FUNCTIONAL CHARACTERIZATION OF STYR-3 NON-PROTEIN-CODING RNA (npcRNA) IN *Salmonella enterica* serovar Typhi

by

THIVIYAA M. OTHAYA KUMAR

Thesis submitted in fulfillment of the requirements for the degree of Master of Science

December 2014

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest and heartfelt gratitude to my main supervisor, Assoc. Prof. Dr. Tang Thean Hock for his endless support, guidance and patience throughout this research. His valuable advices and constructive comments have always been an encouragement to me in completing this work. Next, I also owe my sincere gratitude to my co-supervisor, Dr. Venkata Suresh Chinni for his continuous guidance as well as support. I appreciate his inspirational advices, beneficial inputs and endless effort in teaching me. I would also like to express my sincere thankfulness to Dr. Timofey Rozhdestvensky from University of Muenster for his guidance. I am utmost grateful to him for his willingness to share his vast knowledge and to spend his precious time in order to guide me. I thank Dr. Subash Gopinath for the critical reviewing of this thesis.

I am indebted to Universiti Sains Malaysia and Advanced Medical and Dental Institute (AMDI) for providing excellent facilities for conducting my research besides providing USM Graduate Research Assistant Scheme (GRA) and AMDI Postgraduate Research Scheme respectively. I am also thankful to Ministry of Higher Education (MOHE) for providing me financial assistance via MyMasters scholarship, under the MyBrain15 plan.

I would also like to take this opportunity to thank all the members of Infectomics Cluster (formerly known as Infectious Disease Cluster), AMDI, especially, my lab mates; Madam Siti Aminah Ahmed, Dr. Hoe Chee Hock, Mr. Citartan, Ms. Lee Li Pin, Ms. Nithya, Madam Priya and Ms. Toh Saw Yi for their

great assistance, support and co-operation. I wish all of them great luck in their future undertakings.

I am blessed to have wonderful friends and loved ones around me who have been very supportive and helpful in many ways throughout these years. I sincerely appreciate their endless support, motivations and for providing me shoulders to lean whenever I needed it the most.

I owe my sincere most appreciation to my loving parents, Mr. Othaya Kumar and Mrs. Vasenthey and both my brothers, Thilesh Kumar and Yuggesh Kumar who have been continuously showering me with abundance of love, blessings and motivations for me to strive and complete this work successfully. Words simply cannot describe my profound gratitude towards their endless support and encouragement because without them, I wouldn’t have reached this far.

It is absolutely true that God resides everywhere; and He is always with us to guide us through each and every step of life. I am extremely thankful to God for all His blessings and guidance throughout this journey.

Finally, I would like to extend my sincere thanks to all of those who have involved directly or indirectly in this work.

I dedicate this thesis to all my loved ones

for their unconditional love, endless encouragements and their wonderful blessings.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i

TABLE OF CONTENTS iv

LIST OF TABLES x

LIST OF FIGURES xi

LIST OF ABBREVIATIONS AND SYMBOLS xii

ABSTRAK xv

ABSTRACT .... Error! Bookmark not defined.

