Discussion
Pathogenicity is multifactorial and requires any microorganism to fulfill several cardinal conditions. For microorganisms like mycobacteria that do not have a direct access to the host tissues, the bacilli must be able to: i) colonize the mucous surfaces, ii) enter the host cells, iii) multiply in the environment of host tissues, iv) resist or interfere with host defence mechanisms and, v) cause damage to the tissues of the host. The failure to carry out any of these points causes the microbe to lose virulence. That the cell envelope components are important for the pathogenicity of mycobacteria, at least for the first steps of the internalization of bacteria, is understandable from its privileged position at the inter phase between the bacilli and host cells. Knowledge of their roles in “signaling” events, in pathogenesis, and in the immune response is now emerging, sometimes piecemeal and sometimes in an organized fashion (Pieters, 2008). In this study, attempts were made to characterize mycobacterial cell wall associated Lppi, hypothetical probable lipoprotein, and MDP1, mycobacterial DNA binding protein. Attempts were also made to set a platform for studying the specificity/cross reactivity of mycobacterial capsular polysaccharides.

Gene lppi (Rv2046) from Mycobacteria tuberculosis (H37Rv), so far unexplored, was taken for its molecular characterization. The gene was cloned and the expressed protein was purified. Using polyclonal antisera, generated against this recombinant protein, the lppi was localized in different species of mycobacteria whereas for localisation in infected macrophage, mycobacteria expressing GFP-tagged protein were used. Attempts were also made to check its functional significance.

Lppi is a hypothetical, mycobacterial cell wall anchored probable lipoprotein, sharing no conserved domains outside the genus mycobacteria (Sutcliffe and Harrington, 2004). In this study, the gene was PCR amplified, cloned and expressed in E. coli with N-terminal histidine tag (Fig. 6). The amino acid sequence deduced from the full length gene sequence indicated that the protein has a molecular mass of around 22 kDa, whereas the observed mass in SDS-PAGE was 23.5kDa (Fig. 6, panel D). This discrepancy between the molecular mass could be due to post-translational modification owing to the acylation of mycobacterial lipoproteins (Sutcliffe and Harrington, 2004, Rezwan et al., 2007a). The induction was at a low level, even at different temperatures and in Rossetta strain (Fig. 7, panel A and B), and anti-histidine mAb did not recognize the N-terminal his-tag in the expressed protein (Fig. 7, panel C). The common feature of lipoproteins is the presence of a conserved consensus sequence called a lipobox ([LV] [ASTVI] [GAS] C) with a universally conserved cysteine at position +1; the lipobox directs processing of the prolipoprotein to form the mature acylated protein (Sankaran and Wu, 1995). Lipoprotein
biogenesis is dependent on the presence of specific type II signal peptide sequences (Braun and Wu, 1994, Tjalsma et al., 1999, Sutcliffe and Harrington, 2002). At positions 13–17 of the precursor protein of lppi is the sequence IAGeS (Figs. 8 and 9), which is similar to the consensus sequence for lipidation and cleavage by signal peptidase II based on predicted and known lipoproteins of Gram-negative and Gram-positive bacteria (Braun and Wu, 1994, Sutcliffe and Russell, 1995). The particular lipobox is 100% identical with that of lprG (Rv1411c) as envisaged by Dupont et al., 2005. The sequence upstream of the ‘C’ is the predicted signal sequence which is chopped off after transportation and complete acylation of the protein at ‘C’ in the membrane (Sutcliffe and Harrington, 2004, Rezwan et al., 2007a). This, coupled with other observations that acylated lppi with C-terminal his-tag was expressed, at slightly higher molecular size compared to N-terminal his-tagged protein (Fig. 11, panel A), and reacted with anti-his antibody (Fig. 11, panel B) unlike above observed non reactivity of anti-his antibody with the expressed N-terminal his tagged lppi (Fig. 7, panel C), indeed confirmed the chopping off the his-tag along with the signal sequence at N-terminus.

