The PhoPR TCS has attracted attention in the past few years. A number of investigations show that inactivation of *phoP* in Mtb H37Rv leads to significant growth attenuation (Perez *et al.*, 2001; Gonzalo-Asensio *et al.*, 2006; Walters *et al.*, 2006). Also, biochemical studies reveal that PhoP regulates sulphatides, diacyltrehaloses and polyacyltrehaloses and absence of these lipid molecules in the *phoP* mutant is the major reason for its attenuated growth in a mouse model (Gonzalo Asensio *et al.*, 2006; Ludwiczak *et al.*, 2002) (for review, see Ryndak *et al.*, 2008 ). While two independent studies show that a point mutation in *phoP* contributes to avirulence of Mtb H37Ra (Chesne-Seck *et al.*, 2008, Lee *et al.*, 2008), more recently PhoP has been implicated in the ESAT-6 secretion and specific T-cell recognition during virulence regulation of the bacilli (Frigui *et al.*, 2008). Thus, accumulating evidences suggest that PhoP is a key regulator of Mtb. However, molecular mechanism of how it functions remains largely unknown.

Many members of the PhoP subfamily use different mechanisms to regulate their DNA binding and transcription regulation despite having high sequence similarity. In this study we set out to investigate mechanism of action of PhoP as a regulator of more than 100 genes of Mtb. Chapter 2 comprises of detailed study of DNA protein interaction of PhoP. Although, PhoP has been shown previously to bind to its own promoter, we identified a direct repeat sequence as the primary target site for sequence-specific DNA binding by PhoP (Gupta *et al.*, 2006). Here, we showed that two PhoP protomers are recruited on it's target DNA comprising a 9-bp direct repeat motif. We also show (i) that DNA binding stimulates the dimerization of PhoP, and (ii) the two molecules are structurally organized in a specific head-to-head orientation.

The crystal structure of PhoPC clearly shows that the primary DNA binding of the protein involves winged helix-turn-helix motif (PDB ID: 2PMU) and the surface around the PhoP residues comprising the recognition helix (residues Asn212-Tyr224 of α8) display strong positive electrostatic potential, indicating that these residues are likely to be critical in DNA binding and nucleotide sequence recognition (Wang *et al.*, 2007). To this end, we used structure-guided mutagenesis to obtain single alanine substitutions of 10 solvent-exposed residues spanning α8. Our results of rational mutagenesis coupled with DNA-binding affinity study of the α8-DNA interface in the complex formed by PhoP and its cognate DNA demonstrate that most PhoP mutants have significantly reduced DNA-binding affinity while

possessing near wild-type stability. However, alanine substitution of Glu215 of α8 shows major effect on the specificity of DNA recognition. Using structural insights coupled with biochemical analyses, we identify that Glu215 of PhoP appears to establish a base-specific interaction with $(G/C)^9$ of the upstream repeat motif (DR1 of DR1,2) to contribute significantly to the recognition specificity of the regulator. Biochemical experiments corroborate these results showing that DNA recognition specificity can be altered by as little as a single residue change of the protein or a single base change of the DNA.

Another objective was to investigate domain structure of Mtb PhoP and how does it contribute to PhoP's function. The several functions of PhoP are apportioned between a C-terminal effector domain (PhoPC) and an N-terminal receiver domain (PhoPN), phosphorylation of which regulates activation of the effector domain. In the 3rd chapter we show that PhoPN, on its own, demonstrates PhoR-dependent phosphorylation. PhoPC, the truncated variant bearing the DNA binding domain, binds *in vitro* to the target site with affinity similar to that of the full-length protein. To complement the finding that residues spanning Met1 to Arg138 of PhoP constitute the minimal functional PhoPN, we determined Arg150 as the first residue of the distal PhoPC domain capable of DNA binding on its own, thereby identifying an inter-domain linker. We further show that coupling of two functional domains together in a single polypeptide chain is essential for phosphorylation-coupled DNA binding by PhoP.

Chapter 4 originates from the interesting domain structure that was detailed in chapter 3. To better understand inter-domain interaction(s) in effector domain regulation, we sought to investigate domain structure of PhoP. To this end, we identify an 11-residue long inter-domain linker that tethers two functionally-independent domains of PhoP together and regulates inter-domain interactions. While the newly-identified linker region is not required for either domain functions of PhoP, most strikingly, it plays an essential role for phosphorylation-dependent DNA binding to *msl3* promoter, previously suggested to be regulated by PhoP (Walters *et al.,* 2006). Interestingly, biochemical studies reveal that one of the major differences between OmpR and PhoB reside in the inter-domain linker region that tethers together the N-terminal domain with the C-terminal domain (Walthers *et al.*, 2003). Consistent with this view, a previous study had shown that C-terminal DNA binding by OmpR could influence phosphorylation of the N terminus in which the linker region underwent a conformational change (Ames *et al.*, 1999), thus suggesting a key role of the linker region in regulation of inter-domain interaction(s). Together, our results suggest that

the DNA binding energy and specificity of regulator-promoter interactions is
ted primarily (but not entirely) by the C-domain, linker region of the protein likely
ne regulator to adopt a different phosphorylation-dependent conformation enabling it
minate target promoters while it regulates a vast array of genes to either activate or
ranscription.