CHAPTER 1 1

INTRODUCTION 1

CHAPTER 2 3

2.1 Background of Non- Protein Coding RNAs 3

2.1.1 The discovery of npcRNAs 4

2.1.2 Approaches for the identification of npcRNAs 5

2.1.2.1 Biocomputational screens 6

2.1.2.2 Experimental Screening - Microarray detection 7

2.1.2.3 Experimental RNomics 7

2.1.2.4 RNA- Seq 8

2.2 Bacterial npcRNAs and their classifications 9

2.2.1 *trans*-encoded npcRNAs and its biological roles 11

2.2.1.1 Inhibition of translation 11

2.2.1.2 Activation of translation 12

2.2.1.3 Modification of protein activity 13

2.2.2 cis-encoded npcRNAs 13

2.2.2.1 Termination of transcription 15

2.2.2.2 Modulation of Translation 15

2.2.2.3 Alteration of Target RNA Stability 16

2.3 NpcRNAs involved in bacterial stress responses and virulence mechanisms.. 18

2.3.1 npcRNAs in Envelope/ Outermembrane Protein Stress 18

2.3.2 NpcRNAs in iron stress 20

2.3.3 NpcRNAs in Quorum sensing 21

2.3.4 NpcRNAs as regulators of bacterial virulence 22

2.3.5 NpcRNAs involved in modulation of bacterial drug resistance 25

2.4 Multidrug resistance in *S.* Typhi 26

2.4.1 Genus *Salmonella* 26

2.4.2 *Salmonella enterica* serovar Typhi 26

2.4.3 Treatment of *S.* Typhi infections 27

2.4.4 Molecular mechanisms of MDR in *S.* Typhi 28

2.4.4.1 Plasmid-mediated resistance 28

2.4.4.2 Target gene mutation 29

2.4.4.3 Increased Efflux Pump Expression 29

2.5 NpcRNAs in *Salmonella* 30

2.5.1 Discovery of npcRNAs in *Salmonella* 31

2.5.2 Biological significance of npcRNAs in *Salmonella* 32

2.6 Research Objectives 37

CHAPTER 3 39

MATERIALS AND METHODS 39

3.1 Materials 39

3.1.1 Chemicals and Reagents 39

3.1.2 Buffers/ Solutions 40

3.1.3 Bacterial strains 41

3.1.4 Plasmids 41

3.1.5 Culture Media 41

3.1.6 Electrophoresis 42

3.1.7 Membranes 42

3.1.8 Radioisotopes 42

3.1.9 Oligonucleotides 42

3.1.9.1 Oligonucleotides used for amplification and cloning of StyR3 gene sequence 43

3.1.9.2 Oligonucleotides used for StyR-3 northern blot analysis 43

3.1.9.3 Oligonucleotides used for generation of single stranded DNA (ssDNA) probe by radiolabelled asymmetric PCR 43

3.1.10 Autoradiography 44

3.2 Methods 44

3.2.1 Bacterial Culture Conditions 44

3.2.2 Preparations of *S.* Typhi cell pellets 45

3.2.3 Preparations of Nucleic Acids 45

3.2.3.1 Genomic DNA Isolation 45

3.2.3.2 Total RNA Extraction 46

3.2.3.3 Plasmid DNA Isolation 46

3.2.4 Cloning of StyR-3 fragment 47

3.2.4.1 Polymerase Chain Reaction (PCR) and Agarose Gel-Electrophoresis47

3.2.4.2 Restriction enzyme digestion 48

3.2.4.3 Purification of plasmid DNA/ PCR products.................................... 48

3.2.4.4 Ligation of StyR-3 PCR product into PCR Topo 2.1 vector 49

3.2.4.5 Transformation.. 49

3.2.4.6 PCR Screening 50

3.2.4.7 DNA Sequencing 50

3.2.4.8 Preparation of *S.* Typhi electrocompetent cells and transformation of cloned StyR-3 npcRNA 50

3.2.5 Expression Analysis of StyR-3 npcRNA and the mRNA levels of *ramA* and *ramR*.. 51

3.2.5.1 Northern blot analysis 52

3.2.5.2 End-labeling of Oligonucleotides for Northern Hybridization 53

3.2.5.3 Generation of Single Stranded DNA (ssDNA) Probe........................ 53

3.2.6 Antibiotic susceptibility testing 54

2.2.7 Databases 55

CHAPTER 4 56

RESULTS 56

4.1 Bioinformatics analysis of StyR-3 candidate 56

4.1.1 Sequence conservation analysis via BLAST 56

4.1.2 Secondary structure prediction of StyR-3.. 56

4.1.3 Search for promoters and terminators.. 60

4.1.4 Target RNA predictions... 61

4.2 Experimental analysis of StyR-3 candidate.. 63

4.2.1 Northern blot expression of StyR-3 candidate.. 63

4.2.2 Northern blot analysis of *ramA* and *ramR* mRNA 67

4.2.2.1 Northern blot analysis of *ramA* mRNA 67

4.2.2.2 Northern blot analysis of *ramR* mRNA 67

4.3 Antibiotic susceptibility testing.. 68

CHAPTER 5 72

DISCUSSION 72

5.1 Functional characterization of npcRNAs 72

5.1.1 Biocomputational analysis towards characterization of StyR-3 73

5.1.1.1 Sequence conservation pattern of StyR-3 73

5.1.1.2 Secondary structure prediction of StyR-3 74

5.1.1.3 Screening for promoters and terminators... 75

5.1.1.4 Prediction of Target RNA 76

5.2 Experimental Analysis towards Characterization of StyR-3 78

5.2.1 Cloning and over- expression of StyR-3 npcRNA 78

5.2.1.1 Cloning of StyR-3 78

5.2.1.3 Expression analysis of StyR-3 in *S.* Typhi ST001 79

5.2.3 Unveiling the Potential Role of StyR-3 in MDR Mechanism of *S.* Typhi 81

5.2.3.1 Determining the effect of StyR-3 overexpression on the *ramA* mRNA level 82

5.2.3.1 Determining the effect of StyR-3 overexpression on the *ramR* mRNA level 83

5.2.3.2 The effect of StyR-3 overexpression on the *ramA* and *ramR* mRNA level 84

5.3 Phenotypic analysis of StyR-3 overexpression 85

5.3.1 Disk diffussion assay 86

CHAPTER 6 88

CONCLUSIONS & FUTURE WORKS 88

CHAPTER 7 90

REFERENCES 90

APPENDICES 108

Appendix A: Preparation of Buffers & Reagents 108

Appendix B: Map of PCR 2.1 Topo Vector 113

Appendix C: Curricular Vitae 114

LIST OF TABLES

Table		Page
Table 4.1	Antibiotic susceptibilities of *S*. Typhi ST001 and *S*. Typhi ST001 + pStyR3	70

LIST OF FIGURES

Figures		Page
Figure 2.1	Schematic representation of *cis*- and *trans*-encoded npcRNAs	10
Figure 2.2	Schematic representation of gene regulation mediated by *trans*-encoded npcRNAs	14
Figure 2.3	Overview of regulatory mechanism employed by *cis*-encoded npcRNAs	17
Figure 2.4	Control of virulence traits in *S. aureus* by RNAIII	24
Figure 2.5	*Salmonella enterica* serovar Typhi	27
Figure 2.6	Schematic representation of the genomic localization and orientation of StyR-3 in the genome of *S*. Typhi and the regulation of MDR cascade via *ramA*-RamR interaction	35
Figure 2.7	Schematic representation of the orientation and genomic localization of StyR-3 in the genome of *S*. Typhi Ty2	36
Figure 4.1	Screenshot from NCBI-BLAST analysis of StyR-3 sequence	57
Figure 4.2	Multiple sequence alignment of StyR-3 sequence	58
Figure 4.3	RNA structural alignment of StyR-3 and sequence alignment from related sequences of Salmonella Spp. And secondary structure prediction using LocARNA software	59
Figure 4.4	Promoter prediction for StyR-3 via BPROM software	60
Figure 4.5	StyR-3 target RNA predictions via RNApredator web server	62
Figure 4.6	Northern blot analysis of StyR-3 npcRNA	64
Figure 4.7	ImageJ analysis representing expression of StyR-3 northern blot band intensity from three different growth stages	66
Figure 4.8	mRNA levels of *ramA* and *ramR*	69
Figure 4.9	Antibiotic susceptibility of *S*. Typhi ST001 and *S*. Typhi ST001 + pStyR3	71

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviation	Meaning
A	Adenine
APS	Ammonium persulfate
ATP	Adenosine 5’-triphosphate
Bis	N, N’-methylene bisacrylamide
BLAST	Basic Local Alignment Search Tool
Bp	Base pair(s)
C	Cytosine
°C	Degrees Celsius
$CaCl_2.6H_2O$	Calcium Chloride Hexahydrate
cDNA	Complementary DNA
C-terminal	Carboxy-terminal
dCTP	Deoxycytidine triphosphate
ddH_2O	Double-distilled water
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
et al.	and others
EtBr	3, 8-diamino-5-Ethyl-6-phenyl phenanthridinium Bromide
fRNA	Functional RNA
FQ	Fluoroquinolones
G	Gravitational acceleration
G	Gram
G	Guanine
GTP	Guanosine 5’-triphosphate
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hr	Hour(s)
IGR	Intergenic Region
IPTG	Isopropyl- -D-thiogalactopyranoside
KCl	Potassium chloride
KH2PO4	Monopotassium phosphate
K2HPO4	Dipotassium phosphate
kV	Kilovolts
LB	Luria Bertani medium
M	Mol/Liter, molar
MDR	Multidrug Resistance
Mg2+	Magnesium ion
MgCl2.6H2O	Magnesium Chloride Hexahydrate
Min	Minute(s)

miRNA	Micro RNA
Ml	Milliliter
mM	Millimolar
mRNA	Messenger RNA
Na+	Sodium ion
NaCl	Sodium chloride
NaOAc. 3H2O	Sodium acetate trihydrate
NaOH	Sodium hydroxide
ncRNA	Non coding RNA
Ng	Nanogram
nM	Nanomolar
Nt	Nucleotide(s)
npcRNA	Non-protein coding RNA
N-terminal	Amino-terminal
OD_{600}	Optical density at 600nm wavelength
OH	Hydroxyl
OM	Outer Membrane
OMP	Outer Membrane Protein
ORF(s)	Open Reading Frames
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
QS	Quorum sensing
RBS	Ribosomal Binding Site
R.E	Restriction enzyme
RNA	Ribonucleic acid
RNase	Ribonuclease
RNA-Seq	RNA Sequencing
RNP	Ribonucleoprotein
Rpm	Rotations per minute
rRNA	Ribosomal RNA
RT	Room temperature
RT-PCR	Reverse transcription-PCR
S	Second(s)
SD	Shine Dalgarno sequence
SDS	Sodium dodecyl sulfate
siRNA	Small interfering RNA
snmRNA	Small non-messenger RNA
sRNA	Small RNAs
ssDNA	Single-stranded DNA
S. aureus	*Staphylococcus aureus*
S. Typhi	*Salmonella enterica* serovar Typhi
T	Thymine
TAE	Tris–Acetic Acid–EDTA
TBE	Tris-Boric Acid-EDTA
TEMED	N,N,N’,N’-Tetramethylethylenediamine
TIR	Translation Initiation Region
Tris	Tris-(Hydroxymethyl)-Aminomethane
tRNA	Transfer RNA
U	Unit
UTP	Uridine 5’-triphosphate

UTR	Untranslated region
utRNA	Untranslated RNA
u.v.	Ultraviolet
V	Volt (s)
V. cholerae	*Vibrio cholerae*
v/v	Volume per volume
w/v	Weight per volume
X–gal	5’-Bromo-4’-Chloro-3’-Indolyl- -D-galactoside
μg	Microgram
μF	Microfarad
l	micro liter
μM	micro molar
32P	Phosphorus-32
	Ohm
	Alpha
	Beta
Γ	Gamma
Δ	Delta
Σ	Sigma

PENCIRIAN FUNGSI STYR-3
'RNA TIDAK-BERKOD-PROTEIN' (npcRNA) DI DALAM
***Salmonella enterica* serovar Typhi**

ABSTRAK

RNA tidak-berkod-protein (npcRNA) ialah RNA yang tidak mengodkan protein tetapi memainkan peranan dalam modulasi pelbagai proses pengawalan sel, termasuk mekanisme kekebalan terhadap pelbagai antibiotik 'MDR'. Dalam kajian ini, satu npcRNA baru, StyR-3 yang ditemui di genom *S*. Typhi telah dipilih untuk penentuan fungsinya melalui kaedah bioinformatik dan eksperimen. 'Intergenic' npcRNA yang sepanjang 144-nt ini berada di *ramRA* regulon *S*. Typhi Ty2. Kaedah bioinformatik memaparkan tahap konservasi yang tinggi dalam kalangan spesis *Salmonella*, selain mempunyai struktur sekunder yang stabil, seperti dijangka oleh LocaRNA software. Menurut BPROM software, StyR-3 diramal mempunyai promoter sendiri. RNApredator, sebuah program untuk meramal "target" RNA menunjukkan bahawa StyR-3 boleh menjalin interaksi yang stabil bersama *ramR* mRNA; justeru mencadangkan bahawa ianya berkemungkinan mempunyai potensi dalam memodulasikan MDR *Salmonella*. Berikutan, pengklonan dan pengekspresian berlebihan StyR-3 di dalam strain klinikal *S*. Typhi (ST001) dilaksanakan untuk menentukan kesannya terhadap tahap mRNA *ramA* dan *ramR*. Keputusan 'Northern blot' menunjuk bahawa pengekspresian berlebihan StyR-3 tidak memberi kesan genotipik yang ketara pada tahap ekspresi mRNA *ramA* mahupun *ramR*. 'Disk diffusion assay' yang telah dilakukan untuk menentukan kesan fenotip akibat pengekspresian berlebihan StyR-3 menunjukkan bahawa kedua-dua strain, rekombinan dan bukan rekombinan adalah sensitif terhadap 7 daripada 8 antibiotik yang diuji. Secara keseluruhannya, keputusan ini mencadangkan bahawa

kemungkinan besar StyR-3 tidak terlibat dalam proses modulasi MDR tetapi mempunyai fungsi biologi yang lain dalam bakteria *Salmonella*.

