1. INTRODUCTION & REVIEW OF LITERATURE

Bacteria inhabit a variety of ecological niches and adapt to a very dynamic environment. Hence, they have evolved several sensory and regulatory systems that regulate expression of genes in response to a variety of environmental and host generated stimuli. Much of this regulation was thought to be carried out by proteins at the level of transcription (Romby et al. 2006) and translation initiation (Kozak 2005). Findings that antisense RNAs control replication of some plasmids gave the hint that RNAs could also act as regulators (Storz and Haas 2007). For example, a 108 nt long plasmid encoded RNA, namely RNA I regulated ColE1 plasmid replication by base pairing with the RNA, that post its cleavage, acts as the replication primer (Tomizawa et al. 1981). Other examples of *cis*-encoded antisense RNAs were identified in plasmid addiction systems that ensured survival of the plasmid containing cells (Storz et al. 2005). In 1984, the first chromosomally encoded small RNA regulator MicF was reported, that inhibits the translation of the mRNA encoding the major outer membrane porin OmpF in *Escherichia coli* (Waters and Storz 2009). Since this path-breaking discovery several small regulatory RNAs have been discovered in diverse bacterial genera where they help the cell to face environmental pressures by regulating expression of key proteins.

Bacterial small regulatory RNAs (sRNA) or non-coding RNAs (ncRNAs) comprise a very crucial class of regulators, ranging in size from 40-500 nucleotides in length. Earlier they were called as non-coding RNAs as it was believed that they do not code for proteins. However, studies on RNAs like SgrS (Wadler and Vanderpool 2009) and RNAIII (Morfeldt et al. 1995) revealed that small RNAs might as well code for small peptides. Thus, this class of RNA molecules is now called as small regulatory RNAs (Li et al. 2012) and ncRNAs when they do not code for peptides.

Bacterial sRNAs are evocative of the eukaryotic microRNAs (miRNAs) but hold certain points of differences as well (Gottesman and Storz 2011). sRNAs are transcribed as single transcripts and are mostly not processed into smaller transcripts like miRNAs of eukaryotes. In very few cases where they are processed, they do not form transcripts as small as the miRNAs. Both sRNAs and miRNAs need protein scaffolds to be presented to the mRNA targets but the protein machinery is more complex in the case of miRNAs. Most sRNAs that bind mRNAs, interact with their targets in the 5'UTR or in the ORF whereas miRNAs act by base pairing at the 3'UTRs of the mRNA targets (Bloch et al. 2017).

Moreover, in contrast to miRNAs, that act as negative regulators, binding of the sRNAs to their target mRNA results in both positive and negative effect on the stability and/or translation of the transcripts (Choi et al. 2017). In examples of negative regulation, sRNAs might act stoichiometrically and get degraded with their targets (Gottesman and Storz 2011).

1.1 Regulation of gene expression by sRNAs

sRNAs function either by base pairing with other mRNAs or by binding to proteins and modifying their activity (Ternan 2013;Michaux et al. 2014). They modulate various steps of gene expression like transcription, mRNA translation and RNA stability (Levine and Hwa 2008) leading to the regulation of metabolism, growth processes and stress adaptation (Bobrovskyy and Vanderpool 2013;Michaux et al. 2014). Besides these roles, involvement of sRNAs in regulation of bacterial pathogenesis has gained a lot of focus in the last two decades since the discovery of the role of RNAIII, a small regulatory RNA of *S. aureus*, in targeting at least five mRNAs that encode virulence factors (Hébrard et al. 2012). Base pairing between sRNA and mRNA involves a seed region of at least 6-8 contiguous base pairs, although longer seed regions have also been observed (Gottesman and Storz 2011). This pairing leads to various regulatory outcomes depending upon whether the 5'UTR, 3'UTR or the coding region of mRNA is involved in base pairing with the sRNA.

1.1.1 Translational repression: Most of the sRNAs reported till date, bring about regulation of their mRNA targets by translational repression. sRNAs base pair with their respective targets at the 5'UTR occluding the ribosome binding site, preventing the 30S ribosomal subunit from binding to the mRNA and inhibiting translation (Figure 1A). The sRNA-mRNA duplex is then frequently subjected to degradation by RNase E (Waters and Storz 2009). For example, FnrS of *E.coli*, induced under anaerobic conditions, represses expression of genes like *maeA*, *gpmA* (involved in central metabolism); *sodB* (involved in oxidative stress); and *folE* and *folX* (involved in folate metabolism) by the above mechanism (Durand and Storz 2010).

1.1.2 Translational activation: Several examples of mRNA stabilization and translational activation have come up in recent years, changing the common perception about sRNAs as negative regulators of gene expression (Papenfort and Vanderpool 2015). Two mechanisms have been reported to cause translational activation.

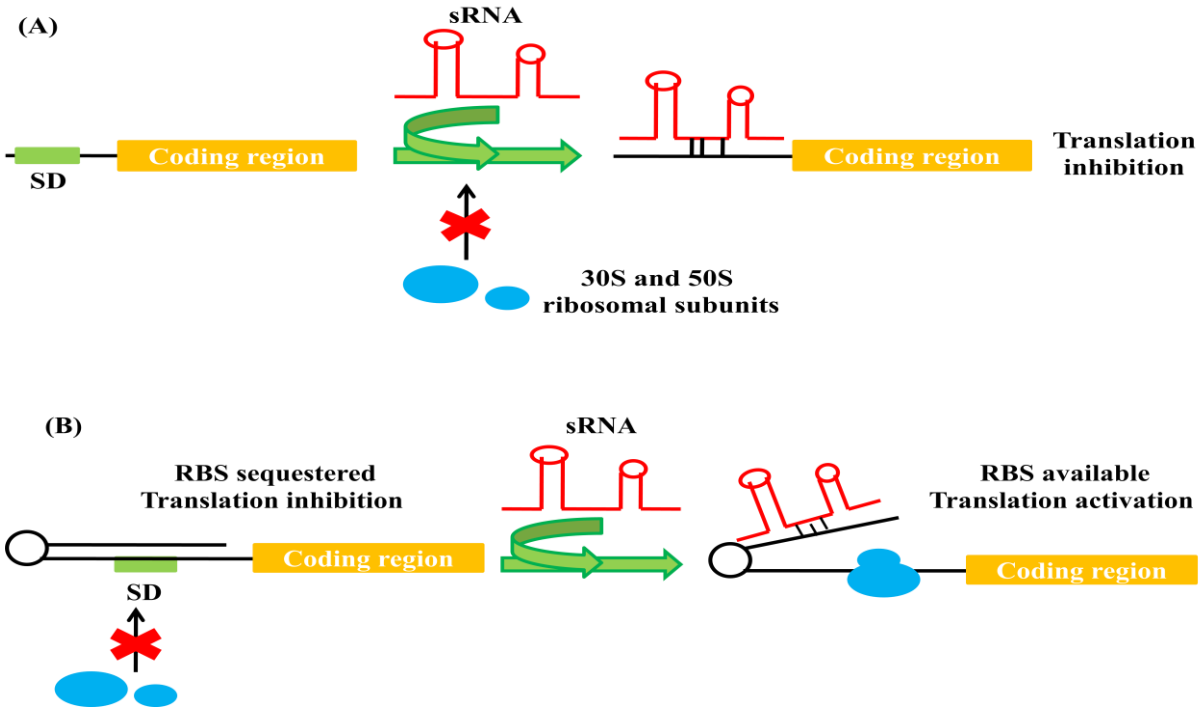

Figure 1: Regulation of gene expression by sRNAs
(A) Translational inhibition: Base pairing of the sRNA (red) with the mRNA at the 5'UTR occludes the Shine Dalgarno (SD) sequence and prevents ribosome (blue) binding and translational initiation. (B) Translational activation: An intramolecular secondary structure formed at the 5'UTR of the mRNA sequesters the Shine Dalgarno (SD) sequence. sRNA (red) binding at the 5'UTR brings about a conformational change making the SD sequence available for ribosome (blue) binding thus leading to translational activation.

In the first mechanism, called as the anti-antisense mechanism, binding of the sRNA to the 5'UTR opens up an intrinsic stem loop formed within the mRNA, improving the access of ribosome to the Shine-Dalgarno sequence (Figure 1B). This leads to enhancement of translation via improved ribosome access to the mRNA. For example, PhrS, in *P. aeruginosa* activates expression of *pqsR* by base pairing with a short **u**pstream leader **o**pen reading **f**rame (*uof*) to which *pqsR* is translationally coupled. The single ribosome binding site (RBS) of *PqsR* and *uof* is partially masked by a secondary structure which opens up when PhrS binds to the *uof* thus allowing ribosome to access the RBS and initiating translation (Sonnleitner et al. 2011). In the second mechanism, sRNA interferes with the ribonucleolytic decay of the mRNA transcript, leading to improved mRNA stability and activation of translation (Papenfort and Vanderpool 2015). Small RNA RydC of *Salmonella enterica*, stabilizes *Cfa* mRNA by base pairing with its 5'UTR and interfering with RNase E-mediated decay in the 5' untranslated region (Fröhlich et al. 2013).

