Introduction and Review of Literature

Chapter-1

1. INTRODUCTION

1.1. Regulation of gene expression in prokaryotes

Plasticity of gene expression is essential for survival with the frequently changing and often extreme environments, encountered by an organism (Murphy *et al.,* 2014). Cellular mRNA and protein levels are finely regulated continuously for needs of an organism. The most common mechanism of gene regulation in bacteria includes modulation of the rate of transcriptional initiation by transcriptional factors, attenuation through the modulation of forming mRNA secondary structure, mRNA stability by ribosomal binding, mRNA translation and protein stability, mRNA decay by ribonucleases (Picard *et al.*, 2009). In recent years, it has become evident that, in addition to, these transcriptional regulatory programs, small regulatory RNAs have also been identified to play an important role in the post transcriptional regulation of many genes. Small RNAs (sRNA) or noncoding RNAs (ncRNA) are the modulators of gene expression and were once recognised as the junk of the genome. Initially, sRNAs were considered as repressors; however, the discovery gradually changed our viewpoint on the complexity of biological regulatory network. RNA molecules that act as regulators were known in bacteria for years before the first microRNAs (miRNAs) and short interfering RNAs (siRNAs) were discovered in eukaryotes (reviewed in Waters & Storz, 2009; Zhou and Xie, 2011).

1.2. Small noncoding RNAs as post transcriptional regulators

Post transcriptional regulation involves modulation of mRNA stabilization or degradation that subsequently affects mRNA's translational efficiency into protein. Small regulatory RNAs in bacteria have been coupled with the regulation of transcription, stability or translational efficiency by base pairing with mRNA targets or proteins, resulting in substantial changes in the physiology of an organism (reviewed in Argaman *et al.,* 2001; Wassarman *et al.,* 2001; Storz *et al.,* 2005). They have emerged as regulatory hierarchy of gene expression, controlling diverse biological functions including cell envelope architecture, cell cycle, metabolism, bacterial cell to cell communication, oxidative stress, biofilm formation, antibiotic resistance and the expression profile of various virulence factors (Ortega *et al.,* 2014). Earlier, the regulation of virulence gene expression was accredited to the activity of transcription

factors that switch on or off relevant sets of genes in response to environmental cues. In addition, small RNAs have emerged as regulators of bacterial pathogenicity as evident in the research of last 2 decades (Papenfort & Vogel, 2010). Historically, sRNAs in bacteria are untranslated, i.e. do not code for proteins, usually range from 70 to 500 nt in length, with an exception of RNAIII, 514 nt and SSR42, 891 nt in *S. aureus* (Morrison *et al.*, 2012), some sRNAs also act as bifunctional RNAs, that encode for small peptides too (reviewed in Bobrovskyy & Vanderpool, 2013).

1.3. Mechanism of regulation of small noncoding RNAs

1.3.1. Requirement of bases for pairing sRNA with mRNA

sRNA controls the fate of mRNAs by base pairing with the target mRNAs (Wang *et al.*, 2015). Base pairing between sRNA and mRNA generally involves a minimum seed region of 6-8 contiguous base pairs (reviewed in Gottesman & Storz, 2010; Storz *et al.*, 2011) leading to a number of regulatory outcomes (Fig. 1). In some cases significantly longer base pairing regions have also been predicted.

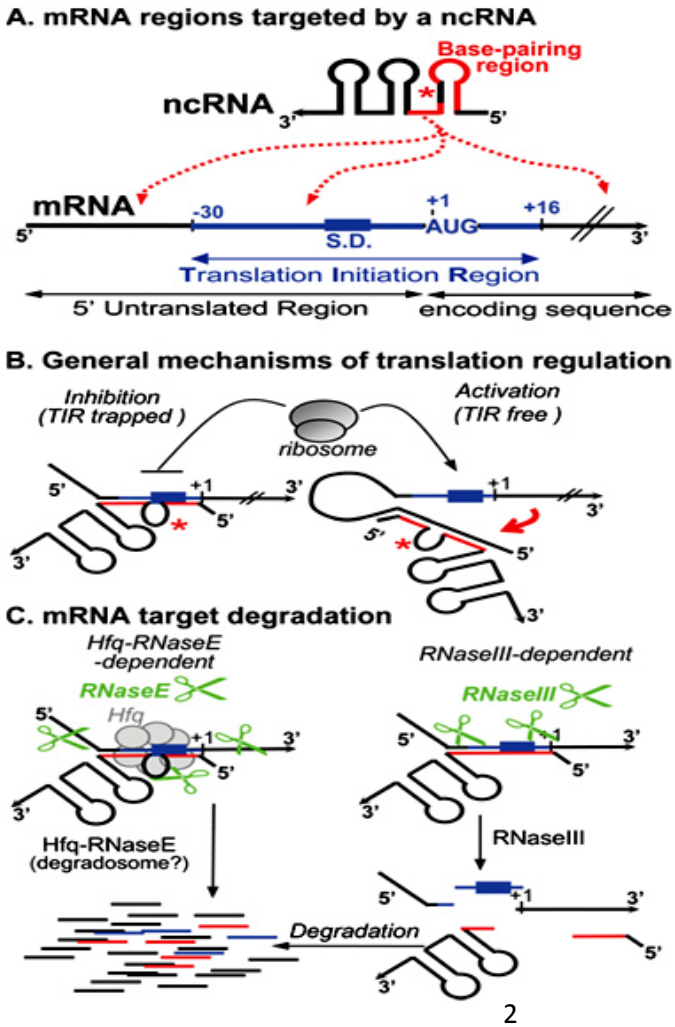

Fig. 1. General features and consequences of the ncRNA/ mRNA interaction.

A. Different regions of mRNA that can be targeted by an ncRNA
B. Mechanisms of translation regulation mediated by ncRNAs
C. RNase E and RNase III dependent mRNA degradation mediated by ncRNAs (Adapted from Repoila & Darfeuille, 2009).

The extent and region of base pairing with sRNAs impart target mRNAs for differential regulation (Mitarai *et al.*, 2007). sRNA can base pair at the upstream or downstream of 5' or 3' untranslated region, translation initiation region (TIR), transcription terminator region or even within the coding sequence of mRNA.

1.3.2. Translational repression

Many of the sRNAs base pair around the Shine-Dalgarno sequence (ribosome binding site (RBS)), thereby preventing the ribosome from binding, resulting in translational repression (Fig. 1). Translation may also be blocked when the length of pairing is 50 or more nucleotides upstream of the RBS (Gottesman & Storz, 2010). sRNA mediated mRNA destabilization and degradation is not only dependent on binding at the 5' UTR of mRNA, but also on the RNA degradosome, an enzyme complex containing endoribonuclease E or RNase II. The interacting RNA complexes then become substrates for RNases and get degraded. For example, in *S. typhimurium*, MicC sRNA silences outer membrane protein *ompD* mRNA by base pairing within the coding sequence (CDS) and accelerating RNase E-dependent ompD mRNA decay (Pfeiffer *et al.*, 2009). Iron responsive sRNA RyhB of *E. coli* is another example that base pairs with the RBS and the start codon of *iscS*, present within the *iscRSUA* operon encoding proteins for biosynthesis of Fe-S clusters (Desnoyers *et al.*, 2009).

1.3.3. Translational activation

The inherent intramolecular base pairing within mRNA by self pairing prevents the ribosomal binding for translation. Base pairing of sRNA with such a mRNA overcomes the formation of an inhibitory secondary structure formed by intramolecular base pairing (Morfeldt *et al.*, 1995). This duplex formation of sRNA:mRNA leads to remodelling of the mRNA structure which unblocks the TIR making it accessible to ribosomes and allowing translation (Fig. 1) (Majdalani *et al.*, 2005). A recent study indicated that sRNA when bound to the 5' UTR of the target mRNA, can elevate its expression by bringing about the nucleolytic cleavage on the 5' UTR which in turn stabilizes the mRNA and allow translation. In *C. perfringens* 3' end of VR small RNA binds to the leader sequence of toxin encoding collagenase *colA* mRNA, resulting in cleavage of nucleotides by ribonucleases and a conformational change of the RNA secondary structure which in turn upregulates its expression by bringing about stabilization of mRNA transcripts (Obana *et al.*, 2010).

1.3.4. Transcription termination

sRNA can cause transcriptional termination/anti-termination, leading to increase or decrease in the levels of expression of downstream genes (Waters & Storz, 2009). Attenuators are 5'-*cis* acting regulatory regions which fold into one of the two alternative RNA structures determining the success of transcription. The folding is modulated by producing either a Rho-independent terminator, resulting in interrupted transcription or an anti-terminator structure, resulting in a functional RNA transcript. Rho independent transcription terminators are characterized by a stable G/C-rich stem-loop structure followed by a stretch of unpaired Us that destabilize the RNA polymerase, causing its dissociation from DNA. Transcription attenuators involve a wide range of mechanisms and are used to couple the sensing of an environmental change with transcription termination. These include riboswitches, T-boxes (tRNA sensing elements) and transcription attenuation by RNA binding proteins (Desnoyers *et al.*, 2013). For example, the interaction between *cis*-encoded antisense sRNA and the *fatDCBAangRT* mRNA leads to transcription termination after the *fatA* gene, thus reducing expression of the downstream *angRT* genes (Stork *et al.*, 2007).

1. REVIEW OF LITERATURE

1.4. Identification of small noncoding RNAs in bacteria

RNAs are now documented well as major regulators of gene expression, affecting the fate of mRNAs at different levels. These sRNAs participate in regulatory pathways that allow the bacteria to adapt to various stresses, alter the metabolism need for cell growth and express timely virulence genes that enable pathogenic bacteria to adapt during the infection process. In a last few decades, with an extensive *in silico* and experimental strategies to identify regulatory RNAs in intergenic region have lead to the discovery of a large number of diverse small noncoding RNAs in several bacteria.

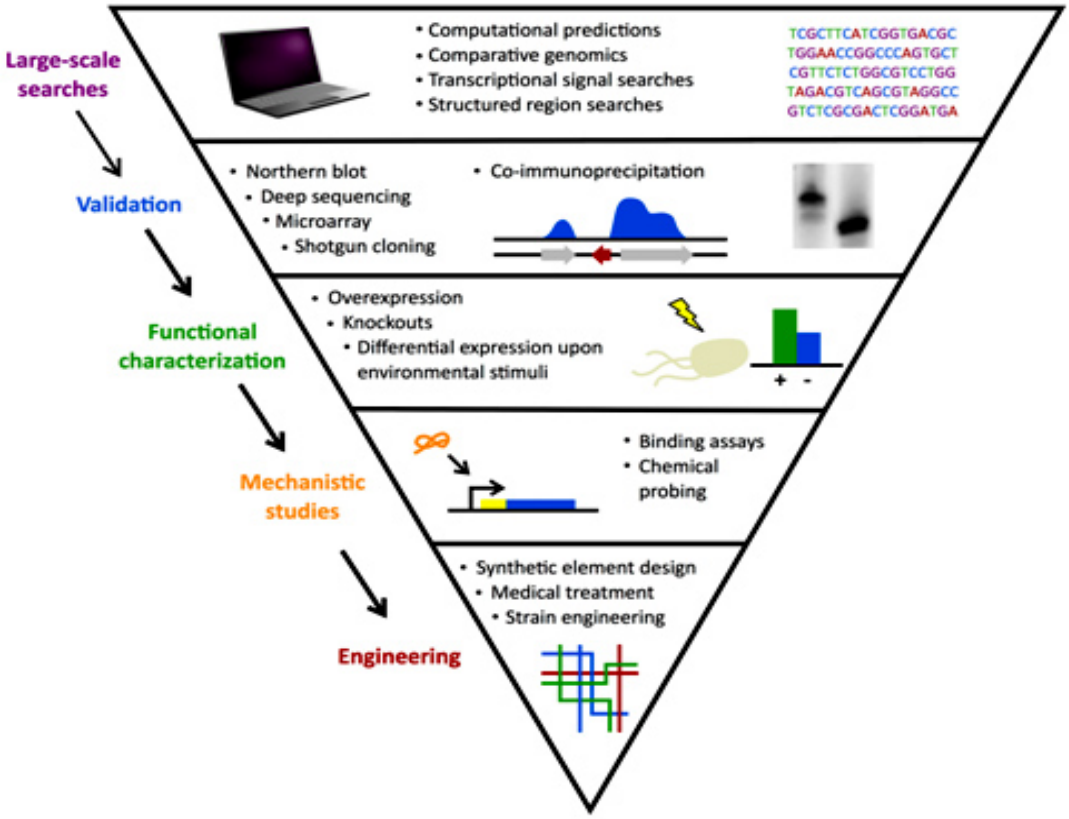

Fig. 2. Approach of sRNA identification and characterization (adapted from Haning *et al.*, 2014). Discovery of sRNAs begins with large scale computational searches followed by experimental validation. Functional characterization of confirmed candidates identifies their gene or protein targets and mechanistic studies elucidate their methods of action. Finally, sRNAs can be used in engineering efforts to develop useful applications from synthetic elements to medical treatments.

The first bacterial sRNA was discovered almost 50 years ago in *Escherichia coli* MRE600 (ATCC29417), but its function remained unknown for three decades, till the complete genome of *E. coli* K-12 was published (Hindley, 1967; Wassarman and Storz, 2000). At this time, 10 sRNAs were known in *E. coli*, mostly discovered by chance during studies of individual genetic systems. However, availability of the fully sequenced *E. coli* K-12 genome led to the booming of computational approaches, enabling large scale systematic searches of intergenic regions for sRNAs.