FUNCTIONAL CHARACTERIZATION OF STYR-3 NON-PROTEIN-CODING RNA (npcRNA) IN *Salmonella enterica* serovar Typhi.

ABSTRACT

Non-protein-coding RNAs (npcRNAs) are RNA molecules that do not encode for proteins but are involved in modulating a myriad of cell regulatory processes, including bacterial drug resistance mechanisms. In this study, a novel npcRNA, StyR-3; discovered from the genome of *S*. Typhi was selected for further characterization via bioinformatics and experimental-based approaches. This 144-nt intergenic npcRNA is located in the *ramRA* regulon of *S*. Typhi Ty2. The bioinformatics-based characterization showed that it is highly conserved among *Salmonella* species; besides having a stable secondary structure as was predicted by LocaRNA software. According to BPROM software, StyR-3 was predicted to possess its own promoter. Furthermore, RNApredator; a target RNA search program showed that StyR-3 could form a stable interaction with the *ramR* mRNA, which further suggest its potential role in the regulation of MDR cascade. This was then followed by cloning and overexpression of StyR-3 in an *S*. Typhi clinical strain (ST001) in order to determine its effect upon the mRNA levels of *ramA* and *ramR*. The Northern blot expression result suggest that the overexpression of StyR-3 does not incur significant genotypic effect over the *ramA* and *ramR* mRNA levels. Moreover, disk diffusion assay that was carried out to determine the phenotypic effect of StyR-3 overexpression showed that both; the recombinant and the non-recombinant strains were susceptible to 7 out of 8 drugs tested. Collectively, these results suggest that StyR-3 is most probably not involved in MDR regulation but it could possibly exert other biological roles.

CHAPTER 1

INTRODUCTION

Non-protein-coding RNAs (npcRNAs) are heterogenous RNA molecules that do not encode for proteins. In bacteria, they exist as relatively short transcripts (~50-400 nts) and are involved in the regulation of various crucial mechanisms such as regulation of transcription, translation, plasmid replication and response to various stress conditions (Masse et al., 2003, Wassarman et al., 1999, Hershberg et al., 2003, Storz, 2002, Gottesman and Storz, 2011, Waters and Storz, 2009, Zhou and Xie, 2011). Several identified npcRNAs have also been implicated with the regulation of bacterial drug resistance mechanism (Delihas and Forst, 2001; Vogel and Papenfort, 2006; Delcour, 2009, Nishino et al., 2011).

In year 2010, Chinni et al, reported the identification of 97 novel npcRNAs from the transcriptome of *S*. Typhi, the causative agent of typhoid fever. Interestingly, one particular candidate; StyR-3, was hypothesized to be involved in the regulation of the pathogen's drug resistance mechanism (Chinni et al., 2010). This intergenic npcRNA was reported to be located in the *ramRA* regulon; intergenic between *ramA* which encodes for RamA (the transcriptional activator of *acrAB* and *tolC*) and *ramR* which encodes for RamR (the transcriptional repressor of *ramA*) (Chinni et al., 2010, Abouzeed et al., 2008, Bailey et al., 2010, Ricci and Piddock, 2009). Previous studies have reported that the overexpression of *ramA* could lead to MDR in *Salmonella* via induction of *acrAB* and *tolC* (Bailey et al., 2008, Chollet et al., 2004, George et al., 1995, Schneiders et al., 2003, van der Straaten et al., 2004). Since StyR-3 is located in this particular regulon, it is interesting to unveil its function. Therefore, the main aim of this study is to characterize the function of this

npcRNA candidate.

Characterization of npcRNAs is usually done using bioinformatics approach and also experimental approach. Bioinformatics/ biocomputational based functional analysis employs various softwares or databases to determine the possible functions of these molecules. On the other hand experimental analysis includes gene 'knock-out'/ mutational studies and overexpression analysis.

In this study, both; experimental and biocomputational analysis were employed in order to determine the biological role of StyR-3. Several softwares/ databases which are available online such as LocaRNA, RNApredator and BPROM were utilized for the biocomputational analysis, where as experimental characterization was carried out via overexpression analysis. In addition, the phenotypic effect of the StyR-3 candidate upon the drug resistance pattern of the pathogen was investigated using disk diffusion assay.

CHAPTER 2

LITERATURE REVIEW

2.1 Background of Non- Protein Coding RNAs

In general, cellular RNAs can be divided into two major classes depending on their functions. The first class represents the messenger RNAs (mRNAs), which are RNA molecules that encodes genetic information required for protein synthesis. On the other hand, the latter class accounts for non-coding RNAs (ncRNAs) (Eddy, 1999). As the name suggests, ncRNAs do not encode for proteins as they do not contain open reading frames (ORFs). However, this does not mean that they do not contain information nor have function. These regulatory RNAs which are also widely known as small RNAs (sRNAs), non-protein-coding RNAs (npcRNAs), functional RNAs (fRNAs), untranslated RNAs (utRNAs) and small non-messenger RNAs (snmRNAs) play many indispensable roles in bacterial cells (Tang et al., 2002a, Huttenhofer et al., 2002, Chinni et al., 2010, Davis et al., 2005, Heidrich et al., 2006, Vogel et al., 2003, Tjaden et al., 2006, Eddy, 1999). Besides prokaryotes, this group of RNAs are also found in eukaryotes and archaea (Tang et al., 2002a, Tang et al., 2002b, Yuan et al., 2003, Mattick, 2001, Mattick and Makunin, 2005, Makarova and Kramerov, 2007). Throughout this thesis, the term npcRNA will be used.

Traditionally, RNAs were regarded as a passive information carrier, while proteins were regarded as the only control units of bacterial gene expression. However, the discovery of npcRNAs certainly redefined this perception (Vogel, 2009). NpcRNAs, which are heterogenous molecules with various sizes and structures, have been recently recognized as important gene expression regulators

in cell. These RNAs can act by itself or in association with accessory proteins in ribonucleoprotein (RNP) complexes. They were shown to be involved in a myriad of cell regulatory processes; i.e RNA processing and modification, mRNA stability and translation, transcriptional regulation, plasmid replication and protein stability and secretion (Masse et al., 2003, Wassarman et al., 1999, Hershberg et al., 2003, Storz, 2002, Gottesman and Storz, 2011). In addition, npcRNAs also play vital regulatory roles in bacterial virulence and quorum sensing besides being involved in many aspects of microbial physiology such as, regulation of outer membrane protein, iron homeostasis and nutrient metabolism (Zhou and Xie, 2011). This certainly exhibits the importance of these regulatory RNAs for bacterial survival in a host.

2.1.1 The discovery of npcRNAs

NpcRNAs were firstly observed in *E. coli* as early as 1970's, much before the discovery of their subtypes, microRNAs (miRNAs) and short interfering RNAs (siRNAs) in eukaryotes. These works showed that the fractionation of *in vivo* ^{32}P-labelled *E. coli* total RNA resulted in the observation of a low molecular weight species of RNA, which have not been described previously. However, neither the genes encoding them nor their functions were determined at that time (Griffin, 1971, Ikemura and Dahlberg, 1973). NpcRNAs can either be plasmid-encoded or chromosomally encoded (Brantl, 2012).

