H. pylori has a panmictic population structure due to high genetic diversity. The homoplasy index for *H. pylori* is 0.85 (where 0 represents a completely clonal organism and 1.0 indicates a freely recombining organism) which is much higher than homoplasy index for *E. coli* (0.26) or naturally competent *Neisseria meningitides* (0.34). It undergoes both inter as well as intra strain transformation. Intergenomic recombination is subject to strain specific restriction in *H. pylori*. Hence, a high homoplasy index means that competence predominates over restriction in *H. pylori*. Annotation of the genomes of *H. pylori* strains 26695 and J99 show the presence of nearly two dozen R-M systems out of which 16 were postulated to be Type II for J99.

H. pylori has been described to be an ideal model system for understanding the equilibrium between competing tension of genomic integrity and diversity (42). R-M systems allow some degree of sexual isolation in a population of competent cells by acting as a barrier to transformation. The mixed colonizing population of *H. pylori* has a polyploidy nature where each *H. pylori* strain adds to 'ploidy' of the colonizing population. Maintenance of polyploidy nature of mixed colonizing population in a selective niche of stomach needs a barrier to free gene flow. Restriction barrier maintains a polyploidy nature of *H. pylori* population which is considered as yet another form of genetic diversity helping in persistence of infection. Thus, according to the model proposed by Kang and Blaser, where *H. pylori* are considered as perfect gases like bacterial population, transformation and restriction both add to genetic diversity of the organism. Again, restriction barriers are not completely effective, which could be due to cellular regulation of restriction system. Thus, a perfect balance between restriction and transformation in turn regulates the gene flow to equilibrate competition and cooperation between various *H. pylori* strains in a mixed population.

RecA, DprA and DprB have been shown to be involved in the presynaptic pathway for recombination substrates brought in through the Com system. Biochemical characterization of HpDprA, during this study revealed its ability to bind to ssDNA and dsDNA. Binding of HpDprA to both ssDNA as well as dsDNA results in large nucleoprotein complex that does not enter the native PAGE. However, DNA trapped in the wells could be released by the addition of excess of competitor DNA, illustrating that the complex are formed reversibly and do not represent dead-end reaction products. Transmission electron microscopy for SpDprA interaction with ssDNA established that a large nucleoprotein complex consisting of a network of several DNA molecules bridged by DprA is formed which is retained in the well. A large DNA-protein complex that sits in the well has also been observed with other DNA binding proteins like RecA. It has been observed for ssDNA binding protein (SSB) that they

bind non-specifically to dsDNA under low salt condition (20 mM NaCl) in the absence of Mg^{2+}. The non specific binding of SSB to dsDNA was prevented under high salt conditions (200 mM NaCl) or in the presence of Mg^{2+}. HpDprA interaction with both ssDNA and dsDNA was stable under high salt condtion (200 mM NaCl) and in the presence of Mg^{2+} indicating that these interactions are specific. The interaction of HpDprA with dsDNA is significant since dsDNA plays an important role in natural transformation of *H. pylori*. The pathway of transformation by dsDNA is highly facilitated (nearly 1000 fold) as compared to ssDNA. However, dsDNA is a preferred substrate for REases which are a barrier to horizontal gene transfer. This implies that the decision of 'restriction' or 'facilitation for recombination' of incoming DNA might be taken before the conversion of dsDNA into ssDNA. The incoming DNA has been shown to be in the double-stranded form in periplasm and in single-stranded form in cytoplasm. Hence, the temporal and spatial events surrounding endonuclease cleavage remain to be understood. Taken together, these results suggest a very important role of dsDNA in natural transformation in *H. pylori*. Hence, binding and protection of dsDNA by HpDprA is possibly of crucial importance in the success of natural transformation process of the organism.

DprA is characterized by presence of a conserved DNA binding domain. The DNA binding domain adopts a Rossman fold like topology spanning most region of the protein. Rossman fold consists of alternating alpha helix and beta strands in the topological order of β-α-β-α-β. It generally binds to a dinucleotide in a pair as a single Rossman fold can bind to a mononucleotide only. All homologous DprA proteins characterized till date show that in addition of the prominent Rossman fold domain they consist one or more smaller domains. RpDprA consists two more domains other than the Rossman fold domain i.e., N- terminal SAM (sterile alpha motif) domain and a C-terminal DML-1 like domain. SpDprA consist of an N-terminal SAM domain other than Rossman fold domain. While the main function of Rossman fold is to bind DNA, the supplementary domains are highly variable in sequences and functions. For example, the SAM domain in *S. pneumoniae* plays a key role in shut-off of competence by directly interacting with ComE~P. HpDprA consist of an N-terminal Rossman fold domain and a C-terminal DML-1 like domain. Both these domains are found to be prominently α-helical in nature. Amino acid sequence analysis of the protein suggests that NTD is basic and CTD is acidic in nature. NTD is sufficient for binding with ssDNA and dsDNA, while CTD plays an important role in formation of higher order polymeric complex with DNA.

