Chapter 4: Discussion

4.1. Bi-layered regulation of horizontally acquired genes

Inappropriate expression of horizontally acquired genes needs to be controlled by transcription regulators to maintain the fitness of the organism. Presence of multiple regulators of horizontally acquired genes shows the necessity of tight regulation of the acquired genes [1]. H-NS is a core regulator of horizontally acquired genes [2]. Other co-regulators of H-NS, such as StpA, Hha and YdgT have been reported to interact with H-NS and modify its function [3]. StpA has been shown to substitute the function of H-NS by regulating the H-NS targets in various microorganisms such as E. coli [4] and Shigella sp [5]. In Salmonella sp, a missense mutant variant of StpA has been reported to functionally substitute H-NS [6]. Chapter 2 of this thesis investigates the importance of StpA, Hha and YdgT in the gene regulatory function of H-NS and also explores the selection forces that drive H-NS, and StpA to select its targets. Under the conditions used for testing, we didn’t find any significant contribution of Hha and YdgT towards the gene regulation of the H-NS regulon. Whereas, we observed that StpA-acts as a backup of H-NS to regulate a subset of H-NS regulon. We classified the genes of H-NS regulon into two classes, based on the ability of StpA to regulate these genes i. epistatic genes and ii. unilateral genes. Epistatic genes are a set of genes within the H-NS regulon, that are regulated by StpA in the absence of H-NS, whereas, unilateral genes are not regulated by StpA in the absence of H-NS. These results demonstrate that the back-up function of StpA is limited to epistatic genes. Our results agree with a chip-ChIP study, which reports the reduction of two-thirds of StpA binding sites in the cells lacking H-NS (Δhns) [4]. Controlled over-expression of StpA in the hns knockout results in repression of the pap,bgl and proU operon, thus adding evidence to its role as a backup of H-NS [7]. Though StpA is over-expressed in Δhns, StpA confers only a partial backup, as its level in the cell is significantly less than that of H-NS, also the absence of H-NS leads to rapid degradation of StpA by the Lon protease enzyme [7], [8]. In order to account for such a limitation only a subset of H-NS targets are kept repressed by StpA in the absence of H-NS.

4.2. Factors that drives H-NS and StpA to selects their targets

We further analyzed our data to understand, what makes StpA to select ‘epistatic genes’ over ‘unilateral genes’ when the system lacks H-NS? Detailed analysis of the epistatic genes brought the following features to light i. the ability of these genes to undergo higher transcription in the absence of the gene silencing system– we call it ‘higher transcribability’ ii. the presence of densely packed H-NS binding motifs within these genes and iii. the dispensability of these genes. The presence of these features makes epistatic genes good candidates for repression by StpA, in the cells lacking H-NS.
Epistatic genes are genes with higher transcribability: In the wild type E. coli epistatic genes are located in the ‘transcriptionally silent region’ of the genome, where transcription is repressed by the silencing system. Removal of the silencing system makes these regions ‘transcriptionally amenable’ thus favoring the transition from transcriptionally ‘OFF mode’ to ‘ON mode’. Such transitions, not only always leads to the accumulation of full length functional transcripts, but also might lead to the production of the pervasive transcripts, which are unusual transcripts of very rare function, these are most often produced as a result of antisense and intragenic transcription [9],[10]. Accumulation of pervasive transcripts negatively affects the fitness of the cell [10]. As de-repression of epistatic genes disturbs the gene expression homeostasis of the cell to a higher extent than that of unilateral genes, high transcribability of epistatic genes possibly acts as a selection force, which drives StpA to prefer silencing epistatic genes over unilateral genes.

Epistatic genes are enriched with densely packed high affinity binding motifs of HNS: Sequence analysis of the epistatic genes revealed the presence of H-NS binding motifs in high density. Many promoters of the H-NS regulon tend to have high density of H-NS binding motifs [11] and H-NS is known to suppress the intragenic promoters [9]. H-NS is shown to block pervasive transcription from such spurious promoter by blocking RNAP binding to these promoter, thus facilitating RNAP binding to the specific promoter to make functionally meaningful transcripts [12]. From these observations, we speculate that the absence of H-NS and StpA may increase the possibility of misreading these motifs as a promoter to initiate transcription, which results in pervasive transcription from intragenic transcription, hence StpA selects these targets for silencing in the absence of H-NS.