The first plasmid-encoded antisense RNA was identified in early 1980s; RNA I (~108 nts in length), was found to control the ColE1 plasmid-copy number in *E. coli* (Stougaard et al., 1981, Tomizawa and Itoh, 1981). This was then followed by the discovery of other antisense RNAs of mobile elements that control the life cycle or copy number of bacterial plasmids, phages, and transposons. One such example is

a ~70 nucleotide RNA which is transcribed from the pOUT promoter of the Tn10 transposon and represses transposition by preventing translation of the transposase mRNA (Simons and Kleckner, 1983).

MicF RNA, discovered in 1984, was the first chromosomally encoded npcRNA identified. Base-pairing of MicF to the leader sequence of *ompF* mRNA led to inhibition of translation of this mRNA, which encodes the major outer membrane porin, OmpF (Mizuno et al., 1984, Andersen et al., 1989).

During that period of time, only ten of such regulatory RNAs were identified; mostly on the basis of their abundance and via serendipity. Majority of those identified were housekeeping RNAs, namely RNase P RNA, tmRNA, and SRP RNA, which participates in tRNA maturation, ribosome rescue, and protein translocation, respectively (Wassarman et al., 1999). Following that, several bioinformatics based npcRNA studies led to the identification of many new npcRNA genes in *E. coli* (Argaman et al., 2001, Chen et al., 2002, Wassarman et al., 2001, Rivas et al., 2001). Hitherto, the technical advancements in various screening methods have allowed the identification and prediction of a burgeoning list of npcRNAs in wide range of bacteria (Altuvia, 2007, Pichon and Felden, 2008, Sharma and Vogel, 2009).

2.1.2 Approaches for the identification of npcRNAs

It is a challenge to develop a universal method for the systematic detection of all classes of npcRNAs due to their structural and functional diversity. The usage of biochemical and genetic methods for the detection of npcRNAs seems to be difficult because bacterial npcRNAs lack characteristic features like a poly(A)-tail and have mostly only a small size; making them a poor target for mutational screens.

Furthermore, since npcRNA genes do not get translated into proteins, they are resistant to frameshift and nonsense mutation (Altuvia, 2007, Hershberg et al., 2003). Besides that, it has been shown that genes encoding npcRNAs are often overlooked via computational based identification as most of the conventional protein-gene finding programs search for features like open reading frames (ORFs) and exon/intron boundaries, which are inapt to npcRNAs. However, the development of newer strategies; both bioinformatics and experimental-based approaches have shed the light in discovering many novel npcRNAs in various bacterial species. A few of these approaches are explained below.

2.1.2.1 Biocomputational screens

The systematic genome-wide searches of bacterial npcRNAs were initially based on biocomputational screens (Argaman et al., 2001, Rivas et al., 2001, Wassarman et al., 2001). The availability of almost 2,200 completed bacterial genome sequences in public databases have certainly facilitated the development of many in silico based approaches for computational npcRNA predictions at genomic level (Sridhar and Gunasekaran, 2013).

There are four major computational approaches for the prediction of npcRNA locations from bacterial genome sequences, which include (a) comparative genomics, (b) secondary structure and thermodynamic stability, (c) 'Orphan' transcriptional signals in intergenic regions and (d) ab initio methods which do not consider either sequence or structural similarity (Argaman et al., 2001, Rivas et al., 2001, Wassarman et al., 2001). These bioinformatics-based tools were usually utilized to locate the putative genomic npcRNA regions which will be further validated experimentally (Sridhar and Gunasekaran, 2013).

2.1.2.2 Experimental Screening - Microarray detection

Microarray has been recognized as a powerful method for simultaneous monitoring of gene expression on a genome-wide scale. Recently, it has also been successfully used for detection of new bacterial npcRNAs. However, majority of the standard microarray assays are limited to detect RNAs expressed from the ORFs or at best include tRNA and rRNA genes. The first study that reported the use of microarray in detection of npcRNAs from *E. coli* included probes specific for IGRs of this genome in addition to strand-specific probes for all mRNA, tRNA, and rRNA regions (Selinger et al., 2000).

Thus far, this method in combination with comparative genomics has led to the identification of various npcRNA genes in other microorganisms besides *E. coli;* such as *Staphylococcus aureus* and *Bacillus subtilis* (Pichon and Felden, 2005, Wassarman et al., 2001, Silvaggi et al., 2006).

2.1.2.3 Experimental RNomics

Experimental RNomics (also called shotgun cloning) is another method which led to the successful identification of many npcRNAs in all domains of life (Huttenhofer et al., 2001, Tang et al., 2002a, Tang et al., 2005, Yuan et al., 2003, Marker et al., 2002, Sonnleitner et al., 2008, Vogel et al., 2003, Chinni et al., 2010). Key steps in a typical RNomics protocol involves size-fractionation of total RNA on polyacrylamide gels, which is followed by directional cDNA cloning and sequencing of the resulting libraries. This method is aimed to comprehensively identify RNAs that are expressed from a particular genome under a given set of conditions. The RNAs detected are regardless of whether they are primary or processed transcripts.

Experimental RNomics have several advantages over other methods because it has the potential to identify npcRNAs that lacks interspecies conservation. This method also has the ability to detect *cis*-encoded antisense RNA from the 5' and 3'-UTR of annotated ORF (Kawano et al., 2005, Vogel et al., 2003, Altuvia, 2007).

2.1.2.4 RNA- Seq

RNA-Seq (RNA Sequencing), which pioneered in a study conducted in yeast is a new strategy which uses deep-sequencing technologies for transcriptomic profiling (Nagalakshmi et al., 2008). This approach consist of several critical steps which includes isolation of total RNA, mRNA enrichment, fragmentation (optional), and conversion to cDNA. Subsequently, those cDNA will be amplified and sequenced via high throughput sequencing (Wang et al., 2009). Depending on the sequencing platforms, millions of short reads (ranging between 25 to 300 bp) could be generated (Oshlack et al., 2010, Shendure and Ji, 2008).

RNA-Seq has several advantages over other methods. Firstly, the detection of RNA transcripts by RNA-Seq does not depend on the existing genomic sequence, unlike hybridization-based methods. Besides that, RNA-Seq also generates highly reproducible results, for both technical and biological replicates. Moreover, in comparison to DNA microarrays, RNA-Seq has relatively low, if any background signals and it is highly accurate in quantification of expression levels. Besides that, it also requires small amount of initial RNA sample as there are no cloning steps involved (Wang et al., 2009). Even though RNA-Seq offers the mapping of RNA at single nucleotide resolution and is especially handy in identifying RNAs of low abundance (Lee et al., 2010, Shendure and Ji, 2008), not many researchers could

afford to use this technology as the cost required to process, analyze, and manage the data is relatively high (Schadt et al., 2010).

2.2 Bacterial npcRNAs and their classifications

In bacteria, a plethora of npcRNAs exist as relatively short transcripts (~50-400 nts) (Waters and Storz, 2009, Zhou and Xie, 2011). These regulatory molecules have recently been recognized as a new class of bacterial gene expression regulators. Despite the successful identification of a myriad of npcRNAs in various species of bacteria, the functions of most of these npcRNAs remain to be elucidated. Nonetheless, some of the identified npcRNAs have been shown to regulate posttranscriptional gene expression either by acting as antisense RNAs, by interacting with target proteins or by binding to complementary sequences of target RNAs and in some cases mimics the structures of other nucleic acids (Argaman et al., 2001, Storz et al., 2005).

In general, npcRNAs can be classified into two distinct classes based on their genomic locations: (a) *trans*-encoded npcRNAs, and (b) *cis*-encoded npcRNAs (Figure 2.1, page 10). The basis of this classification is described below with some examples of known npcRNA.

