6.1 Summary

Tuberculosis is still one of the major killer diseases throughout the world. The emergence of multi drug-resistant strains and lacunae in complete effectiveness of BCG vaccine has severely threatened the efficacy of TB treatment. New vaccines and drugs are needed to stem the worldwide epidemic of TB that kills two million people each year. To rationally develop new anti-tubercular agents, it is utmost essential to thoroughly study the genetics and physiology of *Mtb*. It is equally important to understand that how the tubercle bacilli circumvent the host defenses and cause disease? Although, the human immune response against *Mtb* is highly effective in controlling the primary infection, the organism is almost never eradicated. *Mycobacterium tuberculosis* is foremost among bacterial pathogen in its ability to establish and maintain latency, a period during which the infected person does not have a clinical symptom but harbors *Mtb* which are able to reactivate at a later stage. Therefore, one of the keys to understand the overall molecular pathogenesis of *Mtb* is the identification of the important molecules and the major cellular events which assures its survival within the host. Exploration of mycobacterial genome sequences in recent years has brought many novel insights into the biology of *Mtb* leading to the identification of several previously unknown but important genes including presence of seven *whiB* genes. Therefore, the objective of the present work was "to study the structural and functional properties of WhiB proteins of *Mtb* especially in context of their role in redox signaling".

The seven *whiB*-like genes of *Mtb* were identified based upon their sequence similarity to the *whiB* gene of *Streptomyces coelicolor*, where it has been shown to be involved in the process of sporulation (Davis and Chater, 1992) and were predicted to encode for a transcription factor like protein however, even after 16 years of their discovery, there is not a single direct evidence to show that WhiB proteins really work as a transcription factor. As none of the mycobacterial species is known to sporulate therefore, presences of these genes have always been intriguing. Despite of severe retrogressive evolution, *M. leprae* has retained five of the *whiB*-like genes pointing towards important biological function(s). Although, WhiB proteins of *Mtb* share good amino acids sequence homology, the literature collected over last 5-7 years suggest that the WhiB proteins of *Mtb* are involved in different cellular processes such as cell division, fatty acid metabolism, antibiotic resistance, stress sensing and overall pathogenesis (Reviewed in chapter 1). In spite of several genetic studies on *whiB* genes, detail biochemical properties of WhiB proteins have not been worked out. Multiple sequence alignment of WhiB proteins from

various actinomycetes (as they are unique to actinomycetes) show the presence of four conserved cysteines out of which two are present as a CXXC motif (Fig. 1.2). The unique reactivity of cysteine is mirrored by the myriad of functions it fulfills *in vivo* which includes structural stabilization, catalysis, redox-activity, metal-binding *etc.* (Giles *et al.*, 2003). The CXXC motif has been implicated in proteins with diverse functions *e.g.* protein disulfide oxido-reductase activity, redox sensing, and in the co-ordination of several metal cofactors (Fig. 1.3). Therefore, we assumed that the key to unravel the properties and biochemical function(s) of WhiB proteins lies in four conserved cysteines and in the CXXC motif. To unravel the molecular function(s) of WhiB protein(s), biochemical approach was adopted. To begin with *whiB4*/Rv3681c was chosen as a candidate gene for detail characterization. The gene was PCR amplified, cloned and over-expressed as C-terminal 6×His-tagged fusion protein in heterologous host *E. coli* Bl21(DE3) (Fig. 2.5). The WhiB4-6×His protein was purified to homogeneity by Ni^{2+}-NTA affinity chromatography in a single step (Fig. 2.7). A simple, rapid and efficient matrix-assisted refolding method was also developed to obtain native protein from inclusion bodies. Briefly, the protein was non-covalently immobilized to Ni^{2+}-NTA affinity matrix under denaturing conditions. Protein was allowed to refold by washing the column with a buffer where the urea concentration was gradually decreased. Finally, the protein was eluted in a buffer without urea. After analyzing series of biochemical and biophysical properties which includes secondary structural elements (Fig. 2.5B), surface hydrophobicity, hydrodynamic volume (Fig. 2.9), iron-sulfur content (Fig.3.5A, Table 3.2) and enzymatic activity (Table 4.1), it was concluded that the protein was successfully refolded and attained the native conformation.