Recent developments of experimental approaches such as high density microarrays and RNA sequencing technologies have allowed genome wide searches of sRNAs and their functions in an extensive manner. Now, more than hundreds of sRNAs are known to exist that are involved in regulating expression of various genes in *E. coli* and other organisms such as *S. aureus* (Sassi *et al.,* 2015), *Salmonella sp.* (Hebrard *et al.,* 2012), *Mycobacterium sp.* (Haning *et al.,* 2014). These sRNAs show significant sequence conservation among other organisms, including pathogenic bacteria. Thus the powerful approaches used to detect sRNAs that include biocomputational screening and experimental techniques (Fig. 2) have allowed researchers to discover hundreds of sRNAs in several organisms studied so far are discussed below.

1.4.1. Bioinformatic approaches to identify noncoding RNAs

A biocomputational approach was used as the first systematic genome wide screen to predict sRNA genes (Altuvia, 2007) in the intergenic regions of *E. coli.* The major limitations of computational prediction of sRNA is the lack of known and recognizable features, as do for exon coding genes i.e. conservation of the open reading frame (ORF) and they are diverge in sequences, structures, functions and there are no common identifiers for bacterial sRNAs (Backofen & Hess, 2010).

The basic principles of computational methods for sRNA gene finding in bacteria are:

1. Prediction of RNA transcriptional signals which include:

   a) Orphan promoters b) terminators and c) lack of potential small ORFs

2. Sequence content

   a) Secondary structure stability and b) base composition

3. Comparative genome analysis

a) Conservation of sRNA sequence between related species and b) conservation of secondary structure by compensatory base change (Moore, 2000; Vogel & Sharma, 2005) (Fig. 3).

These computational algorithms rely on transcriptional signals, sequence or structure conservation in IGRs and are likely to miss unique sRNA transcripts that are longer in length, as some sRNAs are processed from larger transcripts (Altuvia, 2007).

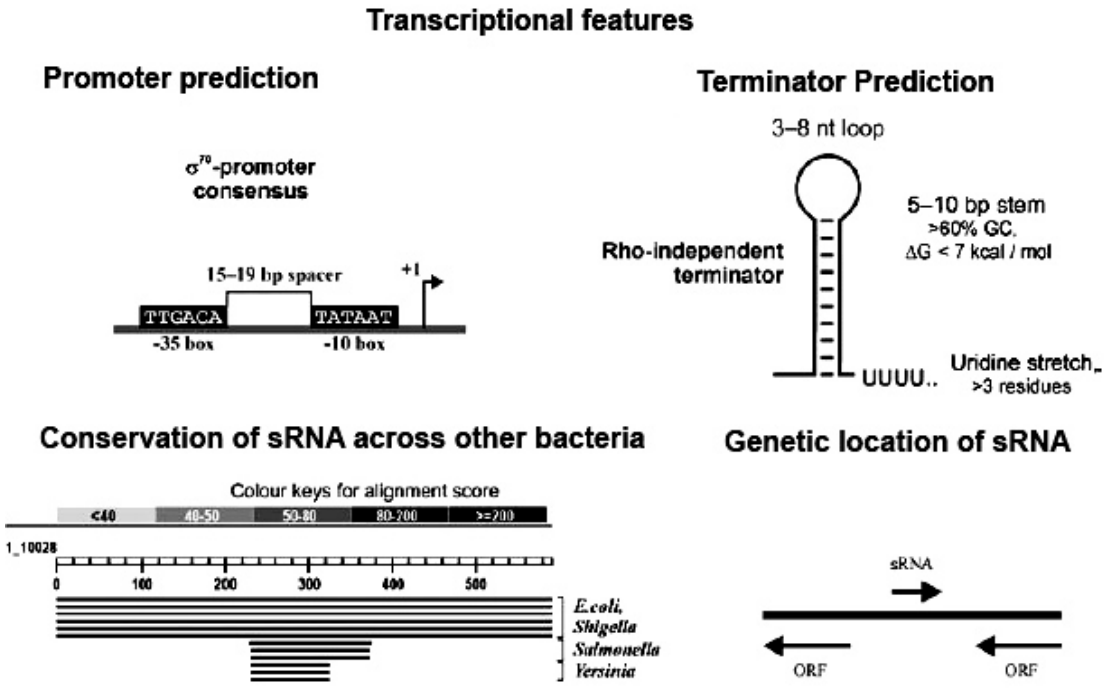

Fig. 3. Biocomputational strategy of sRNA prediction. The strategy is based on the transcription features of non coding RNA genes and includes sequence conservation of intergenic regions (IGRs), the genomic location of putative sRNA genes, scanning of sigma 70 promoter consensus and rho independent terminator sequences having GC rich stem loops followed by atleast four U residues (adapted from Vogel & Sharma, 2005).

Several biocomputational tools are available to predict the sRNAs present in the intergenic regions of the genome. These include QRNA, Intergenic Sequence Inspector, the RNAz, sRNAPredict, SIPHT, TransTermHP programs, sRNAscanner tool, ERPIN (Easy RNA Profile IdentificatioN). These programs basically use the intergenic sequence conservation among related genomes, thermodynamic stability of conserved RNA structures, orphan promoter and rho independent terminator annotations for sRNA predictions (Livny *et al.*, 2005; Backofen & Hess, 2010). ERPIN, an algorithm used to define RNA motifs by using multiple sequence alignments and secondary structure

consensus (Gautheret & Lambert, 2001), predicts the possible stem loop formation of the given sequence.

1.4.2. Experimental RNomics to identify small noncoding RNAs

Several technologies have been used so far for the identification and classification of wide range of small noncoding RNAs, which redesigned the entire gene regulatory mechanism of the cell. Some of the experimental approaches to profile cellular sRNAs include: (i) direct labelling (ii) cDNA cloning followed by sequencing (iii) hybridization based detection (iv) genomic SELEX (Systematic Evolution of Ligands by EXponential enrichment) and (v) RNA direct sequencing (reviewed in Vogel & Sharma, 2005; Huttenhofer & Vogel, 2006; Altuvia, 2007; Sharma & Vogel, 2009; Beaume *et al.*, 2010; Bohn *et al.*, 2010) (Fig. 4).

- **Direct labeling and sequencing:** The first bacterial small RNAs were identified by gel fractionation of metabolically labeled *E. coli* total RNA by radioactive orthophosphate ($^{32}PO_4^{3-}$) or γ-[^{32}P] ATP, [^{32}P] pCp followed by resolving on 1-D or 2-D PAGE and sequence determination by fingerprinting. Further the sequence is analysed directly in BLASTN searches. The major drawback of this method is that it does not distinguish between sRNAs and abundant processed fragments of rRNAs or tRNAs and only the abundant sRNAs are visualized. Besides, the efficiency of labeling may affect RNA structure and can lead to inaccurate representation of the expression pattern of sRNAs. The sRNAs discovered in this way include M1 RNA, RNase P, tmRNA, 4.5S RNA, 6S RNA and Spot 42 RNA.
- **cDNA or shotgun cloning (RNomics) and sequencing**: This method involves the generation of cDNA libraries from total RNA isolated at different stages of an organism and separated based on its size on denaturing PAGE, reverse transcribed and cloned into standard vector, followed by sequencing and assembling of overlapping sequences. In addition, Vogel *et al.,* (2003) pre-selected cDNA clones prior to sequencing by spotting on high density filters and hybridizing with rRNA and tRNA probes to exclude such clones from further study. A major limitation of this approach is that it may exclude the identification of longer ncRNAs, since the cut-off by size is (e.g. 20–500 nt) and it may not always be possible to reverse transcribe an ncRNA into cDNA because of modification of structural bases or its backbone. This method also employs cDNA libraries generated from RNAs, which are not selected by their size but

rather based on their function or binding to a common RNA binding protein, Hfq, which can be isolated by co-immunoprecipitation using an antibody against Hfq. This approach was followed for finding ncRNAs in the eubacteria *E. coli* and *A. aeolicus* and plant *A. thaliana*, fruit fly, *D. melanogaster*, archaeal species *A. fulgidus* and *S. solfataricus* and mouse *M. musculus* (reviewed in Huttenhofer and Vogel, 2006).

- **Microarray analysis**: This is the most powerful and preferred method to monitor the levels of expression of many sRNA transcripts, in parallel with condition dependent sRNA expression patterns of the whole genome of an organism. Microarray for ncRNA identification was first developed for model organism *E. coli* which has intergenic region (IGR) probes for both the coding strands of all ORFs, tRNAs and rRNAs of this genome and later it was also used in other organisms such as *S. aureus* and spore forming *B. subtilis.* Such DNA arrays were also used to identify *E. coli* sRNAs that specifically bind to Hfq using the RNAs isolated by co-immunoprecipitation. The samples used for hybridization are the extracted RNA or the converted cDNA and these probes are generally labelled with fluorescent dyes such as Cy3 or Cy5.

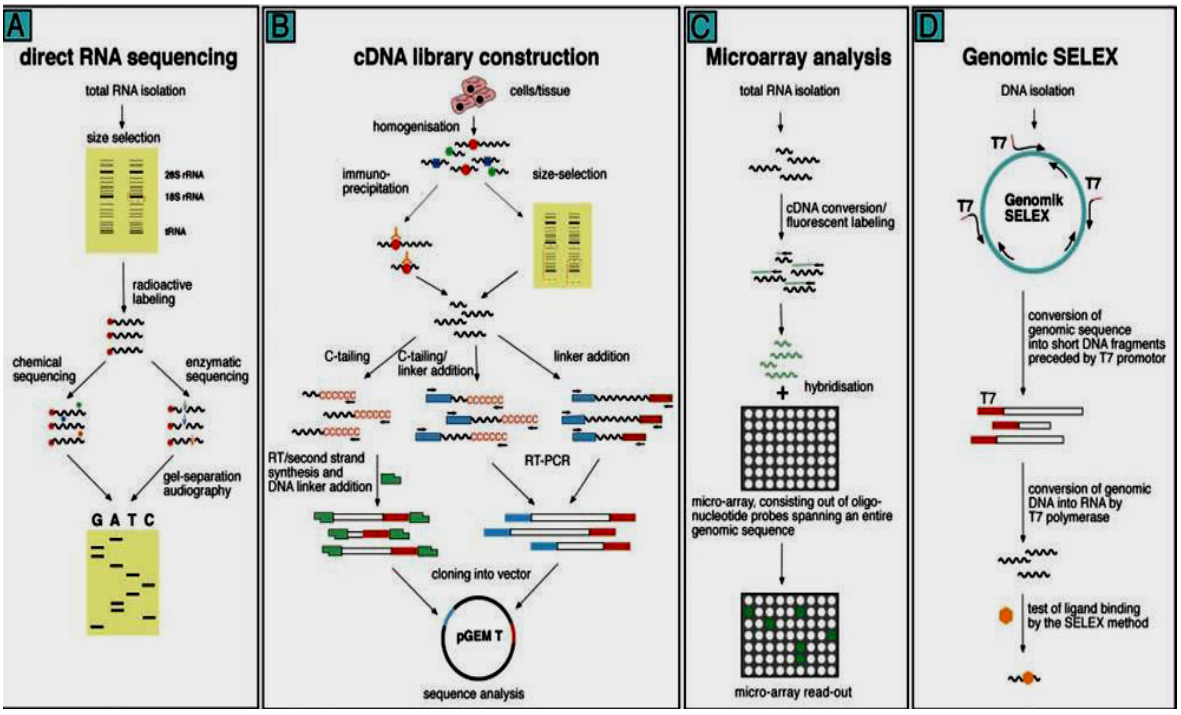

Fig. 4. Experimental approaches to identify ncRNA candidates (Adapted from Huttenhofer & Vogel, 2006).

This technique combined with comparative genome analysis, has led to the identification of several growth phase dependent sRNA genes. The drawback of this method are poor detection of shorter sRNAs (< 50 nt), especially if the probes are not closely spaced, the highly structured or modified sRNAs are likely to be missed as they are poor substrates for amplification and labeling (reviewed in Altuvia, 2007).

- **Genomic SELEX- Protein binding sRNAs:** The genomic SELEX approach which identifies sRNAs that require proteins such as Hfq for their activity to modify the mRNA or protein targets, was studied in the laboratory of Schroeder (Lorenz *et al.*, 2006), to identify Hfq binding sRNAs in *E. coli*. The RNAs CsrB of *E. coli* and RsmZ of *P. fluorescens* were identified by co-purification with their target proteins, CsrA and RsmA respectively. The sequence of a genome is converted into short PCR fragments containing a T7 promoter at their 5' end and subsequently, *in vitro* transcribed into RNA fragments. The generated RNA pool undergoes successive rounds of association by incubating with a given RNA-binding protein, partitioning and re-amplification. As a result, RNA sequences that are stringently bound by the protein partner get enriched. The sequence of the bound RNAs is determined and is searched for matches in the genome. The major drawback of this system is that RNA should remain tightly associated with the protein throughout the purification and co-immunoprecipitation requires highly specific antibodies (Vogel and Sharma, 2005).
- **RNA pyrosequencing:** Pyrophosphate based sequencing technique also called as 454 pyrosequencing avoids the complication of bacterial cloning and thus enables high throughput parallel sequencing of hundreds to thousands of cDNA fragments. The combination of RNomics with the high throughput pyrosequencing technology can therefore be more productive in identifying large number of small RNAs (Altuvia, 2007). The Pyro/deep/next generation sequencing approaches usually use short DNA/RNA sequence mixtures, ligate adapters and consider each fragment separately for sequencing.