Dispensability of the epistatic genes: Another feature of the epistatic genes, that confers StpA selectivity to these genes is the dispensability of these genes. Many of the features of the bacterial evolution have been selected based on the essentiality of genes [13],[14]. For instance, gene strand bias in B. subtilis and E. coli has been shown to be determined by the essentiality of the genes. The presence of coding sequence on the lagging strand results in the truncated mRNA (often non-functional) as a result of head-on-head collision of DNAP and RNAP. Thus a majority of essential genes are present in the leading strand of the DNA [13]. The class of epistatic genes largely consists of non-essential genes, which are dispensable for the survival of the organism under the conditions used. Expression of dispensable genes could have adverse effect on the growth and survival of the organism, as the transcription and translation of the dispensable genes are expensive [11],[15]. Moreover, distribution of the energy and resources to the expression of the unwanted products compromises the fitness of the cell. Epistatic genes are maintained in the genome but silenced, as the cell might require these gene products under certain conditions. Some of these dispensable genes are lost, as soon as the cell lacks H-NS. For instance, in the absence of functional H-NS, excision of Rac prophage was observed [16]. Whereas, some genes are still maintained in the genome and subjected to basal level silencing by
StpA. Thus xenogene-silencing system in bacteria has evolved carefully to silence targets whose de-repression disrupts the overall gene expression homeostasis of the cell, whereas, subset of genes of H-NS regulon, whose de-repression does not have much impact on the overall expression homeostasis of the cell, are not silenced by StpA in the system lacking H-NS.

### 4.3. Importance of having bi-layered regulation of xenogene silencing

Horizontal gene transfer by transformation can occur very often in nature [17]. In response to such acquisition, the host cell utilizes the available pool of transcription factors to regulate the expression of the acquired genes. We speculate that the organism has evolved with multiple layers of regulators to confer an instant response, in situations where the cell has lost one of the regulators or when the regulator level is limited in the cell. For instance, in the *Salmonella* sp, the strain carrying the pSF-R27 plasmid has its own H-NS like protein called Sfh coded in the plasmid, which regulates the AT-rich regions on the plasmid. Deletion of Sfh gene from the plasmid results in a phenotype comparable to that of Δ*hns*, since the genomic H-NS is titrated out to silence the AT-rich regions on the plasmid, which results in an insufficient pool of H-NS to silence its’ genomic counterparts [18], suggesting that, maintaining the cellular pool of transcription factors is necessary for the cells optimum growth and gene regulation of the cells [11]. In case of *E. coli*, horizontally acquired genes are under bi-layered regulation by H-NS and StpA, in the absence of H-NS, StpA can partially back up H-NS to regulate some of the H-NS targets, which have higher transcribability. However, in the absence of StpA, H-NS can regulate almost all the targets of StpA. Hence no significant difference in gene expression and growth phenotype was observed in ∆*stpA* [15]. Thus, in case of *Salmonella* sp, deletion of *hns* itself results in defective growth and it can be compensated by the single mutation in *stpA* [6]. While, high transcription level of the gene is considered as a barrier of horizontal gene transfer [19], acquired genes with higher transcribability are fixed in the genome, only with the aid of the tight regulation by this bi-layered regulation, else, it will be lost from the genome through negative selection.

### 4.4. Adaptive mutations that compensate for the loss of *hns* and *stpA*

Any organism, which is encountering serious threats to its survival or to its fitness, tends to gain mutations to compensate for the loss of fitness caused by the loss of important genes, acquisition of new genes, environmental stresses etc [20]. The severe disruption of gene expression homeostasis in Δ*hns-stpA* is probably responsible for the growth defect. We asked if the cell lacking xenogene silencing system can evolve to rescue from its growth defect. Loss of xenogene silencing system resulted in up-regulation of genes around the terminus (region enriched with the H-NS targets), such up-regulation of the genes around the terminus could limit the availability of the transcriptional machinery to the opposite proximal of the genome called ‘origin or ori’, resulting in down regulation of genes around the origin. Interestingly, we observed diverse evolution strategies adapted by
E. coli- K12 MG1655 lacking ‘xenogene silencing system’ to compensate for the loss of main global regulators. We observed transient amplification of about 40% of the genome and several point mutations that inactivate RpoS - a stationary phase sigma factor. These two diverging strategies adapted by the organism, showed a converging effect on gene expression. This shows that the compensatory mutations are gained to bring the gene expression level to its optimum, which in turn contributes to maintain the fitness of the cell.