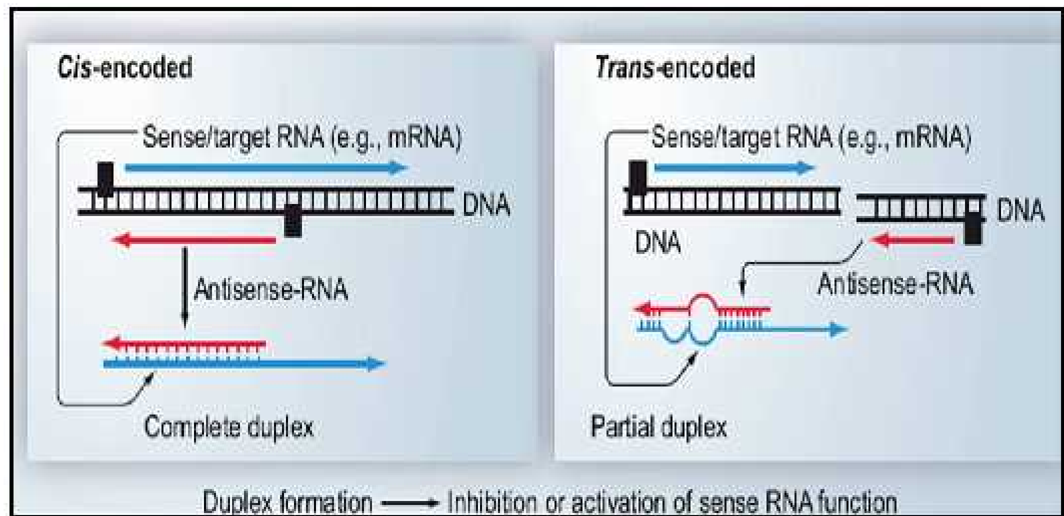

Figure 2.1: Schematic representation of *cis*- and *trans*-encoded npcRNAs. Antisense RNAs are drawn in red arrows and sense RNAs are indicated in blue. The left panel shows the complete complementarity base-pairing between *cis*-encoded npcRNAs with its target. Non-contiguous *trans*-encoded npcRNAs having partial complementarity with its target is shown in right panel (adapted from Brantl, 2009).

2.2.1 *trans*-encoded npcRNAs and its biological roles

Bacterial *trans*-npcRNAs are usually encoded at chromosomal sites distal to those of their target mRNAs. They generally act by base pairing with target mRNAs; often with less complementarity. The regions of complementarity are usually short and non-contiguous. The interaction results in either positive or negative regulation of translation or mRNA stability (Waters and Storz, 2009) and in many cases seems to be functionally analogous to eukaryotic miRNAs (Aiba, 2007, Gottesman, 2005).

Since they usually interact with their targets via imperfect base-pairing, these npcRNAs usually require the aid of RNA chaperone, Hfq. The binding of Hfq is believed to stimulate base pairing interactions by remodelling RNA structures and by increasing local concentrations of mRNA and the npcRNAs (Storz et al., 2004). Intriguingly, most of the chromosomally encoded antisense RNAs are *trans*-encoded. The diverse roles of *trans*-encoded npcRNAs are discussed below.

2.2.1.1 Inhibition of translation

The translation inhibition has been recognized as the main mode of action of *trans*-encoded npcRNAs. This is usually achieved via base-pairing of the npcRNA and target mRNA, which overlap the ribosomal binding site (RBS) or/and the 5' or 3' untranslated region (UTR) of the RBS that leads to inhibition of translation. One such example is the imperfect binding of MicF; the first *trans*-encoded npcRNA identified, with its target *ompF* mRNA, which led to the inhibition of translation of its target (Andersen and Delihas, 1990, Delihas and Forst, 2001). Similarly, the binding of MicA and MicC to the *ompA* and *ompC* RBS, respectively, was also shown to cause an inhibition of the translation of *ompA* and *ompC* (Figure 2.2A,

page 14) (Chen et al., 2004, Udekwu et al., 2005). Besides that, OxyS, which forms base pairs with *fhlA* mRNA was shown to lead to translational repression via formation of two short kissing complexes overlapping the Shine-Dalgarno (SD) sequence and within the coding sequence of *fhlA* mRNA (Argaman and Altuvia, 2000). Some of these npcRNAs such as GcvB and RyhB, performs translational inhibition through base pairing far upstream of the start codon (AUG) of the repressed gene (Sharma et al., 2007, Vecerek et al., 2007).

2.2.1.2 Activation of translation

T*rans*-encoded npcRNAs also function in translational activation. In certain situations, the 5’ UTR of some RNAs fold into an inhibitory secondary structure which sequesters the RBS and thereby prevents translation (Waters and Storz, 2009). The base-pairing of the *trans*-encoded npcRNAs to the complementary region will lead to the melting of the occluded region, which eventually leads to the translation initiation. For an example, the DsrA RNA, found in *E. coli* has been known to serve as a translational activator of the major stress and stationary phase sigma factor, RpoS, under low growth temperatures (Majdalani et al., 1998, Sledjeski et al., 1996). Its target; *rpoS* mRNA has an extraordinarily long 5’ UTR (~ 600 nt), which can fold into a translational-inactive structure by masking the RBS. The base-pairing of DsrA RNA to the complementary region leads to the opening of the stem-loop structure, giving access for ribosome binding to the RBS (Figure 2.2B, page 14). This mechanism which is also called anti-antisense mechanism also exerts translational activation via several other npcRNAs. These include (1) RyhB, which acts on *shiA*, encoding shikimate permease (Prevost et al., 2007), (2) GlmZ which acts on *glmS*, encoding glucosamine-6-phosphate (Urban and Vogel, 2008), (3) RNA III of

Staphylococcus aureus acts on *hla*, which encodes for hemolysin (Morfeldt et al., 1995) and (4) RprA, which acts on *rpoS*; same target as DsrA (McCullen et al., 2010).

2.2.1.3 Modification of protein activity

Some npcRNAs can interact with cellular proteins to modify their activities. These npcRNAs affect the activity of their target proteins by mimicking the structures of other RNAs or DNA. This will in turn affect the transcription, translation or processing of other RNAs. Among them, the CsrB/CsrC RNAs which antagonize the CsrA protein, a global regulator of carbon metabolism are important examples of npcRNAs that reconcile protein activity. Basically, in *E. coli*, CsrB and CsrC carry multiple GCA motifs which mimics the sequences bound by CsrA. The binding of CsrB and CsrC to multiple copies of CsrA decreases its activity by sequestering the protein from its targets (Figure 2.2C, page 14). This in turn affects the virulence and biofilm formation in *E. coli* (Jackson et al., 2002).

2.2.2 cis-encoded npcRNAs

cis-encoded npcRNAs binds with extensive complementarity to their target mRNAs as they are located on the DNA strand opposite to the mRNA targets. This class of npcRNAs constitutes riboswitches, which are part of the 5'UTRs of their target mRNA, that undergoes change in their structures upon binding of a regulatory molecule (Henkin, 2008), as well as *cis*-encoded antisense RNAs, which are encoded on plasmids or in the chromosome. Despite the increasing number of the *cis*-encoded npcRNAs being reported (Georg et al., 2009, Sharma et al., 2010, Toledo-Arana et al., 2009), information on the molecular mechanism of the individual npcRNAs is