1.5. Computational prediction of mRNA targets for sRNA

sRNA regulates the target mRNAs by complementary base pairing at single or multiple locations of the mRNAs (Vogel & Wagner, 2007), bringing about positive or negative effects to the cell. The hybridization energy -ΔG energy of the binding complex is the widely used criteria to predict RNA-RNA interactions (Rehmsmeier *et al.*, 2004; Tjaden

et al., 2006). Several prediction models have been developed based on the above principle. These models first extract the rules from trained data set composed of known, experimentally verified mRNA targets for sRNA in the model organisms for their binding ability and then apply the rules to predict targets for other sRNAs in the same or other organisms whose targets are unknown.

- **TargetRNA**, an online target prediction tool that uses dynamic programming of Smith Waterman algorithm to search for mRNAs in a specified genome and outputs a ranked list of candidate mRNA targets along with the predicted base pairing interaction. The predicted targets have been validated experimentally in several organisms such as *V. cholerae* (Davis & Waldor, 2007), *N. meningitidis* (Mellin *et al.*, 2007) and *Salmonella* (Sharma *et al.*, 2007). However, other factors that may contribute to sRNA:mRNA interactions, such as RNA secondary structure or the role of Hfq protein, are not considered by this program (Tjaden, 2008).
- **TargetRNA2**, a web server that considers a variety of features, including conservation of the sRNA in other bacteria, the secondary structure of the sRNA and each candidate mRNA target, the hybridization energy between them and then gives a ranked list of likely regulatory targets for the given sRNA (Kery *et al.*, 2014).
- **RNApredator** uses the RNAplex dynamic programming approach, which runs faster than RNAup or IntaRNA to compute putative targets. RNApredator considers the accessibility of the targets' ribosomal binding site upon binding of sRNA during the mRNA target search and improves the specificity of the predictions. It provides the hybridization energy for the opening of both the target and the sRNA sequences and the corresponding *Z*-score, which is useful for comparing interactions involving different sRNAs. The formed duplex structure of sRNA-mRNA is represented in dot-bracket format (Eggenhofer *et al.*, 2011).
- **RNAup and IntaRNA:** IntaRNA (**Int**era**c**ting RNAs) (Busch *et al.*, 2008) and RNAup (Muckstein *et al.*, 2006) are based on the RNA folding energy model and predict interacting regions between two RNA molecules by considering the accessibility of both interaction site on both the RNAs and the presence of a seed interaction. The major drawback is the inability to handle more than a pair of sequences at a time, making it difficult for genome wide target search (Busch *et al.*, 2008).

- Other RNA-RNA interaction prediction tools such as sTarPicker (Ying *et al.,* 2011), sRNAtarget (Cao *et al.,* 2009), CopraRNA (Comparative prediction algorithm for small RNA targets) (Wright *et al.,* 2014) or biRNA (Chitsaz *et al.,* 2009) are also being used to predict sRNA targets in bacteria.

1.6. Classification of small noncoding RNAs

Bacterial regulatory RNAs are generally classified into three main groups:

A) *cis*-encoded antisense RNAs (asRNAs), which overlap and are completely complementary to their target genes encoded on the opposite DNA strand at the same genomic locus.

B) *trans*-encoded small RNAs (sRNAs), which are defined as regulators of one or many target genes located elsewhere on the chromosome and elements that are present in the 5′ UTR of the mRNA which they regulate (for example, riboswitches, thermo sensors and pH sensors) and

C) Small RNAs that bind to proteins.

1.6.1. *cis*-encoded sRNAs

cis-encoded sRNAs (also called antisense RNA, asRNA) vary greatly in size and were first discovered in bacterial plasmids, where the RNAs modulate the expression of the genes involved in replication and stable plasmid inheritance (Gisela *et al.,* 2005; Georg & Hess, 2011). Most of the *cis*-encoded sRNAs are expressed constitutively throughout the cell growth (reviewed in Waters & Storz, 2009).

1.6.1.1. Plasmid encoded *cis*-acting sRNAs

In 1981, Jun-Ichi Tomizawa identified the non-coding RNA, RNAI (~108 nt) (Tomizawa & Itoh, 1981; Tomizawa *et al.,* 1981) that controls the copy number of plasmid ColE1 by preventing RNAII (~150 nt) processing that generates replication primers. An increase in the copy number of the plasmid pT181, results in elevated levels of RNAI and RNAII antisense RNAs. These RNAs base pair and stabilize a structure associated with transcription termination upstream of RepC, a protein required for replication initiation. When the plasmid copy number is low, RNAI and RNAII levels decrease, thereby allowing transcription read through, leading to increased RepC levels and renewed replication (Storz *et al.,* 2005).

Later on, Nordstrom and colleagues identified ~ 90 nt long CopA RNA, which controls the copy number of plasmid R1 by regulating the translation of the RepA replication initiator protein (Stougaard *et al.,* 1981). This system is called as the plasmid addiction system, which ensures that plasmid containing cells survive, whereas cells that do not contain plasmid are killed. The R1 plasmid encoded antisense Sok (**S**uppressor **o**f **k**iller) RNA of ~70 nucleotide which binds and represses *hok* (host killing) mRNA which encodes a small protein that damages the bacterial membrane leading to cell death (Storz *et al.,* 2005).

One of the first antisense RNAs discovered was the 70 nucleotide RNA-OUT of the transposon Tn*10* that affected transposition by repressing transposase synthesis. In addition to that, 69 nucleotide Sar RNA of bacteriophage P22 and the 77 nucleotide OOP RNA of bacteriophage λ were reported to repress synthesis of the Ant and cII phage proteins, respectively (Thomason & Storz, 2010).

1.6.1.2. Chromosomally encoded *cis*/*trans*-sRNAs

1.6.1.2.1. Riboswitches

Riboswitch, a part of the coding mRNA molecule and genetic control elements found within 5' UTRs, regulates gene expression in response to binding of small molecules such as metabolites, amino acids or co-factors (Nudler & Mironov, 2004).

Table 1. Riboswitches mediated genetic control in several organisms (adapted from Winkler & Breaker, 2005).

Riboswitch	**Regulatory functions**
Coenzyme B12 riboswitch	cobalamin synthesis and transport, aerobic and anaerobic ribonucleotide reductase, glutamate/succinate fermentation and several uncharacterized genes
Flavin mononucleotide (FMN) riboswitch	riboflavin biosynthesis and transport
Thiamine Pyro Phosphatase (TPP) riboswitch	thiamine synthesis, phosphorylation, and transport
Lysine riboswitch	lysine synthesis and transport, lysine catabolism
S- adenosyl methionine (SAM) riboswitch	methionine/cysteine biosynthesis, methionine recycling, methylene tetrahydrafolate reductase, SAM synthesis, metabolite transport
Purine riboswitches	purine synthesis and transport, and several uncharacterized genes
Glycine riboswitch	glycine catabolism and efflux
Glucosamine 6- phosphate (GlcN6P) riboswitch	synthesis of GlcN6P

It affects transcriptional termination, translation repression/initiation and mRNA stability via changes in RNA structure (Fig. 5). A list of riboswitches and their regulatory function are mentioned in Table 1. Riboswitches are known in bacteria, but have also been found in eukaryotes. Riboswitches may regulate 2% of genes in gram positive pathogens such as *S. aureus* and *L. monocytogenes* in response to metabolites and pH (Caldelari *et al.*, 2013). A class of riboswitches senses the second messenger, cyclic di-guanosine monophosphate (c-di-GMP), which is associated with cell differentiation, biofilm formation, and virulence. It is also associated with GpbA, a protein attaching *V. cholerae* to human epithelial cells and zooplankton, suggesting a role of RNA based sensing of c-di-GMP in cholera pathogenesis (Papenfort & Vogel, 2010).

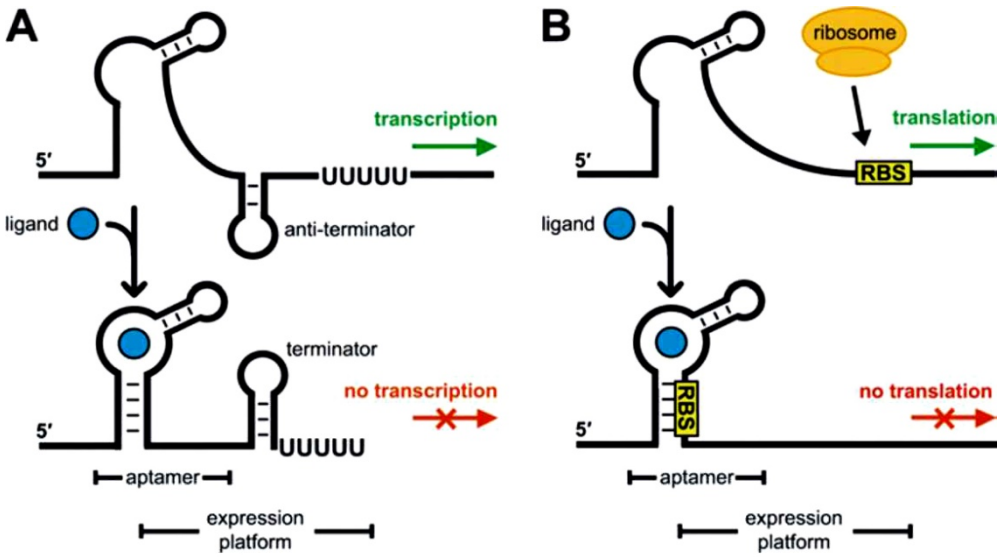

Fig. 5. Structural interchanges by riboswitches to regulate transcription termination and translation initiation (Adapted from Kim & Breaker, 2008). A: pre-mature transcription termination B: inhibition of translation initiation.

1.6.1.2.2. Thermosensors

RNA thermometers (RNATs) usually reside in the 5' UTR of temperature responsive genes and also in intercistronic regions, where they differentially control gene expression (Krajewski and Narberhaus, 2014) and mediate the switch between two distinct structures: a closed conformation formed at low temperature, in which the SD and/or AUG are inaccessible to 30S ribosomes; and the open conformation formed at

high temperature upon melting of the inhibitory structure of RNA around the RBS leading to translational activation (Papenfort & Vogel, 2010).

cis-acting RNA thermosensors were initially discovered in genes encoding heat shock proteins (HSP) and include the *hspA* ROSE1 (repression of heat shock gene expression) element in *B. japonicum*, the *agsA* four U elements (four consecutive uridines pair with the SD) in *S. enterica* and the *hsp17* bipartite hairpin found in *Synechocystis* spp (Klinkert & Narberhaus, 2009). The first RNA thermometer was studied in *Y. pestis*, which resides in the 5' UTR of *lcrF* mRNA encoding a transcriptional activator of virulence genes (Hoe & Goguen, 1993). *L. monocytogenes* is another example where in a 127 nt hairpin in the 5' UTR of *prfA* inhibits translation of the mRNA below 37°C. PrfA is a transcription factor responsible for the switch of the bacterium from being saprophytic to virulent and activate the genes required for bacterial invasion, host cytosolic propagation and transmission to adjacent cells (Freitag *et al.*, 2009).

1.6.2. *Trans*-encoded sRNAs

The most extensively characterized sRNAs are the *trans*-acting sRNAs. Unlike *cis*-encoded sRNA, *trans*-encoded RNAs are expressed under specific stress conditions. Typically *trans*-encoded sRNA base pair with multiple mRNAs as targets, but the major drawback of *trans*-encoded sRNAs is the limited stretch and discontinuous complementarity with their target mRNAs than complete stretch and perfect complementarity which is observed in *cis*-encoded antisense sRNAs. These sRNAs regulate the translation and/or stability of target mRNAs and are functionally analogous to eukaryotic miRNAs.

The majority of the regulation by the known *trans*-encoded sRNAs is negative. Base pairing between the sRNA and its target mRNA usually leads to repression of protein levels through translational inhibition, mRNA degradation, or both. However, *trans*-encoded RNA can also activate expression of its target mRNAs through an anti-antisense mechanism whereby base pairing of the sRNA can disrupt an inhibitory secondary structure, which sequesters the ribosome binding site (reviewed in Waters & Storz, 2009; Papenfort & Vogel, 2010).

