

**Organophosphate degradation (*opd*) island borne esterase responsive small
RNAs (sRNAs) in *E. coli* MG1655 - Characterization and functional
validation**

**Thesis submitted for the award of degree of
DOCTOR OF PHILOSOPHY**

By

Ashok Kumar M
(Enrolment No. 09LAPH12)



**Department of Animal Biology
School of Life Sciences
University of Hyderabad
Hyderabad – 500 046
INDIA**

August, 2016



University of Hyderabad
(A Central University established in 1974 by an act of parliament)
HYDERABAD – 500 046, INDIA

CERTIFICATE

This is to certify that **Mr. Ashok Kumar M** has carried out the research work embodied in the present thesis under the supervision and guidance of Prof. S. Dayananda for a full period prescribed under the Ph.D. ordinances of this University. We recommend his thesis entitled **“Organophosphate degradation (*opd*) island borne esterase responsive small RNAs (sRNAs) in *E. coli* MG1655 - Characterization and functional validation”** for submission for the degree of Doctor of Philosophy of the University.

Head,
Department of Animal Biology

Prof. S. Dayananda
(Research Supervisor)
Department of Animal Biology

Dean,
School of Life Sciences



University of Hyderabad
(A Central University established in 1974 by an act of parliament)
HYDERABAD – 500 046, INDIA

DECLARATION

This is to declare that the work embodied in this thesis entitled **“Organophosphate degradation (*opd*) island borne esterase responsive small RNAs (sRNAs) in *E. coli* MG1655 - Characterization and functional validation”** has been carried out by me under the supervision of Prof. S. Dayananda, Department of Animal Biology, School of Life Sciences. The work presented in this thesis is a bonafide research work and has not been submitted for any degree or diploma in any other University or Institute. A report on plagiarism statistics from the University Librarian is enclosed.

Prof. S. Dayananda
(Research Supervisor)

Department of Animal Biology

Ashok Kumar M



University of Hyderabad
(A Central University established in 1974 by an act of parliament)
HYDERABAD – 500 046, INDIA

CERTIFICATE

This is to certify that **Mr. Ashok Kumar M**, Enrolment No. 09LAPH12, has successfully completed course work as per UGC Regulations 2009 on minimum standards and procedures to be followed for the award of M. Phil/Ph.D prescribed under the Ph.D ordinance of this university.

Head,
Department of Animal Biology

ACKNOWLEDGEMENTS

I express my deep sense of gratitude to **Prof. S. Dayananda** under whose **inspiring, untiring** and **expert guidance**, the present work was carried out. I am indebted to him for the keen interest in my work, constant encouragement and critical suggestions.

I am grateful to Head, Dept. of Animal Biology **Prof. Jagan Pongubala** and former Heads, **Prof. B. Senthil Kumaran** and **Prof. Manjula Sritharan** for allowing me to use the Department facilities.

I am thankful to Dean, School of Life Sciences, **Prof. P. Reddanna** and former Deans **Prof. RP. Sharma**, **Prof. Aparna Dutta Gupta** and **Prof. M. Ramanadham** for allowing me to use school facilities.

I am grateful to my Doctoral Committee members, **Prof. P. Reddanna** and **Dr. K. Gopinath** for their critical assessment, discussions and advise during my research work.

I thank faculty members of School of Life Sciences for their timely and useful suggestions.

I also thank all the non-teaching staff of School of Life Sciences for their timely help.

I specially thank **Dr. Manjula Reddy**, Principal Scientist, CCMB for providing Keio collections strains.

I thank Prof. Utpal Nath, IISc Bangalore for allowing me to conduct northern blot experiments in his laboratory.

I also extend my acknowledgements to late **Prof. P. S. Sastry** for his intellectual interactions and constant motivation.

I thank **Dr. Sailu Yellaboina** and **Dr. Manimaran**, C R Rao AIMSCS for their help in microarray data analysis.

My thanks goes to Proteomics and Genomics facility of School of Life Sciences.

I thank all my lab mates: **Dr. Aleem Basha**, **Dr. E.V Paul**, **Dr. Toshi**, **Dr. Devi Prasanna**, **Dr. Swetha**, **Dr. Sunil**, **Venkateshwar Reddy**, **Uday**, **Ramurthy**, **Harshitha**, **Aparna**, **Hari**, **Annapoorni**, **Ganeshweri**, **Roly Kumari** for their timely help and a pleasant working atmosphere.

Special thanks to **Krishna** and **Rajendra Prasad** for technical help in the lab.

I would like to thank all Department staff **Mr. Ankinedu**, **Ms. Jyoti**, **Ms. Vijaya Lakshmi**, **Mr. Jagan**, **Mr. Babu Rao**, **Mr. Gopi** and **Mr. Srinivas** for their help in office work.

I would like to thank my friends, **Dr. Vinod Ch**, **Mr. Mastan**, **Mr. Papa Rao**, **Mr. Naveen Kumar Chouhan**, **Mr. Kumar Reddy**, **Mr. Mohan raj**, **Mr. Sravan Kumar**, **Mr. Rameshwar family**, **Dr. M Mulaka** and **SFI Hari** for making me happy every time.

I would like to thank my senior friends Dr. Madhu Babu Golla, Kishore Nalam, Dr. Srinivas, Dr. Madhu Prakash, Sujith and Dr. Rajesh

I thank Mr. Ravi Jellepalli and Paparao Vaikuntapu for taking the pictures of microbial plates

I would also like to thank all the Research Scholars in School of Life Sciences for their cooperation and help.

I thank **DST**, **CSIR**, **UPE-I and II**, **DST**, **ICMR**, **UGC**, and **DBT** for funding the lab.

I thank **DBT**, **DBT-CREBB**, **CSIR**, **DST**, **INSA**, **DST-FIST**, **UGC-SAP** for funding to Department and School.

I specially thank **CSIR** for **JRF** and **SRF**

I am immensely grateful to all those individuals who have helped me in making this endeavour possible.

No words of gratitude will repay my debt to my grandparents Mr. Lakshmaiah, Mrs. Sheshamma & Mr. Govinda Rao, Mrs. Jalamma, for their love and support, without which I would not have come to this stage.

I am grateful to my **amma Mrs. Madhu** and my **wife Dr. SubbaLakshami** for their patience, support and Love.

I thank Almighty for giving me strength, support, patience and courage.

Thank You

(Ashok Kumar M)

**DEDICATED
TO
MY MOTHER & MY WIFE**

CONTENTS

	Page No.
Abbreviations	: I
Table of Contents	: II
List of tables	: V
List of figures	: VI
Introduction	: 1-18
Materials and methods	: 19-43
Chapter I: Results and discussion	: 44-57
Chapter II: Results and discussion	: 58-80
Chapter III: Results and discussion	: 81-106
Summary	: 107-108
References	: 109-118
Plagiarism report	

Abbreviations

sRNA :	Small RNA
ncRNA :	Non-coding RNA
SD:	Shine-Dalgarno sequence
RBS:	Ribosomal binding site
AKGDH :	α -keto glutarate dehydrogenase
PDH:	Pyruvate dehydrogenase
ATCC :	American type culture collection
HPP :	Hydroxyl phenyl propionate
ICE :	Integrative conjugative element
IME :	Integrative mobilizable element
IPTG :	Isopropyl β -thiogalactoside
IS :	Insertion sequence
Kb :	Kilobase
KDa :	Kilo Dalton
Lpd :	Lipoamide dehydrogenase
MW :	Molecular Weight
OP :	Organophosphate
OPH :	Organophosphate hydrolase
PDH :	Pyruvate dehydrogenase
PNP :	<i>p</i> -Nitrophenol
PP :	Phenyl propionate
PTE :	Phosphotriesterase
qRT-PCR :	Quantitative real time PCR
SDS :	Sodium dodecyl sulphate
Tn :	Transposon
SPR:	Surface plasmon resonance
K _a :	Association constant
Hrs:	Hours
Min:	Minutes
nmol:	Nanomole

Table of Contents

1. Introduction.....	1-18
1. 1. Two-component regulatory system.....	1
1. 2. Riboswitch mediated gene regulation.....	2
1. 3. Small RNAs as transcriptional regulators.....	2
1. 4. The sRNAs mediated gene regulation	3
1. 4. 1. The <i>cis</i> -regulatory sRNA	4
1. 4. 2. Bacterial <i>trans</i> -acting sRNAs	6
1. 5. sRNA mode of action I - translational inhibition.....	10
1. 6. sRNA mode of action II - Translational activation.....	11
1. 7. The sRNA mode of action III: Sequestration of transcription factors	13
1. 8. Role of Hfq in establishing sRNA and mRNA interactions	13
1. 9. The sRNAs act on multiple mRNA	14
1. 10. Multiple sRNAs act on single mRNA	14
1. 11. sRNA regulation in carbon metabolism.....	15
1. 12. The Orf306 dependent metabolic diversion in <i>E. coli</i> MG1655.....	16
1. 12. 1. The <i>opd</i> island borne <i>orf306</i>	16
1. 13. Objectives.....	18
2. Materials and Methods... ..	19-43
2. 1. Bacterial strains and Plasmids used in this study.....	24
2. 2. Media and growth conditions.....	28
2. 2. 1. Luria-Bertani (LB) medium	28
2. 2. 2. Minimal salts medium for <i>E. coli</i>	28
2. 3. Preparation of antibiotic or chemical stocks.....	28
2. 3. 1. Ampicillin.....	28
2. 3. 2. Chloramphenicol.....	29
2. 3. 3. Tetracycline.....	29
2. 3. 4. Kanamycin.....	29
2. 3. 5. Gentamycin.....	29
2. 3. 6. IPTG.....	29
2. 3. 7. X-Gal.....	30
2. 3. 8. 50% Glucose.....	30
2. 3. 9. Propionate (1 M).....	30
2. 4. DNA Manipulations: Preparation of Solutions and Buffers.....	30
2. 4. 1. Agarose gel electrophoresis.....	30
2. 4. 2. 6X Gel loading buffer.....	31
2. 4. 3. Ethidium bromide stock solution	31
2. 4. 4. TAE Buffer.....	31
2. 4. 5. PCR amplification.....	31
2. 4. 6. Overlap extension PCR.....	31
2. 4. 7. DNA ligation.....	32
2. 4. 8. Preparation of ultracompetent <i>E. coli</i> cells.....	32
2. 4. 9. Transformation	33
2. 4. 10. Isolation of Plasmids	33
2. 5. Solutions for SDS-Polyacrylamide Gel Electrophoresis.....	33
2. 5. 1. Acrylamide Solution.....	33
2. 5. 2. Resolving gel buffer	34
2. 5. 3. Stacking gel buffer for SDS-PAGE	34

2. 5. 4. Tris-glycine electrophoresis tank buffer (pH 8. 3)	34
2. 5. 5. 2X SDS-PAGE loading buffer.....	34
2. 5. 6. Staining solution.....	34
2. 5. 7. Destaining solution.....	35
2. 5. 8. PhoB ^{6HIS} expression and Ni-NTA Purification.....	35
2. 5. 9. SDS-PAGE analysis.....	35
2. 6. RNA analysis.....	36
2. 6. 1. Tris-Borate-EDTA (TBE) buffer.....	36
2. 6. 2. 20XSSC.....	36
2. 6. 3. Isolation of small RNAs.....	36
2. 6. 4. Phenol-Chloroform extraction.....	37
2. 6. 5. DNase I digestion.....	37
2. 6. 6. Synthesis of cDNA.....	37
2. 6. 7. Electrophoretic separation of RNA	38
2. 6. 8. Northern blot analysis.....	38
2. 6. 9. Preparation of labelled oligonucleotide probe.....	39
2. 6. 10. Quantitative Real Time PCR.....	39
2. 6. 11. Absolute quantification of target genes.....	40
2. 7. Reagents for β -galactosidase activity.....	40
2. 7. 1. ONPG (<i>ortho</i> -Nitrophenyl galactoside) solution.....	40
2. 7. 2. 1 M Sodium carbonate.....	40
2. 7. 3. β -Galactosidase Assay.....	40
2. 8. <i>E. coli</i> knockouts by P1 lysate method	41
2. 9. Lipid Extraction.....	41
2. 10. Surface Plasmon Resonance (SPR).....	42

3. Chapter I.....44-57

3. 1. Background Work.....	44
3. 2. The <i>opd</i> gene is highly conserved.....	45
3. 3. Transcriptome Analysis.....	47
3. 4. Analysis of Orf306 responsive sRNAs.....	47
3. 5. Nomenclature of Orf306 responsive sRNAs.....	49
3. 6. The sRNA-target mRNA interactions: <i>in silico</i> predictions.....	50
3. 7. Identification of mRNA targets for <i>seco10054A</i> and <i>seco11136A</i>	51
3. 8. Discussion.....	55

4. Chapter II.....58-80

4. 1. Introduction.....	58
4. 2. Genetic Screen.....	58
4. 3. Construction of target mRNA gene transcriptional fusions.....	59
4. 4. Expression of sRNA, <i>seco10054A</i>	62
4. 5. Induction of sRNA, <i>seco10054A</i> expression.....	63
4. 6. Generation of <i>lac</i> negative knockouts of <i>E. coli</i> MG1655.....	64
4. 7. Co-transformation.....	65
4. 8. Screening of sRNA-mRNA interactions.....	65
4. 8. 1. Qualitative Assay	65
4. 8. 2. Structural complementarity between the sRNA, <i>seco10054A</i> and <i>lpd</i> mRNA.....	68
4. 9. Quantification of transcripts.....	68

4. 10. Generation of sRNA, <i>seco10054A</i> ^{Δ61-78} variant.....	69
4. 11. Expression of sRNA, <i>seco11136A</i>	71
4. 12. The sRNA, <i>seco11136A</i> and targets mRNA interactions.....	72
4. 13. Quantification of transcripts.....	76
4. 14. Discussion.....	77
5. Chapter III.....	81-106
5. 1. Introduction.....	81
5. 2. Construction of <i>seco10054A-lacZ</i> fusion.....	81
5. 3. Construction of <i>seco11136A-lacZ</i> fusion.....	82
5. 4. Two-plasmid assay.....	84
5. 4. 1. The Orf306 repress <i>seco11136A</i> gene expression.....	86
5. 5. Regulation of <i>seco10054A</i> gene.....	87
5. 5. 1. The <i>seco10054A</i> gene is part of PhoB regulon.....	89
5. 5. 2. PhoB binds to the predicted <i>phoB</i> motif of <i>seco10054A</i> gene.....	90
5. 5. 3. Expression and purification of PhoB.....	92
5. 5. 4. PhoB ^{6HIS} interacts with predicted <i>phoB</i> motif.....	93
5. 6. The link between Orf306 and expression of <i>seco10054A</i> and <i>seco11136A</i> genes.....	95
5. 6. 1. Orf306 generates endogenous propionate.....	95
5. 6. 2. Propionate induces <i>seco10054A</i> gene.....	96
5. 6. 3. Propionate represses <i>seco11136A</i> gene	97
5. 7. Propionate activates <i>hcaR</i> transcription.....	98
5. 8. Discussion.....	103
6. Summary.....	107-108
7. References.....	109-118

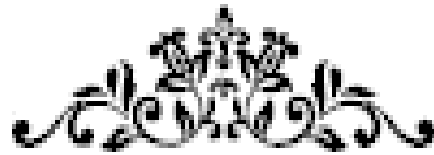
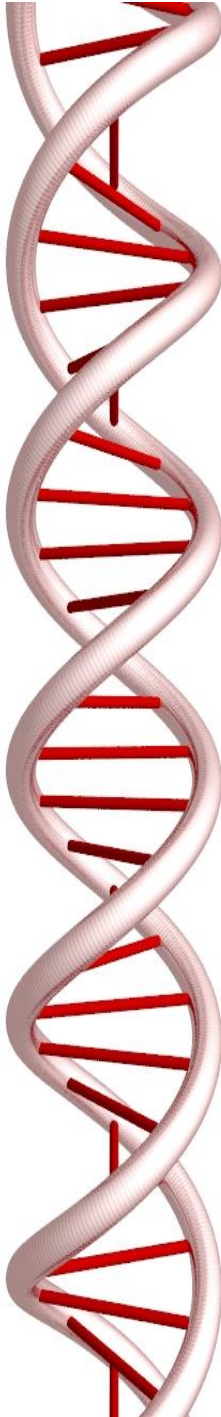
List of Tables

Table 1. A: List of functionally validated sRNAs in <i>E. coli</i>	9
Table 2. A: Chemicals used in the present study	19
Table 2. B: Antibiotics	21
Table 2. C: Restriction enzymes and DNA modifying enzymes	21
Table 2. D: Primers used in the present study	22
Table 2. E: Bacterial strains and plasmids used in the present study	24
Table 2. F: Bioinformatic Tools/Applications used in the present study.....	42
Table 3. A: The predicted seco10054A and targets mRNA interactions	53
Table 3. B: The predicted seco11136A and targets mRNA interactions	54
Table 4. A: The details of the primer sequences and amplicon size to be generated by PCR amplification	60
Table 5. A: PCR primers used to generate <i>mhpR</i> , <i>mhpA</i> and <i>hcaR</i> genes promoter fusions.....	100

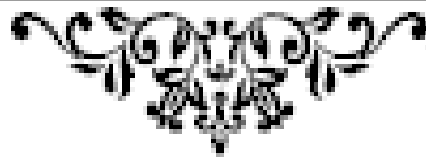
List of Figures

Fig. 1. 01. <i>Cis</i> -acting sRNA.....	4
Fig. 1. 02. <i>Trans</i> -acting sRNA	7
Fig. 1. 03. General scheme indicating regulation of gene expression involving bacterial sRNAs.....	11
Fig. 1. 04. GadY dependent translational activation.....	12
Fig. 3. 01. Organization of plasmid pPDL2 borne <i>opd</i> gene cluster in <i>Sphingobium fuliginis</i> ATCC 27551 and role of <i>opd</i> gene in OP insecticide degradation.....	44
Fig. 3. 02. The <i>opd</i> island borne esterase (<i>orf306</i>) induced metabolic shift in <i>E. coli</i> MG1655.....	46
Fig. 3. 03. The Venn diagram showing the segregation of Orf306 responsive ncRNAs.....	47
Fig. 3. 04. The segregation of ncRNAs found in Sector 6.....	48
Fig. 3. 05. Expression profile of Orf306 responsive ncRNAs in <i>E. coli</i> MG1655 (pSDP10) cells....	48
Fig. 3. 06. Schematic representation of sRNAs nomenclature.....	49
Fig. 3. 07. Organization of Orf306 dependent up-regulated sRNA coding genes.....	49
Fig. 3. 08. Organization of Orf306 dependent down-regulated sRNAs in presence of Orf306.....	50
Fig. 3. 09. Functional segregation of Orf306 responsive mRNAs and their interactions with seco10054A and seco11136A.....	52
Fig. 3. 10. The organisation of <i>hca</i> operon and phenyl propionate degradation pathway.....	56
Fig. 3. 11. The seco10054A / target mRNA interactions.....	57
Fig. 4. 01. Genetic screen used to validate sRNA-target mRNA interactions.....	59
Fig. 4. 02. Cloning strategy used to generate promoter- <i>lacZ</i> fusions.....	62
Fig. 4. 03. Cloning of sRNA, seco10054A coding gene, seco10054A in expression Vector, pMMB206.....	63
Fig. 4. 04. Northern blot analysis to detect pAM10054A coded sRNA, seco10054A.....	64
Fig. 4. 05. Genetic screen.....	65
Fig. 4. 06. Validation of sRNA, seco10054A and target mRNA interactions.....	67
Fig. 4. 07. Structural complementarity between <i>lpd</i> mRNA and seco10054A.....	68
Fig. 4. 08. Relative expression levels of <i>lpd</i> mRNA.....	69
Fig. 4. 09. Construction of <i>seco10054A</i> ^{Δ61-78} variant.....	70
Fig. 4. 10. The sRNA, seco10054A ^{Δ61-78} and <i>lpd</i> mRNA interactions.....	71
Fig. 4. 2. 1. Expression of sRNA, seco11136A.....	72
Fig. 4. 2. 2. Cloning strategy used to generate translation fusions of <i>acek-lacZ</i> , <i>fadR-lacZ</i> and <i>hcaR-lacZ</i> fusions.....	74
Fig. 4. 2. 3. Validation of sRNA, seco11136A and target mRNA interactions.....	75
Fig. 4. 2. 4. Relative fold difference of <i>aceK</i> , <i>fadR</i> , and <i>hcaR</i> specific transcripts in presence of sRNA, seco11136A.....	76

Fig. 4. 2. 5. Role of lipoamide dehydrogenase (Lpd) in carbon catabolic pathways.....	78
Fig. 4. 2. 6. Reactions catalyzed by <i>hca</i> operon coded enzymes.....	80
Fig. 4. 2. 7. Heat map showing Orf306 dependent repression of sRNA, <i>seco11136A</i> expression.....	80
Fig. 5. 01: Nucleotide sequence of intergenic region found between <i>lptD</i> and <i>djlA</i>	83
Fig. 5. 02: The sequence of intergenic region found between <i>icd</i> and <i>ymfD</i> genes.....	84
Fig. 5. 03: Two plasmid assay.....	85
Fig. 5. 04: The Orf306 dependent expression of <i>seco10054A</i>	86
Fig. 5. 05: The Orf306 dependent repression of <i>seco11136A</i> gene.....	87
Fig. 5. 06: Detection of <i>seco10054A</i> by northern blot analysis.....	88
Fig. 5. 07: Nucleotide sequence of <i>seco10054A</i> promoter region.....	91
Fig. 5. 08: Generation of <i>purR</i> , <i>marR</i> , and <i>phoB</i> knockouts in <i>lac</i> negative <i>E. coli</i> MG1655.....	91
Fig. 5. 09: Purification of PhoB ^{6HIS}	92
Fig. 5. 10. PhoB ^{6HIS} with consensus <i>phoB</i> binding motif identified overlapping -10 region of <i>seco10054A</i> gene.....	94
Fig. 5. 2. 0. Analysis of total lipids from <i>E. coli</i> MG1655 (pSDP10) cells and <i>E. coli</i> MG1655 (pMMB206).....	96
Fig. 5. 2. 1. Influence of propionate on expression of <i>seco10054A</i> gene.....	97
Fig. 5. 2. 2. Propionate dependent repression of <i>seco11136A</i> gene.....	98
Fig. 5. 2. 3. Construction of <i>mhpR-lacZ</i> , <i>mhpA-lacZ</i> and <i>hcaR-lacZ</i> fusions.....	101
Fig. 5. 2. 4. Propionate dependent induction of <i>hcaR</i> gene expression.....	102
Fig. 5. 2. 5. Predicted PhoB regulatory network in <i>E. coli</i> MG1655 (pSDP10) cells.....	104
Fig. 5. 2. 6. Proposed model showing sRNAs <i>seco10054A</i> and <i>seco11136A</i> induced metabolic shift in <i>E.coli</i> MG1655 (pSDP10) cells.....	105



INTRODUCTION



1. Introduction

Environmental adaptation is the key phenomenon for a bacterial long lasting survival story. Unlike other forms of life which move away from contaminated sites, the soil bacteria fine-tune their physiological and metabolic responses to survive under stress conditions. (Nickerson *et al.*, 2004; Schimel *et al.*, 2007). Bacteria by sensing these external environmental signals reorient the expressome to facilitate their survival under altered environmental conditions. Complex regulatory networks make the bacteria to efficiently overcome the stress conditions like nutrient limitations, antibiotics etc. (Lopez-Maury *et al.*, 2008; Beisel and Storz *et al.*, 2010). Gene regulation exists both at transcriptional and post-transcriptional level (Wolf *et al.*, 2015; Pérez-Rueda *et al.*, 2015). At the transcriptional level, after receiving a signal from external stimuli, often through a well-structured two-component signal transduction mechanism, the transcription factors either repress or activate expression of bacterial genes (Laub & Goulian 2007; Capra & Laub 2012). Post-transcriptionally, the gene expression is regulated either through i) the existence of a riboswitch, a small sequence motif present at the 5' region of mRNA (Du Toit 2014) or ii) through the involvement of an RNA molecule with a size of 50 to 500 nt. These RNA molecules, otherwise known as bacterial small RNAs (sRNAs), either inhibit or activate translation of their cognate mRNAs (Storz *et al.*, 2011; Vogel 2009; Altuvia 2007; Vogel & Wagner 2007). A brief description is given below on regulation of bacterial gene regulation.

1. 1. Two-component regulatory system

Two-component regulatory system is one of the well characterized signal transduction pathways characterized in bacteria. It senses external stimuli and modulates gene expression to facilitate the survival of bacteria in altered environmental conditions. The two-component

regulatory system consists a sensor kinase and response regulator. The sensor kinase is in an autokinase. Upon sensing the signal it mediates its own phosphorylation by transferring the phosphate group to a conserved histidine residue. The phosphorylated sensor kinase then transfers the phosphate group to the response regulator. The phosphorylated response regulator then binds to the upstream activating sequences (UAS) found at the 5' region to activate the transcription of its cognate genes (Ortet et al., 2015) . The two-component regulatory system is widespread in bacteria and it is also found in yeast and in certain filamentous fungi (Fassler & West 2013; Yoshimi et al., 2005).

1. 2. Riboswitch mediated gene regulation

Riboswitch is a regulatory switch of mRNA that binds to the small molecules. Such interactions induce changes in the structure of mRNA resulting in modulation of expression of the protein coded by the mRNA. There exists direct interaction between mRNA and the molecules, whose synthesis is dependent on the protein coded by the mRNA. The Riboswitch is found in the 5' region of mRNA and the sequence that serve as riboswitch is designated as an aptamer and the small molecule that interacts with aptamer sequence are known as effector molecules (Garst et al., 2011; Du Toit 2014). Interactions between effectors and aptamers induce changes in mRNA structure. The changed mRNA structure either inhibits or promotes the translation of mRNA (Winkler et al. 2002; Nudler & Mironov 2004; Tucker & Breaker 2005).

1. 3. Small RNAs as transcriptional regulators

Investigations into plasmid copy number control mechanisms revealed involvement of RNA in regulatory functions (Tomizawa 1990). The sRNAs play multifunctional roles and justify a number of inexplicable regulatory effects (Storz et al., 2011). However, in recent years, the involvement of sRNAs at post-transcriptional regulation of gene expression has acquired great

attention and the phenomenon is well conserved in all kingdoms of bacteria (Storz et al., 2004). The sRNA-mediated gene regulation is advantageous over transcriptional factor mediated gene regulation (Shimoni et al., 2007). It adds several advantages to the cell. Since they are non-coding, the sRNA-mediated regulation reduces metabolic cost (Beisel & Storz 2010). Since, they are small in size, the genes coding sRNAs occupy a small region in DNA and require less energy to transcribe. In contrast, the transcriptional factors which are larger in size, occupy more space in DNA and they need to be translated into protein to observe the functional effect (Beisel & Storz 2010). Further, the sRNA-mediated regulation adds an extra faster layer of gene regulation at the post-transcriptional level either by base pairing with specific mRNAs or by titrating the regulatory proteins (Bobrovskyy & Vanderpool 2013). In fact, the sRNA-mediated regulation is stoichiometric in nature, where the production of sRNA dominates over mRNA expression (Gottesman 2004). Therefore it completely down-regulates the expression of target mRNA because all the mRNA are bound by the sRNAs. Contrastingly, if the production of mRNA dominates the sRNA, the free mRNA undergo translation. In any study-state-levels where the cells need to respond to environmental conditions they have to face these characters and hence sRNA-mediated regulation is believed to be advantageous to the cell (Beisel & Storz 2010; Bobrovskyy & Vanderpool 2013).

1. 4. The sRNAs mediated gene regulation

The sRNAs regulate gene expression by interacting with specific mRNA by complementarity base pairing. The sRNA-target mRNA interactions either down-regulate the translation of target mRNA or up-regulates target mRNA expression by unzipping the otherwise masked ribosomal binding site (Thomason & Storz 2010). The sRNA genes either reside within the genes coding for target mRNA or exist elsewhere in the genome. Based on their location they

are named either as *cis*-regulatory RNAs or *trans*-regulatory RNAs (Vogel 2009; Desnoyers et al., 2013).

1. 4. 1. The *cis*-regulatory sRNA

The sRNAs transcribing from the same genomic context of its target mRNA is called *cis*-regulatory sRNAs. These *cis*-acting sRNAs transcribe from the same genomic locus of its cognate mRNA but always have an opposite transcription orientation (Brantl 2007; Georg & Hess 2011). This divergent transcription facilitates extended antisense complementarity between *cis*-sRNA and mRNA. The duplex formation between the *cis*-regulatory sRNA and mRNA are very stronger because of high negative energy and extended sequence complementarity which results in either the translation inhibition/activation of a target mRNA. Most of the *cis*-acting sRNAs are found in plasmids and transposable elements to maintain copy number and stringent regulation (Brantl 2007; Wagner & Simons 1994).

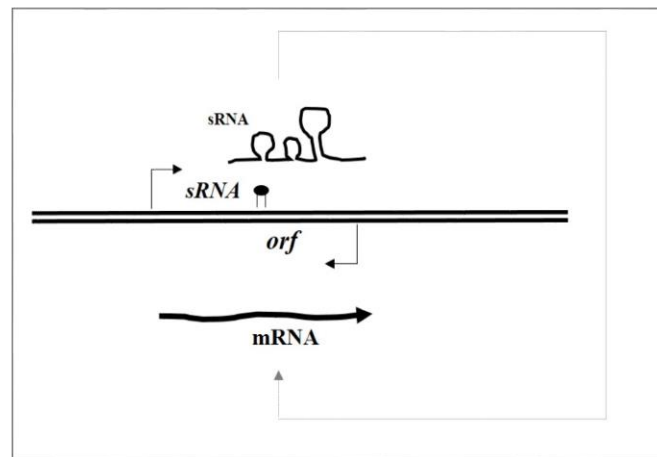


Fig. 1. 01. *Cis*-acting sRNA

The first bacterial antisense RNA with a regulatory function was identified in plasmids (Tomizawa 1990). The antisense RNA, RNAI, is shown to involve in control the copy number of plasmids having ColE1 type of replicative origin (Tomizawa and Itoh et al., 1981). The ColE1

plasmid *ori* encodes an RNA pre-primer known as RNA II which forms DNA-RNA hybrid and is encoded from a sense promoter P2. This hybrid is cleaved by RNase H to produce a 3'-OH end and this is used to initiate the leader strand synthesis which is led by DNA polymerase. This leads to the initiation of replication fork and the replication is completed by DNA replicative enzymes. ColE1 replication is regulated by antisense sRNA, RNAI which is of 108 nt in length and is encoded by promoter P1. The sRNA, RNAI inhibit the formation of pre-primer RNAII which is involved in the ColE1 DNA replication because the antisense promoter is stronger than sense primer and the result is specific copy number per cell under given conditions (Brantl 2007).

Especially the *cis*-acting sRNAs are also involved in regulation of Toxin and Antitoxin (TA) module by blocking the translation of toxic proteins/ helper proteins required for the translation of toxins. TA module codes for a stable toxin whose over-expression leads to the death of the cell or growth retardation. It also encodes an unstable antitoxin which has a role in neutralizing the toxin. Till now five toxin and antitoxin systems are known. In type II, IV and V Toxin–Antitoxin systems, both toxin and antitoxin are peptides, whereas in type I and type III Toxin–Antitoxin systems, antitoxin is sRNAs and regulates the expression toxin proteins. The encoding of Toxin and Antitoxin is transcribed almost always on opposite strands of the DNA (Brantl & Jahn 2015).

In type I Toxin-Antitoxin systems, the antitoxin is a *cis*-sRNA which directly base pairs and blocks the translational of the mRNA which codes for a toxin peptide. The example of type I is the well-studied *hok/sok* TA system. The *hok* (host killing) gene codes for a toxin protein and *sok* (suppression of killing) gene codes for an antisense sRNA. The *hok/sok* Toxin-Antitoxin system is a post-segregational killing system applied by an R1 plasmid in *E. coli*. It is also the

first type I Toxin-Antitoxin system where the toxin is neutralized by an antisense RNA, instead of toxin protein. The *hok* gene encodes a 52 amino acids long toxic protein which causes cell death by depolarizing the cell membrane, in a manner similar to that of the bacteriophage holin, a protein produced before cell-lysis. The *sok* is an antitoxin sRNA which transcribes with antisense to the toxin mRNA and it will directly base pairs with the *hoK* mRNA and blocks the translation. Interestingly in type I Toxin-Antitoxin system, a third protein also plays a key role in functioning of toxin antitoxin system. In case of well-studied *hok/sok* Toxin-Antitoxin system, a protein called Mok (modulation of killing) facilitates the translation of HoK protein. The antitoxin sRNA, Sok binds to the SD sequence of *mok* mRNA and inhibits its translation. Such translation inhibition of *mok* gene indirectly prevents *hoK* mRNA translation. There are quite a few examples of type I Toxin-Antitoxin systems in *E. coli* (*ldr/Rdl*, *tisB/IstR1*, *ibs/Sib*, *shoB/OhsC* and *symE/SymR* (Fozo et al., 2008; Kawano et al., 2012; Fozo et al., 2012).

In type III Toxin-Antitoxin system antitoxin sRNAs neutralize the toxins by binding to the protein. In type III Toxin-Antitoxin system, the toxin and antitoxins are co-transcribed from the same promoter present upstream of antitoxin gene. The ToxN/ToxI (ToxN is toxin and ToxI is antitoxin) is the first type III TA system and is plasmid-encoded toxin/antitoxin system. It is found on plasmid, pECA1039 of *Pectobacterium atrosepticum* (Blower et al., 2011). The sRNA, ToxI folds into pseudoknot where three molecules of ToxI binds to three antitoxin ToxN proteins. This trimeric ToxI/ToxN complex makes the inhibition of ToxN function (Blower et al., 2011).

1. 4. 2 Bacterial *trans*-acting sRNAs

Unlike the *cis*-acting sRNAs, *trans*-acting sRNAs transcribe distant from its target mRNA. This *trans*-acting sRNAs shares a limited sequence complementarity with its target

mRNAs. The *trans*-acting sRNA regulation is advantageous to the cell. Due to limited sequence complementarity towards the target mRNA, the sRNA can directly base-pair near translational initiation region of mRNAs. The potential interaction of sRNA-mRNA depends on the seed region (Maki et al., 2010). Most of the *trans*-acting sRNAs are dependent on the Hfq protein for the successful duplex formation with the target mRNA (Vogel & Luisi 2011).

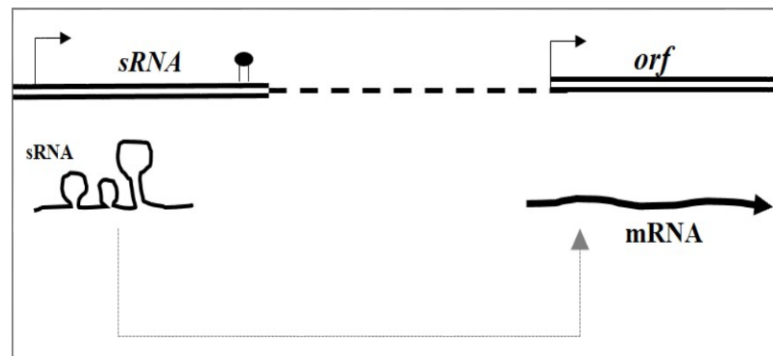


Fig. 1. 02. *Trans*-acting sRNA

In *E. coli* most of the characterized *trans*-acting sRNAs acts at the 5' UTR region or near the translational initiation site of target mRNA with limited sequence complementarities (Li et al., 2012). Such interaction leads to either the translation repression or activation. The *trans*-acting sRNAs, the interacting region or seed region with target mRNAs mostly range between 10-25 nt (Papenfert et al., 2010; Peer & Margalit 2011). In contrast to the *cis*-acting sRNAs, where most of them transcribe constitutively, a majority of *trans*-acting sRNAs transcribe under stress or different environmental conditions (Beisel & Storz 2010).

A well-studied example of *trans*-acting sRNA, MicF, which represses the translation of target mRNA. In 1984, Mizuno et al. reported the first chromosomally encoded sRNA. The MicF is a *trans*-acting sRNA. The discovery of MicF sRNA brought true strength to the sRNA-mediated gene regulation. The sRNA, MicF is of 93 nt length, Hfq-dependent and is transcribed

from the upstream region of the *ompC* gene and its expression is regulated by several transcription factors (Mizuno *et al.*, 1984). Especially H-NS and Lrp, repress its expression and the transcription factors like OmpR, MarA, SoxS contribute for *micF* transcription activation (Huang *et al.*, 1990; Ferrario *et al.*, 1995; Deighan *et al.*, 2000). The sRNA, MicF is shown to inhibit translation of *E. coli ompF* mRNA encoding outer membrane protein OmpF (Mizuno *et al.*, 1984). The OmpF highly expressed under low osmolality and represses in high osmolarity. The OmpF modulates passive diffusion across the membrane and its expression is regulated by the transcription factor, OmpR. In addition to its role in the regulation of OmpF, the OmpR activates the transcription of sRNA, MicF. The MicF inhibits translation of OmpF mRNA. This kind of regulatory network is called feed-forward loop (Ferrario *et al.*, 1995; Deighan *et al.*, 2000).

Most of the *trans*-acting sRNAs are known to down-regulate their target mRNA expression and promote their degradation. Of all these, the role of sRNA, SgrS is a classic example. Glucose-phosphate responsive sRNA, SgrS (Sugar transport related sRNA) is of 227 nt in length and depends on Hfq and it is activated by the transcription factor SgrR when the cell is grown under glucose-phosphate stress (Rice & Vanderpool 2011). The phosphoenolpyruvate phosphotransferase system (PTS) components EIICB^{Glc}, coded by *ptsG* and EIICD^{Man}, coded by *manY* and *manZ*, are responsible for uptake and phosphorylated glucose molecules. Accumulation of these sugar-phosphates in the cell induce glucose-phosphate stress (Maki *et al.*, 2010; Papenfort *et al.*, 2013). The transcription factor SgrR then activates the transcription of *sgrS* under glucose -phosphate stress, which in turn inhibits translation of *ptsG* and *manXYZ* mRNAs coding for phosphorylated sugar transporters. The sRNA, SgrS is unique in nature, in contrast with other sRNAs, the 3' end of the SgrS is engaged with down-regulation

of *ptsG* mRNA, the 5' end of it encodes a 43 amino acid protein called SgrT which is also involved in the glucose-phosphate stress. Along with *ptsG* mRNA, SgrS down-regulates *manXYZ* (Rice and Vanderpool 2011; Rice, Balasubramanian and Vanderpool 2012) and *sopD* (Papenfort et al., 2012). Interestingly, the SgrS also contributes for activation of gene expression. It contributes to *yigL* gene activation (Papenfort et al., 2013) and this aspect is described elsewhere in the introduction chapter.

Table 1. A: List of functionally validated sRNAs in *E. coli*.

sRNA	Target mRNA	Regulated protein / function	Reference
MicF	<i>ompF</i> mRNA	porin, OmpF	Andersen et al., 1987
MicC	<i>ompC</i> mRNA	porin, OmpC	Chen et al., 2002
MicA	<i>ompA</i> mRNA	porin, OmpA	Bossi et al., 2007
DsrA	<i>hns</i> mRNA	thermoregulation	Sledjeski et al., 1996
RprA	<i>rpoS</i> mRNA	general stress	Majdalani et al., 2001
OxyS	<i>fhlA</i> mRNA	oxidative stress	Altuvia et al., 1997
RyhB	<i>sodB</i> mRNA etc.	iron homeostasis	Massé & Gottesman 2002
IstR	<i>tisAB</i> mRNA	SOS response	Darfeuille et al., 2007
RdlD	<i>idrD</i> mRNA	purine metabolism	Kawano et al., 2002
GadY	<i>gadX</i> mRNA	acid stress	Opdyke et al., 2004
DicF	<i>ftsZ</i> mRNA	cell division	Bouche et al., 1989
GcvB	<i>oppA+dppA</i> mRNA	peptide transport	Sharma et al., 2007
Spot42	<i>galK</i> mRNA etc	sugar metabolism	Ikemura et al., 1973
SgrS	<i>ptsG</i> mRNA	sugar metabolism	Vanderpool et al., 2004
CsrB CsrC	CsrA protein	carbon metabolism, virulence	Weilbacher et al., 2003

1. 5. The sRNAs contribute for translational inhibition

When *E. coli* cells are growing in iron sufficient conditions, expression of the iron-uptake genes is repressed by the transcription factor, Fur (ferric uptake regulator protein). Fur repress transcription of iron-regulated genes by binding to the promoter region (Fur box) after interacting with ferrous-iron (Fe^{2+}) as a cofactor (Massé & Gottesman 2002). When cells encounter iron-limiting conditions, in the absence of iron the Fur turns inactive. Since iron free Fur fails to bind fur box, the iron-responsive genes get activated under iron limiting conditions. Transcription factor, Fur also acts as a positive regulator for the *acnA*, *fumA* (encode iron-binding enzymes of the TCA cycle), *ftnA*, *bfr*, and *sodB* (*sodB* encodes a Fe-superoxide dismutase). The sRNA, RyhB coding *ryhB* gene is part of Fur regulon. It is negatively regulated under iron-rich conditions. Under iron-limiting conditions, RyhB is de-repressed by Fur. The RyhB represses the translation of multiple mRNAs involved iron metabolism. Especially the mRNAs coding for iron binding proteins like AcnA, FumA, FtnA, Bfr, and SodB are translationally inhibited by RyhB (Massé & Gottesman 2002; Gottesman 2004; Wilderman et al., 2004). The down-regulation of iron-containing proteins facilitates iron accessibility for essential proteins under iron-limiting conditions (Storz et al., 2011).