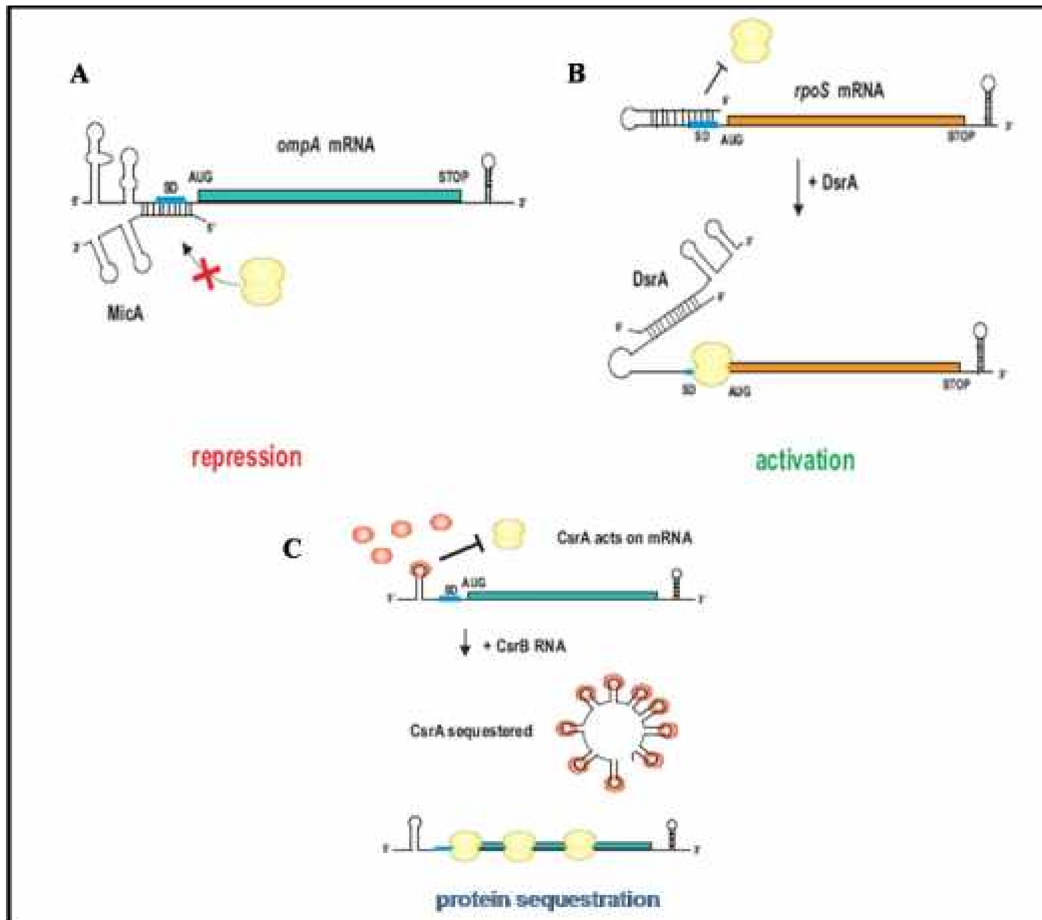

Figure 2.2: Schematic representation of gene regulation mediated by *trans*-encoded npcRNAs. (A) NpcRNA- mediated translational repression. The binding of MicA to the 5' UTR of *ompA* mRNA masks the SD sequence and thereby leads to inhibition of *ompA* translation. (Udekwu et al., 2005). **(B)** npcRNA- mediated translational activation. The *rpoS* 5'UTR usually folds into an inhibitory stem-loop closely upstream of the AUG start codon which occludes the SD sequence and thereby inhibits translation (Brown and Elliott, 1996). The interaction of DsrA RNA with *rpoS* mRNA leads to translational activation by melting the inhibitory structure of *rpoS* mRNA in which the ribosome binding site is masked (Majdalani et al., 1998). Secondary structures of sRNAs and mRNAs are presented schematically. The SD sequences are indicated in blue, and ORFs in light blue (*ompA*) and orange (*rpoS*) (adapted from Sharma and Vogel, 2009). **(C)** Sequestration of CsrA protein by CsrB npcRNA. The binding of CsrA protein (red circles) to hairpins in mRNAs leads to its translational inhibition. When CsrB RNA is expressed, multiple stem-loops in the npcRNA bind to CsrA proteins and, hence, the protein activity is sequestered; allowing translation of otherwise repressed mRNAs. (adapted from Wassarman, 2007).

less understood. In spite of that, these npcRNAs employs unique mechanisms which will be described in the following.

2.2.2.1 Termination of transcription

The best example for this mechanism is the regulation of iron transport biosynthesis operon in *Vibrio anguillarum*, which resides in the virulence plasmid pJM1. Basically, the operon consists of four ferric siderophore transport genes (*fatDCBA*) and two siderophore genes (*angR* and *angT*). Under iron deficiency, the expression level of *fatDCBA* mRNA becomes ~17 times higher than the full-length mRNA (*fatDCBA-angRT*); although it is part of the same polycistronic mRNA. The interaction between the RNA antisense npcRNA and the *fatDCBA* mRNA leads to transcription termination via formation of potential hairpin after the *fatA* gene, thus reducing expression of the downstream *angRT* genes (Stork et al., 2007). Similarly, in *Shigella flexneri*, an npcRNA known as RnaG was shown to promote premature transcriptional termination of *icsA* mRNA, which encodes for an invasion protein that is vital for its host colonization (Giangrossi et al., 2010).

2.2.2.2 Modulation of translation

Cis-encoded npcRNAs can regulate translation as well. An example is the regulation of SOS response-induced protein, SymE in *E. coli* by chromosomally – encoded *cis*-npcRNA, SymR. Similar to another toxin (MazF), SymE generally exerts its toxicity by repressing the global translation within the cell (Gerdes and Wagner, 2007). This 77-nt npcRNA is encoded *cis* to the RBS and AUG start codon of *symE*. Therefore, the interaction between *symE* and SymR disrupts the access of 30S ribosomal subunit to RBS; and thereby prevents the initiation of translation. In

short, SymR serves an antitoxin function by repressing SymE (an endonuclease toxin) synthesis (Figure 2.3B, page 17) (Kawano et al., 2007).

2.2.2.3 Alteration of target RNA stability

Majority of the *cis*-encoded npcRNAs with known functions, negatively regulate translation and promote the degradation of the complementary mRNA (Opdyke et al., 2011). However, in *E. coli*, GadY, a 105 nt length npcRNA, has been recognized as a positive regulator of the *gadX* mRNA levels (Figure 2.3A, page 17). It is located at an intergenic region between *gadX* and *gadW,* which encodes for AraC/XylS-like transcriptional regulators, GadX and GadW respectively. These regulators are involved in glutamate-dependent acid resistance. Interestingly, the *gadW* gene located immediately downstream from *gadX* can be transcribed with *gadX* or from its own promoter as an independent transcript (Ma et al., 2002, Tramonti et al., 2008, Tramonti et al., 2006). GadY is encoded in *cis* to the 3'-untranslated region (UTR) of the *gadX* gene. During acid stress, GadY will induce the cleavage of the bicistronic *gadXW* transcript, hence increasing the stability of the *gadX* mRNA, and in turn to accumulation of the GadX protein. Binding of GadY to the 3'-UTR of *gadX* could form a double stranded RNA duplex and impede digestion by RNase E (Opdyke et al., 2004) In line with this, a reduction in survival rate and decreased amounts of GadY and *gadX* transcripts were obtained in an RNase E knockout strain under acidic condition (Takada et al., 2007). In contrast, Opdyke et al. (2011) reported that only RNase III are involved in GadY dependent cleavage, suggesting that RNase involvement in the cleavage machinery could be growth condition-dependent (Opdyke et al., 2011).

