicin, the two most powerful anti-TB drugs. While drug-resistant TB is generally treatable, it requires extensive chemotherapy (up to two years of treatment) that is often very expensive (often 100 times more expensive than the treatment of drug-susceptible TB), and is also more toxic to patients.

The above scenario prompted the scientific community to look for newer strategies to combat this disease. However, it will not be possible to tame this pathogen without a thorough knowledge of the mechanisms involved in its virulence and pathogenesis. In the last 15 years, the development of efficient vectors and allelic exchange systems has rendered the tubercle bacillus more amenable to molecular genetic analysis, and this has greatly facilitated studies of its biochemistry, physiology and pathogenicity (Jacobs Jr et al., 1991; Jacobs Jr and Bloom, 1994). The availability of the complete genome sequence and its detailed in silico analysis has provided us with a wealth of new information, knowledge and understanding of the biology of Mycobacterium tuberculosis (Cole et al., 1998). From the analysis of the genome, 58% of the genes could be attributed a function with some level of confidence, the product of a further 27% are conserved hypothetical (conserved within the mycobacterium species only); the remaining 15% are unrelated to any known genes or proteins. Approximately 40% of predicted open reading frames (ORFs) have no known function. Presently the most important task at hand is to interpret the genome
in context of the host-pathogen interaction, and then translate the resulting knowledge into a new generation of drugs and vaccines in the fight against this ancient disease. One key motive of this ongoing research is to understand the host-pathogen protein-protein interactions. It is believed that an understanding of the protein-protein interaction network would help us fight this bug more effectively.

A key feature of protein-protein interactions is their diversity. Proteins represent the most complex biomolecules found in nature and owe this diversity to their amino acid side chains. This structural diversity allows them to get involved in multitude of interactions with other biomolecules and hence participate in essential life processes. Most drugs that are currently being designed and developed target the important binding site of a protein. These drugs can either act as competitive inhibitors or bind to an unrelated area of the protein resulting in allosteric inhibition of the interaction. Also, since the binding of such peptide binders is not necessarily dependent on the active site of the target protein, the problem of developing resistance is rather less prevalent using this method.

The next task at hand is to choose crucial protein-protein interactions. Rubin and coworkers have carried elegant work in the area of identification of essential *Mycobacterium tuberculosis* proteins (Sassetti et al., 2001). Using a novel transposon based methodology, called transposon site hybridization (TraSH), they have identified set of proteins that are essential for *Mycobacterium tuberculosis*, ranging from general growth of the pathogen to its growth inside the macrophage (Sassetti and Rubin, 2003; Sassetti et al., 2003). These studies have identified proteins that are required for growth and virulence of mycobacteria and thus, have provided us with candidate proteins that could be targeted for drug development. Small molecules could be designed in a way so as to either bind and inhibit the essential function of these proteins or disrupt the crucial host-pathogen interactions in which these proteins are involved.

Protein-protein interactions are known to play a crucial role in most biological processes and thus represent an attractive target for drug development. These interactions can be targeted by small molecule inhibitors and by peptides, which represent an alternative to conventional drugs. Peptides are extremely potent, show high specificity, and display no or minimal toxicological problems. Also, they do not
accumulate in organs or suffer from drug-drug interactions unlike conventional drugs (Marx, 2005). Such peptide-based drugs can be used as therapeutic agents, or as a starting point for developing peptidomimetics and small molecular weight inhibitors (Marx, 2005).

Many antimicrobial peptides found in nature have been shown to be rich in one or more particular amino acids, like, cysteine, proline-arginine, proline, histidine, phenylalanine, glycine etc. (Diamond et al., 1991; Shi et al., 1994; Harwig et al., 1995; Lee et al., 1997a; Lee et al., 1997b; Lorenzini et al., 2003). Miyakawa et al. have shown the antimicrobial activity of human neutrophil peptide 1 (HNP-1) on \textit{in vitro} growth of \textit{Mycobacterium tuberculosis} (Miyakawa et al., 1996). Later, HNP-1 was also shown to clear \textit{Mycobacterium tuberculosis} from the lungs, liver and spleen of the infected mice (Sharma et al., 2001). Linde et al. have shown the inhibitory effect of PR-39, a proline-arginine-rich peptide, on \textit{in vitro} growth of susceptible and multidrug-resistant \textit{Mycobacterium tuberculosis} (Linde et al., 2001). It has been shown that antimicrobial peptides, in addition to having membrane permeabilizing effect, could also bind and inhibit various intracellular targets (Andreu and Rivas, 1998; Otvos Jr, et al., 2000). This suggests that peptide-mediated cell death may occur as a result of several independent or cooperative mechanisms of action; the latter phenomenon has been referred to as a "multi-hit process" (Zhang et al., 2000).

Our lab has developed a novel method of "codon-shuffling" that may be used to design protein/peptide-based binders against various target proteins of \textit{Mycobacterium tuberculosis}, which could then be use to disrupt such interactions (Chopra and Ranganathan, 2006). The advantage of "codon-shuffling" method is that dicodons can be assembled in a skewed fashion to obtain protein/peptide libraries rich in one or more particular amino acids. Such libraries then can be screened for desired antimicrobial activity \textit{in vivo}. Indeed, in a proof-of-concept experiment, we were able to show the ability of one such \textit{de novo} designed protein inhibitor to adversely affect the growth of model organism, \textit{Mycobacterium smegmatis} (Rao et al., 2007).
1.1 Aims and Objectives

The aim of present study is to design and discover protein/peptide-based inhibitors against selected target proteins of *Mycobacterium tuberculosis*. The broad objective of this project is to understand the host-pathogen protein-protein interaction and design *de novo* protein binders to disrupt these interactions. However, considering the fact that the mycobacterium cell wall is nearly impregnable, we decided to choose culture filtrate proteins of mycobacteria that are known to essential for its survival and virulence. We employed bacterial two-hybrid system to screen interacting partners between human lung cDNA library and rationally picked mycobacterial culture filtrate target proteins. The selected mycobacterial target proteins, Esat6 and Cfp10 are two key proteins co-secreted by Esx-1 secretion system encoded by genes of RD1 locus that play an important role in virulence and pathogenesis of mycobacteria. The choice of screening lung libraries was obvious since lungs are the primary and most prevalent site of *Mycobacterium tuberculosis* infection (Backlock, 1947). The objectives of this study can broadly be categorized as follows:

1. Construction of *de novo* protein libraries from shuffled dicodons
2. Bacterial two-hybrid studies to discover *de novo* protein/peptide binders against Esat6
3. Bacterial two-hybrid studies to discover interacting partners of Esat6 from human lung cDNA library
4. Purification and *in vitro* pull-down assay of Esat6-Cfp10 and Esat6-HcI1
5. Effect of HcI1 on intracellular survival of *Mycobacterium tuberculosis*.
6. Expression of hcl1 in mycobacteria and its effect on *in vitro* growth
7. Effect of HcI1 on cell shape and morphology of *Mycobacterium tuberculosis*
8. Microarray analysis to study the effect of HcI1 on global transcription in *Mycobacterium tuberculosis*