Summary

Vibrio cholerae is a gram negative human pathogen responsible for the pandemic and epidemic cholera, a severe diarrhoeal disease. Cholera is still a trouble in developing nations as a result of poor water supplies and sanitation. *V. cholerae* is an inhabitant of brackish and estuarine water throughout the world and is associated with zooplankton and other aquatic flora and fauna in these environments during the interepidemic periods (Heidelberg *et al.*, 2000). To survive in such two dramatically different environment i.e human host and natural reservoir, *V.cholerae* uses a number of coordinated transcriptional regulatory events (Faruque *et al.*, 1998).One of these regulatory systems is the cell density dependent bacterial communication called quorum sensing.

Quorum sensing is used by both gram negative and gram positive bacteria to regulate various physiological events including environmental survival, pathogenesis and transmission. In *V.cholerae* there are three quorum sensing systems working in concerted fashion. These three quorum sensing circuits include CAI-1/CqsS, AI-2/LuxPQ and VarS/A- CsrA/BCD. At low cell density (in the absence of autoinducers), LuxU activates LuxO through phosphorylation. This phosphorylated LuxO in concert with σ^{54} dictates the expression of cascades of small RNAs called quorum regulatory rRNA (Qrr 1-4), which further cause the destabilization of *hapR* mRNA transcript. At high cell density (in the presence of autoinducers), dephosphorylation of LuxO impairs its ability to activate the expression of small RNAs, which in turn promotes HapR expression (Lenz et al., 2004).

HapR is the master regulatory protein of *V.cholerae* quorum sensing system. It regulates a large number of physiological events including protease production, virulence, biofilm formation, natural competence , protozoan grazing etc. by binding to the target genes (Kovacikova & Skorupski, 2002; Matz *et al.*, 2005; Meibom *et al.*, 2005; Silva & Benitez, 2004; Tsou *et al.*, 2009). HapR is homodimeric and it belongs to TetR family of proteins (Ramos *et al.*, 2005). Each monomer consists of nine α helices, the first three helices form the N-terminal DNA binding domain while the rest six helices form the dimerization domain. The HTH motif of the N-terminal DNA binding domain is composed of α2 and α3. The helix (α3) acts as recognition helix and the residues of which interact with the major groove DNA bases(De Silva *et al.*, 2007).

HapR regulates nearly 70 genes by binding in promoter regions which are further categorized into two distinct binding motifs (Tsou *et al.*, 2009). Further analysis has identified certain crucial residues in the N-terminal region whose alteration affect the binding ability of HapR to its cognate DNA. For example, Phe 55 is predicted to form contact with DNA. Alteration of F55A completely abolishes the DNA binding capacity of HapR to the *aphA* promoter which further suggests the importance of this residue for high affinity binding of HapR to the *aphA* promoter (De Silva *et al.*, 2007). A great deal of work has been dedicated to understand various structural and functional aspects of HapR. Being a master regulatory protein, absence of functional HapR will significantly affect various physiological events in *V.cholerae*. But it is interesting to know that the loss of function or natural mutations occur more frequently in HapR (Joelsson *et al.*, 2006; Talyzina *et al.*, 2009; Wang *et al.*, 2011).

Recently, our lab has identified a non functional natural variant of HapR from V. *cholerae* strain V2, O37. Further genetic analysis has revealed that the functional impairment of $HapR_{V2}$ was due to a natural mutation of glycine to aspartate at position 39 that lies in the linker region ($R^{33}GIGRGG^{39}$) connecting α1 and α2. This mutation did not affect folding and assembly of variant $HapR_{V2}$ as confirmed by gel filtration chromatography and circular dichroism but it drastically changed the binding affinity towards its cognate promoters as proved by gel shift assay. Consequently the variant $HapR_{V2}$ could not rescue protease production in V2. Additionally, small/wide angle x-ray scattering analysis showed the change in global structure of variant $HapR_{V2}$ from the functional HapR without changing in global energy significantly. This is the first structural evidence on the critical contribution of a linker glycine residue in quorum sensing related DNA binding regulatory proteins of *Vibrio* spp. Thus, sequences in a linker region play significant role for proper functioning of the protein. Some reports have shown comprehensively for the LacI/GalR bacterial family that even substitution of similar residues within the linkers may have dramatic functional consequences (Tungtur *et al.*, 2011). Thus, we can conclude that evolution has optimized the linker composition for efficiency.

