Chapter 6

Summary

Helicobacter pylori is a Gram-negative microaerophilic bacteria known to infect as much as 80% of some populations with an average morbidity range of around 50% of the world population. It has been recognized as a definitive carcinogen (Type I). *H. pylori* shows extraordinary genetic diversity and this property is critical to its success as a human pathogen. High genetic diversity and interstrain variations seen in *H. pylori* is attributed to its remarkable ability to take foreign DNA by natural transformation. Natural transformation in *H. pylori* is governed to some extent by the presence of Restriction-Modification systems (R-M systems). Three types of DNA methylation are associated with R-M systems in bacteria, N6-adenine (m6A), C5-cytosine (m5C) and N4-cytosine (m4C). Recent studies in pathogenic bacteria have shown the epigenetic roles of m6A in virulence, gene regulation and genetic evolution of the organism. This is in contrast to eukaryotes where m5C is known to be the epigenetic signal. In mammals and plants DNA cytosine methyltransferases epigenetically regulate the gene expression through the precise epigenetic modification of certain cytosine residues with a methyl group. Moreover, aberrant methylation patterns are embryonic lethal in mammals, and can also lead to diseases including cancer. In plant it can result in pleiotropic morphological defects.

The role of cytosine methylation in bacteria is not very well known. A recent study has shown that the loss of m5C in *H. pylori* strains alters the expression of genes involved in motility, adhesion, and virulence. Another study in *E. coli* has shown the role of m5C in stationary phase stress regulation. However, no physiological role of the other form of cytosine methylation (m4C), aside restriction protection is known in bacteria. Genome sequences of various strains of *H. pylori* reveal an abundance of R-M systems. Typically 25-34 R-M

systems are present in different *H. pylori* strains. Methylome analysis of *H. pylori* 26695 strain has revealed the presence of numerous m6A and m5C methyltransferases. *H. pylori* 26695 strain harbors a phase variable type IIS HpyAII R-M system. This R-M system is composed of two exocyclic methyltransferases, M1.HpyAII (m6A) and M2.HpyAII (m4C) and one type IIS phase variable endonuclease (HpyAII). HpyAII recognizes the sequence 5' GAAGA 3' / 3' CTTCT 5' and cleaves eight bp downstream on the top strand and seven bp downstream on the bottom strand. HpyAII is a novel phase-variable restriction endonuclease containing multiple repetitive stretches of adenine residues in the ORF. M1.HpyAII methylates the final adenine residue of GAAGA sequence, whereas M2.HpyAII methylates the first cytosine of the complementary TCTTC sequence. M2.HpyAII is the only N4-cytosine (m4C) MTase present in *H. pylori* strain 26695.

The aim of the present study is to understand the potential epigenetic role of m4C modification by understanding the roles of HpyAII R-M system in virulence, gene expression and natural transformation of *H. pylori*. Understanding the biochemical properties of the novel phase variable HpyAII endonuclease can reveal critical information about the regulation of natural transformation in *H. pylori*.

Bioinformatics analysis shows that HpyAII is an HNH catalytic motif containing endonuclease. The biochemical study on HpyAII indicates that the enzyme prefers two-site substrate over a one-site substrate for maximal activity. A strong preference for two-sites was observed with supercoiled plasmid and oligonucleotide duplex DNA. Cofactor analysis revealed the preference of R.HpyAII for transition metals (Ni^{2+}, Cd^{2+}, and Co^{2+}) over alkaline earth metals (Mg^{2+}, Ca^{2+}) for maximal cleavage activity. Mutational analysis of the conserved residues of the HNH motif in HpyAII confirmed the presence of functional HNH motif. Interestingly, mutation of first His residue (general acid) of the HNH motif to Ala does not abolish the enzymatic activity but instead causes loss of fidelity compared to wild type HpyAII. The H328A mutant displayed promiscuous DNA cleavage activity on different DNA substrates. The novelty of this observation lies in the fact that mutation of first His residue (general acid) of the HNH motif in other known HNH motif containing enzymes

has always abolished enzymatic activity. Mutation at a single amino acid residue leading to the loss of fidelity provides insights into the regulation of fidelity and evolution of restriction enzymes by point mutation.

Previous studies have shown that HpyAII R-M system is flanked by an 80 bp repeat sequence in *H. pylori* strain 26695. A spontaneous recombination event can either delete or restore HpyAII R-M system in *H. pylori* genome. This novel mechanism of regulation is not observed for other R-M systems in any organism. Human migration across the globe has resulted in phylogeographical evolution of *H. pylori* strains into several clusters. Distinct populations of *H. pylori* are present in different geographical locations. Interestingly *H. pylori* strains from Indian population are different from other Asian *H. pylori* strains. Moreover, Indian *H. pylori* strains share similarity with European *H. pylori* strains. Many R-M systems in *H. pylori* are strain-specific and show geographical distribution. The presence of strain-specific R-M systems has been linked to virulence regulation.

