Summary

V. cholerae, a human pathogen is the known causative agent for severe diarrheal disease, cholera. Cholera remains a persistent problem for the developing and under-developed countries due to their poor sanitation and hygiene practices. This disease is known to be associated with rice watery stools, hypovolemic shock and even death in many cases. *V. cholerae* is a Gram negative bacterium that belongs to Vibrionaceae family and found to be responsible for causing many pandemic and epidemic cases of cholera. When conditions are not favorable for its survival in human host (interepidemic period), this bacterium prefers to reside in its natural aquatic habitats like fresh, estuarine and brackish water. In such aquatic environments, *V. cholerae* is found either in free living form or associated with biotic substrates (copepods, crustaceans and zooplanktons) and abiotic substrates (surface of rocks and sediments). For its successful survival and pathogenesis, *V. cholerae* has to shuttle between two distinct environments of human host and natural reservoirs to coordinate its transcriptional regulatory events through the process of quorum sensing.

Quorum sensing is a process of cell density dependent bacterial communication that relies on the production and detection of signaling molecules known as autoinducers. This phenomenon is known to exhibit by both Gram negative as well as Gram positive bacteria. The quorum sensing network of *V. cholerae* comprises of three systems (CAI-1/CqsS, AI-2/LuxPQ and VarS/A-CsrA/BCD) that functions in a collaborative manner to regulate various physiological events. At condition when cell density is low and autoinducers are absent, LuxU dictates the activation of LuxO through phosphorylation. Soon after its activation, LuxO works in conjunction with σ^{54} to control the expression of cascade of small RNAs known as quorum regulatory rRNA (Qrr 1-4). These Qrr (1-4) further cause the destabilization of *hapR* mRNA transcript. Howbeit, the condition

reverses at high cell density as high concentration of autoinducers leads to destabilization of LuxO. As a result, the expression of small RNAs is restricted which in turn promotes *hapR* expression (Lenz *et al.*, 2004).

Proteins HapR and LuxO remain the two main key players of *V. cholerae* quorum sensing system as HapR regulates the physiological processes at high cell density and LuxO at low cell density. HapR is a homodimeric protein and it belongs to TetR family of proteins (Ramos *et al.*, 2005). The crystal structure of HapR has detailed that each monomer composed of nine α-helices where first three helices form the N-terminal DNA binding domain and rest six helices form the dimerization domain. HapR has been referred as a master regulatory protein as it regulates the expression of nearly 70 genes in *V. cholerae*. Moreover, HapR can act as both transcriptional activator as well as repressor. HapR represses the production of primary virulence factors and biofilm development. Conversely, it stimulates the production of hemagglutinin/protease (Zhu *et al.*, 2002), promotes the chitin induced competence, increase resistance to protozoan grazing and enhances the survival against oxidative stress (Meibom *et al.*, 2005; Ng and Bassler, 2009; Matz *et al.*, 2005; Joelsson *et al.*, 2007).

Previous work from lab has accentuated the importance of natural variants in exploring various structural and functional aspects of proteins, HapR and LuxO. In a continuing effort, we wanted to instigate our work employing the natural variants for master regulatory protein, HapR. In 2010, Wang and colleagues has examined 602 naturally quorum sensing defective strains isolated from clinical and environmental samples of China. These variants were reported to have frameshift mutations, point mutations and multiple mutations in HapR. Our interest was inclined towards the variants with mutations residing in the dimerization domain. In our study, we investigated the

molecular basis of functional inertness of a HapR natural variant harboring a substitution of lysine in place of a conserved residue glutamate at position 117 ($E^{117}K$). According to crystal structure, this residue plays an important role in dimer formation by maintaining the proper orientation as well as stability of dimer interface. Our results indicated that this mutation did not affect the secondary structure of variant $HapR_{V2G}$ and maintained its dimeric state as confirmed by circular dichroism and gel filtration chromatography respectively. However, variant $HapR_{V2G}E^{117}K$ failed to bind the promoter regions *aphA*, *hapA* and *VC0900*, thereby denoting a compromise in DNA binding ability. Western blot analysis revealed that both wild type and variant HapR were stable *in vivo* but variant $HapR_{V2G}E^{117}K$ failed to rescue protease production in a *V. cholerae* protease negative strain. A combined *in vivo* and *in vitro* analysis indicated a compromise in the DNA binding ability of variant HapR but the cause for its molecular inertness was unclear. Therefore, to gain some further insight in the global architecture of variant protein, SAXS analysis was performed which revealed that mutation E117K introduce a deformity in the DNA binding domain of HapR. This deformity resulted in a deviation of its typical "Y" shape structure with DNA binding domain somewhat "fell apart" in variant as compared to its wild type protein structure. Therefore, it can be concluded from our lab studies that any deviation from this "Y" shape may render HapR non-functional.

While working with QS defective natural variants, we came across a classical strain VC014-99 from China, which showed a protease negative phenotype. The sequence analysis of HapR VC014-99 revealed that a deletion of three amino acid residues in the α2 helix of DNA binding domain had rendered it inactive. To identify whether the helix length or the composition was important, we carried out mutational analysis of the three residues $D^{42}IA^{44}$. The mutational analysis included substitution as well as deletion of all the three residues. The functionality of all the variants generated was examined through

protease assay and it was found that the length of the helix was playing an important role in the functioning of HapR rather than the composition of helix. We also performed western blot analysis and observed that all the variants were stable *in vivo*. To further shed some light on the cause for the molecular inertness of variants (HapR$_{V2G}$ D42del, HapR$_{V2G}$ I43del and HapR$_{V2G}$A44del) we tried to perform further experiments but obstructed by protein purification.