1.6.2.1. Hfq mediated *trans*-sRNA regulation

The functions of *trans*-encoded base pairing RNAs are generally dependent on the RNA chaperone Hfq, the RNA binding protein responsible for maintaining the function and/or stability of this family of sRNAs in gram negative bacteria. Hfq generally binds to an A/U-rich single stranded region often located adjacent to a stem-loop structure (Storz *et al.,* 2011). Hfq binds both sRNAs and mRNAs, *in vitro*, and stimulates their pairing. However, a number of questions remain unanswered about how this doughnut shaped hexamer facilitates interactions between the RNAs (Gottesman & Storz, 2010). This protein is a homolog of the Sm and Sm-like proteins that form the core of splicing and mRNA degradation complexes found in eukaryotic and archaeal cells. Hfq binding promotes structural changes in ncRNA or mRNA accessibility to different RNases such as RNase E. The involvement of Hfq in stabilizing the interaction of sRNA and mRNA was mostly observed in gram negative organisms such as *N. meningitidis, Salmonella, V. cholerae* and *P. aeruginosa*. Hfq has been shown to be required for pathogenesis control and deletion of hfq leads to strong virulence defects in many bacterial pathogens (Chao & Vogel, 2010; Liu *et al.,* 2010). The binding of Hfq protein to noncoding RNAs in gram positive bacteria, is also reported. Although Hfq in *S. aureus* binds to interaction complexes such as RNAIII- *spa* mRNA (Huntzinger *et al.,* 2005), SprA1-SprA1$_{AS}$ (Sayed *et al.,* 2011), ArtR- *sarT*, ArtR- *hla* (Xue *et al.,* 2014), it is not involved in the regulation of mRNA targets as their regulation is unaffected in *hfq* deletion mutants, notably in the strains RN6390, COL and Newman (Storz *et al.,* 2005; Bohn *et al.,* 2007; Liu *et al.,* 2010).

1.6.2.2. Protein binding *trans*-sRNAs

This class of untranslated RNA molecules acts by interaction, not with a messenger RNA (mRNA), but with a protein. **C**arbon **s**torage **r**egulator CsrB and CsrC, the two regulatory sRNAs regulate the activity of the 61 aa CsrA, an RNA binding protein by sequestering up to 18 molecules of CsrA, thus preventing it from binding to its target mRNAs (Lucchetti-Miganeh *et al.,* 2008) (Fig. 6). When CsrB and CsrC are present at high levels, CsrA binds to them and therefore is not available for interacting with the target mRNAs (Gottesman & Storz, 2010). CsrA acts as a post transcriptional regulator mediating switch between gluconeogenesis and glycolytic growth, inhibiting glycogen synthesis (Babitzke *et al.,* 2009). CsrA prevents translation of several target mRNAs by

binding to the Shine-Dalgarno sequence, thus blocking ribosome binding and facilitating mRNA decay. CsrA has also been shown to act as a positive regulator by stabilizing and subsequently increasing the translation of target mRNAs (Lucchetti-Miganeh *et al.*, 2008).

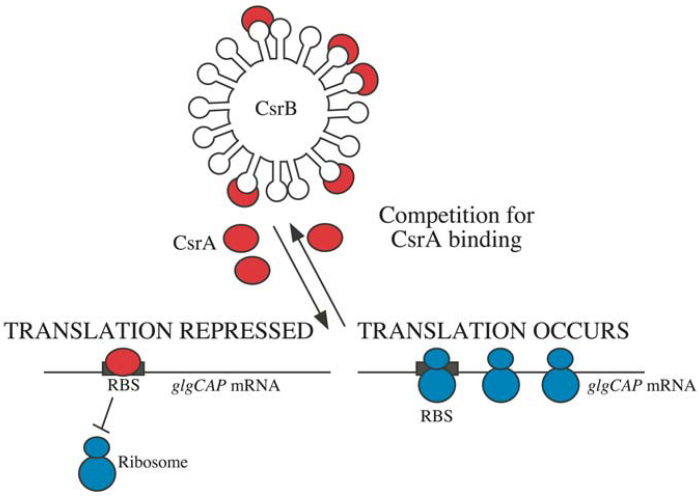

Fig. 6. Regulatory mechanism of CsrB/C with CsrA. Higher levels of CsrB compete successfully for binding to CsrA, hence allow the mRNA translation (Adapted from Majdalani *et al*., 2005).

This Csr system was first described in 1993 by T. Romeo and colleagues in *E. coli*. Its homologs also exist in broad range of γ-proteobacteria. Csr system mediates adaptive physiology and timed virulence trait expression in extracellular pathogens (*P. aeruginosa*, *V. cholerae*, *H. pylori*) and intracellular pathogens (*L. pneumophila*, *S. enterica serovar typhimurium*) at different stages of infection (colonization, persistence and pathogenicity), where they impact secondary metabolism, quorum sensing, biofilm formation, flagellar synthesis and motility, chemotaxis, stress resistance, macrophage infection, intracellular replication and epithelial cell invasion (Babitzke & Romeo, 2007; Lapouge *et al.*, 2008; Lucchetti-Miganeh *et al.*, 2008).

1.6.3. CRISPR RNAs

The prokaryotic adaptive immune system encoded by **C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats (CRISPR) and their associated proteins (Cas) provides resistance to bacteriophages and prevents plasmid conjugation. CRISPR sequences are highly variable DNA regions that consist of a 550 bp leader sequence constituting an array of repetitive sequences separated by equally short unique intervening (30-40 bp long) sequences called spacers. The computational search for spacer sequences on

genebank, matches within the short regions of the genome of viruses and plasmids of bacteria and archaea (Charpentier and Marraffini, 2014). The CRISPR arrays are transcribed as a long RNA that associates with cascade (Cas) proteins and forms ribonucleoprotein complexes (crRNPs), that target the DNA or RNA of invading genomes (adaptive immune function against mobile genetic elements, bacteria and archaea) and lead to the degradation of the exogenous DNA and are also reported to target mRNAs (regulatory function in endogenous gene expression e.g., *F. novicida*). The first experimental evidence of an adaptive immune function of this system was demonstrated in type II-A in *S. thermophilus* (Barrangou *et al.*, 2007).

CRISPR loci CRISPR-Cas is classified into three main types with eleven subtypes. Type I and III CRISPR-Cas systems can be found in bacteria and archaea, that require a large number of ribonucleoprotein complexes which includes the CRISPR RNAs (crRNA) and a second additional *trans*-acting RNA called tracrRNA (Deltcheva *et al.*, 2011) with the association of CRISPR- associated or Cas9 proteins to degrade foreign DNA. Type II CRISPR-Cas, which has only been identified in bacteria, requires only a single Cas protein (Cas9) and two small RNAs (the crRNA guide and the tracrRNA) that work together as ribonucleoprotein complexes to interfere with the invading nucleic acids.

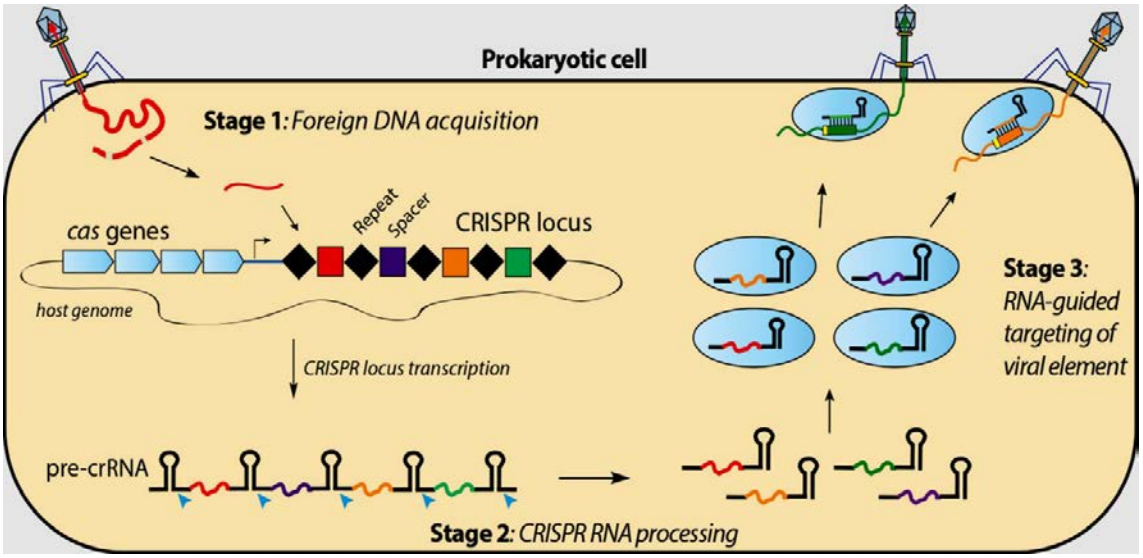

Fig. 7. Bacterial Adaptive Immunity. In the acquisition phase, foreign DNA is incorporated into the bacterial genome at the CRISPR loci. CRISPR loci is then transcribed and processed into crRNA during crRNA biogenesis. During interference, Cas9 endonuclease complexes with a crRNA and cleaves foreign DNA containing a 20-nucleotide crRNA complementary sequence (Adapted from The Doudna lab http://rna.berkeley.edu/crispr.html).

Similar to the restriction-modification systems, CRISPR-Cas systems provide a tool for sequence-specific cleavage of nucleic acids (Fig. 7). The type II CRISPR-Cas9 influences the virulence potential of the human pathogens such as *S. pyogenes and F. novicida* by limiting the acquisition of virulence genes carried on temperate phages. CRISPR systems share similarities with eukaryotic siRNA driven gene silencing. CRISPR-Cas9 provides a methodology for the introduction of mutations, genome targeting and genome editing in bacteria, site specifically silence mRNAs in cells. It may be beneficial in the understanding and treatment of human genetic disorders, cancers, HIV and other infectious diseases, thus proving to be powerful tool for genetic engineering (Jinek *et al.,* 2012; Charpentier, 2015).

1.7. Dual role of sRNAs

Many sRNAs in bacteria have dual functions. In addition to base pairing with mRNA targets they also serve to encode protein products. One of the first bifunctional RNA to be identified was RNAIII, a 514 nt in *S. aureus* that functions by base pairing with several mRNA targets and also encodes the small protein delta hemolysin (Novick *et al.,* 1993; Morfeldt *et al.,* 1995; Boisset *et al.,* 2007). The other example is of SprA1 sRNA which is encoded in the pathogenicity island of *S. aureus* and encodes a lytic peptide which is predicted to be a delta hemolysin.

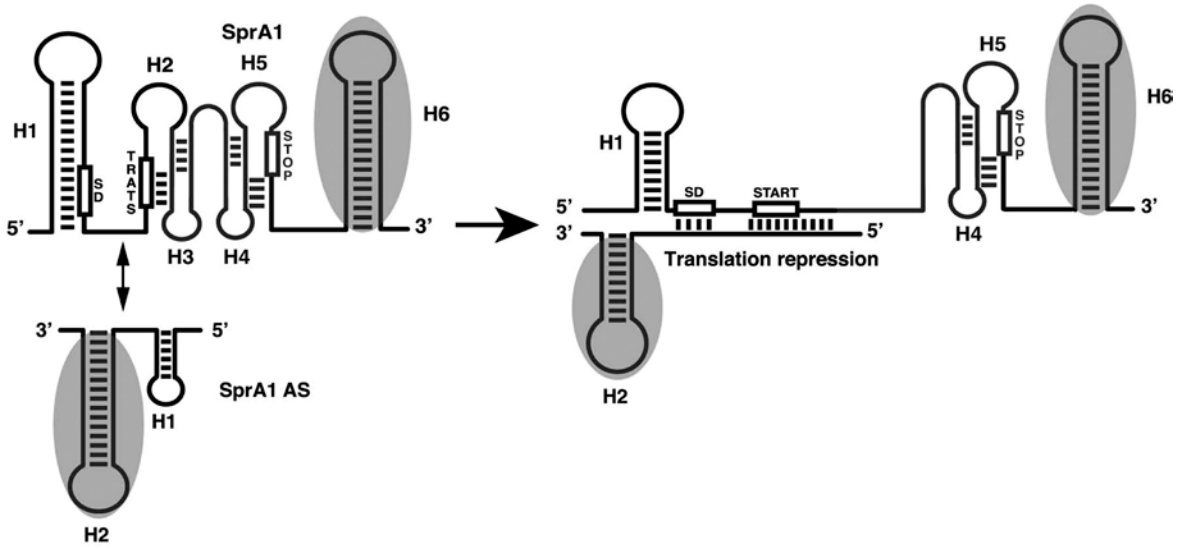

Fig. 8. Dual role of small RNA SprA1. Translation of the SprA1 encoded peptide is repressed by base pairing at internal translation initiation signals (SD sequence and start codon) with $SprA1_{AS}$ RNA (Adapted from Romilly *et al.,* 2012).

Translation of the SprA1 encoded peptide is downregulated by base pairing at internal translation initiation signals which includes SD sequence and start codon with *cis*-encoded antisense $SprA1_{AS}$ (Sayed *et al.*, 2011) (Fig. 8). Bifunctional sRNAs have also been reported in other bacteria. SgrS, a 227 nt sRNA is expressed during glucose phosphate stress in *E. coli* and encodes a 43 aa functional polypeptide namely SgrT whose function is to promote recovery from stress and negatively affect glucose transport (Wadler & Vanderpool, 2007).