Both heterologous and homologous expression systems have been used for the cloning, expression and characterization of the mycobacterial lipoproteins. LpqH (Garbe et al., 1993), MBP83 (Vosloo et al., 1997), 22kDa lipoprotein (Dupont et al., 2005) and LprA (Pecora et al., 2006) had been cloned, expressed and purified in M. smegmatis as expression host, followed by immunological characterization. Besides, standard E. coli expression system has also been employed for the expression and purification of number of mycobacterial lipoproteins like 38 kDa lipoprotein of M. tuberculosis (Singh et al., 1992), MK35 lipoprotein of M. kansasii (Armoa et al., 1995), lppX of M. tuberculosis (Sulzenbacher et al., 2006), lpp34 of M. avium (Gioffre et al., 2006) and Rv0679c, 16.5kDa lipoprotein of M. tuberculosis/M. bovis (Matsuba et al., 2007). Secondly, acyl modification of bacterial lipoproteins and lipopeptides is thought to be important for their ability to signal through TLR2 (Noss et al., 2001, Lee et al., 2002, Giambartolomei et al., 2004), although peptide sequence can also affect TLR2 agonist activity (Yamashita et al., 2004, Buwitt-Beckmann et al., 2005) and a nonacylated (NA) TLR2 ligand has been reported (Yamashita et al., 2004). Investigations with M. leprae LpK variants revealed that both acyl residues and peptide sequences are required for elicitation of an immune response. Thus, another important aspect of expression of mycobacterial lipoproteins is their expression in acylated or non acylated forms.
Attempts were made to clone and purify both acylated and non-acylated forms of lppi (Acyl.lppi and NAcy.lppi), both in heterologous system (E. coli) and homologous (M. tuberculosis Ra) system. In E. coli, expression of Acyl.lppi with C-terminal his-tag could be observed (Fig. 11, panel A and B). However, Acyl.lppi appeared at 25.5 kDa in SDS-PAGE as compared to the predicted size of 22 kDa, characteristic of bacterial lipoproteins (Sutcliffe and Harrington, 2004, Rezwan et al., 2007a) due to the post translational modification (acylation) in lipoproteins. This reduced mobility in SDS-PAGE gels, characteristic of bacterial lipoproteins (Sutcliffe and Harrington, 2004), has been reported for many mycobacterial lipoproteins like MK35, a predicted lipoprotein from M. kansasii (Armoa et al., 1995), the 27 kDa lipoprotein of M. bovis (Bigi et al., 1997) and lpp34 lipoprotein of M. avium (Gioffre et al., 2006) expressed in E. coli.

Although in E. coli, Acyl.lppi could be seen in soluble fraction at different temperatures (Fig. 11, panel C) yet it could not be purified using Ni-NTA column (Fig. 11, panel D). This low affinity binding with the column might be due to either post translational modification of acylation or hydrophobic nature of the protein as reported by Maeda et al., 2002 for the purification of 33kDa lipoprotein, Lpk of M. leprae, from E. coli. The addition of 0.5 - 1% triton X-100 in binding/ lysis and wash buffers led to effective purification of Acyl.lppi (Fig. 12, panel A) as reported for other membrane lipoproteins (Mcbride et al., 2003, Sklar et al., 2007). Also, this observation confirmed the lipid modification of this recombinant lppi (Acyl.lppi), in the heterologous expression host E. coli, in line with other studies showing the expression of mycobacterial lipoproteins in E. coli (Singh et al., 1992, Armoa et al., 1995, Bigi et al., 1997, Maeda et al., 2002, Gioffre et al., 2006).

Bacterial lipoproteins are characterized by the presence of at least one N-terminal cysteine (after cleavage of their leader peptides) which is modified by fatty acylation and serve to anchor the respective proteins in the cell wall (Sutcliffe and Harrington, 2004, Rezwan et al., 2007a). The difference in the mobilities of Acyl.lppi in SDS-PAGE, run under non-reducing and reducing conditions, suggested presence of intra molecular disulphide linkages in Acyl.lppi, unopened in non reducing sample buffer making it move slightly ahead of reducing sample buffer (Fig 12, panel B).

With respect to the expression of non acylated form of lppi (NAcy.lppi), lacking signal sequence and lipobox, in E. coli, the same could not be achieved in initial attempts (Fig. 13, panel A and B). Even induction at different temperatures failed to induce NAcy.lppi (Fig. 13, panel C) as observed for many proteins (Gambhir, 2007, Kaur,
Nevertheless, incorporation of wobble changes at the 5’ end of the gene led to the expression and the purification of NAcyl(W)lppi from *E. coli* (Fig. 14) (Sanli et al., 2001, Sadaf et al., 2008). However, in line with earlier observations on the purification of non-acylated forms of MPB83 (Vosloo et al., 1997) and Rv0679c (Matsuba et al., 2007) of mycobacteria from *E. coli*, the purification of NAcyl.lppi from the cytosolic fractions did not require the presence of TrironX-100, thus confirming the absence of lipid moiety in non-acylated lppi as expected theoretically. SDS-PAGE of purified Acyl.lppi and NAcyl.(W)lppi proteins revealed, only minor difference in the mobilities of the two (Fig. 17, panel A) which might be attributed to the acylation whereas Vosloo et al., 1997 and Matsuba et al., 2007 documented no difference in the mobilities of acylated and non-acylated forms. Further purified acylated and non-acylated lppi were found to be antigenically similar with respect to the reactivity with rabbit anti-acylated lppi in a western blot (Fig. 17, panel B).