In 2007, De Silva and group revealed the crystal structure of HapR$_{\textit{V. cholerae}}$ highlighting the residues that plays an important in maintaining the dimer interface. These residues are important for HapR structure as they are involved in interactions like salt bridge, disulphide bond, water mediated hydrogen bond and hydrophobic contacts. Hence, to evaluate the contributory role of each dimerization domain residue we performed an alanine scanning mutagenesis and checked the effect of each substitution on HapR activity. All the point variants generated did not affect the HapR activity as shown by protease assay. Intriguingly, some combination of substitutions had severely affected the HapR activity. For example, the hextuple variant S7R$_{V2G}$-CERLYL showed almost a 25% drop in the activity, which further increased to 75% drop when activity for septuple and octuple variants was checked. Western blot analysis was performed and all combination of variants was found to be stable *in vivo*. Further, a western blot was performed with non-reducing SDS-PAGE to check the dimeric status of variants. Surprisingly, only the mutant with a double substitution maintained its dimeric status like wild type protein, whereas rest of the variants showed to exist in a monomeric state. Therefore, to validate this analysis we thought of performing MALDI but unfortunate to obtain purified protein for hextuple or septuple variants and hence, proceeded with triple variant (HapR$_{V2G}$ CER). The triple mutant was showing a monomeric form in MALDI analysis which further needs to be confirmed by performing other experiments. The

MALDI analysis alone may not be taken in to consideration as the protein may undergo denaturation while performing MALDI. Essentially, our results underscore the significance of dimerization domain residues and their combinatorial mutational effect on HapR. Further structural details of each variant with subsequent substitution may help in correlating HapR structural change with HapR activity.

Apart from study of variants, protein homologues have also proved to be a useful tool in studying the structural and functional aspects of a protein. To gain further insights in to HapR, we took keen interest in a HapR homologue isolated from a novel species of *Photobacterium*. Genus *Photobacterium* also belongs to the family Vibrionaceae. Being the member of same family, *Vibrio* and *Photobacterium* shares many common features but intriguingly there are certain features that are specific to genus *Photobacterium*. For example, most species of *Photobacterium* are endowed with a remarkable property of sustaining under high pressure, thus known as piezophiles. Therefore, it would be interesting to know how niche specialization plays a role in modulating the genetic content of an organism. Hence, HapR from a novel species *P. marinum*, was amplified, sequenced and compared to the $\text{HapR}_{\textit{V. cholerae}}$. The sequence thus obtained was submitted to gene bank with accession number KF752599.1. The sequence analysis indicated that the maximum conservancy was visible in the DNA binding domain when HapR homologues from both *Photobacterium* and *Vibrio* species were compared. Moreover, the phylogenetic analysis revealed that HapR from both *P. marinum* and *V. cholerae* are distantly related. Although, sharing very less sequence similarity with $\text{HapR}_{\textit{V. cholerae}}$, $\text{HapR}_{\textit{P. marinum}}$ was able to rescue protease production when a protease negative strain of *V. cholerae*, used as a model organism. As evident from the previous data that Phe 55 is critical for the functioning of $\text{HapR}_{\textit{V. cholerae}}$, the corresponding residue in *P. marinum* was found to play an equivalently important role as substitution of this residue severely

affected the HapR$_{P.\ marinum}$ function. Interestingly, the purified proteins of HapR$_{P.\ marinum}$ and its variant F54A when run on a 12%SDS-PAGE showed an aberrant molecular weight. This discrepancy in molecular weight may attribute to factors like protein tertiary structure, SDS loading on protein, increase of net negative charge on protein, high content of proline in protein and post translational modifications like phosphorylation. In order to examine the molecular weight of proteins (HapR$_{P.\ marinum}$ and HapR$_{P.\ marinum}$ F54A), we performed MALDI and analysis specified that their molecular weight were in accordance with molecular weight of HapR$_{V.\ cholerae}$ and both the proteins maintained their dimeric status. Therefore, it can be concluded that during the course of evolution, functionally or evolutionary related proteins maintain the active core of protein by conserving important residues required for function.

Additionally, we have also studied about the low cell density master regulator, LuxO. LuxO belongs to the NtrC family of proteins and required to regulate a number of processes at low cell density. It comprises of N-terminal dimerization domain and C-terminal DNA binding domain separated by a AAA$^+$ATPase domain. Till date the crystal structure of this protein is not available and hence, not much is known in detail. Taking this as an opportunity, we aimed our study on the DNA binding domain of LuxO. To decipher the critical residue of this domain we performed an alanine scanning mutagenesis and checked the activity of all the variants by protease assay. After analysis it was clear that the residues can be categorized into three classes. All these residues severely, moderately or not at all affected LuxO function. But interestingly, the residues which severely affected the LuxO activity were not completely conserved among the *Vibrio* species. Moreover, residue R413 was not conserved among the species but played a major role in activity. This piece of information can further be supported by *in vitro* study and can be useful for delving in to the structural and functional details of LuxO.

Collectively, this study identified two natural variants of HapR$_{\textit{V. cholerae}}$ containing mutations in their dimerization domain and DNA binding domain respectively, that resulted in a defective quorum sensing circuit. Taking in to consideration, the epidemiological significance of HapR, it would be interesting to observe how these mutations attribute in the pathogenesis of *V. cholerae*. Moreover, these natural quorum sensing mutants and novel homologues would provide a tool for exploring different structural aspects of HapR protein.