1. 6. sRNA mode of action II - Translational activation

sRNA regulators are shown to activate the translation of its target mRNAs. In some of the mRNAs, the 5' UTR regions are self-complementary in nature and form loops making ribosomal binding sites inaccessible to the translational binding site (Fig. 1. 03, Panel D). The sRNA sequences base pair by antisense mechanism with the UTR of cognate mRNAs to liberate the

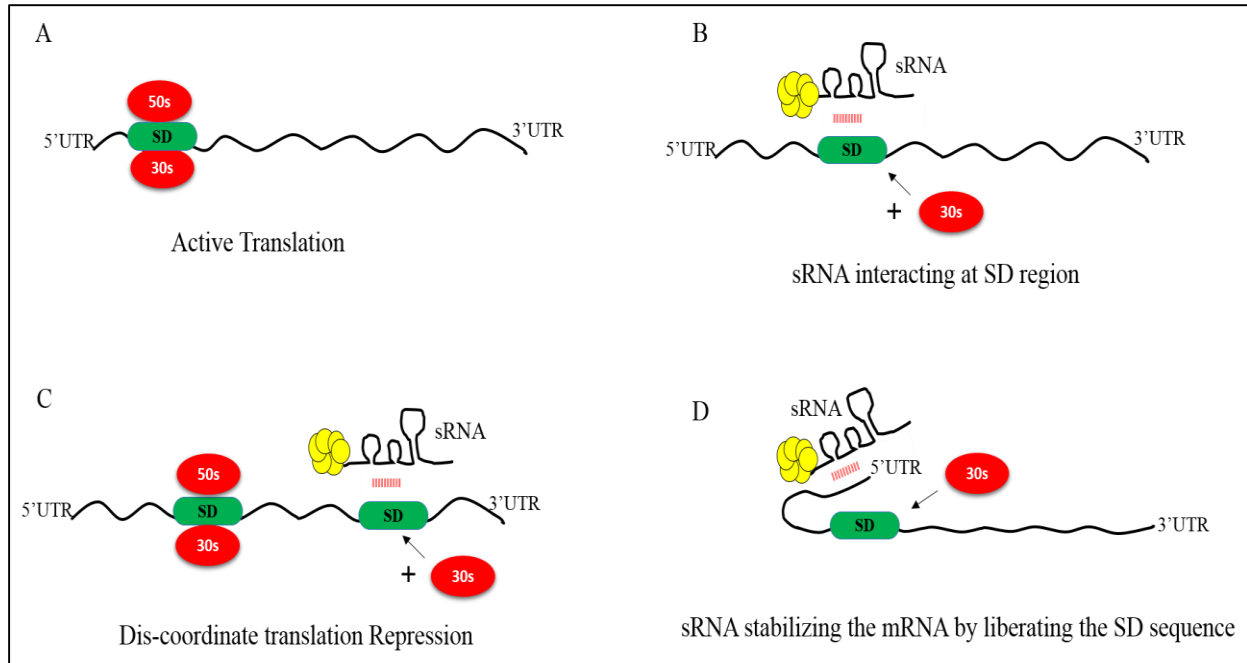


Fig. 1.03. Bacterial sRNAs and target mRNA interactions: The mRNA and ribosome interactions are shown in panel A. The sRNA-target mRNA interaction at ribosomal binding site is shown in panel B. The sRNA interaction at 3' region of mRNA is shown in panel C. Interaction of sRNA at 5'UTR of mRNA is shown in panel D.

ribosomal binding site facilitating translational activation of cognate mRNAs (Papenfort & Vanderpool 2015; Prévost et al., 2007). The sRNA, GadY-mediated activation of target mRNAs by base pairing at the 3' end is a well-studied example of the translational activation mechanism by sRNAs (Opdyke et al., 2004). The *cis*-acting sRNA, GadY is 105 nt in length and is encoded by a *gadY* gene located in the *gadX* and *gadW* intragenic region. The GadX and GadW are two transcription factors. They play a role in controlling the glutamate-dependent acid resistance system (Papenfort & Vanderpool 2015). Over-expression of sRNA, GadY activates the synthesis of *gadX* mRNA. As the GadY transcribes exactly in opposite strand of target mRNA *gadX* the sRNA, GadY shares an extensive sequence complementarity with *gadX-gadW* mRNA. This interaction promotes the processing at the complementary region of the sRNA, GadY and *gadX-gadW* by RNase III. The GadY-mediated splitting of *gadX* and *gadW* mRNA expression levels

were up-regulated because the processed fragments are more stable when compare to the unprocessed *gadX–gadW* mRNA (Opdyke et al., 2004).

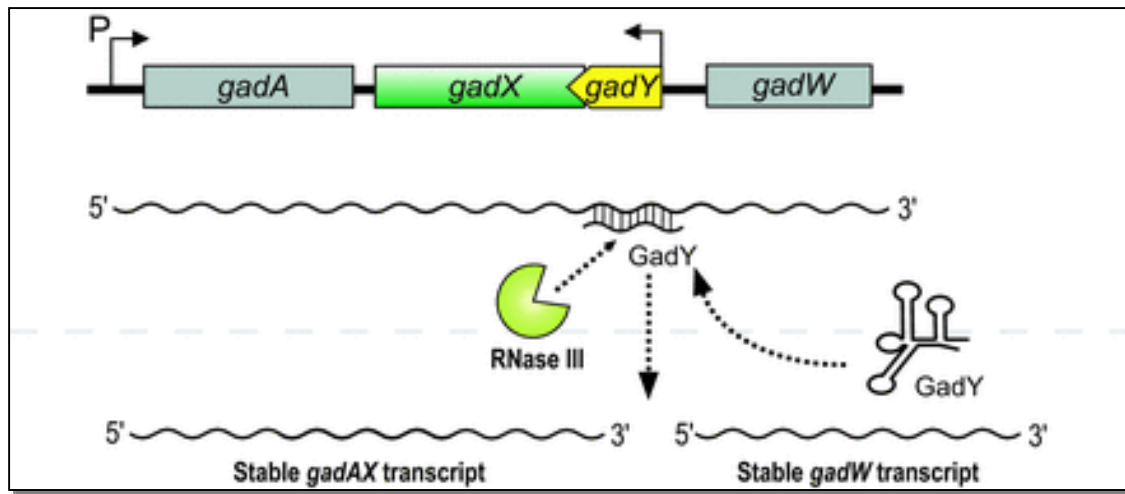


Fig. 1. 04. GadY dependent translational activation: The *cis*- encoded sRNA, GadY pairs at the 3' end of the *gadX* mRNA with sequence complementarity leads to the RNase III-dependent cleavage which generates a stable individual *gadX* and *gadW* transcripts than full length *gadAXW* mRNA (Papenfort & Vanderpool 2015).

The sRNA, SgrS dependent activation of *yigL* mRNA (which encodes haloacid dehalogenase-like phosphatase) is one of the best-studied examples for the translation activation mechanism by sRNAs. The *yigL* gene is located downstream of *pldB*, both of them are part of a bicistronic operon. The sRNA, SgrS mediates the discoordinate expression of *pldB–yigL* mRNA by only activating the expression of YigL. This is achieved by base pairing at the coding sequences of *pldB* mRNA 3'-end region. After the binding of SgrS to the *pldB–yigL* mRNA it is subjected to the degradation by RNase E which yields an mRNA species (*pldB–yigL*) where *pldB* mRNA is truncated leading to the activation of *yigL* mRNA expression. In the sugar-phosphate stress, the sRNA, SgrS is not only limiting the synthesis of new sugar transporters, it also influences the dephosphorylation of the accumulated sugar phosphates (Papenfort et al., 2013; Papenfort & Vanderpool 2015).

1. 7. The sRNA mode of action III: Sequestration of transcription factors

The above two modes of action by sRNAs requires the base complementarity towards its target mRNAs. Whereas another mode of sRNA-mediated action is sequestration of proteins or transcriptional factors. When *E. coli* reaches the stationary phase the sRNA, designated as 6S RNA is induced, whose structure mimics the σ^{70} dependent promoter. Therefore, it sequesters the concentration of otherwise constitutively expressed sigma factor σ^{70} . Under such conditions formation σ^S containing RNA polymerase is favored to facilitate transcription of σ^S dependent genes (Wassarman & Storz 2000; Wassarman 2007).

1. 8. Role of Hfq in establishing sRNA and mRNA interactions

The 102 amino acids Hfq is RNA chaperone protein and was first identified in *E. coli* as a host factor required for the phage Q β RNA replication (Salim et al., 2012). In most of the bacteria, Hfq contains homo hexameric rings with two major RNA binding sites known as a proximal and distal sites, which directly bind to single-stranded AU-rich sequences of sRNAs. After binding to sRNAs it can alter the secondary structures that can facilitate the base-pairing between the sRNAs and its target mRNAs. The Hfq plays a key role in coordinating several physiological responses by modulating the gene expression by stabilizing the sRNAs and facilitating the sRNA-mRNAs base pairing (Vogel & Luisi 2011). The Hfq after binding to the sRNA forms an Hfq-sRNA-mRNA complex and facilitates for RNase E-dependent degradation of target mRNAs (Gottesman 2004; Vogel & Luisi 2011; Wagner 2013).

1. 9. The sRNAs act on multiple mRNAs

Ability to base-pair with multiple mRNAs by sRNAs is because of the existence of short/limited seed complementarity with its specific targets. Modulation of multiple mRNAs expression by a single sRNA could be an advantage to the cell to encounter a particular

physiological response by spending least amount of energy. The sRNA, GcvB is expressed in exponential phase and the same was not seen in stationary phase. The expression of GcvB is regulated by GcvA/GcvR, the major transcription factors of the glycine cleavage system (Argaman et al., 2001; Urbanowski et al., 2000). The sRNA, GcvB modulates the expression of seven mRNAs encoding ABC transporters involved in amino acid transport/metabolism (Sharma et al., 2007). Similarly, the previously described *E. coli* sRNA, RyhB repressed by Fur protein under iron-rich conditions. When it is de-repressed, under iron-limiting conditions, the RyhB, represses expression of multiple iron-containing proteins (AcnA, FumA, Bfr, SodB etc.) that are considered to be a burden under such physiological conditions.

1. 10. Multiple sRNAs act on single mRNA

In contrast to the aforementioned situation several sRNAs act on a single mRNA target. The regulation of RpoS, encoding an alternate sigma factor which activates the transcription of multiple stress-response genes in *E. coli*. The 5'UTR of *rpoS* mRNA is approximately 600 nucleotides in size. Because of its long 5'UTR, *rpoS* forms a self-inhibitory secondary structure that shields the SD sequence and translational initiation site. (Majdalani et al., 1998). The sRNAs, DsrA and RprA upon base pairing at the upstream region of *rpoS* mRNA, disrupt its translationally inactive structure and makes RBS is accessible for the ribosome binding (Majdalani et al., 1998; Lease & Belfort 2000). However, DsrA and RprA are expressed under different physiological conditions. The DsrA expresses under low-temperature whereas the RprA is induced under cell surface stress. Along with the sRNAs, DsrA and RprA, the sRNA, ArcZ which responds to aerobic and anaerobic growth conditions is also shown to regulate expression of *rpoS* positively (Altuvia 1998; Majdalani et al., 2001; De Almeida Ribeiro et al., 2012). In contrast with the sRNAs mediated positive regulation of RpoS, the sRNA, OxyS expressed under

oxidative conditions down-regulates the translation of RpoS by engaging the Hfq protein which is implicated in positive regulation of RpoS. Converging of multiple sRNA on single mRNA leads to a similar outcome. The RpoS, stress responsive sigma factor is subjected to tight transcriptional regulation but the modulation of gene expression is achieved depending on the different external stimuli by differentially expressed sRNAs (Majdalani et al., 2000).

1. 11. The sRNA and metabolism of carbon compounds

Bacteria sense and utilize organic compounds for their survival. In general, bacteria grows in different physiological conditions where it experiences multiple carbon sources. Bacteria choose preferred carbon compounds over non-preferred ones by using carbon catabolite repression (CCR) mechanism. Such preferential usage of sugars is possible either by turning off the genes coding for alternate carbon catabolic pathways or by repressing the expression of transporters involved in the transport of alternate sugars (Görke & Stülke 2008). Transport of non-preferred/alternate carbon enter into the cell through their cognate transporters. These transporters are expressed only when the concentration of glucose in the culture medium is completely depleted. Depletion of glucose leads to increased intracellular cAMP levels and phosphorylation/dephosphorylation state of global regulatory proteins. Such metabolic conditions stimulate the formation of the cAMP-CRP complex, the global transcriptional switch required to turn on/off genes needed for degradation of alternative carbon sources. The carbon substrate to be catabolized acts as a specific transcriptional switch facilitating the transcription of genes coding for its degradative enzymes.

The sRNA, Spot 42 plays a critical role in catabolite repression. The sRNA, Spot 42 was first reported in 1979 (Sahagan & Dahlberg 1979). The 109-nt length sRNA expression levels gets increased in *E. coli* cells when it is grown under glucose sufficient conditions. Its expression

is down-regulated when the cAMP is added to the media (Polayes et al. 1988). The sRNA, Spot 42 is involved in down-regulation of genes involved in galactose metabolism. The galactose operon contains four genes, *galE* (encodes UDP-galactose epimerase), *galT* (encodes galactose-1 Phosphate uridyltransferase), *galK* (encodes galactokinase), *galM* (encodes galactose-1-epimerase). Since GalK is required for phosphorylation of galactose, it is only required when cell is grown using galactose as sole carbon source. However, GalE is an essential enzyme. It is required for cell growth, either when it is grown on glucose. The sRNA, Spot 42 repress the translation of *galK* by directly base pairing at SD region of *galK* third position of polycistronic *galETKM* mRNA. The sRNA, Spot 42 specifically down-regulates the expression of GalK, the key enzyme required for galactose metabolism and does not disturb translation of other regions of *gal* polycistronic mRNA. The *spf* gene, coding for Spot 42 is negatively regulated by CRP-cAMP complex (Polayes et al. 1988). Therefore, the Spot 42 level increases when cells are grown on glucose. Further, its expression leads to the translation inhibition of *galK* without perturbing *galE* translation (Møller et al. 2002). The GalK enzyme is needed only in the presence of galactose but the presence of GalE is always required as it is the second enzyme in synthesis of UDP-galactose, a building block for the cell wall.

1. 12. The Orf306 dependent metabolic diversion in *E. coli* MG1655

1. 12. 1. The *opd* island borne *orf306*

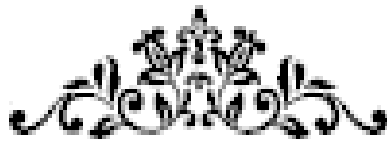
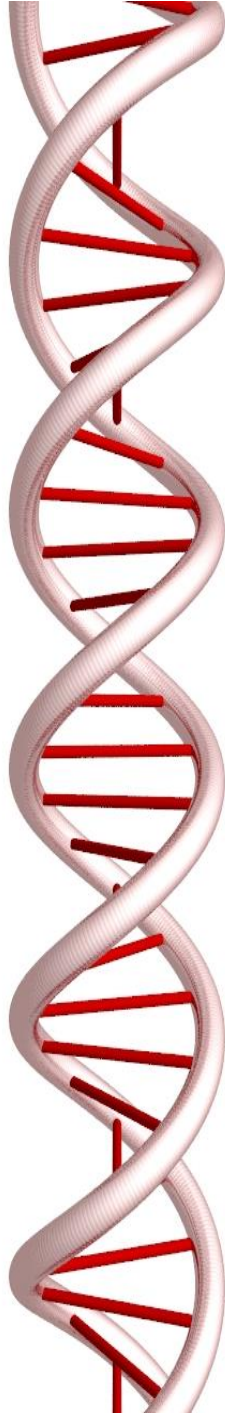
Our laboratory has been working on the microbial degradation of organophosphate (OP) compounds. Recently, we have shown the organization of indigenous plasmid pPDL2 borne *opd* cluster as Integrative Mobilizable Element (IME) (Pandeeti et al., 2012). In the IME, the *opd* gene coding for organophosphate hydrolase (OPH) and an adjacently located *orf306* coding for a

protein with an esterase domain are flanked by mobile elements *IS21* and *Tn3* (Fig. 2. 01). These two genetically linked genes are highly conserved among soil bacteria degrading OP insecticides (Mulbry et al., 1987; Pandeeti et al., 2011). Our lab has conducted several experiments to find out if there exists functional significance behind such genetic linkage. The *opd* gene codes for organophosphate hydrolase (OPH) and it catalyzes hydrolysis of triester bond found in a variety of OP compounds, including nerve agents (Siddavattam et al., 2003; Pakala et al., 2007). It hydrolyzes OP insecticides like methyl parathion into *p*-nitrophenol and diethyl phosphorothioate (Pakala et al., 2007). However, the function of *orf306* is unknown. Since it codes for 306 amino acids long protein, the ORF is designated as *orf306*. As the Orf306 showed weak homology to the proteins coding for aromatic hydrolases like TodF and CumD (Khajamohiddin et al., 2006). Our lab has done several experiments to establish its role in aromatic compound degradation (Khajamohiddin et al., 2006). While evaluating its function, the *orf306* has been cloned and expressed in *E. coli*. Interestingly, the *E. coli* MG1655 cells expressing Orf306 have grown in *p*-nitrophenol (PNP) using as a sole carbon source. While explaining such an unusual growth phenomenon our laboratory conducted both *in vitro* and *in vivo* experiments. The *in vitro* studies have shown esterase/ lipase activity for Orf306. However, it failed to act on *p*-nitrophenol. The *in vitro* assays performed with purified Orf306 using PNP as substrate. In the reaction mixture, the concentration of PNP remained unaltered even after prolonged incubation. However, the assay performed under identical conditions using a typical esterase substrate, nitrophenyl acetate, has shown a very good activity suggesting that Orf306 remained active under assay conditions. Interestingly, the resting cells of *E. coli* MG1655 expressing Orf306 have shown PNP depletion and generation of PNP metabolites such as nitrocatechol and benzenetriol (Chakka et al., 2015). Such observation suggested Orf306

dependent induction of pathways coding for PNP degradation. The transcriptome analysis was done for *E. coli* cells expressing Orf306. It has revealed induction of alternate carbon catabolic operons and inhibition of glycolysis and part of TCA cycle in Orf306 induced *E. coli* cells. Such a shift in catabolic pathways suggest existence of regulation at multiple steps in carbon catabolic pathways. The bacterial sRNAs are known to regulate carbon catabolic pathways in several bacterial strains. The transcriptome analysis has clearly indicated Orf306 dependent differential expression of non-coding RNA sequences in *E. coli* MG1655 (pSDP10) cells. The present study focused to assess the role of noncoding RNAs in the observed Orf306 dependent metabolic shift by formulating the following objectives.

1. 13 Objectives

- 1) Identification of Orf306 responsive sRNAs and their target mRNAs in *E. coli* MG1655
- 2) Elucidation of interactions between Orf306 responsive sRNAs and target mRNAs
- 3) Regulation of genes, *seco10054A* and *seco11136A* coding novel sRNAs.
- 4) Role of sRNA, *seco1136A* on regulation of expression of transcription factors HcaR and FadR.



MATERIALS

AND

METHODS



2. Materials and Methods

TABLE 2. A: Chemicals used in the present study

S.No	Name of the Chemical	Name of the supplier
1	Agarose	Amersham Pharmacia Biotech, Chennai, India.
2	Copper sulphate	Sigma Aldrich Chemicals, Bangalore, India
3	DEPC (Diethyl pyrocarbonate)	Sigma Aldrich Chemicals, Bangalore, India
4	TEMED (Tetra ethyl methylene diamine)	Sigma Aldrich Chemicals, Bangalore, India
5	ONPG (<i>O</i> -nitrophenyl- β -D-galactopyranoside)	Sigma Aldrich Chemicals, Bangalore, India
6	Methylene blue	Sigma Aldrich Chemicals, Bangalore, India
7	β -mercaptoethanol	Sigma Aldrich Chemicals, Bangalore, India
8	PMSF (phenyl methane sulfonyl fluoride)	Sigma Aldrich Chemicals, Bangalore, India
9	X-gal	Sigma Aldrich Chemicals, Bangalore, India
10	PerfectHyb™ Plus Hybridization Buffer	Sigma Aldrich Chemicals, Bangalore, India
11	2X RNA loading buffer	Thermo Fisher Scientific, Bombay, India
12	Trizol	Sigma Aldrich Chemicals, Bangalore, India
13	Magnesium sulphate	Qualigens, Delhi, India
14	Methanol	Qualigens, Delhi, India
15	Potassium acetate	Qualigens, Delhi, India
16	Potassium chloride	Qualigens, Delhi, India
17	Sodium acetate	Qualigens, Delhi, India
18	Sodium chloride	Qualigens, Delhi, India
19	Sodium citrate	Qualigens, Delhi, India

20	Sodium dihydrogen orthophosphate	Qualigens, Delhi, India
21	Coomassie Brilliant Blue G-250	Sisco Research Laboratories (SRL) Mumbai, India
22	EDTA (Ethylene diamine tetraacetic acid)	SRL, Mumbai, India
23	Ethidium bromide	SRL, Mumbai, India
24	Manganese chloride SRL	SRL, Mumbai, India
25	Magnesium sulphate	SRL, Mumbai, India
26	Methanol	SRL, Mumbai, India
27	Glycine	SRL, Mumbai, India
28	Imidazole	SRL, Mumbai, India
29	Isopropanol	SRL, Mumbai, India
30	Isopropyl thiogalactopyranoside (IPTG)	SRL, Mumbai, India
31	Manganese chloride	SRL, Mumbai, India
32	Magnesium sulphate	SRL, Mumbai, India
33	N,N'-Methylene bis acrylamide	SRL, Mumbai, India
34	Sodium chloride	SRL, Mumbai, India
35	Sodium dodecyl sulfate	SRL, Mumbai, India
36	Sodium hydroxide	SRL, Mumbai, India
37	Acrylamide	Merck, Mumbai, India
38	Peptone	HIMEDIA, Mumbai, India
40	Tryptone	HIMEDIA, Mumbai, India

Table 2. B: Antibiotics

Name of the antibiotic	Name of the Supplier
Ampicillin sodium salt	HIMEDIA, Bombay, India
Kanamycin Sulfate	HIMEDIA, Bombay, India
Tetracycline hydrochloride	HIMEDIA, Bombay, India
Chloramphenicol	HIMEDIA, Bombay, India
Streptomycin	HIMEDIA, Bombay, India
Gentamycin	HIMEDIA, Bombay, India

Table 2. C: Restriction enzymes and DNA modifying enzymes

Name of the enzyme	Name of the Supplier
<i>Bam</i> HI	Thermo Fisher Scientific, Bombay, India
<i>Bgl</i> II	Thermo Fisher Scientific, Bombay, India
<i>Eco</i> RI	Thermo Fisher Scientific, Bombay, India
<i>Hind</i> III	Thermo Fisher Scientific, Bombay, India
<i>Not</i> I	Thermo Fisher Scientific, Bombay, India
<i>Nde</i> I	Thermo Fisher Scientific, Bombay, India
<i>Pst</i> I	Thermo Fisher Scientific, Bombay, India
<i>Sal</i> I	Thermo Fisher Scientific, Bombay, India
<i>Sma</i> I	Thermo Fisher Scientific, Bombay, India
<i>Xho</i> I	Thermo Fisher Scientific, Bombay, India
Proteinase K	Thermo Fisher Scientific, Bombay, India
RNase A	Thermo Fisher Scientific, Bombay, India
DNAase	Thermo Fisher Scientific, Bombay, India
T4 DNA Ligase	Thermo Fisher Scientific, Bombay, India
<i>Taq</i> DNA polymerase	Thermo Fisher Scientific, Bombay, India
Phusion® High-Fidelity DNA Polymerase	New England Biolabs, Gurgaon, India
<i>Pfu</i> DNA polymerase	Thermo Fisher Scientific, Bombay, India
T4 Polynucleotide Kinase (T4 PNK)	Thermo Fisher Scientific, Bombay, India
GeneRuler 1kb DNA Ladder	Thermo Fisher Scientific, Bombay, India
HyperLadder 1kb DNA Ladder	Bioline, Indore, India
Unstained Protein Ladder, 10-200kDa, 2X250ul	Genetix, Delhi, India

Table 2. D: Primers used in the present study

S.No	Primer name	Sequence (5'- 3')	Purpose
1	AMPHOB:KANF AMPHOB:KANR	GATGCCAGTCTGAGGTGT GAAGGATG AGCAAAGTAGCAGCTCCA GCACC	Primers used to amplify <i>phoB::kan</i> region from the genomic DNA of <i>E. coli</i> MG1655 <i>phoB</i> knockouts
2	AMPURR:KANF AMPURR:KANR	TCCACGCTTACACTATTTG CGTACT TCGGCATGTTGTCTGACAT GGC	Primers used to amplify <i>purR::kan</i> region from the genomic DNA of <i>E. coli</i> MG1655 <i>purR</i> knockouts
3	AMMARR:KANF AMMARR:KANR	TGTTATCCTGTGTATCTGG GTTAT AGTGCGGATTCCAGGTTGT CCT	Primers used to amplify <i>marR::kan</i> region from the genomic DNA of <i>E. coli</i> MG1655 <i>marR</i> knockouts
4	AM16SRTF AM16SRTR	AGTCGAACGGTAACAGGA AGA GCAATATTCCCCACTGCTG	16s rRNA specific primers. Used to quantify 16srRNA as an internal control during qPCR experiments.
6	AMACEKRTF AMACEKRTR	GAAAGAGCACGATCGCGT GG GCAAGCTGGCGAATAGCG TT	Primers used to quantify <i>aceK</i> specific transcript through qPCR
7	AMFADRRTF AMFADRRTR	ACCGCTAATGAAGTGGCC GA CTGCACAACGCCGACAGT TT	Primers used to quantify <i>fadR</i> specific transcript through qPCR
8	AMHCARRTF AMHCARRTR	CAGACAGCCAGACACCTT GA CGGATCGGTACTGACGAA A	Primers used to quantify <i>hcaR</i> specific transcript through qPCR
9	AMLPDRTF AMLPDRTTR	GGCTAACACCCTGGAAGT TGAAGG GACCGATGATACCGCCAC CCATTAC	Primers used to quantify <i>lpd</i> specific transcript through qPCR
10	AMENOTFF AMENOTFR	GGCCTGGAATTCGATCAG TTGGTCGAGATC ACGACGGGATCCGCAGCT GCCATACCGAC	Used to amplify 5' region of <i>eno</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>eno-lacZ</i> translational fusions using promoter probe vector pRS552.
11	AMGLTATFF AMGLTATFR	GGCCTGGAATTCGTGTTCC GGAGACCTGGC ACGACGGGATCCGTAACC GAGAGTACGGA	Used to amplify 5' region of <i>gltA</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>gltA-lacZ</i> translational fusions using promoter probe vector pRS552.
12	AMICDTFF AMICDTFR	GGCCTGGAATTCAGAGAT TATGAATTGCCGCA ACGACGGGATCCGTAAGGA ACGTTGAGTTTGCC	Used to amplify 5' region of <i>icd</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>icd-lacZ</i> translational fusions using promoter probe vector pRS552.

13	AMPFKATFF AMPFKATFR	ACTAGTTGAATTCAGTAT AAAAGAGAGCCAGACC GTGATCAGGATCCAGATA GCCGTCATAAATACC	Used to amplify 5' region of <i>pfkA</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>pfkA-lacZ</i> translational fusions using promoter probe vector pRS552.
14	AMLPDTFE AMLPDTFR	GGCCTGGAATTCATGA ATGTATTG ACGACGGGATCCGGGCCT GCCCAAGTAC	Used to amplify 5' region of <i>lpd</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>lpd-lacZ</i> translational fusions using promoter probe vector pRS552.
15	AMFADRTFF AMFADRTFR	GGCCTGGAATTCGGCCAT CGCCAGAGTGAA ACTAAGGGATCCTTGAAT GGTCAACCAGC	Used to amplify 5' region of <i>fadR</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>fadR-lacZ</i> translational fusions using promoter probe vector pRS552.
16	AMACEKTFF AMACEKTFR	GCCTGGAATTCCTGACCG GCTCCACTGAA CGACGGGATCCACCGGAG GTCACCTCGAGGAATC	Used to amplify 5' region of <i>aceK</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>aceK-lacZ</i> translational fusions using promoter probe vector pRS552.
17	AMHCARTFF AMHCARTFR	GGCCTGGAATTCCTTGGG ATCTGGCTTTTCG ACTAAGGGATCCACAGTT TTCAAGATCGC	Used to amplify 5' region of <i>hcaR</i> gene as <i>EcoRI</i> and <i>BamHI</i> fragment to construct <i>hcaR-lacZ</i> translational fusions using promoter probe vector pRS552.
18	AMSECO10054APRO BE	CTGACCGTTTGTACGCGC AACGTTACCGATGAT	Used as radiolabelled probe to detect <i>seco10054A</i> sRNA expression by performing northern blot.
19	AM10054A220F AM10054A220R	GCTACGTCGAATTCGGTC ATAGCTTGGCACGCCCA AGCTACCTGCAGTTTGTG ACGCGCAACGTTAC	Used to amplify 5' region of <i>seco10054A</i> gene to construct <i>seco10054A-lacZ</i> fusions as <i>EcoRI</i> and <i>pstI</i> fragment using promoter test vector pMP220
20	AM1136A220F AM1136A220R	AGATGAATTCCTTCATT GATTAAGACATCCCCA CATGACTGCAGTTGCACC CAATGCCTGTTGGCT	Used to amplify 5' region of <i>seco11136A</i> gene to construct <i>seco11136A-lacZ</i> fusions as <i>EcoRI</i> and <i>PstI</i> fragment using promoter test vector pMP220
21	AM10054AF AM10054AR	AGCTACGAATTCGCGGT GGTCAGCATTGA GCTAAGGATCCGTCAGT TACCACCGCCAGTTT	Used to amplify <i>seco10054A</i> gene as <i>EcoRI</i> - <i>BamHI</i> fragment to be cloned in expression vector, pMMB206.
22	AMΔ10054AF1 AMΔ10054AR1 AMΔ10054AF2	AGCTACGAATTCGCGGT GGTCAGCATTGA GACAGGATTACACTAGCG GTTAGTCTCTGAC CGCTAGTGTTAATCCTGTC CAATAG	Used to amplify <i>seco10054A</i> variant <i>seco10054A^{Δ61to78}</i> as <i>EcoRI</i> - <i>PstI</i> fragment to be cloned in expression vector pMMB206.

	AMΔ10054AR2	AGCTACCTGCAGGGCTAC GGGCTTTATCA	
23	AM11136AKSF AM11136AKSR	AGATAGAATTCGGTACCT GCCGGTTTTGTATCAACTA C CATGACTGCAGGTACTAA AGATGTATGTGAAGG	Used to amplify <i>seco11136A</i> as <i>EcoRI</i> and <i>PstI</i> fragment to be cloned in expression vector pR6KS.
24	AMPHOB23BF AMPHOB23BR	GAATTCATATGGCGAGA CGTATTCTGGT CTGCAGCTCGAGAAAGCG GGTTGAAAAACGATAT	Used to amplify <i>phoB</i> gene as <i>NdeI-XhoI</i> fragment to be cloned in Expression vector pET23B
25	AMSPR10054AΔPHO BF AMSPR10054AΔPHO BR	GGCAATCATGGTGGCCAG GAGAGTGGGGATACGTTT GATTTTA TAAAATCAAACGTATCCCC ACTCTCCTGGCCACCATGA TTGCC	Used to generate <i>seco10054A</i> promoter region without PhoB binding site.
26	AMSPR10054AF AMSPR10054AR	GGCAATCATGGTGGCCAG GAGAGTGGGGATACGTTT TTTCATACGTTGATTTTA TAAAATCAACGTATGAAA AAACGTATCCCCACTCTCC TGGCCACCATGATTGCC	Used to generate <i>seco10054A</i> promoter region with PhoB binding site.

2. 1. Bacterial strains and Plasmids used in this study

Table. 2. E: Bacterial strains and plasmids used in this study.

S. No	Strain or Plasmid	Genotype or Phenotype	Reference or Source
Strains			
1	<i>E. coli</i> DH5α	<i>supE44, lacU169 (lacZ, M15) hsdR17 recA1 endA1 gyrA96 thi1 relA1</i>	Hanahan, 1983
2	<i>E. coli</i> BL21 DE3	<i>hsdS gal(λclts857 ind1 sam7 nin5 lac UV5 T7 gene 1</i>	Studier and Moffat, 1986
3	<i>E. coli</i> pir-116	<i>F- mcr Δ(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ-rpsL (StrR) nupG pir-116(DHFR)</i>	Episcentre Biotechnologies, USA
4	<i>E. coli</i> K-12 MG1655	<i>F-, λ-, rph-1</i>	Blattner, 1997
5	<i>E. coli</i> K-12 MG1655 AM001	MG1655 with <i>lacZ</i> deletion ,	This study

6	<i>E. coli</i> K-12 MG1655 AM002	MG1655 AM001 with <i>phoB</i> deletion ,	This study
7	<i>E. coli</i> K-12 MG1655 AM003	MG1655 AM001 with <i>purR</i> deletion	This study
8	<i>E. coli</i> K-12 MG1655 AM004	MG1655 AM001 with <i>marR</i> deletion	This study
Plasmids			
10	pTZ57R/T	Amp ^R , <i>lacZ</i> ⁺ , Easy T cloning vector	Fermentas.
11	pBluescript KS(II)	Amp ^R , <i>lacZ</i> ⁺ , cloning vector	Fermentas.
12	pET-23b	Amp ^R , expression vector	Novagen.
13	pMP220	Tet ^R , promoter probe vector	Spaink, 1987
14	pMMB206	Cm ^R , Low copy, broad host range, a mobilizable expression vector.	Morales, 1991
15	pRS552	Promoter-probe vector used to generate promoter- <i>lacZ</i> translational fusions. Ap ^r , Km ^r , Tet ^R .	Simons, 1987
17	pR6KS	Km ^R , Km ^R gene cloned as <i>Dra</i> I fragment to replace <i>cat</i> gene in pMMB206.	Swetha, 2014
18	pSDP10	Cm ^R , encodes Orf30 ^{6His} , <i>orf306</i> cloned as <i>Bgl</i> II fragment under the control of the inducible <i>tac</i> promoter of pMMB206.	Chakka et al., 2015
19	pAMP10054A	Cm ^R , Generated by cloning <i>seco10054A</i> as <i>Eco</i> RI/ <i>Bam</i> HI fragment in pMMB206. Codes <i>seco10054A</i> , sRNA from an inducible <i>tac</i> promoter.	This study
20	pAM10054AΔ61to78	Cm ^R , Generated by cloning <i>seco10054A</i> ^{Δ61to78} as <i>Eco</i> RI/ <i>Pst</i> I fragment in pMMB206. Codes <i>seco10054A</i> ^{Δ61to78} , sRNA from an inducible <i>tac</i> promoter.	This study
21	pAM11136A	Km ^R , Generated by cloning <i>seco11136A</i> as <i>Eco</i> RI- <i>Pst</i> I fragment in pR6KS. Codes	This study

		seco11136, sRNA from an inducible <i>tac</i> promoter.	
22	pAM16SRRNART	The <i>16s rRNA</i> gene of <i>E. coli</i> K-12 MG1655 cloned in pTZ57R/T easy cloning vector	This study
24	pAMACEKRT	The <i>aceK</i> gene of <i>E. coli</i> K-12 MG1655 cloned in pTZ57R/T easy cloning vector	This study
25	pAMFADRRT	The <i>fadR</i> PCR product amplified from <i>E. coli</i> K-12 MG1655 and cloned in pTZ57/R easy cloning vector	This study
26	pAMHCARRT	The <i>hcaR</i> gene of <i>E. coli</i> K-12 MG1655 cloned in pTZ57R/T easy cloning vector	This study
27	pAMLDPRT	The <i>lpd</i> gene of <i>E. coli</i> K-12 MG1655 cloned in pTZ57R/T easy cloning vector	This study
28	pAMENOTF	Amp ^R , Km ^R , Tet ^R , <i>eno-lacZ</i> translation fusion generated by cloning <i>eno</i> promoter region plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552.	This study
29	pAMGLTATF	Amp ^R , Km ^R , Tet ^R , <i>gltA-lacZ</i> translation fusion generated by cloning <i>gltA</i> promoter region plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552. Amp ^R , Km ^R , Tet ^R .	This study
30	pAMICDTF	Amp ^R , Km ^R , Tet ^R , <i>icd-lacZ</i> translation fusion generated by cloning <i>icd</i> promoter region plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552.	This study
31	pAMPFKATF	Amp ^R , Km ^R , Tet ^R , <i>pfkA-lacZ</i> translation fusion generated by cloning <i>pfkA</i> promoter region plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552.	This study
32	pAMLPDFTF	Amp ^R , Km ^R , Tet ^R , <i>lpd-lacZ</i> translation fusion generated by cloning <i>lpd</i> promoter region	This study

		plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552.	
33	pAMFADRTF	Amp ^R , Km ^R , Tet ^R , <i>fadR-lacZ</i> translation fusion generated by cloning <i>fadR</i> promoter region plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552.	This study
34	pAMACEKTF	Amp ^R , Km ^R , Tet ^R , <i>aceK-lacZ</i> translation fusion, generated by cloning <i>aceK</i> promoter region plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552.	This study
35	pAMHCARTF	Amp ^R , Km ^R , Tet ^R , <i>hcaR-lacZ</i> translation fusion generated by cloning <i>hcaR</i> promoter region plus predicted seed region as an <i>EcoRI/BamHI</i> fragment in pRS552.	This study
36	pAM10054A220	Tet ^R , <i>seco10054A-lacZ</i> fusion generated by cloning <i>seco10054A</i> promoter as <i>EcoRI/PstI</i> fragment in pMP220.	This study
37	pAM11136A220	Tet ^R , <i>seco136A-lacZ</i> fusion generated by cloning <i>seco1136A</i> promoter as <i>EcoRI/PstI</i> fragment in pMP220.	This study
38	pAMPHOB23B	Amp ^R , Generated by cloning 690bp <i>phoB</i> gene in pET23b as <i>NdeI / XhoI</i> fragment. Codes for PhoB ^{6HIS}	This study
39	pAMHCARPF	Tet ^R , <i>hcaR-lacZ</i> fusion generated by cloning <i>hcaR</i> promoter as <i>EcoRI/PstI</i> fragment in pMP220.	This study
40	pAMACEKPF	Tet ^R , <i>aceK-lacZ</i> fusion generated by cloning <i>aceK</i> promoter as <i>EcoRI/PstI</i> fragment in pMP220.	This study
41	pAMFADRPF	Tet ^R , <i>fadR-lacZ</i> fusion generated by cloning <i>fadR</i> promoter as <i>EcoRI/PstI</i> fragment in pMP220.	This study

2. 2. Media and growth conditions

E. coli cells were grown at 37°C in Luria-Bertani (LB) medium or minimal media. When necessary antibiotics ampicillin (100 µg/mL), chloramphenicol (12.5 µg/mL), tetracycline (20 µg/mL), kanamycin (20 µg/mL) were added to the growth medium. For IPTG inducible promoters, cultures were supplemented with 0.5 mM to 1 mM IPTG.

All materials used to prepare media were either autoclaved for 20 minutes at 121°C before use or sterilized by filtration. The following growth media were prepared and used for the propagation of *E. coli* strains. For the preparation of solid media 2 g of agar was added to 100 mL of broth. Required concentrations of antibiotic and chemical stocks were added after cooling the medium to 40⁰ - 50⁰C.

2. 2. 1. Luria-Bertani (LB) medium

The Luria-Bertani (LB) medium (100 ml) was prepared by dissolving 1.0 gm of peptone, 0.5 gm of yeast extract and 1.0 gm of NaCl in 50 ml of distilled in water. The contents were stirred and finally made up to 100 ml with distilled water. The pH of the medium was adjusted to 7.0 with 1 N NaOH. The LB agar plates were prepared by adding 2% agar to LB broth and sterilized by autoclaving.

2. 2. 2. Minimal salts medium for *E. coli*

Minimal salts medium used to grow *E. coli* was prepared by dissolving (6 g of Na₂HPO₄, 3 g of KH₂PO₄, 1 g of NH₄Cl, 0.5 g of NaCl) in 1 Litre of double distilled water. This medium was autoclaved as described above and brought to room temperature by placing it under the laminar hood. After cooling an aliquot of MgSO₄.7H₂O and CaCl₂ solutions were added from stock to obtain a final concentration of 0.12 g of MgSO₄.7H₂O and 0.01 g of CaCl₂ per 1000 mL. The required quantity of sterile glucose (50% from stock) or propionate (20 mM from stock) solutions was added as a carbon source.

2. 3. Preparation of antibiotic or chemical stocks

2. 3. 1 Ampicillin

Ampicillin stock solution was prepared by dissolving 1g of ampicillin powder in 10 mL of sterile milli-Q water. The solution was filtered sterilized by using through filter (0.2 µM) filter and stored in 1 mL aliquots at -20 freezer until further use. Working concentrations of

100 µg/ mL of ampicillin was prepared by adding appropriate amounts of stock solution to the medium.

2. 3. 2. Chloramphenicol

Chloramphenicol stock solution was prepared by dissolving 300 mg of chloramphenicol powder in 10 mL of 70% ethanol (V/V) and filter sterilised by passing through 0.2 µM filter fitted to syringe. Working concentrations of 30 µg/ mL of chloramphenicol was prepared by adding appropriate amounts of stock solution to the medium.

2. 3. 3. Tetracycline

Tetracycline stock solution prepared by dissolving 200 mg of tetracycline powder in 10 mL of 70% ethanol (V/V) and filtered sterilised by passing through 0.2 µM filter fitted to syringe. Working concentrations of 20 µg/ mL of tetracycline was prepared by adding appropriate amounts of stock solution to the medium.

2. 3. 4. Kanamycin

Kanamycin stock solution was prepared from dissolving 500 mg of kanamycin sulfate powder in 10 mL of milli-Q water, filtered sterilised by passing through 0.2 µM filter fitted to syringe. Working concentrations of 20 µg/ mL of kanamycin was prepared by adding appropriate amounts of stock solution to the medium.

2. 3. 5. Gentamycin

Gentamycin stock solution was prepared by dissolving 200 mg of Gentamycin sulphate powder in 10 mL of sterile milli-Q water. The solution was filter sterilized by passing through filter (0.2 µM) fitted to syringe. Working concentrations of 20 µg/ mL of Gentamycin was prepared by adding appropriate amounts of stock solution to the medium.

2. 3. 6. IPTG

1 M IPTG stock solution was prepared by dissolving the 236.8 mg of IPTG in 1 mL of sterilised milli-Q water. The solution was stored in aliquots of 100 µl at -20°C. When required, appropriate amounts of the stock solution was added to the medium to get 1 mM IPTG.

2. 3. 7. X-Gal

X-gal stock solution was prepared by dissolving 40 mg of X-gal in 1 mL of N, N'-dimethylformamide. Required amounts of X-gal was taken from stock solution and added to the medium after cooling it to 45°C.

2. 3. 8. 50% Glucose

Glucose (50%) stock solution was prepared by adding 25 g of dextrose in 35 mL of double distilled water. The contents were then dissolved by warming the solution. Finally, the contents were made up to 50 mL with double distilled water. The solution thus prepared was sterilized by autoclaving. The 50% glucose solution prepared in this manner stored at room temperature until further use. This stock solution was added to minimal salts medium to get a final concentration of 0.05 g/100mL.

2. 3. 9. Propionate (1 M)

Propionate stock solution was prepared by dissolving 7.4 mg of propionate in 100 mL of double distilled water and the solution was filter sterilised by passing through 0.2 µm filter fitted to syringe and stored at room temperature until further use. The stock solution was added to the medium to get a final concentration of 20 mM propionate in the culture medium.

2. 4. DNA Manipulations: Preparation of Solutions and Buffers

2. 4. 1. Agarose gel electrophoresis

Agarose gels were used to separate DNA fragments. While preparing gels appropriate amounts of agarose was dissolved in 100 ml of 1X TAE the contents were allowed to boil till the added agarose is completely dissolved to give 0.8 to 2% (w/v) agarose solution. The molten agarose was allowed to cool to approximately 45°C and ethidium bromide (0.5 µg/ml) was added before pouring the contents into casting unit. A 1 mm thickness comb was placed into a casting unit and the assembly was allowed to solidify. Samples mixed with appropriate amounts of loading buffer were loaded carefully in the wells formed after solidification of agarose solution. Electrophoresis was performed at a constant voltage of 100-120 V.

2. 4. 2. 6X Gel loading buffer

DNA loading buffer was prepared by dissolving 12.5 mg of bromophenol blue, 12.5 mg xylene cyanol and 7.5 g of Ficoll in 75 mL of sterile water. The contents were finally made up to 50 mL and stored at room temperature.