Lipoproteins like lprA, lprG are exclusively expressed in pathogenic species whereas others like lpqH, lpqX, lpqW, lpqM etc. have homologues in both pathogenic and non-pathogenic species which may differ either in gene length or amino acid compositions (Rezwan et al., 2007a, Nguyen et al., 2009). Although homologues of lppi gene have been reported in different mycobacterial species and vary in their sequence and length among themselves (Fig. 8), yet there is a lack of information in general on lppi protein expression during *in vitro* growth. Utilising the specific rabbit sera for lppi, expression of the lppi in the cell wall proteins of different virulent and nonvirulent mycobacterial species under both non-reducing (Fig. 18, panel A) and reducing conditions (Fig. 18, panel B) was checked as done by Bigi et al., 1997 for 27kDa lipoprotein of *M. bovis* in the cell extracts of different mycobacterial species. Under nonreducing conditions, in *M. microti* the diffused band was highlighted at 24-29kDa, *M. smegmatis* at 37 and 24kDa, *M. tuberculosis* H37Rv at 66, 45 and 24kDa, *M. tuberculosis* H37Ra at 24kDa and *M. bovis* at 66, 29 and 24kDa. Whereas, under reducing conditions a band at 25.5kDa, analogous to Acyl.lppi expressed in *E. coli*, in both *M. tuberculosis* complex members viz. *M. microti*, *M. tuberculosis* H37Rv and *M. bovis* and non virulent species *M. tuberculosis* H37Ra and *M. smegmatis* was observed. However, in *M. smegmatis*, another band at 37kDa was recognized. This recognition of multiple bands in the cell wall proteins might be due to: i) cross reactivity of the polysera, ii) higher molecular forms of the proteins, as reported for MBP70 (Fifis et al., 1991) and MPB83 (Wiker et al., 1996) in *M. bovis*, iii) due to post translational modification i.e. lipidation or glycosylation (Garbe et al., 1993, Vosloo et al., 1997, Pecora et al., 2006),
iv) covalent interactions of lppi with other cell wall proteins due to the presence of 7 cysteines in lppi, as demonstrated for Vir B7 lipoprotein of *A. tumefacienis* (Spudich et al., 1996) and 27kDa lipoprotein of mycobacteria expressed in *E. coli* (Bigi et al., 1997), v) pre/prolipoprotein form of the protein in the cell wall as documented by Maeda et al., 2002 for LpK of *M. leprae* and Hovav et al., 2003 for 27 kDa lipoprotein of mycobacteria. Due to the presence of the leader peptide (signal sequence), pre/prolipoproteins have a slightly (2–3 kDa) higher molecular mass than the mature lipoproteins. Also the lowest band recognized in case of reducing condition, almost 1.5kDa above than that of non reducing condition, complemented the earlier observations of mobilitity shift of Acy.lppi ran under reducing and nonreducing conditions (Fig. 18, panel B) and further confirming the presence of intra molecular disulphide bonds in native lppi.

Results (Fig. 18) also revealed that lppi is conserved and expressed in both virulent and nonvirulent mycobacterial species as reported for many lipoproteins of mycobacteria like lpqH, lppX, lpqW, lpqM etc (Rezwan et al., 2007a, Nguyen et al., 2009). Also, as documented by Vosloo et al., 1997 for lipoprotein MPB83 on *M. bovis* (BCG) cell surface, lppi was also found on the cell surface of both *M. tuberculosis* complex members viz. *M. microti*, *M. tuberculosis* H37Rv and *M. bovis* and non virulent species *M. tuberculosis* H37Ra and *M. smegmatis* (Fig. 19).

Homologous expression system has been used to study the native antigen especially for those exhibiting posttranslational modification of glycosylation as reported for lpqH, lprA, MBP83 etc. (Garbe et al., 1993, Vosloo et al., 1997, Pecora et al., 2006), since the heterologous system of *E. coli* lacked the machinery for the same. PCR amplification, cloning, expression of lppi in *M. tuberculosis* H37Ra (Fig. 16) revealed the presence of Acy.lppi in the pellet/membrane fraction whereas NAcy.lppi was detected in the soluble/cytosolic fraction which was what theoretically expected and supported by earlier reports in case of mycobacterial lipoproteins, MPB83 in *M. smegmatis* (Vosloo et al., 1997), lprA in *M. smegmatis* (Pecora et al., 2006), Rv0679c in *M. bovis* BCG (Matsuba et al., 2007), lpqM in *M. smegmatis* (Nguyen et al., 2009). Also, expressed Acy.lppi of *M. tuberculosis* Ra appeared at 25.5kDa (Fig. 16, panel C), having similar molecular size as that of *E. coli* origin Acy.lppi thus negating existence of glycosylation in lppi in mycobacteria and complemented the recognition of 25.5KDa band in the cell wall proteins of different mycobacterial species by utilizing rabbit anti-Acy.lppi antisera in western blotting (Fig. 18, panel B). Garbe et al., 1993, showed glycosylation of 19kDa lipoprotein since the protein from *E. coli* origin, in which there is no glycosylation, moved
at 4kDa lower molecular size in SDS-PAGE as compared to that from *M. smegmatis*. Glycosylation of lipoproteins in mycobacterium has been implicated in protecting it from proteolytic cleavage (Herrmann et al., 1996) and secretion (VanderVen et al., 2005). Interestingly, the mycobacterial NAcyl.lppi was found to be at much lower molecular weight as compared to NAcyl.(W)lppi protein of *E. coli* origin (Fig. 16, panel C and D) despite the fact that Acy.lppi of both the origins appeared at same molecular weight. Difference in the mol. wt. of acylated and non-acylated forms of lpqH, 19kDa lipoprotein expressed in *M. vaccae* was also reported by the Neyrolles et al., 2001. This difference might be attributed to the susceptibility of non-acylated forms to the proteolytic cleavage in mycobacteria, in the absence of signal sequence.