Compensatory mutations which arises in response to loss of gene or its function does not necessarily occur in the gene belongs to the functionally related network of the lost gene, rather it can occur in a gene which is totally unrelated to the network [21]. For instance, in T3 bacteriophage, compensatory mutations raised against the loss of DNA ligase, were observed in both functionally related genes involved in DNA metabolism, as well as in the functionally unrelated genes, which code for virion protein [21]. In this study, we observed a point mutation that inactivated RpoS – a stationary phase stress response σ-factor in response to the loss of xenogene silencing system. Inactivation of RpoS has been reported in a variety of stress conditions, including the long-time storage as a stab culture[22]. Such loss of function mutation is one of the prevalent modes of adaptation, which gives fitness benefits to the cell. There are many instances where the loss of function of genes on adaptation has been reported [23],[24]. A possible mechanism behind the fitness advantage conferred by the loss of function of RpoS, in the cells lacking xenogene silencing system has been discussed below.

4.5. Loss of function of RpoS: Adaptive response against elevated RpoS and nutrient scavenging?

RpoS or σ^{38} - stress response sigma subunit of the RNAP, is kept low in the exponential phase of the bacterial growth cycle and its level shoots up when the cell enters the stationary phase or under stress conditions like nutrient deprivation, high or low pH, osmolality change or temperature shock [reviewed in 23]. RpoS level in the cell is controlled at the level of its transcription, translation and also at the level of degradation by proteolysis [26]. H-NS is shown to directly or indirectly regulates the expression of RpoS at multiple levels [26], as demonstrated by the fact that RpoS level in the cell is elevated in the absence of H-NS [27]. We didn’t see an increase in the rpoS mRNA in the Δhns-stpA transcriptome. This indicated that RpoS expression is not affected by H-NS at the level of transcription, rather it is regulated at the post transcriptional level [27].

The RpoS protein undergoes rapid degradation in the exponential phase by ClpXP protease pathway. RpoS degradation is facilitated by the binding of RpoS to the RssB- an adapter protein, which exposes RpoS to ClpXP proteases [28]. H-NS indirectly affects RpoS level by inhibiting the expression of anti-adapter proteins such as IraD and IraM [28]. In the absence of H-NS and StpA, anti-adapter proteins are de-repressed resulting in the stabilization of RpoS [28]. So, in the absence of H-NS and StpA, RpoS level in the cell is
elevated, which causes gene expression imbalance in the cell. High RpoS stability leads to the induction of RpoS (σ38) target genes, which limits the availability of the RNAP to transcribe the σ70 dependent targets [29]. Over-expression of RpoS and its target genes have been reported to be involved in nutrient scavenging [30],[31]. This could explain, why the spent media of Δhns-stpA collected at the mid exponential phase of growth had failed to support the growth of wild type and addition of nutrient supported the growth (Annexure-II Figure 12), while, the spent media collected from the wild type could support the growth of the cell. From these results we suspect that, rapid nutrient scavenging in the Δhns-stpA as a result of overexpression of RpoS and its targets could be responsible for the growth defect experienced by Δhns-stpA. As the RpoS degradation is dependent on ATP (Adenosine triphosphate), any nutrient deprivation influence the level of RpoS in the cell by affecting the degradation of RpoS by ClpXP pathway, thus resulting in elevated level of RpoS [32]. This is a complex phenomenon, as how the sequence of above-mentioned events leading to increased RpoS level in the cell is not yet clear. Perhaps, to eliminate the possibility of homeostasis disruption caused by the elevated RpoS level, adaptation leads to inactivate RpoS in the cells.