2. 4. 3. Ethidium bromide stock solution

Ethidium bromide (50 mg) was added to 5 mL of sterile water taken in an amber colour bottle. The contents were kept overnight for shaking to ensure complete dissolution of added ethidium bromide. This solution was further diluted to get a final working stock solution of 1 mg/mL. While preparing agarose gel 2.5 µL of working stock solution was added to 50 mL of gel solution to get a final concentration of 0.05 µg/mL.

2. 4. 4. TAE Buffer

The stock solution of 50X TAE (Tris-Acetate- EDTA) buffer was prepared by dissolving Tris (121 g) , acetic acid (28.6 ml) and 0.5 M EDTA (50 ml) (pH 8. 0) in 1000 ml of sterile water . The pH of the buffer was adjusted to 8.3 and stored at room temperature. When the necessary adequate volume of the stock buffer was diluted to get 1 X TAE with sterile water and used for preparing agarose gels

2. 4. 5. PCR amplification

All the amplifications were carried out in the thermal cycler (Biorad) by suitably adjusting the PCR programme depending on the amplicon size and T_m of the primers. PCR reactions were performed in a 50µl volume containing $MgCl_2$ (2.5 mM), dNTP mix containing 200 µM each of dATP, dCTP, dGTP and dTTP mix), 10 picomole of each forward and reverse primers, 1.0 Unit of *Taq* polymerase or *pfu* DNA polymerase, 10-15 ng of plasmid / genomic DNA used as a template. Amplification products were analysed by 0.8-1% agarose gel electrophoresis.

2. 4. 6. Overlap extension PCR

Overlap extension PCR was performed to construct large DNA fragments from two smaller DNA fragments by a ligase-independent methodology. The sRNA, *seco10054A*^{Δ61-78} variant was generated by performing overlap PCR. The DNA region flanking to the seed1 regions of the *seco10054A* was amplified as fragment A and B using primer sets

(ASDΔ10054AF1: 5'-AGCTACGAATTC TCGCGTGGTCAGCATTGA-3', ASDΔ10054AR1: 5'-GACAGGATTACACTAGCGGTTAGTCTCTGAC-3', and ASDΔ10054AF2: 5'-CGCTAGTGTTAATCCTGTCCAATAG-3', ASDΔ10054AR2: 5'-AGCTACCTGCAGGGCTACGGGCTTTATCA-3'). By using primers ASDΔ10054AF1 appended with *EcoRI* and ASDΔ10054AR2 appended with *PstI*, full-length *seco10054A*^{Δ61-78} fragment was amplified with the help of complementary sequence found between fragment A and fragment B. Then it was cloned in a low copy expression vector as *E. coli* *coRI/PstI* fragment. The deletion of seed region was confirmed by sequencing the amplicon obtained by overlap extension PCR. The resulted recombinant expression plasmid is designated as pAM10054A^{Δ61-78}.

2. 4. 7. DNA ligation

Unless otherwise stated all the ligations reactions were carried at 16°C in a 20 µl reaction volume. Before performing ligation reaction the concentration of vector and insert were estimated by using Nanodrop (The Eppendorf BioPhotometer D30) and were mixed in a ratio of 1:3 (vector:insert) in a sterile 1ml Eppendorf tube. One unit of T4 DNA ligase and 2 µl of 10X ligation buffer were added and the reaction volume was adjusted to 20µl with milli-Q water. Ten micro litre (10 µl) of ligation mixture was taken to transform *E. coli* competent cells.

2. 4. 8. Preparation of ultracompetent *E. coli* cells

In order to make the DH5α ultra-competent, the cells were grown at 18°C in 250 ml of LB broth taken in the 1L flask by inoculating with a primary inoculum of 1% overnight culture. The cultures were allowed to grow at 37°C till the cell density reached to 0.4-0.5 OD at 600 nm. The culture was then chilled on ice for 20 min and centrifuged at 2500rpm for 10 min at 4°C to harvest the cells. The cell pellet was then suspended in 80 ml of ice cold Inoue transformation buffer (55 mM MnCl₂·4H₂O, 15 mM CaCl₂·2H₂O, 250 mM KCl, 10 mM PIPES [0.5 M, pH 6.7]) and centrifuged at 2500 rpm for 10 minutes at 4°C. The cell pellet was then re-suspended into 20 ml of ice-cold Inoue buffer and incubated on ice for 10 min after adding 1.5 ml of DMSO. Working quickly, the cells were dispensed into sterile microfuge tubes in 100 µl aliquots and the tightly closed tubes were then snap frozen by placing them in a bath of liquid nitrogen. The tubes were stored at -70°C until further use.

2. 4. 9. Transformation

Stored DH5 α competent cells were thawed by placing them in ice bath. About 10 μ l of ligation mixture or plasmids were added and incubated on ice for 30 min. After 30min, the cells were given heat shock at 42 $^{\circ}$ C for exactly 90 sec and immediately chilled on ice for 2 min. After giving heat shock about 800 μ l of LB broth was added to the transformed cells and incubated in a shaker incubator at 37 $^{\circ}$ C for 45 min. The cells were collected by centrifugation and re-suspended in 200 μ l of LB broth before plating on LB agar plates containing appropriate antibiotic. The plates were then incubated at 37 $^{\circ}$ C for 12 h to observe transformed colonies.

2. 4. 10. Isolation of Plasmids

E. coli cells grown overnight in LB medium were harvested by centrifuging at 6000 rpm for 5 min. The cell pellet was resuspended in 100 μ l of TEGL buffer before incubating for 5 min at room temperature. Freshly prepared lysis solution (200 μ l) containing 1% SDS (Sodium Dodecyl Sulphate), and 2 N NaOH in a ratio of 1:1 was added and the samples were incubated on ice for 5 min. Neutralising solution (150 μ l) containing 3M CH₃COONa pH 4.8 was added before incubating on ice for additional 10 min. The contents were centrifuged at 13000 rpm for 10 min and the supernatant was taken carefully into the new Eppendorf tube. The supernatant was extracted first with an equal volume of phenol and chloroform and later with chloroform and isoamyl alcohol (24:1). The upper phase was transferred to a fresh tube and double the volume of the cold absolute alcohol (100%) and 1/10th volume of sodium acetate (pH 4.8) was added before incubating for 3h at -20 $^{\circ}$ C. After incubation, samples were centrifuged at 13000 rpm for 15 min and then the precipitated plasmid pellet was washed by using 70% ethanol. The pellet was dried for 10 min at 37 $^{\circ}$ C. The DNA Pellet was dissolved in autoclaved water and stored at 4 $^{\circ}$ C until further use.

2. 5. Solutions for SDS-Polyacrylamide Gel Electrophoresis

2. 5. 1. Acrylamide Solution

Acrylamide (Merck) 30% stock solution was prepared by mixing 30 g of acrylamide and 0.8 g of N, N' - methylene-bis-acrylamide (Sigma-Aldrich Chemicals) in 70 mL of sterile water . The contents were kept for mild stirring till the contents were dissolved completely. Finally, the volume of the solution was made up to 100 mL with milli-Q water and the contents were stored at 4 $^{\circ}$ C until further use.

2. 5. 2. Resolving gel buffer

181.71 g of Tris base was dissolved in 800 mL of deionized water and pH of the solution was adjusted to 8.8 by adding 30 mL of con. HCl. Finally, the volume of the buffer solution was made up to 1L using deionised water. The contents were autoclaved for 20 min at 15 psi (1.05 kg/cm²) on liquid cycle and stored at room temperature until further use. When required the stock was added to the gel components to get a final concentration of 390 mM Tris-Cl, pH 8.8.

2. 5. 3. Stacking gel buffer for SDS-PAGE

121.14 g of Tris base was dissolved in 800 mL of deionized water and the pH of the solution was adjusted to 6.8 by adding 42 mL of con. HCl and finally the volume of the buffer was made up to 1L with deionised water. This solution was autoclaved and stored at room temperature until further use. When required the stock was added to the gel components to get a final concentration of 130 mM Tris-Cl, pH 6.8.

2. 5. 4. Tris-glycine electrophoresis tank buffer (pH 8. 3)

10X stock solution of Tris-glycine buffer was prepared by mixing 30.03 g of Tris base, 144 g of glycine (electrophoresis grade) and 10 g of SDS in 1 L of milli-Q H₂O and stored at room temperature. While running SDS-PAGE, 1 X Tris-glycine buffer containing 25 mM Tris base, 250 mM glycine, and 0.1% SDS was used.

2. 5. 5. 2X SDS-PAGE loading buffer

2X stock solution of SDS loading buffer was prepared by mixing 1 mL of 1 M-Tris-Cl pH 6.8, SDS (2 g), bromophenol blue (10 mg) and 10 mL of glycerol in 25 mL of sterile distilled water and kept at 45⁰C for 10 min for complete solubilization of SDS. After bringing the contents to room temperature β-Mercaptoethanol (0.699 mL) was added and finally made up to 50 mL with sterile distilled water. Aliquots of this buffer were distributed into screw-caped 10 ml tubes and stored at 40⁰C. When the necessary equal volume of this buffer was added to the protein samples.

2. 5. 6. Staining solution

The SDS-PAGE gels were stained with coomassie brilliant blue R to visualize protein bands. Coomassie brilliant blue R-250 (0.2 g) was dissolved in 100 ml sterile water containing

methanol (50 ml) and acetic acid (7.5 mL). The contents were filtered through coarse filter paper and stored at room temperature in an amber colour bottle until further use.

2. 5. 7. Destaining solution

The destaining solution was prepared by adding methanol (50 mL) and 7.5 mL of glacial acetic acid to 100 mL of double distilled water. The contents were freshly made and used once or twice to de-stain the SDS-PAGE gels.

2. 5. 8. PhoB^{6HIS} expression and Ni-NTA Purification

The *E. coli* BL21 DE3 competent cells were transformed with confirmed recombinant plasmid (pAMPHOB23B) and plated on LB-Ampicillin (100 µg/ml) plate. A single colony was inoculated into 3 ml of LB broth containing ampicillin and incubated overnight in an incubator set at constant shaking (180 rpm) and temperature (37°C). The overnight culture was diluted to 1% with fresh LB broth with ampicillin and incubated at 37°C till the culture OD₆₀₀ is reached to 0.5. The cultures were then induced with 0.5 mM IPTG to express the PhoB^{6HIS}. Induced cells were harvested after 3 h of induction. The centrifuged cell pellet was suspended in Ni-NTA equilibration buffer (50 mM NaH₂PO₄, 100 mM NaCl, and 10 mM imidazole at pH 8.0). The cells were lysed by sonication at 20% amplitude with 30×30 sec pulses (with 30 sec delay between pulses) on ice, with a Vibra cell Ultrasonic Processor, converter model CV33, equipped with a 3 mm probe (Sonics, Newtown, CT, USA). The expressed protein was purified using Ni-NTA column as the expressed protein was having C-terminal His-tag. Protein preparations were compared with uninduced controls to check the expression of induced proteins 12. 5% SDS-PAGE. Four ml of Ni-NTA agarose was packed in a polypropylene column and equilibrated with 20 ml of Tris-NaCl buffer with 10 mM imidazole. The supernatant fraction collected from the above procedure was applied to the column and flow through was stored at -20°C. The column was washed with 20 ml of washing buffer (Tris-NaCl buffer with 20 mM imidazole). Finally, the protein was eluted with 10 ml of elution buffer (Tris-NaCl buffer with 50, 100, 150 and 200 mM imidazole) and fractions were collected. The purity of the protein was assessed by performing SDS-PAGE.

2. 5. 9. SDS-PAGE analysis

The protein samples were analysed in 12. 5% SDS-PAGE vertical gels according to following standard procedures (Laemmli 1970). The 1 mm thick SDS-PAGE gels were

prepared by using vertical gel apparatus (Bio-Rad). The running gel of 7 cm height prepared using appropriate amounts of acrylamide (30%) solution in running gel buffer. After solidification of running gel, 4.5 % stacking gel was prepared taking appropriate amounts of acrylamide solution in stacking gel buffer. A 1 mm thickness comb was inserted to generate wells. After solidification of stacking gels. The gel apparatus was fitted to the buffer tank and the tank was filled with tank buffer. The protein samples mixed with loading buffer and the samples were boiled before loading in the wells. The electrophoresis was performed at constant voltage (120V).

2. 6. RNA analysis

2. 6. 1. Tris-Borate-EDTA (TBE) buffer

A stock solution of 10 X TBE buffer was prepared by dissolving 108 g of Tris, 55 g of boric acid in 900 ml of distilled water. Once the contents were dissolved 40 ml of 0.5 M EDTA (pH 8.0) was added and finally the volume of the contents was adjusted to 1000 ml with distilled water. The buffer was then stored at room temperature. 1X TBE was prepared by adding an adequate volume of the stock buffer. TBE 1X was used to run the 7 M urea-acrylamide RNA gels

2. 6. 2. 20XSSC

The 20XSSC (Saline-Sodium Citrate) buffer was prepared by dissolving sodium citrate (88.2 gm) (0.3 M) and sodium chloride (175.3 gm) (3 M) in 500 ml of sterile DEPC treated distilled water. The solution was made up to 1000 ml then stored at room temperature.

2. 6. 3. Isolation of small RNAs

Obtaining high-quality, intact RNA is a critical step in performing northern blot analysis for detecting sRNAs. Homogenisation of bacterial cells is required for efficient cell lysis and membrane disruption. Therefore, the dry cell pellet was subjected to freeze-thawing for 2 to 3 cycles. After freeze-thawing cycles, 50 µl of Trizol reagent was added and mixed thoroughly for 2 min by using micro-piston. Subsequently, 700 µl of Trizol reagent was added and the contents were mixed by pipetting up and down for several times and then incubated at room temperature for 10 min. After incubation, chloroform was added at the rate of 200 µl per every 750 µl of Trizol reagent. The contents were subjected for rapid vortex for 15 sec before centrifuging at 12000 rpm for 10 min. The clear aqueous layer was then carefully taken into a

new sterile tube and the RNA was precipitated by adding 500 µl isopropanol. After incubating at room temperature for 10 minutes, the RNA was precipitated by centrifuging at 12000 rpm for 10 min. After removing the supernatant, the RNA pellet was washed using 70 % ethanol and was dried for 3-4 min. The RNA pellet was stored at -80°C after dissolving in 30µl of DEPC treated water.

2. 6. 4. Phenol-Chloroform extraction

The RNA samples were subjected to Phenol-Chloroform extractions to remove proteins after ethanol precipitations. An equal volume of Phenol: Chloroform: Isoamyl alcohol (25:24:1) was added to RNA containing solution and the contents were then vortexed vigorously for a minute before centrifuging at 13000 rpm for 5 min. The upper aqueous phase was transferred to a new tube and the RNA was precipitated by adding 1/10 volume of sodium acetate (3 M) and 2-3 volumes of absolute alcohol. Then RNA was precipitated by placing the tubes at -20°C for 3 to 4 hrs and the precipitated RNA was collected by centrifuging at 13000 rpm for 30 min on a tabletop centrifuge cooled to 4°C. The supernatant was discarded and the pellet was washed with 700 ul of 70% DEPC treated ethanol. The washed RNA sample was collected by centrifuging the sample at 13000 rpm for 5 min on a table top centrifuge cooled to 4°C. The collected RNA sample was dissolved in DEPC treated water. The pellet was air dried for 3-5 minutes to remove ethanol.

2. 6. 5. DNase I digestion

Following resuspension of RNA pellet in DEPC treated milli-Q water, the RNA samples were treated with 1 unit of DNase I per 1 µg of RNA for 30 min at 37°C. Before treating with DNase I, the RNA was denatured at 65°C for 10 min and immediately cooled on ice for 5 min.

2. 6. 6. Synthesis of cDNA

To convert total RNA into cDNA, RevertAid First Strand cDNA Synthesis Kit (Thermo scientific, Fermentas) was used and performed by following the manufacture's protocol. The reaction mix containing about 3µg of total RNA and 1 µl of random hexamer primer (0.2 µg/µl) in 12 ul of DEPC-treated water was incubated at 70°C for 5 min. The reaction mixture was then transferred to an ice bath and the following components were added in sequential order: 4 µl of 5 X reaction buffer, 1 µl of Ribolock Ribonuclease Inhibitor (20 U/ µl) and 2 µl of 10 mM dNTP mix. The contents were gently mixed and incubated at 25°C for 5 min before adding 200

units of RevertAid H-Minus M-MuLV Reverse Transcriptase. The contents were further incubated for an additional 10 min at 25⁰C and shifted to 42⁰C water bath for 1 h. Finally, the reaction was stopped by heating at 70⁰C for 10 min and stored at -20⁰C until further use.

2. 6. 7. Electrophoretic separation of RNA

The electrophoretic separation of RNA was done on 7 M urea-10% acrylamide (19:1 acrylamide: bis-acrylamide) gel. The gel was prepared by using the following components

Components		
1	40% Acrylamide (19:1)	15 ml
2	Urea (g)	25.2 g
3	10X TBE Buffer (ml)	1.2 ml
4	10%Ammonium persulfate (ml)	360 ul
5	Tetramethylethylenediamine (ml)	21 ul

While preparing RNA samples, 30 µg of RNA was mixed with equal volume of RNA gel loading buffer and kept it for 20 min at 65⁰C. Before performing electrophoresis the gel was subjected to pre-run for 30 min and the wells were flushed with 0.5x TBE. After loading, sRNA samples, the gel was subjected to electrophoretic separation with 1X TBE buffer. After the completion of electrophoresis, the gel was rinsed with DEPC-treated 0.5X TBE buffer followed by staining with ethidium bromide (0.5 g/ml) for 10 min. The gel was then photographed before proceeding for northern blot analysis.

2. 6. 8. Northern blot analysis

After successful electrophoretic separation of RNA, the gel containing sRNA was electroblotted onto hybond-N⁺ membrane by using Transblot semi-dry transfer (BioRad) for 1.5 h at 20 V. Before transferring RNA form gel to Hybond N⁺ membrane, length of the gel was measured, then 4 Whatman 3mm filter papers and membrane was taken according to the gel size and placed in 0.5x TBE buffer for 15 min. Initially, the soaked whatman filter papers (two) were placed on the graphite plate without trapping any air bubbles. The gel and the Hybond N⁺ membrane were placed on the tap of these two filter papers. Finally, two more whatman filter papers soaked in 1 XTBE were added on to the membrane before placing the anode plate on the assembly. The transfer unit was then connected to a power pack and the

transfer was allowed to continue for 1 h at constant voltage 18v. After transfer cross-linking of sRNAs was done by placing Hybond N⁺ membrane in the Stratagene UV cross-linker for 2 min at 1200 kJ.

Prehybridization step was performed with 10 ml of PerfectHyb Plus hybridization buffer (Sigma-Aldrich Chemicals) for 1 h at 65°C. After pre-hybridization, the radiolabelled probe was added to the pre-hybridization buffer and hybridization was performed at 40°C for 16 h. After hybridization, the solution was discarded and the membrane was washed 3 times with 30 ml of 2x SSC for every 15 min at 40°C. Radioactivity was measured directly from the membrane using Geiger counter. The membrane emitting moderate activity was wrapped with clean film and exposed to phosphorimaging screen.

2. 6. 9. Preparation of labelled oligonucleotide probe

The DNA oligonucleotide complementary to the sRNA was end labelled following standard procedures (Sambrook et al., 1989). While performing labelling reaction, the DNA oligos, (10 pmol) were taken in a 15 µl DEPC treated water and 25 µCi of ³²P-γ-ATP, 1 unit of polynucleotide kinase (PNK), and 2 µl of 10X PNK buffer was added. The contents were made up to 20 µl with DEPC treated water before incubating for 45 min at 37°C. The free nucleotides were removed from the reaction mix by spin columns provided with QIAquick Nucleotide Removal Kit.

2. 6. 10. Quantitative Real Time PCR

When multiple gene expression analysis was performed, 16s rRNA gene was taken as internal control for the expression calibration of other genes. DNA sequences of target genes of *E. coli* were retrieved from GeneBank. The Primer 3 program available online was used for primer design (<http://www.ncbi.nlm.nih.gov/tools/primerblast/primertool.cgi>). For each target gene, a primer pair capable of amplifying a portion of the coding region of the gene of about 150-250 bp was chosen and commercially synthesized at Sigma-Aldrich Pvt. Ltd, Bengaluru, India. Quantitative PCR (qPCR) reactions were conducted in an Eppendorf qPCR machine operating with Realplex 2.2 software using 30 ng cDNA template, 0.25 µM primers, and Brilliant SYBR Green reagents (Bio-Rad). Data were normalized to 16sRNA and analyzed by absolute quantification by comparing the CT value of the test sample to a standard curve.

2. 6. 11. Absolute quantification of target genes

In order to quantify the expression levels of a target gene, the amplicons of each target gene were cloned into vector pTZ57R/T. By using these constructs as a template, a 10-60 fold dilution series was made resulting in a set of standards containing 10^2 - 10^7 copies of the target gene. The standards and test samples were assayed in the same run. A standard curve was constructed, with the logarithm of the initial copy number of the standards plotted along the X-axis and their respective Ct values plotted along the Y-axis. Finally, the copy number of the target gene in the test sample was obtained by interpolating its Ct value against the standard curve.

2. 7. Reagents for β -galactosidase activity

2. 7. 1. ONPG (*ortho*-Nitrophenyl galactoside) solution

About 400 mg of ONPG was dissolved in 100 ml of phosphate buffer prepared by dissolving ($\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ (1.61 g) and $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (0.55 g) in 90 mL sterile water (pH 7.0). This solution transferred to an amber colour bottle and stored at 4°C until further use.

2. 7. 2. 1 M Sodium carbonate

The 1 M sodium carbonate was prepared by dissolving 12.4 g of sodium carbonate in 100 mL of deionized water. The solution was stored at room temperature.

2. 7. 3. β -Galactosidase Assay

β -Galactosidase activity was performed according to the protocol described elsewhere (Miller, 1972) with minor modifications. Overnight cultures were diluted 1/100 in fresh medium and allowed to grown to mid-log phase. The cultures were then kept on ice bath for 20 min to stop growth. Cell pellet obtained from at least 2 mL of culture was resuspended in the same volume of chilled Z-buffer (40 mM Na_2HPO_4 , 60 mM NaH_2PO_4 , 10 mM KCl, 1 mM MgSO_4 and 50 mM β -mercaptoethanol pH 7.0). The cells were lysed by adding 100 μL chloroform and 50 μL 0.1% SDS (sodium dodecyl sulfate). The contents were thoroughly mixed and incubated at 28°C water bath for 5 min to bring the temperature of lysate to 28°C. The reaction was started by adding 200 μL of O-nitrophenyl- β -D-galactoside (ONPG; 4 mg/mL) and soon after observation of the formation of yellow colour the reaction was stopped by adding 0.5 mL of 1 M Na_2CO_3 . The debris and chloroform were removed by centrifugation at maximum speed for 5 min and the clear supernatant was used to measure the optical density at

420 nm and 550 nm respectively. The β -Galactosidase activity was calculated by using the formula given below and expressed in miller units. $\text{Miller units} = 1000 \times [(\text{OD}_{420} - 1.75 \times \text{OD}_{550})] / (T \times V \times \text{OD}_{600})$

T = time of the reaction in minutes.

V = volume of culture used in the assay in mLs.

The units give the change in $A_{420}/\text{min/mL}$ of cells/ OD_{600} .

2. 8. *E. coli* knockouts by P1 lysate method

To generate desired mutant strains of *E. coli* MG1655, an overnight culture of donor *E. coli* K12 W1153 having a mutation in desired gene (0.3 mL) was mixed with 10 μL of P1 phage and allowed to incubate at 37°C without shaking to facilitate phage adsorption. The P1 phage infected donor strain was then added to 10 mL of LB broth having CaCl_2 (5 mM). The culture was incubated with shaking at 37°C for 5-6 h to facilitate complete lysis of donor cells. The lysate was treated with 200 μL of chloroform to prevent further growth of the cells. The clear lysate was obtained after centrifuging the contents at 6000 rpm for 10 min. The obtained supernatant containing transducing particles was stored at 4°C until further use. When required, 100 μL phage lysate was taken in CaCl_2 (5 mM) and was used to infect the *lac* negative *E. coli* MG1655 cells. After addition of phage lysate, the cells were incubated at 37°C for 15 min and centrifuged for 10 minutes at 6000 rpm to eliminate unabsorbed phage particles. The LB broth (5 mL) containing 5 mM sodium citrate was added to the infected *E. coli* cells and allowed to incubate for 45 min at 37°C . Finally, the cells were pelleted and resuspended in fresh LB (0.3 mL) broth and plated on LB media with 5 mM sodium citrate and 30 $\mu\text{g/mL}$ of kanamycin. The kanamycin resistant colonies were then used to confirm the existence of mutation by performing colony PCR using gene specific primers.

2. 9. Lipid Extraction

The *E. coli* MG1655 (pSDP10) cells and the *E. coli* MG1655 (pMMB206) cells were allowed to grow independently in minimal media with glucose as carbon source till the culture reached mid-log phase. Subsequently, the expression of Orf306 was induced for 2 hours by adding 0.5 mM IPTG to the cultures. After induction, the OD of the cultures was measured and the equal amount of cells were taken both from cultures expressing Orf306 and cultures having expression vector pMMB206. The cells were harvested from the cultures and dissolved in 3.5

ml of chloroform: methanol (1:2 v/v). After dissolving the cultures the contents were added with 1.5ml of chloroform and 1.5 ml of double distilled water. The contents were thoroughly mixed and centrifuged for 5 min at 13000 rpm. The organic layer was slowly taken with the help of syringe fitted with a long needle. After collection, the organic layer was dried and the total lipids found in organic layer were redissolved in minimal amounts of methanol and analysed using GC-MS (Shimadzu 2010). The mass spectrum for unique peaks found in total extracted lipids of *E. coli* MG1655 (pSDP10) cells compared with the mass spectrum profile of the reference compounds found in the data base.

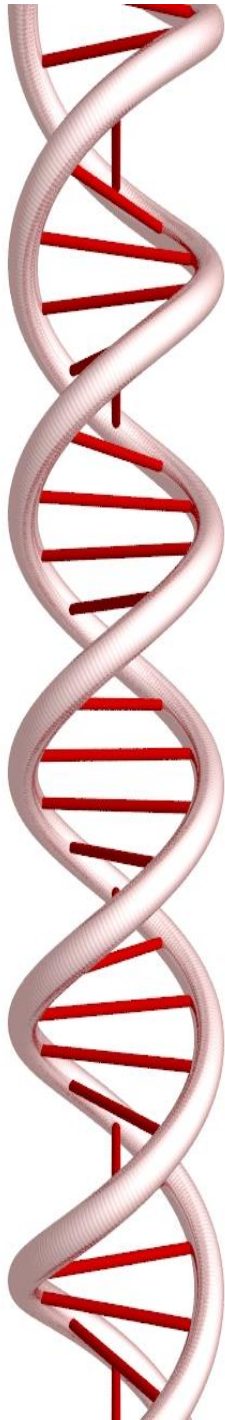
2. 10. Surface plasmon resonance (SPR)

The DNA-protein interactions were studied following surface plasma resonance technique. Biacore T200 (GE Healthcare) for performing SPR experiments. Two complementary oligonucleotides having predicted transcription factor binding sites were taken in an equimolar ratio. About 20 picomoles of each oligo taken in 30 ul of milli-Q water were placed in boiling bath for five minutes and left at room temperature for additional 30 minutes to facilitate annelation. Then the double-stranded DNA containing transcription factor binding motif was biotinylated by using (EZ-Link Sulfo-NHS-Biotin) biotinylation kit following manufacturer's protocols. Similarly treated oligos generated without transcription factor binding sites served as control DNA. These two DNA molecules were then immobilized on the streptavidin chip surface. The buffer containing 20 mM/L Tris-HCl, pH 7.5, 200 mM/L NaCl and 0.005% Tween 20 was arranged to flow continuously on the surface of the chip at the flow rate of 30 μ L/min. Purified transcription factor in a dilution series of 0.5 nM, 0.25 nM, 0.125 nM, 0.0625, 0.015 was passed over the chip surface for 1 min. Kinetic parameters were analyzed using a global data analysis program (BIA evaluation software), and the final graph was plotted with the help of an inbuilt software.

Table 2. F: Bioinformatic Tools/Applications used in the present study

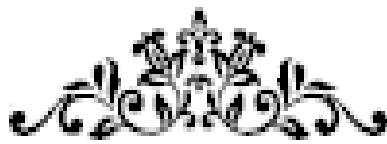
S. No.	Web server	Web address	Used for	Reference
1	Bprom	http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb	This algorithm was used to predict bacterial potential transcription	Solovyev V et al., 2011

			start sites of σ^{70} dependent promoters.	
2	Arnold	http://rna.igmors.u-psud.fr/toolbox/arnold/	This algorithm was used to predict Rho-independent terminators found downstream of genes coding non-coding RNAs.	Naville et al., 2011.
3	RNAfold	http://unafold.rna.albany.edu/?q=rfold/RNA-Folding-Form/	This algorithm was used to predict the secondary structure of sRNAs/mRNAs.	Zuker, 2003.
4	IntaRNA	http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp),	This algorithm was used to predict sRNA and target mRNA interactions	Busch et al., 2008
5	RNAPredator or	(http://nibiru.tbi.univie.ac.at/cgi-bin/RNAPredator/target_search.cgi)	This is an advanced algorithm currently available to predict sRNA-mRNA interactions. This web server was used to confirm the predicted sRNA-mRNAs interactions by using IntaRNA web server.	Eggenhofer et al., 2011
6	EcoCyc	http://ecocyc.org/	This biological database was used to perform literature-based curation of <i>Escherichia coli</i> K-12 MG1655.	Keseler et al., 2013.



CHAPTER-I

Results and Discussion



Identification of Orf306 responsive sRNAs
and their target mRNAs in *E. coli* MG1655



3. 1. Background Work

Our laboratory has been working on biodegradation of organophosphates using *Sphingobium fuliginis* ATCC 27551 as a model organism. Our previous studies have reported organization of plasmid-borne organophosphate degradation (*opd*) gene cluster both in *Sphingobium fuliginis* ATCC 27551 and *Brivundimonas diminuta* (Siddavattam et al., 2003, Pandeeti et al., 2011; Pandeeti et al., 2012). In *Sphingobium fuliginis* ATCC 27551 *opd* gene is found as part of Integrative Mobilizable Element (IME). Adjacent to the *opd* gene, in an opposite transcription orientation, an ORF was identified with a potential to code for a protein of 306 amino acids. These two genes, *opd* and *orf306*, are flanked by mobile genetic elements and exists as a complex transposon in *Sphingobium fuliginis* ATCC 27551. As shown in Fig. 3. 01, upstream of *opd* gene an *IS* element that shows similarity of IS21 region was identified. Similarly, in the downstream region of *orf306* exists a sequence that shows significant homology to the transposon, Tn3 (Siddavattam et al., 2003).

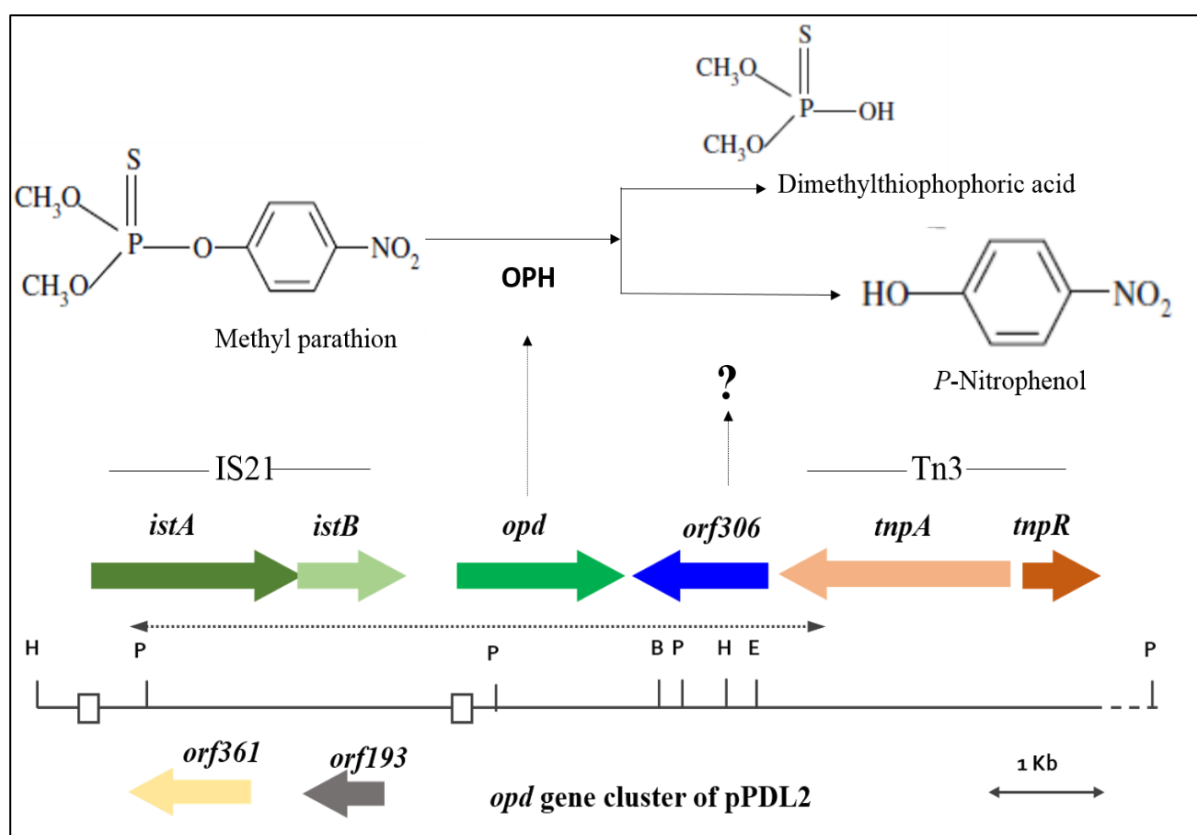


Fig. 3. 01. Organization of plasmid pPDL2 borne *opd* gene cluster in *Sphingobium fuliginis* ATCC 27551 and role of *opd* gene in OP insecticide degradation.

3. 2. The *opd* gene is highly conserved

The 7 kb *opd* gene cluster is highly conserved among soil bacteria. Identical *opd* genes were identified on otherwise dissimilar plasmids isolated from taxonomically unrelated bacterial strains located in diverse geographical regions (Mulbry et al., 1987; Somara et al., 2002). Such distribution of identical *opd* gene clusters suggested lateral transfer of *opd* genes among soil bacteria. A number of studies were conducted to elucidate the functions of various genes found in *opd* clusters. The *opd* gene codes for a triesterase that hydrolyses triester link found in structurally diverse group of organophosphate (OP) insecticides and nerve agents, (Fig. 3. 01) However the function of *orf306* is unknown. When searched for homologues for Orf306 it has shown a very weak homology with aromatic hydrolases TodF and CumD, the enzymes that participate in degradation of aromatic compounds. The Orf306 consists a well conserved esterase domain with a catalytic triad with Serine (Ser58), Asparagine (Asp183) and Histidine (His215) at the predicted active site. In consistence of its structural similarity, the Orf306 has shown esterase activity. Further, as revealed in our previous studies the Orf306 has hydrolysed metafission products generated during biodegradation of aromatic compounds (Khajamohiddin et al., 2006). While gaining further evidence on its function, our laboratory expressed Orf306 in *E. coli* and assessed its involvement in degradation of *p*-nitrophenol (PNP), the lone aromatic compound generated during degradation of certain prominent organophosphate insecticides (Fig. 3. 01). Surprisingly the *E. coli* MG1655 cells expressing Orf306 have grown using PNP as source of carbon. There was no such growth in Orf306 negative background (Chakka et al., 2015). The *in vitro* studies have shown no direct influence of Orf306 on PNP degradation. The PNP incubated with pure Orf306 remained unaltered, even after incubating the reaction mixture for a prolonged period at the optimum pH and temperature. However, under similar conditions the Orf306 has showed maximum esterase activity indicating that the purified Orf306 is active. Since *in vitro* studies suggested no direct role of Orf306 in PNP degradation, our laboratory has performed *in vivo* studies expressing Orf306 in *E. coli* MG1655 cells. Both proteome and transcriptome of *E. coli* MG1655 cells were analysed in Orf306 positive and negative background (Chakka et al., 2015). The proteomic analysis clearly showed up-regulation of proteins involved in degradation of aromatic compounds like phenyl propionate. Supplementing the proteomics data the transcriptome analysis, as revealed through DNA microarray studies, clearly indicated up-regulation of genes coding for degradation of alternative carbon catabolism. The heat-maps generated using microarray data clearly indicated down-regulation of glycolysis and part of

3. 3. Transcriptome Analysis

In the genome of *E. coli* MG1655 there are 2240 noncoding regions. The RNAs transcribed from these noncoding regions were detected by designing two probes. One of them is specific to sense stand and the second one correspond to the antisense strand. In our previous studies our laboratory has generated heat map for identified Orf306 responsive mRNAs (Fig. 3. 02). The heat-map clearly indicated up and down-regulated mRNAs and the effected catabolic pathways in presence of Orf306 (Fig. 3. 02). In this chapter, a similar study is conducted to identify expression pattern of noncoding RNAs (ncRNAs) in the presence and obscene of Orf306.

3. 4. Analysis of Orf306 responsive sRNAs

The expression profile of non-coding RNAs (ncRNAs) was obtained for *E. coli* MG1655 (pSDP10) cells. The cells were induced for expression of Orf306 and the total RNA was isolated at three (0 hrs, 1.5 hrs, and 3hrs) different time intervals. The expression profile was obtained from the generated microarray data. The expression profile of ncRNAs was then analyzed by generating Venn diagram. The Venn diagram consisted are seven sectors. The ncRNAs whose expression remained unaltered in presence of Orf306 as shown in sector 1. The ncRNAs commonly found in un-induced and cultures induced for 1.5hrs and 3hrs were represented in sectors 2, 5, and 4. The sectors 3 and 7 represent ncRNAs exclusively expressed in cultures induced for 1. 5hrs and 3hrs. The ncRNAs commonly expressed both in 1.5 and 3 hrs induced cultures was shown in sector 6.

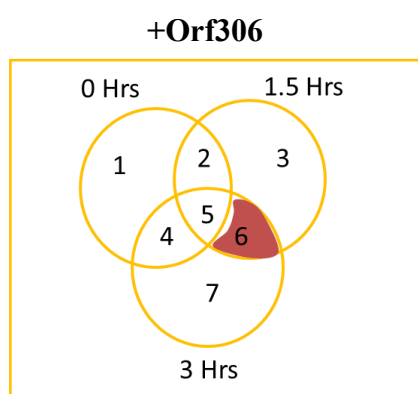


Fig. 3. 03. The Venn diagram showing the segregation of Orf306 responsive ncRNAs. Detailed description is given in the text.

After segregating the expression profile of ncRNAs, further analysis was done only for ncRNAs exclusively found in sector 6. This ncRNAs found in sector 6 were again separated into two subgroups, A and B. The subgroup A represents up-regulated ncRNAs whereas subgroup B indicates down-regulated ncRNAs. Out of 310 differentially expressed ncRNAs

found in sector 6, about 290 ncRNAs are up-regulated and hence kept them under subgroup A. The remaining 22 down-regulated ncRNAs are kept in subgroup B (Fig. 3. 04).

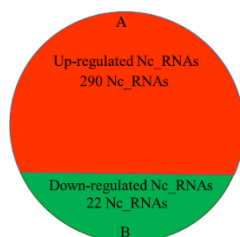


Fig. 3. 04. The segregation of ncRNAs found in Sector 6. The up and down regulated ncRNAs are shown in panels A and B.

These differentially expressed ncRNAs were further screened by using the following stringent parameters. All of them were screened to assess (i) if the difference in expression is more than 0.75 fold, (ii) if the genes coding ncRNAs have an independent promoter and Rho-independent terminator elements, (iii) if their length is anywhere between 50-500 nt. The ncRNAs satisfying all the above conditions were alone considered as small RNAs (sRNAs). The rest of the ncRNAs are not taken into consideration while conducting further studies. The stringent screening conditions have reduced the number of Orf306 responsive sRNAs to five. Out of these five sRNAs, three of them are up-regulated and two sRNAs got down-regulated.

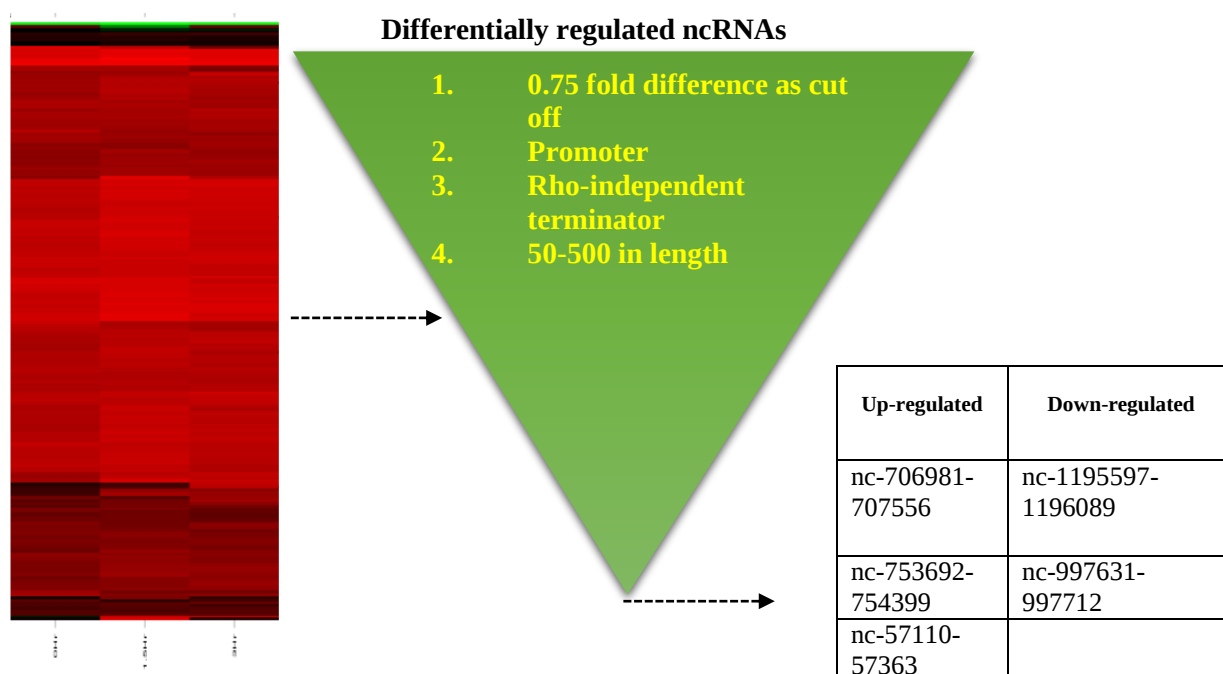


Fig. 3. 05. Expression profile of Orf306 responsive ncRNAs in *E. coli* MG1655 (pSDP10) cells. The red and green color indicate up and down regulation of ncRNAs respectively. The stringent conditions used to screen Orf306 responsive ncRNAs are shown in inverted triangle. The Orf306 responsive ncRNAs shown in Table.

