

Identifying *Escherichia coli* Factors that Selectively Bind the mRNA of Secreted Proteins

by

Walaa Abdulhadi H. Hakeem

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science (MSc) in Biology

The Faculty of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

© Walaa Hakeem, 2014

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE
Laurentian Université/Université Laurentienne
Faculty of Graduate Studies/Faculté des études supérieures

Title of Thesis
Titre de la thèse Identifying Escherichia Coli Factors That Selectively Bind the mRNA of Secreted Proteins

Name of Candidate
Nom du candidat Hakeem, Walaa

Degree
Diplôme Master of Science

Department/Program
Département/Programme Biology Date of Defence
Date de la soutenance September 26, 2014

APPROVED/APPROUVÉ

Thesis Examiners/Examineurs de thèse:

Dr. Mazen Saleh
(Supervisor/Directeur(trice) de thèse)

Dr. Robert Lafrenie
(Committee member/Membre du comité)

Dr. Kabwe Nkongolo
(Committee member/Membre du comité)

Dr. Carolyn Kapron
(External Examiner/Examineur externe)

Approved for the Faculty of Graduate Studies
Approuvé pour la Faculté des études supérieures
Dr. David Lesbarrères
M. David Lesbarrères
Acting Dean, Faculty of Graduate Studies
Doyen intérimaire, Faculté des études supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

I, **Walaa Hakeem**, hereby grant to Laurentian University and/or its agents the non-exclusive license to archive and make accessible my thesis, dissertation, or project report in whole or in part in all forms of media, now or for the duration of my copyright ownership. I retain all other ownership rights to the copyright of the thesis, dissertation or project report. I also reserve the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that this copy is being made available in this form by the authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws without written authority from the copyright owner.

Abstract

It is well accepted that the majority of secreted proteins are targeted to the secretory pathway through amino acid signal sequences located at the N-terminus of the pre-protein during the initial stages of translation on the ribosome. This is true in both eukaryotic and prokaryotic systems. These signal sequences display distinct structural features such as charged/hydrophilic residues at each termini and a continuous hydrophobic stretch of amino acids between them. The signal-recognition particle (SRP), which targets the ribosome to the membrane translocon for secretion of the protein, is hypothesized to recognize and these features and therefore differentiate these proteins from non-secretory proteins. Recent work in other laboratories has suggested a role for the mRNA itself, rather than only the amino acid sequence of the N-terminus of the pre-protein as playing a role in targeting the pre-protein-ribosome complex to the translocon within the membrane. To test this hypothesis, direct interaction between mRNAs encoding secreted proteins and the *E. coli* SRP equivalent (Ffh) was pursued using pull down assays. The mRNA's used as bait corresponded to the N-terminal 40 amino acids of secreted and cytosolic proteins including periplasmic propyl isomerase chaperone SurA (as a model secretory protein with a cleavable signal peptide) and the cytoplasmic protein 3-isopropyl malate dehydrogenase (IsodH). Additionally, the mRNA of two other proteins, PhoA (secreted) and GMP (cytoplasmic), were used but in these mRNA the 5' UTR were also included in case these regions were involved in SRP recognition. Following extensive optimizations and modifications of these experiments, the Ffh protein (the *Escherichia coli* SRP homolog) could not be isolated from cytoplasmic extracts of *E. coli* with the pull down assays. One interesting finding however was that the mRNA of the IsodH protein was pulled down using its cognate mRNA transcript as

bait. This implies a role for this enzyme in regulating its own levels in the cell by binding to and potentially modulating its translation. Other factors involved in DNA and RNA binding were also isolated and include RNase and ribosomal proteins, amongst others. It can therefore be concluded that under these experimental conditions, the mRNA hypothesis for targeting protein secretion could not be supported.

Acknowledgments

Neither words nor actions can express my gratitude to my God and to all those who supported me during this scientific journey. This thesis would not have been possible without the help of my supervisor, Dr. Mazen Saleh and his guidance throughout the thesis planning, lab work and writing.

I would like to extend my gratitude to my committee members, Dr. Robert Lafrenie and Dr. Kabwe Nkongolo for believing in me since the start of my study and for encouraging me to think positive.

Also, I would like to express a special thanks to Ministry of High Education Saudi Arabia, which support me financially to complete this project.

It would not have been possible to write this thesis without the help and support of kind people around me, my parents Abdulhadi Hakeem and Eman Mahmoud and my whole family and friends. My appreciation also goes to Mr. Paul Michael (Ph.D. candidate) for his patience, advice, and technical assistance.

Table of Contents

Abstract	iii
Acknowledgments.....	v
Table of Contents	vi
List of Tables	x
List of Figures	xi
List of Abbreviation.....	xii
Chapter 1: Literature Review	1
1.1 Protein Translation and its Regulation in <i>Escherichia coli</i>	1
1.2 Sec-dependent Secretion and The Signal Peptide Hypothesis.....	4
1.3 Logistical Challenges for the Cell with the Present Signal Peptide Hypothesis.....	5
1.4 The mRNA Hypothesis in Describing Membrane Targeting of the Pre-protein/Ribosome Complex	8
1.5 Evidence of Specific mRNA Recognition and Targeting in Eukaryotic System	9
1.6 Evidence of mRNA Targeting in Prokaryotes	12
1.7 Research Rationale.....	14
1.8 Objectives	15
Chapter 2: Experimental Methods	17
2.1 Bacterial Cultures.....	17
2.2 Extraction of Genomic DNA	17

2.3	Quantitation of Nucleic Acids	18
2.4	Genomic DNA Amplification.....	18
2.5	Transcription by T7 RNA Polymerase	19
2.6	Agarose Gel Electrophoresis.....	19
2.7	Biotin Labeling of mRNA	21
2.8	Chromogenic Detection of Biotinylated mRNA and Evaluation of Labeling Efficiency	21
2.9	Protein Extraction	22
2.10	Protein Assay	22
2.11	UV Crosslinking	23
2.12	Pull-Down Assay	23
2.13	Effect of Inhibitors of mRNA Degradation on The Pull Down Assay	23
2.14	Protein Capture Using Nylon Membranes.....	24
2.15	RNA–Protein Interaction Assay	26
2.16	SDS PAGE.....	26
2.17	Protein Identification Using Matrix-Assisted Laser Desorption Time-of-Flight (MALDI-TOF) Mass Spectrometry.....	27
Chapter 3:	Results	29
3.1	PCR Amplification of Genomic DNA.....	29
3.2	Transcription of Amplicons	29
3.3	Quality of Transcription Product	32

3.4 mRNA Structure of Different Transcripts	32
3.5 Estimation of Labeling Efficiency	37
3.6 Detection of Biotinylated RNA	37
3.7 The Streptavidin – Biotin System	37
3.8 Pull-Down Assay	41
3.9 Effect of Inhibitors Against RNA Degradation	46
3.10 Protein Capture Using Nylon Membranes	46
3.11 RNA-Protein Interaction Assay	46
3.12 Protein Identification by Mass Spectrometry	49
Chapter 4: Discussion	54
4.1 mRNA Structure of Different Transcripts	55
4.2 Factors Identified Through The Pull-Down Assay	56
4.3 Different Transcripts Construct (5' UTRs of phoA and gMP genes).....	59
4.4 The Streptavidin-Magnetic Beads Capacity	62
4.5 The Dynabeads Beads Blocking	62
4.6 Potential Degradation of mRNA in the Procedures	63
Chapter 5: Conclusion.....	65
5.1 Limitations of Study	65
5.2 Future Directions	66
References.....	67

Appendix..... 77

List of Tables

Table 1	A summary of the factors involved in protein synthesis in prokaryotes (Benelli, et al 2009).	3
Table 2	Primers for the PCR reactions.	20
Table 3	mRNA labeling efficiency.	38
Table 4	Dynabeads streptavidin-biotin assay.	40
Table 5	Proteins identified from the pull down assay with surA mRNA using MALDI-MS identification.	51
Table 6	Proteins identified from pull down assay of isodH mRNA using MALDI-MS identification.	52
Table 7	Proteins identified from pull down assay of phoA mRNA using MALDI-MS identification.	53

List of Figures

Figure 1	A diagram depicting the overall process of protein translation in prokaryotes cells.	2
Figure 2	Diagram of Sec dependent secretion and the signal hypothesis.....	6
Figure 3	An alternate model of targeting secreted proteins to the Sec machinery.	16
Figure 4	Depiction of the pull down assay.	25
Figure 5	Amplification of surA and isodH genes fragments.	30
Figure 6	Amplification of phoA and gMP genes fragments.....	31
Figure 7	<i>In vitro</i> RNA transcriptions from surA and isodH genes fragments.	33
Figure 8	Confirmation of the presence of transcripts for phoA and gMP genes fragments using MEGAscript kit.....	34
Figure 9	Analysis of RNA transcription products using the Agilent Bio-analyzers.....	35
Figure 10	Simulated secondary structures of surA (A), isodH (B), phoA (C), and gMP (D). ..	36
Figure 11	Detection of biotinylated mRNA on PVDF membrane.	39
Figure 12	SDS - PAGE gel from the pull down assay of surA and isodH transcripts.....	42
Figure 13	Silver stained gels of proteins from the pull down assay using biotinylated probes from surA and isodH baits.	43
Figure 14	SDS -PAGE gel from optimizing the pull down assay of surA and isodH transcripts.	44
Figure 15	Separation of proteins from the pull down assay using phoA and gMP transcripts as baits.	45
Figure 16	The effect of inhibitors on pull-down assay with phoA and gMP mRNAs.	47
Figure 17	RNA- protein interaction assay.	48

List of Abbreviation

μ J	Microjoule
μ l	Microliter
μ g	Microgram
μ M	Micromolar
Abs	Absorbance
AmBic	Ammonium bicarbonate
AMP	Adenosine Monophosphate
ATP	Adenosine Triphosphate
bp	Base pairs
C	Carboxyl-terminal of protein
cm	Centimeters
DEPC	Diethylpyrocarbonate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetracetic acid
<i>E. coli</i>	Escherichia coli
DNA	Deoxyribonucleic acid
h	Hours
IM	Inner membrane
kDa	Kilo Dalton
L	Liter
LB	Luria Bertani
M	Molar
mg	Milligram
min	Minutes
ml	Milliliters
mM	Millimolar
MP	Membrane protein
mRNA	messenger RNA
MW	Molecular weight
N	Amino-terminal of protein
nmol	Nanomole
NCBI	National Center for Biotechnology Information
OM	Outer membrane
PAGE	polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pmol	Pico mole
RBP	RNA binding protein
RNA	Ribonucleic acid
RPM	Rotations per minute
RT	Room temperature
sec	Second
Sec	Secretory
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SRP	Signal recognition particle

Ta	Annealing temperature
TAE	Tris-acetate-EDTA buffer
TEMED	Tetramethylethylenediamine
Tm	Melting temperature
Tris	Tris(hydroxymethyl) aminomethane
U	Unit
v/v	volume per volume

Chapter 1: Literature Review

1.1 Protein Translation and its Regulation in *Escherichia coli*

Within all cells, protein biosynthesis usually occurs on a specialized super molecular complex called the ribosome via a translation mechanism. In this event, the genetic information that encodes these proteins from the DNA is carried to the ribosome through mRNA as shown in Figure 1 (Laursen et al., 2005). Decades ago, translation factors were acknowledged as highly conserved factors that are required in both eukaryotes and prokaryotes. During translation these translation factors (Table 1) are involved in facilitating initiation, elongation, and termination of translation (Kaczanowska et al., 2007). In prokaryotes the initiation factors IF1, IF2, and IF3 bind to 30 S ribosomal subunit, tRNA, and mRNA to create a supra-molecular complex required for the creation of the translation machinery. The start of translation appears to involve an interacting between a sequence in the mRNA called the Shine-Dalgarno (SD) sequence and the 16S rRNA component of the 30S ribosome which results in expulsion of IF2 and the binding of the large 50S ribosome subunit to create the functional ribosome (Benelli et al., 2009; Kaczanowska et al., 2007; Malys et al., 2011; Marintchev et al., 2004). Once translation has been initiated a group of elongation factors, including EF-TU, EF-Ts and EF-G, promote recruitment and addition of aminoacylated tRNA to the complex and allows incorporation of the appropriate amino acid into the growing peptide chain depending on interactions between the mRNA codon and tRNA anticodon sequences. Lastly, the translation termination process is mostly carried out in the presence of RF1 and RF2 factors which bind appropriate mRNA stop codons and promote disassembly of the mRNA, 30S ribosome and 50S ribosome complex (Marintchev et al., 2004; Kaczanowska et al., 2007).

Table 1 A summary of the factors involved in protein synthesis in prokaryotes (Benelli, et al 2009).

Processes in Translation	Factors	Function
Initiation	IF1	Stabilizes 30S subunit.
	IF2	Binds fmet-tRNA to 30S-mRNA complex: bind to GTP and stimulates hydrolysis.
	IF3	Binds 30S subunit to mRNA
Elongation	EF-Tu	Binds GTP; brings aminoacyl-tRNA to the A site of ribosome
	EF-Ts	Generates active EF-Tu
	EF-G	Stimulates translation; GTP-dependent
Termination	RF1	Catalyzes release of the polypeptide chain from tRNA and disassociation of the translocation complex; specific for UAA and UAG termination codons
	RF2	Behaves like RF1; specific for UGA and UAA codons
	RF3	Stimulates RF1 and RF2

1.2 Sec-dependent Secretion and The Signal Peptide Hypothesis

In gram-negative bacteria, up to 40% of cellular proteins are exported to the periplasm, are present in the outer membrane (OM), or are freely secreted through a conserved secretion system referred to as the Type II secretion or Sec system. This is a universal pathway first described by studying protein secretion in *E. coli* (Lee & Schneewind, 2001; Beckwith, 2013; Collinson et al., 2001; Tseng et al., 2009). The core complex of the protein-conducting channel, which contains the SecY, SecE, and SecG proteins (Sec YEG complex or translocon), is necessary for mediating polypeptide movement through the inner membrane and/or into the periplasmic space during protein secretion (Collinson, 2001; Lee & Schneewind, 2001; Papanikou et al., 2007; Silhavy et al., 1983; Mori & Ito, 2001; Kudva et al., 2013).

The polypeptides that translocate across the inner membrane usually contain signal peptides (20–30-residue) as part of their translated sequence. These signal sequences are often cleaved from the pre-protein during localization to generate the mature protein. The signal sequences typically consist of three distinct domains (as shown in Figure 2 A) including the N terminal hydrophilic domain (1-5 residues long) which is flanked by positively charged amino acids, followed by a large domain comprised of hydrophobic amino acids, named the H region (about 8-12), and finally a polar carboxy-terminal domain, the C-terminal region (3-7 residues) which contains the signal peptidase cleavage site (Kudva et al., 2013; Low et al., 2013; Papanikou et al., 2007; Peterson et al., 2006; Lee & Schneewind, 2001; Edman et al., 1999; Fekkes et al., 1999; Lammertyn et al., 1998; Nesmeyanova et al., 1997; Collier, 1994; Pugsley, 1993; Silhavy et al., 1983). These signal peptides play a decisive role in targeting the protein to the Sec machinery

and direct the protein across the inner membrane in bacteria (Edman et al., 1999; Low et al., 2013).

