Synopsis of the thesis

To maintain the genomic integrity, cell has evolved various DNA repair pathways. Base Excision Repair pathway (BER) is one such DNA repair pathway which is dedicated to protect DNA from small lesions such as oxidation, alkylation, deamination and loss of bases. Uracil is a promutagenic base which appears in the genome as a result of misincorporation of dUTP or due to oxidative deamination of cytosine. Uracil-DNA glycosylases (UDGs) are DNA repair enzymes that initiate multistep base excision repair (BER) pathway to excise uracil from DNA. Excision of uracil generates an abasic site (AP-DNA). AP-sites are cytotoxic and mutagenic to the cell. AP endonucleases act downstream to UDG in this pathway and generate substrates for DNA polymerase to fill in the correct bases. The cytotoxicity of AP-sites raises the question whether uracil excision activity is coupled to AP endonuclease activity. Also, there is transient formation of single stranded DNA (ssDNA) during DNA metabolic processes such as replication, repair and recombination. ssDNA is more prone to various nucleases and DNA damaging agents. All the living organisms encode single stranded DNA binding protein (SSB) that binds to ssDNA and protects it from various damages. In addition, SSB plays a vital role during DNA replication, repair and recombination. Studies on SSBs from prototype *Escherichia coli* and an important human pathogen, *Mycobacterium tuberculosis* have shown that despite significant variations in their quaternary structures, the DNA binding and oligomerization properties of the two are similar.

My PhD thesis consists of four Chapters. Chapter 1 summarizes the relevant literature review on DNA damage and repair with an emphasis on uracil DNA glycosylase and its interacting protein, SSB. Chapters 2 and 3 describe my studies on the mechanism of uracil excision repair in *E. coli*. Chapter 4 describes my findings on the structure-function relationship of single stranded DNA binding proteins from *E. coli* and *M. tuberculosis*. Specific details of my research are summarized as follows:

(1) Analysis of the impact of allelic exchange of *ung* with a mutant gene encoding Uracil DNA Glycosylase attenuated in AP-DNA binding in the maintenance of genomic integrity in *Escherichia coli*.

There are five families of UDGs. Of these, Ung proteins (family 1 UDGs) represent highly efficient and evolutionary conserved enzymes. Structural and biochemical analysis of Ung proteins has identified two conserved motif, motif A (^{62}GQDPY^{66}) and motif B (^{187}HPSPLS^{192}) in *E. coli* that are important for the catalysis by Ung enzyme. Y66 of motif A is in van der Waals contact with the C5 position of the uracil and prevents entry of other bases. Earlier study from the laboratory showed that the Y66W and Y66H mutants of Ung were compromised by ~7 and ~170 fold, respectively in their uracil excision activities. However, unlike the wild-type and Y66H proteins, Y66W was not inhibited by its product (uracil or AP-DNA).

In this study, by fluorescence anisotropy measurements I have shown that compared with the wild-type protein, the Y66W mutant is moderately compromised and attenuated in binding to AP-DNA. Allelic exchange of *ung* in *E. coli* with *ung*::*kan*, *ungY66H*:*amp* or *ungY66W*:*amp* alleles showed ~5, ~3.0 and ~2.0 fold, respectively increase in mutation frequencies. Analysis of mutations in the rifampicin resistance determining region (RRDR)

of *rpoB* revealed that the Y66W allele resulted in an increase in A to G (or T to C) mutations. However, the increase in A to G mutations was mitigated upon expression of wild-type Ung from a plasmid borne gene. Biochemical and computational analyses showed that the Y66W mutant maintains strict specificity for uracil excision from DNA. Interestingly, a strain deficient in AP-endonucleases also showed an increase in A to G mutations. These findings have been discussed in the context of a proposal that the residency of DNA glycosylase(s) onto the AP-sites they generate shields them until recruitment of AP-endonucleases for further repair. It is proposed that an error prone replication against AP-sites (as a result of uracil excision activities on A:U pair) may result in A to G mutations.

2. Mechanism of appearance of A to G mutations in *ungY66W:amp* strain of *Escherichia coli*.

In this part of my study, I have investigated the role of error prone DNA polymerases in the mutational specificity of *ungY66W:amp* strain. It was observed from various studies in *E. coli* that, DNA polymerase IV (Pol IV) and DNA polymerase V (Pol V) are involved in error-prone replication on damaged or AP-site containing DNA. *E. coli* strains containing deletion of either *dinB* (encoding DNA Pol IV) or *umuDC* (encoding DNA Pol V) were generated and used to study mutation frequency and mutation spectrum. Deletion of DNA Pol V resulted in a decrease in A to G mutations in *ungY66W:amp E. coli* strain, suggesting that increase in A to G mutations were a consequence of error prone incorporation by DNA Pol V.

3. Structure and Function studies on Single Stranded DNA Binding Proteins from *Escherichia coli* and *Mycobacterium tuberculosis*.

SSB from *M. tuberculosis* (*Mtu*SSB) has similar domain organization as the *Eco*SSB. Moreover, the biochemical properties such as oligomerization, DNA binding affinity and minimum binding site size requirements were shown to be similar to *Eco*SSB. However, structural studies suggested that quaternary structures of these two SSBs are variable.

In this study I have used X-ray crystal structure information of these two SSBs to generate various chimeras after swapping at various regions of SSBs. Chimeras mβ1, mβ1'β2, mβ1-β5, mβ1-β6, and mβ4-β5 SSBs were generated by substituting β1 (residues 6-11), β1'β2 (residues 21-45), β1-β5 (residues 1 to 111), β1-β6 including a downstream sequence (residues 1 to 130), and β4-β5 (residues 74-111) regions of *Eco*SSB with the corresponding sequences of *Mtu*SSB, respectively. Additionally, mβ1'$\beta2_{ESWR}$ SSB was generated by mutating the *Mtu*SSB specific 'PRIY' sequence in the β2 strand of mβ1'β2 SSB to *Eco*SSB specific 'ESWR' sequence. Biochemical characterization revealed that except for mβ1 SSB, all chimeras and a control construct lacking the C-terminal domain (ΔC SSB) efficiently bound DNA in modes corresponding to limited and unlimited modes of binding. The mβ1 SSB was also hypersensitive to chymotrypsin treatment. The mβ1-β6, *Mtu*SSB, mβ1'β2 and mβ1-β5 constructs complemented *E. coli* Δ*ssb* in a dose dependent manner. Complementation by the mβ1-β5 SSB was poor. In contrast, mβ1'$\beta2_{ESWR}$ SSB complemented *E. coli* as well as *Eco*SSB. Interestingly, the inefficiently functioning SSBs resulted in an elongated cell/filamentation phenotype of *E. coli*. Taken together, our observations suggest that specific interactions within the DNA binding domain of the homotetrameric SSBs are crucial for their biological function.