3. 5. Nomenclature of Orf306 responsive sRNAs

The identified five sRNAs were given specific names by following established conventions (Lamichhane et al., 2013). All the sRNAs were given a four letter code “*seco*” where ‘s’ stands for sRNA, ‘eco’ stands for both the genus and species name. This four letter code is followed by a numeric locus tag of the gene found upstream of sRNA coding sequence. If sRNA coding genes are identified in the intergenic region, to indicate a new numerical locus tag for sRNA gene, Number ‘1’ is inserted in between the four letter code and numeric locus tag. If the sRNA is transcribed from a complementary strand, a suffix ‘c’ is added after the numerical locus tag. If the sRNA is discovered for the first time letter ‘A’ is added immediately after the locus tag. The identified five Orf306 responsive sRNAs were named as *seco10054A*, *seco10679Ac*, *seco10719A*, *seco11136A* and *seco10937A*. The detailed method followed to name the sRNAs is shown in (Fig. 3. 06). The organization of Orf306 responsive sRNAs are shown in Fig. 3. 07 and 3. 08

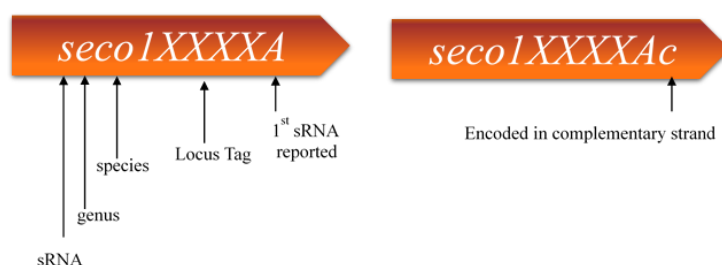


Fig. 3. 06 Schematic representation of sRNAs nomenclature

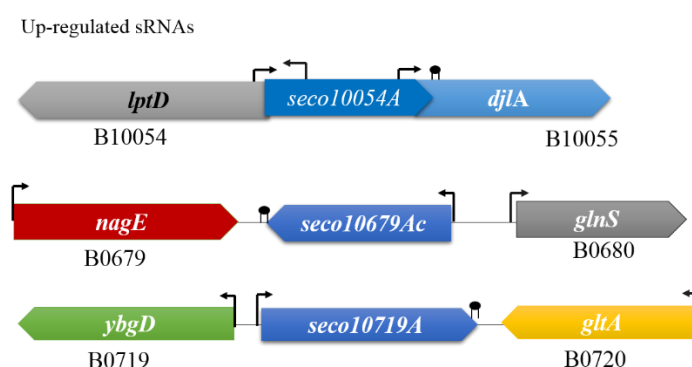


Fig. 3. 07. Organization of up-regulated sRNAs in presence of Orf306. The sRNA, *seco10054A* is located between the *lptD* (polypeptide outer membrane lipopolysaccharide transport and assembly complex) and *djIA* (co-chaperone with DnaK) genes. The sRNA, *seco10679Ac* is identified in the intergenic region found between *nagE* (N-acetylglucosamine PTS permease) and *glnS* (Glutaminyl-tRNA synthetase) genes. Similarly, in the intergenic regions of the *ybgD* (predicted fimbrial-like adhesin protein) and *gltA* (Citrate synthase) genes the sRNA, *seco10719A* is identified. The putative promoter and Rho-independent terminators are shown with bent arrows and filled circles respectively.

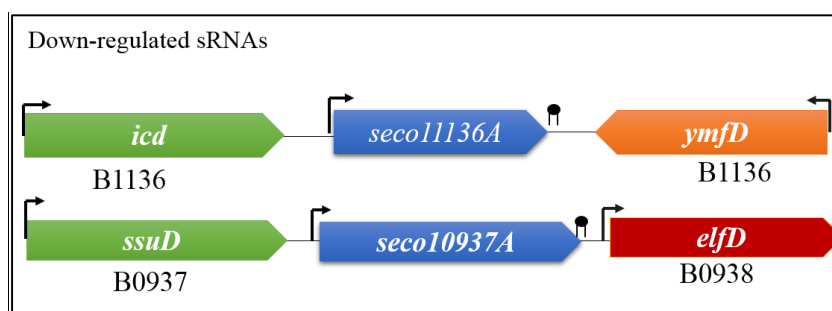


Fig. 3. 08. Organization of down-regulated sRNAs in presence of Orf306. The sRNA, *seco11136A* gene is identified in the intergenic region found between *icd* (Isocitrate dehydrogenase) and *ymfD* (SAM-dependent methyltransferase) genes. Similarly, the sRNA, *seco10937A* gene is located in the intergenic regions of *ssuD* (NADPH-dependent FMN reductase) and *elfD* (Fimbrial-like adhesin protein) genes. The putative promoters are shown with bent arrows. The Rho-independent terminator identified downstream of sRNA coding genes are shown with filled in circles.

3. 6. The sRNA-target mRNA interactions: *in silico* predictions

The sRNAs regulates gene expression by interacting either with target mRNAs or by sequestering the transcription factors involved in regulation of gene expression (Majdalani et al., 2002; Thomason et al., 2010). The sRNAs interacting with target mRNAs either inhibit or promotes translation of their cognate mRNAs (Urban & Vogel 2007; Bandyra et al., 2012). The sRNAs inhibiting translation of proteins interact at the ribosomal binding site of their cognate mRNAs and prevent initiation of translation process. Contrary to this mode of action certain sRNAs unzip the otherwise masked mRNA translation initiation site of their cognate mRNAs and facilitate their translation process. Since the length of sRNAs are fairly long, often they have multiple interaction sites within the same mRNA or with mRNA coding altogether a different protein. Therefore, the sRNA can have multiple mRNA targets. Successful sRNA-mRNA interactions require a minimum of ‘7’ base pairs length continuous base pairing (Bobrovskyy et al., 2014). The online tools that predict sRNA-mRNA interactions have been developed incorporating all these features. However, these predictions end up giving erroneous results. In this study, two prediction tools, RNApredator (http://nibiru.tbi.univie.ac.at/cgi-bin/RNApredator/target_search.cgi) and IntaRNA (<http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp>) were used to predict Orf306 responsive sRNA targets. As mentioned before, out of five identified Orf306 responsive sRNAs, three of them are up-regulated two of them are down-regulated. Assuming that the up-regulated sRNAs have targets in the mRNAs down-regulated in presence of Orf306, the up-regulated sRNAs were screened against down-regulated Orf306 responsive mRNAs and similar screening was done to predict interactions between the down-regulated sRNAs and up-regulated Orf306 responsive mRNAs.

Since the computational predictions of sRNA-mRNA interactions generate false positives. The following stringent parameters were applied to obtain valid interactions between sRNA and target mRNAs, (i) the interacting region should be around the ribosomal binding site (-30 to +20 where +1 is 'A' of translational start codon ATG), (ii) the seed region should be more than seven base pairs in length. Screening sRNA-mRNA interactions following aforementioned stringent conditions has generated a set of mRNA targets for the all five Orf306 responsive sRNAs. Out of these five sRNAs only seco10054A and seco11136A are predicted to have mRNA targets coding for enzymes involved in carbon catabolic pathways. Precisely these are the pathways that are found differentially regulated in presence of Orf06. In order to have in-depth analysis on observed Orf306 induced catabolic shift further studies were conducted only on these two sRNAs.

3. 7. Identification of mRNA targets for seco10054A and seco11136A

As stated before, while determining the sRNA target mRNAs, the Orf306 responsive mRNAs were divided into two groups based on their expression pattern. These two groups were further segregated based on their function using online tool PANTHER (<http://pantherdb.org/>). The functionally segregated down-regulated mRNAs are shown in Fig. 3. 09; Panel A. As shown in Panel A, more number of mRNAs coding for glycolysis and TCA cycle are down-regulated. Similarly, the mRNAs coding for the catabolism of alternate carbons such as phenyl propionate, phenyl ethyl amine, methyl isocitrate were up-regulated (Fig. 3. 09, Panel B). Coinciding this expression pattern, the sRNA, seco10054A up-regulated in presence of Orf306, has shown more number of target mRNAs coding for glycolysis and TCA cycle (Fig. 3. 09, Panel C). Similar analysis was done for the up-regulated Orf306 responsive mRNAs. As shown in Fig. 3. 09, Panel B more number of mRNAs coding in either phenyl propionate degradation or methylisocitrate pathway enzymes were found up-regulated (Chakka et al., 2015). Interestingly, one of the sRNA, seco11136A found a maximum number of mRNA target mRNAs coding for phenyl propionate degradation pathway (Fig. 3. 09, Panel D). This *in silico* predictions were made using online tool RNApredator were further strengthened by using second online tool IntaRNA. The predicted interactions between the target mRNAs and sRNAs were shown in Table 3. A and Table 3. B. The sRNA, seco10054A up-regulated in presence of Orf306, has shown significant interactions with mRNAs transcribed from genes coding for phosphoglucose isomerase (*pgi*), phosphofructokinase II (*pfbK*), phosphofructokinase (*pfbA*), phospho- glycerate kinase (*pgk*),

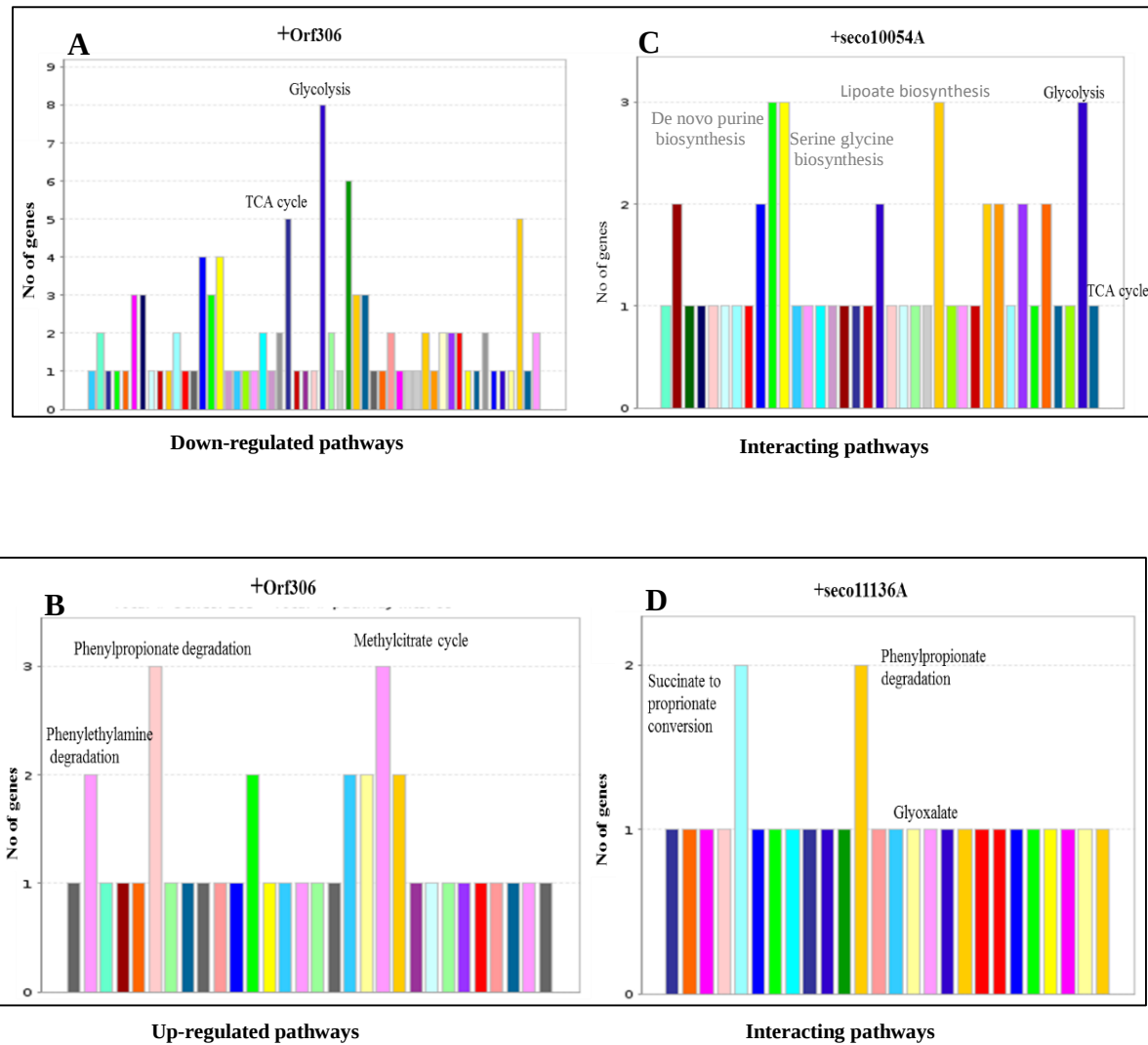


Fig. 3.09. Functional segregation of Orf306 responsive mRNAs and their interactions with seco10054A and seco11136A. Graphs were plotted taking functions on X-axis and number of genes on Y-axis. Panel A and Panel B represent functional segregation of Orf306 dependent up-regulated mRNAs and down-regulated mRNAs. The number of targets found in down-regulated mRNAs for sRNA, seco10054A are shown in panel C. The panel D indicates mRNA targets found for sRNA, seco11136A in Orf306 responsive down-regulated mRNAs.

glucose PTS permease - PtsG subunit (*ptsG*), glyceraldehyde 3- phosphate dehydrogenase (*gapA*), and enolase (*eno*). Out of these predicted interactions the seed region found between mRNA coding for lipoamide dehydrogenase and seco10054A appears to be highly stable. The seed region found at translational start codon extends from (-8 to +12) (Table 3. A). As shown in Table 3. A, the maximum sRNA, seco10054A targets are found to be involved either in glycolysis/TCA cycle. Interestingly, the sRNA seco11136A has shown interactions with mRNAs encoding HcaR, the transcriptional dual regulator. In addition to this mRNA the seco11136A has also shown interactions with mRNAs coding for isocitrate dehydrogenase phosphatase/isocitrate dehydrogenase kinase (*aceK*) and FadR, the DNA-binding

transcriptional dual regulator (*fadR*). Interestingly, *seco11136A* target mRNAs/operons are involved in the alternate carbon catabolism. The other three Orf306 responsive sRNA, *seco10679Ac*, *seco10719A* and *seco10937A* have found no targets coding for alternate carbon catabolism pathways. Therefore, they were not taken into consideration.

Table 3. A: The predicted *seco10054A* and target mRNA interactions. The seed region is numbered to indicate interacting region. The translational start site of mRNA is shown in bold case.

Gene Name	Protein	Target mRNA	Energy [kcal/mol]	Seed region
<i>pgi</i>	Phospho glucose isomerase	-210 to -177	-12.42	<pre> -210 -177 5'-CCG...UAAAG GC U AUGGG...AAA-3' GGUGA GGGGCGGU GUCAACG : : : : : : : : : : CUACU CCUUGUUA UAGUUGC 3'-ACC...AAUGG A UUU A-5' 26 1 </pre>
<i>pfbB</i>	Phosphofruc tokinase II	+19 to +34	-12.04	<pre> +19 +34 5'-UUA...UACGU U CUCUC...ACA-3' UGACACU GCGCCCU : : : : : : : AUUGUGA UGCGGGG 3'-ACC...UCCGG UUUUC...GCA-5' 330 315 </pre>
<i>pfkA</i>	phophofruct okinsase	+11 to +42	-10.22	<pre> +11 +42 5'-UCA...AAAUUC U U C G GCCAG...AUG-3' GGUG GU GA AAGCGGCG UGAUGC : : : : : : : : : : CCAC CG UU UUUGUCGC ACUGCG 3'-A U GG AGUGC...GCA-5' 426 400 </pre>
<i>pgk</i>	Phospho- glycerate kinase	+21 to +36	-9.82	<pre> 21 38 5'-GAC...UGCAA UUU AGAGG...UUU-3' AAC UAGAAUCAACG : : : : : : : UUG AUUUUAGUUGC 3'-ACC...CUACC UU A-5' 18 1 </pre>
<i>ptsG</i>	Glucose transporter	-136 to 125	-9.38	<pre> -136 -125 5'-AUG...UCAAA UGAAU...AUG-3' CAAAUUGGCAC : : : : : : GUUUAGCUGUG 3'-ACC...AAGGG UGCCU...GCA-5' 247 235 </pre>
<i>gapA</i>	Glyceraldeh yde 3- phosphate dehydrogen ase	-119 to 109	-9.18	<pre> -119 -109 5'-UGA...GCUUA C UUGUC...GUU-3' GG ACAGGAUUGAU : : : : : : : CC UGUCCUAAUUG 3'-ACC...AUAAG UGAUC...AUA-5' 100 86 </pre>

<i>eno</i>	Enolase	+24 to +40	-6.59	<pre> +24 +40 5'-GUA...UCAUC UC U ACUCC...GGU-3' GG GUGAAAUCA CG : CC UAUUUUAGU GC 3'-ACC...UACUA UUUGU U A-5' 20 1 </pre>
<i>Icd</i>	Isocitrate dehydrogenase	-44 to -28	-11.2	<pre> -44 -28 5'-UAU...CAACG C AACCA...CAA-3' UGGUGGCAGA GAGCA : : : ACCACCGUUU UUUGU 3'- CGCGG...GCA-5' 426 411 </pre>
<i>gltA</i>	Citrate synthase	-98 to -89	-9.11	<pre> -98 -89 5'-UGG...AGAGC AAUAA...GAA-3' GGCGAGCCA : CCGCUUGGU 3'-ACC...ACCAA GCGGU...AUA-5' 392 382 </pre>
<i>lpd</i>	Lipoamide dehydrogenase	-8 to +12	-8.72	<pre> -8 +12 5'-GUA...AUUAU GUC UGAAA...GCG-3' UAGAG AUGAUGAGUAC : AUCUC UACUGCUAUG 3'-ACC...AUCGC CAAUC...GCA-5' 78 61 </pre>
<i>aceB</i>	Malate synthase A	+7 to +30	-7.75	<pre> +7 +30 5'-UCU...AACAG AACAA ACUGG...UAU-3' GCAAC ACCGAUGA CGUUG UGGCUACU 3'-ACC...GUGCG CAA ACCUU...GCA-5' 37 20 </pre>

Table 3. B: The predicted seco11136A and target mRNA interactions. The seed region is numbered to indicate interacting region. The translational start site of mRNA is shown in bold case.

Gene Name	Protein Name	Target mRNA	Energy [kcal/mol]	Seed region
<i>hcaR</i>	HcaR transcriptional dual regulator	-6 to +2	-4.03	<pre> -6 +2 5'-GUG...AGGGG GGAA...UGC-3' AAGGUGAU UUCCACUA 3'-AAG...UGCGA ACGUG...CGU-5' 74 65 </pre>

<i>aceK</i>	Isocitrate dehydrogenase phosphatase / isocitrate dehydrogenase kinase	-2 to -34	-9.87	<pre> -2 +34 5'-GCA...AA<u>AUG</u> AUUUAU UCAAA...UCG-3' CCGCGUG GCCUGGA UGAUUGC : : : GCGGUAC UGGACUU ACUAACG 3'-AA A GCGAUUCC UGGGU...CGU-5' 93 62 </pre>
<i>fadR</i>	FadR-DNA-binding transcriptional dual regulator	+20 to +44	-11.82	<pre> +20 +44 5'-GAU...CAAAG UU C AA CAUUA...UCC-3' CCC GGCGGGU CG GG GAGUA : GGG CCGCUCA GC CC CUCAU 3'-AAG...GACAG A GG A CAACU...CGU-5' 42 18 </pre>

3. 8. Discussion

As stated before, previous studies from our laboratory have clearly demonstrated the existence of shift in catabolic pathways in *E. coli* MG1655 cells expressing Orf306. The TwoD gel and quantitative PCR experiments supported the transcriptomic data showing the clear up-regulation of propionate and phenyl propionate catabolic operons. Similarly, the transcripts coding for glycolytic pathway and part of TCA cycle enzymes were found down-regulated in the presence of Orf306 (Chakka et al., 2015). Such a metabolic shift clearly suggests transcriptional regulation at multiple steps. The bacterial sRNAs are known to regulate gene expression (Hoe et al., 2013; Storz et al., 2011). Since, single sRNA can have multiple mRNA targets, repression or activation of sRNA coding genes are shown to have influence on expression of various proteins/enzymes involved in similar function (Sharma et al., 2007; Richter and Backofen 2012). The role of sRNA, RyhB mediated iron uptake regulation is a classical example of sRNA elicited regulation affecting iron acquisition (Hoe et al., 2013). Likewise, the sRNAs also inhibit translation of mRNAs that code for transcriptional factors. Translational inhibition of transcription factors coding mRNAs influence the expression of multiple genes that affect a particular function. Interestingly, TwoD gel analysis of proteins isolated from *E. coli* MG1655 (pSDP10) cells have shown up-regulation of HcaR, the dual transcription factor that activates/represses a number of genes, including the *hca* operon coding for phenyl propionate degradation enzymes (Fig. 3. 10 Panel A and Panel B) (Turlin et al.,

2001). The *hca* operon is also shown to degrade PNP in presence of *opd* island encoded Orf306 (Chakka et al., 2015).

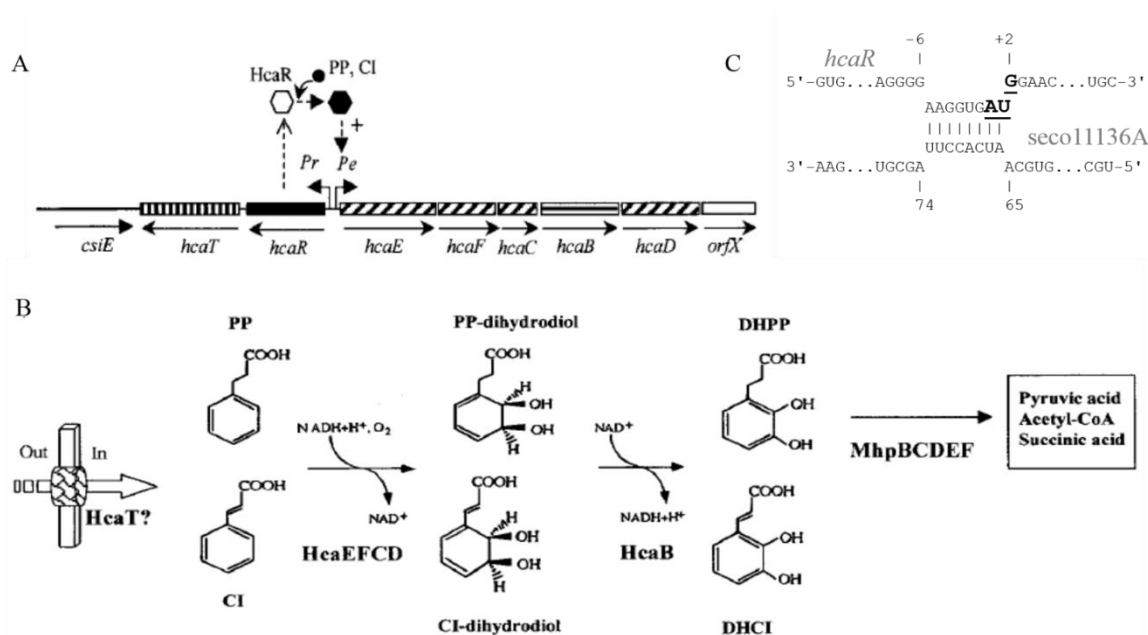


Fig. 3. 10. The organisation of *hca* operon and phenyl propionate degradation pathway as reported by Diaz and associates are shown in panels A and B (Diaz et al., 2001). The predicted interactions between sRNA, seco11136A and *hcaR* mRNA are shown in panel C

Interestingly, the sRNA, seco11136A down regulated in the presence of Orf306 has shown good seed region for mRNA coding transcription factor HcaR (Fig. 3. 10, Panel C). The HcaR repress *hca* operon, when *E. coli* cells are grown in the presence of glucose (Turlin et al., 2001). In the absence of HcaR the *hca* operon gets activated. The *hca* operon is involved in degradation of phenylpropionate (Fig. 3. 10 Panel A and Panel B). If predicted sRNA-mRNA interactions are real, the reciprocal regulation of sRNAs seco10054A and seco11136A is expected contribute repression of Lpd protein and activation of HcaR under Orf306 positive background. Lipoamide dehydrogenase (LpD) subunit is part of multienzyme complex pyruvate dehydrogenase and α -ketoglutarate dehydrogenase. Therefore its repression has potential for the down-regulation of glycolysis and TCA cycles (Fig. 3. 11). Similarly absence of sRNA, seco11136A promotes translation of HcaR which in turn activates *hca* operon which is shown to involve in degradation of PNP (Chakka et al., 2015). The sRNA, seco10054A as revealed through bioinformatic predictions shown good seed region with mRNA coding for enzymes involved in glycolysis and TCA cycle. These two catabolic pathways are down-regulated in Orf306 positive background. Since, sRNA, seco10054A is predicted to have interactions with mRNAs coding for glycolysis and TCA cycle, it is assumed that the observed down-regulation of these pathways is due to Orf306 dependent up-regulation of sRNA, seco10054A. Though the existence of seco11136A is reported in *E. coli* as C.0293, its physiological relevance is unknown (Shinhara et al., 2011). However,

the presence of seco10054A is reported only in the prescient study. The subsequent chapters of the dissertation is devoted to validate bioinformatic predictions and to draw a meaningful conclusions on the role of the novel sRNAs induced metabolic shift in *E. coli* MG1655 (pSDP10) cells.

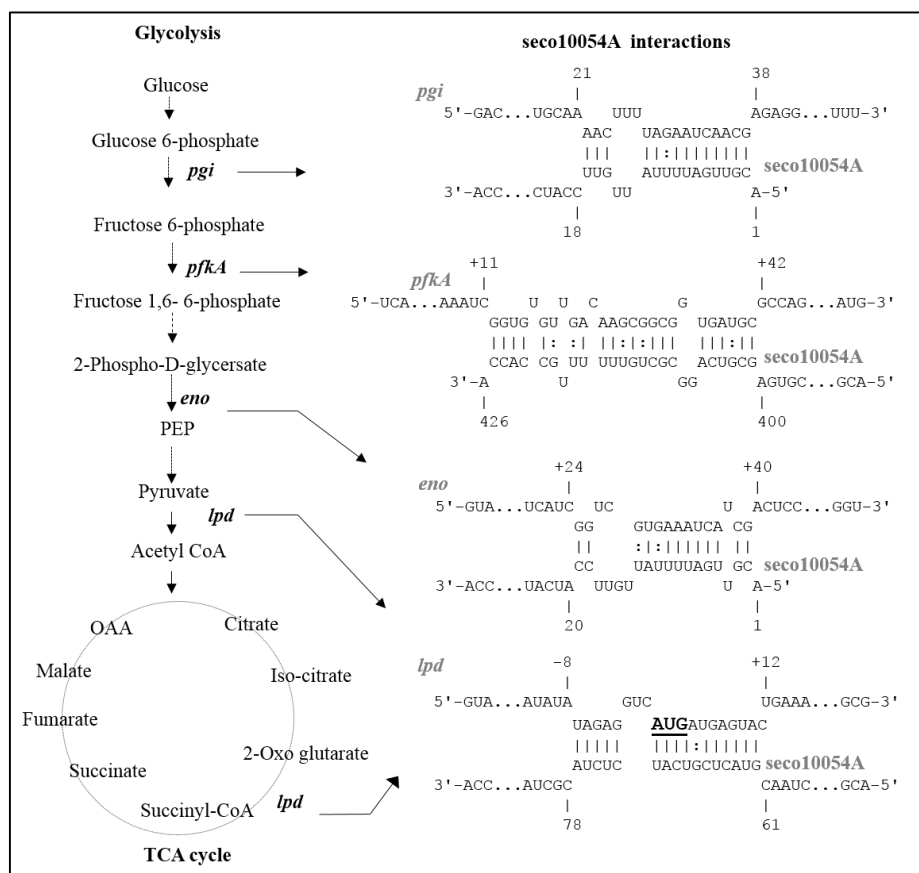
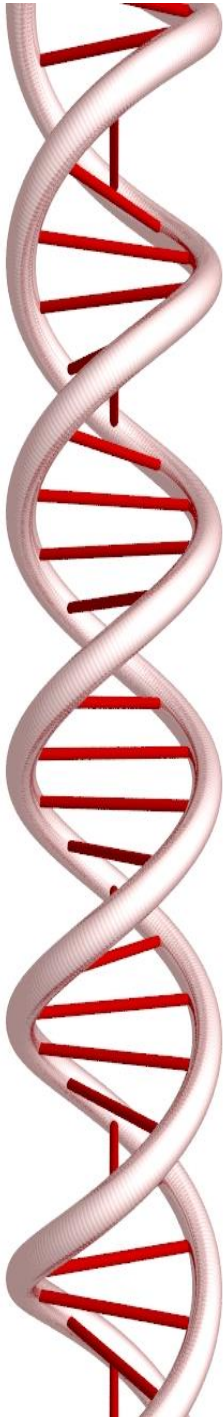


Fig. 3. 11. The seco10054A / target mRNA interactions: Arrows indicate the enzymatic steps of glycolysis and TCA cycle that are predicted to be down-regulated due to translation inhibition of sRNA, seco10054A target mRNAs.

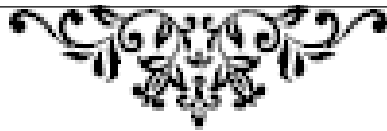


CHAPTER- II

Results and Discussion



Elucidation of interactions between Orf306
responsive sRNAs and target mRNAs



4. 1. Introduction

In the preceding chapter, online tools were used to predict Orf306 dependent expression patterns and interactions of sRNAs with target mRNAs. The predictions have clearly suggested existence of interactions between *seco10054A* and mRNAs coding for phosphoglucose isomerase (*pgi*), phosphofructokinase II (*pfkB*), phosphofructokinase (*pfkA*), phosphoglycerate kinase (*pgk*), glucose PTS permease - PtsG subunit (*ptsG*), glyceraldehyde 3-phosphate dehydrogenase (*gapA*), enolase (*eno*) and lipoamide dehydrogenase (*lpd*). The prediction fits to the proposed hypothesis, which states that sRNA, *seco10054A* up-regulated in the presence of organophosphate degradation (*opd*) island borne Orf306 inhibits translation of mRNA coding for key glycolytic and TCA pathway enzymes (Fig. 3. 02). Similarly, the predicted interaction between the sRNA, *seco11136A*, down-regulated in the presence of Orf306 and mRNA coding for transcription factor HcaR supports the kind of catabolic shift noticed in *E. coli* MG1655 (pSDP10) cells (Fig. 3. 02) (Chakka et al., 2015). The present chapter is devoted to gain experimental proof to the *in silico* predictions made in the first chapter.

The online predictions have shown existence of interactions between sRNA, *seco10054A* and mRNAs codes for phospho glucose isomerase (*pgi*), phosphofructokinase II (*pfkB*), phosphofructokinase (*pfkA*), phosphoglycerate kinase (*pgk*), glucose PTS permease - PtsG subunit (*ptsG*), glyceraldehyde 3-phosphate dehydrogenase (*gapA*), enolase (*eno*) and lipoamide dehydrogenase (*lpd*). If the predicted interactions are carefully examined the interactions found between sRNA, *seco10054A* and mRNA coding for *Lpd* are alone found at the ribosomal binding site (RBS). The rest of the interactions are found outside of RBS. There are published reports for suggesting the existence of inhibitory effect if interactions are found even outside RBS (Sharma et al., 2007). Therefore, in this study, a genetic screen was developed to validate predicted interactions and to eliminate false positives generated through online tools.

4. 2. Genetic Screen

The genetic screen developed in this study contains two compatible plasmid vectors. One of them is a low-copy number, broad host range expression vector, pMMB206 (Morales et al., 1991). The second plasmid belongs to promoter probe vector, pRS552 which facilitates the cloning of a target mRNA coding gene in frame of the *lacZ* reporter gene (Simons et al., 1987). Such cloning drives transcription of a chimeric mRNA, which contains 5' region of

target mRNA and *lacZ* specific sequence at the 3' end. Since these two plasmids are of different incompatibility groups, they can be co-transformed into single *E. coli* host. If the sRNA coding gene is cloned into an expression vector, its expression can be regulated by placing it under the control of an inducible promoter. If the expression of sRNA is induced, the expressed sRNA targets constitutively expressed chimeric mRNA and inhibits its translation leading to the generation of *lacZ* negative phenotype (Fig. 4. 01). The genetic screen can successfully be used to observe the expected translation inhibition by performing both quantitative and qualitative assays. The qualitative screening can be done on X-gal plates and the quantitative assays can be performed by measuring β -galactosidase activity (Miller 1972).

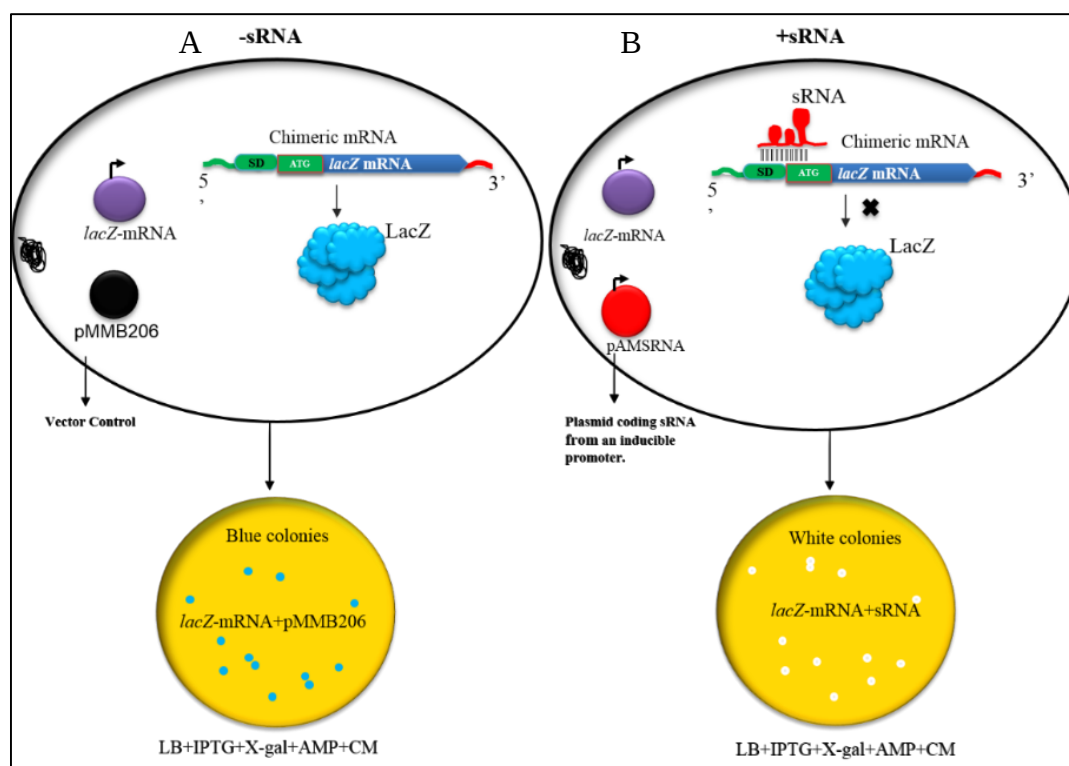


Fig. 4. 01. Genetic screen used to validate sRNA-target mRNA interactions.

4. 3. Construction of target mRNA gene translational fusions

As stated before the sRNA, *seco10054A* targets mRNAs transcribed from *pgi*, *pfkB*, *pfkA*, *pgk*, *ptsG*, *gapA*, *eno* and *lpd*. While validating the *seco10054A* and mRNAs coding for aforementioned genes. Translational fusions of these genes were generated based on the seed region (-30 to +20) and by cloning the 5' regions of the genes including their native promoters and predicted seed regions. While cloning the respective genes (*eno*, *pfkA*, *gltA*, *icd*, and *lpd*) they were amplified by designing a forward primer taking the sequences found

upstream of the promoter motif. The reverse primer was essentially designed by taking the sequence found about 50 bp downstream of the predicted interacting regions. Both forward and reverse primers were appended with restriction sites *EcoRI* and *BamHI* respectively. While generating restriction sites necessary care was taken so as to facilitate in-frame fusion of cloned fragment with the *lacZ* reporter gene of the promoter probe vector pRS552. Such cloning facilitates generation of a chimeric mRNA (Fig. 4. 01). The primers sequences used and the size of target mRNA coding gene amplicons are used in Table 4. A.

Table 4. A: The details of the primer sequences and amplicon size to be generated by PCR amplification.

Name of the Primer	Primer sequences	Primers used for	Size of the Amplicon in Base pairs
AMENOTFF AMENOTFR	GGCCTGGAATTCGATCAGTTGGTCGAGATC ACGACGGGATCCGCAGCTGCCATACCGAC	Used to amplify 5' region of <i>eno</i> gene to construct <i>eno-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	350bp
AMPFKATFF AMPFKATFR	ACTAGTTGAATTCAGTATAAAAGAGAGCCA GTGATCAGGATCCAGATAGCCGTCATAAATACC	Used to amplify 5' region of <i>pfkA</i> gene to construct <i>pfkA-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	300bp
AMGLTATFF AMGLTATFR	GGCCTGGAATTCGTGTCCGGAGACCTGGC ACGACGGGATCCGTAACCGAGAGTACGGA	Used to amplify 5' region of <i>gltA</i> gene to construct <i>gltA-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	610bp
AMICDTFF AMICDTFR	GGCCTGGAATTCAGAGATTATGAATTGCCG ACGACGGGATCCTCAGGAACGTTGAGTTTGCC	Used to amplify 5' region of <i>icd</i> gene to construct <i>icd-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	260bp

AMLPDTF AMLPDTR	GGCCTGGAATTCATGAATGTATTG ACGACGGGATCCGGGCCTGCCCCAAGTAC	Used to amplify 5' region of <i>lpd</i> gene to construct <i>lpd-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	330bp
AMFADRTF AMFADTR	GGCCTGGAATTCGGCCATCGCCAGAGTGAA ACTAAGGGATCCTTGAATGGTCAACCAGC	Used to amplify 5' region of <i>fadR</i> gene to construct <i>fadR-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	430bp
AMACEKTF AMACEKTR	GCCTGGAATTCCTGACCGGCTCCACTGAA CGACGGGATCCACCGGAGGTCACTTCGAGG AATC	Used to amplify 5' region of <i>aceK</i> gene to construct <i>aceK-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	350bp
AMHCARTF AMHCARTR	GGCCTGGAATCTTTGGGATCTGGCTTTCG ACTAAGGGATCCACAGTTTTCAAGATCGC	Used to amplify 5' region of <i>hcaR</i> gene to construct <i>hcaR-lacZ</i> translational fusions by cloning the amplicon as <i>EcoRI</i> and <i>BamHI</i> fragment in promoter probe vector pRS552.	220bp

As shown in Table 4. A and Fig. 4. 02, Panel IV- A, a 350 bp amplicon containing 5' region of *eno* gene was amplified. This amplicon was digested with *EcoRI* and *BamHI* and cloned in promoter probe vector pRS552 (Fig. 4.02 Panel II) digested with *EcoRI* and *BamHI*. The resulting recombinant plasmid is designated as pAMENOTF (Fig. 4. 02, Panel III-A). The generated plasmid was digested again with *EcoRI* and *BamHI* to check the existence of insert in promoter test vector pRS552. The plasmid which has insert was alone retained and used for further studies.

Similarly, PCR amplification has generated amplicons containing 5' regions of *pfkA* (300 bp) (Fig. 4. 02, Panel IV-B), *gltA* (610 bp) (Fig. 4. 02, Panel IV-C), *icd* (260bp) (Fig. 4. 02 Panel IV-D), and *lpd* (330bp) (Fig. 4. 02 Panel IV-E) genes. All these amplicons were independently cloned in pRS552 vector as *EcoRI* and *BamHI* fragments to generate respective translational fusions and the generated translational fusions were designated as pAMPFKATF

(*pfkA-lacZ*), pAMGLTATF (*gltA-lacZ*), pAMICDTF (*icd-lacZ*), and pAMLPTDF (*lpd-lacZ*) respectively (Fig. 4. 02, Panel III- B, C, D and E). All these translational fusions were independently digested and analyzed on agarose gels to confirm existence of insert in right orientation (Fig. 4. 02, Panel IV).

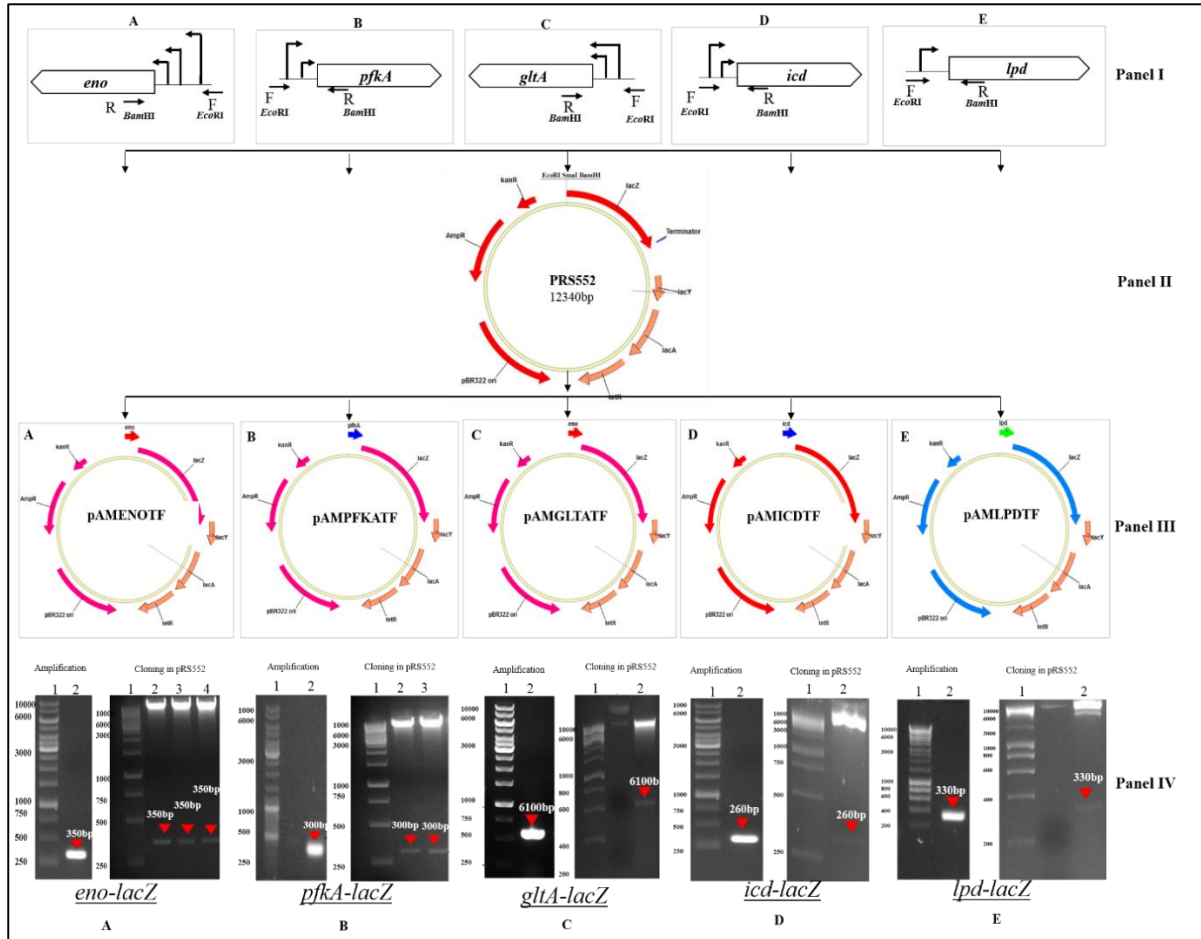


Fig. 4. 02: Cloning strategy used to generate promoter *lacZ* fusions. Panel I indicates sRNA, *seco10054A* target mRNA coding genes. The bent arrow indicate promoter motif and direction of transcription orientation. The primers used for generating promoter amplicons were shown with arrows. Panel II shows promoter test vector, pRS552. The generated translational fusions are shown in Panel III. PCR amplicons and clone confirmations were shown in Panel IV. Panel IV A, B, C, D and E represent agarose gels showing existence of correct insert in promoter fusions, *eno-lacZ*, *pfkA-lacZ*, *gltA-lacZ*, *icd-lacZ* and *lpd-lacZ*.