It is well recognized that the pre-proteins destined for export through the inner membrane are recognized by factors such as the ribonucleoprotein signal recognition particle (SRP). The SRP carries the nascent amino acid chain/ribosome complex to the translocon within the plasma membrane. Once the SRP pre-protein complex is bound to SRP receptors and the ribosome is docked on the SecYEG pore, the polypeptides are translocated across the membrane, as shown in Figure 2 B (Lee & Schneewind, 2001; Papanikou et al., 2007; Mori & Ito, 2001; Beckwith, 2013; Kudva et al., 2013; Pugsley, 1993). The SRP complex consists of a 4.5S RNA and a 54-kDa Ffh protein (a homologue of eukaryotic SRP). During translation the SRP-nascent polypeptide complex is formed and reacts with the FtsY SRP-receptor (docking protein) on the cytoplasmic side of the membrane (Beckwith, 2013; Mori & Ito, 2001). Once translocation has occurred and the mature protein is released into the periplasmic space, the signal peptides are subsequently removed by signal peptidase on the exterior side of the membrane (Lee & Schneewind, 2001; Mori & Ito, 2001; Kudva et al., 2013).

1.3 Logistical Challenges for the Cell with the Present Signal Peptide

Hypothesis

The bacterial cell faces several challenges in the function of the protein secretion system based on a signal peptide. After a protein is synthesized on the ribosome and before it reaches its final destination, the process of proteins sorting, which involves a number of interactions with the Sec

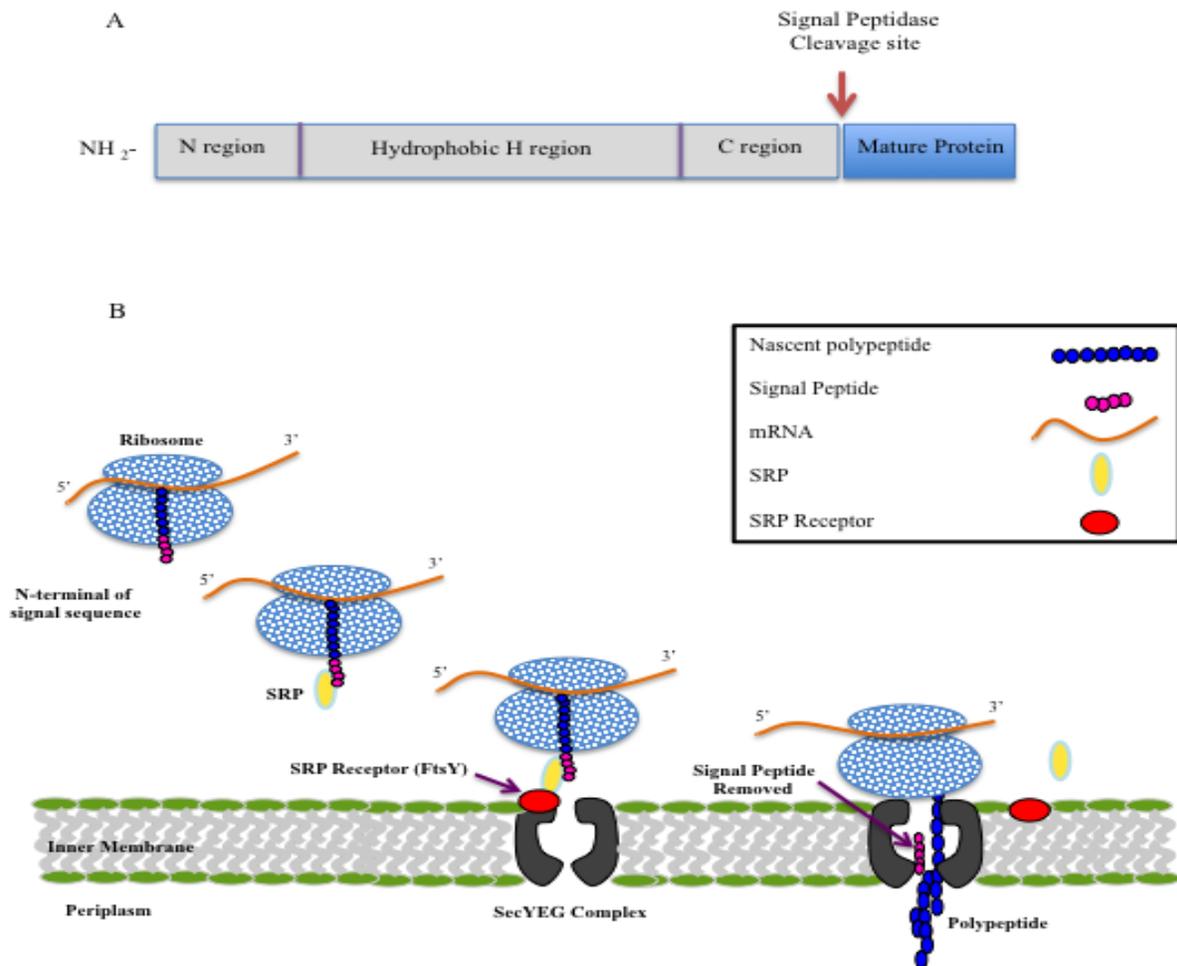


Figure 2 Diagram of Sec dependent secretion and the signal hypothesis.

(A) Structure of the signal peptide with three basic regions (N, H and C regions) and the cleavage site is identified with the arrow. (B) In the signal recognition particle (SRP) pathway the nascent amino acid chain/ribosome complex targets for secretion via the co-translational pathway.

machinery must occur (Low et al., 2013). Specifically, SecB chaperones and the ribonucleoprotein signal-recognition particle (SRP) play a significant role in mediating protein translocation. These proteins are characterized as piloting factors that recognize the nascent pre-proteins at the first stage of protein sorting and targeting (Papanikou et al., 2007). It has been well documented that SRP targets the highly hydrophobic signal peptides of nascent secretory proteins and membrane proteins than SecB chaperones to the Sec translocase (Low et al., 2013) (Papanikou et al., 2007). The cytoplasmic SecB chaperone function by creating a transport-competent state when it is bound to the pre-protein and it interacts with membrane-bound SecA (Low et al., 2013) (Papanikou et al., 2007). Nevertheless, SecB is not available in all bacteria (Papanikou et al., 2007). SecB binds to the mature region of the pre-protein, and some signal peptides delay pre-protein folding. Finally the SecB-pre-protein and SRP-pre-protein complexes bind to the SecYEG channel at the membrane (Papanikouetal, 2007). The Sec machinery can distinguish secreted from non-secreted proteins by binding to the signal peptide. Different sequences of these signal sequences were identified in prokaryotes and eukaryotes cells (Low et al., 2013). Experimental studies comparing different signal peptides has shown that the sequence properties of the signal peptide, such as charge, length, and hydrophobicity, were highly variable and rapidly evolving in different groups of organisms (Edman et al., 1999). The diversity of the signal peptides can control the differences in the ability of these signals to drive protein export (Low et al., 2013). For instance, blocking or impairing the secretion of *Escherichia coli* and *Bacillus subtilis* can result from disrupting the H region of the signal peptide (Collier, 1994). Recent findings also have shown that cytoplasmic proteins can be highly secreted when fused to signal peptides. For example, Manduca diuresin (MD) and the winter flounder antifreeze protein in *E.coli* can be secreted when fused to the OmpA signal peptide (Low et al., 2013).

There is a significant variety in the sequence of signal peptides, which makes it difficult to determine which alteration can enhance the performance of targeting secreted proteins to the Sec system.

1.4 The mRNA Hypothesis in Describing Membrane Targeting of the Pre-protein/Ribosome Complex

There have been several studies that show that changing the structure of the signal peptide can eliminate or reduce the efficiency of protein secretion (Van der Wolk et al., 1998; Puziss et al., 1989; Lehnhardt et al., 1988; Lino et al., 1987; Michaelis et al., 1986). However, other alterations in the signal peptide sequence can enhance protein sorting or have no effect (Petersen et al., 2006; Blachly-Dyson and Stevens, 1987; Kaiser and Botstein, 1986; Perlman et al., 1986; Koshland et al., 1982; Nevo-Dinur, et al., 2011). For example, deletions and two substitution mutations in the signal peptide sequence of the secreted *Saccharomyces cerevisiae* enzyme invertase still allowed normal secretion and expression of this enzyme which indicates this sequence can tolerate significant changes (Kaiser and Botstein, 1986). Although this protein secretion system is present in yeast, the structure of signal sequences in eukaryotes and prokaryotes are similar. Indeed, these findings suggest a central question: why do alterations that disrupt the basic features of signal peptides have no impact on protein secretion in some cases while in other cases some alterations appear to significantly influence secretion without significantly altering signal peptides structure? This leads to a possible alternative, mRNA hypothesis. The mRNA hypothesis proposes that the mRNA that codes for the secreted protein can actively participate in protein targeting rather than or in addition to the N-terminal amino acid signal sequence. (Habyarimana et al., 2013; Sorg et al., 2005; Wilhelm, et al., 1993; Okita, et al., 2002; Nevo-Dinur, et al., 2011; Cuia, et al., 2012). The mRNA hypothesis was established because mRNA targeting and subcellular localization have been shown to exist in

eukaryotes and prokaryotes (Wilhelm, et al., 1993; Okita, et al., 2002; Nevo-Dinur, et al., 2011; Cuia, et al., 2012).

It is known that the type 3 effector proteins can be trafficked from the cytoplasm of the bacterium into the cytoplasm of the host cell. A particular signal sequence within the first 30 amino acids of the protein structure is sufficient for protein targeting to the type 3-secretion system (T3SS). However, the type 3 effectors are still secreted even when a frame shift mutation in the signal sequence is introduced indicating that the amino acid sequences are unnecessary for targeting. Surprisingly, this observation confirmed the hypothesis that the required signal might be in the nucleotide sequences (mRNA), rather than the amino acid sequences of the signal peptide. It has been further reported that *fliC* export by the flagellar system in *E. coli* requires 5' untranslated RNA (UTR) sequences. Moreover, it was observed that the C-terminal region of *Y. enterocolitica* protein *yopR*, requires a nucleotide sequence, as an essential element for its export confirming the importance of an mRNA secretion signal (Habyarimana et al., 2013). In the *Yersinia* species, 12 additional types of Yop proteins are exported into the eukaryotic cytosol, where they are required due to their toxic function. These polypeptides translocate across the inner and outer membrane through the type III secretion pathway even though they do not possess a common secretion signal within the amino acid sequence of the target polypeptides (Anderson et al., 1997).

1.5 Evidence of Specific mRNA Recognition and Targeting in Eukaryotic System

In the last decade, a number of mRNAs have been shown to be specifically localized to different cellular regions in animal, plant, and yeast cells using *in situ* hybridization techniques. The

intracellular localization of mRNA can have distinct functions in regulating translation and targeting proteins to specific regions where they are required (Palacios & Johnston, 2001). Controlling gene expression is the primary advantage of mRNA transport (Glisovic et al., 2008; Martin & Ephrussi, 2009). It is also obviously a useful mechanism to localize copies of proteins in response to translations of localized mRNAs occur (Martin & Ephrussi, 2009).

The majority of mRNA targeting to the membrane requires becomes the presence of *cis*-acting elements or localization elements (Nevo-Dinur et al., 2011; Martin & Ephrussi, 2009). These *cis*-acting elements are mostly located in the 3'untranslated region (UTR). Subsequent work has shown that there are RNA-binding proteins (RBP) that recognize and bind to the *cis*-acting elements (Palacios & Johnston, 2001; Martin & Ephrussi, 2009; Slobodin, et al., 2010). These RBPs are considered to be the *trans*-acting elements and contribute significantly to the transcript localization and translational regulation mechanisms (Palacios & Johnston, 2001; Bashirullah et al., 1998; Martin & Ephrussi, 2009). The well-studied RBPs reveal how mRNAs can reach their destinations (Martin & Ephrussi, 2009; Palacios & Johnston, 2001; Bashirullah et al., 1998). For instance, RNA-RBP complexes, named ribonucleoproteins, can be transported along cytoskeletal elements by motor proteins (Martin & Ephrussi, 2009; Slobodin, et al., 2010; Kloc, et al., 2002). The ribonucleoprotein granule (RNP) also contributes to mitochondria, endoplasmic reticulum ER or peroxisomes to localization of mRNA toward distinct sites (Slobodin, et al., 2010).

Over 100 RNAs are known to localize to different intracellular regions have been characterized in eukaryotes. There are also 25 transcripts that are differentially localized in yeast (Palacios, 2007). For instance, cell division of yeast, the ASH1 mRNA is specifically transported to the bud tip of daughter cell in *Saccharomyces cerevisiae* as an assembly with myosin (Myo4) and actin (Glisovic et al., 2008; Martin & Ephrussi, 2009; Slobodin et al., 2010). The She2 and She3

proteins are also essentially in the localization of ASH1 mRNA. The process initiates when the She2 protein binds to the *cis*-acting elements in the 3' UTR of ASH1 mRNA. A high affinity binding between She2 and the C-terminus of She3 is also created. The She3-Myo4 complex further binds to the N-terminus of Myo4 and as a result, the ASH1 protein is localized (Glisovic, et al., 2008; Martin & Ephrussi, 2009; Palacios, 2007). The previous mechanism is needed for the repression of *HO* endonuclease gene transcription, which causes the repression of a mating type protein in the yeast bud (Glisovic et al., 2008).

A particularly clear example of transcript localization has been provided by the observation that β -actin mRNA targets to the lamella region in fibroblasts (Martin & Ephrussi, 2009). The interaction depends on the zipcode-binding protein (ZBIP), a 68-kDa protein, which forms one RNA recognition motif (RRM) and four (hnRNP K homology) KH domains (Kindler et al., 2007). ZBIP binds to the “RNA-zipcode,” a 54-nucleotide sequences, that exists on the 3' UTR site of the β -actin mRNA and forms a complex called ribonucleoproteins (RNP) (Glisovic et al., 2008; Martin & Ephrussi, 2009; Kloc et al., 2002). A KH domain facilitate the formation of an RNP- β -actin mRNA complex and also is involved in the association with actin microfilaments while the localization of the β -actin RNP complex is achieved by the RRM domains (Martin & Ephrussi, 2009).

Another example of RNAs trafficking has been described within the growing oocyte or syncytial embryo in drosophila. RNAs targeting is important in the definition of the oocyte and the specification of embryonic axes in the developing embryo (Lasko, 1999). *Staufen*, a double-stranded RNA-binding protein homologue, is important in mediating RNA localization and plays translation control in diverse cell types during embryogenesis (Kindler et al., 2007; Roegiers, 2000). In Oocytes, the localization of *Oskar*, *bicoid* and *nanos* mRNAs are fundamental for

initiating protein gradients in the anterior and posterior domains required for proper development (Roegiers, 2000; Lasko, 1999; Du, 2007). Whereas *bicoid* is directed to move along the anterior end of the embryo, the *oskar* mRNA is directed to the pole plasma at the posterior of the embryo (Du, 2007; St Johnston, 1995).

