Chapter 7

SUMMARY
AND
CONCLUSIONS
Summary and Conclusions

*M. tuberculosis* causes approximately 8.9 million new cases of tuberculosis (TB) and 1.7 million deaths each year (Korenromp *et al.*, 2009). With its uniquely lipid rich cell wall, a generation time of close to 20 hours, an arsenal of almost 4000 genes and its ability to spread through aerosol infection, *M. tuberculosis* (Mtb) has been perhaps one of the most devastating pathogens in human history. Mtb is a facultative intracellular pathogen with the human alveolar macrophage as its most preferred host cell. Although cellular processes regulating the binding of Mtb to macrophages and its subsequent endocytosis by macrophages have been characterized in several studies, precise information on host factors putatively implicated in regulating a stable infection is lacking (Monack *et al.*, 2004). In order to identify the distinct cellular responses elicited by the Mtb, we adopted a global approach where time dependent changes in the global proteome and transcriptome profiles were generated after infecting human monocytic cell line THP-1 with H37Rv or H37Ra strains.

Combined proteomic approaches (MALDI and ESI MS/MS) and relative quantitation by the iTRAQ labeling chemistry- capable of labeling the N-terminal of digested peptides and offering a greater tolerance over other labeling methods at protein levels - identify 947 and 907 proteins from H37Rv and H37Ra-infected host cells, respectively. Finally, relative quantitation revealed 457 proteins were differentially regulated and common between H37Rv and H37Ra infected host cells. Comparing these 457 *M. tb* genes that were differentially regulated in human macrophage cell line with data available from published literature (Nau *et al.*, 2002; Rago *et al.*, 2001; and Greenwell-Wild *et al.*, 2002) we observed not only considerable overlap, but more importantly also unique proteins with distinct immuno regulatory patterns. For example, we found that Glycolytic enzymes from host cells showed distinct and opposite expression patterns upon infection with virulent H37Rv and avirulent H37Ra strains. Apart from these, we identified proteins having a role in cytoskeletal rearrangement and reorganization including Calpain, Cofilin, Meosin, Calgizarrin. These proteins are temporally regulated for cytoskeletal remodeling, and promoting cell migration. Proteins identified were also known to play an important role in maintaining redox regulation, lysosome acidification, protein synthesis, cell proliferation, apoptosis,
Summary & Conclusions

and homeostasis against reactive oxygen species (ROS). Those were, MnSOD, CORO1A, Elongation factor 1A, Cathepsin D, H⁺ ATPase, Calmodulin etc. Comparison of virulent and avirulent strains of *M. tuberculosis* induced gene expression in parallel macrophage cultures has provided important insight into virulence-related genes, although, in addition to the host factors, environmental factors and bacterial genotype and phenotype also influence the outcome of infection and evolution of disease.

In order to analyze the dynamic differences in the cellular responses of commonly identified 457 proteins, as they showed divergent pattern of expression profile, a correlation analysis was carried out between the proteome data from H37Rv- and H37Ra-infected cells. This analysis revealed that 128 proteins showed similar time dependent kinetics (correlation coefficient > 0.5) while 130 identified proteins showed contrasting patterns (correlation coefficient < -0.5). The former set therefore constituted products implicated in generic host mechanisms while the latter one represented those from the specific response machinery.

The application of transcriptional profiling in this study of the host response to pathogens has revealed several general themes. Hierarchical cluster analysis revealed differences in gene expression between H37Rv- and H37Ra-infected THP-1 cells. Cluster analysis deciphered that six (6) and three (3) groups of genes clusters were involved in H37Rv- and H37Ra-infected cells respectively. This observation supports the well-known strategy of virulent mycobacteria to co-opt host physiological processes and establishing of a dynamic equilibrium between the host and pathogen machinery. Cells responded with a broadly common transcriptional programme, components of which were preferentially differentially regulated and pathogens can interfere with these responses, which can enhance their virulence in the host.

The proteins we identified in both H37Rv and H37Ra infected cells provide a unique platform to look at the global events interlinked with exclusive intricate pathway distinguishing a virulent and avirulent infection. In order to analyze the dynamic differences and to identify the intracellular network in the cellular responses of commonly identified 457 proteins, Protein-Protein Interaction network was built that represented known network of 200 nodes out of 457 commonly identified molecules along with their interacting partners. However, comparative analysis of microarray and
proteomic data revealed relatively low levels of agreement between transcript and protein level. Therefore, we looked at the system level perturbations at host cellular network. Integration of these two data sets (proteomic as well as transcriptomic) provided snapshots of the time dependent changes at the cellular network level upon infection with H37Rv and H37Ra in host cells; and we concluded that modulation of host cell proteome itself is distinct and unique in terms of their specific host-pathogen interaction.

The host response is likely to represent the concerted actions of tens, perhaps hundreds, of transcriptional regulators. Clusters of genes that are regulated in a cell-type or pathogen-specific manner are likely to represent specific gene expression programmes, which are regulated by a unique combination of transcription factors (TFs). Thus assuming the transcription control as the major determinant for the observed variations, we identified cumulative set of TFs that could be together regulating transcriptional control of the entire H37Ra- or H37Rv-induced data sets. In total, 173 TFs were identified in this exercise; more specifically we found distinctly 105 and 73 TFs involved in H37Ra- and H37Rv- modulated host responses, respectively.

In conclusion, we have shown that iTRAQ protein labeling reagents can be employed to successfully identify proteins in which expression is potentially modified. The iTRAQ methodology produced high quality, reproducible data regarding relative expression levels in the time-course experiments. Comparison of the iTRAQ data with cDNA microarray data showed low degree of similarity; however, combined analysis provided additional insights in to the metabolic window underlying complex biological systems. This window presumably represents a phase in which the pathogen is either establishing or has established a sustainable equilibrium with the host intracellular milieu. Several host factors whose differential regulation significantly influenced the load of the intracellular pathogen could be consequently identified. Importantly, these proteins could be incorporated into a functional association network in which the component modules together defined the entire range of cellular processes that participate in the interplay between the host cell and the pathogen. Thus, in summary, this work has indicated that global proteome profiling of host molecules involved during the stabilization of M. tb infection could provide insight into unique and shared inducible host factors involved in the differential virulence between the virulent H37Rv and avirulent H37Ra strains of M. tuberculosis.