4. 4. Expression of sRNA, *seco10054A*

The gene *seco10054A* gene coding for the sRNA, *seco10054A* was amplified from genomic DNA of *E. coli* MG1655 using forward (5'-AGCTACGAATTCTCGCGTGGTCAGCATTGA-3') and reverse primers (5'-GCTAAGGATCCTGCCAGTTACCACCGCCAGTTT-3') appended with restriction site

EcoRI and *BamHI* (Fig. 4. 03, Panel A). The amplicon with a size of 1.1 kB was cloned in expression vector pMMB206 (Fig. 4. 03, Panel D). The resulting recombinant plasmid is designated as pAM10054A (Fig. 4. 03, Panel C). The detailed cloning strategy and confirmation of clone is shown in Fig. 4. 03 Panel D.

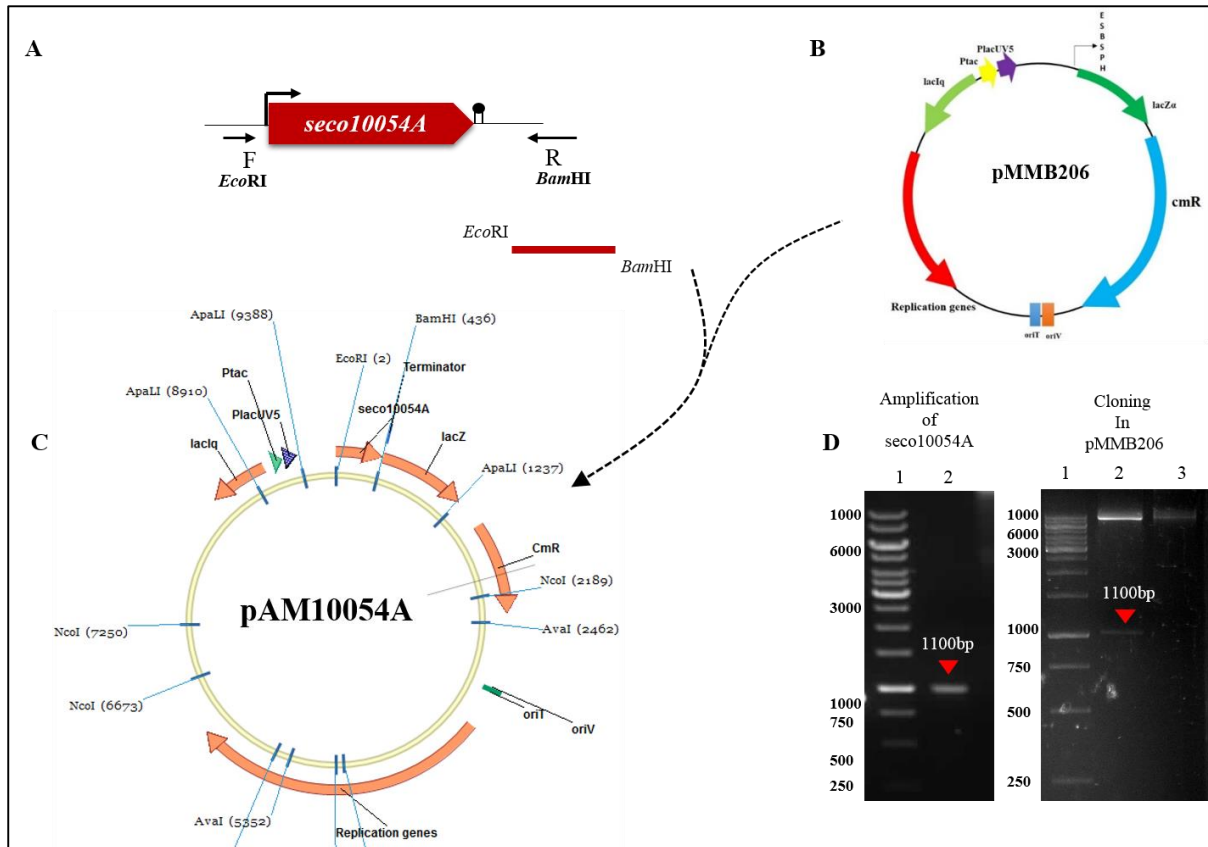


Fig. 4. 03. Cloning of sRNA, *seco10054A* coding gene *seco10054A* in expression vector pMMB206. Panel A indicates intergenic region found between *lptD* and *djlA* genes. Primers used to amplify *seco10054A* were indicated with arrow. The filled in circle shows transcription terminator. The bent arrow indicate promoter sequence and direction of transcription orientation. Panel B and C shows expression vector pMMB206 and expression plasmid pAM10054A, coding for sRNA, *seco10054A*. Panel D shows agarose gel picture containing existence of insert in vector pMMB206.

4. 5. Induction of sRNA, *seco10054A* expression

The *E. coli* MG1655 (pAM10054A) cells were grown to mid-log phase and expression of *seco10054A* was induced by adding IPTG (1mM). After inducing the expression of sRNA, *seco10054A* for 1hr, the total RNA was extracted and analyzed on 7% Acrylamide-urea gel as indicated in Material and Methods section. Subsequently, the gels were bolted on hybond-N⁺ membrane and existence of sRNA, *seco10054A* was detected by probing the blot with 5' end labeled oligonucleotide, AMSECO10054APROBE, (5' CTGACCGTTTGTACGCGCAACGTTACCGATGAT-3') complementary to the sRNA,

seco10054A sequence. The total RNA isolated from uninduced *E. coli* MG1655 and *E. coli* MG1655 (pAM10054A) cultures with expression plasmid pMMB206 served as controls. As shown in Fig. 4. 04, Panel B, a clear signal at predicted size of 430nt is observed only in induced cultures and such signal was not seen in uninduced cultures. These results clearly indicate expression of sRNA, seco10054A from vector specific *tac* promoter. The plasmid (pAM10054A) was also sequenced to ascertain the existence of sRNA, *seco10054A* gene in pAM10054A. After obtaining expression, sequence identity the plasmid pAM10054A was used for further studies.

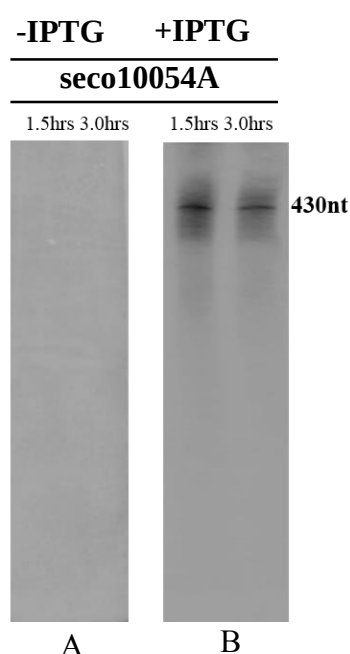


Fig. 4. 04. Northern blot analysis to detect pAM10054A coded sRNA, seco10054A. Panel A represents 10% acrylamide-urea gel used to analyse total RNA isolation from un-induced (pAM10054A) cells, Panel B represents similar gel used to analyse total RNA extracted from induced cultures.

4. 6. Generation of *lac* negative knockouts of *E. coli* MG1655

Since *E. coli* MG1655 is a model strain, the genetic screen was developed using MG1655. The *E. coli* MG1655 is a *lac* positive strain. Since the promoter test vector has *lacZ* as a reporter gene. Usage of wild-type *E. coli* MG1655 is not feasible. Therefore the *lac* negative strain of *E. coli* MG1655 cells were used for developing the proposed genetic screen (gifted by Mnajula Reddy, CCMB, Hyderabad).

4. 7. Co-transformation

The *lac* negative *E. coli* MG1655 cells were transformed with plasmids pAM10054A coding sRNA, seco10054A. The *E. coli* MG1655 (pAM10054A) cells were then grown in chloramphenicol containing LB medium and made them competent to transform with respective gene translational fusions. The co-transformed *E. coli* MG1655 cells (seco10054A coding plasmid, pAM10054A and a plasmid coding for target gene translational fusion) were maintained in culture medium containing chloramphenicol and ampicillin until further use.

4. 8. Screening of sRNA-mRNA interactions

4. 8. 1. Qualitative Assay

While performing qualitative assays the ampicillin, chloramphenicol, X-gal and IPTG-containing LB plate was prepared and the plate was divided into four sectors. Sector A is used to grow *E. coli* MG1655 Δlac cells containing target gene *lacZ* fusion and expression vector pMMB206. Sector B is used to grow cells containing promoter test vector and expression vector pMMB206. Sector C is used to grow cells having sRNA, seco10054A coding plasmid, pAM10054A and target gene *lacZ* fusions. Sector D contained cells having expressing plasmid pAM10054A coding for sRNA, seco10054A and promoter probe vector pRS552. A pictorial representation of plating pattern is shown in Fig. 4. 05

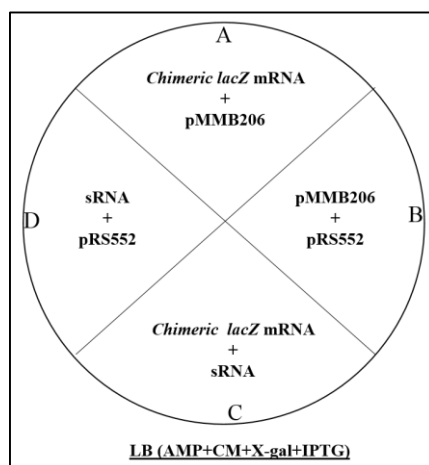


Fig. 4. 05. Genetic screen: The LB plate shown plasmids in various combinations.

As shown in Fig. 4. 06, cells containing *eno-lacZ* fusion (Fig. 4. 06, Panel I), *pfkA-lacZ* fusion (Fig. 4. 06, Panel II), *gltA-lacZ* fusions (Fig. 4. 06, Panel III), *icd-lacZ* fusions (Fig. 4.

06, Panel VII) and *lpd-lacZ* fusions (Fig. 4. 06 Panel VIII), with appropriate controls were plated on LB plates containing ampicillin, chloramphenicol, X-gal and IPTG. After allowing them to grow for 12hrs the sectors were carefully observed for the appearance of blue colored colonies. If sRNA *seco10054A* is inhibiting translation of chimeric mRNA coding for respective translational fusion, the sector C of the plate devoted for plating *lac* negative *E. coli* MG1655 cells having target gene *lacZ* fusions and sRNA, *seco10054A* coding expression plasmid, pAM10054A should give pale colored colonies. If it is not inhibiting the translational of chimeric mRNA, the cells plated in Sector C should appear blue in color. Interestingly, the cells containing *eno-lacZ* fusion (Fig. 4. 06, Panel I, Sector C), *pfkA-lacZ* fusion (Fig. 4. 06, Panel II, Sector C), *gltA-lacZ* fusions (Fig. 4. 06, Panel III, Sector C) and *icd-lacZ* fusions (Fig. 4. 06, Panel VII, Sector C) showed no *lac* negative phenotype suggesting the induced sRNA, *seco10054A* has no inhibitory effect on translation of the chimeric mRNA transcribed from these *lacZ* fusions. This data contradicts online tool predicted interactions described in Chapter-I and microarray generated transcription profile of *E. coli* MG1655 (pSDP10) cells (Chakka et al., 2015). However, cells grown in sector C of plate devoted to growing *lpd-lacZ* translational fusions appeared pale in color (Fig. 4. 06, Panel VIII, Sector C). In the absence of sRNA, *seco10054A*, the *lpd-lacZ* fusions containing cells appeared blue in color (Fig. 4. 06, Panel VIII, Sector A). This clearly suggests that the chimeric *lpd-lacZ* mRNA is made and translated successfully from the *lpd-lacZ* translational fusions. Therefore, the cells containing *lpd-lacZ* fusions gave *lac* positive phenotype. This data clearly suggests that sRNA, *seco1005A* has an inhibitory effect on the translation of *lpd-mRNA*. In order to gain quantitative data the *E. coli* MG1655 cells having respective *lacZ* fusions (*eno-lacZ*, *pfkA-lacZ*, *gltA-lacZ*, *icd-lacZ* and *lpd-lacZ*) and expression plasmid coding sRNA, *seco10054A* were grown and used for assaying β -galactosidase activity. As shown in Fig. 4. 06, Panel IV, V, VI and IX, the presence of sRNA, *seco10054A* showed no influence on amounts of β -galactosidase produced by respective *lacZ* fusions (*eno-lacZ*, *pfkA-lacZ*, *gltA-lacZ* and *icd-lacZ*). Re-confirming the results obtained in the qualitative assay, the *lac* negative *E. coli* MG1655 containing *lpd-lacZ* fusions produced insignificant amounts of β -galactosidase in the presence of sRNA, *seco10054A*. However, the cells produced 450 miller units of β -galactosidase in the absence of sRNA, *seco10054A*. In its presence, it has produced less than 50 miller units of β -galactosidase activity. Encouraged by these results further experiments were designed to gain secondary evidence on sRNA, *seco10054A* and *lpd-lacZ* fusions.

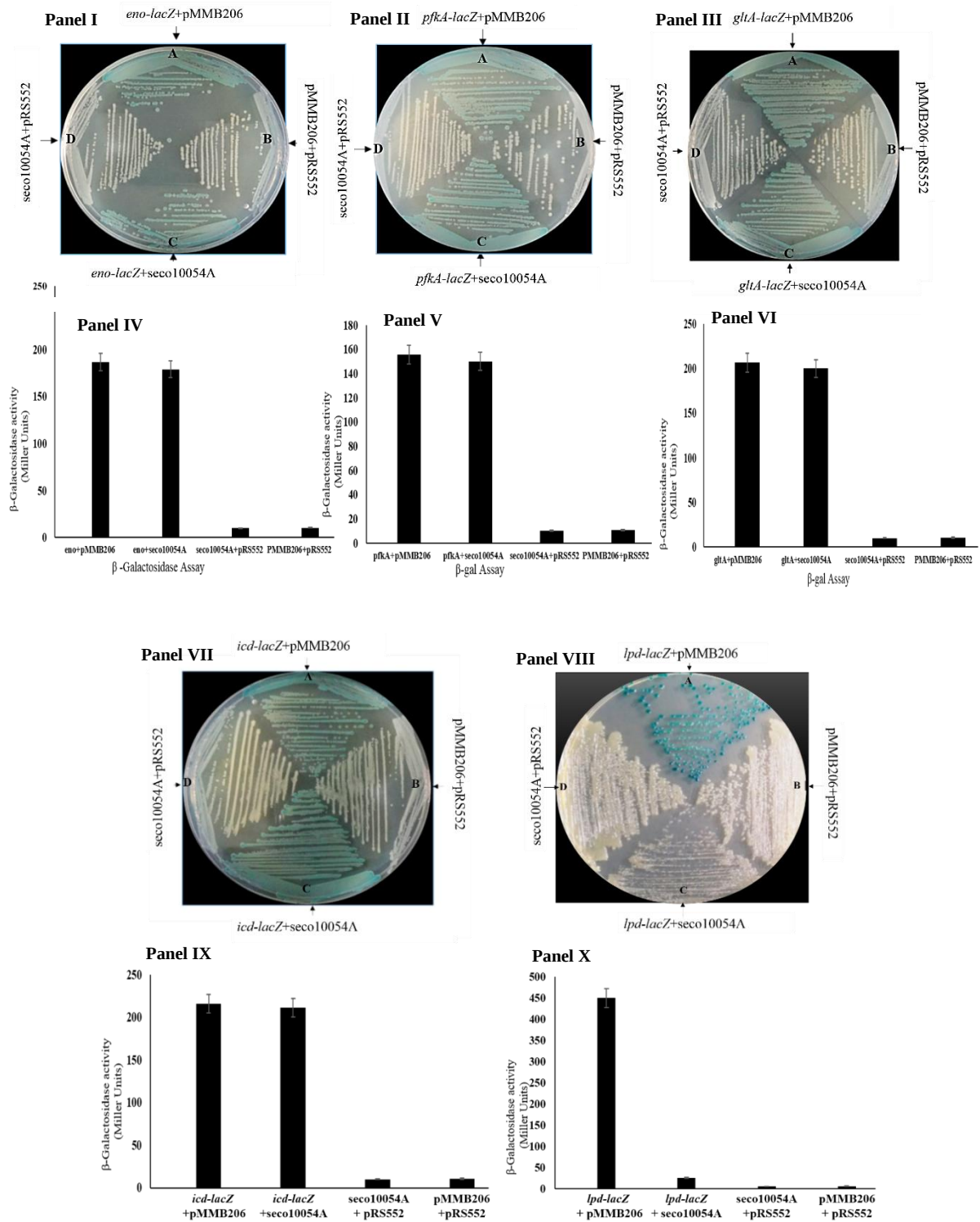


Fig. 4.06. Validation of sRNA, seco10054A and target mRNA interactions. Panels I, II, III, VII and VIII represent qualitative assays indicating sRNA, seco10054A interactions with chimeric mRNAs coded by *eno-lacZ*, *pfkA-lacZ*, *gltA-lacZ*, *icd-lacZ* and *lpd-lacZ* fusions respectively. The *lac* negative *E. coli* MG1655 cells having pMMB206 + respective *lacZ* fusion (Sector A), pMMB206 and promoter probe vector, pRS552 (Sectors B), the sRNA, seco910054A coding, pAM10054A + respective *lacZ* fusion (Sectors C) and pAM10054A + pRS552 (Sectors D) were grown on IPTG, X-gal plates. Quantitative assays performed for cultures grown under similar conditions were shown in Panels IV (*eno-lacZ*), V (*pfkA-lacZ*), VI (*gltA-lacZ*), IX (*icd-lacZ*) and X (*lpd-lacZ*). The sRNA, seco10054A inhibited only chimeric mRNA transcribed from *lpd-lacZ* fusions (Sector C, Panel VII).

4. 8. 2. Structural complementarity between the sRNA, seco10054A and *lpd* mRNA

Initially, the secondary structures were generated for both *lpd* mRNA and sRNA, seco10054A by using mfold (<http://unafold.rna.albany.edu/?q=mfold/rna-folding-form>). These structures were then examined to assess existence of structural complementarity between these two molecules, especially at the interaction sites. As shown in Fig. 4. 07, the predicted structures have shown the existence of structural complementarity between these two molecules at the predicted interaction sites.

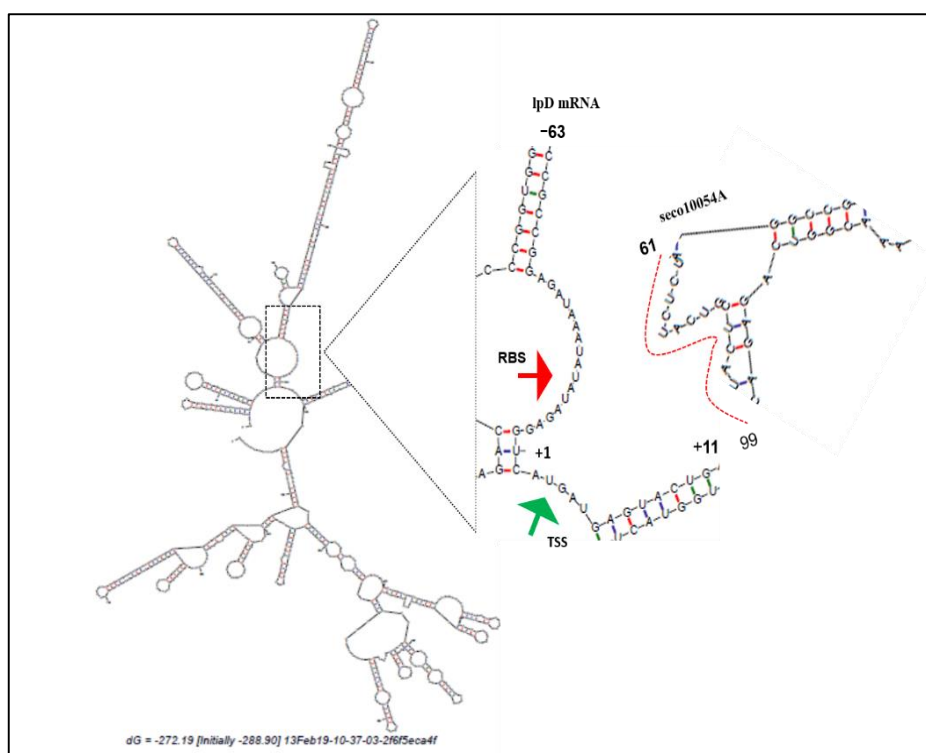


Fig. 4. 07. Structural complementarity between *lpd* mRNA and seco10054A. The online tool mfold predicted secondary structures and predicted complementarity at the seed region is shown with arrow mark.

4. 9. Quantification of transcripts

In order to gain further evidence of translational inhibition, the *lpd* transcripts were quantified in the presence of sRNA, seco10054A. While achieving this objective the *E. coli* MG1655 cells transformed with plasmid pAM10054A and expression of seco10054A was induced as stated in material and methods section. After inducing the cells for 1.5 hrs the total RNA was isolated and used to quantify *lpd* transcript expression levels by performing qPCR. Supporting the results obtained through the genetic screen, *lpd* specific transcript got significantly reduced in the presence of sRNA, seco10054A (Fig. 4. 08). These two

independent experiments have provided strong evidence to suggest *seco10054A* dependent translational inhibition of *lpd* mRNA expression levels.

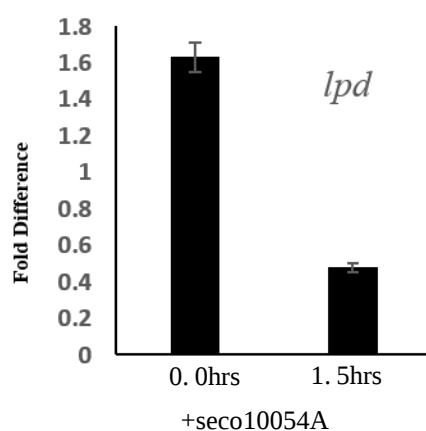


Fig. 4. 08. Relative expression levels of *lpd* mRNA.

4. 10 Generation of sRNA, *seco10054A*^{Δ61-78} variant

In the light of afore-described results, the predicted interactions between the *lpd* mRNA and sRNA, *seco10054A* were examined. The interactions between these two molecules exist from -63 to +11 nucleotides (where +1 is 'A' of translational start codon ATG) of *lpd* mRNA. This seed region is divided into three positions, seed1, seed2 and seed3 (Fig. 4. 09 Panel A). The seed1 exists exactly at translational initiation site of *lpd* mRNA (Fig. 4. 09 Panel A). Assuming that the sequence of sRNA, *seco10054A* contributing for the seed1 region is influencing translational inhibition of *lpd* mRNA, this region was deleted by performing overlapping PCR (Fig. 4. 09, Panel B). The deletion derivative of sRNA, *seco10054A*^{Δ61-78} gene, then cloned in expression vector, pMMB206 (Fig. 4. 09, Panel D) to express *seco10054A*^{Δ61-78} which does not contain the sequence predicted to be interacting at RBS of *lpd* mRNA and repeated experiments to know the influence of sRNA, *seco10054A* variant on the translation of *lpd* mRNA.

The *E. coli* MG1655 cells having *lpd-lacZ* translational fusions were then co-transformed independently with plasmids coding sRNA, *seco10054A* and its deletion derivative, *seco10054A*^{Δ61-78}. As shown in Fig. 4. 10, Panel I, Sector E, the sRNA,

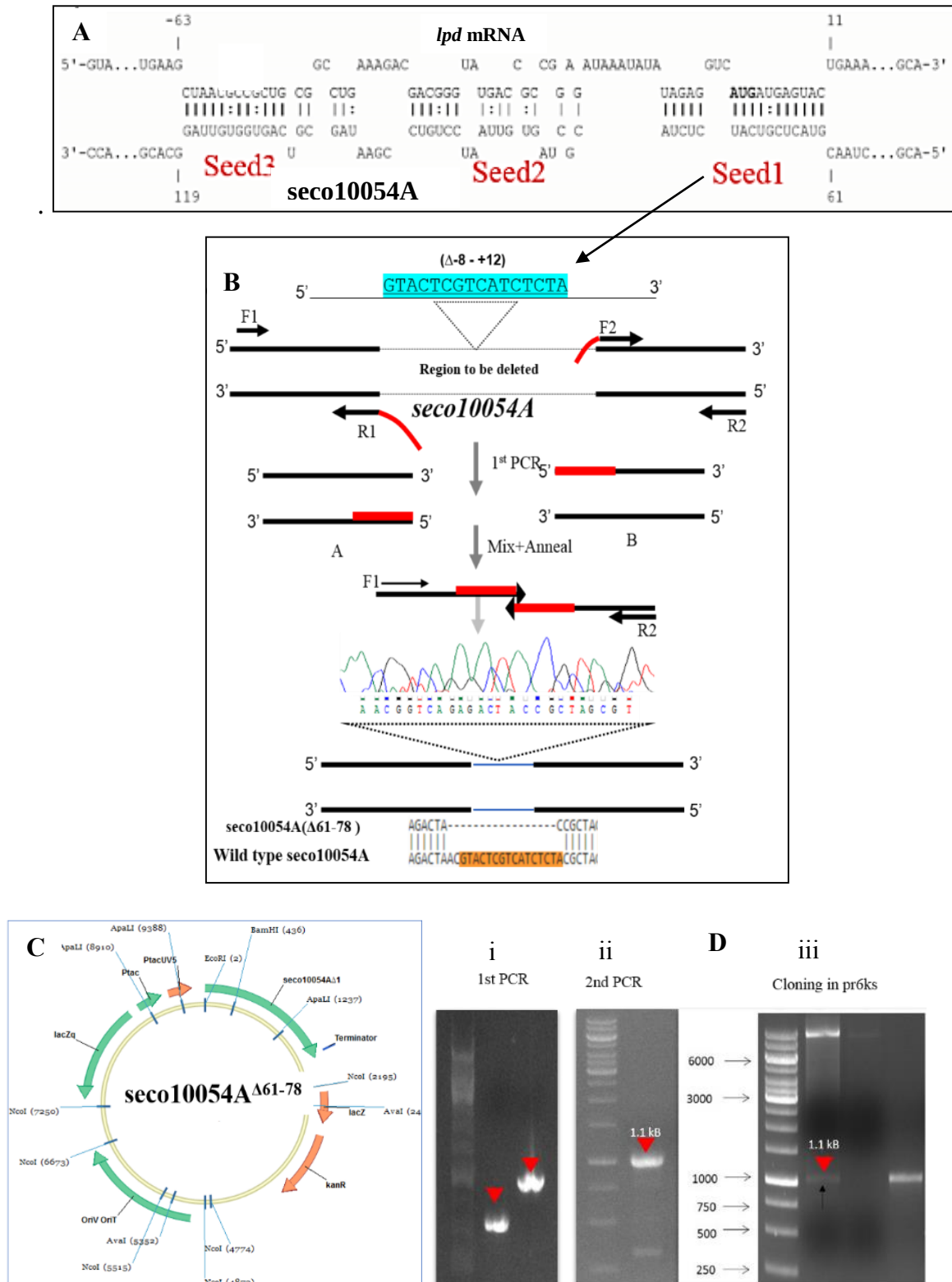


Fig. 4. 09. Construction of *seco10054A* ^{Δ 61-78} variant. Panel A indicates the sequence complementarity between the sRNA, *seco10054A* and *lpd* mRNA. Strategy used to delete predicted seed region of sRNA, *seco10054A* by performing overlapping PCR is shown in (Panel B). Panel C indicates the pAM10054A ^{Δ 61-78} plasmid map. Agarose gel showing amplification of fragment A and B flanking to the predicted seed region of *seco10054A* gene is shown in Panel D-i). The sRNA variant, *seco10054A* ^{Δ 61-78} generated by performing overlapping PCR is shown in Panel D-ii). Cloning of *seco10054A* ^{Δ 61-78} in expression vector pMMB206 is shown in Panel D-iii).

seco10054A variant coding seco10054A^{Δ61-78} lacking seed1 region failed to inhibit the translational of *lpd-lacZ* chimeric mRNA (Fig. 4. 10, Panel A, Sector E). The quantitative assays were performed to re-confirm the results obtained through qualitative assays. As expected the cells expressing wild-type sRNA, seco10054A cells produced insignificant amounts of β-galactosidase. However, cells expressing sRNA variant seco10054A^{Δ61-78} produced similar amounts of β-galactosidase that are produced by the *lpd-lacZ* fusion in the absence of sRNA, seco10054A (Fig. 4. 10, Panel II).

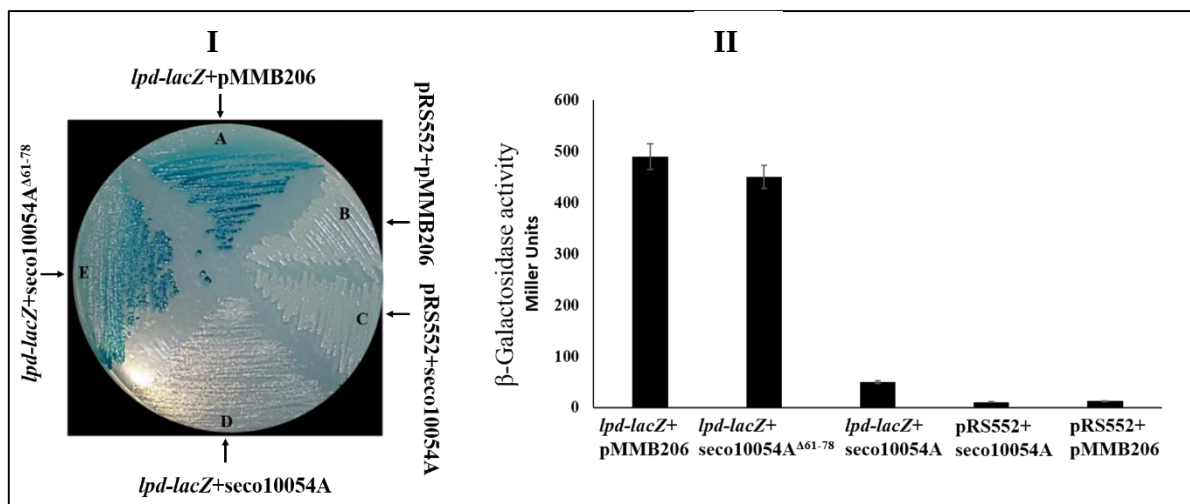


Fig. 4. 10. The sRNA, seco10054A^{Δ61-78} and *lpd* mRNA interactions. Panel A represents qualitative β-galactosidase assay of *lac* negative *E. coli* MG1655 cells containing *lpd-lacZ* translational fusions with plasmids coding sRNA, seco10054A (Sector D), or its variant seco10054A^{Δ61-78} (Sector E), or expression plasmid pMMB206 (Sector A) are plated in X-gal, IPTG plates. The sectors B and C indicate control cultures containing vectors (pRS552 + pMMB206) and promoter test vector, pRS552 + pAM10054A coding sRNA, seco10054A. Quantitative assays performed for similarly grown cultures are shown in panel II.

4. 11. Expression of sRNA, seco11136A

The sRNA, seco11136A coding gene *seco11136A* (C0293) was amplified from genomic DNA of *E. coli* MG1655 using forward (AGATAGAATTCGGTACCTGCCGTTTTGTATCAACTAC) and reverse primers (CATGACTGCAGGTACTAAAGATGTATGTGAAGG) appended with *EcoRI*, *KpnI* and *PstI*. The 140 bp amplicon containing *seco11136A* gene was digested with *EcoRI*, and *PstI* further cloned in expression vector pR6KS (Swetha, 2015) digested with similar enzymes. The cloning strategy facilitated placing of *seco11136A* gene under the control of inducible *tac* promoter. The resulting plasmid designated as pAM11136A (Fig. 4. 2. 1 Panel C). Since the amplicon is of 140 bp in size it was difficult to visualize on an agarose gel. The presence of insert was,

a good seed region was detected between sRNA, *seco11136A* and mRNA coding for isocitrate dehydrogenase kinase/phosphatase (*aceK*). In this section of chapter II, the developed two-plasmid genetic screen was used to validate the predicted interactions. Before validation of predicted interactions between the sRNA, *seco11136A* and target mRNAs, the genes coding these mRNAs were amplified along with their native promoters, translational signals and predicted seed regions. These amplifications were then translationally fused to the *lacZ* reporter gene of promoter test vector pRS552. The detailed cloning strategies, clone confirmations are shown in the Fig. 4. 2. 2. The generated translational fusions were designated as pAMACEKTF, pAMHCARTFF, and pAMFADRTF. The plasmid pAMACEKTF contains the promoter along with the putative seed region coding for *aceK* gene. Similarly, the plasmids pAMFADRTF and pAMHCAR contain promoters and respective seed regions of the genes *hacR* and *fadR* respectively. These translational fusions were then independently transformed into *lac* negative strains of *E. coli* MG1655. The respective translational fusions pAMACEKTF, pAMHCARTFF, and pAMFADRTF code for chimeric mRNA.

The plasmid pAM11136A codes for sRNA, *seco11136A* from an inducible *tac* promoter. If the expression of sRNA, *seco11136A* inhibits the translation of chimeric mRNA, the cells containing translational fusion display *lac* negative phenotype. If such inhibition is not seen, the cells show *lac* positive background. As shown in Fig. 4. 2. 3, the LB containing kanamycin, ampicillin, X-gal and IPTG (1mM) were divided into 4 sectors. Sectors were designated as A, B, C and D. Sector C is used to plate cells containing translational fusion and pAM11136A coding sRNA, *seco11136A* from an inducible promoter. The sector B is used to plate cells having expression vector pMMB206 and promoter test vector pRS552. Sector A is used to plate cells having *lacZ* fusions and expression vector pMMB206. Sector D is used to plate the cells containing pAM11136A coding for sRNA, *seco11136A* and promoter test vector pRS552. Sectors-A, B and D serve as controls whereas sector-C indicates sRNA, *seco11136A* and mRNA interaction by displaying *lac* phenotype. As shown in Fig. 4. 2. 3, Panel A sector-C, the sRNA, *seco11136A* failed to inhibit chimeric *aceK-lacZ* mRNA. Therefore, the cells displayed *lac* positive phenotype. The quantitative β -galactosidase assays also revealed no influence of sRNA, *seco11136A* on the translation of *aceK-lacZ* mRNA. Both in presence and absence of sRNA the *aceK-lacZ* fusions showed identical β -galactosidase activity (Fig. 4. 2. 3. Panel I, Sector C and Sector A). Interestingly, the sRNA, *seco11136A* inhibited translation of

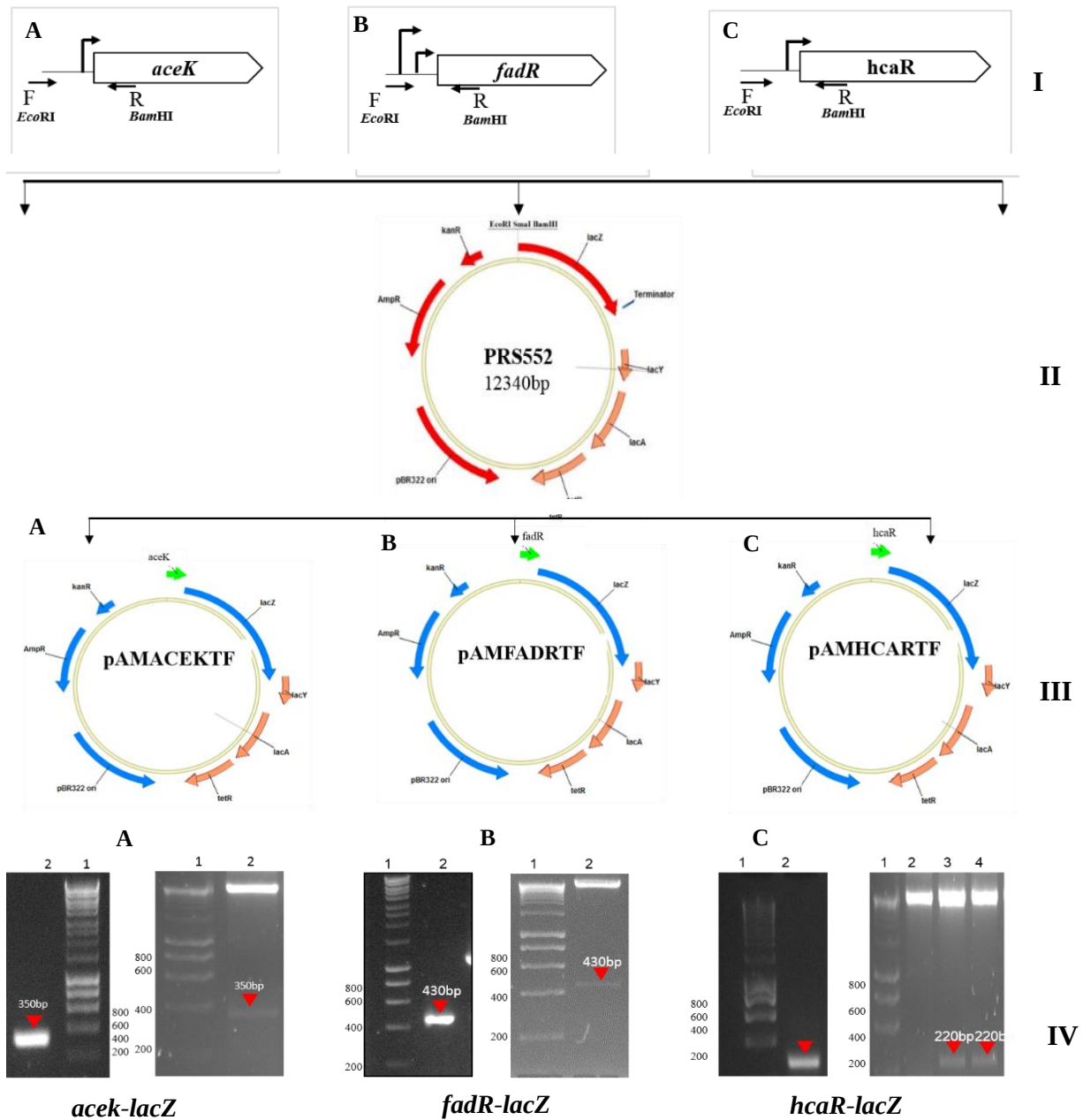


Fig. 4. 2. 2. Cloning strategy used to generate translation fusions of *aceK-lacZ*, *fadR-lacZ* and *hcaR-lacZ* fusions. The primers used for generating promoter amplicons were shown with arrows. The bent arrow indicate promoter motif and direction of transcription orientation. The primers used for generating amplicons used to generate translational fusions were shown with arrows (Panel I). Panel II shows promoter test vector pRS552. The generated translational fusions pAMACEKTF, pAMFADRTF and pAMHCARTF are shown in Panel III. PCR amplicons and clone confirmations were shown in Panel IV. Panel IV A, B, and C represent agarose gels showing existence of translation fusions of *aceK-lacZ*, *fadR-lacZ* and *hcaR-lacZ*, fusions.

chimeric mRNAs transcribed from *fadR-lacZ* and *hcaR-lacZ* fusions. As shown in sector-C, the *E. coli* MG1655 cells containing *fadR-lacZ* fusion (Fig. 4. 2. 3, Panel II, Sector-C) and *hcaR-lacZ* translation fusions showed *lac* negative phenotype in the presence of sRNA,

seco11136A (Fig. 4. 2. 3, Panel III, Sector-C). The quantitative assays showed no significant β -galactosidase activity in *E. coli* MG16655 cells (pAM11136A) having either pAMFADRTF or and pAMHCARTF (Fig. 4. 2. 3, Panel II, Sector-C and Fig. 4. 2. 3, Panel III, Sector-C). This results clearly suggests the involvement of sRNA, seco11136A in translational inhibition of mRNAs coding for transcription factors, FadR and HacR.