Attempts to purify *M. tuberculosis* H37Ra expressed recombinant Acy.lppi and NAcyl.lppi in sufficient quantity for immunological characterization were not successful. *E. coli* expressed few mycobacterial lipoproteins have also been shown to be immunologically identical to native mycobacterial antigens e.g. 38kDa lipoprotein of *M. tuberculosis* (Singh et al., 1992) and 33kDa lipoprotein from *M. leprae* (Maeda et al., 2002). In this study, *E. coli* expressed Acy.lppi, exhibited strong IgG1 response, with IgG1/IgG2a ratio of 10:1 (Fig. 20, panel A), indicative of good Th2 type immune response whereas in case of NAcyl.(W)lppi, IgG1/IgG2a ratio was drastically reduced to that of 3.7:1 in second booster sera at dilution of 1:10k (Fig.20, panel B). Lipid moiety of lipoproteins has been found to enhance antibody production, since the elimination of the lipid reduced the immune response of the host against the antigen (Akkoyunlu et al., 1997). Although in this study detailed immunological characterization of lppi was not done, yet IgG1/IgG2a ratio seen upon immunization with Acy.lppi was indicative of strong Th2 type of immune response, which warrants further studies. Rao et al., 2005 evidenced that increased expression of *M. tuberculosis* 19kDa lipoprotein obliterates the protective efficacy of BCG by polarizing host immune responses to the Th2 subtype. Unlike above observation, many other lipoproteins e.g. LprG, LprA, are presumed to be IL-12 inducers from the host cell and facilitate the development of Th1 cells (Vordemeier et al., 1991, Sieling et al., 1994, Pecora et al., 2006).

There is a paucity of information on functional properties of mycobacterial lipoproteins. Only few lipoproteins have been explored for their functional role. To quote a few, LprA, a TLR2 agonist, induces expression of TNF-α, IL-10 and IL-12 (Pecora et al., 2006). LprG is a novel TLR-2 ligand that inhibits human macrophage class II MHC antigen processing (Gehring et al., 2004). LpqH (19 kDa lipoprotein), a TLR-2 agonist is involved
in bactericidal response in macrophages (Thoma-Uszynski et al., 2001). LppX (Rv2945c) is involved in translocation of complex lipids, the phthiocerol dimycocerosates (DIM), to the outer membrane (Sulzenbacher et al., 2006). Recently, Nguyen et al., 2009, reported that LpqM, a mycobacterial lipoprotein-metalloproteinase, is required for conjugal DNA transfer in \textit{M. smegmatis}. Several putative lipoproteins, e.g. LppP (Rv2330c), LprO (Rv2290), LprK (Rv0173), LpqT (Rv1016c), LpqY (Rv1235), LpqZ (Rv1244), LprG (Rv1411) and LppX (Rv2945c), are required for optimal growth in vivo (Sassetti and Rubin, 2003, Bigi et al., 2004, Rengarajan et al., 2005). The pathogenesis of \textit{M. tuberculosis} is complex, and involves an elaborate interaction with the host (Pieters, 2008). Key factors, including the ability to survive in macrophages, the predilection for the lung, the formation of granulomas and long-term persistence, are poorly understood and difficult to model to carry out more focused experiments and \textit{in vitro} conditions are used that might mimic in vivo situations (Kendall et al., 2004). Thus, the intra-macrophage environment has been associated with low pH, low nutrients, nitrogen and oxygen stress and general stresses (Kendall et al., 2004).

