

# Chapter 4

## *Discussion*

Tuberculosis remains one of the major causes of death from a single infectious agent worldwide. Of great concern for tuberculosis control is the emergence of drug resistance since there is no cure for some multidrug-resistant strains of *M. tuberculosis*, and there is concern that they may spread around the world, stressing the need for additional control measures such as new diagnostics and better drugs for treatment

A variety of mechanisms have been suggested that contribute to the survival of *M. tuberculosis* inside the host, including inhibition of phagosome-lysosome fusion (Nossal *et al.*, 1975), inhibition of the acidification of phagosomes (Sturgil *et al.*, 1994), resistance to killing by reactive oxygen intermediates (Nathan *et al.*, 1983) and reactive nitrogen intermediates (Chan *et al.*, 1992 ; MacMicking *et al.*, 1997) and modification of the lipid composition of the mycobacterial cell membrane, there by altering its capacity to interact with immune or inflammatory cells. Little progress has been made towards the identification of the mycobacterial products that promoted intracellular survival.

The study of interaction between pathogenic mycobacteria and components of the host immune system at the molecular level could provide basic understanding of mycobacterial immunity and pathogenesis. This in turn might facilitate the development of reagents useful for diagnosis and prophylaxis of tuberculosis caused by *M. tuberculosis*. (Bull WHO, 1983)

Several techniques have been developed to facilitate identification of bacterial genes expressed specifically during infection. These techniques include in vivo induced expression technology (IVET; Mahan *et al.*, 1993), signature tagged mutagenesis (STM; Hensel *et al.*, 1993), differential fluorescent induction (DFI; Valdivia and Falkow, 1997) transcriptional and proteome profiles (Ramachandran *et al.*, 2004)

Several immunogenic screening has been used from time to time using sera from patients infected with pathogen of interest to probe protein expression library to find out important antigens (Shalka *et al.*, 1978; Anderson *et al.*, 1979). A modification of earlier immunogenic technique is in vivo induced antigen technology (IVIAT). This technique has been used to identify those gene that specifically induced in vivo during human infection (Handfield *et al.*, 2002). The IVIAT technique has been successfully used in our laboratory to identify those genes, which are specifically induced in vivo (Dev *et al.*, 2002). Sera from

several infected patients were collected, pooled and absorbed with in vitro grown *M. tuberculosis* H37 Rv. This absorption step removed antibodies that bind antigens expressed during in vitro growth condition. Rest of the antibody seems to be from the antigens which are specifically induced during in vivo growth. A genomic library expressing proteins of *M. tuberculosis* was created and screened with the absorbed patient sera. The reactive bacteriophage was isolated and sequenced. Rv3878 is one of the ORFs which is located in the RD1 region of *M. tuberculosis* genome (Cole *et al.*, 1998)

Full length ORF of Rv3878 consists of 840 bp and codes for a protein of molecular weight 27 kDa. When recombinant protein Rv 3878 was run on SDS-PAGE, it moved as a single band, of 45 kDa in molecular weight. This anomalous behaviour on SDS-PAGE was also reported for several other proteins of *mycobacteria* as well as in different organisms.(Dougelet *et al.*, 2003). Generally proteins rich in amino acids like alanine, glycine, proline and glutamic acid, assume nonglobular elongated structure and become more acidic (Sundaram *et al.*, 2004). Such acidic proteins display lesser affinity to SDS and therefore show delayed mobility on SDS -PAGE. Rv3878 is also acidic protein. Its theoretical pI is 4.11.

Secondary structure of proteins can be well characterized using CD spectrum. In aqueous buffer (20 mM phosphate buffer at pH 7.0) Rv3878 protein showed a strong negative band at 200 nm and positive band at 212 nm, indicating a predominantly random coil conformation. However, in TFE/water mixture, the CD spectrum showed maxima at 193 nm and double minima at 208 and 222 nm, which are indicative of a highly  $\alpha$  helical conformation. Cosolvents such as alcohols are frequently employed to stabilize the equilibrium intermediates (Bhakuni *et al.*, 1998). One of the alcohols, trifluoroethanol (TFE), stabilized  $\alpha$  helical structure in peptides/proteins (Lehrmann *et al.*, 2000). TFE and other alcohols generally weakens the hydrophobic interactions of proteins by creating the hydrophobic environment (Thomas, 1993). With the increasing concentration of TFE (10 to 90 %), signals at 192, 208 and 222 nm were intensified which indicates that the helicity of the Rv3878 protein increased in more hydrophobic environment. Peptides and proteins having amphipathic helix show random coil in aqueous solution and more ordered  $\alpha$  helix in TFE/water mixture. (Lehrmann *et al.*, 2000). Rv3878 protein showed random coil in aqueous buffer but in TFE/water buffer showed helix formation, which increased with the