The brownish appearance of freshly purified protein indicated that the WhiB4 may bind a chromophore (Fig. 3.1). Indeed, the absorption spectroscopy and metal analysis revealed the presence of a [2Fe-2S] cluster which was bound non-covalently to WhiB4 (Fig. 3.2, Table 3.2). Although, the purified protein contained iron but the amount was sub-stoichiometric therefore, the iron-sulfur cluster of WhiB4 was reconstituted *in vitro* and the biochemical and biophysical properties of reconstituted protein was studied. The iron-sulfur reconstituted protein contained 2.6-3.5 Fe atoms per monomer (Table 3.2). In contrast to freshly purified protein, the UV-visible absorption spectra (Fig. 3.4A) and 'Electron Paramagnetic Resonance' (EPR) spectra of reconstituted WhiB4 suggested that the protein coordinated a [4Fe-4S] cluster (Fig. 3.4B). Rapid loss of iron-sulfur cluster in oxidizing environment (incubation in the presence of air, GSSG or H_2O_2) and its stability under

reducing conditions (Fig. 3.6 and 3.7) (incubation in the presence of GSH or DTT) suggested that the metal cluster of WhiB4 is redox sensitive thus may have a regulatory role. Further, the rate of cluster loss was directly proportional to the oxidizing strength of the agent and also on the exposure time. It appears that the [4Fe-4S] cluster of WhiB4 being highly oxidation labile is first converted into [2Fe-2S] cluster, a relatively stable intermediate but ultimately disassembled from the protein under oxidizing environment leading to the generation of apo protein (Fig. 3.8). Although, the iron-sulfur cluster of WhiB4 was oxidation labile but was stably co-ordinated to the protein, as even high concentration of chaotropic agents (8 M urea & 6 M GdnHCl) (Fig. 3.3, Table 3.2) which otherwise unfolded the protein (Fig. 3.5) significantly but could not remove the cluster. The involvement of cysteine residues in the co-ordination of several metal ions including iron-sulfur cluster have been observed in many proteins therefore, we assumed that the four-conserved cysteine residues are possibly involved in the binding of iron-sulfur cluster. A site directed mutagenesis approach to decipher the role of cysteines in the cluster co-ordination. Cysteines (Cys^{37} Cys^{59} Cys^{62} and Cys^{68}) were substituted by serine and several single, double, triple and quadruple cysteine mutants were prepared. The importance of each cysteine residue in the cluster binding was evaluated by comparing the iron-sulfur cluster binding properties of mutants with each other and with wild type WhiB4, using both absorption spectroscopy and total iron analysis (Fig. 3.9 and 3.10, Table 3.3). The data suggested that all four cysteine residues are involved in cluster co-ordination and the cysteines present in the CXXC motif (Cys^{59} and Cys^{62}) were more important. This study revealed **'$CX^{21}CXXCX^{5}C$'** arrangement as a novel [4Fe-4S] cluster co-ordinating sequence.

The apo WhiB4 showed retarded mobility both on a non-reducing 15% SDS-PAGE and 15% native PAGE, when treated with DTT, β-ME and GSH (Fig. 4.1), raising a possibility for the presence of intra-molecular disulfide bond(s). To demonstrate that intra-molecular disulfide bond(s) are formed in apo WhiB4, the protein was alkylated with iodoacetamide, both under oxidized and reduced conditions and was analyzed by mass spectrometry (Fig. 4.2). Various double and quadruple cysteines mutants were used to decipher and validate the number and arrangement of intra-molecular disulfide bonds in apo WhiB4. The data concluded that oxidation of the iron-sulfur cluster and its concomitant disintegration lead to the formation of two intra-molecular disulfide bonds, wherein the first

is between the Cys^{59} and Cys^{62} (CXXC motif) and the second is between the Cys^{37} and Cys^{68} (Fig. 4.3).

