# Abstract

The prokaryotic genome, though lacks a membrane bound organelle for its housing, is restricted to only about 25% of the cytoplasmic space called the nucleoid. The dramatic compaction required for the genome to fit in is mainly achieved by the combined action of three key factors, macromolecular crowding, DNA supercoiling and binding by nucleoid associated proteins (NAPs). NAPs are the prokaryotic counterparts of histones. They are the bacterial 'genome architects' and play vital roles in genome compaction and organization. By their varying DNA binding modes like bridging, wrapping, coating, bending and their changing concentrations during the bacterial growth, they shape the dynamicity of the nucleoid. Their versatile DNA binding properties also make them global transcription regulators. In *Escherichia coli* (*E. coli*), more than a dozen NAPs have been identified. Although *Mycobacterium tuberculosis* (*Mtb*) has comparable genome size to that of *E. coli*, there is a dearth of annotated NAPs in the *Mtb* genome. Hence, we have explored the possibility of the existence of yet unidentified NAPs and additional mechanisms of genome organization in the present thesis.

The **first** chapter of the thesis provides a general introduction on the structure of the bacterial nucleoid and the main players involved in its organization and packaging with emphasis on NAPs. The role of NAPs as bacterial genome architects and global gene expression regulators with specific examples has been described. Further, an overview of *Mtb* as a pathogen, its mode of survival and pathogenesis and the epidemiology of the tuberculosis disease has been discussed. Attempts to discover new drugs to counter the rising threat of drug resistant TB and the importance to understand in greater detail the physiology of *Mtb* to arrive at novel drug targets has been described. Finally, a detailed description of the mycobacterial NAPs, their properties and role in genome structuring has been presented.

In the **second** chapter of the thesis, the physical and functional interaction between HU and Lsr2 has been demonstrated. HU and Lsr2 are two of the principal NAPs in *Mtb*. HU is essential for *Mtb* survival and is one of the most abundant NAP in the nucleoid. It is different from the other eubacterial HUs, having long flexible lysine and arginine rich carboxy terminal domain (CTD). Lsr2, the functional *Mtb* analogue of the eubacterial H-NS, plays a role in genome organization and gene regulation. It binds to and regulates A/T rich parts of the otherwise G/C rich mycobacterial genome. Using co-immunoprecipitation assays and mass spectrometric analysis we have demonstrated the physical interaction between these two NAPs. This interaction occurs primarily through carboxy terminal domain (CTD) of HU and the acidic amino terminal domain (NTD) of Lsr2. The resulting complex has DNA binding properties markedly distinct from that of either HU or Lsr2. Further, by transcription assays it has been shown that the HU-Lsr2 complex is also a regulator of gene expression in a mode different from the individual NAPs. This physical and

functional interaction between the two NAPs HU and Lsr2, hitherto not observed, seems to be important for DNA organization and gene expression in *Mtb*.

Although *Mtb* has a genome comparable in size to that of *E. coli*, fewer number of ORFs encoding NAPs have been identified in its genome. In our search for yet uncharacterized NAPs in its genome, the protein Rv0430 which has NAP like features was identified. *rv0430* is the first gene of a five gene operon harboring regulators of virulence, *virR* and *sodC*. In the **third** chapter it has been demonstrated that Rv0430 is a bonafide NAP. Its ability to bind to DNA in a length and topology dependent fashion has been shown by binding assays and its preference for A/T rich DNA sequences has been demonstrated using SELEX. It has also been shown to be able to protect DNA from DNA damaging agents. Its ability to modulate DNA topology and autoregulate its own promoter in a topology dependent manner has been demonstrated. Its architecturing property has also been studied using AFM. At lower concentrations it can bridge distant DNA segments while at higher concentrations it coats the DNA forming inflexible rods. Overexpression of the NAP leads to altered nucleoid morphology. Thus, Rv0430 is a distinct NAP in *Mtb* adding to the repertoire, also doubling up as a transcriptional regulator of an operon involved in regulating virulence.

In the previous chapter the biochemical properties of Rv0430 was studied. The **fourth** chapter delineates the attempts to understand the function of Rv0430 and its *Msm* homologue MS_0833. To elucidate their physiological role, their expression was altered in *Msm*. Overexpression leads to alteration of pleotropic phenotypes and changes in the composition of lipids of the cell wall indicating its possible role as regulator of genes involved in lipid metabolism in mycobacteria. Also, to analyses the binding sites of Rv0430 and understand its regulatory impact in *Mtb*, FLAG tagged strain was generated for ChIP experiments. With these tools and the on-going experiments, the *in vivo* role of Rv0430 can be uncovered.