REVIEW OF LITERATURE

1.2 Classification of sRNAs

The bacterial small regulatory RNAs act by two mechanisms: either by base pairing with the mRNAs and affecting their translation and/or stability or by binding to and modifying the activity of proteins. The base pairing sRNAs can further be catalogued into two classes: *cis* encoded and *trans* encoded sRNAs (Liu and Camilli 2010).

1.2.1 sRNAs regulating mRNA targets

1.2.1.1 *cis*-encoded sRNAs: are encoded at the same genetic location as their targets but on the opposite strand and therefore share complete complementarity with their targets and show perfect, continuous base-pairing. Mostly these types of sRNAs occur in mobile genetic elements like transposons, plasmids and phages with a few of them identified in the chromosome also (Bobrovskyy and Vanderpool 2013).

1.2.1.1.1 Plasmid encoded RNAs: Plasmid encoded antisense RNAs have served as models in the study of sRNAs. These base pair with, and regulate the transcription, stability or translation of mRNAs encoding proteins involved in plasmid replication and maintenance. Although they express constitutively, they are metabolically unstable and the changes in the plasmid copy number are reflected in the levels of these RNAs. One such example is of RNAI (~85 nt) and RNAII (~150 nt) of the plasmid pT181. These RNAs base pair with and stabilize a structure responsible for transcription termination of an upstream gene *repC*. When the plasmid number goes down, levels of RNAI and II reduces, leading to a transcription read through and increased RepC levels, in turn initiating the plasmid replication.

The other type of plasmid encoded antisense RNAs are involved in the toxin-antitoxin system (also called as plasmid addiction system), which ensures that only the plasmid containing cells survive and others get killed. The famous *hok-sok* genes of plasmid R1 are representatives of this type. The *hok* encodes a killer protein that makes pores in the cell membrane and causes cell death. The Sok RNA base pairs with the *Hok* mRNA leading to its translational repression. When the plasmid is lost from the cell, Sok RNA degrades faster than the *Hok* RNA due to the differential stability of the two RNAs. This allows translation of *Hok* mRNA and expression of Hok protein which then kills the cell (Storz et al. 2005).

1.2.1.1.2 Chromosomally encoded antisense RNAs: A very few examples of this type of RNAs are reported in literature. Although it is hypothesized that most of the bacterial encoded *cis* RNAs would belong to the prototype of toxin-antitoxin system and lead to translational repression, GadY RNA of *E. coli* is an example of RNA that increases the expression of its *cis*-encoded target mRNA *GadX*. It is seen that binding of GadY RNA to the 3'UTR of *GadX* mRNA leads to increased stability and accumulation of this mRNA in the cell (Storz et al. 2005).

1.2.1.2 *trans*-encoded sRNAs: These sRNAs are located at genomic loci distinct from their mRNA targets and share imperfect, non-contiguous base pairing with each other (Waters and Storz 2009). This imperfect base pairing facilitates one sRNA to base pair with multiple mRNA targets (Waters and Storz 2009) and regulate their translation or stability. The sRNAs bind with their targets with limited complementarity in discontinuous patches and although interaction regions as long as 10-25 nucleotides have been predicted, it has been observed that a continuous base pairing of 6-8 nucleotides is sufficient for regulation (Gottesman and Storz 2011). It is a general observation that the regions of sRNAs that base pair with multiple mRNA targets are highly conserved. For example, SdsR sRNA regulates synthesis of the carbon utilization regulator CRP, the nucleotide associated chaperone StpA, and the antibiotic resistance transporter TolC along with an outer membrane porin OmpD, in *Salmonella enterica* (Frohlich et al. 2016).

trans-encoded sRNAs are known to bring about negative regulation of their mRNA targets and reduction in protein levels by translation inhibition and/or mRNA degradation (Gottesman and Storz 2011). Binding of the sRNA at the 5'UTR of the mRNA causes occlusion of the ribosome binding site leading to the inhibition of translation and reduction in the protein levels (Figure 1A). For example, Qrr4 sRNA represses type VI secretion system (T6SS) in *V. cholerae* by base pairing at the 5'UTR of the large T6SS cluster (Shao and Bassler 2014).

Several sRNAs especially in gram negative bacteria are found to function in coordination with the Sm-like Hfq protein. Hfq facilitates base pairing between the sRNA and the mRNA which leads to translational inhibition and destabilization of the untranslated mRNA exposing it to the RNases. Although sRNA mediated translational inhibition is sufficient for gene silencing, degradation of the translationally inactive mRNAs makes the silencing irreversible (Morita et al. 2006).

Although several bacterial sRNAs were shown to act as negative regulators, an increasing number has been identified, that activates translation and promotes stabilization of mRNA targets. These sRNAs base pair at the 5'UTR, coding region or at the 3'UTR of the mRNA targets; and function in coordination with Hfq protein and cellular ribonucleases. Translational activation can be carried out by two mechanisms: the first one called as the anti-antisense mechanism and the second one involving interference with the ribonucleolytic decay (Papenfort and Vanderpool 2015). In the anti-anti sense mechanism, base pairing of the sRNA to the mRNA opens up an intramolecular inhibitory stem loop within the target mRNA that was sequestering the Shine-Dalgarno sequence. Opening of the stem loop allows ribosome binding and translation initiation (Figure 1B).

RydC sRNA, which is conserved in few of the enteric bacteria like *Escherichia*, *Shigella*, *Klebsiella*, *Salmonella*, *Enterobacter*, and *Citrobacter*, activates the expression of *cfa* gene and increases protein production by interfering with the ribonucleolytic cleavage of the *Cfa* transcript. RydC activates Cfa synthesis by base pairing at the RNase E cleavage site in the 5'UTR of the mRNA thus preventing its decay and in turn stabilizing it (Figure 2) (Fröhlich et al. 2013).

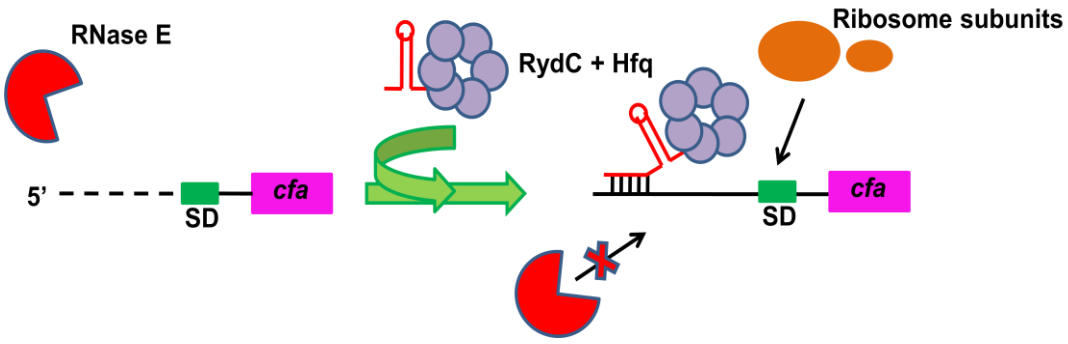

Figure 2: Translational activation by interference in the ribonucleolytic activity.
RNase E cleavage of the *Cfa* mRNA prevents the synthesis of Cfa protein. RydC sRNA binding at the 5' end of *CfA* protects it from RNase E degradation and leads to ribosomal binding and translation initiation.

1.2.2 sRNAs that modify protein activity

These sRNAs act by binding with proteins that are post transcriptional repressors of mRNAs. Sequestration of these proteins by sRNAs, makes them unavailable to bind with their natural mRNA targets thus allowing translation of these mRNAs (Gottesman and Storz 2011). CsrA

(**c**arbon **s**torage **r**egulator), a 61 amino acid dimeric protein of *E. coli* with homologs in other bacteria, regulates carbon utilization, glycogen synthesis, motility and biofilm formation by regulating expression of mRNA targets. CsrA has a high affinity towards the GGA motifs in the 5'UTR of its mRNA targets. Once bound to the mRNA, it prevents ribosome binding and inhibits translation initiation. CsrB/C sRNAs of *E. coli* with 18 GGA motifs in their sequence sequester the CsrA protein molecules, in turn activating translation of the downstream mRNA targets (Figure 3) (Waters and Storz 2009). Homologs of CsrA in *Erwinia carotovora* (RsmA) and in *Pseudomonas fluorescens* (RsmA and E), in conjunction with RsmX/Y/Z sRNAs, control the secretion of extracellular metabolites including pathogenicity factors (Jensen et al. 2006). CsrA and CsrB/C/D sRNAs regulate quorum sensing in *Vibrio cholerae* (Lenz et al. 2005). CrcZ sRNA of *P. aeruginosa* binds the Crc (**c**atabolite **r**epression **c**ontrol) protein with its five CA motifs thus releasing the catabolite repression of the degradative enzymes when grown on non-preferred carbon sources like mannitol (Sonnleitner et al. 2009). The classification of the sRNAs has been broadly summarized in Figure 4.