The presence of CXXC motif and confirmation that a disulfide bond is formed between the two cysteines of the motif led to assume that WhiB4 might have properties similar to thiol-disulfide oxido-reductases. The CXXC is an active site motif of proteins of thioredoxin superfamily (Holmgren, 1985). They all have an active site disulfide bond between cysteine residues of CXXC motif, a common structural fold and are involved in thiol-disulfide exchange reactions (Arner and Holmgren, 2000). Despite primary sequence divergence from thioredoxin family proteins, the apo WhiB4 showed several properties, similar to Trxs. These are: (a) thermal stability of secondary structure (Fig. 4.4B) (b) fluorescence quenching of tryptophan by disulfide in oxidized form (Fig. 4.4C) and (c) ability to catalyze the reduction of insulin disulfide *in vitro* (Fig. 4.5, Table 4.1), a signature assay to confirm that the protein has a disulfide reductase activity. In an attempt to confirm that the disulfide reductase activity of WhiB4 is governed by the CXXC motif, a curious observation was made where conversion of CXXC into SXXS led only ~50% loss of activity (Table 4.1). Similar loss of activity was also observed when the CXXC motif was intact but the other two cysteines (Cys^{37} and Cys^{68}) were mutated to serine (Table 4.1) suggesting that apart from the cysteine thiol of CXXC motif the distantly placed thiol pair also contributes equally to the enzymatic activity of WhiB4. The data suggested that the non-CXXC cysteine residues came close enough in the three-dimensional form and are involved in the exchange of disulfide during thiol-disulfide exchange reaction. This is the first report to show that distantly placed thiol pair can also participate in thiol-disulfide exchange reactions (Alam *et al.* 2007).

Thioredoxin system is composed of Trx, TrxR and NADPH, wherein the recycling of Trx catalytic power is carried out by transfer of electron by NADPH *via* TrxR. TrxR takes electron from NADPH and transfers to the active site disulfide of thioredoxin (Holmgren, 1985). Since, all initial studies indicated that WhiB4 has thioredoxin like property; we were tempted to speculate that it may be recognized and reduced by *Mtb* TrxR. *Mycobacterium tuberculosis* has single TrxR ORF, *trxB2*/Rv3913 which has been shown to code for a functional dimeric protein (Akif *et al.*, 2005). The ORF coding for TrxR of *Mtb* was PCR amplified, cloned in pET-29A, over-expressed in *E. coli* and purified by Ni^{2+}-NTA chromatography upto 98% of purity (Fig. 4.7). TrxB2 appeared yellowish and showed absorption spectra, typical of FAD containing protein (Veine *et al.*, 1998) with

peaks at 370 nm, 450 nm and a shoulder at 470 nm (Fig. 4.8A). Biochemical assay using DTNB as a substrate suggested that TrxB2 was functionally active (Fig. 4.8B). We examined whether oxidized WhiB4 will be reduced by TrxB2 in the presence of NADPH by monitoring NADPH oxidation at 340 nm. The rate of NADPH oxidation was very slow and comparable to negative control (Fig. 4.8C). The accelerated rate of NADPH oxidation in a reaction, where mammalian thioredoxin reductase and *E. coli* thioredoxin were used (Fig. 4.8C), ruled out the possibility of experimental artifacts. Therefore, the data suggested that WhiB4 might not be a substrate of TrxB2.

The vulnerability of iron-sulfur clusters to oxidative destruction is sometimes associated with the modulation of function of a protein depending upon the redox state of the cell. Reversible cluster rearrangement or disassembly often facilitates alterations in protein structure, which in turn, regulate the function of protein as observed in SoxR, Fnr (Beinert and Kiley, 1999), human Grx2 (Lillig *et al.*, 2005) *etc.* The role of iron-sulfur cluster on the conformation of WhiB4 was studied by ANS binding assay using fluorescence spectroscopy. Fluorescence intensity of apo protein was ~5 fold higher at 499 nm than the holo protein suggesting for a conformational change in WhiB4 after the loss of Fe-S cluster (Fig. 4.10). To study a direct relation between the loss of iron-sulfur cluster and change in the protein conformation, the holo protein was air oxidized at 25 °C and ANS fluorescence spectra were recorded at different time intervals till 24 h. The data showed considerable increase in the ANS fluorescence with the increase in exposure time (Fig. 4.10) which clearly suggested the exposure of some of the buried hydrophobic patches of the protein due to the removal of iron-sulfur cluster. In order to test the role of iron-sulfur cluster on the disulfide reductase activity of WhiB4, enzymatic activity of apo and holo WhiB4 was compared. The holo WhiB4 was enzymatically inactive suggesting that [4Fe-4S] cluster of WhiB4 has regulatory function (Fig. 4.9, Table 4.2). Further, exposure of hydrophobic patches of WhiB4 upon cluster disassembly strengthened the fact that the cluster removal is possibly essential for the interaction and subsequent transfer of reducing equivalents from WhiB4 to its target protein(s).