The green fluorescent protein (GFP) has become a popular reporter system for use in both prokaryotes and eukaryotes. The study of GFP fusion protein targeting in eukaryotes has been used to identify proteins targeted to cellular locations such as the plasma membrane (Yokoe and Meyer, 1996), the endoplasmic reticulum (Miyawaki et al., 1997), and the phagosome (Maniak et al., 1995). In prokaryotes, the \textit{gfp} gene has been used primarily as a reporter of promoter activity by creating transcriptional fusions in a wide variety of bacterial species, including \textit{Brucella suis} (Kohler et al., 1999), \textit{Salmonella typhimurium} (Valdivia et al., 1996), and Mycobacterium species (Dhandayuthapani et al., 1995, Kremer et al., 1995, Valdivia et al., 1996). Indeed, the list of GFP transcriptional fusion vectors for use in mycobacteria has grown in the last 5 years, and several studies have demonstrated that GFP is extremely useful as a method for monitoring gene expression in mycobacteria, to identify \textit{in vivo} induced genes, such as the strategy termed differential fluorescence induction (DFI) (Valdivia and Falkow, 1996), including localization of the organism within the host macrophage (Dhandayuthapani et al., 1995, Barker et al., 1998, Teitelbaum et al., 1999). The analysis of GFP Fluorescence of the recombinant bacteria of interest by fluorescence-activated cell sorting (FACS) was applied to \textit{S. typhimurium} for the isolation of pH-regulated promoters and to identify promoters showing differential activity outside and within host cells (Valdivia and Falkow, 1996, Valdivia and Falkow, 1997). The observation that GFP is efficiently expressed in
mycobacteria and that fluorescent bacteria can be directly observed in infected macrophages (Dhandayuthapani et al., 1995, Kremer et al., 1995) suggested that such a strategy may be applicable to the study of gene induction by mycobacteria in host cells.

Promoter activity, when checked in the 1kb region upstream of lppi coding region in *M. tuberculosis* Ra harboring lppi promoter-**gfp** fusion construct, showed two fold higher peak than positive control of pSC301, harboring SOD promoter-**gfp** construct (Fig. 23, panel A). Even when the activity was checked under different *in vitro* stress conditions, mimicking intracellular environment of infected macrophages viz., nitrosative stress, oxidative stress, acidic environment, and starvation conditions, the lppi promoter showed higher expression as compared to that of the SOD promoter (Fig. 23, panel A, Table 4.1). When an overlay of the lppi promoter activity under different stress conditions over the activity at 37°C was plotted (Fig. 23, panel B), the promoter activity was found to be significantly inducible, almost by 5 fold (Table 4.1), under acidic condition at pH 4.0. Statistical significance of the experiment was established using analysis of variation (ANOVA) at a *P* value of <0.05. Complementing *in vitro* experimental results of induction of lppi promoter activity under acidic conditions, the confocal images of macrophages infected with *M. tuberculosis* Ra harboring lppi promoter-**gfp** transcriptional fusion construct at different post infection time period, till 96 hrs, showed strong lppi promoter activity (Fig. 24), indicating that lppi have some role in intracellular conditions, particularly when pathogen is exposed to acidic pH residing in phagolysosomes of macrophages. A similar system was used to identify promoters of *M. marinum* which showed elevated levels of expression within macrophages compared to the extracellular environment (Barker et al., 1998). Cowley and Gay, 2001 utilized **gfp** transcriptional fusion constructs and revealed that incubation at 42°C and exposure to low pH decreased the transcription levels of the promoter of pknH, a serine threonine kinase to a significant extent. Similarly, Cowley et al., 2002 reported that the *M. tuberculosis* ptpA promoter is induced in the stationary phase of *M. bovis*, and upon infection of human monocytes. Saviola et al., 2003, also documented that lipF and Rv0834c are transcriptionally upregulated by exposure to acidic growth media by similar approach of **gfp** reporter gene. Again, Pawaria et al., 2008, implemented the identical strategy and revealed that the glbN promoter activity increased substantially during stationary phase and nitrosative stresses unlike the glbO promoter activity which was significantly induced under oxidative stress however, both the promoters were found to be active in infected macrophages.
Mycobacterial glycolipids are exported from the immature phagosome during live infection (Xu et al., 1994, Beatty et al., 2000). Also in one of the study, evidence from confocal microscopy, electron microscopy, and subcellular fractionation suggested that the 19-kDa lipoprotein is indeed exported from the mycobacterial phagosome during the first hour after phagocytosis of live mycobacteria, dependent on the presence of the lipid tail on the protein (Neyrolles et al., 2001). This was not the case in IpP, as neither Acy.IpP nor NAcy.IpP were found to be trafficked inside macrophages and remained localized on the bacteria inside phagolysosome (Fig. 25, panel B and C).

Thus, in nut shell, cloning and expression of IpP (Rv 2046) in E. coli with both N- and C- terminal histidine tags revealed the existence of N-terminal signal sequence in IpP. The lipobox of IpP was found to be 100% identical to that of lprG of mycobacteria. IpP is localized not only on the cell surface but also present in SDS soluble cell wall proteins of different mycobacterial virulent and avirulent species. Significant difference in the molecular weight of Acy.IpP and NAcy.IpP, expressed in mycobacteria, was observed unlike E. coli. Contrary to what is reported for many lipoproteins, acylated IpP exhibited strong Th2 response. Further, significant promoter activity of IpP was found to be acid inducible at pH 4.0, thus suggesting its role in intracellular survival of mycobacteria.