1.8. Small RNAs in pathogenic bacteria

sRNAs facilitate pathogenicity by mediating the switch from saprophytic to virulent lifestyle and a fast adaptation to changing environments in the host. During infection, bacterial pathogens have to adapt rapidly to the changing environmental conditions in the host, and ensure survival strategies in specific niches, avoidance of exposure to the immune system and systemic toxicity. Among the multiple mechanisms of pathogenicity that bacterial pathogens have developed, regulatory small RNAs are considered as signal transducers of environmental cues by participating in the precise coordination of gene expression (Michaux *et al*., 2014). The sRNAs discovered in several pathogens thus exert far diverse functions that include regulation of transcription factors, virulence genes, quorum sensing and outer membrane dynamics in response to a variety of environmental stress such as temperature, pH, metabolites, oxidative and anaerobic stresses (Haning *et al.*, 2014). A list of small RNAs and their role in several bacterial pathogens is mentioned in Table 2. The application of synthetic RNA regulators to manipulate expression of essential genes or virulence pathways might provide a novel handle to limit human specific bacterial infections (Sharma & Heidrich, 2012). The transcriptome of a pathogen uncovers a wide range of novel noncoding RNAs with a specific pattern of gene expression *in vitro* and during infection. The abundance of sRNAs in infected tissues suggest that they can be used as biomarkers for disease identification (Arnvig *et al.*, 2011; Ignatov *et al.*, 2014).

Table 2. List of small RNAs involved in bacterial pathogenesis.

List of small noncoding RNAs involved in the pathogenecity of several organisms				
Organism	**sRNAs**	**Target mRNAs**	**Regulatory functions**	**References**
Clostridium sp.	VR-RNA	*colA, plc, ptp, cpd, ycgJ, metB, cysK, ygaG*	Transcriptional regulation of toxin encoding genes which codes for collagenase and alpha toxin, play a crucial role in the pathogenesis in *C. perfringens*	(Shimizu *et al.*, 2002)
	sCAC610	*CAC0528*	Upregulation of the expression of a putative ABC transporter, which in turn can bring about clindamycin resistance in *C. botulism, C. acetobutylicum* and *C. beijerinckii*	(Chen *et al.*, 2011)
Helicobacter pylori	HPnc5490	*tlpB*	Antisense repressor of chemotaxis receptor mRNA	(Sharma *et al.*, 2010)
	ureB	*ureAB*	Enhances the truncation of gastric acid acclimation operon *ureAB* at neutral pH, but releases its control in acidic environments to allow survival	(Wen *et al.*, 2012)
Leginonella pneumophila	RsmY,Z	RsmA	RsmY and RsmZ additively affect replication in macrophages via RsmA	(Rasis & Segal, 2009; Sahr *et al.*, 2009)
	6S	RNAP	Required for replication in macrophage and amoeba, regulates type IV secretion	(Faucher *et al.*, 2010)
	RliB	*lmo2104*	rliB mutation increases colonization of spleen in mice	(Toledo-Arana *et al.*, 2009)
	Rli38	Unknown	rli38 mutant is attenuated in oral mouse infection	(Toledo-Arana *et al.*, 2009)
	LhrA	*chiA*	Role in pathogenesis by targeting *chiA* gene which mediates the degradation of glycoproteins on the host cells and plays a role in infection	(Mraheil *et al.*, 2011)
	Rli50, Rli112	Unknown	Survival inside macrophages/ attenuation in the mice and the *Galleria mellonella* insect model	(Mraheil *et al.*, 2011)
	LhrC	*lapB*	Downregulation of *lapB* mRNA, an adhesin required for bacterial entry into mammalian cells and for virulence, affects the colonization of the intracellular niche	(Sievers *et al.*, 2014)

Listeria monocytogenes	Rli33, Rli33-1, Rli31	Unknown (Rli33-1), *pgdA*, *pbpX* (Rli31)	Regulates the intracellular growth in macrophages by affecting the pathogen's ability to propagate intracellularly Regulate enzymes involved in modification of peptidoglycan structure	(Mraheil *et al.*, 2011; Burke *et al.*, 2014)
	Rli55	ethanolamine utilization pathway genes	Controls the expression of ethanolamine utilization pathway genes involved in infection	(Mellin *et al.*, 2014)
Mycobacterium tuberculosis	ASdes	*desA1*, *desA2*	Induced upon bacterial uptake, regulate fatty acid desaturase genes , essential gene for the growth of mycobacterium	(Arnvig & Young, 2009)
	ASpks	*pks12*	Regulate polyketide synthase gene, implicated in bestowing antigenicity to the contagion	(Arnvig & Young, 2009)
	B11, G2, F6		Regulates the protein in cell division, overexpression prevents growth and lethal to the cell	(Arnvig & Young, 2009)
	MTS2823, MTS0997, MTS1338		Increased accumulation in the lungs of chronically infected mice, supports a role in the infection	(Ignatov *et al.*, 2014)
	RybB-1	*ompC*, *ompW*	Repression of porin synthesis	(Papenfort *et al.*, 2006)
	GcvB	*oppA*, *dppA*, *gltl*, *livK*, *livJ*, *argT*	Repression of peptide transport	(Sharma *et al.*, 2007)
	InvR	*ompD*	Invasion gene island (SPI-1) encoded sRNA targeting porin synthesis	(Pfeiffer *et al.*, 2007)
	IsrJ	unknown	Implicated in i) invasion of the *Salmonella* sp. into non phagocytic cells ii) injection of bacterial effector proteins into the host cytosol	(Padalon-Brauch *et al.*, 2008)
	IsrC	*msgA*	Antisense regulator of msgA virulence gene	(Padalon-Brauch *et al.*, 2008)
	SgrS	*ptsG*, *sopD*	Repressor of sugar uptake and also regulates secreted virulence factor	(Wadler & Vanderpool, 2009)
	tmRNA	damaged mRNA	Mutant attenuated in murine and macrophage infection	(Julio *et al.*, 2000; Ansong

Salmonella typhimurium			models	*et al.*, 2009)
	AmgR	*mgtBC*	Impacts on magnesium homeostasis and virulence in mice	(Lee & Groisman, 2010)
	IsrM	*hilE, sopA*	Upregulation in mouse organs, required for invasion and intracellular proliferation inside macrophages.	(Gong *et al.*, 2011)
	RyhB-1 (sRNA)	*cyoABC, cydB, cybC, nirBCD*	Upregulation of mRNA genes leading to increased sensitivity to nitrosilating and oxidative agents	(Padalon-Brauch *et al.*, 2008; Ortega *et al.*, 2012; Calderon *et al.*, 2014a)
	RyhB-2 (asRNA)	*yeaQ, cyoABC, cydB, cybC, nirBCD*	Upregulation of mRNA genes leading to increased sensitivity to nitrosilating and oxidative agents	(Padalon-Brauch *et al.*, 2008; Ortega *et al.*, 2012; Calderon *et al.*, 2014b)
Streptococcus pyogenes	FasX	*fasBCA*	Regulates i) fasBCA regulon required for the repression of adhesins FBP54 and MRP ii) activation of virulence factors such as streptokinase	(Kreikemeyer *et al.*, 2001)
	tmRNA	damaged mRNA	Increased expression upon contact with antibiotics	(Steiner & Malke, 2001)
	Pel	*emm, sic, speB*	Bifunctional RNA that also encodes the SagA protein	(Mangold *et al.*, 2004)
	FasX	*fpbA, mrp, ska, pel*	Increases the interaction of *S. pyogenes* with epithelial cells	(Klenk *et al.*, 2005)
	RivX	*mga*	Regulates expression of virulence transcription factors	(Roberts & Scott, 2007)
Vibrio cholera	CsrB/CsrC /CsrD	CsrA protein	Quorum sensing control via CsrA, involved in virulence	(Lenz *et al.*, 2005)
	Qrr1-4	*hapR, vca0939*	Quorum sensing biofilm regulator and de-repression of virulence genes	(Lenz *et al.*, 2004; Hammer & Bassler, 2007; Zhao *et al.*, 2013)
	VrrA	*ompA*	Outer membrane vesicle synthesis, colonization of mouse intestine	(Song *et al.*, 2008)

(reviewed in (Romby *et al.*, 2006; Toledo-Arana *et al.*, 2007; Papenfort & Vogel, 2010; Iyer *et al.*, 2012; Michaux *et al.*, 2014; Ortega *et al.*, 2014).

1.9. *Staphylococcus aureus* - A human pathogen

The gram positive bacterium, *S. aureus* is a prominent human pathogen and a leading cause of community and hospital acquired bacterial infections worldwide (Malachowa *et al.,* 2011). It is estimated that 20-30% of the human population is carrier of *S. aureus*. *S. aureus* harbors an arsenal of virulence factors to facilitate tissue adhesion, immune evasion and host cell injury. These factors cause inflammation, impair immune cell function, alter coagulation and comprise vascular integrity in the bloodstream. (Powers and Wardenburg, 2014). The pathogenesis of staphylococcal disease usually ranges from skin infections (e.g., wound infection, furuncle and cellulitis) to metastatic deep rooted infections such as endocarditis, osteomyelitis and septic arthritis, respiratory infections (Pneumonia), meningitis, toxic shock syndrome (TSS), arthritis, scalded skin syndrome and food poisoning (Schjorring *et al.,* 2002; Plata *et al.,* 2009).

1.9.1. Methicillin Susceptible *Staphylococcus aureus* (MSSA) strain Newman- a clinical isolate

S. aureus strain Newman was first isolated in 1952 from a case of secondarily infected tubercular osteomyelitis in man (Duthie & Lorenz, 1952) and has been widely used in studying the staphylococcal disease in animal models due to its virulent phenotypes. Genomic islands of Newman consist of prophages, staphylococcal cassette chromosome and pathogenicity islands which carry genes for integrase, DNA recombinases and several virulence factors. The family of staphylococcal pathogenicity islands that carry genes for superantigen toxins (SaPIs) are 15-20 kb elements located at constant positions in the chromosome (Plata *et al.,* 2009). The major pathogenicity island of strain Newman νSaα and νSaβ have played a major role in the evolution of this pathogen. The genome of strain Newman contains four prophages, ΦNM1 to ΦNM4. *S. aureus* Newman variants that lacked NM3 or NM1, NM2, and NM4, or all four prophages (NM1 to NM4) displayed dramatic reductions in their ability to form organ specific abscesses after intravenous infection of mice, suggesting their important roles during the pathogenesis of staphylococcal infections.

The chromosome of *S. aureus* Newman is 2,878,897 bp in size and it encodes 2,614 open reading frames. When compared to MRSA, *S. aureus* Newman harbors only a small number of insertion sequences (IS) and lacks known antibiotic resistance determinants. The absence of drug resistance gene SCC*mec* renders it susceptible to

methicillin and β-lactam antibiotics conferring *S. aureus* Newman a Methicillin susceptible *Staphylococcus aureus* (MSSA) (Baba *et al.*, 2008).

1.10. Virulent mechanism of *S. aureus*

The success of *S. aureus* as a pathogen is primarily due to its ability to produce a large number of virulence factors which participate in the pathogenesis of infection (Kubica *et al.*, 2008). Staphylococcal virulence factors are multifactorial (Fig. 9), categorized according to their functions: (i) surface proteins that promote adhesion, internalization, and colonization (ii) toxins and enzymes that promote tissue damage, inflammation, and invasion and dissemination (iii) surface factors that affect phagocytosis by leukocytes and enhance survival in phagocytes or (iv) super antigens and other molecules that modulate the immune system by altering the function of lymphocytes and antigen presenting cells (Malachowa *et al.*, 2011). An important feature of staphylococci is that a single virulence factor may have several functions in pathogenesis and multiple virulence factors may perform the same function as well (Harris *et al.*, 2002). Some of the virulence factors and their characteristics are listed:

1.10.1. Surface Adhesins/molecules

To initiate invasive infection, the pathogen *S. aureus* adheres to extracellular matrix of eukaryotic cells such as host fibrinogen, fibronectin, collagen and von Willebrand factor by asset of different surface proteins called adhesins MSCRAMM (stands for **m**icrobial **s**urface **c**omponents **r**ecognizing **a**dhesive **m**atrix **m**olecules) which promote the attachment of bacteria to extracellular matrixes, enables the bacteria to colonize and establish infection (McAleese *et al.*, 2001).

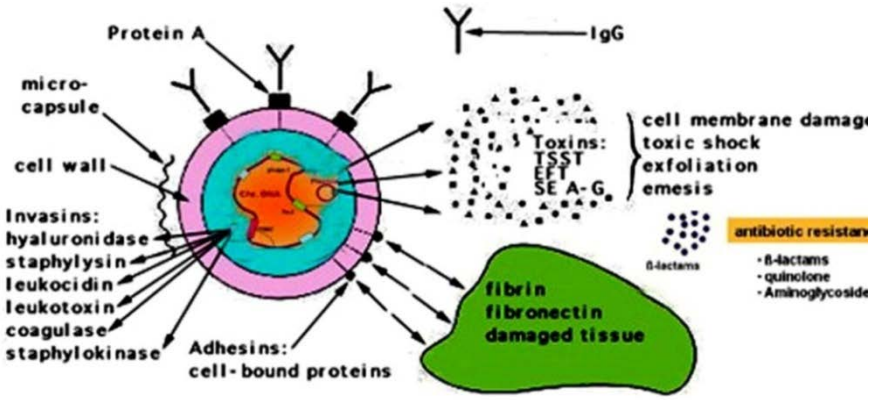

Fig. 9. Pathogenic factors of *Staphylococcus aureus* with surface and secreted products. (Web Review of Todar's Online Textbook of Bacteriology. *"The Good, the Bad, and the Deadly"*).