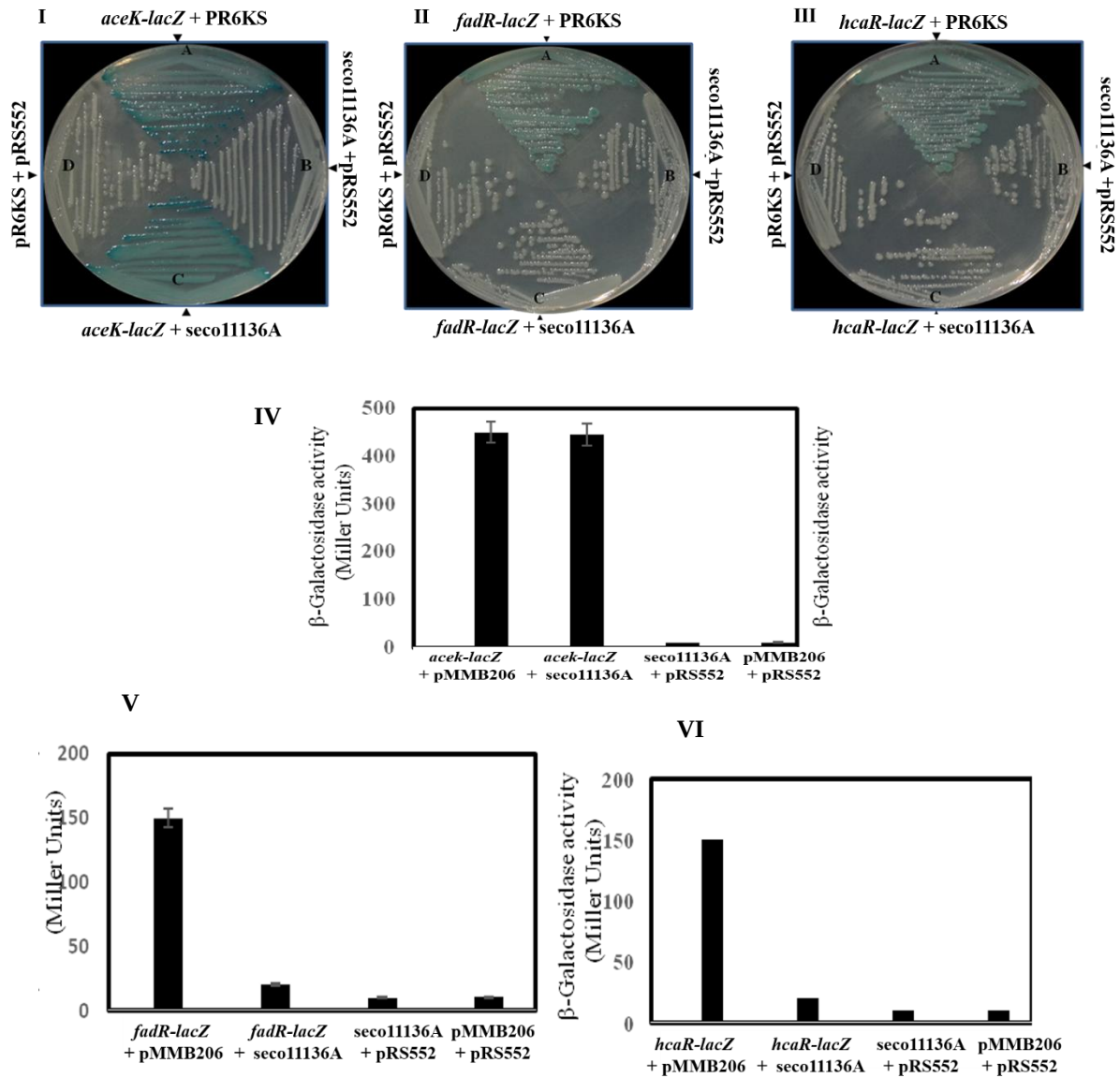


Fig. 4. 2. 3. Validation of sRNA, seco11136A and target mRNA interactions. Panels I, II and III, represent qualitative assays indicating sRNA, seco11136A interactions with chimeric mRNAs coded by *aceK-lacZ*, *fadR-lacZ* and *hcaR-lacZ* fusions respectively. The *lac* negative *E. coli* MG1655 cells having expression vector, pR6KS + respective *lacZ* fusion (Sector A), pAM11136A + pRS552 (Sectors B), the sRNA, seco11136A coding plasmid, pAM11136A + respective *lacZ* fusion (Sectors C) and , pR6KS + promoter probe vector, pRS552 (Sectors D) were grown on IPTG, X-gal plates. Quantitative assays performed for cultures grown under similar conditions were shown in Panels IV (*aceK-lacZ*) , V (*fadR-lacZ*), and VI (*hcaR-lacZ*).

4. 13. Quantification of transcripts

In order to gain secondary evidence on translational inhibition, the transcripts *aceK*, *fadR*, *hcaR* were quantified in the presence of sRNA, *seco11136A*. While achieving this objective the *E. coli* MG1655 cells were transformed with plasmid pAM11136A and expression of *seco11136A* was induced as stated in material and methods section. After inducing the cells for 1.5 hrs the total RNA was isolated and used to quantify *aceK*, *fadR*, *hcaR* specific mRNAs by performing qPCR. Supporting the results obtained through genetic screen *fadR* and *hcaR* specific transcripts got significantly reduced in the presence of sRNA, *seco11136A* (Fig. 4. 2. 4. Panel B and Panel C). Interestingly, such reduction was not seen in the case of *aceK* transcript (Fig. 4. 2. 4, Panel A). These independent experiments provide strong evidence to suggest *seco11136A* dependent translational inhibition of *fadR* and *hcaR* mRNAs.

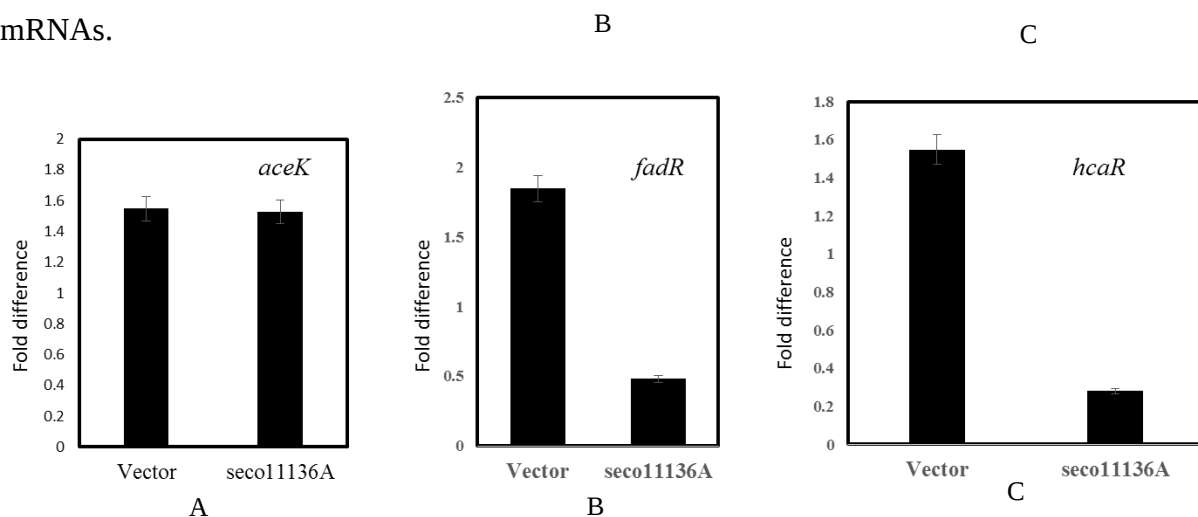


Fig. 4. 2. 4. Relative fold difference of *aceK*, *fadR*, *hcaR* specific transcripts in presence of sRNA, *seco11136A*.

4. 14. Discussion

The chapter I devoted to describe identification of target genes for sRNAs, *seco10054A* and *seco11136A*. In the second chapter experimental evidence is gained to show the interactions between sRNA, *seco10054A* and mRNA coding for Lpd protein. Likewise, the interactions between sRNA, *seco11136A* and mRNA coding for *fadR*, *hcaR* were also established.

The *opd* island borne esterase induced metabolic shift in *E. coli* MG1655 (pSDP10) cells clearly indicates down-regulation of glycolysis and TCA cycle and up-regulation of alternate catabolic operons coding for propionate and phenylacetate degradation (Chakka et al 2015). In Orf306 positive background up-regulation of methylcitrate cycle involved in the metabolism of single carbon compound is also clearly visible (Chakka et al., 2015). In the light of these observations, an attempt was made to draw correlations between the observed metabolic shift in *E. coli* MG1655 (pSDP10) cells and sRNA-mediated translation inhibition of mRNAs coding for *lpd*, *fadR*, and *hcaR*.

The lipamide dehydrogenase (Lpd) is part of the multienzyme complex, pyruvate dehydrogenase (PDHc). The PDHc enzyme complex consists three enzymes such as E1, E2, and E3. The E1 is of pyruvate dehydrogenase and it removes CO₂ from pyruvate. The enzyme E2, also known dehydrolipoate acyltransferase transfers the remainder of the substrate, acetyl group, to coenzyme A. The third enzyme dihydrolipoate dehydrogenase (Lpd) transfers the retained H₂ to NAD⁺ (Wilkinson & Williams 1981). Each PDHc complex contains multiple copies of three enzyme subunits. The E1 and E2 are present 24 copies each, whereas 12 copies of E3 is present in *E. coli* pyruvate dehydrogenase (PDHc) (Izard et al., 1999). The enzyme E3, Lipamide dehydrogenase (Lpd) is, therefore, an important component in PDHc. The complex structure and reactions catalysed by PDHc, as adopted from <http://flipper.diff.org/apprulesitems/pathways/7618> is shown in Fig. 4. 2. 5, Panel A.

Interestingly, the Lpd is also part of α -ketoglutarate dehydrogenase (KGDH) which is involved in the conversion of α -ketoglutarate to succinyl Co-A. The multi-enzyme complex α -ketoglutarate dehydrogenase catalyzed reactions, as represented by Pettit and Reed showed in Fig. 4. 2. 5, Panel B (Pettit & Reed 1967). Both PDHc and α -ketoglutarate dehydrogenase are regulatory enzymes. The PDHc is a link between glycolysis and TCA cycle. Similarly, α -ketoglutarate dehydrogenase is the rate limiting enzyme and regulates TCA cycle.

Interestingly, the transcriptome profile of *E. coli* MG1655 (pSDP10) cells indicated an identical shift in carbon catabolic pathways (Chakka et al., 2015). This kind of catabolic shift identified in *lpd* null mutants of *E. coli* coincides with the catabolic shift identified in *E. coli* MG1655 (pSDP10) cells. If this catabolic shift is seen together with the Orf306 dependent induction of sRNA, *seco10054A* and its translational inhibition of *lpd*, it clearly points towards sRNA, *seco10054A* dependent translation inhibition of *lpd* mRNA as the primary cause for the observed catabolic shift in *E. coli* MG1655 (pSDP10) cells. Since, the Lpd is also a critical component in α -ketoglutarate dehydrogenase and glycine cleavage multi-enzyme (GCV) systems, its influence on TCA cycle and one carbon compound catabolism can easily be predicted. Though the relation between the presence of Orf306 and induction of sRNA, *seco10054A* is unknown, the role of sRNA in the observed Orf306 dependent catabolic shift is clearly established in this study. The chapter-III devoted to understanding the relationship between the Orf306 presence and sRNA, *seco10054A* induction.

The second sRNA, *seco11136A* down-regulated in Orf306 positive background inhibited translation of transcription factors FadR and HcaR. Though online tool has suggested its interactions with mRNA coding for *aceK*, the genetic screen generated in this study failed to establish such interaction between the sRNA, *seco11136A* and mRNA coding for *aceK*. In the light of this background, an in-depth analysis was made between Orf306 dependent repression of sRNA, *seco11136A* and PNP dependent growth of *E. coli* MG1655 (pSDP10) cells.

The transcription factor HcaR is a dual regulator, it activates *hca* (hydroxycinnamic acid) operon and *mhp* (m-hydroxyphenylpropionate) operons involved in degradation of phenyl propionate and hydroxyl phenyl propionate (Turlin et al., 2001). The *hca* operon contains five open reading frames. *hcaE*, *F*, *C*, *B* and *D*. The HcaE, F, C and D form an enzyme complex called phenyl propionate dehydrogenase and the *hcaB* codes for 2, 3-dihydroxy-2, 3-dihydrophenylpropionate dehydrogenase. The *hca* operon encoded enzymes that converts to phenyl propionate to dihydroxy phenyl propionate, which is further degraded by *mhp* operon (Fig. 4. 2. 6).

When the Orf306 dependent PNP growth and catabolic shift in *E. coli* MG1655 (pSDP10) cells are critically examined, the up-regulation of *hca* and *mhp* operons are the only

up-regulated operons that contributed for aromatic compound degradation. In the light of this observation an attempt was made to establish link between *hca* operon and PNP catabolism. Our laboratory has placed *hca* operon under the influence of an inducible promoter and tested if the induced *hca* operon has a role in PNP degradation. Interestingly, research done in our lab showed *hca* induction dependent PNP degradation (Chakka et al., 2015). It is therefore proposed that PNP is a fortuitous substrate for enzymes encoded by *hca* and *mhp* operon. The data presented in this chapter clearly suggest Orf306 dependent repression of sRNA, *seco11136A*. Under glucose-grown conditions, the sRNA, *seco11136A* is present in sufficient concentrations. However, a five-fold decrease is seen in Orf306 positive background (Fig. 4. 2. 7). Such repression indicates avoidance of sRNA, *seco11136A* dependent translational inhibition of mRNA coding for *hcaR*, leading to the production of transcription factor HcaR. The presence of HcaR induces *hca* operon which is shown to contribute for degradation of PNP. The relationship between Orf306 presence and repression of sRNA, *seco11136A* is unknown. The work presented in next chapter provides experimental evidence to understand the relationship between Orf306 and *seco11136A* expression. The influence of FadR inhibition by sRNA, *seco11136A* is not properly understood in this present context. Further studies are needed to explain its physiological relevance in *E. coli* MG1655 (pSDP10) cells.

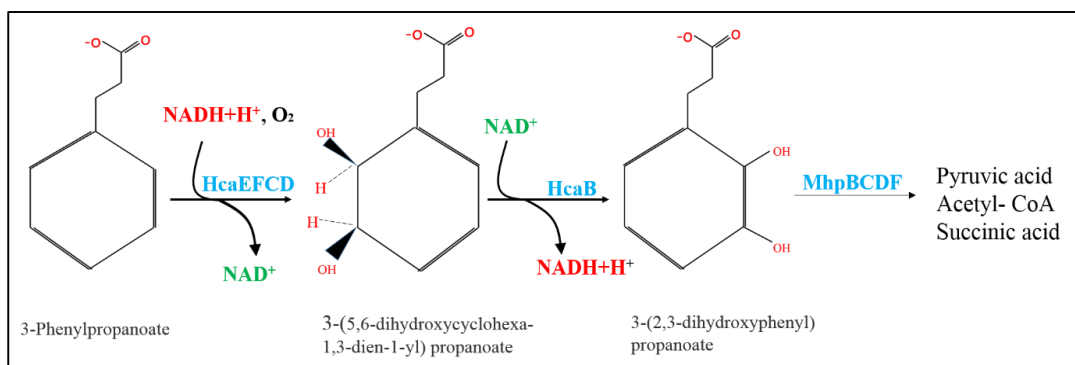
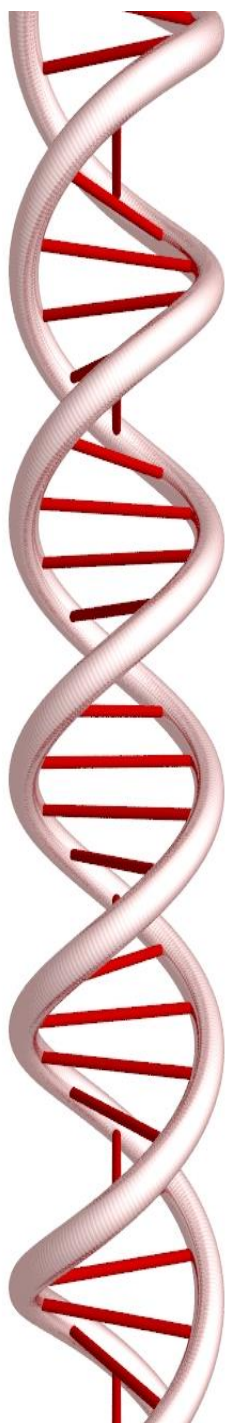


Fig. 4. 2. 6: Reactions catalyzed by *hca* operon coded enzymes. The HcaEFFCD complex converts 3- phenyl propionate to 3-(5, 6-dihydroxycyclohexa-1, 3-dien-1-yl) propanoate. The HcaB converts 3-(5, 6-dihydroxycyclohexa-1, 3-dien-1-yl) propanoate to 3-(2,3-dihydroxyphenyl) propanoate.



Fig. 4. 2. 7: Heat map showing Orf306 dependent repression of sRNA, *seco11136A* expression.



CHAPTER-III

Results and Discussion



Regulation of expression of
seco10054A and *seco11136A* genes



5. 1. Introduction

As stated before, the present study is designed to unravel the molecular basis of the observed Orf306 dependent catabolic shift in *E. coli* MG1655 (pSDP10) cells. Initial leads were collected from the heat maps generated from the transcriptome data generated from the *E. coli* MG1655 (pSDP10) cells. The heat maps have clearly indicated down-regulation of glycolysis and TCA cycle (Chakka et al., 2015). Further, the heat maps have also shown up-regulation of alternate carbon catabolic pathways involved in propionate, phenyl propionate degradation (Chakka et al., 2015). The present study has employed both *in silico* and *in vivo* approaches to understanding the molecular basis of the observed metabolic shift. The results described in chapter -I and chapter –II, clearly indicate, i) up-regulation of sRNA, seco10054A, ii) down-regulation of sRNA, seco11136A, iii) interactions between sRNA, seco10054A and mRNA coding for Lpd, subunit of pyruvate dehydrogenase complex iv) interactions between sRNA, seco11136A and mRNAs coding for transcription factors FadR and HcaR. Detailed discussions were also presented in first and second chapters inferring the roles of sRNAs, seco10054A, and seco11136A in observed shift in catabolic pathways in *E. coli* MG1655 (pSDP10) cells. The present chapter is devoted to gain experimental evidence required for explaining the link between Orf306 and expression of these novel sRNAs. As part of this detailed investigation, promoter fusions were generated to monitor the expression pattern of these two novel sRNAs in presence and absence of Orf306.

5. 2. Construction of *seco10054A-lacZ* fusion

The *seco10054A* gene is identified in the intragenic regions found between *lptD* and *djlA* genes (Fig. 5. 01). As shown in Fig. 5. 01, the 5' region of *seco10054A* overlaps with the 5' region of the *lptD* gene, due to the opposite transcription of these two genes. In fact, the predicted σ^{70} dependent promoter of *seco10054A* gene is found in the coding region of *lptD* gene (Fig. 5. 01). While generating amplicon containing putative promoter region of *seco10054A* gene, the forward (AM10054A220F, GCTACGTCGAATTCGGTCATAGCTTGGCACGCCCA) and reverse (AM10054A220R, AGCTACCTGCAAGTTTGTCACGCGCAACGTTAC) primers were generated to amplify putative promoter region (Fig. 5. 01). As the forward and reverse primers were appended restriction sites *Eco*RI and *Pst*I respectively, the 220 bp amplicon was then digested with *Eco*RI and *Pst*I and ligated to the promoter test vector pMP220 (Spaink et al., 1987). The resulted recombinant

plasmid was designated as pAM10054A220. In plasmid pAM10054A220, the putative promoter of *seco10054A* gene is transcriptionally fused to the promoter less *lacZ* gene of the promoter test vector pMP220 (Spaink et al., 1987). The resultant *seco10054A-lacZ* fusion is designated as pAM10054A220. The existence of the cloned promoter sequence of *seco10054A* gene in plasmid pAM10054A220 is confirmed by determining its sequence.

5. 3. Construction of *seco11136A-lacZ* fusion

The *seco11136A* gene is found in the intergenic region identified between *icd* and *ymfD* genes. In fact, the presence of *seco11136A* gene is reported by Rosenow group (Tjaden et al., 2002). The authors have named sRNA, *seco11136A* as C0293. In the present study, it is designated as *seco10054A* following established convention (Lamichhane et al., 2013). Though the existence of sRNA *seco11136A* is known its physiological role is not known. In this study, as described in chapter-II, its involvement in translational inhibition of mRNAs coding for transcriptional factors FadR and HcaR is experimentally proved. The *seco11136A* contains a putative σ^{70} dependent promoter (Fig. 5. 02). Interestingly, the putative promoter sequence has also shown sequence motifs that shows similarity to NarL, PdhR, and Crp binding sites were identified (Fig. 5. 02). In order to validate its function and to establish the reasons for observed Orf306 dependent repression, the *seco11136A-lacZ* transcriptional fusions were generated by fusing the predicted promoter motif to the *lacZ* gene of pMP220 (Fig. 5. 02). The promoter region of *seco11136A* gene was amplified using forward (AM11136A220F, 5' AGATAGAATTTCCTTCATTGATTAAGACATCCCCA 3') and reverse (AM11136AA220R, 5' CATGACTGCAGTTGCACCCAATGCCTGTTGGCT 3') appended with *EcoRI* and *PstI* (Fig. 5. 02). The generated amplicon was then digested with *EcoRI* and *PstI* and cloned in promoter test vector pMP220. The resulting plasmid designated as pAM11136A220 was sequenced to confirm the existence of insert in the right orientation.

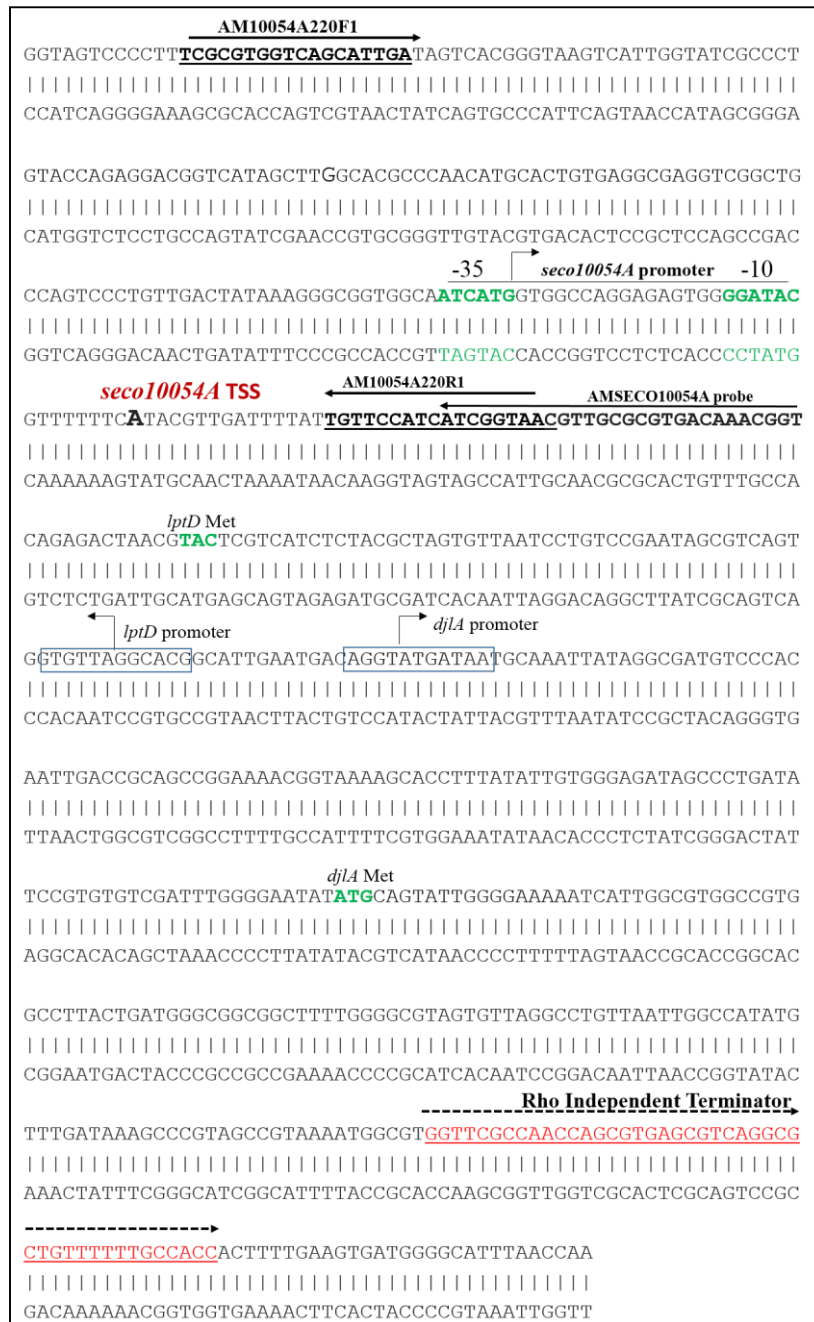


Fig. 5. 01. Nucleotide sequence of intergenic region found between *lptD* and *djlA*: Putative Rho-independent terminator sequences of *seco10054A* gene is underlined. The -10 and -35 promoter sequence motifs are shown with bold case. The regions used to design primers for amplification of promoter sequence of *seco10054A* are underlined and shown with bold case. The promoter sequences of *lptD* and *djlA* are shown with open boxes. The translation initiation codons of *lptD* and *djlA* are shown with bold case. The oligo AMSECO10054A probe used as probe for detection of sRNA, *seco10054A* is indicated with solid arrow.

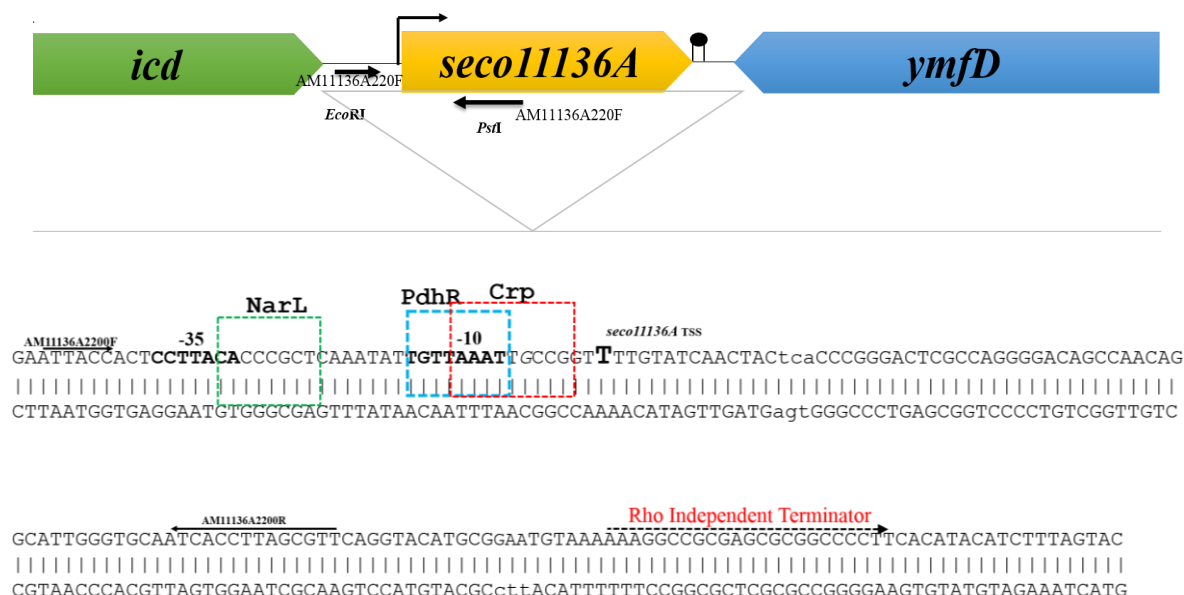


Fig. 5. 02. The sequence of intergenic region found between *icd* and *ymfD* genes. The putative promoter motif of *seco11136A* is shown with bold case. The rho-independent terminator is shown with dotted line. The putative transcription factor binding sites are shown with open boxes. The primers used to amplify promoter region of *seco11136A* are shown with solid arrows.

5. 4. Two-plasmid assay

A two plasmid assay was developed to assess the influence of Orf306 on the expression of *seco10054A* and *seco11136A* genes. As shown in Fig. 5. 03, the *lac* negative *E. coli* MG1655 was transformed with plasmid pSDP10, which codes Orf306 from an inducible promoter. In *E. coli* MG1655 (pSDP10) cells, the Orf306 is induced from an IPTG-inducible *tac* promoter (Chakka et al., 2015). These cells were then used to transform with either transcriptional fusion pAM10054A220 or with pAM11136A220 (Fig. 5. 03). It is clearly known from microarray data the Orf306 contributes to the transcriptional activation of *seco10054A* gene. Further, it is also shown to represses transcription of *seco11136A* gene. In the two plasmid assay, if Orf306 induces *seco10054A* promoter, the *E. coli* MG1655 (pSDP10 plus pAM10054A220) should display *lac* positive phenotype. In the absence of Orf306 i. e. *E. coli* MG1655 (pAM10054A220 plus pMMB206) should generate *lac* negative phenotype (Fig. 5. 03).

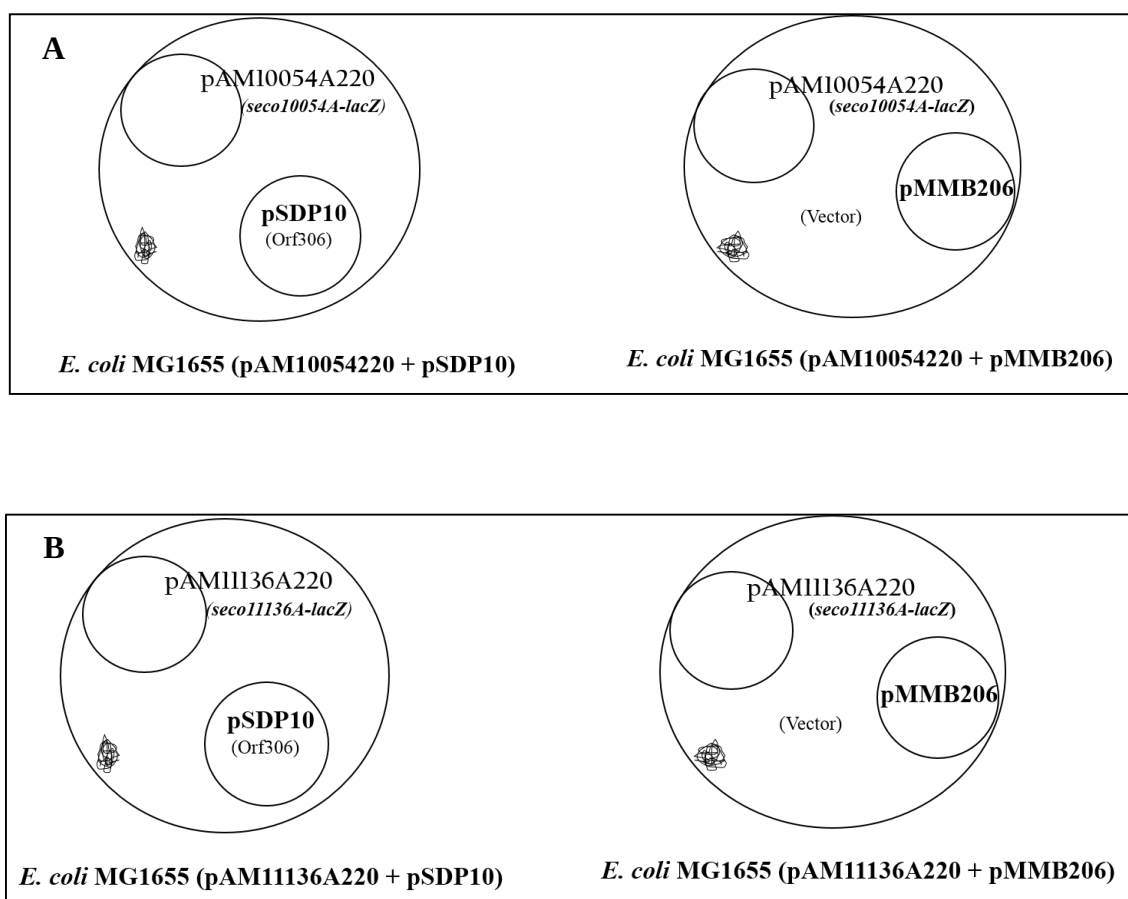


Fig. 5. 03: Two plasmid assay: Panel A shows the *lac* negative *E. coli* MG1655 (pAM10054A220) cells either with expression plasmid pSDP10 coding Orf306^{6HIS} or expression vector pMMB206. Similar cells having *seco11136A-lacZ* fusions, pAM11136A220 either with expression plasmid pSDP10 coding for Orf306^{6HIS} or vector pMMB206 is shown in panel B. The Orf306 dependent activation or repression of respective *lacZ* fusions can be observed on X-gal plates

Supporting the proposed hypothesis the cells containing *seco10054A-lacZ* fusion in a combination of expression plasmid pSDP10 coding Orf306 generated *lac* positive phenotype (Fig. 5. 04, Panel A, Sector-i). However, the *seco10054A-lacZ* fusion containing *E. coli* MG1655 cells in the presence of expression vector pMMB206 gave *lac* negative phenotype. (Fig. 5. 04, Panel A- Sector-ii). The quantitative assay also supported the Orf306 dependent activation of *seco10054A-lacZ* fusion (Fig. 5. 04, Panel B). The *seco10054A-lacZ* fusion containing *E. coli* MG1655 (pSDP10) cells gave 450 units of β -galactosidase activity in the presence of Orf306, as against 50 units of β -galactosidase activity in its absence. This clearly suggested that the predicted promoter upstream of *seco10054A* gene is functional and is positively regulated by Orf306.

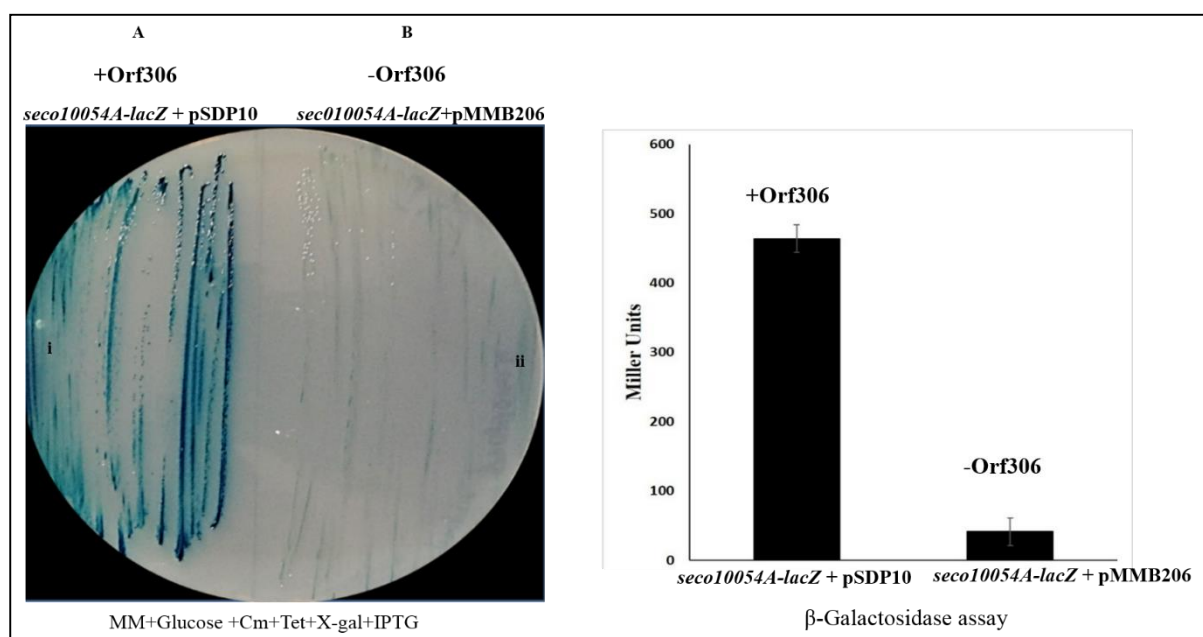


Fig. 5. 04: The Orf306 dependent expression of *seco10054A*: The *seco10054A-lacZ* fusions (pAM10054A220) containing *E. coli* MG1655 cells having either pSDP10, coding for Orf306^{6HIS} or expression vector, pMMB206 are plated in minimal medium, X-gal, IPTG, tetracycline and chloramphenicol. In presence of Orf306, the *lac* negative *E. coli* MG1655 (pAM10054A220) cells have shown *lac* positive phenotype (Panel A, Sector I). Quantification of β-galactosidase in these cells are shown in Panel B.

5. 4. 1. The Orf306 repress *seco11136A* gene expression

The two plasmid assay was also used to assess the influence of Orf306 on the expression of *seco11136A* gene. Here the *E. coli* MG1655 (pAM11136A220) cells having *seco11136A-lacZ* fusion were transformed with either pSDP10, plasmid coding Orf306^{6HIS} or with expression vector pMMB206. After co-transformation, these cells were independently plated on a minimal medium plate containing glucose, chloramphenicol, tetracycline, X-gal and IPTG. If Orf306 repress the expression of *seco11136A* gene it should display *lac* negative phenotype in the presence of Orf306 and in its absence it should show *lac* positive phenotype (Fig. 5. 05, Panel A). Both quantitative and qualitative assays supported the proposed hypothesis (Fig. 5. 05). About 600 units of β-galactosidase activity was observed in the absence of Orf306. In its presence, the *seco11136A-lacZ* fusion containing cells have made just 170 units of β-galactosidase activity (Fig. 5. 05, Panel B) providing a strong indication on Orf306 dependent repression of *seco11136A* gene.

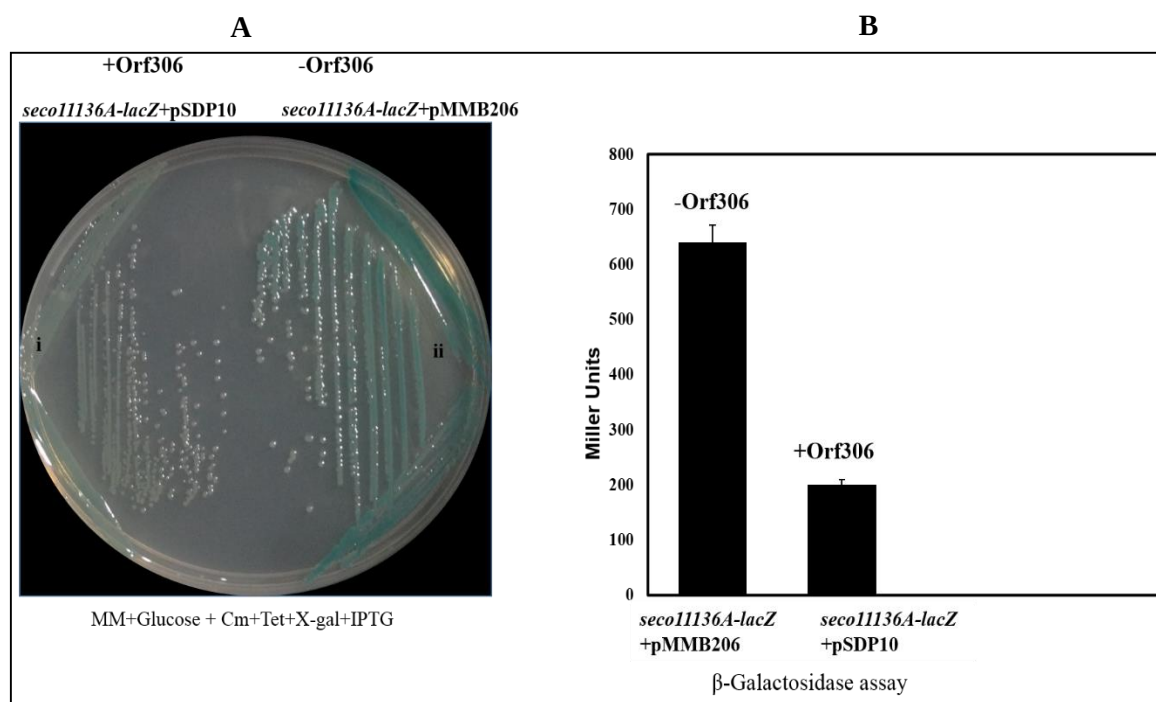


Fig. 5. 05: The Orf306 dependent repression of *seco11136A* gene: The *E. coli* MG1655 having *seco11136A-lacZ* fusions (pAM11136A220), with either expression plasmid pSDP10 coding Orf306^{6HIS} or expression vector pMMB206 were plated on minimal media glucose plates having tetracycline, chloramphenicol, X-gal and IPTG. The cells are showing *lac* negative phenotype in presence of Orf306 (Panel A, Sector-I). The quantitative β -galactosidase in these cells are shown in Panel B.

5. 5. Regulation of *seco10054A* gene

After observing Orf306 dependent activation of *seco10054A* gene, further experiments were performed to detect Orf306 dependent induction of *seco10054A* gene. Initially, the total RNA extracted from *E. coli* MG1655 (pSDP10) cells were analyzed on denaturing polyacrylamide gels. The gels were then blotted and probed with end labelled oligo, (AMSECO10054APROBE, 5'-CTGACCGTTTGTACGCGCAACGTTACCGATGAT-3'), complementary to sRNA, *seco10054A* (Fig. 5. 01). A clear signal specific to sRNA, *seco10054A* was seen after 1.5 hrs of Orf306 induction and the intensity of sRNA, *seco10054A* signal has increased with the increase of induction time (Fig. 5. 06, Panel I-B, Lanes; 3, 4, 5, and 6). Such sRNA, *seco10054A* specific signal was not seen in the lanes loaded with total RNA isolated from *E. coli* MG1655 (pMMB206) cells (Fig. 5. 06 Panel I-C, Lanes; 7, 8, 9, and 10). This is indeed a clear confirmation to show Orf306 dependent induction of sRNA, *seco10054A*.

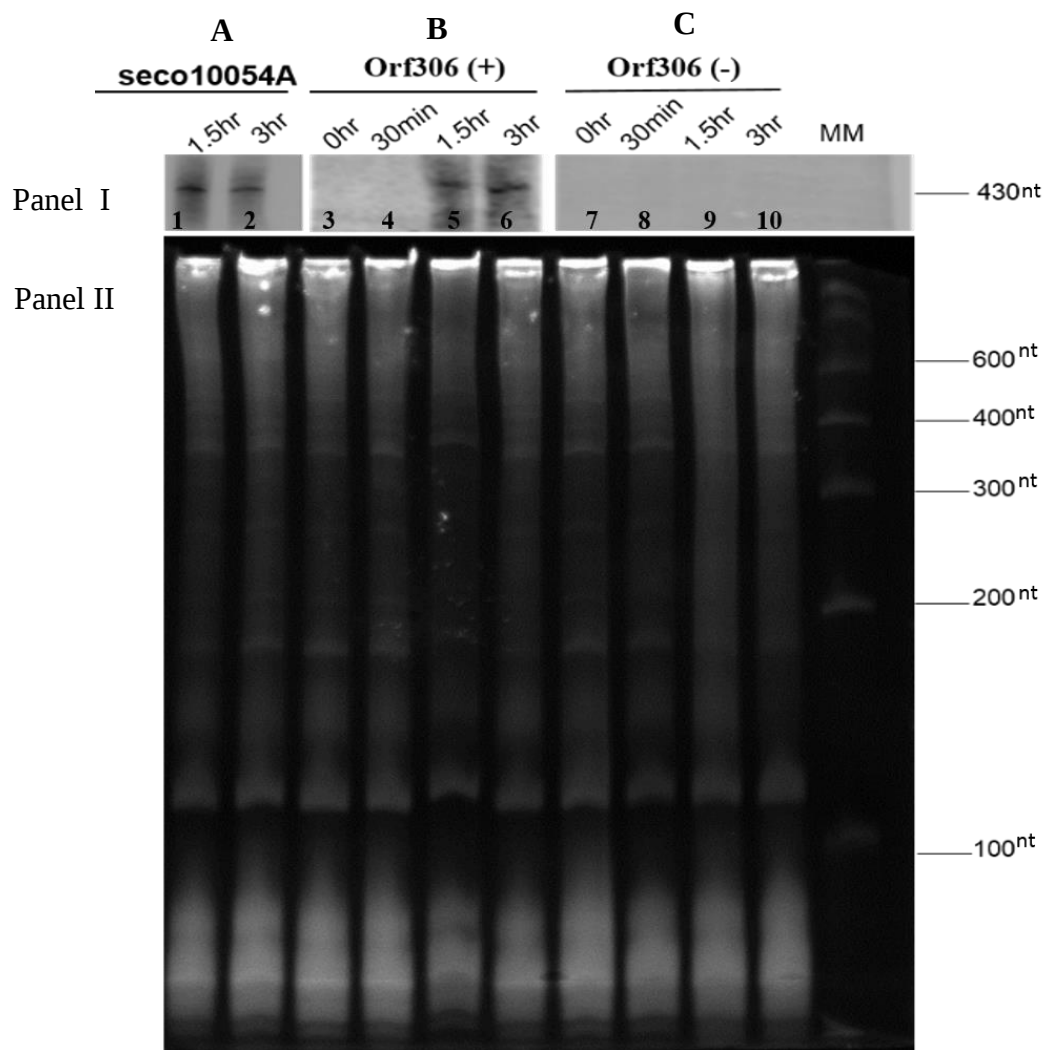


Fig. 5. 06: Detection of *seco10054A* by northern blot analysis: The total RNA from *E. coli* MG1655 (pMMB206), *E. coli* MG1655 (pSDP10) and *E. coli* MG1655 having expression plasmid pAM10054A coding sRNA, *seco10054A* from inducible *tac* promoter was isolated and analysed in 15% acrylamide-7M urea gel (Panel II). The corresponding northern blot showing sRNA, *seco10054A* specific signal was shown in Orf306 positive background (Panel I-B) and cells expressing *seco10054A* from an inducible promoter (Panel I-A). No sRNA signal was seen in Orf306 negative background (Panel I-C).