The surface of M. tuberculosis is rich in carbohydrate-containing fractions and a polysaccharide-containing layer, possibly representing a capsule (Daffe and Draper, 1998, Daffe and Etinne, 1999). For most of the pathogens, including mycobacteria, different strains express capsular polysaccharides (CPS) of different structures, resulting in a number of different serotypes or serogroups (Jones, 2005, Tsang et al., 2005). Virulence and pathogenicity may be serotype or serogroup dependent, or there may be geographic differences in the clinically relevant serotypes (Racoosin et al. 1998, Mayer et al. 2002).

Thus, these carbohydrates on bacterial cell surface represent species/strain specific as well as cross reactive epitopes (Yongye et al., 2008). Analogous to a hapten-carrier system, predominant antigenic determinants of polysaccharides often consist of short oligosaccharides (1-5 sugar long) at the nonreducing end linked to a large nonantigenic polysaccharide backbone, which is important for the immunogenecity, like carriers for haptens (Kabat, 1976, Paul, 2008). With a long term goal to check the specificity/cross reactivity of these carbohydrate determinants on mycobacteria surface, collaborative efforts were initiated to make database with entries for CPS of pathogenic microorganism and respective antibodies (Figs. 26-28).
The main components of the capsule of mycobacteria with respect to the polysaccharides are neutral and lipid-free D-glucan, D-arabino-D-mannan and D-mannan, corresponding to molecular masses of 120, 13 and 4kDa, respectively (Ortalo-Magne et al., 1995). The glucan, represented up to 90% of the polysaccharides, is composed of repeating units of five or six $\rightarrow 4-\alpha-D-Glcp-1\rightarrow$, residues and a $\rightarrow 4-\alpha-D-Glcp$ substituted at position 6 with an $\alpha-D-Glcp$, indicating a glycogen-like highly branched structure whereas the arabinomannan consisted of a mannan segment composed of a $\rightarrow 6-\alpha-D-Man-1\rightarrow$ core substituted at some positions 2 with an $\alpha-D-Manp$ (Lemassu and Daffe, 1994). As a test case, when checked for glucan, database revealed that it is not only present in different mycobacterial species but also in many human pathogenic fungi. Further, it has been considered as vaccine component for Candida albicans and Aspergillus fumigates (Torosantucci et al., 2005). Anti-glucan monoclonal antibody 24c5 (Schewbach et al., 2002) reacted equally well with M. microti and M. tuberculosis H37Rv in ELISA and IFA (Figs. 29 and 30) without any observed reactivity with Candida albicans in ELISA (Fig. 31), thus indicating that with respect to the epitope of mAb 24c5, glucans of these two species of mycobacteria are same, but differ from that of Candida albicans. On the other hand, anti-arabinomannan monoclonal antibody 9d8 (Glatman-Freedman et al., 1996) did not react with M. microti, unlike M. tuberculosis H37Rv (Figs. 29 and 30) indicating that arabinomannan differs in these two species. These findings were further corroborated by the observations that: i) unlike mAb 9d8 (anti-arabinomannan antibody), mAb 24c5 (anti-glucan antibody) coating of M. microti overcomes the mycobacteria mediated inhibition of NFkB activation (Fig. 32), thus modulating the intracellular survival of mycobacteria and, ii) glucans when given along with M. bovis BCG for infection in to Balb/c mice reduced the level of infection occupying the sites meant for the entry (Hetland et al., 1998). Although, elaborative studies could not be carried out, yet the database described here might find wide application in structural and functional characterization of capsular polysaccharides not only in mycobacteria but also in other microorganisms.

The mAb E2B9 (IgG1) reacted strongly with the cell wall proteins of M. microti in ELISA and in western blot (Fig. 33). The antibody not only reacted with M. microti cell surface but also recognized M. tuberculosis complex members like M. tuberculosis H37Rv, M. bovis, M. avium and avirulent species like M. smegmatis and M. tuberculosis H37Ra (Figs. 34-36). In western blotting with cell wall fractions of M. microti, M. smegmatis, M. bovis and M. avium this antibody revealed, a common band at 29kDa, along with additional bands except for M. bovis (Fig. 37). Thus, above result (Fig. 37) indicated that the
component recognized by the antibody is conserved among mycobacterium species with minor individual variations. When checked in infected macrophages, using GFP tagged *M. tuberculosis* H37Ra, mAb E2B9 reactive component could be localized intracellularly (Fig. 44). Interestingly, the antibody also reacted in the nuclear region of normal cells, barring those few cells which differed morphologically (Figs. 44 and 45).