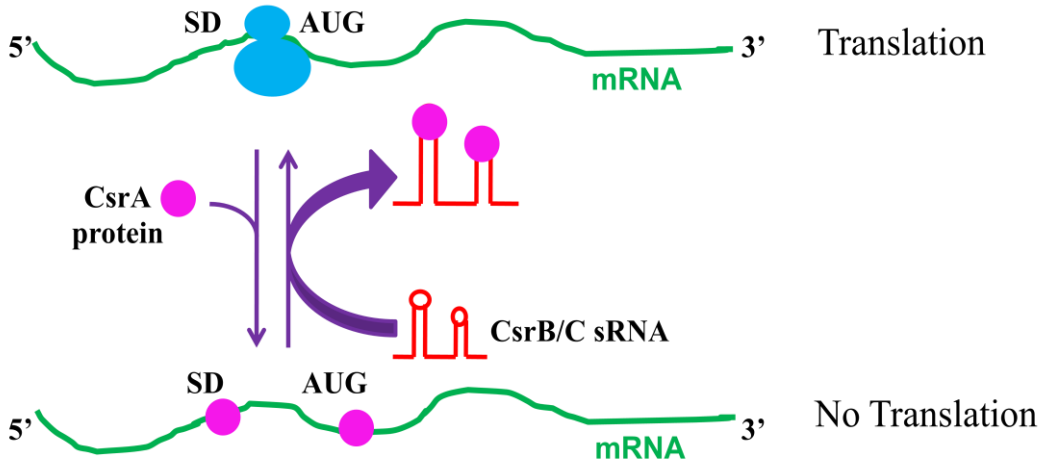

Figure 3: Protein binding sRNAs.
Translational repressor proteins (pink circle) bind to the 5'UTR of the mRNAs (green) and prevent ribosome binding resulting in translational inhibition. Protein binding sRNAs (red) sequester these proteins. The ribosomes (blue) bind to the RBS and initiate translation.

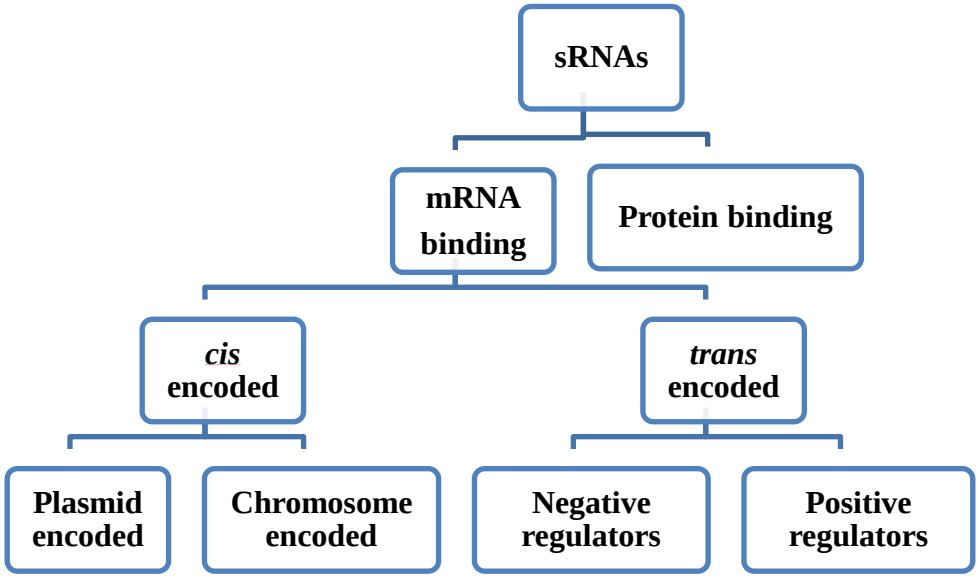

Figure 4: Classification of sRNAs.

1.3 Role of Hfq protein in sRNA regulation

Most of the sRNAs that act via limited complementarity have been found to require the RNA chaperone protein Hfq (Gottesman and Storz 2011). Hfq is a homolog of the Sm and Sm like proteins that are involved in splicing and degradation of mRNAs in eukaryotic and archaeal cells. It is present in 50% of the bacterial species and its deletion leads to serious defects in growth and virulence of the bacteria (Kavita et al. 2017).

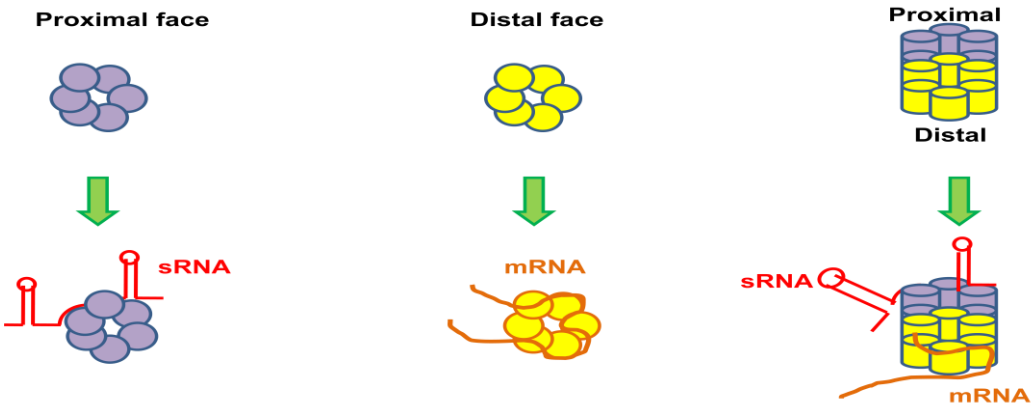

Figure 5: Hfq mediated regulation.
The proximal face of Hfq (purple) interacts with sRNA (red) while the distal face (yellow) interacts with mRNA (orange). Hfq is important for sRNA binding and its stabilization. It also binds to the mRNAs and facilitates sRNA-mRNA interaction.

Although gram-positive bacteria also contain Hfq homologs, there is no evidence for their role in riboregulation (Vanderpool et al. 2011). Hfq is known to promote the base pairing between the sRNAs and their mRNA targets by binding to the AU rich single-stranded regions of the RNA molecules. The sRNA and the mRNA bind at different faces to the Hfq molecule and the binding site of Hfq coincides to that of the sites vulnerable to RNA cleavage by RNases (Figure 5) (De Lay et al. 2013). Therefore, Hfq binding prevents the degradation of the RNAs by blocking the RNase E cleavage sites. It changes the secondary structure of the RNAs and facilitates base pairing between the sRNA and its target mRNA by opening up a certain stem loop involved in base pairing with the target (Storz et al. 2004;Aiba 2007).

1.4 sRNAs with dual functions

As mentioned earlier, small regulatory RNAs when discovered were called as non-coding RNAs as they did not code for any proteins and functioned by base pairing with other mRNAs. Dual function sRNAs came into picture with the discovery of RNAIII in *S. aureus* (Morfeldt et al. 1995). These sRNAs not only regulate the expression of other genes by base pairing mechanism but also code for small peptides that have a role to play either in the same or in a different pathway as that of the sRNA that codes for them (Gimpel and Brantl 2017). So far, only 10 sRNAs are reported to have a role in regulation as well as code for a peptide. Five of these; RNAIII and Pmc-mec of *S. aureus*, Pel sRNA of *S. pyogenes*, SgrS of enteric bacteria and SR1 of *B. subtilis* encode a peptide with known functions. Other than the above mentioned sRNAs, five more sRNAs were reported to have an ORF but only the peptide encoded by PhrS sRNA of *P. aeruginosa* was shown to be translated. For the rest of the sRNAs, neither the translation of the ORF has been proved nor could a function be assigned to the peptide encoded by them. Table 1 gives a summary of the dual function sRNAs reported in literature.

1.5 sRNAs regulating transcription in bacteria

Although most of the sRNAs reported are known to act as post-transcriptional regulators, a review of literature on sRNAs shows examples where sRNAs act upon the process of transcription elongation and termination thus acting as transcriptional regulators as well. 6S sRNA of *E. coli* (Lee et al. 1978), with homologs in several other gram-negative and gram-positive bacteria, is a 180-200 nt long sRNA. It forms a double stranded hairpin structure with a critical bulge that resembles the DNA in an open promoter complex.