1.10.1.1. Clumping factor A/B

S. aureus expresses two structurally related surface associated fibrinogen binding MSCRAMMs, clumping factor ClfA and ClfB (McAleese *et al.*, 2001). These proteins are covalently attached to the cell wall and mediate adherence of bacteria to immobilized fibrinogen, blood clots, conditioned biomaterial *ex vivo* and to thrombi on damaged heart valves in a rat model of endocarditis (Moreillon *et al.*, 1995). They are important mediators of *S. aureus* induced platelet aggregation. A recent study indicates that the staphylococcal cell wall protein clumping factor B (ClfB) is required for biofilm formation under calcium depleted conditions (Abraham & Jefferson, 2012). ClfB acts as an important colonization factor by promoting adhesion to squamous epithelial cells *in vitro* (Wertheim *et al.*, 2008). When mice were immunized intranasally with *S. aureus* cells bearing mutant *clfB*, they showed reduced nasal colonization. Therefore, it may serve as an important component for vaccine development (Schaffer *et al.*, 2006). It also promotes clumping of bacteria in the presence of soluble fibrinogen. The great impact of ClfA expression on the virulence has been demonstrated in a murine model of septic arthritis (Josefsson *et al.*, 2001). ClfA expression also significantly protects *S. aureus* against macrophage phagocytosis (Palmqvist *et al.*, 2004). *clfA* and *clfB* mutants display defects in survival in blood, resistance to phagocytosis and lead to reduced staphylococcal load in organ tissues (Palmqvist *et al.*, 2004; Josefsson *et al.*, 2008).

1.10.1.1.1. Biofilm

Biofilm is one of the important defense mechanism of *S. aureus* and is formed on host tissue and medical implants such as stents, ventilators, intravenous catheters, cardiac defibrillators, artificial heart valves and prosthetic joints, aspirators, pacemakers, stitch materials, penile implants and orthopaedic devices which are in direct contact with normal blood and is very critical for chronic infections (Archer *et al.*, 2011). Biofilm formation involves multistep processes: initial adherence of cells to a surface, proliferation and accumulation to form multilayered cell clusters and finally detachment (Fig. 10). Biofilm formation can occur through the surface associated polysaccharide intercellular adhesin (PIA), which is synthesized by the icaADBC encoded proteins or PIA independent ways. Teichoic acids, proteinaceous factors, autolysin (*atl*), cell wall associated proteins SasG, Pls, ClfA/B, α-hemolysin, FmtA (a penicillin binding like protein that plays a role in methicillin resistance), extracellular DNA (eDNA) (DNA

from lysed cell), environmental factors such as glucose, ethanol, osmolarity, temperature and antibiotics such as tetracycline also contribute significantly to the attachment and/or biofilm formation and are required for cell-cell interactions (Resch *et al.*, 2005; Seidl *et al.*, 2008; Boles *et al.*, 2010; Archer *et al.*, 2011; Den Reijer *et al.*, 2016).

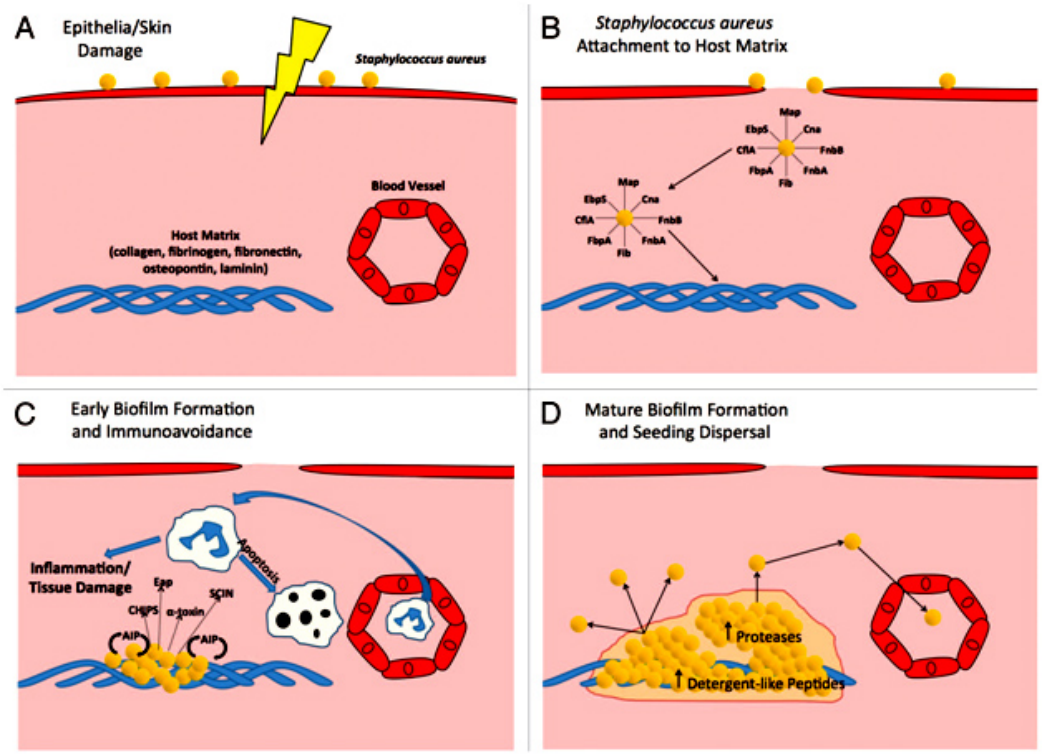

Fig. 10. Colonization of *Staphylococcus aureus* onto host epithelial surfaces (A) Colonization occurs by dissemination onto skin and various other host epithelial surfaces (B) Skin damage allows *S. aureus* to breach the epithelial layer and bind to host matrix via surface expressed colonization factors includes ClfA, clumping factor A and other factors (C) Initial attachment and cell division produces an early *S. aureus* biofilm and produces the quorum sensing compound auto-inducing peptide (AIP) leading to the upregulation of SarA and produces virulence. When AIP reached threshold levels, it results in *agr* activation and a downregulation of adhesins and an upregulation of virulence factor expression that cause damage to the host and evade the immune response. (D) Mature biofilm is encapsulated by polysaccharide intercellular antigen (PIA), protein and extracellular DNA (eDNA). Subsequently, protease and detergent like peptides are secreted into the biofilm and promote seeding dispersal and planktonic bacterial cells migrate from the mature biofilm into the circulatory system to repeat the cycle (adapted from Archer *et al.*, 2011).

Other factors include CcpA (catabolite control protein A), which regulates gene expression in response to the carbon source, tricarboxylic acid cycle (TCA) genes (*citZ*, *citB)*, *arlRS*, *mgrA*, *rbf*, staphylococcal accessory regulator (*sarA)* that are important to influence biofilm accumulation (Seidl *et al.*, 2008; Beenken *et al.*, 2003). Transcriptomic

profiling of entire genome of *S. aureus* during biofilm formation indicated differential gene expression under different growth phases of static and planktonic conditions (Resch *et al.,* 2005). Bacterial biofilms are resistant to host immune responses and antibiotic treatments (Joo & Otto, 2013).

1.10.1.2. Capsular polysaccharide

Capsular polysaccharide (CP) production was first recognized in 1930 by Gilbert. Over 90% of *S. aureus* strains produce atleast 11 capsular serotypes, but only type 1, 2, 5 and 8 have been chemically characterized. Most of the clinical isolates of *S. aureus* belong to either capsular types 5 or 8 (Lee *et al.,* 1994; Thakker *et al.,* 1998; Portoles *et al.,* 2001; O'Riordan & Lee, 2004) which makes the organisms resistant to phagocytic uptake and enhances virulence in animal models of infection. *cap5* and *cap8* are highly expressed on solid media than in liquid media. The ability of *S. aureus* to survive within the bloodstream and in abscesses, where phagocytic cells abound, has been linked to the expression and antiphagocytic properties of capsular polysaccharides. The global accessory gene regulator (*agr*) positively regulates capsular production of CP5 and CP8 both *in vitro* and *in vivo* (Van Wamel *et al.,* 2002).

1.10.1.3. Staphyloxanthin

The golden pigment staphyloxanthin (STX) of *S. aureus* is the product of a C30 triterpenoid carotenoid biosynthesis pathway. The genes are organized in an operon *crtOPQMN*, positively controlled by the alternative sigma factor σ^B (*sigB*) and cold shock protein (*cspA*) (Morikawa *et al.,* 2001). Carotenoid functions as antioxidant which shields the microbe from oxidation based attack by quenching singlet oxygen within the phagosome, impairs the antimicrobial action of neutrophils and enhances innate host immune response, thus protecting against oxidative stress (Xiong & Kapral, 1992; Clauditz *et al.,* 2006; Mishra *et al.,* 2011).

Carotenoid pigments stabilize the *S. aureus* cell membrane's (CM) fluidity and increase rigidity by ordering the alkyl chains of intra and extracellular leaflets of the CM lipid bilayer during infection and pathogenesis (Mishra *et al.,* 2011). The disruption of metabolic genes such as the tricarboxylic acid (TCA) cycle genes (*citZ, citG, SAV2365),* purine biosynthetic genes (*purN, purH, purD,* or *purA*), decreased expression of oxidative phosphorylation (*qoxB, ctaA*), cold shock protein (*cspA*) and other genes (*SAV1117, SAV1108, hemL*) lead to increase in pigment production (Lan *et al.,* 2010;

Katzif *et al.,* 2005). Deletion of a gene encoding the STX biosynthesis enzyme, such as *crtM*, renders the bacterium more susceptible to killing by human and mouse polymorphonuclear leukocytes (PMNs) or whole blood cells (Mishra *et al.,* 2011).

1.10.1.4. Other virulent factors

S. aureus secretes two clotting factors, coagulase (Coa) and von Willebrand factor binding protein (vWbp) that bind to prothrombin forming a complex of staphylothrombin and convert fibrinogen into fibrin which promotes clotting of plasma or blood. Both Coa and vWbp are essential for *S. aureus* strain Newman abscess formation and persistence in host tissues (Cheng *et al.,* 2010). The coagulase expression is both positively and negatively controlled by an *agr* dependent mechanism. Agr elevates the *coa* mRNA level at the early exponential phase and causes a strong decrease at the post exponential phase of growth (Lebeau *et al.,* 1994). The immunoglobulin binding protein (Sbi) is an immune evasion factor that promotes bacterial survival in blood by binding to IgG and beta-2 glycoprotein and avoids neutrophil mediated opsonophagocytosis. The inactivation of *agr* increases *sbi* expression *in vivo*. (Burman *et al.,* 2008; Smith *et al.,* 2011; Jones *et al.,* 2008). A list of virulent factors associated with cell surface and secreted components of *S. aureus* are given in Table 3.

1.10.2. Secreted toxins and enzymes

The subsequent task of *S. aureus* after colonization is to invade tissues and spread the infection. One of the important characteristics of *S. aureus* is its ability to secrete toxins that lyse the membranes of host cells by forming *β*-barrel pores in the cytoplasmic membranes (Foster, 2005; Plata *et al.,* 2009). Pore forming toxins are exoproteins that cause cell leakage by the insertion of oligomeric β-hairpins into the cell membrane. *S. aureus* produces several toxins such as exotoxins (α, β, γ, δ- hemolysins), enterotoxins, toxic shock syndrome toxin (TSST-1), exfoliative toxins A/B, bicomponent cytolysin Panton valentine leukocidin (PVL) such as LukF/S and expresses several proteases, lipases, deoxyribonucleases (DNases) and fatty acid modifying enzyme (FAME) (Harris *et al.,* 2002; Pourmand & Namaki, 2009). Most of the bacterial toxins act as super antigens by binding to class II MHC molecule along with T-cell receptor. These activate large number of T-cells and cytokine production which causes capillary leak, epithelial damage and hypotension and weakens the host immune response (Baker & Acharya, 2004). The important toxins are discussed here.

Table 3. List of surface associated and secreted virulent factors of *Staphylococcus aureus* involved in the colonization and invasion of host tissues (Patti *et al.,* 1994; Clarke *et al.,* 2009; Lin and Peterson, 2010; Geoghegan *et al.,* 2010; Bien *et al.,* 2011; Shinji *et al.,* 2011).