4.6. Amplification: An instant response to the stress caused by loss of xenogene silencing system

All organisms respond to stress by employing various adaptation strategies. Amplification is one of the prevalent adaptive strategies that most of the organisms opt in response to both internal stress caused by the loss of genes or function and the external stress caused by the environmental factors such as nutrient deprivation, temperature etc [31], [32]. One of the adaptive mutations we report here in our study is the amplification of about ~2 Mb of genome around the origin of replication. Importance of amplifying the origin is to compensate for the down-regulation of genes around the origin, which is an indirect response to the de-repression of H-NS targets around the terminus, as a result of deletion of xenogene silencing system. Recombination between the insertion elements insC1 and insC5 - the boundaries of the amplification regions sets the boundary to separate the ori and ter macro domains. This kind of quick rescue mechanism by amplification of particular gene [33] ,[34] or a part of genome [6] has been reported in wide variety organisms like in E. coli [33], [35], Salmonella sp [7], yeast [36],[37] and in plants [38]. Amplification is better over the point mutation in some aspects as it is reversible [39], not affecting the original copy of the gene. Amplification of the large part of the genome is one of the transient and instant rescue mechanism which the organism adapts to get out of the stressful condition [36], [39]. Though amplification is one of the first aid mechanisms to rescue the cells under stress, it is transient as the cost of maintaining the duplicated gene could be relatively costlier than the point mutations that arise at the later stages of evolution. Both adaptive point mutations and the amplification are parallel mechanisms of adaptation, occurrence of one is independent of the other [40]. Most of the studies that have reported the amplification of genes as a ubiquitous strategy of adaptation, have also reported the instability of amplification. Instability of the amplification is due to the
deletion of the regions between the repeats [41], either by recombination dependent or by independent manner [41], [42]. In the absence of selection or when the occurrence of beneficial point mutation, the cost of maintaining the amplified gene or genome is considered to be expensive, as a result the cell returns to the non-amplified state.

### 4.7. Instances of gene amplification

The genes and the regions which are duplicated might or might not have a related function to the stress to which the cell is exposed to [43]. In *E. coli* deletion of ΔserB (a phosphatase gene involved in serine biosynthesis) has been shown to be rescued by the overexpression of non-cognate protein YtjZ (hypothetical protein), which has very limited structural similarity than the cognate counterpart of SerB such as HisB (a phosphatase involved in histidine biosynthesis) [43]. An important cold shock protein coding gene *csp* genes, have shown to be duplicated by multiple rounds and the resulting gene can respond to a variety of stresses like cold shock, nutritional deprivation etc [33]. Amplification of β-lactamase gene was reported as one of the preliminary step in the evolution of *Salmonella typhimurium* in gaining the antibiotic resistance, which is followed by point mutation in the genes such as *cpxZ, envZ* that results in the reduced transport of antibiotic into the cell [34]. Tandem genome amplification was also observed in one of the natural clinical isolates of *Streptococcus sp.* which leading to the amplification of genes involved in the dihydrofolate biosynthesis, which is a target of antibiotics such as trimethoprim and sulphonamide [44]. In yeasts amplification of *metallothionein* responsible for the resistance to metals like cadmium and copper has also been reported [45]. Thus amplification results in increasing the copy number of the gene or the regions to maintain the fitness of the organism by giving resistance to antibiotic and by activating the responses to encounter the unfavorable conditions.

### 4.8. Opposing gene expression trends across macro domains

Most of the essential genes which are involved in the cellular processes like transcription, translation, replication, cell division and post transcriptional modifications, are shown to have high expression level in the cell, are enriched around the origin of replication [46],[47] whereas, the non-essential and the horizontally acquired genes are clustered around the terminus of replication and it is considered as a transcriptionally silent protein occupancy zones [48], whose expression is tightly regulated by xenogene silencing system, mainly by H-NS. The terminus region of the chromosome is also enriched for the genes regulated by RpoS.

We noticed that the genes that are de-repressed in the Δhns-stpA are located around the terminus, which is known to have majority of the H-NS targets and the stress response genes, regulated by RpoS, whereas, genes which are down regulated, are enriched around the origin of replication. This explains the link between the gene expression levels and the chromosome organization. To understand this link between the chromosome organization and the gene expression in a better way, we further analyzed the transcriptome data of the
suppressor mutant obtained by evolution of the strain, which lacks xenogene silencing system. In the suppressor strains the gene expression imbalance is partially recovered, but the up and down regulated genes maintain the opposite proximity as seen in the imbalanced strain. Change in gene expression pattern across two macro domains can be explained by accumulation of the components of the transcriptional machinery at one proximal of the chromosome limits the availability of the same to the other end. In the absence of xenogene silencing system, as there is no exclusion of RNAP by H-NS, RNAP accumulates towards H-NS targets around the terminus, which might result in the indirect down regulation of genes around the origin by limiting RNAP. As the high expression of the genes around the origin is observed in the fast growing cells, such indirect down regulation of the genes around the origin can explain the defective growth of ∆hns-stpA. We further asked whether this kind of opposing gene expression pattern across the chromosomal domains is specific to the condition we tested or it can be generalized. Our analysis of publicly available gene expression data generated from ~300 different conditions (data different studies) explains that the observed trend is conserved across many of the conditions, thus generalizing the opposing gene expression trend across macro domains.