Table 1: Summary of dual function sRNAs

Name/length (nt)/ host	Target gene	Function of target gene	Mechanism of action	Peptide encoded by the sRNA	Function of encoded protein
RNAIII (514) *S. aureus*	*spa* *coa* *rot* *hla* *sa1000* *sa2353* *sa2093* *sbi* *lytM* *mgrA* *map*	Immune evasion Fibrin clot formation Repressor of toxins Pore-forming toxins Fibrinogen binding protein SsaA homologue Immune evasion Cell wall metabolism Global regulator Extracellular adherence Cell wall metabolism	TI,RD TI,RD TI,RD TA TI,RD TI,RD TI*,RD* TI TI,RD RS TA*	Hld (26 aa)	PSM-toxin, *δ*-hemolysin
Psm-mec (157) MRSA	*agrA*	Activator of virulence genes	TI, RD independent of TI	PSM-mec (22 aa)	PSM-protein biofilm formation
Pel (459) *S. pyogenes*	*emm* *sic* *nga* *speB*	Virulence determinant Inhibitor of complement NAD glycohydrolase Cysteine protease	TcA TcA TcA PM	SLS (53 aa)	Streptococcal *β*-haemolysin
SR1 (205) *B. subtilis*	*ahrC*	Activator of arginine catabolism	TI	SR1P (39 aa)	Activator of RNase J1
SgrS (227) *E. coli*	*ptsG* *manXYZ* *yigL* *sopD*	Glucose transporter Mannose transporter Phosphatase Secreted virulence protein	TI,RD TI,RD RS TI, RD	SgrT (43 aa)	Repression of PtsG activity
PhrS (212) *P.aeruginosa*	*pqsR*	Quorum sensing regulation	TA	PhrS-ORF (37 aa)	unknown
RivX (291/239/219) *S. pyogenes*	*mga*	Activator of virulence genes	TA*,RS*	RivX protein (47 aa)	unknown
RSs0019 (218) *R. sphaeroides*	Sulfur metaboli-sm	Unknown	Unknown	RSs0019-ORF (50 aa)	unknown
VR-RNA (378) *C. perfringens*	*plc* *colA* *ptp* *cpd* *ycgJ-metB-cysK-luxS*	Phospholipase Collagenase tyrosine phosphatase Cyclic nucleotide phosphodiesterase Cysteine metabolism	Unknown	*hyp7*-ORF (72 aa)	unknown
Scr5239 (159) *S. coelicolor*	*dagA* *metE*	Agar utilization B12-independent methionine synthase	TI TI	scr5239-ORF (33 aa)	unknown

MRSA- Methicillin resistant *S. aureus*; TI-translation inhibition; RD- RNA degradation; TA- translation activation; RS-RNA stabilization; TcA-transcription activation; RP-RNA processing; PM-protein maturation; PSM-phenol-soluble modulin; *- mechanisms/targets are proposed but not experimentally verified; unknown-mechanisms or functions are unknown (Gimpel and Brantl 2017)

It binds to the σ^{70}-RNA polymerase at the region 4.2 which otherwise binds to the open promoter complex on DNA (Figure 6). This inhibits transcription from a subset of housekeeping genes and increases transcription of σ^{s} dependent promoters by altering the competition between σ^{70} and σ^{S} holoenzyme binding to the promoters (reviewed in (Storz et al. 2005;Waters and Storz 2009;Gottesman and Storz 2011). In another example of transcriptional regulation, sRNAs DsrA, ArcZ, and RprA bind to the 5' UTR of *RpoS* and suppress premature rho-dependent transcription termination (Figure 7). *rpoS* gene encodes the σ^{S} subunit of RNA polymerase, the general stress sigma factor that initiates transcription of genes involved in response to nutrient limitation and other adverse conditions. This sRNA-mediated anti-termination efficiently activates transcription of *rpoS* and leads to smooth transition of the cell to the stationary phase, thereby allowing the bacteria to rapidly adjust to the stressful metabolic changes (Sedlyarova et al. 2016).

Figure 6: 6S RNA antagonizes the activity of σ^{70} RNA polymerase.
In logarithmically growing cultures with low abundance of 6S RNA, σ^{70} RNA polymerase (green oval) transcribes the house keeping genes. As the cells approach the stationary phase, level of 6S RNA (red) increases which sequester the σ^{70} RNA polymerase making it unavailable for transcription of certain set of genes.

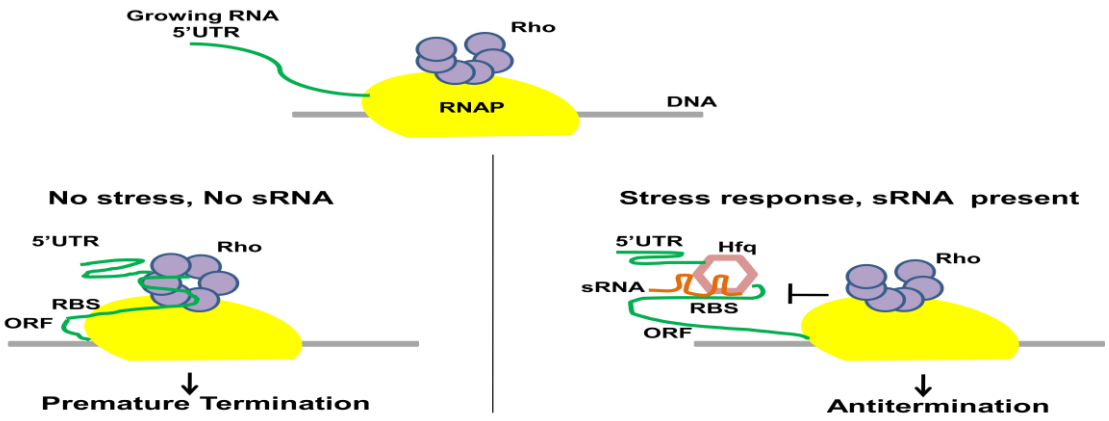

Figure 7: sRNA affects transcription.
Rho transcription terminator acts at the 5'UTR of the mRNA and causes attenuation of gene expression. sRNAs when bind at the 5'UTR interfere with the activity of the Rho factor thus causing antitermination.

1.6 Identification of small regulatory RNAs in bacteria

The first sRNA was identified in *E. coli* serendipitously while conducting other genetic studies (Mizuno et al. 1984). Although sRNAs are mostly untranslated and do not code for any proteins, they do not belong to the class of ribosomal or transfer RNAs and investigation of their functional significance is a very sought after research area. sRNAs being heterogeneous in length and immune to the frame shift and nonsense mutations makes their identification difficult by genetic screens. A combination of bioinformatics algorithms and experimental methods has proved to be the best approach for identification of sRNAs (Li et al. 2012).

1.6.1 Biocomputational methods for sRNA prediction

Computational tools for identification of sRNAs use algorithms based on certain characteristic features of sRNAs like sequence homology, conserved secondary structure, presence of promoters, and rho-independent transcriptional terminators in the intergenic regions, and those based on comparative genomics (Sridhar and Gunasekaran 2013).

sRNA prediction based on comparative genomics involves bioinformatics identification of conserved sequences in intergenic regions of closely related bacteria, clustering them together by pair wise or multiple alignments and deciphering the scores based on predicted secondary structural features (Backofen and Hess 2010;Li et al. 2012). By using this strategy in combination with identification of promoters and rho-independent terminators, authors identified several sRNAs in 21 clostridial genomes (Chen et al. 2011).

Approximately 433 sRNA candidates were predicted in the plant pathogen *Pectobacterium atrosepticum* (Kwenda et al. 2016) when a tool for identification of rho-independent terminators was used in combination with a web interface called SIPHT (**s**RNA **I**dentification **P**rotocol using **H**igh-throughput **T**echnology) (Livny 2012), that identifies sequence conservation in the upstream regions of rho-independent terminators. Six novel sRNAs were identified in large scale screens carried out in *R. prowazekii* using SIPHT web interface in combination with BPROM and TransTermHP web tools for determination of σ^{70} promoters and rho-independent terminators respectively (Schroeder et al. 2015).

1.6.2 Experimental methods for sRNA identification

sRNA candidates identified by biocomputational methods need experimental validation. Methods like sequencing of size fractionated labeled RNAs, microarray analysis of the transcriptome of the bacteria subjected to specific growth conditions, and RNomics followed

by cDNA cloning and Sanger sequencing have been used to identify sRNAs in bacteria reviewed in (Sharma and Vogel 2009). The experimental approaches in this field have undergone constant changes and modifications to yield better results. Following paragraphs describe the techniques used in this field.

1.6.2.1 Direct labeling and sequencing: This method involved sequencing and identification of size fractionated, gel eluted labeled RNA bands. Most abundant sRNAs or the ones with higher turnover rates were easily identified by this method. Housekeeping and regulatory RNAs like 6S RNA, Spot42, tmRNA (Vogel and Sharma 2005) and Hfq associated sRNAs in *L. monocytogenes* (Christiansen et al. 2006) were identified by the metabolic labeling followed by size fractionation and sequencing.

1.6.2.2 Microarrays: These were used for identification of sRNAs and differed from the ones meant to study the gene expression of the entire genome as they were high density tiling microarrays and included probes for the ORFs, tRNAs, rRNAs and strand specificity for the intergenic regions of both the strands of the genome. In combination with comparative genomic screens, sRNAs were identified by this technique in *E. coli* (Wassarman et al. 2001) and *B. subtilis* (Silvaggi et al. 2006).

1.6.2.3 RNomics and RNA sequencing: In this method, size fractionated cellular RNAs, reverse transcribed into cDNA are cloned into plasmid vectors followed by sequencing. The relative expression of RNAs in the cDNA libraries indicates the expression levels of the particular RNA under the studied growth condition. Before the advent of next generation sequencing techniques, sequencing was carried out by the conventional Sanger sequencing method. RNA-seq based RNomics has been carried out in *Salmonella* (Sittka et al. 2008), *Sinorhizobium meliloti* (Mao et al. 2008;Shi et al. 2009), *Vibrio cholerae* (Liu et al. 2009) and several other bacteria since the advent of this technology.