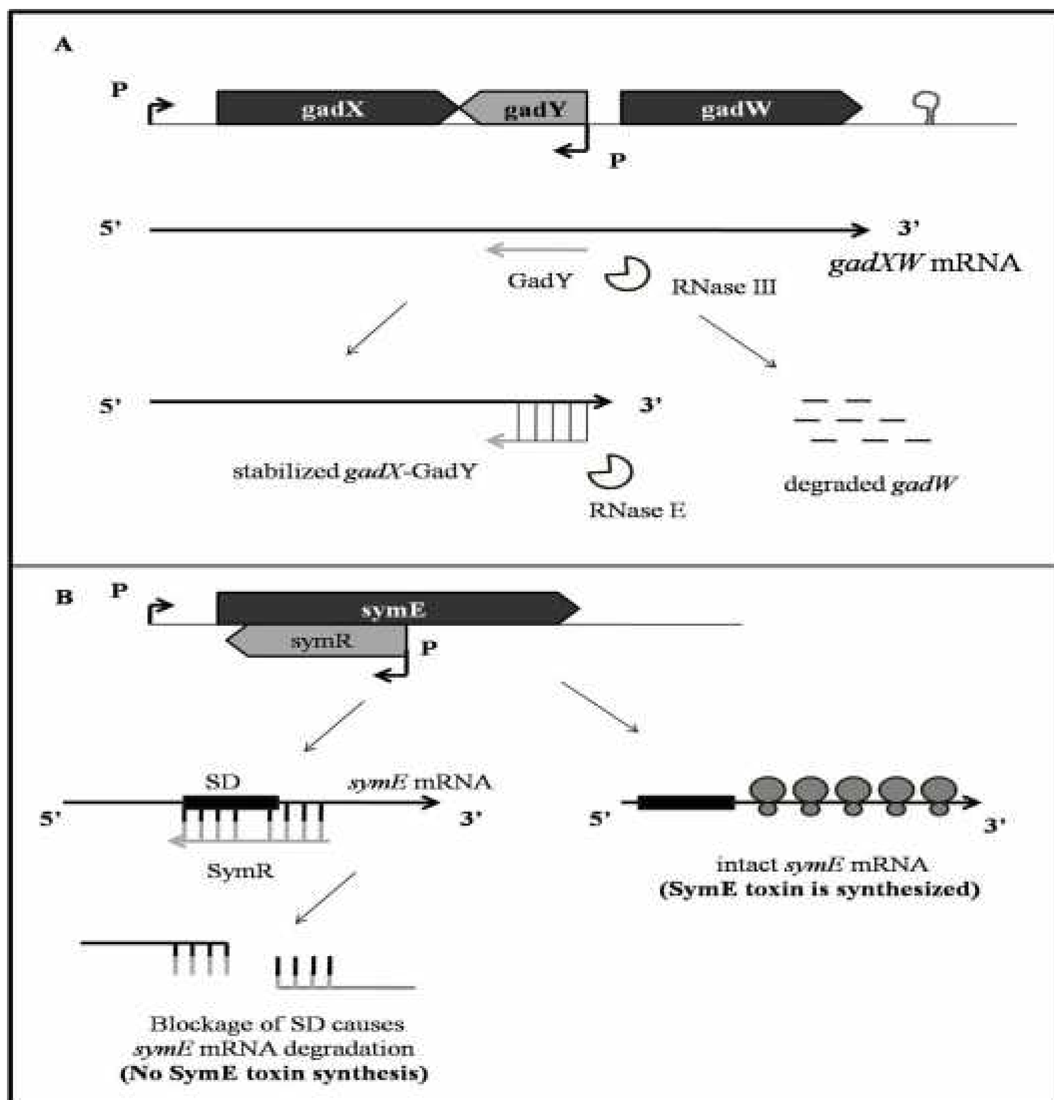

Figure 2.3: Overview of the regulatory mechanism employed by *cis*-encoded npcRNAs. (A) mRNA stabilization mechanism employed by GadY npcRNA. The grey arrow represents the GadY npcRNA. The black arrow shows the *gadXW* mRNA, which encodes for GadXW.In *E. coli*, during acid stress, GadY npcRNA binds to the intergenic region of *gadX-gadW* mRNA. This eventually results in two RNAs. Part of this processing is due to GadY- dependant RNAse III cleavage (adapted from Hoe et al., 2013). **(B)** Translation inhibition mechanism of SymR. The SymR npcRNA is represented by the grey arrow. The black arrow shows the *symE* mRNA,which encodes for SymE endonuclease toxin. The black rectangles represent the SD *symE* and *symR* genes respectively. The grey circular structures denote the ribosomes. In *E. coli*, SymR, a *cis*-encoded npcRNA binds to its complementary region at the 5' of *symE*. This eventually blocks the SD, which results in prevention of translation; thus SymE toxin will not be produced. In case of absence of SymR, intact *symE* mRNA will be transcribed and SymE toxin will be synthesized (adapted from Brantl, 2009).

2.3 NpcRNAs involved in bacterial stress responses and virulence mechanisms

Recently, npcRNAs have been implicated as vital modulators of bacterial gene expression in response to changing environmental conditions. They have been shown to orchestrate many essential stress and physiological responses such as envelope/ outermembrane protein (OMP) stress, oxidative stress, iron homeostasis, temperature fluctuations and pH stress (Hoe et al., 2013). Besides that, these heterogeneous molecules also participate in virulence mechanisms of several pathogens; *S. aureus*, *Salmonella*, *Vibrio* and *Shigella*; to name some of them. The involvement of several npcRNAs in a few selected stress responses and some npcRNAs involved in bacterial virulence are explained below.

2.3.1 npcRNAs in envelope/ outermembrane protein stress

In bacteria, cell envelope serves as a selective barrier that obstructs the entry of many toxic molecules into the cell and ensures bacterial survival in hostile environments. It also acts as a mechanical barrier which protects the cell from internal turgor pressure. In order to survive in constantly changing environment, bacteria have to regularly alter the nature and abundance of envelope components (Scott and Barnett, 2006, Bos et al., 2007). Basically, in Gram-negative bacteria, the cell envelope is composed of an outer membrane (OM) and an inner membrane, which are separated by a periplasmic space containing peptidoglycan; providing structural integrity to the cell.

Outer membrane proteins (OMPs) form channels across the OM and functions as a selective barrier for hydrophilic solutes. Besides functioning as barrier, they also serve as enzymes, adhesins and bacterial surface proteins which often confer bacterial virulence.

OmpA is an OMP which is highly conserved among enterobacteria. The *ompA* mRNA is abundant and has a relatively long half-life throughout bacterial growth. However, it has been reported that upon entry of bacteria into stationery phase, the RNA levels reduce drastically (Nilsson et al., 1984). This effect is mediated by an Hfq-dependant *trans*-encoded npcRNA, MicA (which was previously known as SraD) (Rasmussen et al., 2005, Udekwu et al., 2005). This regulation occurs when a 17 nt stretch in MicA forms a nearly perfect duplex with the 5' UTR of *ompA* mRNA, masking its SD sequence. The resulting MicA-*ompA* mRNA duplex is then subjected to RNase E-mediated decay. Similarly, another npcRNA, RybB was reported by Papenfort et al. (2010) to down-regulate multiple *omp* mRNAs in a E-dependent pathway (Papenfort et al., 2010).

Besides MicA, another npcRNA called VrrA was shown to base pair with 5' region of *ompA* mRNA. It was shown in *Vibrio cholerae* that this binding results in translational repression of *ompA*, without the aid of Hfq protein. Upon translational repression by VrrA, the reduction in OmpA level causes increased production of outer membrane vesicles (OMVs), which in turn relieves the envelope stress (Vytvytska et al., 2000). In addition to that, in *V. cholerae*, this npcRNA also represses the translation of *ompT* by binding to the 5' region of its mRNA, in an Hfq-dependent manner; showing the complexity of the OMP stress regulation (Song et al., 2010).

2.3.2 NpcRNAs in iron stress

In general, iron is an essential metal for most of the living organisms, as it acts as a cofactor for many enzymes which are involved in important biological processes such as respiration and DNA biosynthesis. However, in the presence of oxygen, iron becomes unfavourable as it catalyses the formation of hydroxyl radicals and ion superoxides that lead to oxidative damage (Imlay, 2003, Touati, 2000). As a consequence, the intracellular levels of iron have to be in tight control to avoid toxic effects and at the same time to provide the necessary amount for physiological needs.

In most bacteria, the Ferric uptake regulator (Fur) is the main regulator and sensor of iron homeostasis. It acts as a transcriptional repressor that controls iron metabolism and the use of iron as cofactor (Fe^{2+}) (Hantke, 2001, Moore and Helmann, 2005). During iron limitation, Fur will be inactive and the genes required for iron-uptake will be expressed. On the other hand, in iron-rich condition, Fur will be activated and it represses the iron-uptake genes. During this period, the genes which encode iron-containing and iron-storage proteins are expressed (Hantke, 2001).

An example of npcRNA which is involved in iron homeostasis is RyhB. It was first identified in *E. coli* as an npcRNA expressed under iron limitation conditions, and it was reported to be controlled by Fur protein. In *E. coli*, the regulation occurs via two consecutive repression processes. First, Fur represses the transcription of *ryhB*, which encodes RyhB npcRNA. This is followed by the binding of RyhB to mRNAs of iron-containing proteins, which eventually blocks their translation and promote their degradation (Masse et al., 2007). RyhB is believed to be able to repress more than 56 genes which include those involved in anaerobic and aerobic respiration, TCA cycle, and glycolysis. In short, RyhB enhances the level of

iron in a cell by limiting the expression of iron-using proteins and thus, making it available to iron-requiring proteins (Jacques et al., 2006).