Surface adhesins	**Putative function**
Clumping factor (ClfA/B)	Colonization factor, mediates adhesion of fibrinogen and fibronectin
Plasmin sensitive protein (Pls)	Have both adhesive and antiadhesive functions; mediate adhesion to cellular lipids and glycolipids and promote bacterial cell-cell interactions
Fibronectin binding protein (FnBPA/B)	Responsible for adhesion and internalization; important for *in vitro* and *in vivo* infections resulting in septic death
Collagen adhesin (Cna)	An adhesin, mediates the attachment of *S. aureus* cells to cartilage, important factor for developing septic arthritis
Staphylococcal protein A (SpA)	Plays a major role in the induction of pneumonia; high affinity of binding to IgG/M resulting in depletion of B-cells
Von-Willebrand protein (vWp)	mediates platelet adhesion at sites of endothelial damage
Intercellular adhesin poly N-succinyl β-1,6 glucosamine (IcaADBC)	Mediates colonization, required for cell-to-cell interaction
Iron regulated surface determinant A	Bind to human hemoproteins; remove the heme and use as a nutrient source; promotes adhesion to nasal cells
S. aureus surface protein G (SasG)	Role in adherence and biofilm formation
Secreted Virulent factors	**Putative function**
Toxic shock syndrome toxin 1 (TSS); staphylococcal enterotoxins (SeaA-F)	Activate T cells and macrophages
Cytolysins (α-,β-,γ-,δ- toxins); leukocidins (LukD/E) or Panton valentine leukocidin (PVL)	Induce apoptosis (at low concentration) and lysis of various cell types, including erythrocytes, lymphocytes, monocytes, epithelial cells
Lipase	Inactivate fatty acids
Hyaluronidase (HylA)	Degradation of hyaluronic acid
Serine proteases; cysteine proteases (including staphopains); aureolysin	Inactivate neutrophil proteolytic activity; inactivate antimicrobial peptides
Staphylokinase (Sak)	Plasminogen activation; inactivate antimicrobial peptides
Exfoliative toxins	Act as serine proteases; activate T cells
Chemotaxis inhibitory protein of staphylococci (CHIPS), Staphylococcal complement inhibitor (SCIN)	Inhibit complement C3b formation on the surface of the bacterium and enhance the ability of human neutrophils to phagocytose *S. aureus*

1.10.2.1. Staphylococcal delta hemolysin

S. aureus delta toxin, is a 26 aa surface active protein which acts on hydrophobic membrane structures and is distinguished from other staphylococcal hemolysin by its thermotolerance (heat stability) even at 100°C, specific activity to lyse human and horse RBCs, ability to be inhibited by phospholipids and normal sera (Wiseman, 1975; Kreger *et al.,* 1970). It also damages a variety of cell types including bacterial protoplasts, spheroplasts, tissue culture cells, leukocytes and also organelles such as lysosomes, mitochondria by damaging the cell membranes. Delta toxin is produced by >97% of *S. aureus* isolates and is mitogenic for human lymphocytes, produces dermonecrosis in guinea pigs and rabbits and lethal for mice (Kreger *et al.,* 1970). It exerts proinflammatory effects because of its binding specificity to neutrophils and monocytes which stimulate TNF-α production and acts as co-stimulator of human neutrophil oxidative burst (Schmitz *et al.*, 1997). The delta hemolysin gene *(hld)* is a part of the quorum sensing *agr* system, is encoded within RNA III (Novick *et al.*, 1993) and plays a major role in pathogenicity.

1.10.2.2. Staphylokinase

Staphylokinase (Sak) is a 136 amino acid long, 15.5 kDa bacteriophage encoded protein expressed by lysogenic strains of *S. aureus*, that favours the symbiosis of staphylococci in the host and acts as an important colonization factor (Bokarewa *et al.*, 2006). Staphylokinase is a thrombolytic agent complex with plasminogen and converts it into proteolytic plasmin that binds to the fibrin network around the infectious loci allowing staphylococci to enter into deeper tissues. Plasmin is serine protease, has various other host matrix proteins as substrates such as collagen, elastin. Staphylokinase is produced in the late exponential phase and is positively regulated by *agr* and negatively by *sar*. Sak enhances bacterial resistance to phagocytosis by interacting with short cationic peptides HNPs (Jin *et al.*, 2004) produced by human neutrophils.

1.11. Regulation of virulence factors

Regulation of expression of staphylococcal virulence factors plays a central role in pathogenesis. Several global regulatory loci have been identified in *S. aureus* such as *agr* (**a**ccessory **g**ene **r**egulator), *sarA* (**s**taphylococcal **a**ccessory **r**egulator), *sae* (***S. aureus* e**xoprotein expression), *rot* (**r**epressor **o**f **t**oxin), *ssrAB* or *srhSR* (Staphylococcal respiratory response), *ArlSR* (**A**utolysis **r**elated **l**ocus), *LytRS*, 1E3 transposon, sigma σ^B,

mgrA (master regulator of transcription) and sRNA (**s**mall **RNA**) that regulate the expression of surface proteins, exoproteins and other proteins essential for growth (Wisell, 2000; Bronner *et al.*, 2004; Tomasini *et al.*, 2014).

1.11.1. The Agr regulon

The complex polycistronic *agr* (accessory gene regulator) locus is driven by two divergent promoters P2 and P3, with the help of *sarA*, another global regulator produces a small 514 nt long regulatory RNA molecule called RNAIII (Novick *et al.*, 1993; Somerville & Proctor, 2009) (Fig. 11). The activity of *agr* can be assayed quantitatively by measuring the transcription of RNAIII levels or delta hemolysin protein levels (Sakoulas *et al.*, 2006) since delta hemolysin is encoded by RNAIII.

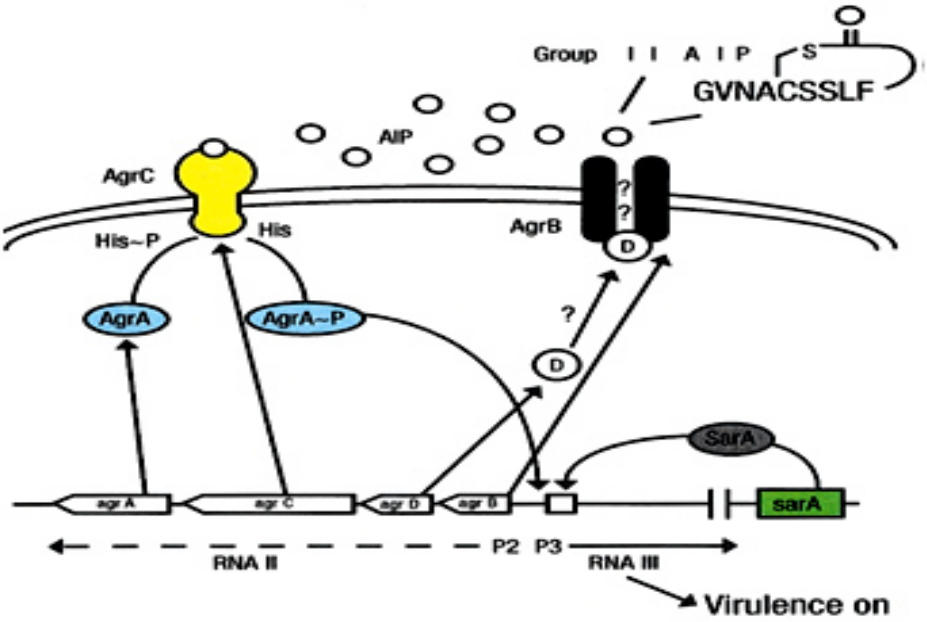

Fig. 11. Model showing the activation of *agr* and RNAIII (adapted from Sakoulas *et al.*, 2006). Expression of *agr* is driven by P2 and P3 promoters. The P2 promoter promotes transcription of AgrA/B/C/D to activate the circuit *via* the aid of *sarA*. AgrC is the transmembrane receptor histidine kinase and AgrA is the response regulator of the *agr* two component regulatory system. AgrD is a small peptide that is processed by AgrB into a autoinducing peptide (AIP) functions as the quorum-sensing pheromone. When the extracellular concentration of AIP reaches a threshold level, it complexes with AgrC, leading to the activation of the kinase domain. Upon the transfer of a phosphate from AgrC to AgrA, AgrA is activated, in turn activates P3 transcription with the help of *sarA*, produces a regulatory RNA RNAIII, the effector molecule of *agr* regulates transcription of various virulence genes (Somerville & Proctor, 2009).

Studies have revealed that *agr* upregulates the expression of exoproteins such as toxic shock syndrome toxin-1, enterotoxin B and C, and V8 protease (*sspA*) and downregulates the synthesis of cell wall associated proteins during postexponential and stationary growth phase (Harris *et al.,* 2002). Several animal models of staphylococcal infection including endocarditis, osteomyelitis, and septic arthritis, showed that *agr* mutant strains are attenuated in virulence compared to corresponding parental strains, indicating that *agr* is important for the pathogenesis of staphylococcal disease (Booth *et al.*, 1997). RNAIII, an effector molecule of *agr* regulates transcription of various virulence genes.

When bacteria are few (e.g. in the blood stream), *agr* system is not activated due to the low concentration of autoinducing peptide (AIP) that function as a quorum sensing pheromone, but when the cell density increases, *agr* quorum sensing is activated, in turn synthesizing the RNAIII molecule which leads to successful colonization and expression of extracellular toxins and enzymes, facilitating tissue degradation and dissemination (Wisell, 2000).

1.11.2. SarA

Staphylococcal accessory regulator (*sarA*) is essential for *agr* dependent regulation (Cheung *et al.*, 1992) and it also stimulates the expression of RNAIII. SarA is transcribed from three distinct promoters (P1, P2 and P3) generating three transcripts *sarA, sarB* and *sarC*. Inactivation of the *sarA* locus led to attenuation of virulence in several animal models of infection (Wisell, 2000). Animal model studies also suggested that *agr* and *sar* interact *in vivo* to control genes that affect the pathogenesis of *S. aureus* (Cheung *et al.*, 1994). A *sarA* mutant shows decreased expression of several exoproteins, such as α, β and δ-hemolysin and increased expression of proteases, Protein A, collagen adhesin (Harris *et al.,* 2002; Bronner *et al.,* 2004).

1.11.3. Rot

Transcription factor Rot (Repressor of toxin), a central global regulator of virulence gene expression negatively regulates the transcription of genes such as lipase, hemolysins and proteases which are postulated to play a role in tissue invasion. Rot positively regulates the expression of a number of genes including those encoding cell surface adhesins such as clumping factor B, coagulase, immunoglobulin G binding protein A precursor (Said-Salim *et al.,* 2003; Killikelly *et al.,* 2014). The regulatory

function of Rot is aided by the *agr* quorum sensing system but the alternative sigma factor (σ^B) has an opposite effect on *rot* expression during the post exponential phase of growth (Hsieh *et al.*, 2008). Rot played a significant role in biofilm formation, as studied in mouse catheter model (Mootz *et al.*, 2015).

1.11.4. Sae

sae (*S. aureus* exoprotein expression), a global regulator first described by Giraudo *et al.*, has been shown to have a role in the regulation of virulence determinants. The *sae* locus is a two-component regulatory system constituted by two co-transcribed genes, *saeR* (687 bp) and *saeS* (1062 bp) and is essential for the transcription of alpha- (*hla*), beta- hemolysin (*hlb*) staphylocoagulase (*coa*), DNase and protein A (Bronner *et al.*, 2004) as mutation of *sae* exhibited decrease in the production of these virulence factors (Harris *et al.*, 2002) and is essential for virulence gene expression *in vivo* (Wisell, 2000; Harraghy *et al.*, 2005).

1.11.5. Sigma σ^B factor

The stress responsive alternative sigma factor σ^B is expressed at stationary phase and σ^B mediated regulation depends on the recognition of consensus sequences found on the upstream of the genes coding virulence determinants. It indirectly regulates several genes that do not contain σ^B consensus promoter, through σ^B gene regulatory systems (Nielsen *et al.*, 2011) and contributes to pathogenesis. Sigma B is involved in the bacterial aggregation by modulating the expression of genes encoding clumping factor (*clfA*) and other adhesins. It also decreases the susceptibility of bacteria to hydrogen peroxide, by regulating catalase (*cat*), thereby protecting against superoxide anions released by polymorphonuclear neutrophils during the oxidative burst (Wisell, 2000; Homerova *et al.*, 2004). σ^B might be involved in antimicrobial resistance, example, in *S. aureus* Col mutant shows less resistant to methicillin than wild type (Bronner *et al.*, 2004). The overexpression of σ^B lead *S. aureus* resistant to lysostaphin and also raised minimum inhibitory concentration (MIC) of the cell wall affecting antibiotics (Morikawa *et al.*, 2001). The protection of bacteria against UV radiation might involve SigB by modulating the pigmentation by carotenoid which acts as antioxidant (Bronner *et al.*, 2004).

1.11.6. Small noncoding RNA

Small noncoding RNAs (sRNAs) are post transcriptional regulatory molecules that tune the expression of mRNAs and proteins at several stages of bacterial growth and essential for *S. aureus* infection. Small RNAs regulate the genes concerned with bacterial pathogenesis and basic physiology of the cell (Majdalani *et al.,* 2005; Fechter *et al.,* 2014). *S. aureus* has emerged as a model organism for the study of bacterial sRNAs. Several novel sRNAs have been found in *S. aureus* in recent years (Tomasini *et al.*, 2014). In *S. aureus*, to date approximately more than 700 regulatory RNAs have been discovered, but until now only a few have been characterized for their physiological functions (Guillet *et al.,* 2013; Sassi *et al.,* 2015).

1.12. Small noncoding RNAs in *Staphylococcus aureus*

1.12.1. Pathogenicity island (PaIs) encoded ncRNAs

Mobile genetic elements designated as pathogenicity islands (PaIs) mediate acquisition of antibiotic resistance genes and are a repository of many toxins, adherence and invasion factors, super antigens and secretion systems acquired through phages, transposons and plasmids. In addition to the protein coding genes, SaPIs encode and express several sRNAs, some of them are present in multiple copies (up to eight copies may be due to repeated events of horizontal transfer as well as gene duplications), scattered in *S. aureus* genome (Schmidt and Hensel, 2004). The sRNAs expressed from SaPIs are expected to regulate expression of genes involved in *S. aureus* pathogenicity, either directly or via intricate regulatory networks including transcriptional regulatory factors (Felden *et al.,* 2011). SprA/B/C/D/E/F/G/X named after **S**mall **p**athogenicity island **r**NA (Pichon and Felden, 2005; Bohn *et al.,* 2010) are examples of SaPI encoded ncRNAs in *S. aureus*.