To survive in an inhospitable world, microbes, like all life forms, must be able to adapt to changing environmental condition. Particularly, intracellular parasites are faced with a new hostile environment in which host cellular defense mechanisms are sophisticated and effective. In response, these pathogens must be able to sense when they have entered a host cell and adapt accordingly. Much of the reason for the success of Mtb as an intracellular pathogen lies in its ability to adapt to its host environments through signal transduction leading to switching on of complex transcriptional programs. It is now known that the major response of the bacterium to environmental changes is through classical TCSs via histidine-aspartate phosphorelay between the sensor kinase and the response regulator. A number of recent studies revealed that PhoP of the PhoPR system controls a variety of functions including synthesis of complex pathogenic lipids, hypoxia response through DosR cross-talking, respiratory metabolism, secretion of the major T-cell antigen ESAT-6, et cetera (Gonzalo-Asensio *et al.*, 2006; Walters *et al.*, 2006; Gonzalo-Asensio *et al.*, 2008b) . Further supporting the role of PhoP in regulation of Mtb virulence, two recent articles suggest that a point mutation in PhoP contributes to avirulence and also accounts for the absence of polyketide-derived acyltrehaloses in Mtb H37Ra (Chesne-Seck *et al.*, 2008; Lee *et al.*, 2008).

PhoP, a member of the *E. coli* OmpR/PhoB subfamily, consists of two functional domains, an N-terminal receiver domain and a C-terminal transactivation domain (also called an effector domain). The N-domain, like other members of the RR family also shares a conserved doubly-wound $(\alpha/\beta)_5$ topology with a phosphorylation site at the N-terminus (Asp71 for PhoP; Gupta *et al.*, 2006). The C-domain of the protein has been structurally characterized (PDB ID: 2PMU; Wang *et al.*, 2007). The structural analysis revealed overall folds similar to those of four other OmpR family proteins, Mtb PrrA (Nowak *et al.*, 2006), *E. coli* PhoB (Okamura *et al.*, 2000), *Thermotoga maritima* DrrD (Buckeler *et al.*, 2002), and *B. subtilis* PhoP (Birck *et al.*, 2003) with a winged-helix-turn-helix DNA binding motif involved in DNA binding. Despite global functional diversity, members of the PhoP family share significant structural homology in their receiver domain as well as in the basic mode of DNA binding. All of the family members utilize a winged helix-turn-helix DNA binding motif,

PhoP has been experimentally shown to bind direct tandem repeat sites (Blanco *et al.*, 2002; Ryndak *et al.*, 2002) and inverted repeats of DNA (Glover *et al.*, 2007). However, there are significant differences in the mechanism to regulate DNA binding activity and modulate transcription. The only reported interaction of PhoP from Mtb H37Rv involves binding of the regulator to its own promoter (Gonzalo-Asensio *et al.*, 2008). Previously, we demonstrated transcriptional autoregulation of *phoP* by sequence-specific interaction of PhoP from Mtb H37Ra to its own promoter (Gupta *et al.*, 2006). Strikingly, these two independent studies show largely similar DNA sequences being recognized by PhoP in DNaseI footprint. However, very little is known about the sequence motif recognized by PhoP and the orientation of the protein(s) on the target DNA to promote transcription regulation. As a step towards understanding how the regulator functions, here we show sequence-specific recognition of 23-bp region of the *phoP* promoter by the protein. We further show that two molecules of monomeric PhoP are recruited on a *phoP*-promoter derived oligonucleotide-based substrate DNA comprising two direct repeat motifs. While our results suggest that DNA binding stimulates dimerization of PhoP, evidence is presented that unlike other members of the subfamily of proteins, PhoP binds to DNA in a head-to-head orientation to project their N-termini towards each other.

Although global gene expression profiling shows that 44 genes are up-regulated and another 70 genes are down-regulated by PhoP in Mtb, the origin(s) of DNA binding affinity and sequence specificity of the regulator remain largely unknown. The crystal structure of PhoPC clearly shows that the primary DNA binding of the protein involves a winged helix–turn–helix motif (PDB ID code: 2PMU) and the surface around the PhoP residues that constitute the recognition helix (α8) (residues Asn212–Tyr224). Also, these residues largely display strong positive electrostatic potential, indicating that these are likely to be critical in DNA binding and nucleotide sequence recognition. Structure-guided mutagenesis was carried out to obtain single alanine substitutions of 10 solvent-exposed residues spanning α8. The results of rational mutagenesis coupled with the DNA binding affinity of the α8–DNA interface in the complex formed by PhoP and its cognate DNA demonstrate that most PhoP mutants have significantly reduced DNA binding affinity while possessing near-wild-type stability. However, alanine substitution of Glu215 of α8 shows a major effect on the specificity of DNA recognition. (Das *et al*, 2010). Using structural insights coupled with

ve further showed that Glu215 of PhoP appears to establish a base-
(G/C)[9] of the upstream repeat motif (DR1 of DR1,2) to contribute
gnition specificity of the regulator.