1.6 Evidence of mRNA Targeting in Prokaryotes

A better understanding of the widespread phenomenon of RNA transport has also been demonstrated in prokaryotes. In a translation-independent manner, mRNAs has been characterized in *E.coli* to reach a specific destination of their encoded proteins, cytoplasm, inner membrane, and poles (Nevo-Dinur et al., 2011; Martin & Ephrussi, 2009; Nevo-Dinur, et al., 2012). Interestingly, RNA localization was first observed in foci near transcribed genes. Investigations into mRNA localization demonstrated that the mRNAs are located at the center of the *E.coli* cells and show a limited amount of movement. This was achieved using a florescent RNA binding protein, MS2. Thus far, most RNA moved to a specific area related to their attachment to RNA polymerase. This area is known as F plasmid localized to the quarter central region of the bacterial cell (Nevo-Dinur et al., 2012; Keiler, 2012).

Likewise in *E.coli*, a parallel approach of localization of RNA to the cytoplasm, membrane, and pole has been further detected by fluorescence protein complementation using the eIF4A RNA-binding protein (Nevo-Dinur et al., 2012; Broude, 2011). Significantly, through the previous approach, the distribution of the *lacZ* mRNA, 5S RNA, and a short artificial untranslated RNA (Nevo-Dinur et al., 2012) which encodes the same plasmid were observed. All of these types of RNAs displayed different patterns of localization (Nevo-Dinur et al., 2012; Broude, 2011). The 5S RNA was predominately localized to the pole or in foci in regions devoid of nucleoids, while

the *lacZ* mRNA was distributed in the cytoplasm (Broude, 2011). In addition to the 5S RNA, the short non-coding RNA was in fact revealed at the cell poles (Nevo-Dinur et al., 2012).

1.7 Research Rationale

In the ribosomes, one or more polypeptides will form after the translation of mRNA that carries genetic information for protein synthesis. Once translated, these peptides are translocated across the inner membrane via the Sec machinery to assemble in the outer membrane. Several studies have proved that mRNA binding proteins (RBP) regulate the translocation process in *E.coli* through association with a single peptide. Using *E.coli* as a model I will identify and characterize unknown RBPs. We will utilize the periplasmic protein propyl isomerase chaperone SurA, which is involved in the biogenesis of outer membrane proteins to identify novel RBP involved in the secretory pathway in *E.coli*. As well as surA, 3-isopropylmalate dehydratase will be used as a control for the non-secreted protein.

Additionally, I will also examine the 5'UTR of alkaline phosphatase phoA and gMP reductase mRNAs that might achieve efficient isolation of such cytoplasmic factors. Our finding suggested that no specific factors were present to bind mRNA of secreted proteins, however, a logic number of RNA/DNA binding proteins were identified using mass spectrophotometer. Moreover, the existence of 3-isopropylmalate dehydratase enzyme may play a wider role in regulating itself in the cytoplasm.

1.8 Objectives

The objectives of the present study were to:

- As outlined in Figure 3B, I propose an alternate model where instead of the SRP only binding the signal sequence of polypeptides it also interacts with the 5' region of the mRNA.
- To use two *E. coli* proteins, one localized to the cytoplasm (3-isopropylmalate dehydrogenase IsodH and GMP reductase) and the other (periplasmic propyl isomerase chaperone SurA and alkaline phosphatase PhoA) secreted through the Sec-dependent pathway.
- Produce transcripts *in vitro* system and to use these as bait for pull down assays to identify a cytoplasmic factor that associate with these mRNAs.
- Look for factors in the cell that may interact with the mRNA of the secreted protein and not in the cytoplasmic protein. This factor could be involved in targeting that mRNA/ribosome complex to the membrane to effect secretion.

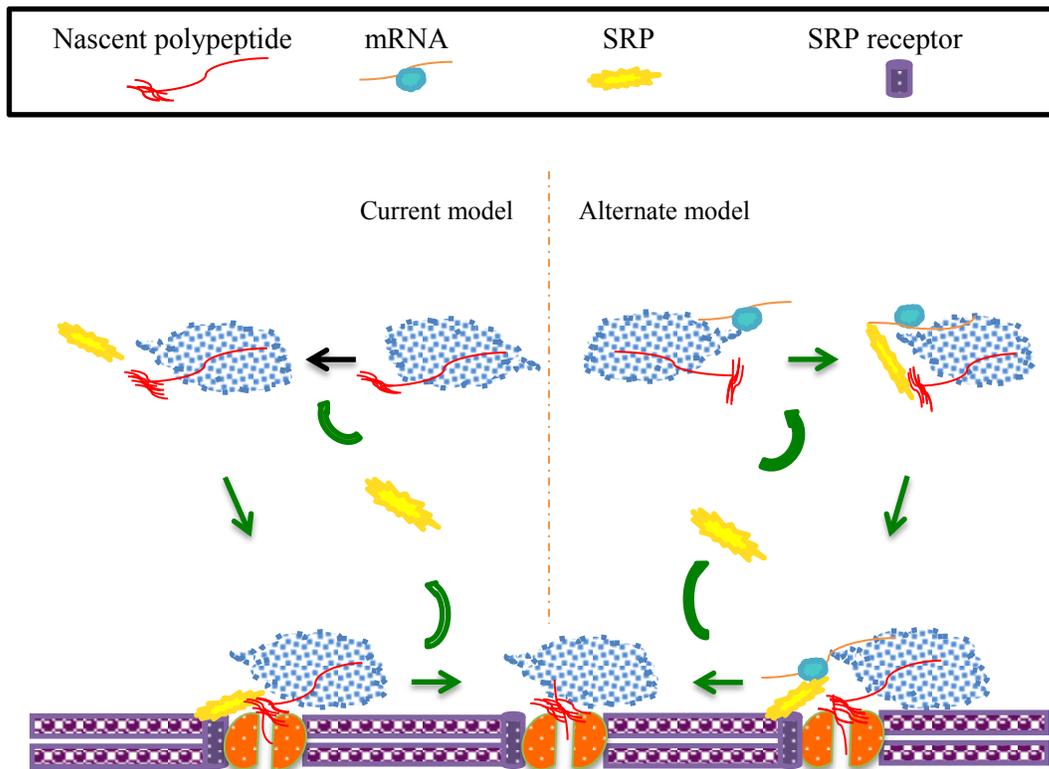


Figure 3 An alternate model of targeting secreted proteins to the Sec machinery.

(A) In the current model, the SRP recognizes the N-terminal signal peptide of the nascent protein and then targets the ribosome-protein complex to the inner membrane. (B) In our alternate model, the SRP recognizes both the N-terminal signal peptide and the mRNA 5' region as it exits the ribosome and targets the ribosome-pre-protein-mRNA complex to the translocation channel located in the membrane.

Chapter 2: Experimental Methods

2.1 Bacterial Cultures

Escherichia coli strains B; ATCC 11303 & BL21 was cultured and maintained in LB (Fisher Scientific, Mississauga) liquid media. For solid media, agar was added to the liquid media at a final concentration of 1%. Long term storage at -80°C in LB media containing 10% glycerol was used to save the stock cultures. Liquid cultures were maintained at 37°C in a rotary shaker (Inova 4000, New Brunswick Scientific) at 200 rpm as a source of bacteria, bacterial DNA, and protein extracts.

2.2 Extraction of Genomic DNA

Twenty ml from an overnight culture of *E. coli* at 37°C/200 rpm were transferred to 50 ml polypropylene tubes containing 1.5 ml TNE buffer (0.1 M Tris-Cl, pH 8.0, 0.15 M NaCl, and 20 mM EDTA) and washed thoroughly by three cycles of spin/resuspend at 2000 xg for 5 minutes/cycle. Following a final spin, the cells were collected and resuspended in 2 ml ice-cold 70% ethanol and incubated on ice for 20 minutes, collected by centrifugation and resuspended in 4.8 ml of TEST/LR buffer [0.1 M Tris-Cl, pH 8.0, 20 mM EDTA, 0.5 M sucrose, 1% (v/v) TritonX-100, 24 mg of lysozyme (30 µl of lysozyme). Following an incubation step on ice for 1 h with occasional shaking, the suspension was transferred to a water bath set at 68°C for 10 minutes. Subsequently, 50 µl of 10% SDS was added and incubation was continued for an additional 15 minutes. Finally, 87 µl of 5 M NaCl and 69 µl of CTAB/NaCl solution (1% N - acetyl-N, N, N -trimethylammonium bromide in 0.73 M NaCl) were added and the tube contents

were mixed by inversion and rotation. The extraction was followed by incubating the tube at 68°C for 15 min and then at -20°C for 30 minutes. The DNA was extracted from the cell lysate by the addition of 1 vol chloroform: isoamyl alcohol (24:1, v/v), mixing for 5 minutes, followed by centrifugation at 10000 rpm for 10 minutes to separate the organic phase from the aqueous phase. Added drop by drop, 2 vol 100% ice-cold ethanol were added to precipitate the genomic DNA for 30 second. A final centrifugation at 12,000 xg for 15 minutes and washing in ice-cold 70% ethanol produced the purified DNA.

2.3 Quantitation of Nucleic Acids

Purified DNA, resuspended in 50–250 μ l of TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) buffer, was quantitated spectrophotometrically using a Shimadzu UV-2401PC Spectrophotometer (Hitachi, Tokyo, Japan). The ratio of absorbance at 260 nm (A₂₆₀) to 280 nm (A₂₈₀) and a standard curve provided a quantitative measure of the concentration of the DNA. For quantitation of RNA, the Qubit fluorometer (Invitrogen) and the Qubit RNA HS Assay Kit (Life technologies) was used. This kit is optimized to quantify mRNA with little to no interference from other RNA, DNA, or proteins.

2.4 Genomic DNA Amplification

To generate the mRNA transcripts, DNA templates from *E.coli* strains BL21 and ATCC11303 (see Table 2 below) were synthesized using the polymerase chain reaction (PCR) mediated by oligonucleotides primers that are 40 base pairs long including the 20 bp of the T7 promoter sequence. PCR was carried out with initial denaturation of 30 second at 98 °C followed by 32 cycles of: 10 sec at 98 °C, 30 sec at optimal annealing temperature (Ta) and 30 sec at 72°C

followed by a final extension period of 10 min at 72°C. The optimal annealing temperatures for PCR amplification are 57-58 °C for surA, 51-52 °C for isodH, 60 °C for phoA, and 56 °C for gMP. Each 25µl reaction contains 12.5 µl PCR-EZ D-PCR Master Mix (BIO BASIC INC), 0.1-1 µM reverse and forward primers, up to 0.5 µg of template DNA. The primers for the PCR reactions and their products are shown in Table 2.

2.5 Transcription by T7 RNA Polymerase

The mRNA transcripts were generated *in vitro* using the MEGAscript T7 transcription Kit (Ambion, Austin, TX). Briefly, 20 µl or 40 µl reactions were assembled using the supplied UTP, ATP, GTP and CTP ribonucleotides and 1 µg of template DNA from *E.coli* strains BL21 and ATCC 11303. The mixture was assembled on ice and the reactions were started by shifting the temperature to 37°C. The reactions were terminated after 3 h. Additionally; DNase treatment was used to remove the DNA template by adding 1 µl TURBO DNase to the reaction once its completed. The RNAs were purified using the Rapid Bacteria RNA Isolation Kit (Bio Basic) and the RNA concentrations were determined spectrophotometrically using a Shimadzu UV-2401PC Spectrophotometer (Hitachi, Tokyo, Japan).

2.6 Agarose Gel Electrophoresis

The PCR amplifications and mRNA transcripts were assessed by electrophoresis on 1% agarose gels containing 0.5 µg/ml ethidium bromide (Fisher Biotech). The gel was subjected to 60V for 1.5 h in TAE (1mM EDTA, 40 mM Tris-acetate) running buffer. The samples were loaded with 6X loading dye (Fermentas) and the gel was visualized using ChemiDoc XRS (BIO-RAD) UV Tran illuminator.

Table 2 Primers for the PCR reactions.

Gene	Gene Fragment Amplified	Primers	Temperature	Fragment size (bp)
surA	<i>surA</i>	Forward: 5'TAATACGACTCACTATAGGGATGAAGAAGCTGGAAAACGCTG-3' Reverse: 3'CAACGTCGCTTTCCAGCAC-5'	55-58°C	127
isodH	<i>isodH</i>	Forward: 5'TAATACGACTCACTATAGGGATGGCTAAGACGTTATACGAAAAA-3' Reverse: 3'TAGCGAAGGTTTTGCCCG-5'	51-52°C	178
pohA	<i>pohA</i>	Forward: 5'TAATACGACTCACTATAGGGAAAAAGTTAATCTTTTCAACAGCTGTC3' Reverse: 3'GCCCCGGTTTTCCAGAACAG-5'	59-60°C	200
gMP	<i>Gmp</i>	Forward: 5'TAATACGACTCACTATAGGGAGTCTTGCGCCGATTGATTA-3' Reverse: 3'GTGAATTGACGTTCCAGTTCA5'	56°C	194

2.7 Biotin Labeling of mRNA

The biotin label enables the RNA probe to be immobilized with streptavidin that is beneficial to study such RNA interactions. RNA 3' End Biotinylation Kit (Pierce) was used to attach a single biotinylated nucleotide to the 3' terminus an RNA strand. The reaction mixture was prepared by combining of 50 pmol of test RNA, 40 U RNase inhibitor, 1X of 10X RNA ligase reaction buffer, 1 nmol of biotinylated cytidine (Bis) phosphate, 40 U of T4 ligase, 15 % of PEG and up to 30µl nuclease free water. The reaction was then incubated for 3h, 6h and 16 h at 16°C. In order to extract the RNA ligase 100µL of chloroform: isoamyl alcohol was added and tubes were centrifuged for 2-3 minutes to separate the phases .10 µl of 5M NaCl, 1 µl of glycogen and 300 µl of ice-cold ethanol were added to the aqueous phase. The samples were then centrifuged at 13,000 rpm for 15 min at 4 °C. The supernatant discarded and the pellet was washed with 300µl ice-cold 70 % ethanol and air-dried the pellet was resuspended in 20µl dH₂O.

2.8 Chromogenic Detection of Biotinylated mRNA and Evaluation of Labeling Efficiency

The Biotin Chromogenic Detection Kit (Fermentas) was used to detect the biotinylated RNA on polyvinylidene difluoride membrane (PVDF). 30 ml of Blocking/Washing Buffer was used [1 volume of 10X Blocking/Washing Buffer with 9ml dH₂O] to wash the membrane for 5min at RT on platform shaker. The membrane was further blocked in 30 ml of Blocking Solution [1%(w/v) 0.3 g of Blocking Reagent in 1X Blocking/Washing Buffer] for 30 min and then the membrane was incubated for an additional 30 min in 20 ml of diluted Streptavidin-AP conjugate [dilute concentrated Streptavidin-AP conjugate 5000-fold in Blocking Solution]. With moderate shaking

the membrane was washed in 60 ml Blocking/Washing Buffer for 15 min and repeated once with fresh Blocking/Washing Buffer. The solution was discarded and then the membrane was incubated with 20 ml of Detection Buffer for 10 minutes using BCIP/NBT as substrates. To determine the labeling efficiency, spectrophotometric measurement of labeled RNA was performed by quantifying the amount of biotin at 500 nm in the samples in order to estimate the biotin concentration that attached on bait in compared with the Biotinylated IRE RNA Control from the Pierce RNA 3' End Biotinylation Kit using Shimadzu UV Spectrophotometer (Hitachi, Tokyo, Japan).