whiB genes are exclusively reported in actinomycetes and till date, genome sequencing has revealed over 270 *whiB* homologues in different actinomycetes (den Hengst and Butner, 2008). Although, WhiB proteins of *Mtb* share relatively high (amino acids) sequence homology, the available literature suggest that the WhiB proteins of *Mtb* are involved in different cellular processes. The *in vivo* functional diversity among seven WhiB

proteins of *Mtb* raises several questions: are all WhiB proteins co-ordinate an iron-sulfur cluster? If yes, then what are their basic properties? Are the iron-sulfur cluster equally oxidation labile? Does iron-sulfur cluster removal lead to the formation of disulfide bond between the cysteines? Are structural features of mycobacterial WhiB proteins similar? Are all WhiB proteins behave like protein disulfide reductase? *etc*. Therefore, to address the above questions we sought to study and compare the biochemical and biophysical properties of all seven WhiB proteins of *Mtb*.

The annotated *whiB*-like ORFs were PCR amplified, cloned, expressed and the recombinant proteins were purified by Ni^{2+}-NTA affinity chromatography (Fig. 5.1). The UV-visible spectroscopy suggested the presence of a [2Fe-2S] cluster in each member of the family (Fig. 5.3). The total iron content of the native and in-column refolded proteins varied between 0.14-0.20 atoms per monomer (Table 5.1). The iron content of proteins purified from soluble fraction; from inclusion bodies under denaturing conditions and after refolding, was similar (Table 5.1) suggesting that the protein fold responsible for holding the iron-sulfur cluster was resistant to the denaturing effect of 8 M urea.

Interestingly, different WhiB proteins responded differently against the oxidative effect of air and GSSG. Iron-sulfur cluster of WhiB1 was found to be the most oxidation labile whereas, WhiB6 and WhiB7 clusters were comparatively stable (Fig. 5.5 and 5.6). In each case the rate of [2Fe-2S] cluster loss were significantly reduced under reducing environments (Fig. 5.6). The differential response showed by *whiB* genes against different stress conditions and our observation that all seven WhiB proteins responded differently to oxidizing environments *in vitro*, argued to conceive the proposition that the seven WhiB proteins of *Mtb* would respond and relay the redox signal differently. We hypothesized that the differential sensitivity of iron-sulfur cluster towards different oxidizing agents is governed by their relative surface accessibility. The assumption gains support from the fact that the extent of iron chelation by EDTA varied significantly among the seven members and could be co-related to their oxidation sensitivity. After 20 h of incubation, iron chelation was ~60% (maximum) in WhiB1, whereas it was 15-20% (minimum) in WhiB6 and WhiB7 (Fig. 5.7). In all of them, oxidation of the [2Fe-2S] cluster was coupled with the generation of apo form in which the cysteines thiol oxidized to form two intra-molecular disulfide bonds (Fig. 5.9). Formation of reversible intra-molecular disulfide bond between the cysteines of CXXC motif is essential for a protein disulfide reductase activity therefore, we speculate that in WhiB proteins, one disulfide bond is formed between the two cysteines

of CXXC motif (CXXXC in case of WhiB5) and the other is between the remaining two cysteines (Fig. 5.9F). The assumption is supported by the WhiB1 and WhiB4 data where the similar arrangements of intra-molecular disulfide bonds have been established (Section 4.3.1). In WhiB6, although one of the intra-molecular disulfide bonds would have formed between Cys^{53} and Cys^{56}, but the involvement of cysteines for the second bond remains to be worked out, as it contains five cysteines (Cys^{12}, Cys^{34}, Cys^{53}, and Cys^{56} and Cys^{62}) unlike other WhiB proteins.