4.9. Conclusion and Future direction

In this study, we investigated how the gene expression homeostasis is affected on deletion of ‘xenogene silencing system – H-NS and StpA’ and its modulators Hha and YdgT ? Additionally, we also investigated how the E. coli lacking xenogene silencing system evolved to compensate for the loss of fitness caused by the disruption of gene expression homeostasis in E. coli? In Chapter 2, we report a bi-layered regulation of horizontally acquired genes by H-NS and StpA. We also show that, StpA selects for the subset of H-NS targets based on the following features to silence in the absence of H-NS i. transcribability of the gene ii. Dispensability of the gene and also iii. Density of the high affinity motifs within the gene.

We have not seen any effect of Hha and YdgT on gene expression under the conditions tested. Nevertheless, recent studies on E. coli and in Salmonella sp , have shown the role of Hha and YdgT on gene expression, especially on the regulation of horizontally acquired genes, under high osmolarity and anaerobic condition [3], [49]. From this, we speculate that the possibility of multi-layered regulation by Hha and YdgT along with H-NS and StpA, we might have missed this possibility, as our experimental system either had Hha or YdgT. The presence of one paralogue partner would have replaced the function of the other. Furthermore, studies with the deletion of both hha and ydgT in the ∆hns-stpA background and other global regulators would help to understand, if the multi-layered regulation of horizontally acquired genes exists under the experimental conditions used in this work.

Chapter 3 of this thesis describes the reversal of gene expression homeostasis in the suppressor strains, which arose during the evolution of cells lacking xenogene silencing system and also reports the opposing gene expression trends across two domains of the
chromosome. This chapter reveals the link between the gene expression status of the cell and the chromosome architecture.

Finally, an interesting observation during this study offered an idea to explore the role of H-NS in the cell division, more specifically on chromosome segregation. Exploring the reason for the filamentous cell morphology of Δhns-stpA observed by us in this study and other studies [7] is of our future interest.

From this observation, we speculate that the growth defect of Δhns-stpA might be due to the perturbation of septum formation as a result of nucleoid occlusion. While studies on H-NS and StpA are mostly exploring the different aspects of gene expression and chromosome architecture, very few studies have investigated the role of H-NS in chromosomal division and in cell division. Though many studies have reported the severe growth defect phenotype of Δhns-stpA, but the precise mechanism behind such phenotype is left unattended.

Division of a single cell into two involves high-level co-ordination between various processes like, chromosome division, chromosome segregation and septum formation. Early studies on chromosomal division have reported the role of H-NS in the initiation of chromosome division and the role of H-NS in reverting the over-initiation phenotype of dnaA mutant [43]. Role of H-NS in the later processes of the cell division such as chromosomal segregation and septum formation is poorly studied. Co-ordination between these two processes is highly mandatory for the proper cell division. Septum formation prior to the chromosomal segregation is blocked by the process called ‘nucleoid occlusion’. Many nucleoid occlusion factors (NOFs) like Noc, MinCD, SlmA etc, which block the septum formation by inhibiting the FtsZ ring formation in its vicinity [50].

An interesting study has reported that the nucleoid occlusion factors are scattered around almost all the regions of the chromosome except for the ‘ter’ region [51]. Overlapping this observation with our study, I hypothesize that enrichment of H-NS around the ‘ter’ region might occlude the nucleoid occlusion proteins in its vicinity, thus leading to nucleoid segregation followed by septum formation. Absence of H-NS and StpA, might employ nucleoid occlusion proteins in the ‘ter’ region of the chromosome that could perturb the cell division by ‘nucleoid occlusion’. Intensive studies on this aspect would help to explore the role of H-NS in the cell division.
4.10. References


[7] B. Sonden and B. E. Uhlin, “Coordinated and differential expression of histone-


R. P. Maharjan, J. Gaffé, J. Plucain, M. Schliep, L. Wang, L. Feng, O. Tenaillon,