1.6.2.4 Copurification of sRNAs with proteins: sRNAs, especially in gram negative bacteria are known to stably associate with Sm-like Hfq protein to achieve intracellular stability and for base pairing with mRNAs (Wagner and Romby 2015). This feature was exploited in the technique RIL-seq and ~2800 Hfq-bound sRNAs and their interacting mRNA targets were identified in *E. coli*. FLAG-Hfq was UV-crosslinked with the interacting sRNAs *in vivo* and coimmunoprecipitation of the Hfq-bound sRNAs was performed. This step was

followed by RNA ligation, isolation, sequencing and mapping of the RNAs to the genomic locations (Melamed et al. 2016).

1.7 Identification of targets for sRNAs

With the ever increasing number of sRNAs being identified in several genera of bacteria, elucidation of their physiological roles becomes very important. Most of the sRNAs act by. binding to their direct mRNA targets via a very small stretch of 6-13 bp (Gottesman and Storz 2011). The base pairing between the sRNA-mRNA pairs is highly imperfect and non-contiguous and different regions of sRNA base pair with the targets (Durand et al. 2017). These features make detection of mRNA targets for sRNAs very difficult. Following paragraphs describe some of the bioinformatics programs used to predict mRNA targets of sRNAs and methods used for the experimental validation of the predicted targets.

1.7.1 Computational methods for target prediction: Algorithms designed for prediction of mRNA targets use experimentally validated sRNA-mRNA pairs as the positive training sets. Allowing non-canonical base pairing, length of base pairing interaction, region of mRNAs (TIR, 5'UTR, coding region, 3'UTR) involved in interaction are some of the parameters that can be fed into the programs for determining the plausible targets (Vogel and Wagner 2007). Following is a short description of few of the programs that can be used for *in silico* prediction of mRNA targets.

1.7.1.1 TargetRNA: A user specified genome sequence is taken as the input to determine the hybridization potential of the submitted sRNA sequence against each of the annotated mRNAs in the selected genome by using a modified Smith-Waterman dynamic algorithm which scores base pairing potential instead of homology potential. Once the hybridization scores are determined, the program determines the statistical significance of each of the potential sRNA- mRNA pair using RNAhybrid. After calculating the statistical significance of binding potential of the sRNA with each of the annotated messages in the genome, TargetRNA gives a ranked list of mRNAs whose binding potential with the sRNA meets the defined threshold (default P-value ≤ 0.01). mRNAs with significant base pairing potential are considered to be putative targets of the sRNA (Tjaden 2008).

1.7.1.2 IntaRNA: Using the RNAplex (Tafer and Hofacker 2008) algorithm this program predicts the interaction between two given RNA species. It allows the user to define the seed

size and considers accessibility of the target sites while making the predictions (Busch et al. 2008).

1.7.1.3 RNA Predator: This program also uses the dynamic RNAplex (Tafer and Hofacker 2008) algorithm to identify putative targets of sRNAs but is faster than the then existing programs in three orders of magnitude. It takes a single sRNA sequence as an input and gives a ranked list of probable mRNA targets. This online web server also gives additional options like enrichment in terms of Gene Ontology and accessibility studies around the site of interaction in the post processing steps (Eggenhofer et al. 2011).

1.7.1.4 CopraRNA: Comparative **pr**ediction **a**lgorithm for small RNA targets works in conjunction with the IntaRNA program. It takes minimum of 3 and maximum of 8 homologous sRNA sequences from distinct organisms as the input data and uses IntaRNA to come up with a list of whole genome mRNA targets for each pair of sRNA-genome. These results are then compared amongst each other and first 100 predictions are listed (Wright et al. 2014).

1.7.1.5 TargetRNA2: was built up by using the already existing algorithms and programs like TargetRNA (Tjaden 2008), RNA Predator (Eggenhofer et al. 2011), IntaRNA (Busch et al. 2008) and a certain observation of Peer and Margalit (Peer and Margalit 2011) which says that the sRNA regions involved in interaction with mRNA targets are better conserved and easily accessible as compared to the remaining sRNA sequence. TargetRNA2 considers four major aspects while determining mRNA targets: a) sRNA sequence conservation, b) sRNA accessibility c) mRNA accessibility and d) hybridization energy (Kery et al. 2014).

Other than the above described online prediction tools, programs like RNAup (Muckstein et al. 2006), sRNATarget (Cao et al. 2009), sTarPicker (Ying et al. 2011), *etc.*, are also available for biocomputational prediction of mRNA targets for bacterial sRNAs.

1.7.2 Experimental identification and validation of targets for sRNAs

Apart from the above described bioinformatics approaches; there are wet lab methods that can be used for identification of the target mRNAs and proteins. The standard approach for sRNA target identification over expresses and/or deletes the sRNA and analyzes the proteome or

transcriptome under its altered levels. However, continuous over expression of the sRNA makes it difficult to discriminate direct targets from the downstream ones and in some cases proves toxic to the cell. Therefore, the concept of pulse expression of the sRNA is used, wherein the sRNA is briefly expressed for 10-15 min followed by transcriptomic analysis to identify the rapid changes caused by the direct sRNA-mRNA interaction (Barquist and Vogel 2015). Use of this method led to the identification of 20 new mRNA targets in *Salmonella* where RNA from the strain pulse-expressing SdsR sRNA and that from a strain containing an empty vector was subjected to microarray analysis (Frohlich et al. 2016).

Microarrays have been rapidly replaced by high throughput RNA sequencing which has emerged as the technique of choice for sRNA and target identification. This has eliminated the requirement of custom probes, has high sensitivity and specificity and has reduced artefacts due to non-specific probe hybridization. Some of the RNA-seq modifications used for target identification are described below:

(a) GRIL-seq (Global sRNA target Identification by Ligation and Sequencing) was used to identify two sRNAs of *P. aeruginosa*, Sr016 and ErsA and their target OprD. In this method, the sRNA under study is over expressed in a strain expressing T4 RNA ligase. Chimeric transcripts of sRNA ligated to its mRNA targets are enriched and the isolated mRNAs are subjected to RNA sequencing and mapping of the ligated transcripts to the bacterial genome thus allowing identification of targets of a specific sRNA (Zhang et al. 2017).

(b) Dual RNA-seq is an impressive technique that performs RNA profiling of the bacterial pathogen and the infected host simultaneously. This technique identified bacterial mRNA targets of PinT sRNA of *Salmonella* and also indicated that PinT affects the RNA profile of the host cells in several ways, one being its effect on the JAK-STAT pathway of HeLa cell lines (Westermann et al. 2016).

(c) Ribosome profiling (Ribo-seq) determines the altered transcript levels as well as the changes caused in translation by the sRNAs. Briefly, the sRNA under study is pulse expressed and ribosome foot printing is performed. The foot printed RNA and total RNA is converted to cDNA and PCRs are performed using this cDNA as template. The PCR amplified DNA libraries are sequenced using next generation sequencing methods and the sequenced data is analyzed to identify the affected mRNA and protein targets of the sRNA under study (Wang et al. 2015).

Once the mRNA targets of the sRNA are identified by any of the above described methods, the next step is to validate the post-transcriptional regulation of the identified mRNAs by sRNA and proving direct base pairing between the two. One of the most common approaches to achieve the same is to estimate the expression levels of the *in vivo lacZ* or *gfp* reporter fusions made on mRNA targets in the presence of normal and mutant sRNA. Alternatively, *in vitro* approaches like gel mobility shift assays are used to prove the base pairing between the sRNA and its mRNA target, employing mutations and compensatory mutations in the interacting regions of either the sRNA or the target mRNA (Barquist and Vogel 2015).

1.8 Complexity in regulation by sRNAs

Research in past three decades has explicated several mechanisms by which sRNAs carry out regulation of gene expression. However, the sRNAs mostly do not hold any characteristic features that indicate or dictate the outcome of regulation, making it difficult to predict the result of base pairing. For example, sRNAs are known to function in coordination with ribonucleases, but whether the recruited RNase would carry out ribonucleolytic decay of the RNA duplex or will play a role in processing of the mRNA target is difficult to predict. One sRNA can regulate the expression of several targets and multiple sRNAs can act upon to regulate a single target. RhyB sRNA of *E. coli* is an example of the first category as it represses the expression of 20 operons under iron limiting conditions (Massé et al. 2005). sRNAs DsrA, RprA and ArcZ are representatives of the second type as all of them positively regulate the expression of *RpoS* mRNA, each of them base pairing at the 5'UTR of the mRNA and opening an inhibitory stem loop (Soper et al. 2010). On the other hand, Qrr3 sRNA of *V. harveyi* uses four different mechanisms to regulate the expression of four different targets *i.e. luxM*, *luxO*, *luxR*, and *aphA* to modulate the quorum sensing response of *V. harveyi* (Feng et al. 2015). McaS sRNA of *E. coli* acts in a unique manner by base pairing with *CsgD* mRNA encoding a regulator of curli and flagella synthesis and also by sequestering the CsrA protein, the translational repressor of exopolysaccharide β-1,6 N-acetyl-D-glucosamine (Jorgensen et al. 2013;Gimpel and Brantl 2017). Examples of the above kind increase the complexity of sRNA regulation.

1.9 Regulatory outcomes of sRNAs

sRNAs modulate the stability/translation of the mRNA targets in response to the ever changing environment (Holmqvist and Wagner 2017). They have myriad roles to play in several cellular processes like metabolism, quorum sensing and biofilm formation, virulence, *etc* (Nitzan et al. 2017).