increasing concentration of TFE in buffer. This phenomenon indicated that this protein might be containing an amphipathic  $\alpha$  helix. Bioinformatic analysis of this protein with protean programme of DNA star software revealed that it has two large stretches of amphipathic  $\alpha$  helix. Amphipathic helix containing proteins are generally found on membrane surface protein (Sundaram *et al.*, 2004).

Rv 3878 protein also showed minima at 208 and 222 nm and maxima at 192 nm in presence of SDS. The helicity of the protein increased up to 3 mM SDS after that it decreased and remained constant up to 30 mM concentration of SDS. SDS is an amphipathic molecule, it interacts with the peptide/protein in a dose dependent manner in buffer. At low concentrations, SDS exists as a monomer, whereas at higher concentrations micelles or vesicles are usually formed. The main driving force required for the formation of micelles or vesicles of SDS is the reduction of hydrophobic segment-water contacts. The binding force in a SDS-protein complex may be both electrostatic and hydrophobic in nature. The interactions between SDS and protein are primarily of hydrophobic character; however at low concentration ( $\ll$ CMC) electrostatic interactions are important. The binding of SDS, especially at higher concentrations often results in considerable structural deformation/denaturation of the protein (Halfman, 1986). However the resulting conformation of SDS denatured protein is still ordered and not random coiled (Lapenje, 1978). The property of protein showing native like secondary structure in SDS with no tertiary structure is due to the exposure of side chains belonging to the hydrophobic core in the hydrophobic environment. Any protein containing amphipathic helix, it is generally expected that increase in the concentration of SDS would increase helicity. There are lots of evidences suggesting that generally amphipathic helix containing protein forms helix whereas non amphipathic proteins do not. (Lark, 1989; Pasta, 1990) In our case we have shown that protein shows increased helicity at increasing concentration of SDS but after CMC level of SDS, it becomes little unstructured and fixed up to 30 mM. Therefore Rv3878 protein appears to gain structure at low concentrations of SDS and denatured at higher concentrations of SDS. This secondary structure formation in presence of SDS is referred to as molten globular structure (Ptitsyn, 1993). Molten globular structure is the intermediate structure between the native and denatured states, and is defined as a compact conformation with a comparable amount of native like secondary structure but with a high enhancement in

the intermolecular motion, i.e., largely disordered tertiary structure ( Ptitsyn, 1995 ). Molten globular structure has been suggested to play a part in the transmembrane trafficking, membrane insertion, protein transport and chaperon assisted refolding (Reshma *et al.*, 2003)

When tyrosine fluorescent measurements were done in presence of TFE and SDS, Rv 3878 protein showed increase in the fluorescent intensity compared to phosphate buffer. TFE and SDS reduce the hydrophobic interactions of the protein. In hydrophobic environment created by the TFE and SDS tyrosine molecule, which were buried inside the hydrophobic patches in water become more exposed. Outer membrane protein OmpA was generally unstructured when present in aqueous buffer but when it was incubated with lipid vesicles or detergent it became folded. When fluorescent spectra were taken in SDS, OmpA protein showed intense fluorescence compared to in unfolded state. Tryptophan molecules were buried inside the hydrophobic environment in aqueous buffer. But when it was incubated in SDS or in lipid vesicles it became more exposed and showed more fluorescent intensity. (Surrey, 1992)

Chemical denaturation showed that protein becomes denatured at very low concentration of GuHCl. So it is assumed that hydrophobic interaction is predominating to stabilize the secondary structure of protein. Proteins were incubated at different pH and its CD spectra were recorded. There was no secondary structural change observed. CD spectrum was also recorded at different pH in presence of 4 mM SDS. In this case there was no spectral change. Protein maintained its secondary structure. Hence pH did not cause any change in the secondary structure of protein. Thermal denaturation in presence of 4 mM SDS and 20% TFE, caused little structure loss. So it is a heat stable protein. Proteins containing amino acids like alanine, proline, glutamic acids are generally heat stable (Sundaram *et al.*, 2004).