The immunoprecipitated antigen, from the cell wall proteins of *M. microti* (Fig. 39, panel A and B) and cytosolic proteins of *M. tuberculosis* H37Ra (Fig. 39, panel C and D), gave similar N-terminus sequence (Fig. 40, panel A). The 15 amino acid sequence retrieved showed 100% homology with Rv2986c, also known as mycobacterial DNA binding protein 1 (MDP1) or eukaryotic histone like proteins (HUP B) (Fig. 40, panel B). As *M. microti* genome is not yet out we blasted the unfinished *M. microti* genome lying at www.sanger.ac.uk with Rv2986c (Fig. 41). The *M. microti* genome consisted of a protein similar to Rv2986c but it lacked the stretch of 18 basic amino acids at C-terminus. This appears a plausible explanation for the immunoprecipitation of lower size of MDP-1 from the cell wall proteins of *M. microti* as observed in Fig. 39, panel A. This observation might have implication towards PCR based diagnostic to differentiate between *M. tuberculosis* complex members *M. tuberculosis* H37Rv, *M. microti* and *M. bovis*. The gene for histone-like protein (*hupB [Rv2986c]*) of *M. tuberculosis* has been identified as a singular target which allows differentiation of two closely related mycobacterial species, namely, *M. tuberculosis* and *M. bovis* of the MTB complex, by a PCR assay by Prabhakar et al., 2004. Sequence analysis indicated that in *M. bovis* there was a deletion of 27 bp (9 amino acids) in frame after codon 128 in the C-terminal part of the *hupB* gene. Although the authors too included *M. microti* in their study but they somehow missed the 54bp (18 amino acids) deletion at the same site as revealed by genomic sequence available at www.sanger.ac.uk (Fig. 41) and experimental evidence reported in this particular study (Fig. 39, panel A).

That the mAb E2B9 is indeed directed towards MDP1 was further confirmed by its reactivity with the recombinant MDP1, although partially degraded during its expression and purification (Fig. 43, panel D). The observed degradation of the recombinant MDP1 might be due to susceptibility of the basic amino acid rich C-terminus of the protein containing potential cleavage sites for the serine endoproteases as reported earlier in literature (Menozzi et al., 1998, Cohavy et al., 1999).

Data presented in this study demonstrate the localization of MDP1 on mycobacteria cell surface and cytosolic fraction (Figs. 34-38). Other groups also found that a histone-
like DNA-binding protein, designated as mycobacterial DNA-binding protein 1 (MDP1), laminin-binding protein, histone-like protein (HLP), or HupB, not only localizes in the cytoplasmic region but also occurs externally or is in the mycobacterial cell wall (Matsumoto et al., 1999, Shimoji et al., 1999, Soares de Lima et al., 2005, Katsube et al., 2007). Despite the surface location of this protein, amino acid sequence analysis revealed the lack of a signal peptide, as also reported for some exported proteins of mycobacteria which do not possess signal sequence (Thole et al., 1995, Menozzi et al., 1998), although the exact mechanism of exporting such proteins is not known.

The observed molecular weight of MDP1, revealed by mAb E2B9 in cell wall proteins of different mycobacterial species, was in the form of doublets in case of M. tuberculosis Rv, M. tuberculosis Ra and M. microti at 29-30 kDa along with 23kDa lower band in M. microti cell wall proteins (Fig. 38). The results can be explained on the basis of literature reports by Menozzi et al., 1998, Matsumoto et al., 1999, Shimoji et al., 1999, who had reported that although deduced amino acid sequence of Rv2986c predicts a 22kDa protein, it migrates in SDS-PAGE as a 6-7kDa larger molecule than the calculated molecular size. This might be because of the high content of Lys and Arg residues, which make MDP1 a highly positively charged molecule, thus allowing it to migrate slowly in SDS-PAGE. Furthermore, the reactive proteins expressed by the different mycobacterial species showed slight variation in apparent molecular mass. These minor differences might reflect variations in amino acid sequence length, posttranslational modifications, or sequence differences as suggested by Cohavy et al., 1999. The doublets in case of M. tuberculosis H37Rv and Ra could be justified based on interesting observation by Calder and Horwitz, 1998, according to which this protein was identified as a major iron-regulated protein of M. tuberculosis, with two forms differing slightly in apparent mass, one form (referred to as Irp28) upregulated by low iron concentrations and the other form (Irp29) up regulated by high iron concentrations, being modified differently when expressed under different iron conditions leading to mobility difference. Also, since the protein sequences is deprived of any cysteine so there is no possibility of intra or inter molecular disulphide bonds.