For HpDprA and SpDprA, dimerization site was mapped in Rossman fold domain. Gel filtration data revealed an important observation that HpDprA can exist as a monomer (dominant species at lower concentration) as well as a dimer (dominant species at higher concentration) in solution. However, the exchange between these two forms is very fast resulting in a single peak of elution. Since, HpDprA binds to DNA in dimeric form, the dimer species will be favoured in presence of DNA. Hence, even at lower concentrations HpDprA will be mainly a dimer in presence of DNA. Interestingly, both domains of HpDprA i.e., NTD and CTD were able to form dimers but no higher oligomeric form. On the other hand, HpDprA was seen to form oligomeric forms higher than dimer in gluteraldehyde cross linking assay. The strength of CTD dimer was much lower that NTD dimer, therefore it could be proposed that there are two sites of interaction present in HpDprA - a primary interaction site (N-N interaction) and a secondary interaction site (C-C interaction). The N-N interaction is responsible for dimer formation but further oligomerization of HpDprA necessitates the interaction of two dimers using C-C interaction site.

It was shown that NTD binds to ssDNA but forms lower molecular weight complex. SPR analysis of DprA and NTD – DNA interaction pointed out that deletion of CTD leads to faster dissociation of the protein from DNA. Concomitantly, reduction in binding affinity was observed for both ss and ds DNA upon deletion of CTD from full length protein. These results suggest that CTD does play an important role in interaction of full length HpDprA with DNA. Two possible roles of CTD were proposed by Wang *et al* (2014) group to explain their observation of formation of lower molecular weight complex in absence of CTD. (i) CTD possesses a second DNA binding site but much weaker than site present in NTD. (ii) CTD is not involved in DNA binding but mediates nucleoprotein complex formation through protein – protein interaction. EMSA and SPR analysis with purified CTD protein confirmed that there is no secondary DNA binding site present in CTD. As discussed above, it was observed that CTD can mediate interaction between two HpDprA through C-C interaction. Since the interaction is weaker it is lesser likely to be responsible for dimer formation but in trimer or higher oligomeric form of HpDprA, the presence of N-N interaction will facilitate and stabilize C-C interaction. These observations together bring forward an interesting model for HpDprA – DNA interaction. HpDprA forms dimer through N-N interaction (favourably in presence of DNA) and many HpDprA dimers bind to DNA owing to their high affinity and sequence independent nature of binding. These dimers interact with each other through C-C interaction resulting in higher molecular weight nucleoprotein complex. HpDprA - DNA

complex formation is slower than NTD – DNA complex but the former one is more stable (Fig. 2). According to the above proposed model there are two binding events (DNA – protein and protein – protein) in case of HpDprA – DNA complex formation and hence it would take longer time than NTD-DNA complex formation which involves only one binding event. But the resulting higher order complex with HpDprA – DNA would be much more stable.

NTD is able to offer equally efficient protection from nuclease to ssDNA and dsDNA (Fig. 7). This shows that NTD alone is sufficient to completely coat single molecule DNA. AFM images confirm the difference in binding pattern of HpDprA full length protein and NTD. As can be seen in Fig. 8F, NTD binds a DNA molecule by entirely occupying all the available space but forms nucleoprotein filaments isolated from each other. In contrast to full length HpDprA, which forms tightly packed, condensed, extensively cross linked poly-nucleoprotein complexes, NTD forms much thinner complexes with DNA. In the electron micrographs of SpDprA – DNA complex, extensive cross filament interaction was observed resulting in a dense molecular aggregate. Similar kinds of complexes with DNA were also observed for *Bacillus subtilis* DprA in atomic force microscope images. Thus, it could be proposed that HpDprA binds to a single DNA molecule (single strand or double strand) mainly as a dimer formed through N-N interaction. Such multiple individual nucleoprotein filaments come together and interact with each other through C- C interaction resulting in dense and intricate poly – nucleoprotein complex.