Rv3878 protein was cross linked with glutaraldehyde treatment and when run on SDS –PAGE. It was found that protein shifted its position and moved as a large molecular weight protein. It seems that Rv3878 protein may be present in a oligomeric state or in aggregate form.

Spatial organization of proteins within bacterial cells is fundamental to many cellular processes, including DNA replication, chromosome segregation, protein secretion,

chemotaxis, adhesion, motility, cell division, cell shape and virulence. A number of proteins are preferentially localized to one or both poles of the cell. These include *Shigella* spp. (Goldberg, 1993) and *Listeria monocytogenes* actin assembly proteins (Kocks *et al.*, 1993), *B. subtilis* cell-cycle proteins (Marston *et al.*, 1999), the conserved signal recognition particle-associated protein YidC (Scotti *et al.*, 2000) as well as components of the *Aquaspirillum* and *C. crescentus* chemotaxis apparatuses (Jenal *et al.*, 1996), *Vibrio cholerae* type II secretion apparatus (Maria *et al.*, 2001), the *Agrobacterium tumefaciens* type IV DNA transport apparatus (Paul *et al.*, 2005), *C. crescentus* cell-cycle regulatory signal transduction complex (Alley, 1992).

In this study we have made recombinant Rv3878 protein and antibody raised against Rv3878. Western blotting was done with the different subcellular fractions of *M. tuberculosis*. It was found that this protein was localized in the cytosol and cell membrane, as also reported by (Marie, 2003). This protein might be synthesized inside the cell and then transported towards the cell membrane for specific purpose.

Fluorescent antibody labeling was done with antibody raised against Rv3878 in wild type *M. tuberculosis*. Fluorescent intensity was confined at the cell pole and some fluorescent was seen through out the cell membrane under fluorescent microscope. So this observation also strengthens that Rv3878 protein was localized at the pole.

Polar localization was also confirmed by exploiting the reporter gene *gfp*. Rv3878 gene was fused to N terminal of the *gfp* gene and its translational fusion was made inside the mycobacterial cell. It was found that Rv3878 protein was synthesized inside the cell in fusion with the GFP reporter protein (confirmed by western blotting using antibody against GFP). Fusion products were specifically localized at the pole. In some cells it was localized at the one pole and in some at both poles as observed by the fluorescence microscopy.

Is this protein ubiquitous in distribution among mycobacteria? Southern blotting was done within the different mycobacterial genomic DNA and probed with the rv3878. A single band was picked up in the *M. tuberculosis* H37 Rv but absent in other species. It would be assumed that this gene was unique in distribution. PCR was done with the primers for full length gene of *M. tuberculosis* Rv3878. A single band was observed in *M. smegmatis*, *M. aurum*, *M. phlei*, *M. fortuitum*, of the same size as in *M. tuberculosis* and a single band of

smaller size was observed in *M. avium* and *M. simiae*. This suggested that rv3878 gene was found in several other mycobacterial species irrespective of its pathogenesis. (We have however not sequenced the amplified PCR product). However western blotting was done with different mycobacterial cell lysates with the antibody against Rv 3878. It was found that this protein was picked up in *M. smegmatis*, *M. bovis*, *M. aurum*, *M. avium*, *M. asiaticum*, *M. phlei* of the same molecular weight as found in *M. tuberculosis*.

We attempted to understand the mechanism of localization of Rv3878 protein towards the poles of mycobacterial cell. Fusion constructs of *gfp* with N-terminal, mid and C-terminal domains of Rv3878 were made. It clearly demonstrated that specificity in mid to C-terminals of the gene was involved in such process. Bioinformatic analysis does not reveal any specific reasons to explain. However, N-terminal region was found to contain one hydrophobic domain whereas C-terminal region contained two hydrophobic domains. We find very difficult to correlate this difference to polar transport of proteins.