Besides that, RyhB is also able to exert feedback, by repressing the translation of *fur* mRNA, which results in alteration of the Fur protein level. During iron scarcity, RyhB exerts post-transcriptional repression on *fur* and reduces the number of Fur proteins (Vecerek et al., 2007). In addition, RyhB also activates some genes involved in iron homeostasis. For an example, it acts on the mRNA of *shiA*, which encodes for a transporter of shikimate, which is needed for siderophore synthesis. The binding of RyhB to the 5' region of *shiA* abolishes the inhibitory secondary structure which sequesters the RBS. This leads to the formation of translation initiation complex, which eventually increases the translation and stability of the shikimate permease protein (Masse et al., 2007, Prevost et al., 2007).

2.3.3 NpcRNAs in Quorum sensing

Quorum sensing (QS) is the mode of communication for some bacteria. It involves a series of processes; from production, secretion and detection of signalling molecules known as autoinducers, which allow the bacterial population to co-ordinate their gene expression simultaneously in response to changes in cell density (Lenz et al., 2005). In *Vibrio* species, the interaction between the QS systems and npcRNAs forms a complex network, which regulates the virulence, biofilm formation and bioluminescence. The network then transduces signals to a transcriptional regulator, LuxO by a phosphorylation cascade, using a relay protein LuxU. The LuxO is indirectly connected to VarS-VarA system (a two-component hybrid sensor kinase) and the Csr system which has three npcRNAs; CsrB, CsrC and CsrD, where as the LuxU is directly connected by phosphorylation to LuxP-Q and

CqsS sensor proteins. The phosphorylated LuxO is needed by four quorum regulatory RNAs (Qrr1, Qrr2, Qrr3 and Qrr4) for the ^{54}- dependent transcription. These Qrr npcRNAs binds in an Hfq-dependent manner to the translation initiation region of *hapR* mRNA, which encodes HapR, a key regulator of many genes, including virulence genes that are repressed by it. The binding of these Qrr npcRNAs results in destabilization of the mRNA (Hammer and Bassler, 2007, Waters and Bassler, 2006, Tu and Bassler, 2007).

Besides that, the Qrr1-4 has the ability to activate the translation of vca0939 mRNA, whose gene encodes for GGDEF domain-containing protein. This activation is achieved via an Hfq-dependent base-pairing mechanism, which disrupts an inhibitory secondary structure that masks the SD (preventing access to ribosome). The disruption allows the binding of ribosome to the SD sequence and therefore activating the translation of vca0939 (Lenz et al., 2005, Lenz et al., 2004).

2.3.4 NpcRNAs as regulators of bacterial virulence

Pathogenic bacteria encounter constantly changing conditions during their infection in a host. Co-ordination of their virulence genes is highly crucial in order for them to adapt to these environments. For an instance, in *S. aureus*, RNAIII is one of the transcriptional units of the *agr* system that controls the virulence traits of the bacteria. RNAIII, which is 514-nt long encodes for δ-haemolysin (hld) besides being an npcRNA which targets at least five mRNAs encoding virulence factors. The base pairing between the 5’ end of RNAIII to the 5’ UTR of *hla* mRNA, which encodes for -haemolysin, promotes its translation (Morfeldt et al., 1995). On the other hand, some stem loops in the 3’ domain of RNAIII base pairs with the TIR of *sa1000*, *spa*, *sa2353* and *tor* which encodes for a fibrinogen-binding protein/ adhesin, cell-surface

protein A, secretory antigen precursor, and a pleiotropic transcription factor of many genes involved in virulence, respectively. This binding eventually results in their translational repression (Boisset et al., 2007, Geisinger et al., 2006, Huntzinger et al., 2005, Said-Salim et al., 2003).

RNA III expression is highly dependent on the cell density (Figure 2.4, page 24). When the cell density is low, the expression of RNAIII is also low; thus the expression of proteins that are required to hide the pathogens from host defences (Protein A, SA2353) and adhesins (SA1000) will be high to facilitate host colonization. Upon reaching a certain cell density, the RNAIII expression will be induced which will eventually switch on the haemolysins and degradative enzymes for the destruction of host tissues and to aid further colonization of the host (Repoila and Darfeuille, 2009).

Other npcRNAs related to virulence are IsrC which targets a virulence related gene, *msgA* in *S. Typhimurium* (Padalon-Brauch et al., 2008) and RnaG which represses plasmid-borne *icsA* mRNA, that encodes for an OMP needed for host cell invasion in *Shigella* (Giangrossi et al., 2010). Besides that, three npcRNAs (DsrA, RprA and ArcZ) have been shown to activate the synthesis of ^{S}, a major sigma factor required for the virulence of *E. coli* and *Salmonella* (Papenfort et al., 2009, Soper et al., 2010).

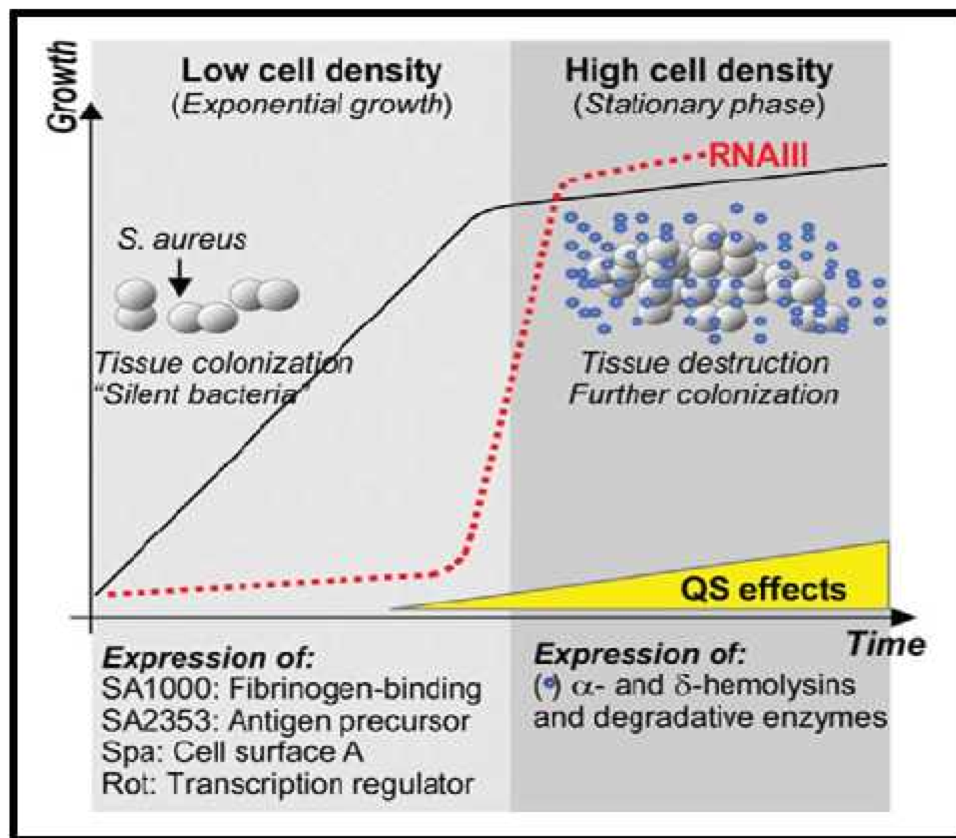

Fig 2.4: Control of virulence traits in *S. aureus* by RNAIII. When the cell density is low, the *agr* system is not induced.The levels of RNAIII are low and the mRNA targets which plays role in host tissue colonization will be expressed (sa1000, sa2352, spa and rot).During this period, the bacteria will be silent for host defences. At high cell density, the level of RNAIII increases as the *agr* system is autoinduced by QS. This leads to the translational inhibition of sa1000, sa2352, spa and rot. Concurrently, the RNAIII activates the expression of and - hemolysins and degradative enzymes. This causes host tissue destruction and enables its further colonization (adapted from Repoila and Darfeuille, 2009).