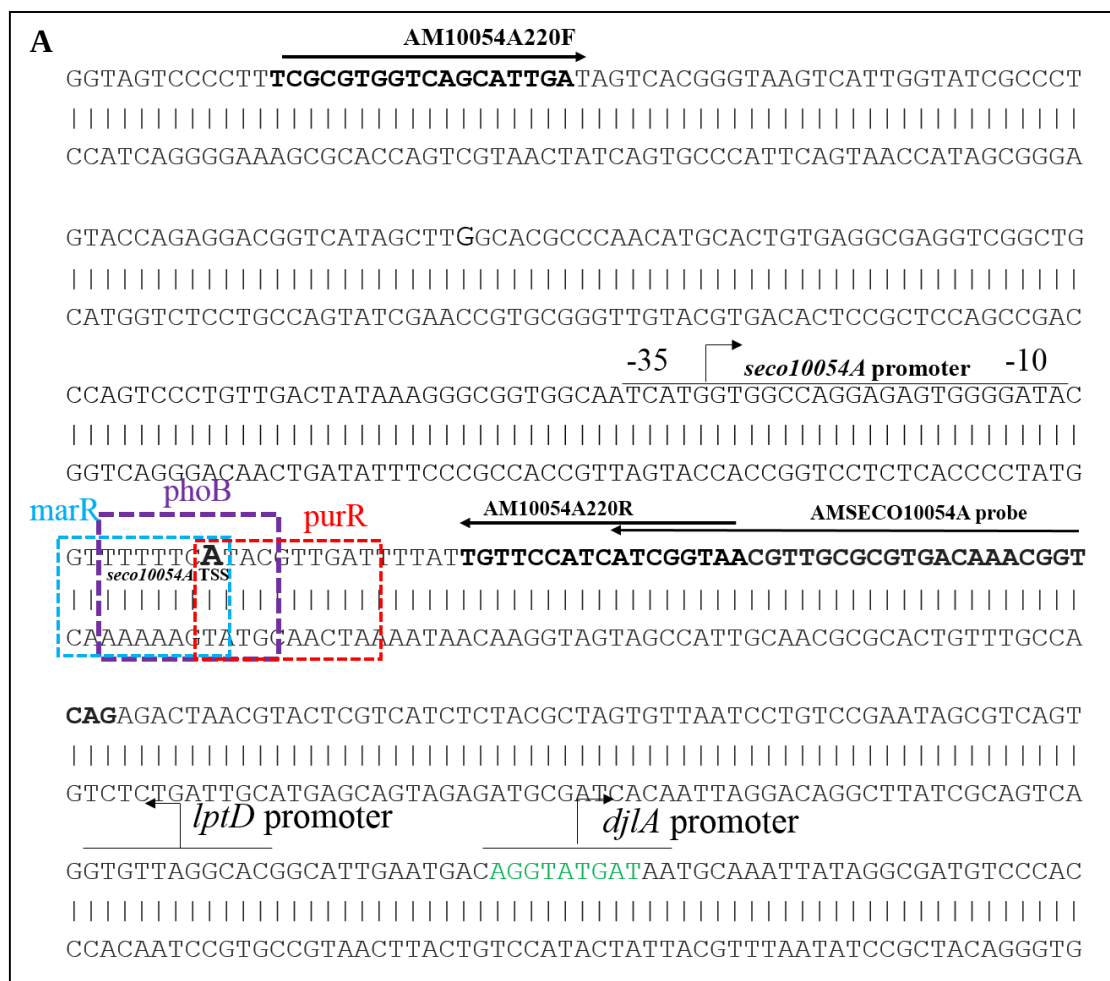
5. 5. 1. The *seco10054A* gene is a PhoB regulon

While gaining further details on Orf306 dependent induction of *seco10054A* gene, The promoter region was re-examined to identify sequence motifs that serve as binding sites to transcriptional factors. The *in silico* analysis (<http://meme-suite.org/tools/meme>) done to predict transcription factor binding sites provided interesting leads. The *in silico* predictions have indicated the existence of MarR, PurR, and PhoB binding sites immediately downstream of the putative promoter sequence (Fig. 5. 07, Panel A and Panel B). The transcription factors MarR, PurR and PhoB are dual transcription regulators (Aleksun and Levy 1997; Meng et al., 1990; Wanner 1993). Depending on the physiological condition they either activate or repress the expression of their cognate genes (Yang et al., 2012). Encouraged by these predictions, further experiments were designed to validate the role of predicted transcription factors, if any, in a regulation of a *seco10054A* gene expression. Initially, the *lac* negative *E. coli* MG1655 cells were used to generate knockouts of genes coding for transcription factors MarR, PurR, and PhoB. As explained in material and methods section, the P1 transducing particles obtained from *marR*, *purR* and *phoB* mutant's strains of *E. coli* K-12 BW25113 were independently infected to *lac* negative *E. coli* MG1655 strains to obtain respective mutants. The colonies appeared on kanamycin-containing LB plates were then screened by PCR using gene specific primers (Table 2. D). The Fig. 5. 08 shows agarose gel indicating deletion of *purR* (Panel A-I), *marR* (Panel A-II) and *phoB* (Panel A-III) genes from *E. coli* MG1655 cells. After generating these mutants, they were transformed with

a plasmid containing *seco10054A-lacZ* fusions (pAM10054A220). In principle, all of them should display *lac* positive phenotype if they are involved in repression of *seco10054A* gene. As shown in the Fig. 5. 08, Panel B, such repression of *seco10054A* gene expression is seen only in *purR* and *marR* null mutants. However, in the *phoB* null background, no such repression of *seco10054A* gene was observed. The *phoB* null mutants containing *seco10054A-lacZ* fusion gave 700 units of β -galactosidase activity (Fig. 5. 08, Panel B). In the *phoB* positive background, the *seco10054A-lacZ* fusions have given just 150 units of β -galactosidase activity (Fig. 5. 08, Panel B). These results clearly suggest the involvement of PhoB in repression of *seco10054A* gene expression and it also indicates that the predicted PhoB binding site downstream of the promoter of *seco10054A* gene is functional.

5. 5. 2. PhoB binds to the predicted PhoB motif of *seco10054A* gene

After establishing PhoB involvement in repression of *seco10054A* gene promoter an attempt was made to gain *in vitro* evidence on predicted *phoB* motif and PhoB interactions by performing Surface Plasmon Resonance (SPR) spectroscopy (vide material and methods section). In order to perform SPR experiments, pure PhoB and PhoB binding motif are required. The complementary oligo's having PhoB binding site were designed to generate the PhoB binding site. Likewise, the affinity purified PhoB^{GHIS} was used while performing SPR studies.



B

Transcription Factor	Motif	E-value
PurR		9.9e+002
MarR		4.7e+003
PhoB		9.2e+002

Fig. 5. 07. Nucleotide sequence of *seco10054A* promoter region: The bent arrow indicates promoter motifs and direction of transcription. The putative transcription factor binding sites are shown with open boxes. (Panel A). The consensus PdhR, MarR, and PhoB binding motifs are shown in panel B.

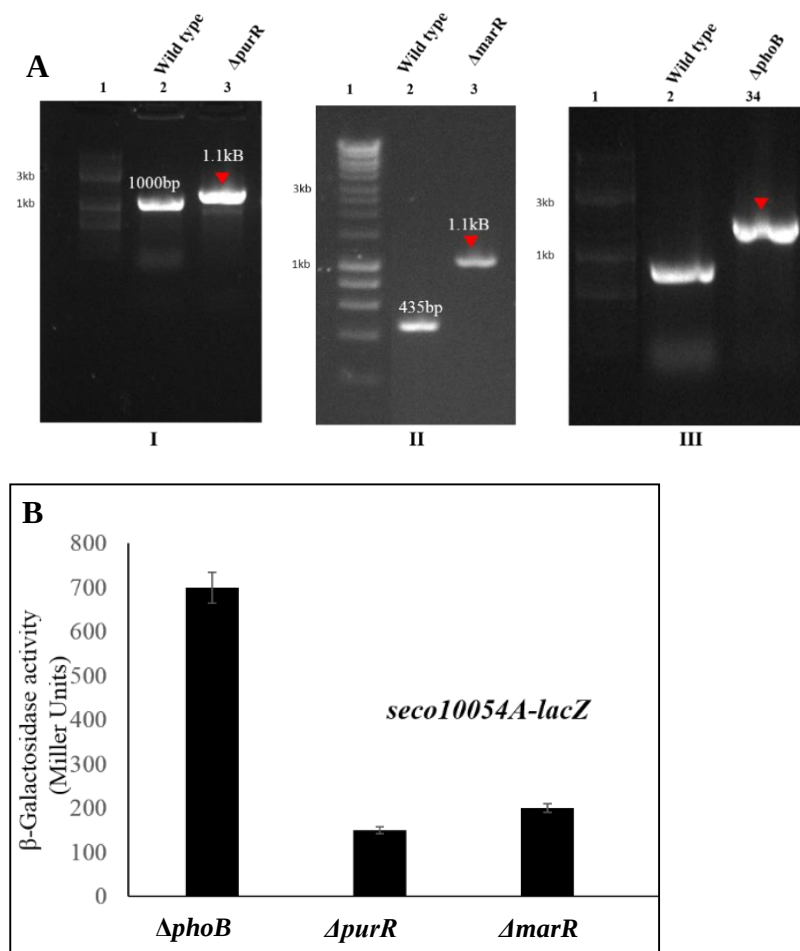


Fig. 5. 08. Generation of *purR* (I) *marR* (II) and *phoB* (III) knockouts in *lac* negative *E. coli* MG1655: Lane 1 is loaded with 1 kb DNA ladder. Lanes 2 and 3 represent size of PCR products generated using chromosomal DNA isolated either from wild type (Lane 2) or mutant (Lane 3) as templates. Panel B: β -galactosidase activity (in miller units) in *lac* negative *E. coli* MG1655 cells containing *seco10054A-lacZ* fusion under *phoB*, *purR* and *marR* negative backgrounds.

5. 5. 3. Expression and purification of PhoB

The *phoB* gene was amplified as *NdeI* and *XhoI* fragment using forward (AMPHOB23BF, 5'-GAATTCCATATGGCGAGACGTATTCTGGT-3') and reverse (AMPHOB23BR, 5'-CTGCAGCTCGAGAAAGCGGGTTGAAAAACGATAT-3') primers appended with *NdeI* and *XhoI*. The amplicon was then digested with *NdeI* and *XhoI* and ligated to pET23B digested with similar enzymes. The resultant plasmid was designated as pAMPHOB23B. As cloning places, *phoB* gene in frame of vector coded His-tag, the resultant plasmid pAMPHOB23B codes for PhoB^{6HIS}. The BL21 *E. coli* (pAMPHOB23B) cells were induced for the expression of PhoB^{6HIS} as mentioned in material and methods section. The SDS-PAGE (12.5%) analysis has clearly shown induction of 26 kDa PhoB^{6HIS} only in induced cultures (Fig. 5. 09, Panel A, Lanes 3). The expressed PhoB^{6HIS} was affinity purified using metal ion chromatography. As indicated in Fig. 5. 09, Panel A, Lane 9, and 10, the purified PhoB contains no other contaminated proteins.

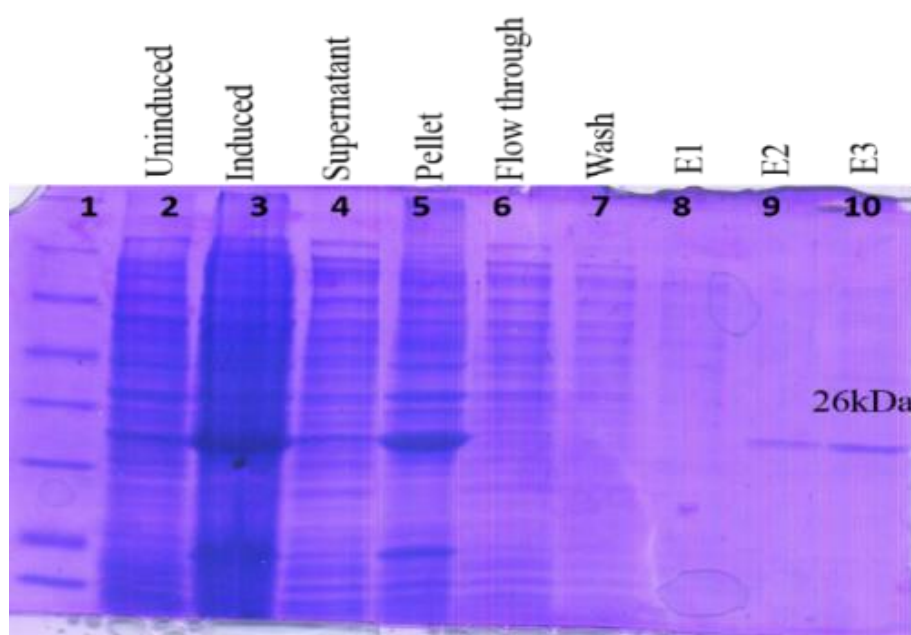


Fig. 5. 09. Purification of PhoB^{6HIS}: Panel A: Purification of PhoB^{6HIS}; Lane 1- Molecular weight marker.

The proteins extracted from uninduced and induced cultures are shown in lanes 2 and 3 respectively. The PhoB found in soluble and inclusion bodies is shown in lanes 4 and 5 respectively. The fractions such as flow through (Lane 6), wash (Lane 7) and eluted fractions E1 (Lane 8), E2 (Lane 9) and E3 (Lane 10) generated during affinity purification of PhoB^{6HIS} were analysed on 12.5% acrylamide gel. Pure PhoB found in elution fractions E2 and E3 are shown in lanes 9 and 10.

5. 5. 4. Determination of *phoB* –PhoB^{6HIS} interactions by SPR

Preparation of oligos with and without PhoB binding site was taken from the promoter region of *seco10054A* and complementary oligos were (AMSPR10054AF, 5'-GGCAATCATGGTGGCCAGGAGAGTGGGGATACGTTTTTTCATACGTTGATTTTA-3' and AMSPR10054AR, 5' TAAAATCAACGTATGAAAAACGTATCCCCACTCTCCTGGCCACCATGATTGCC-3') designed to generate PhoB binding site that exactly resembles the predicted PhoB binding site of *seco10054A* gene. This complementary oligo were taken in eqimolar ratio and annealated to generate PhoB binding site. After annealation the generated *phoB* motif is biotinylated as indicated in material and methods section. The biotinylated *phoB* motif is then immobilised on streptavidin chip. The similarly sized oligo's (AMSPR10054AΔPHOBF, 5'-GGCAATCATGGTGGCCAGGAGAGTGGGGATACGTTTGATTTTA-3' and AMSPR10054AΔPHOBR, 5'-TAAAATCAAACGTATCCCCACTCTCCTGGCCACCATGATTGCC-3') having no *phoB* binding site was treated in a similar manner and used as negative control.

The pure PhoB^{6HIS} taken in concentrations ranging from 0. 15 nm to 0. 5 nm by dissolving appropriate amounts of PhoB^{6HIS} in HBS-EP buffer. The PhoB^{6HIS} containing buffer was passed over 1min on immobilised oligo's having with and without PhoB binding sites. SPR results revealed that the binding affinity (K_D) for PhoB^{6HIS} to the predicted PhoB binding site is $k_D = 8.107E-10$. As shown in Fig. 5. 10, the observed interactions increased with the increase of PhoB^{6HIS} concentrations (Fig. 5. 10, Panel A). When PhoB^{6HIS} concentration was 0.015 nano-mole the refractive index as measured in response units was only 2 RU. With the increase of PhoB^{6HIS} concentration in the buffer, the refractive index has gone up to 20 RU. There exists a linear concentration-dependent increase in binding of PhoB^{6HIS} with *phoB* binding motif suggesting that the observed interactions are specific According to response to concentration plot shown in Fig. 5. 10, Panel B, at least 810.7 picomole to 1621.4 picomole needed to have 100% strength of binding between PhoB^{6HIS} and PhoB binding motif.

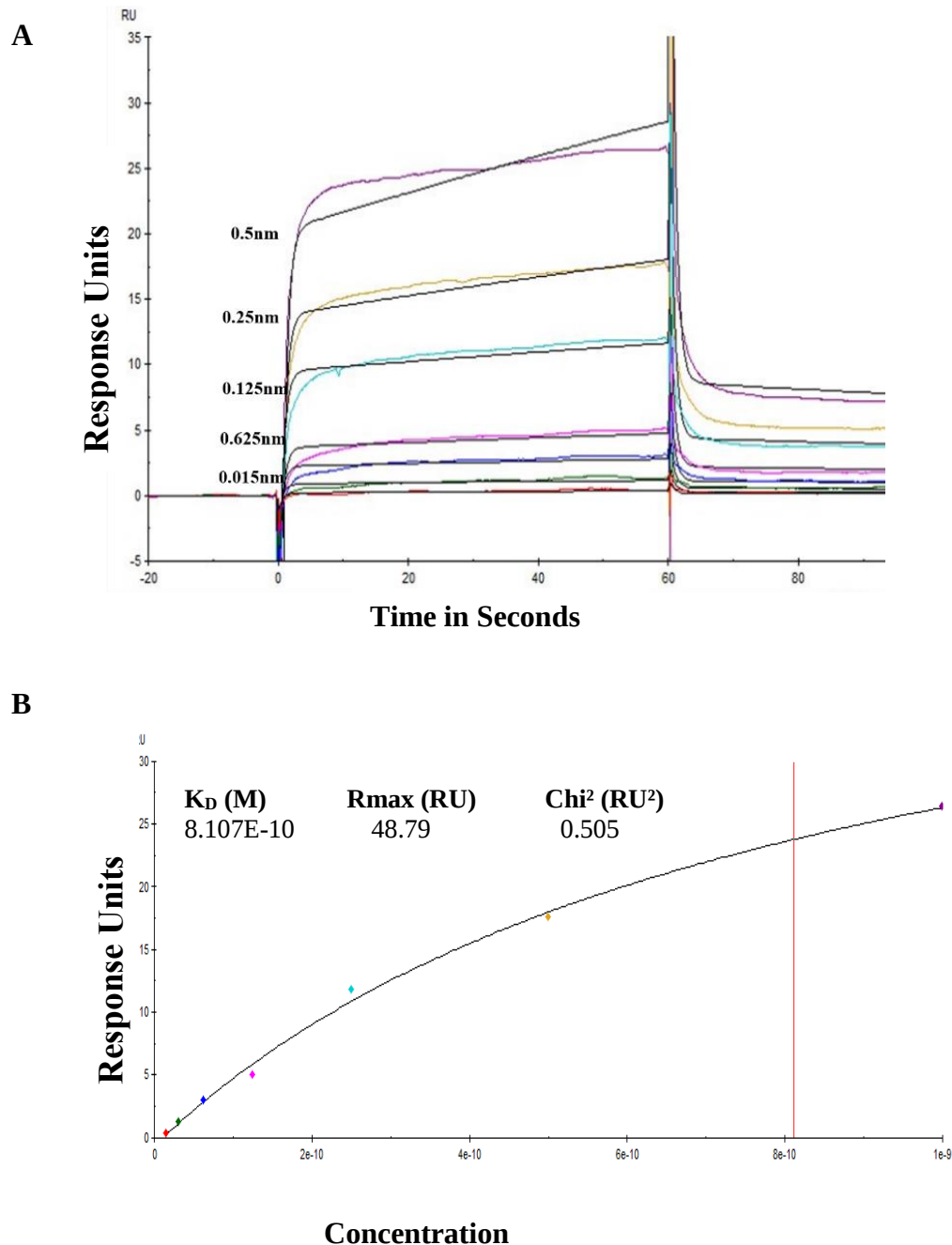


Fig. 5. 10. PhoB^{6HIS} interactions with consensus *phoB* binding motif identified overlapping -10 region of *seco10054A* gene. Panel A indicates visual assessment of sensogram fitted data. The Y-axis indicates response units obtained when different concentrations of PhoB^{6HIS} bound to the ologo containing *phoB* binding site. Five different concentrations of PhoB^{6HIS} was taken while performing SPR at each titration. The black line indicates fitted curves which is overlaid on the sensogram data shown in coloured lines. Binding kinetics as revealed by plotting response units versus concentration of PhoB^{6HIS} is indicated in panel B.

5. 6. The link between Orf306 and sRNA, *seco10054A* and *seco11136A* regulation

The experimental evidence gained in this chapter clearly suggests up-regulation of *seco10054A* gene and down-regulation of *seco11136A* gene in the presence of Orf306. The promoter assays and northern bolts clearly indicated the requirement of Orf306 for induction of *seco10054A* gene. However, the relationship between Orf306 dependent induction of *seco10054A* and repression of *seco11136A* gene are unknown. Initially, the Orf306 sequence was analyzed using online tools to predict the existence of sequence motifs that bind to promoter elements causing either activation or repression of these genes. However, such DNA binding motifs were not found in Orf306. In consistant of its homology with esterases, the pure Orf306 has shown catalytic activity when assayed using nitrophenyl acetate as substrate and suggested that the Orf306 is an esterase/lipase. Based on the enzyme assay a hypothesis was proposed stating that expression of Orf306 is generating endogenous fatty acids due to its esterase/lipase activity. If the generated odd chain fatty acids are metabolized through conventional β -oxidation pathway, propionayl-CoA will be generated as an end product. The proposed hypothesis also fits to the identified catabolic shift in *E. coli* MG1655 (pSDP10) cells where the propionate degradation pathway was found significantly up-regulated in *E. coli* MG1655 (pSDP10) cells (Chakka et al., 2015).

5. 6. 1. Orf306 generates endogenous propionate

While testing the proposed hypothesis, *E. coli* MG1655 (pSDP10) cells and *E. coli* MG1655 (pMMB206) cells are induced to produce Orf306. The *E. coli* MG1655 (pMMB206) cells served as vector control. These induced cells were quickly harvested and total lipids were extracted and analyzed as described in materials and methods section. Interestingly in *E. coli* MG1655 (pSDP10) cells, a new peak was identified at the retention time of 11.5 min (Fig. 5. 2. 0, Panel A). The peak was then subjected to mass analysis to know its identity. The mass spectrum of the new peak perfectly matches with the mass spectrum of propionic acid decyl ester (Fig. 5. 2. 0, Panel B). This data clearly suggests that in the presence of Orf306 propionyl-CoA is generated due to oxidation of odd chain fatty acids. The generated propionyl-CoA in an unknown way appears to have converted to propionic acid decyl ester. Since propionyl-CoA generation in *E. coli* MG1655 (pSDP10) cells is apparent, it is further hypothesized that the generated propionate has a direct influence on the expression of sRNA genes *seco10054A* and *seco11136A*. In order to validate the hypothesis the constructed promoter *lacZ* fusions were used and promoter, assays were performed in the presence of propionate.

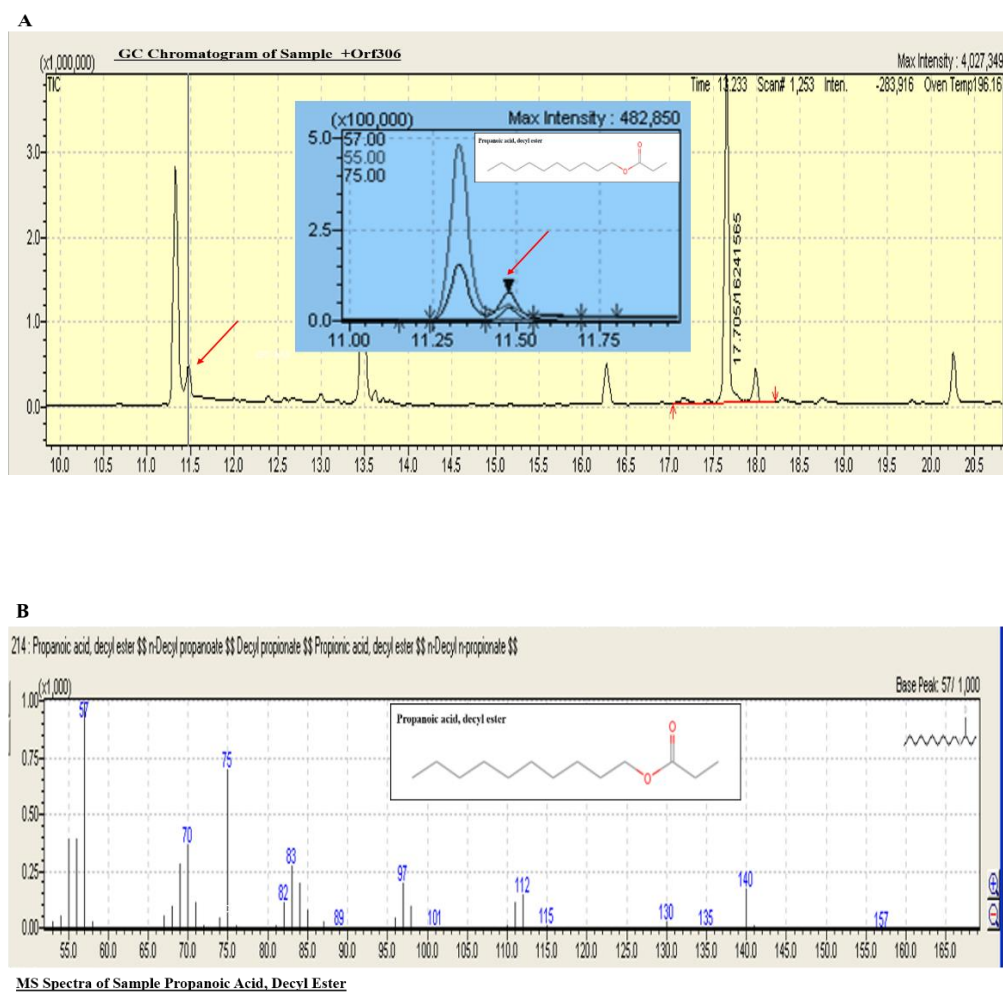


Fig. 5. 2. 0. Analysis of total lipids from *E. coli* MG1655 (pSDP10) cells and *E. coli* MG1655 (pMMB206) cells. Panel A indicates GC analysis of lipidome of *E. coli* MG1655 (pSDP10) cells. The new peak was identified at a retention time of 11.5 min. Panel B shows the mass spectrum of the new peak that perfectly matches with the mass spectrum of propionic acid decyl ester.

5. 6. 2. Propionate induces *seco10054A* gene

The *lac* negative *E. coli* MG1655 cells were transformed with *seco10054A-lacZ* fusion and these cells were then plated on a minimal medium plate containing propionate as a sole source of carbon. The *E. coli* MG1655 (pMP220) served as negative control. The minimal medium plate contains X-gal, the propionate dependent *lacZ* activation can be monitored due to the presence of blue coloured colonies. Interestingly, the *seco10054A-lacZ* fusion containing *E. coli* MG1655 (pAM10054APF) turned into intense blue coloured colonies (Fig. 5. 2. 1, Panel A, Sector-i). However, the cells containing promoter test vector pMP220 remained pale in colour (Fig. 5. 2. 1, Panel A, Sector-ii). This data clearly suggested propionate dependent activation of *seco10054A* gene. The quantitative assay clearly supported the observations

gained under qualitative assays. The propionate-grown *E. coli* MG1655 (pAM10054APF) cultures gave 350 miller units of β -galactosidase activity as against 5 miller units of β -galactosidase activity obtained in glucose-grown cultures (Fig. 5. 2. 1, Panel B).

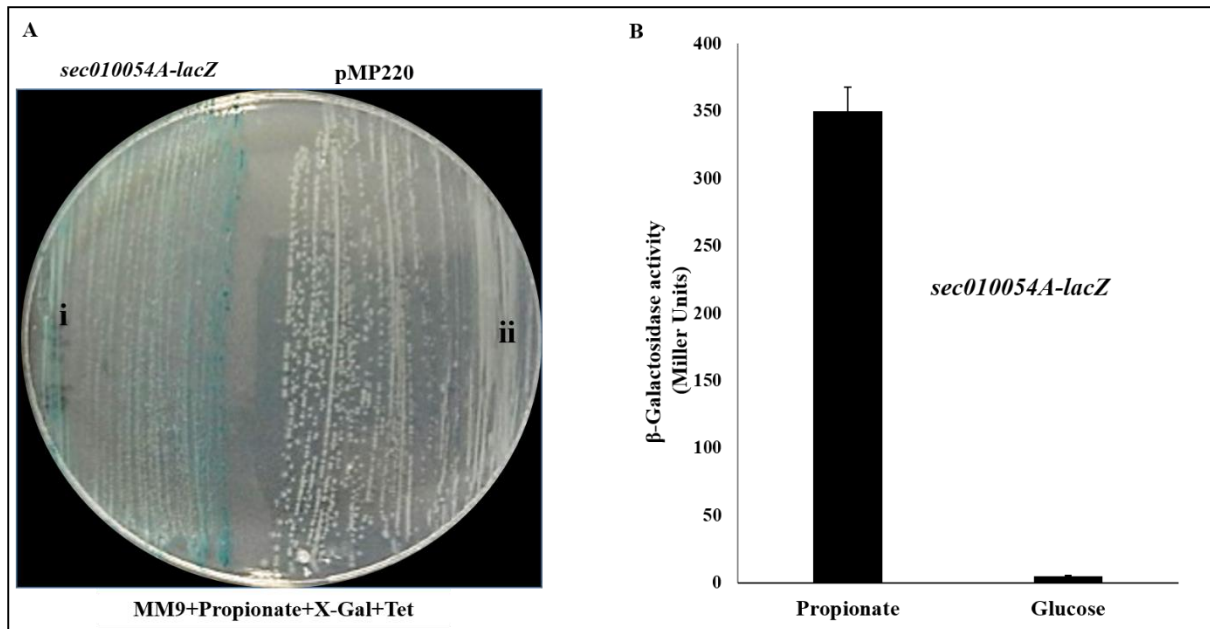


Fig. 5. 2. 1. Influence of propionate on expression of *seco10054A* gene. Panel A shows blue coloured *E. coli* MG1655 (pAM10054A220) cells indicating propionate dependent activation of *seco10054A-lacZ* fusion (Sector-I). The pale coloured *E. coli* MG1655 (pMP220) cells are shown in sector II. Panel B indicates quantitative assay: About 350 miller units of β -galactosidase activity were obtained in *E. coli* MG1655 (pAM10054A220) cultures grown in propionate. Only 5 miller units of β -galactosidase activity were obtained in glucose grown cultures of *E. coli* MG1655 (pAM10054APF) cells.

5. 6. 3. Propionate represses *seco11136A* gene

The aforementioned experiments clearly suggested propionate dependent induction of *seco10054A* gene. As *seco11136A* gene is repressed in the presence of Orf306 it is hypothesized that the generated propionate has a repressive effect on gene expression. Therefore the generated *seco11136A-lacZ* fusions were used to assess the influence of propionate on the expression of *seco11136A* gene. The *lac* negative cells *E. coli* MG1655 containing either *seco10054A-lacZ* fusion or *seco11136A-lacZ* fusion were plated on minimal media plates containing propionate as a sole source of carbon. As stated before *E. coli* MG1655 (pAM10054APF) cells turned blue indicating propionate dependent activation of *seco10054A-lacZ*. However, the *E. coli* MG1655 cells having *seco11136A-lacZ* remained pale in colour (Fig. 5. 2. 2, Panel A, Sector-II) showing propionate dependent repression *seco11136A* gene. The quantitative assays reflected the observations made in qualitative assays. The *E. coli* MG1655 (pAM10054A220) cells gave 295 miller units of β -galactosidase activity as against

15 miller units obtained in *E. coli* MG1655 (pAM11136A220) cells (Fig. 5. 2. 2, Panel B). This clearly suggests that the Orf306 dependent generation of propionate has an influence on expression *seco10054A* and *seco11136A* genes. The propionate induces *seco10054A* gene and represses *seco11136A* gene. As shown in chapter II these two reciprocally regulated sRNAs, *seco10054A* and *seco11136A* inhibit key enzymes of glycolysis and TCA cycle and promote translation activation of transcription factors HcaR and FadR.

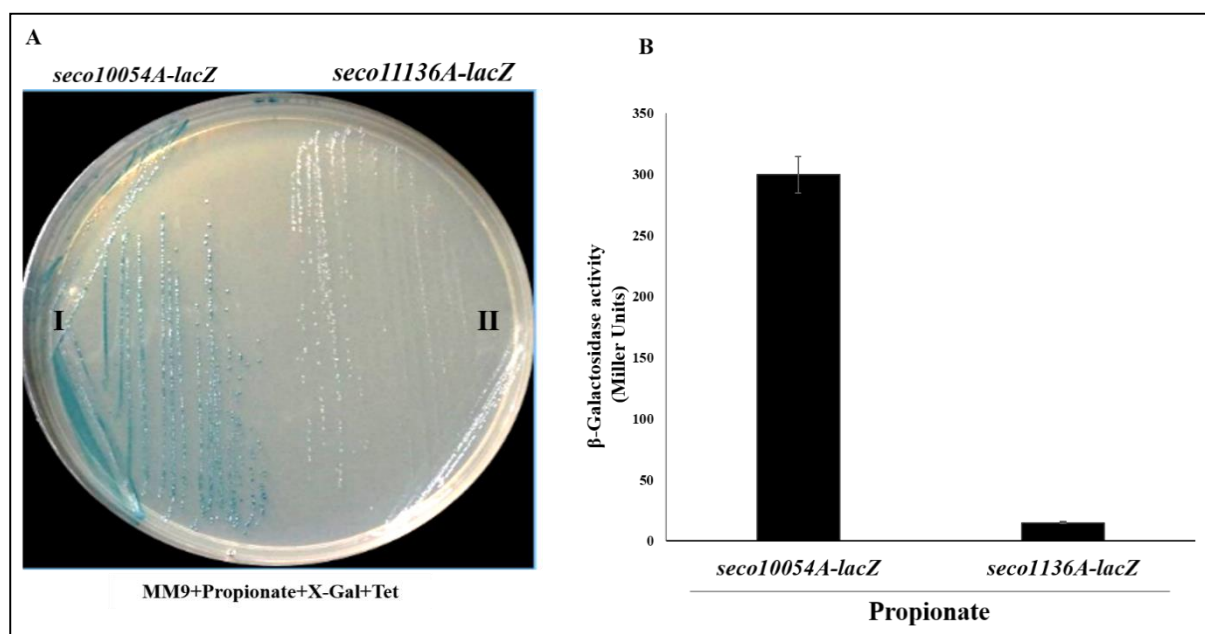


Fig. 5. 2. 2. Propionate dependent repression of *seco11136A* gene: Panel A: Qualitative β-galactosidase assay: *E. coli* MG1655 (pAM11136A220) cells remained pale in colour. (Panel A, Sector-II), whereas the *E. coli* MG1655 (pAM10054A220) cells have shown intense blue colour in Panel A, Sector- I. Quantitative assays have given only 15 miller units of β-galactosidase activity where *E. coli* MG1655 (pAM10054A) cells as against 295 miller units of β-galactosidase activity obtained in

5. 7. Propionate activates *hcaR* transcription

The *hca* and *mhp* operons are activated in *E. coli* MG1655 (pSDP10) cells. These two operons are involved in degradation of phenyl propionate and hydroxy propionate. Our previous studies have shown their involvement in degradation of PNP (Chakka et al., 2015). Since *hca* and *mhp* operons are up-regulated in Orf306 positive background it is further hypothesized that the Orf306 dependent generation of propionate has a role in activation of *hca* and *mhp* operons. The *hca* and *mhp* operons are positively regulated by HcaR and MhpR. Therefore, activation of *hca* and *mhp* requires the presence of active transcription factors HcaR and MhpR. These two transcription factors are found to be repressed in glucose-grown cultures (Kovárová et al., 1996). In order to assess the influence of propionate on the expression of *hcaR*

and *mhpR* genes, the promoter fusions were constructed by cloning the promoter region of these two transcription factor coding genes in promoter test vector pMP220. The *mhpA*, the first gene of *mhp* operon is taken as negative control. The generated transcription fusions are designated as pAMHPRPF, pAMMHPAPF and pAMHCARPF. The detailed cloning strategy followed to generate *mhpR-lacZ*, *mhpA-lacZ* and *hcaR-lacZ* fusions are shown in Fig. 5. 2. 3. This *lacZ* fusions were then transformed into *lac* negative cells of *E. coli* MG1655. The qualitative assays were performed on the minimal medium plate containing propionate as sole source of carbon. The plate was divided into four sectors and each one of these sectors is then devoted to grow *hcaR-lacZ* fusions (pAMHCARPF), *mhpR-lacZ* fusions (pAMMHPAPF) and promoter test vector (pMP220). The plate contained X-gal, therefore the propionate dependent activation of *lacZ* fusions can easily be monitored by observing the colour. As shown in Fig. 5. 2. 4, Panel-I, Sector-A, intense blue colour was seen in the sector of plate used to grow *hcaR-lacZ* fusion. A very mild blue colour was seen in the sector used to grow cells having *mhpR-lacZ* fusion (Fig. 5. 2. 4, Panel-I, Sector-B). Such *lacZ* activation was not seen in the sector used to plate either cells having promoter test vector (pMP220) (Fig. 5. 2. 4, Panel-I Sector-D) or *mhpA-lacZ* fusion (Fig. 5. 2. 4, Sector-C). The same trend was seen in quantitative assays (Fig. 5. 2. 4., Panel-III). The cells having *hcaR-lacZ* fusions gave 180 miller units of β -galactosidase activity as against 50 units in cells containing *mhpR-lacZ* fusions. The rest of the control cultures (pMP220) gave negligible amounts of β -galactosidase activity. The same experiment was repeated by supplementing glucose as a sole source of carbon. Interestingly, none of the cultures have shown *lac* positive phenotype (Fig. 5. 2. 4, Panel II and Panel IV), suggesting propionate dependent induction of *hcaR* gene in *E. coli* MG1655 (pSDP10) cells. Propionate dependent induction of *hcaR* gene explains the Orf306 dependent activation of both *hca* and *mhp* operons.

Table 5. A: PCR primers used to generate *mhpR*, *mhpA* and *hcaR* genes promoter fusions.

Name of the Primer	Primer sequences	Primers used for	Size of the Amplicon in Base pairs
AMMHPRF AMMHPRR	CCGGA <u>GAATTC</u> TTCAGTACCTCACGAC CCGAG <u>CTGCAG</u> ATTAATTGACATTTCTAT A	Used to amplify promoter region of <i>mhpR</i> gene to construct <i>mhpR-lacZ</i> promoter fusions by cloning the amplicon as <i>EcoRI</i> and <i>PstI</i> fragment using promoter test vector pMP220.	190 bp
AMMHPAPF AMMHPAPF	CCGGA <u>GAATTC</u> ATTAATTGACATTTCTAT CCGAG <u>CTGCAG</u> TTCAGTACCTCACGACTC GG	Used to amplify promoter region of <i>mhpA</i> gene to construct <i>mhpA-lacZ</i> promoter fusions by cloning the amplicon as <i>EcoRI</i> and <i>PstI</i> fragment using promoter test vector pMP220.	190 bp
AMHCARPF AMHCARPF	AG <u>GAATTC</u> AAGCTCCAGTTGGTAA AC <u>CTGCAG</u> ACGGGTAAAGTTCAGT	Used to amplify promoter region of <i>hcaR</i> gene to construct <i>hcaR-lacZ</i> promoter fusions by cloning the amplicon as <i>EcoRI</i> and <i>PstI</i> fragment using promoter test vector pMP220.	300 bp

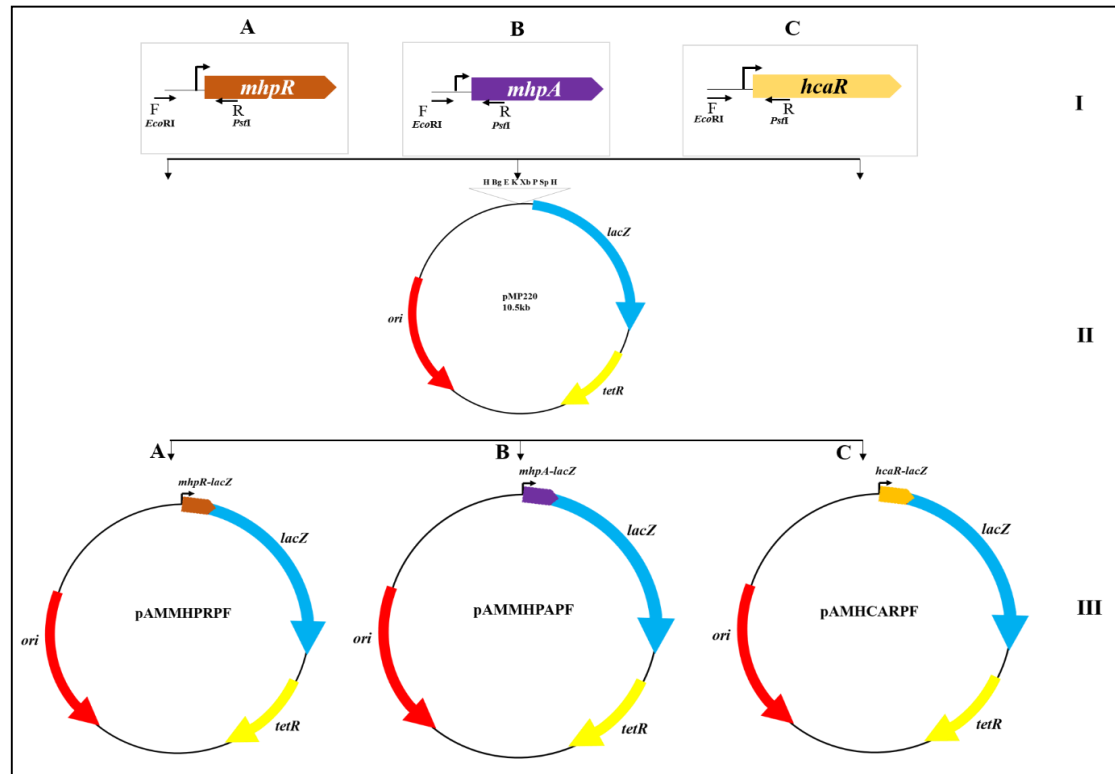


Fig. 5. 2. 3. Construction of *mhpR-lacZ*, *mhpA-lacZ* and *hcaR-lacZ* fusions: Promoter regions of *mhpR*, *mhpA* and *hcaR* genes were amplified by using wild type *E. coli* MG1655 genomic DNA as template and primers (vide Table 5. A) appended with *EcoRI* and *PstI* restriction sites (Panel I-A, B and C). The promoter fusions were constructed by cloning the promoter region of these genes promoter amplicons in promoter test vector pMP220 (Panel II). The generated transcription fusions were designated as pAMMHPRPF, pAMMHPAPF and pAMHCARPF (Panel III).