HpDprA is proposed to undergo conformational changes from closed state to open state in presence of ssDNA. In agreement with this, structural transition (resulting in reduction of α-helicity of the protein) was observed in presence of ssDNA. Similar structural transitions were observed for dsDNA indicating possibly a common mode of interaction for both forms of DNA. Further, mutation of the residues shown to be involved in binding ssDNA from crystallographic data, resulted in decrease of binding affinity with dsDNA as well. The fold reduction in binding affinity of dsDNA was lower than that for ssDNA despite that it is obvious that the same positively charged pocket which is primarily involved in ssDNA interaction is also responsible (atleast partially) for binding with dsDNA. However, the residues crucial for interaction with these two forms of DNA may be different.

Both DprA and R-M systems have been shown to have presynaptic role in natural transformation process. While DprA has a protective role, R-M systems have an inhibitory

role for incoming DNA suggesting a functional interaction between them. Results of this study show that HpDprA interacts with dsDNA, inhibits Type II restriction enzymes from acting on it and at the same time stimulates the activity of MTases resulting in increased methylation of bound DNA. This observation is of significance from the view of genetic diversity as the only way a bacterial cell discriminates between self and nonself DNA is through the pattern of methylation. Binding of HpDprA to incoming DNA inhibits its access to restriction endonucleases but not to methyltransferases. As a result DNA will be methylated with the same pattern as that of the host cell. Hence, it no longer remains a substrate for restriction enzymes. HpDprA thus, effectively alleviates the restriction barrier. However, it remains to be understood as to how DNA in complex with HpDprA, while not accessible to REases or other cellular nucleases, is accessible to a MTase? A possible explanation could be that HpDprA interacts with MTase and recruits it on DNA. It has been shown that there is a overlap between DprA dimerization and RecA interaction interfaces and in presence of RecA, DprA-DprA homodimer is replaced with DprA-RecA heterodimer allowing RecA nucleation and polymerization on DNA followed by homology search and synapsis with the chromosome. A similar scenario can be thought for interaction of HpDprA with the MTase.

R-M systems play an important role in protection of genomic DNA from bacteriophage DNA. Hence, downregulation of restriction barrier by HpDprA may not be desirable by host during the entire life cycle. Therefore, the expression of HpDprA, which is ComK dependent and that which takes place only when competence is achieved is noteworthy. In *H. pylori*, DNA damage induces genetic exchange via natural competence. Direct DNA damage leads to significant increase in intergenomic recombination. Taken together it can be proposed that when genetic competence is induced, R-M systems are down regulated to allow increased genetic exchange and thus, increasing adaptive capacity in a selective environment of stomach.

There is an evolutionary arms race between bacterial genomes and invading DNA molecules. R-M systems and anti-restriction systems have co-evolved to maintain an evolutionary balance between prey and predator. Phages and plasmids employ anti-restriction strategies to avoid restriction barrier by a) DNA sequence alteration, b) transient occlusion of restriction sites and c) subversion of restriction-modification activities. DNA binding proteins have been shown to bind and occlude restriction sites. On the other hand, λ Ral protein alleviates restriction by stimulating the activity of Type IA MTases. The observations of MTase stimulation and site occlusion of restriction sites by HpDprA appears to be analogous to anti-

restriction strategies, otherwise employed by bacteriophages. Thus, DprA could be a unique bacterial anti-restriction protein used by *H. pylori* for downregulating its own R-M systems to maintain the balance between fidelity and diversity.

In conclusion, HpDprA has unique ability to bind to dsDNA in addition ssDNA but displays higher affinity towards ssDNA. Binding of HpDprA to DNA results in a compact complex that is inert to the activity of nucleases. A novel site of oligomerization for HpDprA was observed which suggests the role of C-C interaction in inter-nucleoprotein filament interaction. It would be interesting to further study the effects of CTD deletion on the transformation efficiency of *H. pylori*, to understand these mechanisms better. It has been well demonstrated that R-M systems offer a barrier to incoming DNA, but our understanding of the regulation of R-M systems has been poor. While other factors like regulation of cellular concentration of restriction enzymes and conversion of dsDNA into ssDNA might play crucial roles in striking the perfect balance between genome diversity and integrity, one of the factors that regulate R-M systems could be DprA.