1.9.1 Regulation of cellular metabolism

CsrA protein, the central component of carbon utilization system in *E. coli* that regulates response to carbon starvation, glycogen biosynthesis, motility, etc., is sequestered by two sRNAs CsrB and CsrC during the exponential growth phase to repress the metabolic pathways related to the stationary phase (Revelles et al. 2013). *S. typhi* sRNAs, RfrA and RfrB, are involved in iron homeostasis, resistance to oxidative stress, and optimal growth in host cells (Leclerc et al. 2013).

1.9.2 Regulation of quorum sensing and biofilm formation

The phenomenon of quorum sensing (QS) is complexly regulated by several regulatory mechanisms in which the role of sRNAs is well established (Papenfort and Bassler 2016). For example, PhrS and ReaL sRNAs of *P. aeruginosa* positively affect the synthesis of PqsR, the transcriptional regulator; and Pseudomonas quinolone signal (PQS), the effector; of the *pqs* branch of QS in *Pseudomonas* (Sonnleitner et al. 2011;Carloni et al. 2017).

sRNAs govern the expression of transcriptional regulators responsible for biofilm formation and its maintenance (Chambers and Sauer 2013). CsrB/C sRNAs of *E. coli* were the first examples of sRNAs identified to be involved in biofilm formation (Jackson et al. 2002) following which a number of sRNAs were found to regulate biofilm formation. For example, OmrA/B (Holmqvist et al. 2010), GcvB (Jorgensen et al. 2012), RprA (Mika et al. 2012), McaS (Thomason et al. 2012) sRNAs of *E. coli*; ArcZ (Monteiro et al. 2012) and SdsR sRNA of *S. typhimurium*, CrcZ (Sonnleitner et al. 2009) and RsmY/Z sRNA of *P. aeruginosa* and Qrr sRNAs of *Vibrio* species have important roles to play in biofilm formation.

1.9.3 Stress response and adaptation to growth conditions

sRNAs help the bacteria to cope up with, and survive environmental and host-induced stress (Holmqvist and Wagner 2017). GadY sRNA of *E. coli* positively regulates the expression of GadX, the major acid response regulator (Opdyke et al. 2004). FnrS sRNA of *E. coli* expressed under anaerobic conditions regulates expression of some 32 mRNAs most of which code for proteins involved in energy metabolism and combating oxidative stress (Durand and Storz 2010).

1.9.4 Regulation of pathogenicity by sRNAs

Pathogenic bacteria possess intricate regulatory networks that control temporal and spatial expression of virulence factors allowing them to infect and persist in the ever changing host environments (Caldelari et al. 2013;Michaux et al. 2014). sRNAs have been found to be at the

centre of these regulatory networks (Papenfort and Vogel 2010). sRNAs encoded on pathogenicity islands (PAIs) play central role in virulence (Svensson and Sharma 2016). IsrM sRNA, encoded in the pathogenicity island SPI-1 of *Salmonella*, is indispensable for invasion of host cells and its deletion impairs virulence of this pathogen (Gong et al. 2011). RsaA sRNA of *S. aureus* enhances biofilm formation and reduces capsule formation thus initiating a shift from an acute infection to a chronic one (Romilly et al. 2014). Multicopy LhrC sRNA of *L. monocytogenes* regulates the expression of *lapB* that encodes a cell wall anchored adhesin responsible for invasion of eukaryotic cells (Sievers et al. 2014). Type 6 secretion system in *V. cholerae* which is responsible for secretion of virulence factors to the cell exterior is regulated by Qrr sRNAs (Shao and Bassler 2014). In addition to the above stated examples, a large number of other sRNAs have been identified and characterized for their involvement in regulating the virulence of pathogens like *S. enterica* (Hébrard et al. 2012), *L. monocytogenes* (Mellin and Cossart 2012), *Vibrio* spp. (Nguyen and Jacq 2014), *Yersinia* spp. (Heroven et al. 2012), *S. aureus* (Fechter et al. 2014), *M. tuberculosis* (Arnvig and Young 2012), *H. pylori* (Pernitzsch and Sharma 2012), Streptococcal species (Zorgani et al. 2016), etc.; a few of which are listed in table 2.

Table 2: sRNAs regulating pathogenicity

Pathogen	sRNA	Target	Phenotype/function	Reference
B. abortus	AbcR-1, AbcR-2	?	Infection of macrophages and chronic infection of mice	(Caswell et al. 2012)
B. burgdorferi	DsrA	*rpoS*	Expression of virulence-associated surface proteins	(Lybecker and Samuels 2007)
C. perfringens	VR-RNA	*cola*	Toxin expression	(Obana et al. 2010)
H. pylori	RepG	*tlpB*	Antisense repressor of chemotaxis receptor mRNA	(Pernitzsch et al. 2014)
Group A streptococci	FasX	*ska, cpa*	Streptokinase and pilus expression	(Liu et al. 2012)
L. monocytogenes	RliB	Lmo2104	*rliB* mutation increases colonization of spleen in mice.	(Toledo-Arana et al. 2009)
	LhrA	*chiA*	Chitinase	(Mraheil et al. 2010a)
	Rli38	?	*rli38* mutant is attenuated in oral mouse infection.	(Toledo-Arana et al. 2009)
	Rli27	Lmo0514	Activation of expression of a cell wall protein inside cells	(Quereda et al. 2014)
P. aeruginosa	PhrS	*pqsR*	Key quorum sensing regulator	(Sonnleitner et al. 2011)
S. typhimurium	InvR	*ompD*	Represses a porin of the core genome	(Pfeiffer et al. 2007)
	IsrJ	?	Affects invasion into non-phagocytic cells	(Padalon-Brauch et al. 2008)
	MgrR	*eptB*	Modulator of LPS	(Moon and

			modification	Gottesman 2009)
	DapZ	*oppA, dppA*	Regulator of amino acid / oligopeptide metabolism	(Chao et al. 2012)
	SgrS	*sopD*	Core genome sRNA that regulates a SPI-1 effector	(Papenfort et al. 2012)
	SroA	?	Required for infection of mice	(Santiviago et al. 2009)
S. dysenteriae	RyhB	*virB*	Affects T3SS/effectors and virulence	(Murphy and Payne 2007)
S. enteric	RprA	*rpoS, ricI*	Controls conjugation of pSLT virulence plasmid in response to membrane conditions	(Papenfort et al. 2015a)
S. aureus	RNAIII	*rot, spa, hla, coa,* SA1000, SA2353	Global regulator of QS and virulence gene expression	(Novick et al. 1993), (Chabelskaya et al. 2014)
	SprD	*Sbi*	Immune evasion	(Chabelskaya et al. 2010)
	SprX	*spoVG*	Glycopeptide resistance	(Eyraud et al. 2014;Kathirvel et al. 2016)
S. pneumoniae	csRNA1-5	?	csRNA 4 and 5 regulate stationary phase autolysis	(Halfmann et al. 2007)
S. pyogenes	FasX	*fpbA, mrp, ska, pel*	Increases interaction with epithelial cells	(Klenk et al. 2005)
	RivX	*mga*	Regulates expression of virulence transcription factors	(Roberts and Scott 2007)
	Pel	*emm, sic, speB*	Bifunctional RNA also encodes the SagA protein	(Mangold et al. 2004)
V. cholerae	VqmR	*vpsT, rtx*	Biofilm formation and toxin expression	(Papenfort et al. 2015b)
	TarA	*ptsG*	Major glucose regulator	(Richard et al. 2010)
	TarB	*tcpF*	Regulates a virulence factor and affects colonization	(Bradley et al. 2011)
	Qrr1-4	*hapR, vca0939*	QS control and de-repression of virulence genes	(Lenz et al. 2004), (Hammer and Bassler 2007)
	VrrA	*ompA*	Outer membrane vesicle synthesis & colonization of mouse intestine	(Song et al. 2008)
	CsrB/C/D	CsrA	QS and virulence factor production	(Lenz et al. 2005)
X. campestris	sX13	*hrpX*	Affects levels of T3SS regulator	(Schmidtke et al. 2013)

?= unknown, reviewed in (Romby et al. 2006;Papenfort and Vogel 2010;Caldelari et al. 2013;Michaux et al. 2014;Ortega et al. 2014;Svensson and Sharma 2016)

1.10 *Pseudomonas aeruginosa*- an opportunistic pathogen

Pseudomonas aeruginosa, a gram-negative ubiquitously present bacterium that grows in soil, water, plant and animal tissues is one of the top three causes of nosocomial infections (Balasubramanian et al. 2013). It is an opportunistic human pathogen that causes severe infections in immunocompromised individuals suffering from AIDS and neutropenia due to chemotherapy (Sadikot et al. 2005). While the above conditions make a patient vulnerable to several bacterial and fungal infections, there are three particular cases in which *Pseudomonas aeruginosa* causes severe infections that are very difficult to control. These are: bacteraemia in burn wound patients, chronic lung infections in cystic fibrosis patients and ulcerative keratitis in individuals using contact lenses (Lyczak et al. 2000). The success of this bacterium in causing these infections relies on its large battery of virulence factors, capacity to form biofilm, and inherent and acquired resistance to several antibiotics (Lyczak et al. 2000). Virulence of *P. aeruginosa* is attributed to several pathogenicity factors that could be either cell surface associated like alginate, lipopolysaccharide (LPS), pilus and non-pilus adhesions, flagellum, etc., or secretory in nature like proteases, elastases, exoenzymes, rhamnolipids, pyocyanin, siderophores, exotoxins, etc.(Balasubramanian et al. 2013).