2.9 Protein Extraction

Cells from an overnight culture of *E. coli* at 37°C/200 rpm were harvested by centrifugation at 15,000 rpm for 10 min. The pellet was resuspended in 1.0 ml lysis buffer [50 mM Tris-HCl, pH 7.5 (6.05 g/l), 150 mM NaCl (8.76 g/l)] and Lysozyme at 0.2mg/ml was added. The pellet was then sonicated in an ice bucket using 3 x 10 sec pulses and then 1.0 ml lysis buffer containing 0.1 % Triton X-100 was added. The soluble proteins (supernatant) were prepared by centrifugation (15, 000 xg for 10 min.), which was then used for the pull-down assay.

2.10 Protein Assay

Protein concentrations throughout this work were quantified using the BCA kit (Pierce Scientific). The BCA assay is based on the reduction of Cu^{2+} to Cu^{1+} by protein in the samples in alkaline medium with a highly selective colorimetric detection of Cu^{1+} with a reagent containing bicinchoninic acid. The assay was performed according to the manufacturer's instructions supplied with the kit.

2.11 UV Crosslinking

In an attempt to identify factors that interact with an mRNA sequence, *in vitro* studies have been conducted. *E. coli* extracts were incubated with biotinylated RNA species, and a number of factors that are able to bind the RNA were enriched. The putative interacting factors were physically cross-linked to biotinylated mRNA by exposure to UV-C light at 254nm / 120,000 $\mu\text{J}/\text{cm}^2$ for 5 min (Microprocessor-Controlled UV Cross linkers). Therefore, the mRNAs and the bound interactome can be efficiently captured using pull down assays.

2.12 Pull-Down Assay

An aliquot of magnetic beads (0.8 mg beads /1mg of *E.coli* extract) (Dynabeads, Invitrogen) was washed with TBS buffer, 7.4. The beads were then blocked with 5% casein hydrolysate in TBS [50 mM Tris-Cl, pH 7.6; 150 mM NaCl] for 2 h at 22°C (RT) with gentle rotation followed by 3 washing in TBS buffer. Approximately 40-50 μg of biotinylated mRNA was incubated with 1 mg of *E. coli* protein extract in binding buffer (3ml of 1X TBS [50 mM Tris-Cl, pH 7.6; 150 mM NaCl], 0.05% TritonX-100) at 4°C for 1h .The mixture then was added with coated beads for an additional hour at 4°C. Samples were boiled for 10 minutes following the addition of 15 μl of 2X SDS sample buffer [2 mL Tris (1 M, pH 6.8), 4.6 mL glycerol (50%), 1.6 mL SDS (10%), 0.4 mL bromophenol blue (0.5%), and 0.4 mL β -mercaptoethanol] (Figure 4).

2.13 Effect of Inhibitors of mRNA Degradation on The Pull Down Assay

Since several RNase inhibitors have been shown to have a significant role in the formation of nucleic acid -protein complexes, the effect of these inhibitors were examined. 1 mg of *E. coli* cell lysate was treated with either 10 mM of Aurintricarboxylic acid (ATA) or with

Diethylpyrocarbonate (DEPC) (Sigma) followed by an incubation with 40-50 µg of biotinylated RNA at 4 °C for 1h. The reaction was then incubated with beads that blocked with 5% casein in TBS buffer for 2 h (0.8 mg beads /1mg of *E.coli* extract) at 4°C for an extra hour. After several washes with 1X TBS buffer (50 mM Tris-Cl, pH 7.6; 150 mM NaCl) the reactions then were ran on 12 % SDS-PAGE gel. The gel was stained with Coomassie brilliant blue stain.

2.14 Protein Capture Using Nylon Membranes

Up to 40-50 µg of mRNAs were Immobilized on Hybond N⁺ membrane [positively charged nylon membrane “Amersham Biosciences”] using RNA slot-blotting technique “PR648 Hoefer Scientific”. Samples were carefully loaded to the membrane through the slot and vacuum was turned on (set up at 13-25 cm Hg). Once the samples were completely loaded, 1 ml of 10x DEPC buffer [3 M NaCl, 300 mM tri-sodium citrate, pH 7.0] was added to each well and the liquid was pulled through after resetting the vacuum at 35-50 cm Hg (PR648 Slot Blot Filtration Manifolds). After repeating the previous step for a total three rinses, the mRNAs were then cross-linked on the membrane with UV light (Spectroline UV cross linker) for 5min. The membrane was blocked with 5% milk for 1 h at RT followed by an incubation with 1 mg cell extract for an additional 1 h at RT in the presence of 200 µl binding buffer (3ml of 1X TBS [50 mM Tris-Cl, pH 7.6; 150 mM NaCl], 0.05% TritonX-100). Subsequently after three washing steps the extract areas of the dot-blot on the membrane were boiled in SDS sample buffer to release any proteins that may bind to mRNAs and were ran on SDS-PAGE.

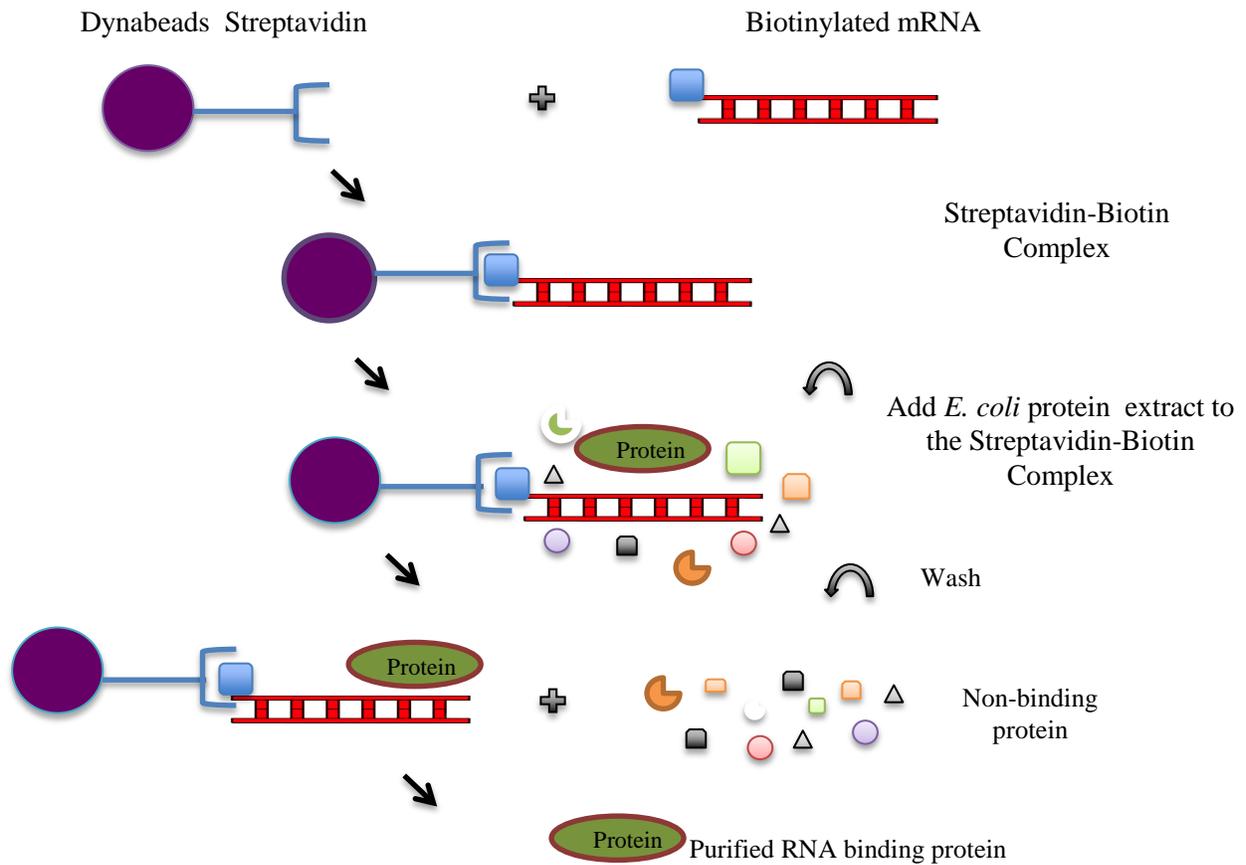


Figure 4 Depiction of the pull down assay.

The mRNA (bait) is immobilized onto magnetic beads via binding of bead surface streptavidin to the biotin on the 3' end of the mRNA. This complex is then used to pull down soluble *E. coli* proteins that would interact with the mRNA.

2.15 RNA-Protein Interaction Assay

In order to capture any binding factors that bind to the mRNA, the RNA-Protein assay was also used. Native polyacrylamide gel electrophoresis was used in this assay to separate proteins based on both proteins charge and its size. These gels were made of a 12 % separating gel (3.4 ml H₂O, 2.5 ml 1.5 M Tris-HCl pH 8.8, 4ml 30% acrylamide / 0.8 % bis-acrylamide (w/v), 60 µl 10% (w/v) ammonium persulfate (APS), and 7 µl tetramethylethylenediamine (TEMED)) and a 4% stacking gel (3.1 ml H₂O, 1.25 ml 0.5 M Tris-HCl pH 6.8, 0.65 ml 30% acrylamide / 0.8 % bis-acrylamide (w/v), 31 µl 10 % (w/v) APS, and 5 µl TEMED). 1mg of *E.coli* extract were prepared in a non-reducing non-denaturing 2X sample buffer (38% glycerol, 0.1 M Tris-HCl pH 6.8, 0.02 % bromophenol blue, 0.1 M DTT) to final concentration of 1X and then loaded into the wells. The gels were ran at a constant voltage of 200 V for 45 min in Mini -PROTEAN Tetra System (BIO-RAD) with 1x native running buffer (144 g glycine, 1/10 dilution of 30 g Tris-HCl in 1 L dH₂O; pH 8.3). Lanes from native-PAGE gel are cut and the strips were incubated in SDS sample buffer (see SDS PAGE below) for 15 min at 22°C (RT). Strips were then rotated through 90° and placed into SDS-PAGE gel (explained in 2.16) that either included phoA bait or has no transcript as control (Nijtmans, et al., 2002). The Pierce Prestained Protein Molecular Weight Marker was used as a protein standard for protein ranging in size from 20 -120 KDa. The gel was stained with Coomassie brilliant blue stain.

2.16 SDS PAGE

Polyacrylamide gels containing SDS were made of a 12 % separating gel (3.4 ml H₂O, 2.5 ml 1.5 M Tris-HCl pH 8.8, 4ml 30% acrylamide / 0.8 % bis-acrylamide (w/v), 50 µl 20% SDS 60 µl 10% (w/v) ammonium persulfate (APS), and 7 µl tetramethylethylenediamine (TEMED) and 4%

stacking gel (3.1 ml H₂O, 1.25 ml 0.5 M Tris-HCl pH 6.8, 25 µl 20% SDS, 0.65 ml 30% acrylamide / 0.8 % bis-acrylamide (w/v), 31 µl 10 % (w/v) APS, and 5 µl TEMED). Protein samples from pull-down assay were combined with 2X sample buffer (3.8 % SDS, 38% glycerol, 0.1 M Tris-HCl pH 6.8, 0.02 % bromophenol blue, 0.1 M DTT) to final concentration of 1X (Laemmli, 1970). The entire contents of the pull down assay were boiled for 10 min in water bath and then loaded into the wells. The gels were performed at a constant voltage of 200 V for 45 min in Mini -PROTEAN Tetra System (BIO-RAD) with 1x SDS running buffer (144 g glycine, 1/10 dilution of 30 g Tris-HCl, 10 g SDS in 1 L dH₂O; pH 8.3). The Pierce Prestained Protein Molecular Weight Marker (20-120 kDa) and the Fermentas Unstained Protein Molecular Weight Marker (14.4-166 kDa) were used as a protein standard. The gel was stained with either Silver or Coomassie brilliant blue stains.

2.17 Protein Identification Using Matrix-Assisted Laser Desorption Time-of-Flight (MALDI-TOF) Mass Spectrometry

Followed the protocol posted by the Advanced Protein Technology Centre (<http://www.sickkids.ca/Research/APTC/Map-Spectrometry/Sample-Protocols/In-Gel-Tryptic-Digestion-Protocols/index.html>) for in-gel digestion and extraction. After placed the gel pieces in micro centrifuge tubes they were first de-stained by washing with 50 µl of 50mM Ammonium bicarbonate (AmBic) for 5 min and supernatant was discarded. The gel was then shrinking with 50 µl 50% acetonitrile/25 mM ammonium bicarbonate for 10min. 30 µL of 10 mM DTT for was added to reduce the gel for 30 minutes at 56°C followed by an alkylation step with 30 µL of 100 mM iodoacetamide for 15 minutes in the dark and the supernatant was discarded once more. About 14 µL of 13 ng/µL trypsin in 50 mM ammonium bicarbonate were used to digest proteins

on ice for 20 min. further 10 μ L of the supernatant was replaced with 20 of 50 mM AmBic and the gel pieces were incubated at 37 °C overnight. Through several washes with AmBic, formic acid and acetonitrile the digested peptides were extracted. After the overnight incubation the supernatant was collected and transferred to a new tube. 20 μ L of 25 mM ammonium bicarbonate was add to the previous collected liquid for 10 min followed by addition of 20 μ L of 5% formic acid for an extra 10 min. This wash was repeated with 20 μ L of 100% acetonitrile for 10 min. both 100% acetonitrile and 5% formic acid washes were then repeated one more. The extracted peptides were stored at -20 °C until ready to dried down using a SpeedVac concentrator (Fisher Electron Corporation, Savant 120) and then sent to the mass spectrometry analysis in Western University and The Hospital for Sick Children.

Chapter 3: Results

3.1 PCR Amplification of Genomic DNA

The optimization of annealing temperatures resulted in successful amplification of the periplasmic propyl isomerase chaperone (SurA) and the cytoplasmic protein 3-isopropyl malate dehydrogenase (isodH) genes from *E.coli* ATCC 11303. The PCR of surA and isodH genes produced bands that migrated between 100–200 bp when screened on a 1% agarose gel. The primer pair of both genes was designed with a T7 promoter sequence. The surA gene amplified best with a Ta range between 57-58 °C while the optimal Ta of (isodH) gene was 51-52°C (Figure 5). Furthermore, the 5' untranslated region (or UTR) of phoA gene from *E.coli* BL21 was amplified by PCR with a Ta of 60 °C while at Ta 56 °C the gMP fragment was amplified (Figure 6).

3.2 Transcription of Amplicons

Since the RNA binding factors were the intended target, a high yield of a particular mRNA was produced by conventional *in vitro* transcription reactions in the presence of the T7 polymerase enzyme. By modifying typical transcription reaction condition the reaction yields a total of approximately 100 µg of RNA. Between 100-200 bp of surA mRNA (Figure 7 lane 3) and isodH mRNA (Figure 7 lane 4) were generated and samples were visualized in 1% agarose gel while lanes 1 and 2 were represent the surA and isodH gene fragment transcripts. Bands in Figure 8 (lanes 2 and 4) were referred to transcription reactions of phoA and gMP gene fragments with a predicted size between 100-200 bp. Lanes 1 and 3 otherwise indicate the phoA and gMP gene fragments.