1.12.2. Current perspective of *S. aureus* noncoding RNAs

Over the past 10 yrs, a large number of small noncoding RNAs have been identified by experimental approaches, RNA sequencing and computational prediction in *S. aureus* (Sassi *et al.,* 2015). The transcriptomic and proteomic analysis gave insights into the functions of sRNA in *S. aureus*. RNAIII, the first regulatory multifunctional sRNA in *S. aureus*, is an intracellular effector of the *agr* quorum sensing system, stimulates the expression of several virulence genes such as alpha and beta hemolysins, toxic shock

syndrome toxin, enterotoxins, lipases, proteases, nucleases and represses the expression of surface proteins such as protein A and coagulase (Benito *et al.,* 2000; Boisset *et al.,* 2007). Besides its regulatory role as sRNA, it is also an mRNA that encodes for delta hemolysin (*hld*) at the 5' end (Novick *et al.,* 1993).

Many sRNAs regulate multiple targets and also influence same target mRNAs. For example, RNAIII collaborates with a second RNA, SprD, to downregulate the common target *sbi* expression to overcome host immune defense mechanism by base pairing at multiple locations of mRNA (Chabelskaya *et al.,* 2014). A list of sRNAs that have been functionally characterised in *S. aureus* is given in Table 4.

Accumulation of sRNA in the cell, above the threshold level, is found to be toxic suggesting an important functional role in *S. aureus.* Example RsaE and SprD, when overexpressed affect cell viability (Geissmann *et al.,* 2009; Bohn *et al.,* 2010). The expression of sRNAs is significantly increased upon exposure to antibiotics. A study reported that differential expression of 409 potential sRNAs was observed in multiresistant *S. aureus* ST239 strain before and after exposure to four antibiotics (vancomycin, linezolid, ceftobiprole, and tigecycline), the major classes of antimicrobials used to treat methicillin resistant *S. aureus* (MRSA) infections. Notably, sRNA356 (also called Teg24, RsaI, RsaOG) a *trans*-encoded sRNA, was reported to have enhanced expression after exposure to vancomycin antibiotic at 6 h (Howden *et al*., 2013).

Several antisense sRNAs (asRNAs) expressed from four PIs and six SCCmec mobile genetic elements, ranging in sizes from 54 to 400 nucleotides, act as gene regulators. Among the well studied pairs are sRNAs, SprG1/SprF1 and SprA1/SprA1$_{AS}$ (Teg152). These sRNA pair with each other and are predicted to form type I "toxin-antitoxin" modules in which SprA1 and SprG1 encodes hydrophobic small peptides which are cytolytic for human cells and have an antimicrobial activity against gram positive and negative organisms (Pichon & Felden, 2005; Sayed *et al.,* 2011; Felden *et al.,* 2011; Pinal-Marie *et al.,* 2014).

Table 4. List of functionally characterized small RNAs in *S. aureus.*

Small RNAs characterized in *S. aureus*			
sRNAs	**Target mRNAs or proteins**	**Mode of action**	**References**
RNAIII (514 nt)	↑ *hla, hlb, ent* ↓*rot, sbi, protein A, Spa, coa, SA1000, SA2353*	Stimulates the expression of extracellular toxins and enzymes such as alpha, beta hemolysins, toxic shock syndrome toxin, enterotoxins and enzymes (lipases, proteases and nucleases) and represses the expression of surface proteins such as protein A and coagulase and increases the virulence in the host	(Novick *et al.*, 1993); Morfeldt *et al.*, 1995; Benito *et al.*, 2000; Boisset *et al.*, 2007)
RsaE (100 nt)	↓*25 genes* ↑*39 genes*	RsaE accumulation lead to the downregulation of 25 genes which includes numerous metabolic enzymes involved in the citrate (TCA) cycle and the folate dependent one carbon metabolism and upregulation of 39 genes including membrane proteins involved in peptide transport, operon for valine, leucine and isoleucine biosynthesis and it is also toxic to the cell	(Geissmann *et al.*, 2009; Bohn *et al.*, 2010)
SSR42 (891 nt)	~ *80 mRNA*↓↑	Modulates the expression of ~80 mRNAs including the expression of several virulence factors such as protein A, capsule, α-hemolysin and Panton Valentine leukocidin (PVL) in a strain dependent manner in MSSA and MRSA. SSR42 contributes pathogenesis by mediating erythrocyte lysis, resistance to human polymorphonuclear leukocyte killing and in a murine model of skin and soft tissue infection	(Morrison *et al.*, 2012)
RsaA (139 nt)	↓*mgrA*	Translational repression of the synthesis of the transcriptional regulator *mgrA* mRNA enhances the production of biofilm and decreases the synthesis of capsule formation. It also regulates several metabolic enzymes. RsaA RNA functions as a suppressor of	(Geissmann *et al.*, 2009; Romilly *et al.*, 2014)

		virulence by attenuating the severity of acute systemic infections and enhances chronic catheter infection in infection model	
ArtR (345 nt)	↓*sarT*	ArtR is regulated by *agrA* by binding to artR sRNA promoter and block its transcription. It positively regulates α toxin expression by binding to 5' UTR of transcriptional regulator *sarT* mRNA	(Xue *et al.*, 2014)
Pathogenicity island encoded small RNAs			
SprD (142 nt)	↓*sbi*	SprD negatively regulates the expression of the immunoglobulin binding molecule (immune evasion) SprD accumulation in *in vivo* is toxic for the cells and reduces bacterial growth and controls the virulence in animal model by impairing both the adaptive and innate host immune responses	(Pichon & Felden, 2005; Chabelskaya *et al.*, 2010)
SprA1/SprA1$_{AS}$ (208/60 nt)	↓SprA1$_{AS}$	SprA1 and SprA1$_{AS}$, both forms a complex *in vivo* through base pairing interactions and prevent internal translation of the SprA1 encoded toxic delta hemolysin peptide	(Pichon & Felden, 2005; Sayed *et al.*, 2011; Felden *et al.*, 2011)
SprF1/ SprG1 (141/312 nt)	↓SprG1	*cis*-antisense encoded SprF1, negatively regulates the expression of SprG1 RNA which encodes the two toxic antimicrobial and hemolytic peptides and thereby prevents the mortality of *S. aureus*	(Pinal-Marie *et al.*, 2014)
SprX (160 nt)	↓*spoVG*	SprX negatively regulates the *spoVG* in operon gene (*yabJ-spoVG*) expression and thereby influencing vancomycin and teicoplanin glycopeptide antibiotic resistance	(Eyraud *et al.*, 2014)
SprC (152 nt)	↓*atl*	SprC negatively regulates autolysin ATL and thereby reduces virulence and phagocytosis by human monocytes and macrophages in the host	(Pichon & Felden, 2005; Le Pabic *et al.*, 2015)

The pathogenicity of *S. aureus* is enhanced in animal infection model when tested with strains bearing altered levels of sRNAs present in staphylococcal pathogenicity island like SprD (Chabelskaya *et al.*, 2010), SprC (Le Pabic *et al.*, 2015) and at other loci like SSR42 (Morrison *et al.*, 2012), RsaA (Romilly *et al.*, 2014), ArtR (Xue *et al.*, 2014) administered through several routes of infection. There is growing evidence that sRNAs are involved in antibiotic resistance. sRNAs influence bacterial resistance to antibiotics by interfering with the activities of antibiotics such as blocking of transcriptional machinery and protein synthesis of bacteria (Eyraud *et al.*, 2014; Lalaouna *et al.*, 2014).

1.12.3. Staphylococcal Regulatory RNA database (SRD)

To manage the abundant RNAs in bacteria, several databases such as fRNAdb (Kin *et al.*, 2007), NONCODE (Xie *et al.*, 2014), Rfam 11.0 (Burge *et al.*, 2013), sRNAMap (Huang *et al.*, 2009), sRNATarbase (Cao *et al.*, 2010), sRNAdb (Pischimarov *et al.*, 2012), BSRD (Li *et al.*, 2012) were developed. The limitations of existing databases are lack of a unique sRNA nomenclature, they are not update and some of the databases focused exclusively on eukaryotes or prokaryotes.

SRD is a storehouse mainly for the Staphylococcal sRNAs (http://srd.genouest.org/) developed recently in 2015. SRD compiles the existing data at a single interface with non repetitive or redundant sequences of small RNAs having a specific nomenclature and document the number and location of staphylococcal regulatory RNAs (srns) in 18 *S. aureus* strains and 10 other Staphylococci (Sassi *et al.*, 2015).

1.13. Use of antisense reagents as possible anti infective therapy

Antisense therapy is a form of treatment where the genes which are known to be causative for a particular disease are inactivated (turned off) by a small synthetic oligonucleotide analogue (RNA or DNA) of 14-25 nt in length. This analogue is designed to bind either to the DNA or to the specific mRNA, thereby inhibiting the transcription of that gene (reviewed in Mraheil *et al.*, 2010). The first antisense reagent approved by the FDA (US Food and Drug Administration) is the antiviral Fomivirsen (marketed as Vitravene), a 21-mer oligonucleotide with phosphorothioate linkages. However there are several disadvantages with the first and second generation of antisense reagents like low binding affinity to complementary nucleic acids and non specific binding to proteins. They also cause toxic side effects that limit many applications.

sRNA based therapeutic treatment of infectious diseases may become a useful tool in the near future (Lalaouna *et al.*, 2014). The third generation of antisense agents has arrived that provides efficient and specific antisense activity *in vivo* without having toxic activity. Stabilization can be accomplished through phosphate backbone modifications via., modified peptide nucleic acids (PNAs), phosphorodiamidate morpholino oligomers (PMOs), locked nucleic acid (LNA) which refers to an oligonucleotide containing one or more bicyclic, tricyclic or polycyclic nucleoside analogues to overcome the *in vivo* degradation (exo- and endo- nuclease and protease), that are commercially available. The other modifications include phosphodiester, phosphorothioate or combinations of both, methylphosphonate, methylphosphorothioate, phosphorodithioate, p-ethoxy and/or combinations, morpholinos (phosphorodiamidate morpholino oligomers, PMOs), 2'-O-Met oligomers, tricycle (tc)- DNAs (Fig. 12).

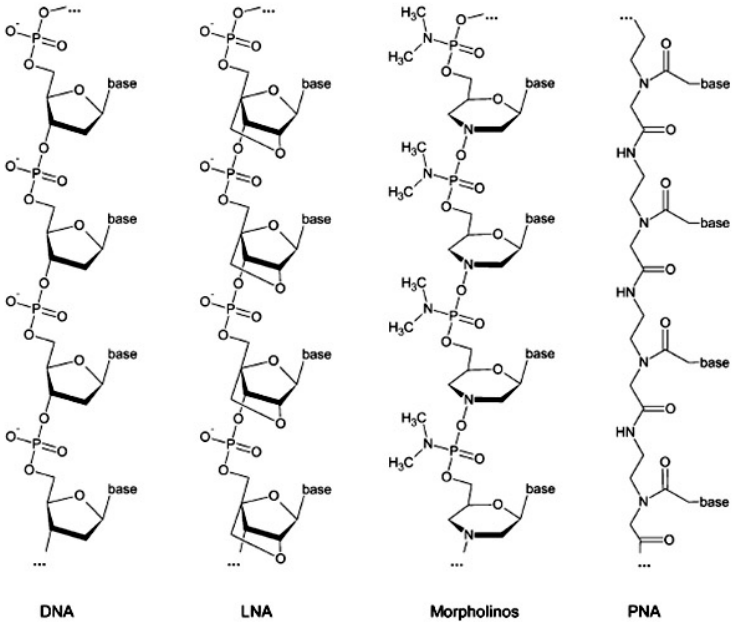

Fig. 12. Chemical structure of analogues of DNA, LNA (locked nucleic acids), PMO (phosphorodiamidate morphino oligomer) and PNA (peptide nucleic acids) (adapted from Mraheil *et al.*, 2010).

These stabilizers hybridize with high affinity and target specificity and forms stable duplexes than RNA: RNA duplexes and need no substrates for enzymatic degradation, thus can be used as novel tool to inactivate potential sRNA (reviewed in (Mraheil *et al.*,

2010). The oligonucleotide may be conjugated to a second molecule that may be able to target the organism. Second molecule can be a cell or bacterial penetrating peptide. The route of administration that does not produce an adverse, allergic or other untoward reaction when administered to mammals is preferably to be made.

sRNAs can also be used as diagnostic markers to identify the diseased state, a useful alternative method because of its regulatory functional role in the host. For example, *S. aureus* infection can be diagnosed by monitoring the expression of SprX (Chabelskaya *et al.*, 2015) and RNAIII and SprD which in turn might help to determine the severity of blood stream infections (Bordeau *et al.*, 2016). Recently, sRNA SprX from *S. aureus* HG001 strain has been filed for patent, for the treatment of a bacterial infection (International Publication Number W02015/075166A1) (Chabelskaya *et al.*, 2015). SprX targets *spoVG* mRNA and influences antibiotic resistance (Eyraud *et al.*, 2014). The present invention proposes to use the recombinant oligonucleotide (also referred as antisense oligonucleotide or SprX) comprising DNA sequence with the range of oligonucleotides from 75 to 115, targeting *spoVG* ribosomal binding site of *S. aureus yabJ-spoVG* mRNA with differential stabilization and administration strategies in the host model of infection (Chabelskaya *et al.*, 2015).