1.10.1 *Pseudomonas aeruginosa* strain PAO1

Pseudomonas aeruginosa strain PAO1 is the most common reference strain used to carry out studies on this bacterium. The complete genome of this strain was sequenced in the year 2000 (Stover et al. 2000). It is a spontaneous chloramphenicol mutant derivative of the original Australian strain PAO that was isolated in 1954 from a burn wound in Melbourne, Australia (Klockgether et al. 2010). With a genome size of 6.4 Mbp and G+C content of 66.6%, it has 5,570 predicted open reading frames, 434 transcriptional regulators, 24 sigma factors and 8% of its genome codes for regulatory proteins (Stover et al. 2000).

1.10.2 Pathogenesis of *P. aeruginosa*

Equipped with a battery of virulence determinants (Figure 8), pathogenesis of *P. aeruginosa* is invasive, toxinogenic and multifactorial. Its infection process can be divided into three distinct phases: (1) bacterial adhesion and colonization (2) invasion and (3) disseminated systemic disease. After the initial colonization that is facilitated by cell surface associated factors, the infection might take an acute or a chronic form. The acute infection is marked by high production of extracellular virulence factors the production of which decreases during chronic phase of infection (Strateva and Mitov 2011).

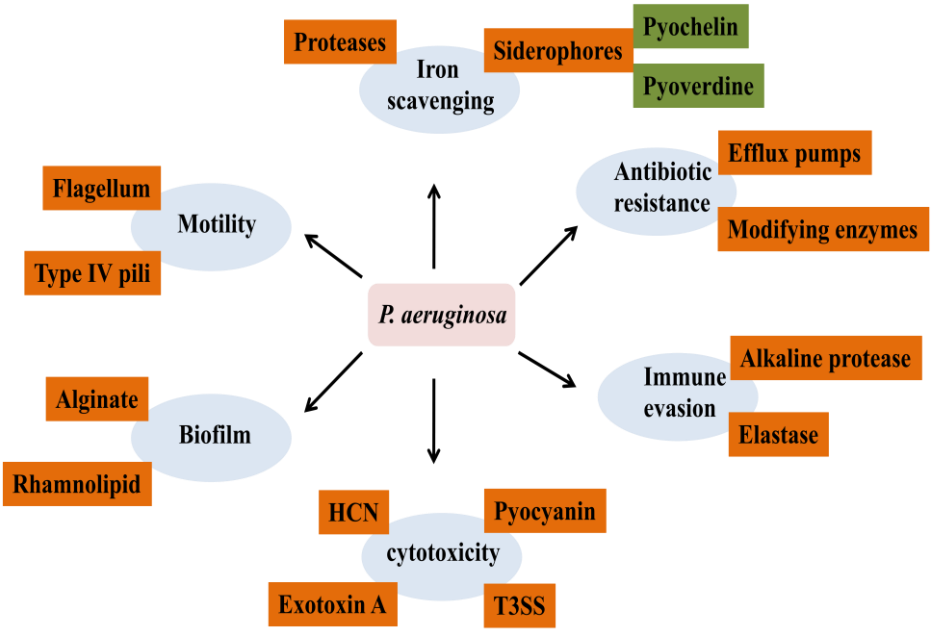

Figure 8: Virulence factors of *P. aeruginosa*

Table 3: Virulence determinants of *P. aeruginosa*

Phase of infection	Cell associated factors	Extracellular factors
Adhesion	Type IV pili, lectins, glycocalix, alginate	-
Adhesion facilitation	-	Neuraminidase
Motility/ chemotaxis	Flagella (swimming motility) Retractile pili (twitching motility)	-
Invasion	None	Elastase, Alkaline protease, Haemolysins (phospholipases and rhamnolipid), Cytotoxin (leukocidin), Pyocyanin, Siderophores
Toxinogenesis	Lipopolysaccharide (endotoxin), lectins	Exotoxin A, ExoS, ExoU, ExoT, ExoY, Enterotoxin
Dissemination		
Antiphagocytic surface properties	Slime layers,Glycocalix, Lipopolysaccharide	-
Defense against serum bactericidal reaction	Slime layers, Glycocalix Lipopolysaccharide	Proteases
Defense against immune responses	Slime layers, Glycocalix	Proteases

Reviewed in (Strateva and Mitov 2011)

Virulence factors of *P. aeruginosa* can be broadly divided into two categories namely, factors involved in acute infections and factors involved in chronic infections. Virulence factors

causing acute infections can be cell surface associated or secretory in nature. For example, pili help in adherence to the epithelium, adhesins like exoenzyme S reinforce the adherence, and proteases cause tissue necrosis and bleeding. Siderophores like pyoverdin and pyochelin that help in iron scavenging inside the host and pseudocapsule of alginate that protects *P. aeruginosa* from phagocytosis and dehydration are examples of virulence factors involved in chronic infections (Ben Haj Khalifa et al. 2011). A list of virulence factors of *P. aeruginosa* in relation to their involvement in different phases of infection is supplied in table 3.

1.10.3 Virulence gene regulation in *P. aeruginosa*

Virulence genes of *P. aeruginosa* are under very stringent controlling mechanisms that govern their timely expression in response to a variety of environmental and host induced stimuli.

a) Two-component systems (TCS): The adaptability of *P. aeruginosa* in a wide range of habitats comes from the 127 TCSs encoded by the genome of this organism (Stover et al. 2000). A TCS consists of a pair of sensor kinase and a response regulator (RR) protein. On sensing a change in the environment, the membrane integrated sensory histidine kinase (HK) relays it into the cell by phosphorylating a specific cytoplasmic transcriptional regulator. Following are few examples of TCS of *P. aeruginosa*. GacSA controls the expression of virulence factors, secondary metabolites, biofilm formation, and QS and has a central role to play when the bacteria switches from acute to chronic stage of infections (Pessi et al. 2001). PmrAB TCS regulates response to the limiting concentrations of cations and resistance against polymyxin B and cationic antimicrobial peptides (McPhee et al. 2006). RocS1/R/A1, a nonclassical two-component system regulates expression of fimbrial adhesins involved in biofilm formation (Kulasekara et al. 2005). Along with imparting resistance to antimicrobial peptides and aminoglycosides like the PmrAB TCS, the PhoPQ TCS also regulates swarming motility and biofilm formation (McPhee et al. 2006) .

b) Quorum sensing in *P. aeruginosa*: QS in *P. aeruginosa* plays a vital role in regulation of expression of virulence genes and biofilm formation (Bjarnsholt and Givskov 2007). QS regulates approximately 10% of the *Pseudomonas* genome, which is responsible for a number of physiological processes and virulence phenotypes (Lee and Zhang 2015). Few of the QS regulated virulence factors of *P. aeruginosa* are listed in table 4.

Table 4: QS regulated virulence factors of *P. aeruginosa*

Gene	Virulence factor	Effect on the host	Benefits to the bacteria	Reference
lasB	Elastase	Degradation of elastin and collagen	Iron acquisition	(Yanagihara et al. 2003)
lasA	Protease	Disruption of epithelial barrier	Host immune evasion and enhanced colonization	(Park et al. 2000)
toxA	ExotoxinA	Cell death	Establishment of infection and enhanced colonization	(McEwan et al. 2012)
aprA	Alkaline protease	Degradation of host complement system and cytokines	Immune evasion and persistent colonization	(Laarman et al. 2012)
rhlAB	Rhamnosyl-transferases (Rhamnolipid)	Necrosis of host macrophage and PMN lymphocytes	Immune evasion and biofilm development	(Jensen et al. 2006)
lecA	Lectin	Paralysis of airway cilia	Establishment of infection and enhanced colonization	(Adam et al. 1997)
hcnABC	Hydrogen cyanide	Cellular respiration arrest and poorer lung function	Enhanced colonization	(Ryall et al. 2008)
phz ABCDEFG, phzM	Pyocyanin	Dampens host cellular respiration and causes oxidative stress, paralysis of airway cilia, delayed inflammatory response to *P. aeruginosa* infections through neutrophil damage	Establishment of infection, enhanced colonization and immune evasion	(Lau et al. 2004)

Reviewed in Lee and Zhang, 2015

The highly adaptable *Pseudomonas* QS gives flexibility to respond to the cues of the environment and those arising during host infection and is composed of at least four interconnected branches namely; *las, rhl, pqs* and *iqs* (Figure 9) (Lee and Zhang 2015). Each branch consists of an autoinducer and its cognate cytoplasmic receptor (Rutherford and Bassler 2012). This autoinducer/receptor complex binds in the promoter region of the downstream genes to initiate their transcription (Rutherford and Bassler 2012). *las* system is at the apex of this hierarchy and its transcriptional regulator LasR in conjunction with its cognate autoinducer 3-oxo-C_{12}-HSL activates the expression of other branches of QS in a cell density dependent manner. RhlR-C_4-HSL complex activates expression of elastase B, rhamnolipids, pyocyanin, and hydrogen cyanide. The third signal molecule, PQS

(Pseudomonas quinolone signal), structurally identified as 2-heptyl-3-hydroxy-4-quinolone, binds to PqsR (also known as MvfR) and regulates synthesis of pyocyanin, rhamnolipids, lectins, and elastase (Reen et al. 2011). Integrated Quorum Sensing (IQS) with its signal molecule 2-(2-hydroxyphenyl)-thiazole-4-carbaldehyde regulates synthesis of PQS, C_4-HSL, pyocyanin, and elastase (reviewed in (Lee et al. 2013)).