Besides linker composition, some literature has also suggested the importance of linker length in protein stability, inter-domain interactions and DNA binding. Accordingly, we proceeded to observe the consequences of alteration of linker length by introducing glycine residues at the C-terminus of this linker ($R^{33}GIGRGG^{39}$) of HapR. Functional analysis of all the linker variants of HapR i.e G39+1, G39+2, G39+3 and G39+4 showed gradual

ıpairment in terms of protease production. Gradual increase in spacing between α1 and α2 / inserting glycine residues did not show any difference in secondary structure contents and sembly but it resulted in gradual loss in affinity of DNA binding. Further analysis by olecular dynamics study has revealed that there was a gradual shift in position of α3 in onformations of linker variants affecting the interactions of amino acid residues including 55 with the major groove DNA bases. Therefore, in accordance with the previous bservations, not only linker composition but also linker length contribute significantly in naintaining HapR, a well functioning master regulatory protein in *V.cholerae*.

Since the glycine rich linker region (R^{33}GIGRGG^{39}) lies entirely in N-terminal DNA binding domain, our investigation on linker composition was also extended to $F^{58}PT^{60}$ linker connecting N-terminal DNA binding domain and C-terminal dimerization domain of HapR. F^{58} is highly conserved among the members of entire TetR family proteins while the P^{59} and T^{60} showed variations and are replaced by different amino acids among the members. Intriguingly, substitution of F^{58} with a small and flexible glycine resulted in loss of function failing to rescue protease production. In contrast to this observation, there was no significant difference in protease production in mutants of P^{59}G and T^{60}G in comparision with wild type HapR.

OpaR, homologue of HapR, is a master regulatory protein of *V.parahaemolyticus* quorum sensing system and it regulates switching from translucent to opaque morphology in this organism. OpaR has 72% identity with HapR and it is also a member of TetR family of proteins. Being a HapR homologue, we also investigated the effect of alteration of glycine rich linker length in OpaR. Like HapR, substitution of glycine^{38} to aspartate in OpaR corresponding to glycine^{39} to aspartate in HapR could not rescue protease production in *V.cholerae*. Surprisingly, alteration of glycine rich linker length (R^{32}GIGRGG^{38}) of OpaR by introducing glycine residues at the C-terminus showed dramatically different result from HapR. Unlike HapR, there was no significant different in protease production of OpaR linker variants having nine extra glycine residues in a row(G38+1, G38+2, G38+3, G38+4, G38+5, G38+6, G38+7, G38+8 and G38+9. Although there is a conservancy of glycine rich linker (R^{32}GIGRGG^{38}) in OpaR, the consequences resulted from the alteration of linker length was profoundly different from HapR. OpaR can accommodate more number of glycine residues at this region without affecting its function in rescuing protease production. In order to

understand the mechanism of adjustment of extra residues in the linker region, crystal structure of OpaR is highly essential.

Apart from analysis of structure function relation of HapR and its homologue in *V.cholerae*, we endeavored to identify two more natural variants of HapR from the same serogroup O37 of *V. cholerae* strain S7 and SC130. Sequence analysis revealed the presence of stop codons in HapR from both the strains at different position and the sequences were deposited to gene bank under accession numbers JX503929 and JX503930 respectively.

Finally, we also studied the differential growth pattern of El Tor variants or hybrid strains in carbohydrate rich media. *V.cholerae* is highly sensitive to acidic pH. El Tor strains having the the capability to metabolize carbohydrate into neutral products like acetoin and 2,3-butandiol can avoid lethal acidification. Consequently, it enhances the survival of El Tor in the environment where *V.cholerae* forms biofilm on chitinous substrate. On the contrary, classical strains acidify the medium and hence it inhibits the growth in carbohydrate rich media. Mekalanos and colleagues suggested better growth and survival in the carbohydrate rich media may be a reason behind the replacement of classical strains with El Tor in seventh pandemic. With the advent of El Tor variants or hybrids strains accompanying characteristics of both El Tor and classical strains, there are many reports of these newly emerged hybrid strains contributing to the ongoing seventh-pandemic in different parts of Africa and Southeast Asia. Out of eleven hybrid strains collected from different parts of Southeast Asia, only one showed the similar pattern of growth with prototype El Tor N16961 during growth curve assay in carbohydrate rich media. It produced acetoin to avoid acidification and it grew better in the glucose containing medium enhancing its own survivability. The rest hybrid strains were found to hinder their growth like prototype classical strain O395 by acidifying the glucose supplemented media. However, unlike classical strain O395, we also observed different phenomena where presence of acetoin was found in the culture media of these hybrid strains during growth assay. The acetoin formed during the metabolism of carbohydrate could not neutralize the acid in these strains. Thus, dynamic process of constant evolution leading to the emergence of new strains has been observed frequently in *V.cholerae*.