Is this protein present in the outer surface or inner surface of bacteria ? This was investigated by the osmotic shock experiment. It has been reported that osmotic shock caused release of proteins which are present on the outer surface of the membrane. In our experiment recombinant bacteria expressing Rv3878 and GFP fusion protein were given osmotic shock by high sucrose concentration, which resulted in complete loss of fluorescence. After osmotic shock most of the bacteria in the population lost its fluorescence. Therefore it is concluded that Rv3878 protein is synthesized in the cytoplasm and localized at the outer surface of the bacterial cell membrane.

Outer membrane localization was also confirmed by antibody labeling. Antibodies are not able to cross the cell membrane but it can recognize the outer surface antigens. So when cells were incubated with antibodies, the latter can bind with the outer membrane antigens. In our experiment different mycobacterial species were incubated with the primary antibody raised against Rv3878 protein followed by secondary antibody labeled with FITC. Cells were observed under fluorescence microscope. Fluorescent labeling of cells was observed. Hence it was conclusive that Rv3878 protein was present on the outer surface of the mycobacterial cells as observed with *M. tuberculosis*, *M. bovis*, *M. smegmatis*, *M. aurum*, and *M. avium*.

Rv3878 protein was found to be co-localized with DNA. When RNase treated cells were stained with nucleic acid binding dye it was found that GFP fusion protein and DNA were colocalized at the pole as well as in the cytosol of cells.

Vast majority of pathogenic mycobacterial strains are slow growers, and intracellular capable of surviving inside eukaryotic cells that can be quiescent for a long period of time, whereas the fast growers are mostly nonpathogenic. How mycobacteria co-ordinate their growth rates is an important question which has not been elucidated. To understand this issue efforts have been given on DNA-binding proteins because it is directly concerned with gene expression, and consequently might control growth rate. A DNA binding protein designated as mycobacterial DNA binding protein 1 (MDP1) was studied in *M. bovis BCG*. It has molecular weight of 28 kDa and estimated by SDS-PAGE to be constituting 8-10% of the total DNA-binding protein. MDP1 was conserved in all mycobacterial strains so far examined including BCG, *M. tuberculosis*, *M. leprae*, *M. smegmatis*, *M. intracellulare*, *M. kansasii*, *M. avium*, *M. haemophilum*, *M. malmoense*, *M. scrofulaceum*, *M. fortuitum*, *M. gastri*, *M. xenopi*, *M. szulgai* and *M. marinum*.( Sohkichi, 2002). Extracellular MDP1 acts as adhesin by binding to GAGs and mediates mycobacterial adherence to A549 human lungs epithelial cells through hyaluronic acid (HA). Adherence of bavillus Calmette-Guerin (BCG) and *M. tuberculosis* to A549 cells was inhibited by addition of HA, DNA, and anti-MDP1 antibody, showing that MDP1 participates in the interaction between mycobacteria-alveolar epithelial cells.

Rv 3878 is annotated as a hypothetical protein without any known function and show very little homology to any known functional protein. When Rv3878 was over expressed in *M. smegmatis*, physiological and morphological changes were observed. It was found that colony morphology of *M. smegmatis* was altered. Colonies were irregular in shape where as colonies from bacteria containing vector alone had smooth margin. It has been reported that due to the overproduction of RD1 region proteins in *M. bovis BCG* change in colony morphology was observed (Pym *et al.*, 2002). Overproduction of protein Rv3878 was confirmed by western blotting. When recombinant cells expressing Rv3878 were grown in liquid medium clumping of cells was more as compared to wild type cells and those containing vector alone. The genetic factors responsible for clumping in *M. smegmatis* are unknown. Recently it been shown that mutation in *impA*, one of the genes involved in the

phospholipids biosynthesis pathway of *M. smegmatis*, caused clumping of cells (Paris *et al.*, 1997). When individual colony was grown on NAT plate recombinant strain expressing recombinant protein were moist and slimy compared to the wild type colony, which might be due to accumulation of some metabolites on the cell surface. It is possible that change in phospholipid biosynthesis might cause clumping of the cells.