However, there are situations where in this hierarchy is challenged and *rhl* takes over the role of *las*. For example, PQS production is delayed in *las* mutants whereas it completely stops in the *lasR-rhlR* double mutant (Diggle et al. 2003). Over production of *rhlR* in the above strain restores the levels of PQS and other *lasR* dependent virulence factors thus concluding that RhlR takes over the functions of LasR in its absence (Dekimpe and Déziel 2009).

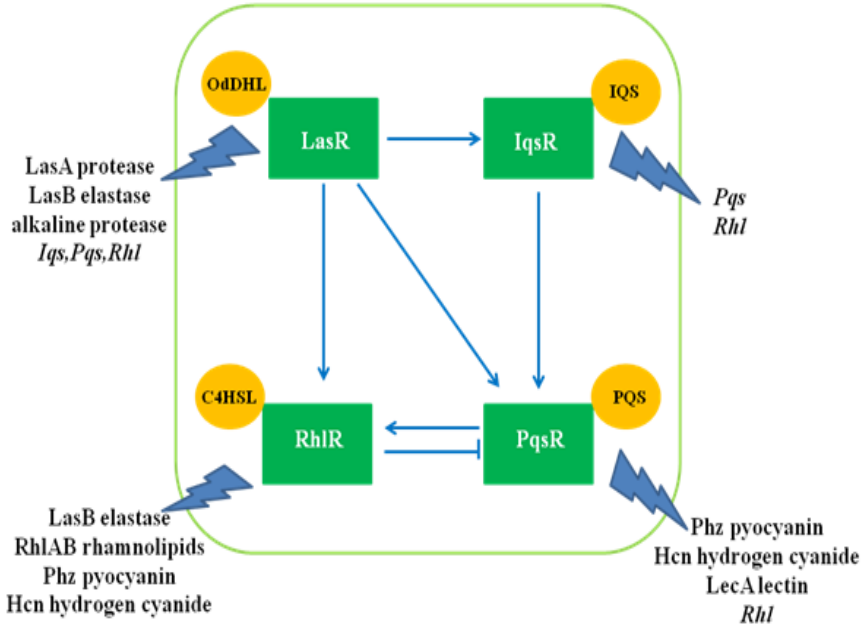

Figure 9: Schematic illustration of the quorum sensing system in *P. aeruginosa*.
Open arrows indicate a stimulatory effect and blunted arrows indicative inhibitory effect.

c) Regulatory small RNAs: Role of sRNAs in regulating the virulence of pathogenic bacteria is well established and is discussed under section 1.9.4. Virulence of *P. aeruginosa* also relies on sRNA mediated post-transcriptional regulation.

For example, sRNAs RsmY and Z sequester the RsmA/RsmF protein which act as post-transcriptional repressors of mRNAs encoding several virulence associated exoproducts (Kay et al. 2006).

Table 5: Functionally characterized sRNAs in *Pseudomonas aeruginosa* PAO1

sRNA (size in nt)	Regulated by	mRNA/ protein targets	Mechanism of action	References
PrrF1/2 (~110)	Fur and iron availability	PA4880, *sodB*	Bind to 5'UTR of targets and down regulates expression under Fe-limited conditions	(Wilderman et al. 2004)
RsmY/Z (~96)	GacS/GacA TCS	RsmA protein	Sequester RsmA protein, prevent translational repression of virulence factors and antifungal metabolite encoding mRNAs	(Kay et al. 2006)
CrcZ (~407)	CbrA/B TCS, RpoN sigma factor	CrcA protein	Catabolite repression of degradative enzymes like aliphatic amidase by binding to CrcA protein & removing it from the target mRNAs	(Sonnleitner et al. 2009)
RgsA (~120)	GacA and RpoS sigma factor	Fis (global transcriptional regulator), AcpP (acyl carrier protein)	Swarming motility, pyocyanin production	(Gonzalez et al. 2008) (Lu et al. 2016)
PhrS (~212)	Oxygen responsive regulator ANR	*pqsR*	Positively regulates the expression of transcriptional regulator PqsR by anti-antisense mechanism under O-limited conditions	(Sonnleitner et al. 2011)
NrsZ (~130,40)	NtrB/C, RpoN sigma factor	*rhlA*	Binds to *rhlA* mRNA in the 5'UTR, opens up the self-inhibitory loop within the mRNA thus increasing rhamnolipid synthesis and swarming motility under N-limited conditions	(Wenner et al. 2014)
ErsA	σ^{22}	AlgC provides sugar precursors for poly-saccharides, *amrZ* transcriptional regulator of alginate production	Promotes biofilm formation by unknown mechanism	(Ferrara et al. 2015;Falcone et al. 2018)
RsmW	?	?	Upregulated under biofilm formation	(Miller et al. 2016)
RsmV	?	RsmA/F	Translational activation of T6SS and inhibition of T3SS	(Janssen et al. 2018)

PhrS sRNA, expressed under hypoxic conditions, activates *pqs* branch of quorum sensing in turn influencing pyocyanin production (Sonnleitner et al. 2011).

Two tandemly placed sRNAs, PrrF1 and PrrF2 regulate iron homeostasis by blocking the expression of iron containing proteins like superoxide dismutase, succinate dehydrogenase, and a bacterioferritin (Wilderman et al. 2004). They also induce production of PQS and expression of several virulence related genes (Reinhart et al. 2017). NrsZ sRNA, under N-limited conditions, increased the production of biosurfactant rhamnolipid. Rhamnolipids act as haemolysins of *Pseudomonas* and are indispensible for its swarming motility (Wenner et al. 2014). Table 5 gives a summary of the sRNAs characterized in *P. aeruginosa* PAO1 to date.

1.11 Potential use of bacterial sRNAs

1.11.1 sRNAs as biomarkers

Small RNAs regulate virulence factors of bacterial pathogens. Their expression levels are often growth phase dependent (Patenge et al. 2012) and sRNAs like RsaA are responsible for the switch from acute to chronic infections of *S. aureus* (Romilly et al. 2014). Therefore, sRNAs have potential use as biomarkers in diagnosis and detection of the severity of infection. When *S. aureus* isolates from infected patients were studied for the presence and the expression of sRNAs, RNAIII expression was lower in strains isolated from patients with septic shock syndrome than in patients with chronic infections. SprD levels in conjunction with RNAIII expression distinguished colonized patients from patients with blood stream infections (Bordeau et al. 2016).

1.11.2 Targeting *trans* sRNAs for antimicrobial therapy

Antisense therapy by sRNAs to treat microbial infections comprises the use of a small synthetic DNA or RNA oligo of 15-25 nt in length to inhibit transcription or translation. Antisense drugs have been studied in case of cancers, diabetes, amyotrophic lateral sclerosis, asthma and arthritis. Updated information about production, regulation, and pathogenic implications of sRNAs can be used to design novel, potential therapeutics based on sRNA-complementary peptide nucleic acids (PNAs). Information about sRNA expression can be used to develop an ultrasensitive diagnostic system to detect extremely low concentrations in a very short time (Mraheil et al. 2010b).

Antisense therapy might not kill the pathogen but might make the bacteria more susceptible to the antibiotic and increase the efficacy of existing antibiotics in a combined therapy. Drugs interfering directly with *trans*-encoded sRNA function *in vivo* are not known but ongoing research indicates that compounds modulating sRNA levels can be employed to boost

antibiotic activity (Dersch et al. 2017). For example, the key metabolite GlcN6P synthesized by GlmS and involved in cell envelope formation in enteric bacteria is regulated by sRNAs GlmY and GlmZ. Depletion of this metabolite in turn stimulates the synthesis of the sRNAs, which provide protection against antibiotics such as bacilysin, that act by inhibition of GlmS. Therefore, the bactericidal action of GlmS inhibitors can be ensured and enhanced by co-application of a non-metabolizable GlcN6P analogue, which will suppress the accumulation of GlmY/GlmZ (Khan et al. 2016).

Biofilms are known to provide protection against antibiotics and host defences as mentioned earlier. *E. coli* sRNA RybB negatively regulates biofilm formation by targeting the mRNA *CsgD*, which codes for a crucial biofilm regulator. Epigallocatechin gallate (EGCG), a polyphenol present in green tea, activates *rybB* expression thus abolishing biofilm formation and exposing the bacteria to antibiotics and host defences. This makes EGCG a promising adjuvant that increases antibiotic susceptibility in combined chemotherapy (Serra et al. 2016).