To understand the physiological role of GAAGA/TCTTC site methylation, distribution of HpyAII R-M system in Indian *H. pylori* clinical isolates was analyzed. PCR-based analysis revealed that HpyAII R-M system is strain specific in Indian clinical isolates (n= 74). Upon analysis of the symptomatic strains (n=49) from Indian population, it was observed that 16% of the strains are positive for the complete HpyAII R-M system, whereas 40% of the asymptomatic strains from healthy volunteers (n=25) were found positive for HpyAII R-M system. This predominant distribution of HpyAII R-M system in asymptomatic strains might suggest a role of GAAGA/TCTTC site methylation in *H. pylori* pathogenesis and outcome of the disease.

To further investigate this relevant observation, deletion of HpyAII R-M system was carried out in two genetically unrelated *H. pylori* strains. In this study *H. pylori* strains 26695 and I-10 were used. The full genome sequence, transcriptome and global methylome of *H. pylori* 26695 strain (isolated in the UK) is available making it the ideal candidate for in-depth analysis. However, I-10 strain used in the present study was isolated from an Indian patient suffering from duodenal ulcer, and its full genome sequence is not yet

available. Step wise deletion of HpyAII R-M system was carried out in 26695 and I-10 strains of *H. pylori*. Since HpyAII R-M system is a type II R-M system deletion of methyltransferases alone will be lethal to the cell. Sequential deletion of HpyAII R-M system was done by first inactivating HpyAII endonuclease followed by deletion of M2.HpyAII or both M1.HpyAII and M2.HpyAII.

The growth profile of deletions strains were compared to the corresponding wild-type strains, and it was observed that deletion of HpyAII R-M system has no effect on the growth of *H. pylori* strain 26695. However, early induction to stationary phase was seen in the I-10 HpyAII R-M system deletion strain. To understand the *in vivo* role of GAAGA/TCTTC site methylation on gene expression, regions upstream of the transcription-start-site (TSS) were analyzed. It was observed that total 66 GAAGA and 103 TCTTC sites are present upstream of TSS in 61 and 99 ORFs respectively. Genes involved in metabolism, natural transformation, transcription, DNA replication, ion-transport, and virulence harbor the GAAGA/TCTTC sites upstream to the TSS.

Natural transformation is linked to genome diversification, virulence and persistent colonization of *H. pylori* in the mouse. Natural transformation capacity of the HpyAII R-M system deletions strains was compared to wild type strains. The deletion strains of *R.hpyAII* did not show any effect on the uptake of A128G point mutated (*rpsL* gene) DNA in either 26695 or I-10 mutant strains. However, deletion of HpyAII R-M system reduced the natural transformation capacity of both 26695 and I-10 strain. To understand the effect of GAAGA/TCTTC site methylation on *H. pylori*-induced inflammation, coculture of gastric epithelial cells (AGS) with HpyAII R-M system deletion strains was done. HpyAII R-M system deletion resulted in significant decrease in pro-inflammatory cytokine (IL-8) secretion. The deletion also impaired the potential to induce apoptosis and inhibited the cell cycle progression of AGS cells. Lipopolysaccharide (LPS) of *H. pylori* acts as a vital determinant of adhesion and pathogenicity. Decreased expression of O-antigen is associated with low adherence of *H. pylori* in the mouse model. It was observed that in the strain 26695, deletion of HpyAII R-M system leads to altered LPS profile

compared to wild-type. In contrast in I-10 strain, no major alteration of LPS profile was seen between wild-type and mutant strains. To determine the implication of altered LPS profile, *in vitro* cell adhesion assay was performed using AGS cells. It was observed that decreased level of LPS in mutant strain was associated with lower level of adherence. Furthermore, comparisons of the global proteome profile of various deletion strains revealed that the loss of GAAGA/TCTTC site methylation resulted in differential expression of proteins. Results from the present study provide the first evidence that *H. pylori* proteome is regulated by m4C modification on the genome.

These results indicate that in addition to its role in cellular defense, HpyAII R-M system regulates virulence and horizontal gene transfer of foreign DNA in *H. pylori*. This study provides the evidence that HpyAII R-M system maintains the balance between natural competence and restriction, leading to fine regulation of genomic diversity and genomic integrity. These results also suggest that the presence of strain specific R-M system provides strain specific DNA methylation pattern which might result in better adaptation of *H. pylori* to the hostile environment.

Future directions

To under the mechanism of m4C modification mediated regulation of virulence, natural transformation and gene expression, complete transcriptome analysis of the *H. pylori* 26695 and its deletions strains will be performed. The differential proteins spots identified in 2D gel-electrophoresis will be identified to understand the target proteins whose expression is altered by the presence of m4C modification. Complete transcriptome analysis of *H. pylori* strain 26695 and the *M2.hpyAII* deletion strain will help in identification of the target genes which are regulated by the presence of m4C modification. This study will provide the first insights into the role of m4C in bacterial gene expression and virulence regulation.