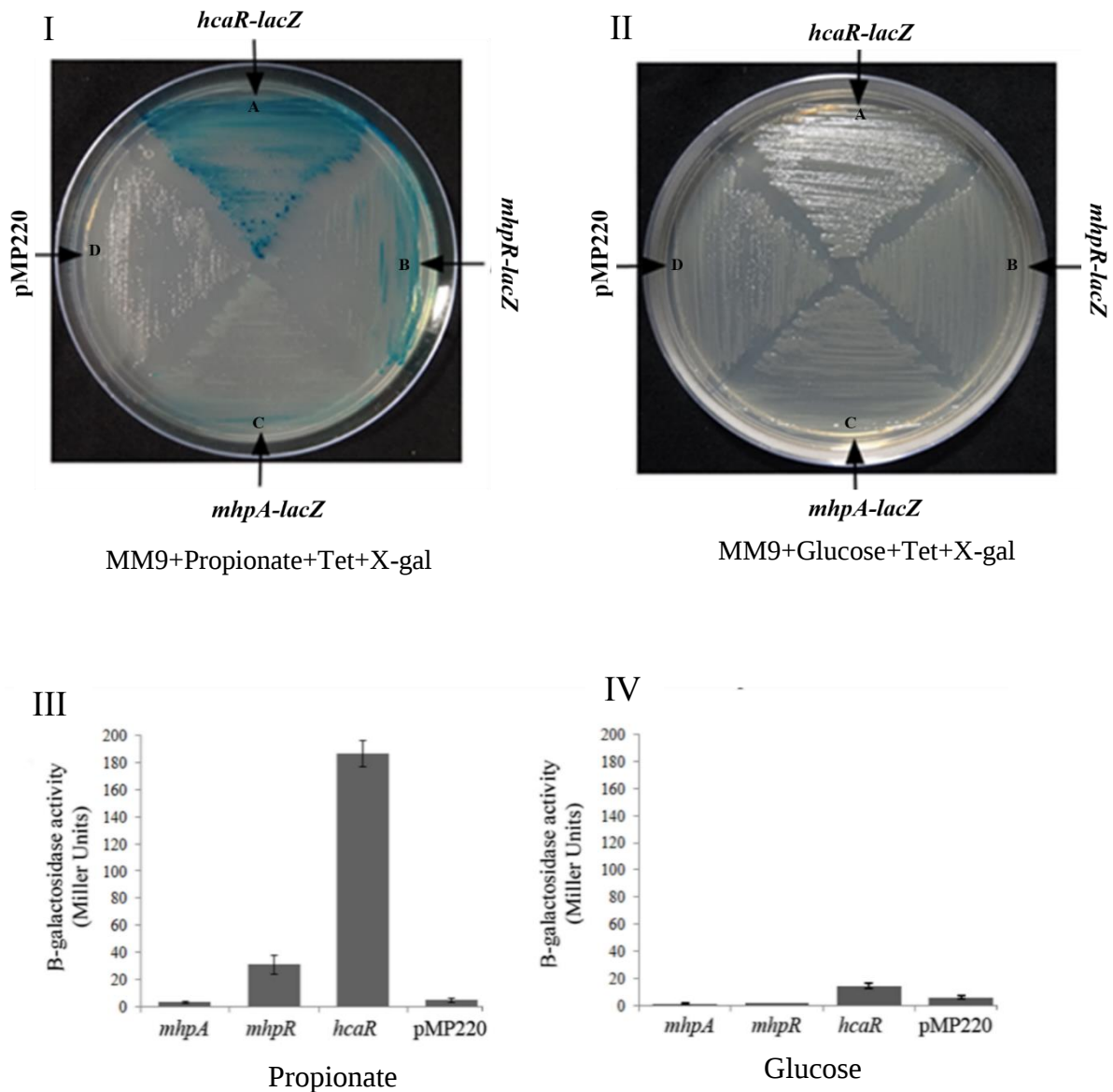


Fig. 5.2.4. Propionate dependent induction of *hcaR* gene expression: Panel I and Panel II indicates growth of the *lac* negative strains of *E. coli* MG1655 containing promoter test vector pMMP220, *hcaR-lacZ*, *mhpR-lacZ* and *mhpA-lacZ* fusions in propionate and glucose plates. Panel III and Panel IV indicate the quantification of β-galactosidase activity for propionate and glucose grown cultures respectively. The propionate dependent activation of the *hcaR-lacZ* and *mhpR-lacZ* shown in Panel I, Sector A and B respectively. Similar cultures grown in glucose remained *lac* negative (Panel II). The quantitative assays performed for similar cultures grown in propionate and glucose are shown in Panel III and IV respectively.

5. 8. Discussion

The *p*-nitrophenol supported the growth of *E. coli* MG1655 (pSDP10) cells ignited interest to analyse molecular basis for such unusual growth phenotype of *E. coli* MG1655 (pSDP10) cells. The transcriptome analysis clearly indicated Orf306 dependent repression of glycolysis and TCA cycle and up-regulation of propionate of phenyl propionate degradation pathways. Our previous studies have shown the involvement of phenyl propionate degradation pathways in PNP catabolism (Chakka et al., 2015). The investigations carried out in this study has clearly shown the involvement of sRNAs, *seco10054A*, and *seco11136A* in an Orf306 dependent metabolic shift in *E. coli* MG1655 cells. The data generated in this study clearly shows the roles of these reciprocally regulated sRNAs, *seco10054A*, and *seco11136A* in the observed inhibition of glycolysis, TCA cycle and activation of *hca* operon. The critical question that remained unanswered was the relationship between the Orf306 and regulation of sRNA genes *seco10054A* and *seco11136A* expression. The Orf306 is an esterase/lipase and doesn't have any signature motifs that clearly suggests its direct involvement in observed activation of *seco10054A* gene or repression of *seco11136A* gene. The only possibility for such Orf306 dependent differential regulation could be due to its observed lipase activity. The present study established generation of propionyl-CoA in *E. coli* MG1655 (pSDP10) cells. Therefore, an attempt was made to gather experimental evidence linking propionate and observed catabolic shift in *E. coli* MG1655 (pSDP10) cells. The *seco10054A-lacZ* fusions clearly got activated in the presence of propionate (Fig. 5. 2. 2, Panel A, Sector-I) suggesting propionyl-CoA generated in the presence of Orf306 is responsible for induction of sRNA, *seco10054A* expression. Since, sRNA, *seco10054A* is made only in the presence of Orf306, its presence prevents translation of Lpd subunit, a critical component of glycolysis and TCA cycle enzymes, pyruvate dehydrogenase and α -ketoglutarate dehydrogenase. Such translation inhibition by sRNA, *seco10054A* appears to be responsible for the down-regulation of glycolysis and TCA cycle.

The present study also demonstrates the existence of *phoB* motif overlapping the -10 of *seco10054A* gene (Fig: 5. 06 Panel A). The PhoB binding site has in fact served as substrate for PhoB. The pure PhoB bound to the predicted *phoB* site implicating its role in the regulation of sRNA gene. In support of *in vitro* data, the sRNA, *seco10054A* gene got significantly up-

cycle. The Orf306 generates propionate and the generated propionate induces sRNA, *seco10054A* which in turn prevent translation of *lpd* mRNA. Precisely this mechanism appears to be the reason for the observed the Orf306 dependent down regulation of glycolysis and TCA cycle (Fig. 5. 2. 6).

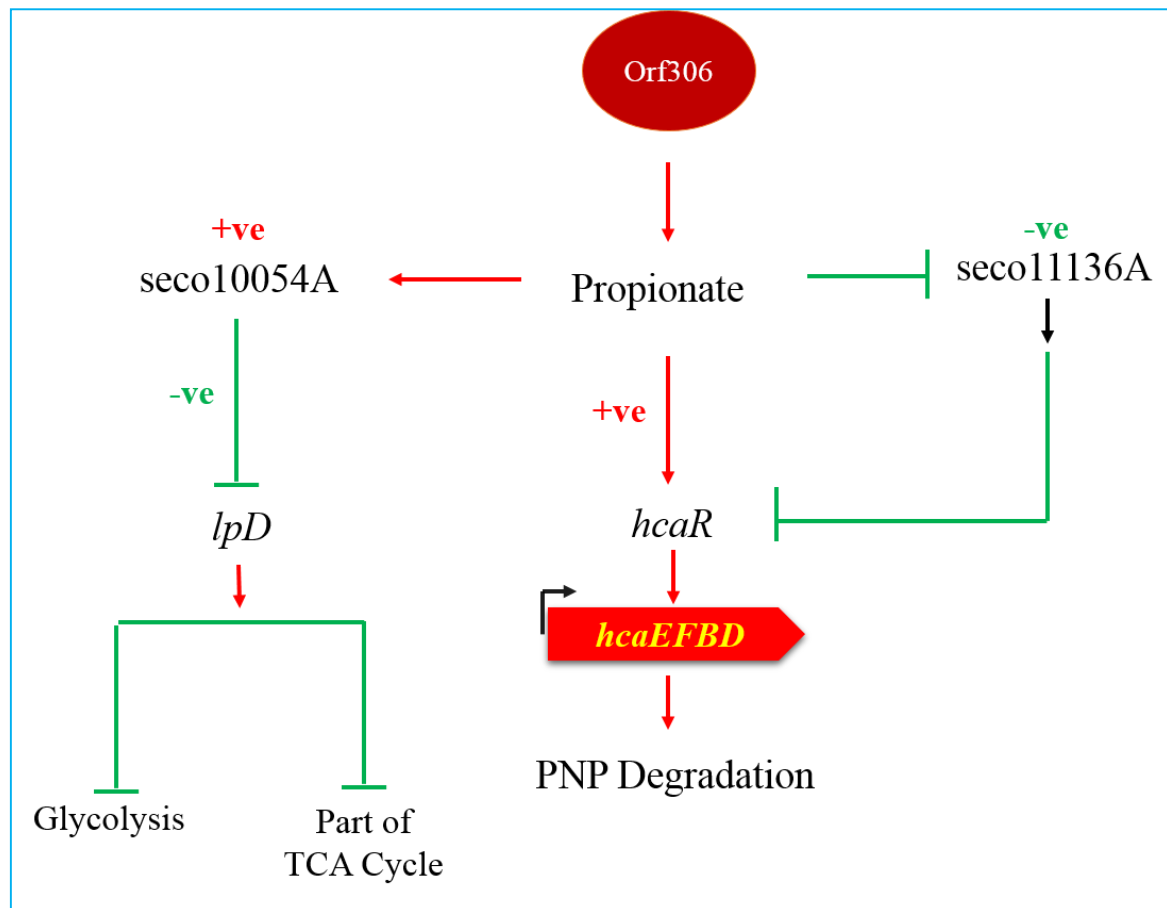
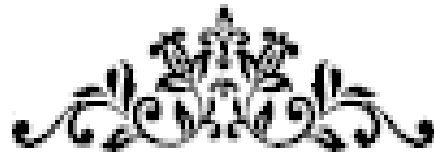
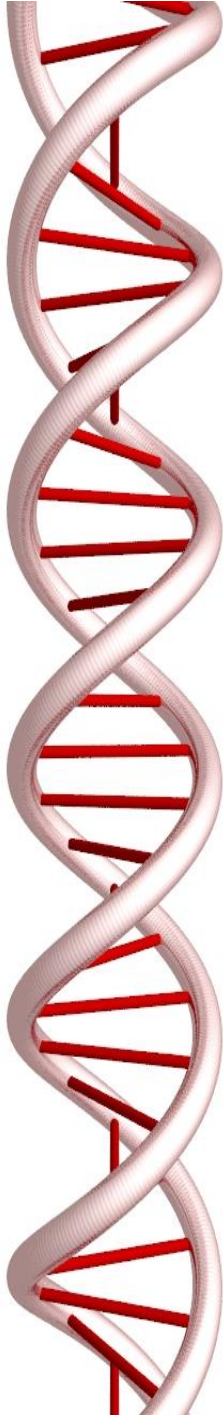


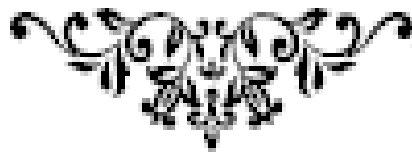
Fig. 5. 2. 6. Proposed model showing sRNAs *seco10054A* and *seco11136A* induced metabolic shift in *E. coli* MG1655 (pSDP10) cells

As shown in Fig. 5. 2. 5, the transcription factor PrpR is also part of PhoB regulation (Yang et al., 2012). PhoB represses the expression of PrpR (Yang et al., 2012). Probably the Orf306 dependent propionate generation might be responsible for derepression of the *prpR* gene. Such derepression is bound to produce active transcription factor PrpR which in turn induces *prp* operon contributing for the up-regulation of propionate degradation pathway. This again suggests that Orf306 dependent generation of propionate is the key signaling molecule that triggers transcription activation of *prp* operon. The most interesting observation is propionate dependent induction of *hcaR* expression. Regulation of HcaR is a well-studied subject. It is repressed in glucose-grown cultures (Turlin et al., 2001; Kovárová et al., 1996).

Again its transcription is activated during stationary phase in a σ^S dependent manner (Turlin et al., 2001). The propionate dependent activation of *hcaR* gene as shown in this study is a new observation. The propionate is playing multiple roles and causing for the genetic reprogramming that facilitates activation of *hca* and *mhp* operons. The present study clearly shows propionate dependent repression of sRNA, *seco11136A*, which inhibits translation of *hcaR* mRNA. Such a situation is a favorable condition for producing maximum amounts of the transcription factor, HcaR. The more amounts of HcaR means activation of *hca* and *mhp* operons that contribute to the degradation of PNP in *E. coli* MG1655 (pSDP10) cells (Chakka et al., 2015). A diagrammatic repression indicating the Orf306 dependent propionate induced genetic reprogramming that facilitates the growth of *E. coli* MG1655 (pSDP10) cells on alternate carbon sources like PNP (Fig. 5. 2. 6).



Summary



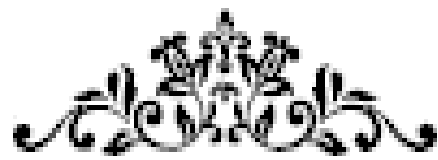
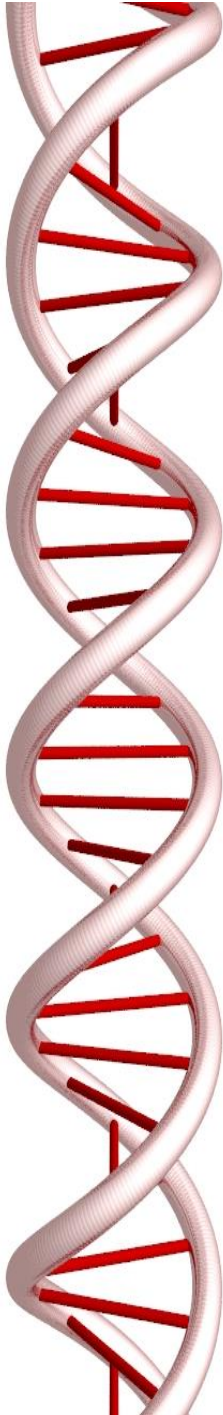
6. Summary

Self-transmissible plasmids, genomic islands, such as integrative mobilizable elements (IMEs). Integrative Conjugative Elements (ICEs) help bacteria to acquire genetic information pertaining to pathogenicity, drug resistance and degradation of toxic and recalcitrant chemicals. The IME having information for organophosphate degradation has showed presence of an ORF, *orf306*, along with organophosphate degradation (*opd*) gene that codes for a phosphotriesterase (PTE). The bacterial triesterases are highly conserved among soil bacteria. They hydrolyse triester bond found in a variety of OP insecticides and nerve agents. During biodegradation of OP insecticides, several recalcitrants and toxic compounds such as nitrophenols are generated as catabolic intermediates. The *orf306* found associated with *opd* gene, when expressed in *E. coli* MG1655 enabled the cells to grow using *p*-nitrophenol (PNP) as a source of carbon. The transcriptome analysis of *E. coli* cells, expressing Orf306 revealed the existence of a shift in carbon catabolic pathways (Chakka et al., 2015). However, no studies were conducted to establish a link between Orf306 and an observed shift in carbon catabolic pathways. The work presented in the dissertation is an attempt to establish a link between Orf306 and observed catabolic shift.

The first chapter described *in silico* studies used to find out Orf306 responsive small RNAs (sRNAs). These sRNAs were then screened against mRNAs coding for glycolysis and TCA cycle enzymes which are found to be down-regulated in the presence of Orf306. In order to know the existence of seed regions between sRNAs and their target mRNAs, the *in silico* tools were employed. The *in silico* prediction have indicated interactions between Orf306 responsive sRNAs and mRNAs coding for key glycolytic/TCA cycle enzymes and transcription factors that stimulate alternate carbon catabolic pathways. The sRNA, seco10054A showed good seed region with mRNA coding for Lpd subunit of pyruvate dehydrogenase and α -ketoglutarate dehydrogenase. Likewise, the sRNA, seco11136A has shown to have sequence complementarity with mRNA coding for transcription factor HcaR. The second and third chapter describe experimental procedures employed to validate the *in silico* predicted interactions. Using a range of molecular tools highly convincing evidence were gathered in support of predicted interactions between seco10054A and *lpd* mRNA and between seco11136A and *hcaR* mRNA.

Further, experimental evidence was also presented in the third chapter on regulation of genes coding these novel sRNAs. The promoter assays and northern blot analysis have shown

existence of clear link between *orf306* and *seco10054A* induction and *seco11136A* repression. This kind of Orf306 dependent reciprocal regulation of these two novel sRNA coding genes and their involvement in translation inhibition of respective target mRNAs, explain the kind of shift observed in Orf306 dependent carbon catabolic pathways in *E. coli* with reasonable evidence. However, the precise reason for Orf306 dependent induction/repression of these sRNAs were not known through these experimental design. Taking the lead from the previous studies which have shown esterase/lipase activity for Orf306, the lipidome analysis was performed for *E. coli* MG1655 cells grown in Orf306 positive and negative background. The total lipid analysis revealed Orf306 dependent accumulation of propionate. Up-regulation of propionate degradation pathways in the presence of Orf306 has supported clearly endogenous propionate generation due to its esterase/lipase activity. Based on these two leads, further experiments were designed to know the influence of propionate on sRNA, *seco10054A*, and sRNA, *seco11136A* expression. The promoter fusions have clearly shown propionate dependent induction of sRNA, *seco10054A*, and repression of sRNA, *seco11136A* expression. The experimental evidence generated in this study has clearly shown involvement of propionate dependent signaling mechanism behind reciprocal regulation of *seco10054A*, and *seco11136A* coding for sRNAs, *seco10054A* and *seco11136A*. Out of these two sRNAs, *seco10054A* promotes down-regulation of glycolysis and TCA cycle, whereas repression of *seco11136A* activates *hca* operon as it is shown to inhibit the translation of mRNA coding for transcription factor HcaR. Finally, the study clearly indicates that the propionate generated due to Orf306 dependent lipase activity as a key signaling molecule for the observed catabolic shift in *E. coli* MG1655 (pSDP10) cells.



References



7. References

1. Alekshun, M. N. & Levy, S. B. The mar regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol.* 7, 410–3 (1999).
2. Alekshun, M.N. & Levy, S.B. Regulation of chromosomally mediated multiple antibiotic resistance: the mar regulon. *Antimicrob Agents Chemother* **10**: 2067 –2075 (1997).
3. Altuvia, S. Identification of bacterial small non-coding RNAs: experimental approaches. *Curr. Opin. Microbiol.* 10, 257–61 (2007).
4. Altuvia, S. The *Escherichia coli* OxyS regulatory RNA represses *fhlA* translation by blocking ribosome binding. *EMBO J.* 17, 6069–6075 (1998).
5. Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L. & Storz, G. A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* 90, 43–53 (1997).
6. Andersen, J. *et al.* The isolation and characterization of RNA coded by the micF gene in *Escherichia coli*. *Nucleic Acids Res.* 15, 2089–101 (1987).
7. Andersen, J., Forst, S. A., Zhao, K., Inouye, M. & Delihias, N. The function of micF RNA. micF RNA is a major factor in the thermal regulation of OmpF protein in *Escherichia coli*. *J. Biol. Chem.* 264, 17961–70 (1989).
8. Argaman, L. *et al.* Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* 11, 941–50 (2001).
9. Baek, J. H. & Lee, S. Y. Novel gene members in the Pho regulon of *Escherichia coli*. *FEMS Microbiol. Lett.* 264, 104–9 (2006).
10. Bandyra, K. J. *et al.* The Seed Region of a Small RNA Drives the Controlled Destruction of the Target mRNA by the Endoribonuclease RNase E. *Mol. Cell* 47, 943–953 (2012).
11. Beisel, C. L. & Storz, G. Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol. Rev.* 34, 866–882 (2010).
12. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–62 (1997).
13. Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37, 911–917 (1959).

14. Blower, T. R. *et al.* A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nat. Struct. Mol. Biol.* 18, 185–90 (2011).
15. Bobrovskyy, M. & Vanderpool, C. K. Regulation of Bacterial Metabolism by Small RNAs Using Diverse Mechanisms. *Annu. Rev. Genet.* 47, 209–232 (2013).
16. Bobrovskyy, M. & Vanderpool, C. K. The small RNA SgrS: roles in metabolism and pathogenesis of enteric bacteria. *Front. Cell. Infect. Microbiol.* 4, 1–8 (2014).
17. Bossi, L. & Figueroa-Bossi, N. A small RNA downregulates LamB maltoporin in *Salmonella*. *Mol. Microbiol.* 65, 799–810 (2007).
18. Bouché, F. & Bouché, J. P. Genetic evidence that DicF, a second division inhibitor encoded by the *Escherichia coli* dicB operon, is probably RNA. *Mol. Microbiol.* 3, 991–994 (1989).
19. Brantl, S. & Jahn, N. sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol. Rev.* 39, 413–427 (2015).
20. Brantl, S. Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr. Opin. Microbiol.* 10, 102–9 (2007).
21. Busch, A., Richter, A. S. & Backofen, R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 24, 2849–56 (2008).
22. Capra, E. J. & Laub, M. T. Evolution of Two-Component Signal Transduction Systems. *Annu. Rev. Microbiol.* 66, 325–347 (2012).
23. Chakka, D. *et al.* The organophosphate degradation (*opd*) island-borne esterase-induced metabolic diversion in *Escherichia coli* and its influence on *p*-nitrophenol degradation. *J. Biol. Chem.* 290, 29920–29930 (2015).
24. Chen, S. *et al.* A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *Biosystems.* 65, 157–77 (2012).
25. De Almeida Ribeiro, E. *et al.* Structural flexibility of RNA as molecular basis for Hfq chaperone function. *Nucleic Acids Res.* 40, 8072–8084 (2012).
26. Deighan, P., Free, A. & Dorman, C. J. A role for the *Escherichia coli* H-NS-like protein StpA in OmpF porin expression through modulation of micF RNA stability. *Mol. Microbiol.* 38, 126–39 (2000).

27. Desnoyers, G., Bouchard, M.-P. & Massé, E. New insights into small RNA-dependent translational regulation in prokaryotes. *Trends Genet.* 29, 92–98 (2013).
28. Diaz, E., Ferrandez, A., Prieto, M. A. & Garcia, J. L. Biodegradation of Aromatic Compounds by *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 65, 523–569 (2001).
29. Du Toit, A. Bacterial transcription: Turning the switch on gene expression. *Nat. Rev. Microbiol.* 12, 657–657 (2014).
30. Eggenhofer, F., Tafer, H., Stadler, P. F. & Hofacker, I. L. RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Res.* 39, W149–W154 (2011).
31. Fassler, J. S. & West, A. H. Histidine Phosphotransfer Proteins in Fungal Two-Component Signal Transduction Pathways. *Eukaryot. Cell* 12, 1052–1060 (2013).
32. Ferrario, M. *et al.* The leucine-responsive regulatory protein of *Escherichia coli* negatively regulates transcription of *ompC* and *micF* and positively regulates translation of *ompF*. *J. Bacteriol.* 177, 103–113 (1995).
33. Fozo, E. M. New type I toxin-antitoxin families from ‘wild’ and laboratory strains of *E. coli*: Ibs-Sib, ShoB-OhsC and Zor-Orz. *RNA Biol.* 9, 1504–12 (2012).
34. Fozo, E. M., Hemm, M. R. & Storz, G. Small Toxic Proteins and the Antisense RNAs That Repress Them. *Microbiol. Mol. Biol. Rev.* 72, 579–589 (2008).
35. Garst, A. D., Edwards, A. L. & Batey, R. T. Riboswitches: structures and mechanisms. *Cold Spring Harb. Perspect. Biol.* 3, a003533–a003533 (2011).
36. Georg, J. & Hess, W. R. cis-Antisense RNA, Another level of gene regulation in bacteria. *Microbiol. Mol. Biol. Rev.* 75, 286–300 (2011).
37. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* 6, 613–624 (2008).
38. Gorla, P., Pandey, J. P., Parthasarathy, S., Merrick, M. & Siddavattam, D. Organophosphate Hydrolase in *Brevundimonas diminuta* Is Targeted to the Periplasmic Face of the Inner Membrane by the Twin Arginine Translocation Pathway. *J. Bacteriol.* 191, 6292–6299 (2009).
39. Gottesman, S. & Masse, E. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. 99, (2002).

40. Gottesman, S. The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu.Rev.Microbiol.* 58, 303–328 (2004).
41. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557–580 (1983).
42. Hoe, C.-H., Raabe, C. A., Rozhdestvensky, T. S. & Tang, T.-H. Bacterial sRNAs: Regulation in stress. *Int. J. Med. Microbiol.* 303, 217–229 (2013).
43. Huang, L., Tsui, P. & Freundlich, M. Integration host factor is a negative effector of in vivo and in vitro expression of ompC in *Escherichia coli*. *J. Bacteriol.* 172, 5293–8 (1990).
44. Ikemura, T. & Dahlberg, J. E. Small ribonucleic acids of *Escherichia coli*. II. Noncoordinate accumulation during stringent control. *J. Biol. Chem.* 248, 5033–41 (1973).
45. Izard, T. *et al.* Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1240–5 (1999).
46. Kamireddy, S. CRYSONS, the CryIEC toxin encapsulated mesoporous SiO₂ nanoparticles, as an effective biopesticide against *Spodoptera litura*. (University of Hyderabad, 2014).
47. Kawano, M. Divergently overlapping cis-encoded antisense RNA regulating toxin-antitoxin systems from *E. coli*: hok/sok, ldr/rdl, symE/symR. *RNA Biol.* 9, 1520–7 (2012).
48. Kawano, M., Oshima, T., Kasai, H. & Mori, H. Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in *Escherichia coli*. *Mol. Microbiol.* 45, 333–349 (2002).
49. Keseler, I. M. *et al.* EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* 41, D605–D612 (2013).
50. Khajamohiddin, S. *et al.* A novel meta-cleavage product hydrolase from *Flavobacterium sp.* ATCC27551. *Biochem. Biophys. Res. Commun.* 351, 675–681 (2006).
51. Kim, S. K. *et al.* Dual transcriptional regulation of the *Escherichia coli* phosphate-starvation-inducible *psiE* gene of the phosphate regulon by PhoB and the cyclic AMP (cAMP)-cAMP receptor protein complex. *J. Bacteriol.* 182, 5596–9 (2000).

51. Kovárová, K., Käch, a, Chaloupka, V. & Egli, T. Cultivation of *Escherichia coli* with mixtures of 3-phenylpropionic acid and glucose: dynamics of growth and substrate consumption. *Biodegradation* 7, 445–53 (1996).
52. Lamarche, M. G., Wanner, B. L., Crépin, S. & Harel, J. The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis. *FEMS Microbiol. Rev.* 32, 461–73 (2008).
53. Lamichhane, G., Arnvig, K. B. & McDonough, K. a. Definition and annotation of (myco)bacterial non-coding RNA. *Tuberculosis* 93, 26–29 (2013).
54. Laub, M. T. & Goulian, M. Specificity in Two-Component Signal Transduction Pathways. *Annu. Rev. Genet.* 41, 121–145 (2007).
55. Lease, R. A. & Belfort, M. A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9919–24 (2000).
56. Li, M., Ho, P. Y., Yao, S. & Shimizu, K. Effect of *lpdA* gene knockout on the metabolism in *Escherichia coli* based on enzyme activities, intracellular metabolite concentrations and metabolic flux analysis by ¹³C-labeling experiments. *J. Biotechnol.* 122, 254–66 (2006).
57. Li, W., Ying, X., Lu, Q. & Chen, L. Predicting snas and their targets in bacteria. *Genomics. Proteomics Bioinformatics* 10, 276–284 (2012).
58. Liu, M. Y. *et al.* The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *J. Biol. Chem.* **272**, 17502–10 (1997).
59. López-Maury, L., Marguerat, S. & Bähler, J. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat. Rev. Genet.* 9, 583–93 (2008).
60. Majdalani, N., Chen, S., Murrow, J., St John, K. & Gottesman, S. Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol. Microbiol.* 39, 1382–94 (2001).
61. Majdalani, N., Cunnig, C., Sledjeski, D., Elliott, T. & Gottesman, S. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc. Natl. Acad. Sci. U. S. A.* 95, 12462–7 (1998).
62. Majdalani, N., Hernandez, D. & Gottesman, S. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. 46, 813–826 (2002).

63. Maki, K., Morita, T., Otaka, H. & Aiba, H. A minimal base-pairing region of a bacterial small RNA SgrS required for translational repression of *ptsG* mRNA. *Mol. Microbiol.* 76, 782–92 (2010).
64. Manso, I. *et al.* 3-Hydroxyphenylpropionate and phenylpropionate are synergistic activators of the MhpR transcriptional regulator from *Escherichia coli*. *J. Biol. Chem.* 284, 21218–28 (2009).
65. Marzan, L. & Shimizu, K. Metabolic regulation of *Escherichia coli* and its *phoB* and *phoR* genes knockout mutants under phosphate and nitrogen limitations as well as at acidic condition. *Microb. Cell Fact.* 10, 39 (2011).
66. Massé, E. & Gottesman, S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 4620–5 (2002).
67. Matsuoka, Y. & Shimizu, K. Metabolic regulation of *Escherichia coli* cultivated under anaerobic and aerobic conditions in response to the specific pathway gene knockouts. *Adv. Biosci. Biotechnol.* 04, 455–468 (2013).
68. McCleary, W. R. The activation of PhoB by acetylphosphate. *Mol. Microbiol.* 20, 1155–63 (1996).
69. Meng, L. M., Kilstrup, M. & Nygaard, P. Autoregulation of PurR repressor synthesis and involvement of *purR* in the regulation of *purB*, *purC*, *purL*, *purMN* and *guaBA* expression in *Escherichia coli*. *Eur. J. Biochem.* 187, 373–9 (1990).
70. Miller, J. H. *Experiments in molecular genetics*. (CSHL Press, 1972).
71. Mizuno, T., Chou, M. Y. & Inouye, M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci. U. S. A.* 81, 1966–70 (1984).
72. Møller, T., Franch, T., Udesen, C., Gerdes, K. & Valentin-Hansen, P. Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev.* 16, 1696–706 (2002).
73. Morales, V. M., Bäckman, A. & Bagdasarian, M. A series of wide-host-range low-copy-number vectors that allow direct screening for recombinants. *Gene* 97, 39–47 (1991).
74. Mulbry, W. W., Kearney, P. C., Nelson, J. O. & Karns, J. S. Physical comparison of parathion hydrolase plasmids from *Pseudomonas diminuta* and *Flavobacterium sp.* *Plasmid* 18, 173–7 (1987).

75. Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A. & Gautheret, D. ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol.* 8, 11–13 (2011).
76. Nickerson, C. A., Ott, C. M., Wilson, J. W., Ramamurthy, R. & Pierson, D. L. Microbial Responses to Microgravity and Other Low-Shear Environments. *Microbiol. Mol. Biol. Rev.* 68, 345–361 (2004).
77. Nudler, E. & Mironov, A. S. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.* 29, 11–7 (2004).
78. Opdyke, J. A., Kang, J.-G. & Storz, G. GadY, a Small-RNA Regulator of Acid Response Genes in *Escherichia coli*. *J. Bacteriol.* 186, 6698–6705 (2004).
79. Ortet, P., Whitworth, D. E., Santaella, C., Achouak, W. & Barakat, M. P2CS: updates of the prokaryotic two-component systems database. *Nucleic Acids Res.* 43, D536–D541 (2015).
80. Pakala, S. B. *et al.* Biodegradation of methyl parathion and p-nitrophenol: evidence for the presence of a p-nitrophenol 2-hydroxylase in a Gram-negative *Serratia sp.* strain DS001. *Appl. Microbiol. Biotechnol.* 73, 1452–1462 (2007).
81. Palmer, L. D., Paxhia, M. D. & Downs, D. M. Induction of the sugar-phosphate stress response allows *Saccharomyces cerevisiae* 2-methyl-4-amino-5-hydroxymethylpyrimidine phosphate synthase to function in *Salmonella enterica*. *J. Bacteriol.* 197, 3554–3562 (2015).
82. Pandeeti, E. V. P. *et al.* Multiple mechanisms contribute to lateral transfer of an organophosphate degradation (*opd*) island in *Sphingobium fuliginis* ATCC 27551. *G3* 2, 1541–54 (2012).
83. Pandeeti, E. V. P., Chakka, D., Pandey, J. P. & Siddavattam, D. Indigenous organophosphate-degrading (*opd*) plasmid pCMS1 of *Brevundimonas diminuta* is self-transmissible and plays a key role in horizontal mobility of the *opd* gene. *Plasmid* 65, 226–31 (2011).
84. Papenfort, K. & Vanderpool, C. K. Target activation by regulatory RNAs in bacteria. *FEMS Microbiol. Rev.* 39, 362–378 (2015).
85. Papenfort, K., Bouvier, M., Mika, F., Sharma, C. M. & Vogel, J. Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20435–40 (2010).

86. Papenfort, K., Podkaminski, D., Hinton, J. C. D. & Vogel, J. The ancestral SgrS RNA discriminates horizontally acquired *Salmonella* mRNAs through a single G-U wobble pair. *Proc. Natl. Acad. Sci. U. S. A.* 109, E757–64 (2012).
87. Papenfort, K., Sun, Y., Miyakoshi, M., Vanderpool, C. K. & Vogel, J. Small RNA-Mediated Activation of Sugar Phosphatase mRNA Regulates Glucose Homeostasis. *Cell* 153, 426–437 (2013).
88. Peer, A. & Margalit, H. Accessibility and Evolutionary Conservation Mark Bacterial Small-RNA Target-Binding Regions. *J. Bacteriol.* 193, 1690–1701 (2011).
89. Pérez-Rueda, E., Tenorio-Salgado, S., Huerta-Saquero, A., Balderas-Martínez, Y. I. & Moreno-Hagelsieb, G. The functional landscape bound to the transcription factors of *Escherichia coli* K-12. *Comput. Biol. Chem.* 58, 93–103 (2015).
90. Pettit, F. H. & Reed, L. J. Alpha-keto acid dehydrogenase complexes. 8. Comparison of dihydrolipoyl dehydrogenases from pyruvate and alpha-ketoglutarate dehydrogenase complexes of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 58, 1126–30 (1967).
91. Polayes, D. A., Rice, P. W., Garner, M. M. & Dahlberg, J. E. Cyclic AMP-cyclic AMP receptor protein as a repressor of transcription of the *spf* gene of *Escherichia coli*. *J. Bacteriol.* 170, 3110–4 (1988).
92. Prévost, K. *et al.* The small RNA RyhB activates the translation of *shiA* mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Mol. Microbiol.* 64, 1260–1273 (2007).
93. Rice, J. B. & Vanderpool, C. K. The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. *Nucleic Acids Res.* 39, 3806–19 (2011).
94. Rice, J. B., Balasubramanian, D. & Vanderpool, C. K. Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2691–8 (2012).
95. Richter, A. & Backofen, R. Accessibility and conservation: General features of bacterial small RNA–mRNA interactions? *RNA Biol.* 9, 954–965 (2012).
96. Sahagan, B. G. & Dahlberg, J. E. A small, unstable RNA molecule of *Escherichia coli*: spot 42 RNA. I. Nucleotide sequence analysis. *J. Mol. Biol.* 131, 573–92 (1979).
97. Salim, N. N., Faner, M. A., Philip, J. A. & Feig, A. L. Requirement of upstream Hfq-binding (ARN)_x elements in *glmS* and the Hfq C-terminal region for GlmS upregulation by sRNAs GlmZ and GlmY. *Nucleic Acids Res.* 40, 8021–8032 (2012).

98. Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
99. Schimel, J., Balser, T. C. & Wallenstein, M. Microbial stress-response physiology and its implications for ecosystem function. *Ecology* 88, 1386–94 (2007).
100. Sharma, C. M., Darfeuille, F., Plantinga, T. H. & Vogel, J. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.* 21, 2804–2817 (2007).
101. Shimoni, Y. *et al.* Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol. Syst. Biol.* 3, (2007).
102. Shinohara, A. *et al.* Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. *BMC Genomics* 12, 428 (2011).
103. Siddavattam, D., Khajamohiddin, S., Manavathi, B., Pakala, S. B. & Merrick, M. Transposon-like organization of the plasmid-borne organophosphate degradation (*opd*) gene cluster found in *Flavobacterium sp.* *Appl. Environ. Microbiol.* 69, 2533–9 (2003).
104. Siddavattam, D., Khajamohiddin, S., Manavathi, B., Pakala, S. B. & Merrick, M. Transposon-like organization of the plasmid-borne organophosphate degradation (*opd*) gene cluster found in *Flavobacterium sp.* *Appl. Environ. Microbiol.* 69, 2533–9 (2003).
105. Simons, R. W., Houman, F. & Kleckner, N. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* 53, 85–96 (1987).
106. Sledjeski, D. D., Gupta, A. & Gottesman, S. The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J.* 15, 3993–4000 (1996).
107. Solov'yev V, S. A. in p. 61–78 (Nova Science Publishers, 2011).
108. Somara, S., Manavathi, B., Tebb, C.C. & Siddavattam, D. Localisation of identical organophosphorus pesticide degrading (*opd*) genes on genetically dissimilar indigenous plasmids of soil bacteria: PCR amplification, cloning and sequencing of *opd* gene from *Flavobacterium balustinum*,. *Indian J Exp Biol.* 40, 774-779 (2002).
109. Soper, T., Mandin, P., Majdalani, N., Gottesman, S. & Woodson, S. A. Positive regulation by small RNAs and the role of Hfq. *Proc. Natl. Acad. Sci. U. S. A.* 107, 9602–7 (2010).

110. Spaink, H. P., Okker, R. J. H., Wijffelman, C. A., Pees, E. & Lugtenberg, B. J. J. Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. *Plant Mol. Biol.* 9, 27–39 (1987).
111. Storz, G., Opdyke, J. A. & Zhang, A. Controlling mRNA stability and translation with small, noncoding RNAs. *Curr. Opin. Microbiol.* 7, 140–4 (2004).
112. Storz, G., Vogel, J. & Wassarman, K. M. Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Mol. Cell* 43, 880–891 (2011).
113. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–30 (1986).
114. Swetha, K. CRYSONS, the CryIEC toxin encapsulated mesoporous SiO₂ nanoparticles, as an effective biopesticide against *Spodoptera litura*. A thesis submitted to University of Hyderabad. (2015).
115. Thomason, M. K. & Storz, G. Bacterial Antisense RNAs: How Many Are There, and What Are They Doing? . *Annu. Rev. Genet.* 44, 167–188 (2010).
116. Tjaden, B. *et al.* Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucleic Acids Res.* 30, 3732–8 (2002).
117. Tomizawa, J. & Itoh, T. Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. *Proc. Natl. Acad. Sci. U. S. A.* 78, 6096–100 (1981).
118. Tomizawa, J. Control of ColE1 plasmid replication. Intermediates in the binding of RNA I and RNA II. *J. Mol. Biol.* 212, 683–94 (1990).
119. Tucker, B. J. & Breaker, R. R. Riboswitches as versatile gene control elements. *Curr. Opin. Struct. Biol.* 15, 342–8 (2005).
120. Turlin, E., Perrotte-piquemal, M., Danchin, A. & Biville, F. Regulation of the early steps of 3-phenylpropionate catabolism in *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* 3, 127–33 (2001).
121. Urban, J. H. & Vogel, J. Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res.* 35, 1018–1037 (2007).
122. Urbanowski, M. L., Stauffer, L. T. & Stauffer, G. V. The *gcvB* gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*. *Mol. Microbiol.* 37, 856–68 (2000).

123. Vanderpool, C. K. & Gottesman, S. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol. Microbiol.* 54, 1076–89 (2004).
124. Vogel, J. & Luisi, B. F. Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9, 578–89 (2011).
125. Vogel, J. & Wagner, E. G. H. Target identification of small noncoding RNAs in bacteria. *Curr. Opin. Microbiol.* 10, 262–70 (2007).
126. Vogel, J. A rough guide to the non-coding RNA world of *Salmonella*. *Mol. Microbiol.* 71, 1–11 (2009).
127. Wagner, E. G. & Simons, R. W. Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* 48, 713–42 (1994).
128. Wagner, E. G. H. Cycling of RNAs on Hfq. *RNA Biol.* 10, 619–26 (2013).
129. Wanner, B.L. Gene regulation by phosphate in enteric bacteria, *J. Cell. Biochem.* 51, 47-54 (1993).
130. Wassarman, K. M. & Storz, G. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101, 613–23 (2000).
131. Wassarman, K. M. 6S RNA: a small RNA regulator of transcription. *Curr. Opin. Microbiol.* 10, 164–8 (2007).
132. Weilbacher, T. *et al.* A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol. Microbiol.* 48, 657–70 (2003).
133. Wilderman, P. J. *et al.* Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis. *Proc. Natl. Acad. Sci.* 101, 9792–9797 (2004).
134. Wilkinson, K. D. & Williams, C. H. NADH inhibition and NAD activation of *Escherichia coli* lipoamide dehydrogenase catalyzing the NADH-lipoamide reaction. *J. Biol. Chem.* 256, 2307–2314 (1981).
135. Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419, 952–6 (2002).
136. Wolf, L., Silander, O. K. & van Nimwegen, E. Expression noise facilitates the evolution of gene regulation. *Elife* 4, (2015).

137. Yang, C. *et al.* Genome-wide PhoB binding and gene expression profiles reveal the hierarchical gene regulatory network of phosphate starvation in *Escherichia coli*. *PLoS One* 7, e47314 (2012).
138. Yoshimi, A., Kojima, K., Takano, Y. & Tanaka, C. Group iii histidine kinase is a positive regulator of hog1-type mitogen-activated protein kinase in filamentous fungi. *Eukaryot. Cell* 4, 1820–1828 (2005).
139. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415 (2003).