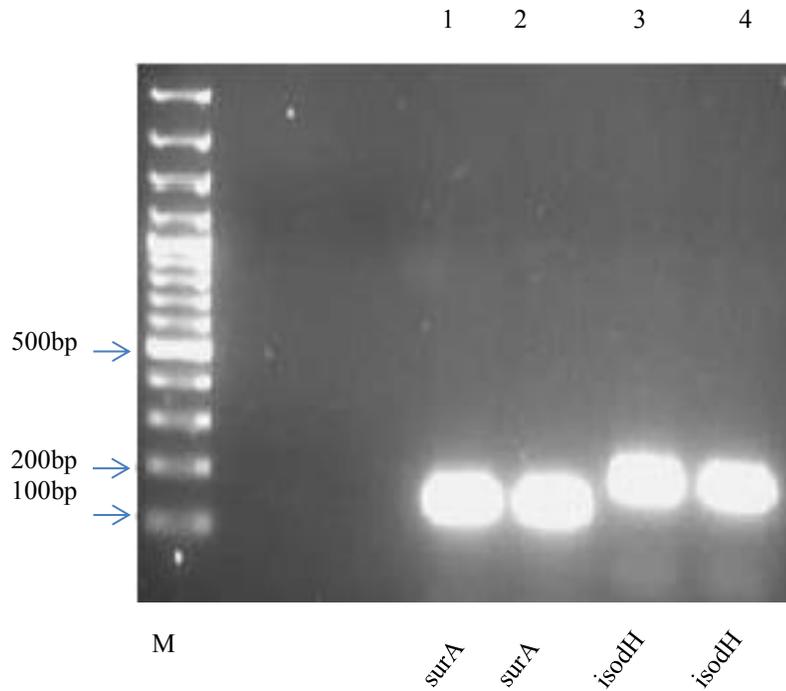


Figure 5 Amplification of *surA* and *isodH* genes fragments.

Agarose gel showing a PCR product of *surA* gene that were amplified at 57 °C in lane 1 and at 58 °C in lane 2 while in lane 3 an amplification of *isodH* gene fragment at 51 °C and at 52 °C were shown in lane 4. Bands shown between 100-200bp are the expected size of both genes fragments amplified reaction (predicted size 127 bp for *surA* and 178 bp for *isodH*). M= 100 bp plus DNA ladder (Fermentas).

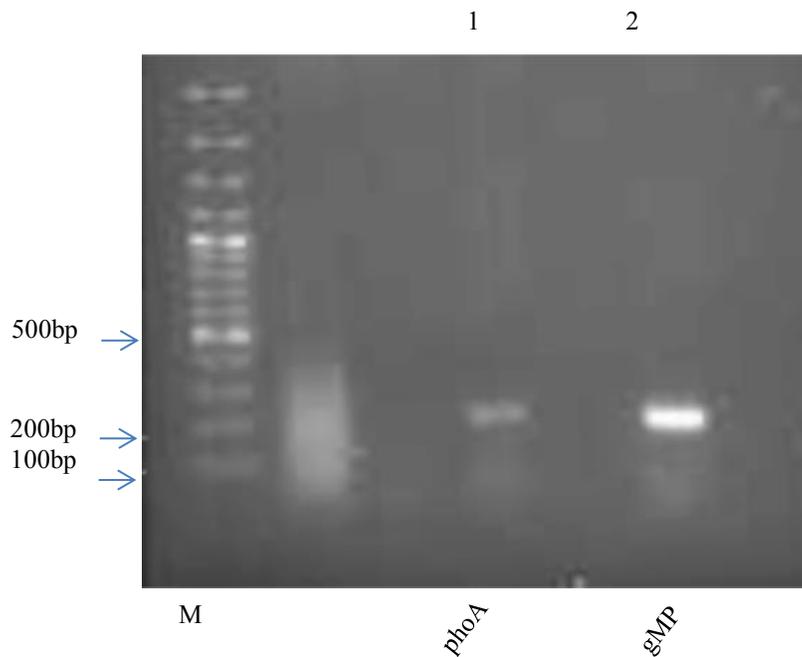


Figure 6 Amplification of phoA and gMP genes fragments.

An amplification of the phoA and gMP gene fragments shown in agarose gel. Distinct band in lane 1 indicate the typical product size between 100 - 200 bp of phoA gene fragment that was amplified at 60 °C. As well as phoA size gMP gene fragment also amplified at 56°C as shown in lane 2. M=100 bp plus DNA ladder (Fermentas).

3.3 Quality of Transcription Product

Characterization of RNA quality was done using the Agilent 2100 Bioanalyzer. It is a simple and alternative method to determine the quantification and data analysis of RNA. After transcriptions of *surA* and *isodH* mRNAs from *E. coli* different concentrations of transcripts were indicated (Figure 9). The analysis determined that *surA* mRNA (23 ng/μl) and *isodH* mRNA (71 ng/μl) concentrations were enhanced after 4 h of incubation (lanes 1 and 2). However the concentrations of *surA* mRNA (9 ng/μl) and *isodH* mRNA (3 ng/μl) decreased after 16 h incubation (lanes 3 and 4).

3.4 mRNA Structure of Different Transcripts

Despite the fact that the sequences of the secreted and cytoplasmic mRNA are different mRNA structure is also an important factor to be consider (Figure 10 A, B, C and D). This fact explains that the proteins that bind to mRNA are actually based not only on the actual sequence but can also bind to structural elements. I followed the process posted by the *RNAfold* Webserver (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) in order to predict secondary structures of mRNA sequences (Szostak et al., 2013).

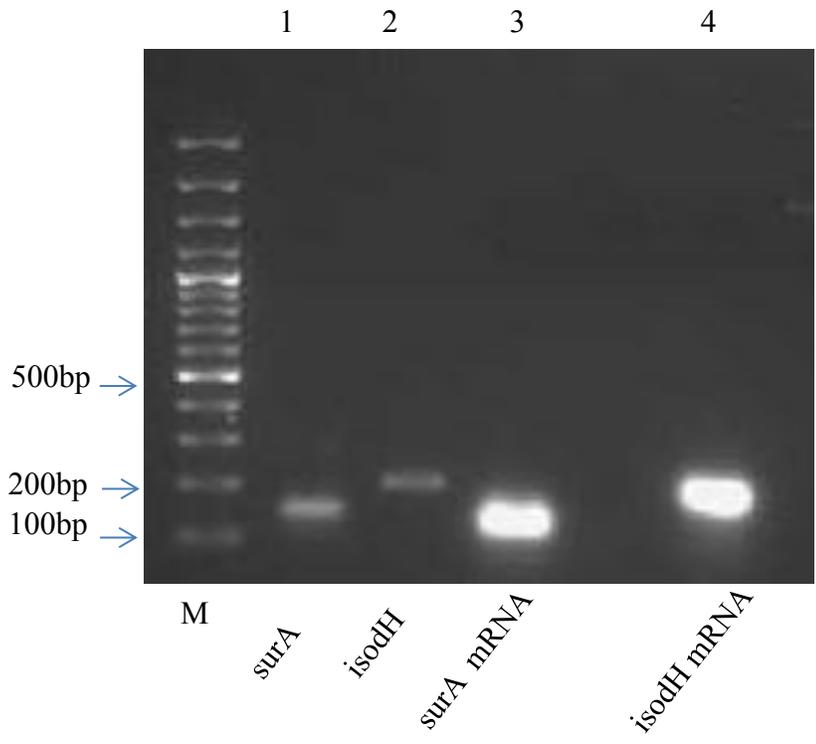


Figure 7 *In vitro* RNA transcriptions from *surA* and *isodH* genes fragments.

Agarose gel showing the *in vitro* transcription reaction (lanes 3 and 4) using the *surA* and *isodH* gene fragments transcripts as template (lanes 1 and 2). The main bands shown between 100 – 200 bp match the expected size of transcription products. M=100 bp plus DNA ladder (Fermentas).

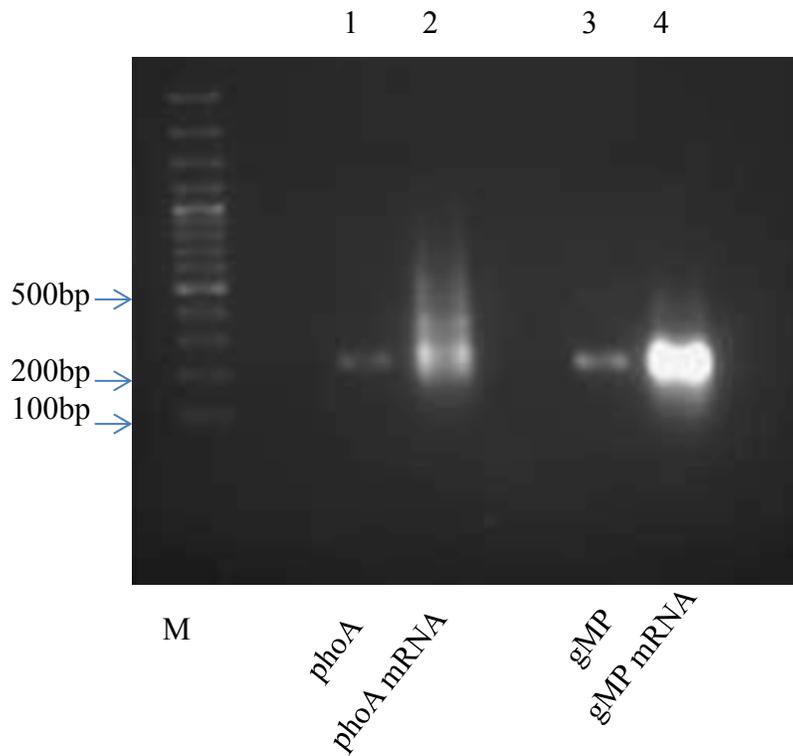


Figure 8 Confirmation of the presence of transcripts for phoA and gMP genes fragments using MEGAscript kit.

Agarose gel showing typical results of the *in vitro* transcription reaction (lanes 2 and 4) using the phoA and gMP gene fragments as template (lane 1 and 3). The bands shown in lane 2 and 4 were between 100-200 bp of transcription products. M=100 bp plus DNA ladder (Fermentas).

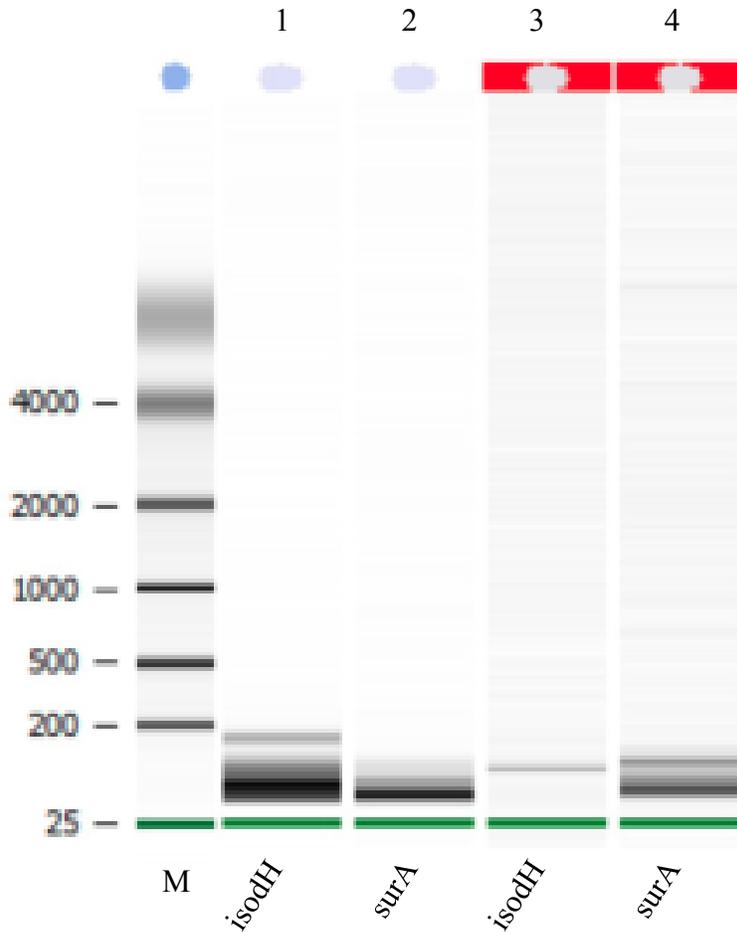


Figure 9 Analysis of RNA transcription products using the Agilent Bio-analyzers.

Bioanalyzer analysis was used to give a size estimate of the *in vitro* transcription reaction (using the *surA* and *isodH* gene fragments transcripts as template to generate RNAs) after incubation for 4 h (lanes 1 and 2) and for 16 h (lanes 3 and 4). This method was also used to determine the quality of *isodH* and *surA* mRNAs. The following concentrations were determined after transcription for 4 h *surA* mRNA (23 ng/ μ l), *isodH* mRNA (71 ng/ μ l) while *surA* mRNA (9 ng/ μ l) and *isodH* mRNA (3 ng/ μ l) were determined after 16 hours of incubation. The gel indicates in a couple of mRNAs samples more than one size of RNA product in lanes 1 and 4.

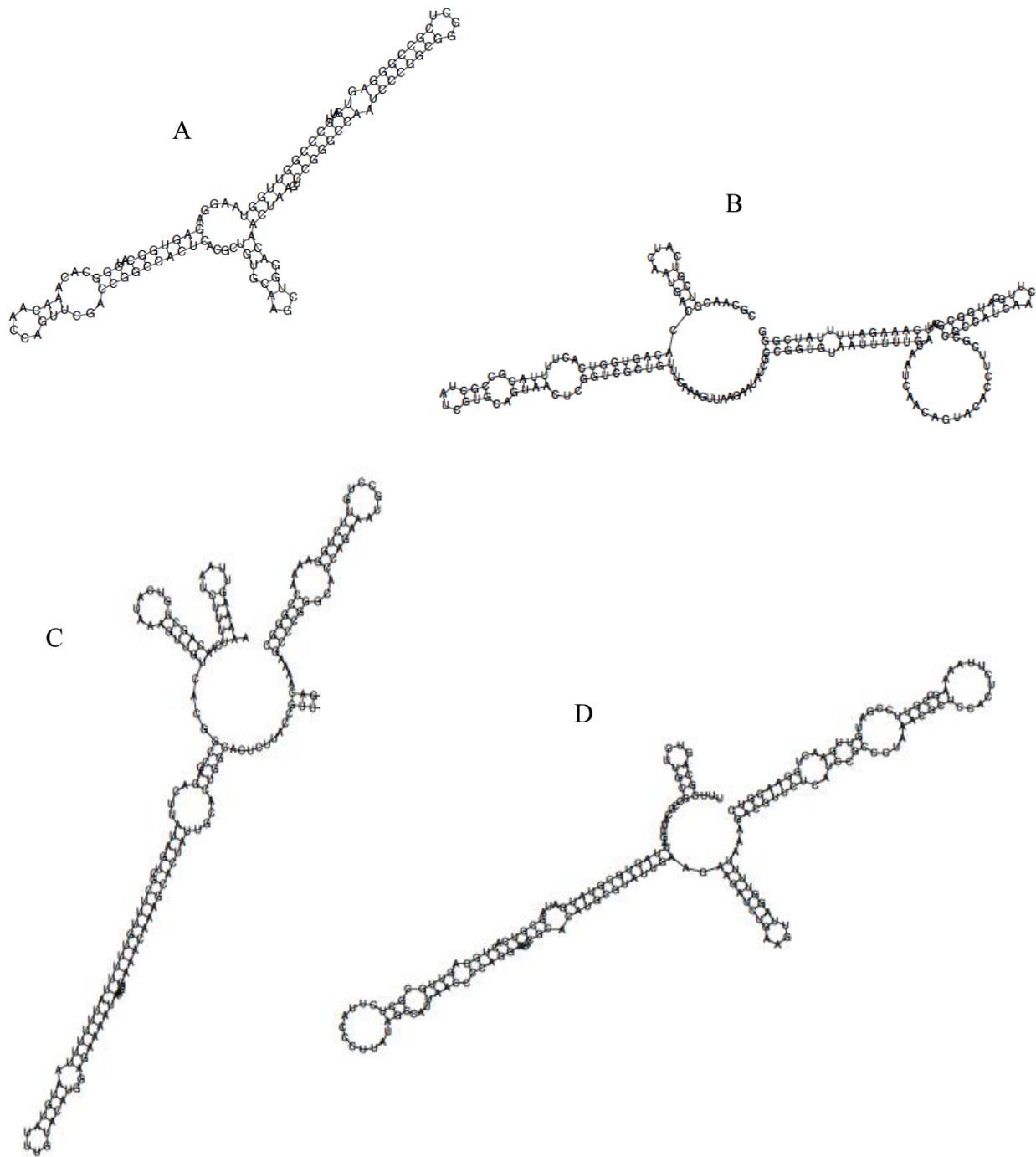


Figure 10 Simulated secondary structures of *surA* (A), *isodH* (B), *phoA* (C), and *gMP* (D).

mRNAs. The simulation was performed using the mfold webserver (<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>).