High serum levels of anti-neutrophil cytoplasmic antibodies (pANCA) is a marker immune response in inflammatory bowel disease (IBD) associated with 60 to 70% of patients with ulcerative colitis (UC) and a subset of Crohn’s disease (CD) patients (Shanahan et al., 1992, Yang et al., 1993, Satsangi et al., 1998). Using a pANCA monoclonal antibody, the C-terminal basic random-coil domain of histone H1 was
identified as a pANCA autoantigen (Eggena et al., 1996, Eggena et al., 1999). BLAST analysis of the peptide databases revealed H1 epitope homologues in open reading frames of the *M. tuberculosis* genome corresponding to conserved 214-amino-acid, an iron-regulated protein termed MDPI/HupB/Rv2986c. Sequence analysis demonstrated its homology with the mammalian histone H1 gene family, and recombinant protein expression confirmed its reactivity with the pANCA monoclonal antibody (Cohavy et al., 1999). These observations provide plausible explanation for the observed reactivity of mAb E2B9 with the perinucleolar region of macrophages, as histone H1 has been localized to perinuclear chromatin by Hutchinson and Weintraub 1985, Clark et al., 1991. The nuclear periphery is believed to be relatively transcriptionally active with an open chromatin conformation, which may favor increased immunoaccessibility. Thus, it is likely that mAb E2B9 cross reacts with the C-terminal domain of eukaryotic histone H1 proteins resulting its nuclear localization. Also, it provides support that epitopic region of mAb E2B9 lies towards the C-terminus of Rv2986c which might be lysine and arginine rich region, more specifically P(AKKA)A motif, responsible for its homology to the mammalian histone H1 proteins.

Briefly, MDPI protein possesses a striking combination of two features: it is a DNA-binding protein which is also surface exposed (Matsumoto et al., 1999). It has been associated with a large variety of cellular functions in different mycobacteria like replication, transcription and translation (Furugen et al., 2001), growth regulation (Matsumoto et al., 2000), adaptation to dormancy (Lee et al., 1998) and adhesion to Schwann and epithelial cells (Shimoji et al., 1999, Aoki et al., 2004). Recently, the MDPI gene from *M. bovis* has been shown to play an important role in cell wall biosynthesis by binding to the antigen 85 (Ag85) and to its substrate trehalose-6-monomycolate (Katsube et al., 2007). The extremely strong expression of MDPI reaching up to 8–10% of the total protein amount in *M. bovis* BCG (Matsumoto et al., 1999) suggests its high importance for mycobacterial growth and survival. Protein structure of this protein as solved by Bhowmick et al., 2009, revealed its Y shaped structure (Fig. 46), very appropriate for DNA binding and acting as a transcriptional regulator as predicted in the literature. Further, Furugen et al., 2001, demonstrated that stretch of 31-50 amino acids from N-terminus of MDPI is responsible for DNA binding and inhibited synthesis of both DNA and RNA in vitro. This stretch was found to be localized in the centre of the cavity in MDPI solved structure (red highlighted region of Fig. 46). This structural analysis suggested that this protein has got appropriate structural topology of a transcription regulator.
Chromatin immunoprecipitation (ChIP) assays are a powerful technique that allows detection of protein-DNA interactions in vivo (Kuo and Allis, 1999, Orlando, 2000). This approach, when applied to MDP1, revealed the presence of upstream promoter region of Rv0688 which codes for putative ferredoxin reductase and Rv3375 (Ami D) which codes for a amidase. Although the fragment obtained were from M. tuberculosis H37Ra but when blasted in NCBI, showed 100% identity with M. tuberculosis H37Rv genome and hence the coding region were derived from the same (Fig. 48). Yeruva et al., 2006, observed the coordinated expression of MDP1/HUPB protein with mycobactin and carboxymycobactin indicating that its expression in vivo may be influenced by iron limitation. The mycobactin and its extracellular counterpart carboxymycobactin are the siderophore which are able to remove iron from host sources of iron such as transferrin and ferritin as well as being able to solubilize it from inorganic sources such as ferric hydroxide or ferric phosphate. Restricting the availability of iron is an important strategy for defense against bacterial infection (Barclay and Ratledge, 1986). The mechanism of the release of iron from mycobactin is by a ferric mycobactin reductase in which the ferric iron is reduced in the presence of NAD(P)H to ferrous iron (Ratledge, 1971, Brown and Ratledge, 1975, McCready and Ratledge 1979). Whether iron regulated protein, Irep-28/MDP1/HUPB, is a receptor for ferricarboxymycobactin is a hypothesis that requires further experimental proof (Yeruva et al., 2006). In light of above observation, Irep-28/MDP1/HUPB, which exists in two different forms depending upon iron concentration in cell wall, regulating at least putative ferredoxin reductase (Rv0688) make sense because it is the enzyme involved in release of iron in to the cytoplasm, to be assimilated for the production of iron binding protein. As both the coding regions, obtained through ChIP (Fig. 48) have not been studied so far, hence nothing could be imparted to their exact physiological location and function. However, the regulation of these proteins or binding of MDP1 to upstream sequences of Rv0688 (putative ferredoxin reductase) / Rv3375 (AmiD), needs to be substantiated by mapping their respective promoter regions and evaluating the binding by EMSA. Nevertheless, this study provides the first indication by ChIP experiments that MDP1 is the transcription regulator.