The multiple sequence alignment of WhiB proteins of *Mtb* showed 49-66% sequence homology and 31-50% identity among themselves (Table 1.2). Therefore, due to the variation in amino acid composition and sequence, there is a possibility of structural variations which may act as an important determinant to their functional properties *in vivo*. Therefore, the structural organizations of each of the *Mtb* WhiB proteins were studied and compared employing several biophysical tools. The far-UV CD spectra of WhiB proteins were dissimilar as their molar ellipticities varied significantly (Fig. 5.11). In WhiB5 and WhiB6, the structure was dominated by α-helix and β-strands. WhiB1, WhiB2 and WhiB4 showed significant amount of random coils (Fig. 5.11). The proportion of all three secondary structural elements in WhiB3 and WhiB7 appeared almost similar. The secondary structure contents of WhiB proteins remained unaffected due to disulfide bond formation, except in WhiB6 and to some extent in WhiB5 (Fig. 5.11). Unlike others, the secondary structures of WhiB5 and WhiB6 started to melt at 50 °C and 77 °C respectively (Fig. 5.12). At 70 °C, WhiB5 lost almost all secondary structures, whereas in WhiB6 some secondary structure was maintained even at 95 °C. In order to estimate the level of β-sheets structures in different WhiB proteins, ThT binding assay was performed where the binding was measured by fluorescence spectroscopy. The fluorescence intensities of each protein varied with respect to each other (Fig. 5.13). It appears that there is a major contribution of crossed β-sheet structures in WhiB5 and WhiB6, but the proportion is relatively low and almost similar in other WhiB proteins.

In order to investigate the differences at tertiary structure level, surface topology of each of the WhiB proteins was studied. Presence of total exposed hydrophobic patches was probed by ANS binding using fluorescence spectroscopy. The ANS fluorescence of various WhiB proteins was significantly different (Fig. 5.14). The fluorescence intensity was maximum in WhiB3 and minimum in WhiB6. Based on ANS fluorescence data, the surface hydrophobicities of various WhiB proteins fall in the following order:

WhiB3>WhiB1>WhiB4>WhiB2>WhiB5 >WhiB7>WhiB6 (Fig. 5.14). Further, the relative cross reactivity of polyclonal antibodies was used as a measure to probe the conformational variations at tertiary structure level amongst different WhiB proteins. Polyclonal antibodies against all WhiB proteins were raised in rabbit and their cross-reactivity were tested by ELISA. We observed considerable differences in the antibody cross-reactivity (Fig. 5.15) which suggested that the antigenic sites on different WhiB proteins are heterogeneous. We assume that the conformational difference among various WhiB proteins is the possible cause for this heterogeneity. Together, the data demonstrated that there are significant structural differences among different WhiB proteins of *Mtb*. Similar to WhiB4, apo form of each protein, except WhiB2, showed protein disulfide reductase activity however, the efficiency of reduction varied significantly. Despite the presence of several overlapping molecular properties, each member of WhiB family appears to be significantly different and may represent a novel redox system of *Mtb*.

6.2 Conclusions

In the available genome sequence of *Mtb* the seven WhiB proteins are annotated as putative transcription factors. Bioinformatic analyses of the proteins using recent tools show that they lack a standard DNA binding motif. Various data have been taken as circumstantial evidence that WhiB proteins are transcription factor. However, till date, there is no direct observation that these protein indeed work as transcription factor. We adopted a biochemical approach to delineate the molecular function of WhiB proteins of *Mtb*. The present work identifies WhiB proteins (except WhiB2) to work as protein disulfide reductase in the apo form. The cysteine bound oxidation labile iron-sulfur cluster was found to act as a regulatory switch. Based upon the available data we assume that WhiB proteins of *Mtb* act as a redox sensing regulatory proteins which possibly functions under stress to relieve and maintain the redox homeostasis of *Mtb* especially during infection and pathogenesis.

Based on the results described in the present study, we propose a functional model of WhiB4 (Fig. 6.1). We assume that WhiB4 remains as a holo protein in normal growth conditions but once *Mtb* invades the host, it encounters oxidative and nitrosative stress in the form of ROIs and RNIs. These agents attack the iron-sulfur proteins including WhiB4 which leads to the disassembly of [4Fe-4S] cluster, generation of apo WhiB4 and the formation of two intra-molecular disulfide bonds. The process also brings conformational