Several studies have showed that clumping in the cells occurred due to the change in the hydrophobicity of the cell surface (Miyamoto, 2004). We measured hydrophobicity of the Rv 3878 producing cells with the hydrophobic indicator dye Congo red and xylene adherence assay. Congo red binding assay showed that Rv3878 overproducing cells were more hydrophobic compared to vector alone. It was demonstrated that due to the increased hydrophobicity of the membrane of *M. smegmatis* more dye was bound to the cell surface (Cangelosi *et al.*, 1999). Increased hydrophobicity due to the overproduction of this protein was also confirmed by xylene binding assay. Over producing cells had more affinity to xylene than the cells containing vector alone.

Increased production of Rv3878 protein in *M. smegmatis* caused irregular colony morphology. Earlier studies showed that smooth or rough colonies and aggregate formation was due to the change in the lipid profile of the cells and GPL deficient mutant of *M. smegmatis* become rough and more aggregate compared to the wild type ( Etienne *et al.*, 2002 ). Total lipid and GPL were extracted and compared by TLC. An additional band was seen on TLC plate in both total lipid and GPL profiles of R3878 producing cells which was absent in case of cells with vector alone.

Overproduction of Rv3878 protein caused filamentous structure of cells and cells becomes more filamentous compare to cells containing vector alone. Some of the filamentous cells contained buds and were branched like outgrowth on either ends of the filaments. Some cells contained bulbous structure at one end or both ends of the cells. This type of structural changes was also observed in *M. smegmatis* due to the overproduction of protein FtsZ (Dziadek, 2002).

Protein -protein interaction is a simple and powerful technique to identify interacting partners of any known or unknown proteins. This technique also demonstrates the interaction of two known interacting partners. We have used this technique to explore and

identify proteins interacting with Rv3878. A random expression library of *M. tuberculosis* was created in one of the vectors of two hybrid system and Rv3878 was cloned in the other vector. Positive interacting partner was selected from the library. There were 15 individual clones which were characterized by DNA sequencing. These clones were showing different metabolic functions as mentioned in (Table I below). Some of these clones were involved in the lipid breakdown as lipase and esterase involved in breakdown and synthesis of cell membrane. One of the interacting partner codes for the ATPase. When these proteins were analysed through bioinformatics it was predicted that more than 60% proteins were localized in the cell membrane (Grady *et al.*, 2005). Interestingly our analysis shows that Rv 3878 protein is also a membrane protein.

Table I

No.	Rv no	Homology search	Localization prediction
1	1006	Hypothetical protein, function unknown	Not known
2	2275	Hypothetical protein, no homology found	Bacterial cytoplasm
3	1288	Annotated in tuberculist data bank as a hypothetical protein. It has LysM domain, found in variety of enzymes involved in bacterial cell wall degradation. This domain may have a general peptidoglycan binding function.[ May be involved in cell enveloped biogenesis, outer membrane	Bacterial membrane
4	3094	Conserved hypothetical protein, has hology with acyl-coa dehydrogenase ( ACAD). Both mitochondrial acyl-coa dehydrogenases ( ACAD) and peroxisomal acyl-coa oxidases (AXO) catalyze the alpha, beta dehydrogenation of the corresponding trans-enoyl- coa by FAD, which becomes reduced. The ACAD family includes the eukaryotic beta-oxidation enzymes, short (SCAD),	Bacterial membrane

		medium (MCAD), long (LCAD) and very-long (VLCAD) chain acyl-coa dehydrogenase.	
5	0896	Probable citrate synthase I GLTA2	Bacterial cytoplasm
6	1515	Annotated as conserved hypothetical protein. It has functional similarity with UBiG, 2-polyprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinol methylase	Bacterial cytoplasm
7	3264	D-alpha-D-mannose-1-phosphate guanyltransferase ManB, involved in novel pathways for biosynthesis of nucleotide- activated glycerol-manno-heptose precursors of bacterial glycoproteins and cell surface polysaccharides	Bacterial membrane
8	3095	ARSR subfamily of helix-turn helix bacterial transcription regulatory proteins (winged topology), includes several proteins that appear to dissociate from DNA in the presence of metal ions.	Bacterial cytoplasm
9	1159	Conserved transmembrane protein.	Bacterial membrane
10	1493	Possible methylmalonyl- coa mutase large subunit MutB	Bacterial membrane
11	1366	Annotated as conserved hypothetical protein. Homology found with Rel_Spot region of protein Rel/ Spot proteins. This region of unknown function is found in RelA and SpoT of <i>E. coli</i> , and their homologues in plants and in other eubacteria. RelA is a guanosine 3',5'- bis pyrophosphate (ppGpp) synthetase ( EC:2.7.6.5 ) while Spot is thought to be a bifunctional enzyme catalyzing both ppGpp synthesis and degradation (ppGpp 3'-	Not known