3.5 Estimation of Labeling Efficiency

Biotinylation of the transcripts in this study was performed not for detection but to be used as a target for the pull down assay. Addition of the biotin to the transcripts allowed the use of streptavidin-coated beads to specifically bind and pull down the transcripts along with any bound target protein. Results shown that 41 % of surA, 36% isodH %, 43% of phoA and 38.5% gMP mRNAs were labeled at 37°C for 2h as shown in Table 3.

3.6 Detection of Biotinylated RNA

Biotinylated surA and isodH transcripts were detected with streptavidin coupled to alkaline phosphates (AP) on PVDF membrane. Streptavidin –AP conjugates bind to the biotin – labeled mRNA and then visualized using chromogenic substrate for alkaline phosphatase BCIP/NBT, which produces a blue-purple precipitate (Figure 11).

3.7 The Streptavidin – Biotin System

In order to improve the of binding between the streptavidin-coated beads and biotin (to model binding to biotinylated RNA) I decided to examine different amount of beads (0.2, 0.4, 0.6 and 0.8 mg/ reaction) each incubated with 1nmol of biotin. I expected that increasing amounts of beads would increase the biotin binding to the beads. A comparison of the biotin molecules in the supernatants after incubation with the streptavidin-coated beads is presented in Table 4. Hence I can clearly see successful streptavidin-biotin binding confirming the pull-down system. This data show that increasing the amount of streptavidin-coated beads above 0.2 mg / reaction decreased the amount of free biotin present in the supernatant.

Table 3 mRNA labeling efficiency.

To determine the labeling efficiency the following equation was used: labeled RNA/ total RNA X 100.

Sample ID (mRNA)	Label Efficacy %
surA	41
isodH	36
phoA	43
gMP	38.5

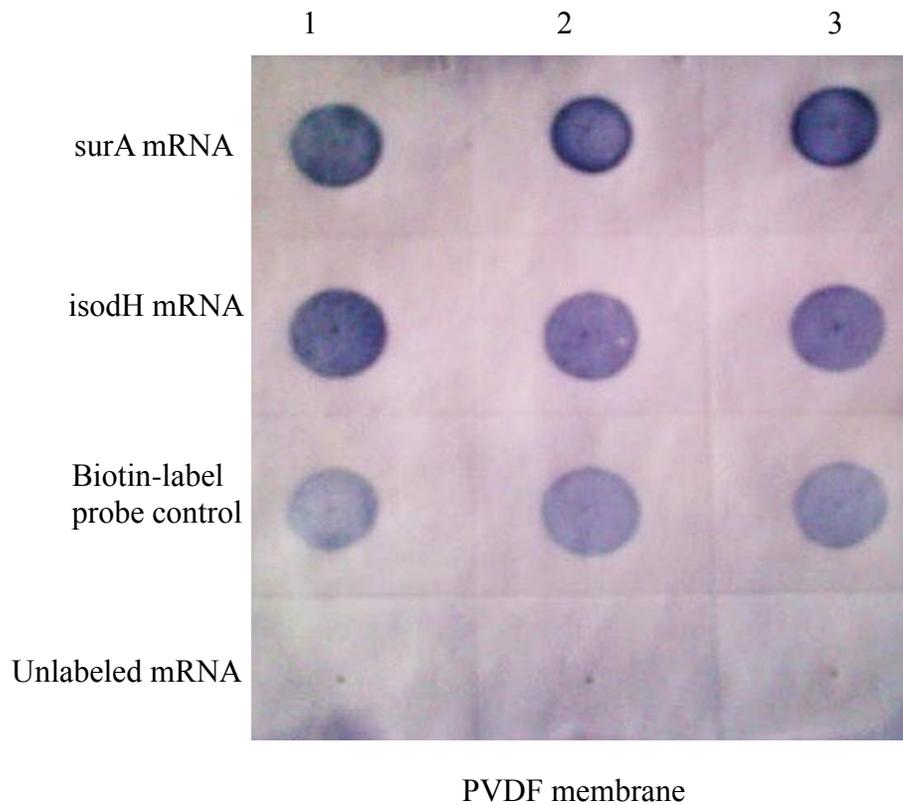


Figure 11 Detection of biotinylated mRNA on PVDF membrane.

The biotinylated reaction of surA and isodH mRNAs was performed for 3, 6 or 16 hours as shown in columns 1, 2 and 3 respectively. Biotin-label transcript control and unlabeled mRNA were detected using alkaline phosphatase-conjugated streptavidin from the biotin chromogenic detection kit.

Table 4 Dynabeads streptavidin-biotin assay.

In vitro assay of supernatant collected after the incubation of 1 nmol of biotin with different concentration of Dynabeads Streptavidin. The above data indicated that the amount of beads conjugated to biotin increased with the amount of Dynabeads.

mg of Beads/Reaction	%Free Biotin	% Biotin Bound to Beads
0	100	0
0.2	93	7
0.4	54	46
0.6	28	72
0.8	18	82

3.8 Pull-Down Assay

A pull-down assay employing biotinylated RNA to characterize and identify RNA binding proteins from *E. coli* was performed. To stabilize the nucleic acid protein interaction, UV crosslinking was used. Finally the purified targeted proteins were visualized after electrophoresis on an SDS-PAGE gel. In Figure 12 and 13 the lanes of crude extract of *E. coli*, SurA and IsodH were identical and there was not a unique band generated from pull down assay. Another attempt at purifying RNA binding factors was performed. A duplicate amount of biotinylated RNA was cross linked to an *E. coli* proteins extract followed by an extra incubation at 4 °C for 2h. Proteins were collected as before after binding the extract with Streptavidin-Coupled Dynabeads. Analysis of the proteins by electrophoresis on SDS-PAGE did not identify any target protein that specifically bound to surA mRNA or isodH mRNA although a number of proteins in the *E. coli* extracts were identified that bound non-specifically to the Dynabeads (Figure 14).

The *phoA* and *gMP* transcripts were shown to interact with several protein species as indicated by the presence of specific bands following analysis on 15% SDS -PAGE of pull down assays (Figure 15 A). Additional experiments to isolate binding proteins in the *E. coli* extracts following an extra incubation with the a duplicate amount of same *phoA* and *gMP* transcripts, showed similar results with identification of similar protein bands as shown in Figure 15 B.

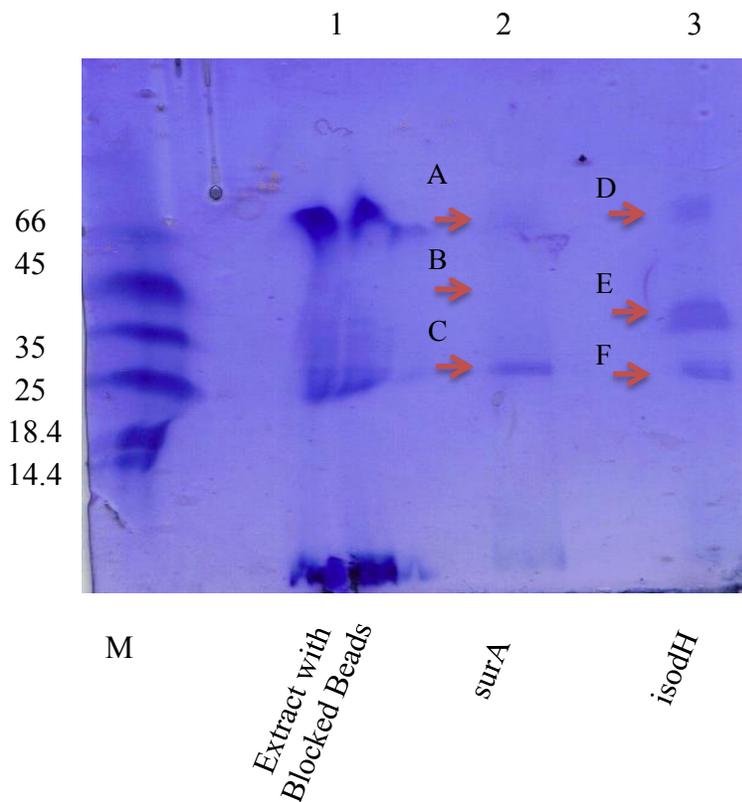


Figure 12 SDS - PAGE gel from the pull down assay of *surA* and *isodH* transcripts.

The biotinylated *surA* and *isodH* mRNAs were incubated with 1 mg of *E. coli* protein extract in binding buffer (3ml of 1X TBS [50 mM Tris-Cl, pH 7.6; 150 mM NaCl], 0.05% TritonX-1) at 4°C for 1h followed by the addition of blocked beads (blocked with 5% casein for 2 h) for an additional hour at 4°C. The bands that were generated from pull down assays with two transcripts are shown in this image. The bands identified with arrows and labeled (A, B, C, D, E & F) refer to the protein identification that is shown in Table 5. Lane 1 is referred to the *E. coli* extract that incubated with blocked beads (with 5 % casein in TBS for 2 hours) that was used as control. M=Marker (kDa).

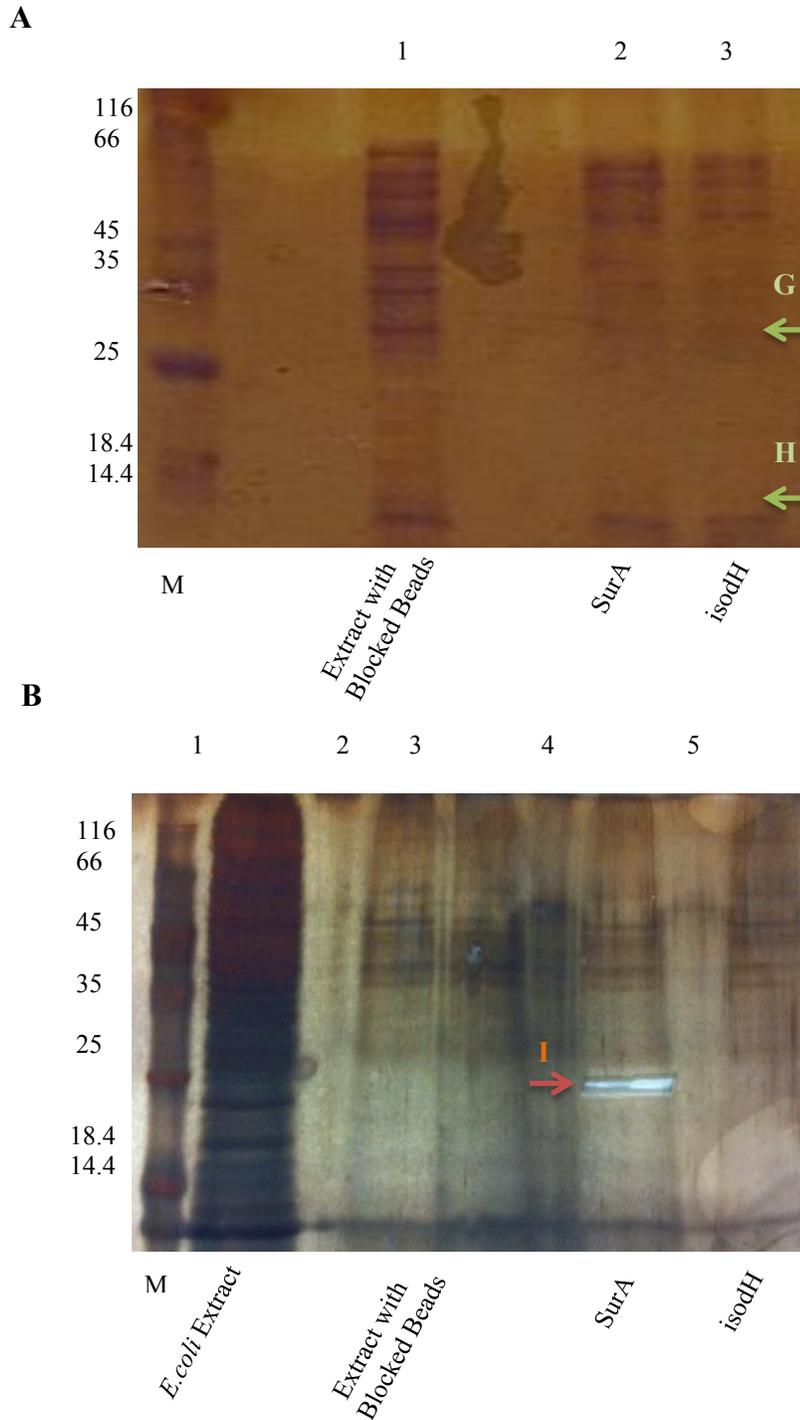


Figure 13 Silver stained gels of proteins from the pull down assay using biotinylated probes from *surA* and *isodH* baits.

Letters and arrows indicate the protein bands that were selected for tryptic digestion and identification using MALDI-TOF (Tables 5 and 6).