		pyrophosphohydrolase ( EC: 3.1.7.2).	
12	1383	Probable carbamoyl phosphate synthase glutamine chain	Bacterial cytoplasm
13	1087	PE-PGRS, cell membrane protein , may be involved in antigenic variation.	Bacterial membrane
14	0103	Probable cation transporter P-TYPE ATPase B CTPB,transport or detoxify heavy metals, for example , the heavy metal ATPase and copper chaperones.	Bacterial membrane
15	3203	Possible lipase lip V, these lipases are esterases that can hydrolyze long-chain acyl-triglycerides into di-and monoglycerides, glycerol, and free fatty acids at a water/lipid interface. The active site of a lipase contains a catalytic triad consisting of ser-his-ASP/glu, but unlike most serine proteases, the active site is buried inside the sturcture. A lid or flap covers the active site, making it inaccessible to solvent and substrate access to the active site.	Bacterial cytoplasm

Antisense approaches were taken to inhibit the expression of Rv3878. Sense and antisense constructs of Rv3878 were constructed and transformed in *M. tuberculosis*. Over production and knock out of Rv3878 in sense and antisense containing cells were confirmed by western blotting. When mice were challenged with sense and antisense strains and compared with wild type *M. tuberculosis*, there was no significant deference in virulence in mice model. This indicated this protein probably had no direct role in pathogenesis of disease. Earlier it was reported that clinical strain deleted in Rv3878 were not attenuated in virulence (Pym, 2003 ). Therefore it may be concluded that Rv3878 might have no role in virulence.

Understanding changes in gene expression is crucial to investigating the adaptive responses of mycobacteria to environmental stress and corresponding changes in the gene expression during infection. Transcriptional control of protein synthesis is a primary control strategy used in living cells. The rate of formation of transcripts is primarily determined by the frequency of initiation, which is directly related to promoter initiation. Both are major forms of gene regulation and results in altered rates of protein synthesis. Studies of promoter sequence and regulation are important for understanding global gene regulation. Expression of gene was monitored by cloning intergenic region of Rv3878 inframe with reporter gene *gfp* and expression was recorded by measuring the GFP expression level. GFP expression was maximum in acidic condition as compared to neutral pH and other stress conditions. Recently up regulation of *lipF* of *M. tuberculosis* was reported during in vitro acidic condition (Saviola *et al.*, 2003). *M. tuberculosis* faces acidic condition during infection inside the alveolar macrophage. The response of the Rv3878 promoter (P<sub>1500</sub>) was monitored in nutrient starvation condition. Expression of GFP was measured. Since mycobacteria reside in granulomas, they are likely to face stresses induced due to depletion of nutrients such as nitrogen, phosphorous and carbon. Indeed, the expression of GFP was regulated in respective starvation condition which appeared associated with phase of growth. When GFP expression driven by intergenic region was analyzed ex-vivo, there were no changes occurring in expression level.

When biophysical and bioinformatic analysis were combined, it appears that Rv3878 protein is an unusual protein. It is alanine rich, and contains in excess some unusual amino acids like proline, glycine and glutamic acid. Protein is acidic with theoretical pI of 4.11 and migrates slowly in SDS-PAGE. In this respect it resembles membrane binding proteins or MARKS family of proteins (Sundaram *et al.* , 2004). Due to the properties associated with Rv3878 protein, it is logical to conclude that this protein is a membrane bound protein. It has 37.4% alpha helix and 49.64% random coil. There were two amphipathic alpha helices. This amphipathic alpha helix is characteristic of membrane bound protein.