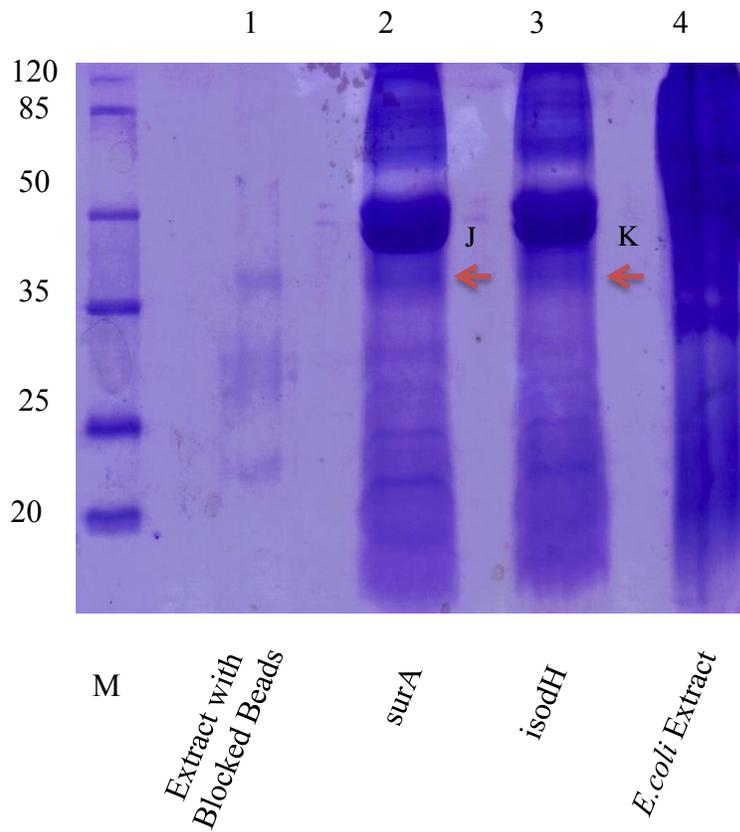


Figure 14 SDS -PAGE gel from optimizing the pull down assay of *surA* and *isodH* transcripts.

A duplicate amount (80-100 μg) of *surA* and *isodH* mRNAs were incubated longer with *E.coli* supernatant proteins in binding buffer (3ml of 1X TBS [50 mM Tris-Cl, pH 7.6; 150 mM NaCl], 0.05% TritonX-100). The baits then were pulled-down using blocked Dynabeads (blocked with 5% casein in TBS for 2 h). There was no specific protein band in lanes 2 (*surA* mRNA) and lane 3(*isodH* mRNA) whereas distinct bands marked (J and K) were subjected to trypsin digestion identification using mass spectrometry analysis (Table 5 and 6). Lane 4 is referred to *E.coli* protein extract while lane 1 refers to the *E.coli* extract that was incubated with blocked beads (with 5 % casein for 2 hours) that was used as controls. M=Marker (kDa).

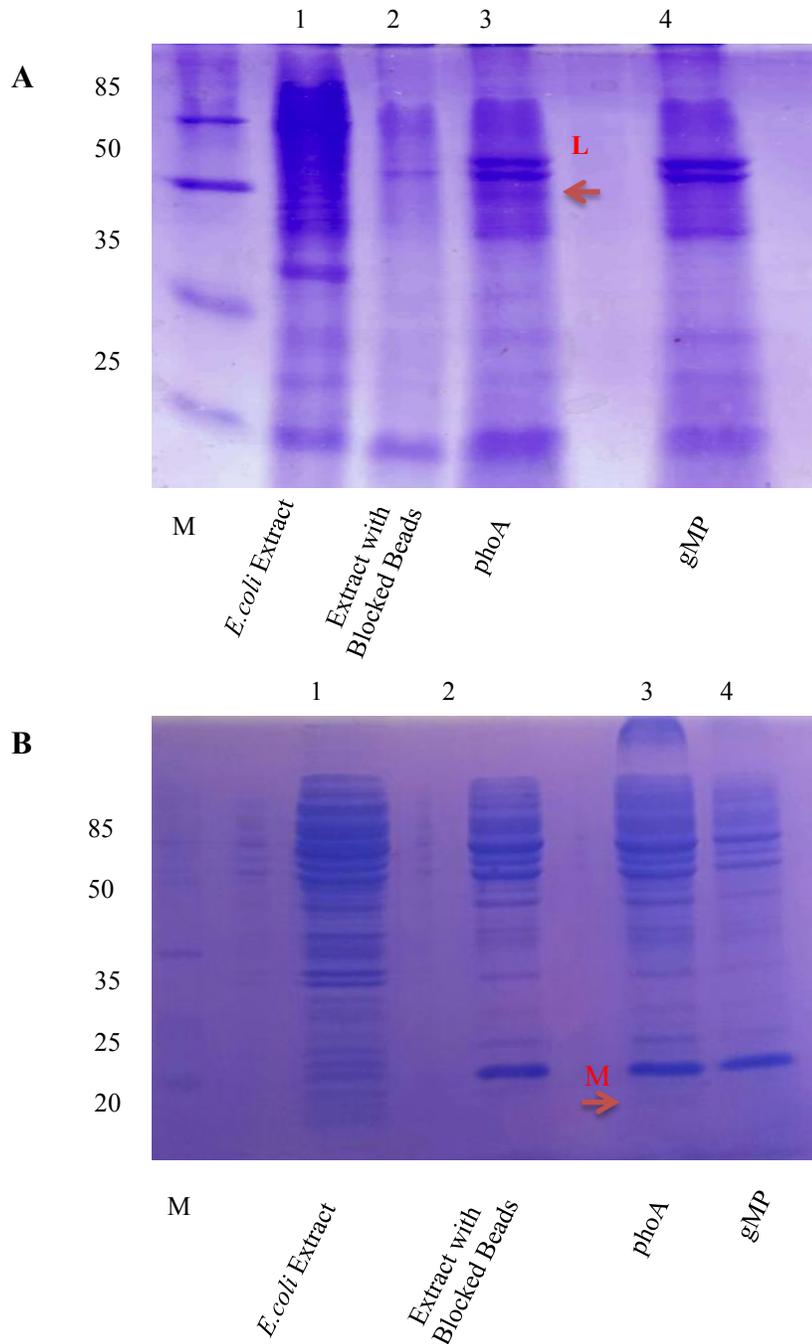


Figure 15 Separation of proteins from the pull down assay using *phoA* and *gMP* transcripts as baits.

The pull down assay with an extra incubation of transcripts with *E. coli* protein extract in binding buffer (3ml of 1x TBS [50 mM Tris-Cl, pH 7.6; 150 mM NaCl], 0.05% TritonX-100) is represented in panel B rather than panel A. Distinct bands marked (L) and (M) were subsequently subjected to trypsin digestion and identification using mass spectrometry (Table 7).

3.9 Effect of Inhibitors Against RNA Degradation

The SDS PAGE gel showed that there was no effect of ATA and DEPC inhibitors on the pull down assays for *phoA* and *gMP* mRNAs when compared against the non-treated pull-down assay control (Figure 16). The generated protein bands from SDS PAGE would support the fact that these chemicals have no effect on pull down assay. Subsequently, lanes 4, 5, 6 and 7 in Figure 16 from pull down assays of *phoA* and *gMP* mRNAs that were treated with ATA and DEPC are similar to the assays that were not treated (Lanes 2 and 3). Lane 1 is refers to an *E.coli* protein extract while all other lanes are refer to mRNAs pull down assays of *E.coli* protein extracts incubated with blocked beads (with 5 % cosine in TBS buffer).

3.10 Protein Capture Using Nylon Membranes

Beside the pull-down assay with Dynabeads, RNA slot-blotting technique was also used to detect any proteins that bind to the *phoA* mRNA of secreted protein rather than *gMP* mRNA of cytoplasmic protein. SDS PAGE showed that there was no specific protein bind to mRNA of secreted protein in comparison with mRNA of cytoplasmic protein.

3.11 RNA-Protein Interaction Assay

After the separation of *E.coli* protein extract in native PAGE gels, each lane was cut and placed horizontally on denaturing gel with a 4% stacking and 12% resolving gel containing either *E.coli* extract as control (Figure 17 A) or contain the transcript of *phoA* (Figure 17 B). The gels were

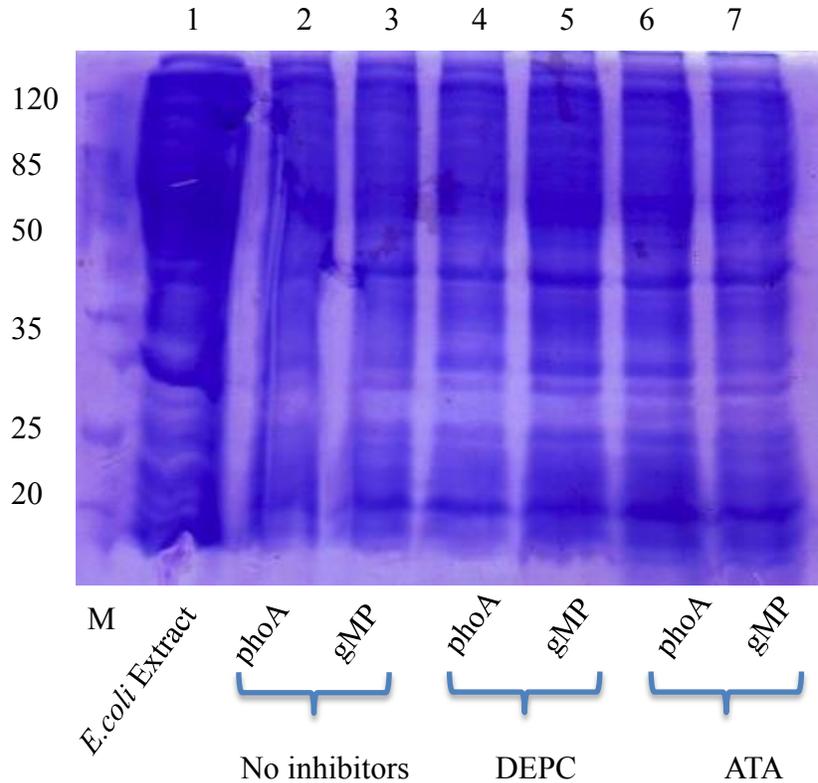


Figure 16 The effect of inhibitors on pull-down assay with *phoA* and *gMP* mRNAs.

E.coli protein extracts were treated with 10mM of Aurintricarboxylic acid (ATA) and Diethylpyrocarbonate (DEPC) inhibitors then incubated with *phoA* and *gMP* transcripts in binding buffer (3ml of 1X TBS [50 mM Tris-Cl, pH 7.6; 150 mM NaCl], 0.05% TritonX-100). The reactions then proceed to the pull-down assay with blocked beads (blocked with 5% casein in TBS for 2 h). SDS-PAGE gel shows that DEPC (lanes 4 and 5) and ATA (lanes 6 and 7) inhibitors have no effect on pull-down assay of *phoA* a transcript from secreted protein and *gMP* a transcript from cytoplasmic protein compared with non - treated pull down assays (lanes 2 and 3). M=Marker (kDa).

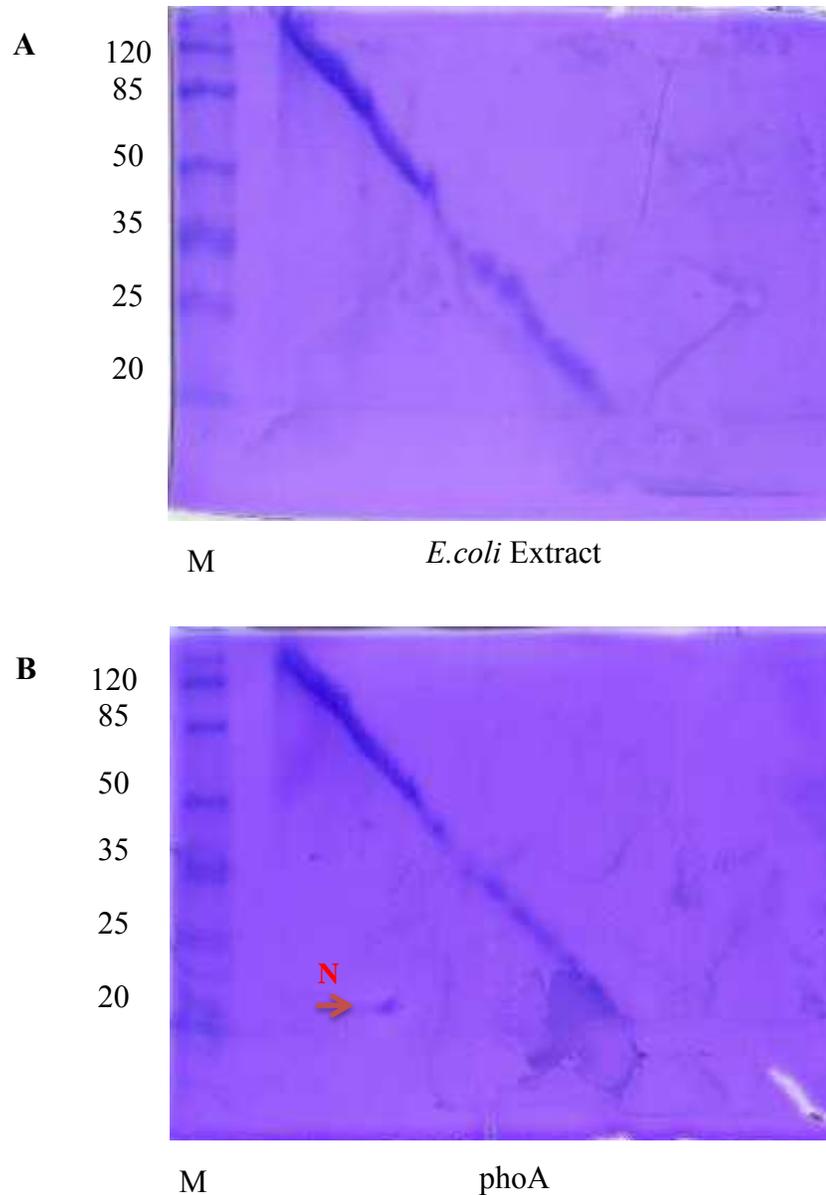


Figure 17 RNA- protein interaction assay.

After the 12 % native polyacrylamide gels were performed with *E.coli* extract the cut lanes were pressed to 12 % SDS-PAGE gel in order to capture any aggregation that might refer to RNA - protein complex. Electrophoreses of *E.coli* extract in gels contain either control (A) or contain the transcript of *phoA* generated from *in vitro* transcription reaction (B). A distinct spot marked (N) was subsequently subjected to trypsin digestion identification using mass spectrometry analysis. M=Marker (kDa).

stained with Coomassie brilliant blue and a diagonal migration of proteins was generated for both gels. An aggregation in Figure 17 B (spot N) refers to the binding of the mRNA to proteins in the extract. The binding of the mRNA to protein will change the effective migration of the protein away from the diagonal pattern seen in the extract alone. MALDI MS identification of spot (N) assigned it as Putative type II secretion system L-type protein YghE .

3.12 Protein Identification by Mass Spectrometry

Throughout all the RNA labeling and pull down steps for the RNA binding factors there was one major band (band A in the surA pull down assay Figure 12) that consistently appeared at approximately 45-50 kDa. The band appeared on several of the gels and was tentatively putatively identified as a signal recognition particle (SRP) component, FtsY-like protein or DEAD/DEAH box helicase following MALDI-TOF-dependent characterization. In addition at protein bands identified in the surA pull-down experiments between 20-40 kDa (band B and C), respectively, were putatively identified as the Nucleoid-associated protein YbaB, and putative protein RhsE (or other possible proteins) (see Table 5 below). As a final attempt to isolate the RNA binding factor protein bands (I and J) present in the surA mRNA pull down assays were further excised for MALDI identification after a trypsin digestion step. The results of this analysis indicated that the proteins captured by surA pull downs included different candidates possibly identified as the HTH-type transcriptional regulator ZntR and the ABC transporter ATP-binding protein that might have several functions in *Escherichia coli*.

On the other hand, the MALDI mass spectroscopy results showed that the proteins that bound to the isodH mRNA (band D, E and F) included candidate proteins such as the Fused protein chain

initiation factor 2 (IF-2), 3-isopropylmalate dehydrogenase and Transcriptional regulatory protein YehT. In additional pull down attempts with the isodH mRNA, proteins bands (G, H and K) were identified using MALDI-MS as the DEAD/DEAH box helicase, Acetyltransferase, GNAT family and ATP-dependent helicase Lhr (Table 6).

For the isolation of a specific protein factor that binds the *phoA* mRNA rather than *gMP* mRNA the pull down assay showed a unique band following SDS-PAGE for the *phoA* mRNA in Figure 15 A and B that was not present in extracts interacted with *gMP* mRNA or in the total *E. coli* extract. The MALDI mass spectroscopy identified the band (L) as the Poly (A) polymerase I at 50 kDa. The smaller 20 kDa band (band M) was similarly identified as the 30S ribosomal protein S10, the 50S ribosomal protein L28, and the fumarate and nitrate reduction regulatory protein in the second *phoA* mRNA pull down (band M). Furthermore, beside all previous identification, the digestion of spot (N) in Figure 17 B was identified as Putative type II secretion system L-type protein YghE (see Table 7 below).

Table 5 Proteins identified from the pull down assay with surA mRNA using MALDI-MS identification.

Protein Identified from Pull Down Assay of surA mRNA	Protein Function
<i>Staphylococcus aureus</i> SRP	Targeting membrane proteins and eventually to the co-translational secretion pathway
<i>Staphylococcus aureus</i> FtsY-like protein	An essential protein component of the secretory complex
DEAD/DEAH box helicase proteins	Ribosome biogenesis, RNA metabolism, mRNA decay and possibly translation initiation in <i>E.coli</i>
Nucleoid-associated protein, YbaB	Binds DNA
Multifunctional conjugation protein TraI	DNA helicase

Table 6 Proteins identified from pull down assay of isodH mRNA using MALDI-MS identification.

Protein Identified from Pull Down Assay of isodH mRNA	Protein Function
3-isopropylmalate dehydratase enzyme	Might belong to any of the translation factors
Fused protein chain initiation factor 2 (IF-2)	Encouraging the initiation of protein synthesis in prokaryotes
DEAD/DEAH box helicase proteins	Ribosome biogenesis, RNA metabolism, mRNA decay and possibly translation initiation in <i>E.coli</i> .
Acetyltransferase GNAT family protein	Involved in resistance to some antibiotics
ATP-dependent helicase Lhr	Associated in RNA turnover, processing, translation control, and ribosome biogenesis
Multifunctional conjugation protein TraI	DNA helicase

Table 7 Proteins identified from pull down assay of phoA mRNA using MALDI-MS identification.

Protein Identified from Pull Down Assay of phoA mRNA	Protein Function
Poly (A) polymerase I enzyme (PAPI)	Polyadenylation
30S ribosomal protein S10	Facilitate binding of tRNA to the ribosomes
50S ribosomal protein L28	Ribosome assembly
Putative type II secretion system L-type protein YghE	Involved in a type II secretion system for the export of proteins
Fumarate and nitrate reduction (FNR) regulatory protein	Transcription factor

Chapter 4: Discussion

Protein secretion is an essential, and therefore evolutionarily conserved, cellular process. The importance of this process is further displayed by the finding that there are of up to eight different protein secretion systems in prokaryotes (Saier, 2006). By far, the most conserved and best recognized as the general secretory system is the Sec-dependent system. In prokaryotes, the signal peptide hypothesis elegantly maps how pre-secretory proteins are recognized by the cell, targeted to the Sec translocon within the inner membrane, and eventually exported outside the inner membrane. In certain studies however, this hypothesis appeared not to be held as some experimental findings, could not be explained through this hypothesis (Habyarimana et al., 2013; Samander et al., 2013). Recent studies on the secretion of the effectors of Type III secretion in *Yersinia* species have provided a plausible alternate hypothesis whereby the pre-secretory proteins are recognized and tagged for secretion through signals in the 5' region of the proteins' mRNA rather than through signals in the protein's amino acid signal sequence (Habyarimana et al., 2013). This became known as the mRNA hypothesis. In an effort to explain the experimental observations where the signal peptide hypothesis did not hold for certain Sec-dependent secreted proteins, I hypothesized that in addition to recognizing the signal peptide, the SRP also requires simultaneous recognition of the 5' mRNA of the secretory protein. I called this the alternate hypothesis (Figure 3 B).

In this study I examined the possibility that the Sec-dependent secreted proteins are targeted to the Sec system via signatures in the mRNA structure rather than/or in addition to, signatures (signal peptides) in the protein structure. I looked at Sec targeting indirectly by establishing interactions between the mRNA test constructs and the SRP equivalent in *E. coli* lysates,

including the fifty four homologue (Ffh) cytoplasmic factor (Schneewind, et al., 2012; Braig, et al., 2011; Yosef, et al., 2010; Rosch, et al., 2008; Angelini, et al., 2006) Association between the test mRNA constructs and Ffh or other cytoplasmic factors was followed using a pull down assay. In this assay, the mRNA was generated using *in vitro* transcription reactions of DNA segments representing the 5' UTR of secreted (test) and cytoplasmic (control) gene products as well as segments representing the signal peptide or its equivalent in a cytoplasmic protein (N-terminal 25 amino acids). These transcripts were then used to pull down cytoplasmic factors from total cell lysates from *E.coli* that are then identified using mass spectrometry. Using this approach, several cytoplasmic factors were pulled down and identified (Tables 5, 6 and 7).

The surA protein was used as the secretory model protein. SurA is formally identified as a periplasmic propyl isomerase chaperone of *E. coli*, which is required for proper biogenesis of outer membrane proteins (Lazar et al., 1996, 1998; Bitto et al., 2003; Sklar et al., 2007). This protein contains a typical signal sequence in its N-terminus and could potentially provoke further insight into characterization of the RNA binding proteins (RBPs) in *E. coli*. Moreover, it could open new avenues to understand the function of RBPs in the translocation machinery that is located in the cytoplasmic membrane. The 3-isopropylmalate dehydratase enzyme was used in this study as a control for non-secreted proteins, given that it is a cytoplasmic enzyme. This enzyme is actually involved in the biosynthesis of the amino acid leucine (Velíšek et al., 2006; He et al., 2010).

4.1 mRNA Structure of Different Transcripts

In considering the alternate hypothesis and to propose a potential role for the mRNA in targeting and secretion of pre-secretory proteins, one must define the features of the mRNA that may be

involved in this targeting. A previous study has pointed to the possible role that the ratio of A (T)/U could play in mediating specific recognition of the mRNA by Ffh (Samander et al., 2013). Despite the fact that the sequences of the mRNA of the secreted and cytoplasmic proteins are different, mRNA secondary structure is also an important factor to be considered (Figure 10).

What is evident from these simulated structures is that all four mRNA's are different, even the two mRNA's of the secretory proteins (SurA and PhoA). Lack of an obvious common feature between the mRNA coding for the secretory proteins or between the cytoplasmic proteins' mRNAs (IsodH and GMP) does not support the proposition that these mRNA's would be different or similar enough to bind to specific cellular factors that segregate them from each other and target them to different cellular locations (secreted vs. cytoplasmic). It can be argued however, that these are only simulated structures that were prepared using one simulation program. The actual structures may in fact have features sufficient for the cell to segregate them apart from each other. Alternatively, the proposition of Samander et al. (2013) may prove to be correct by identifying specific linear sequences as the signal recognized by secretory factors such as Ffh and not secondary structures. These questions remain to be answered and would shed more light onto the specific features of the mRNA needed for targeting.

4.2 Factors Identified Through The Pull-Down Assay

Pull-down assays using the synthetic mRNAs for surA and isodH followed by analysis of enriched proteins by SDS-PAGE identified several different bound proteins (Figure 12 and 13). Hence, there does not seem to be any protein that specifically binds to one mRNA over the other. Following optimization of the assay, distinct SDS-PAGE patterns were shown to distinguish

binding between the two mRNA baits (Figure 14). In order to establish the identity of these factors, the bands were excised for in-gel trypsin digest followed by MALDI MS identification.

Of the proteins identified from the pull down assay with *surA* mRNA (Table 5), proteins corresponding to the *Staphylococcus aureus* SRP protein and *S. aureus* FtsY-like protein are perhaps the most curious. These proteins, which have been discovered in Gram-positive as well as Gram-negative bacteria, are involved in membrane proteins targeting and eventually to the co-translational secretion (Schneewind, et al., 2012;). Repeated experiments with this *surA* mRNA construct did not show any specific association with *E. coli* SRP.

Remarkably, other proteins identified in this assay include the 3-isopropylmalate dehydratase enzyme; the fused protein chain initiation factor 2 (IF-2), DEAD/DEAH box helicase protein and Acetyltransferase GNAT family protein. Some of these proteins, as in the case of the *isodH* mRNA pull down assay, have established functions related to mRNA unfolding, processing, translation control, or decay. Particularly in *E. coli*, most of the proteins would be expected to bind the mRNA of the 3-isopropylmalate dehydratase enzyme as those are the ones involved in the critical steps in normal translation of mRNA (Marintchev et al., 2004; Kaczanowska et al., 2007). The predicted proteins isolated via pull-down assay should belong to any of the prokaryote translation factors that were identified in Table 6.

Interestingly, MALDI MS identified the presence of the 3-isopropylmalate dehydratase enzyme (a 39 kDa protein) in extracts pulled down by the mRNA (Table 6). This result could potentially provide insight into the fact that this type of enzyme might belong to any of the translation factors (Table1). Thus, we can establish that this type of binding interaction involves the regulation of the mRNA product.

The fused protein chain initiation factor 2 (IF-2) has been identified from MALDI MS analysis in Table 6. It is a GTPase that has a significant role in promoting the initiation of protein synthesis in prokaryotes. IF2 enhances the binding of 30S ribosomal subunit to the initiator tRNA and thus controls the delivery of tRNA into the ribosome (Table 1) (Milon, 2010). This result would actually support the interactions between translation factors and RNA sequences of cytoplasmic proteins, which is believed to be subsequently involved in protein biosynthesis (Milon, 2010).

In addition to IF-2, DEAD/DEAH box helicase proteins have been identified. This protein family contributes to ribosome biogenesis, RNA metabolism, mRNA decay and possibly translation initiation in *E.coli*. Most of DEAD box proteins participate in the assembling RNA or ribonucleoprotein (RNP) structure *in vitro*. This is related to its function as ATP-dependent RNA helicases and RNA-dependent ATPases (Iost et al., 2006). The presence of DEAD/DEAH box helicases from the pull-down assay is logical according to their function. It has been proven that the Acetyltransferase GNAT family proteins in bacteria appear to be involved in resistance to some antibiotics. However, there is no previously identified function for these protein types in mRNA translation or processing (Vetting et al., 2005). I can conclude that participation of the GNAT in mRNA recognition is not known.

Some of the other proteins identified, however, are interesting in that they appear only in the pull down assay using the *surA* transcript and are not present in the *isodH* mRNA pull-downs. For example, the Nucleoid-associated protein, YbaB, specifically interacts with the *surA* mRNA and has been previously identified as a factor that binds DNA and thus changes its conformation (Table 5). Additionally, this protein appears to be involved in regulating gene expression

(Cooley et al., 2009). Proteins that bind DNA have been found also to bind RNA in many instances as has been demonstrated for the DNA binding protein H-NS (Brescia, et al., 2004).

Different proteins identified in pull down assay with the IsodH mRNA include DNA-dependent helicase, Lhr, which is a SF2 helicase and one of the longest known proteins in *E. coli* containing motifs from the helicase superfamily II (Ordonez et al., 2013; Reuven et al., 1995). Further study proved that this polypeptide also contains an amino region similar to the DEAD family helicases (Reuven et al., 1995). Ordonez et al (2013) have shown that some of the helicases have established functions related to RNA turnover, processing, translation control, and ribosome biogenesis.

In a similar manner the Multifunctional conjugation protein TraI (Tables 5 and 6) has been identified as a DNA helicase I. Furthermore, TraI has been shown to have a major role in binding and cleaving single-stranded DNA oligonucleotides containing an oriT sequence in plasmid transfer experiments (Street, et al., 2003; Dna et al., 2001) TraI also functions in the initiation and/or termination of plasmid transfer (Dna et al., 2001).

4.3 Different Transcripts Construct (5' UTRs of *phoA* and *gMP* genes)

It is quite possible that binding of a translation factor is dependent not only on the 5' translated region of the mRNA but in fact requires sequences in the untranslated region (UTR) (Berg et al., 2009). Direct evidence supporting the importance of the UTR comes from the recent observation that the ribosomal protein S1 (Hajnsdorf et al., 2012; Berg et al., 2009; Sørensen et al., 1998; Kalapos et al., 1997), which is essential for *in vivo* translation in *E. coli* (Sørensen et al., 1998) can bind to an 11 nucleotides region in the mRNA immediately upstream of the Shine-Dalgarno (SD) sequence (Berg et al., 2009; Sengupta, et al 2001). The SD sequence is a short sequence of

mRNA that interacts with anti-SD sequences in the 3' end of 16S rRNA which is involved in the initiation of protein translation (Malys, 2012; Berg et al., 2009; Laursen et al., 2005). The SD sequence-S1 protein interaction further plays a significant role in effecting gene expression (Berg et al., 2009). Additionally, it has been demonstrated that there is a parallel approach in *Pseudomonas aeruginosa* that further supports a role for the 5' UTR in *E. coli* translation regulation (Sevo et al., 2004). This evidence indicates that S1 protein preferentially binds within the first 78 nucleotides of the 5' UTR of the *rpoS* mRNA and thus may contribute to the translational regulation of *rpoS* mRNA (Sevo et al., 2004).

Interestingly, since the UTR sequence is considered to be important for targeting the location of mRNA to specific cellular locations via interactions with specific proteins, it seems likely that protein binding to the UTR of different transcripts may have other functional roles. It was therefore of interest to establish an artificial transcript that includes the 5' UTR of *phoA* and *gmp* genes for use in the pull-down experiments. Using this construct as bait I was able to pull down a protein, which has been previously shown to be a polyadenylation-regulating protein. SDS-PAGE shows a unique 50 kDa band that was present in the pull-down experiment using the *phoA* mRNA bait but not in the controls (Figure 15 A). The band was identified as the poly (A) polymerase I enzyme (PAPI) that is actually responsible for the addition of adenylyl residues to the 3' hydroxyl termini of new mRNA molecules (Sarkar, 1997). The mRNA bait constructs that were used in these experiments do not have poly A tails since the *in-vitro* transcription reaction I used does not generate poly A tails. However, it is possible that the PAPI proteins are able to bind to other parts of the mRNA. For example, it is possible that the poly (A) polymerase I protein is able to bind first to the RNA and then initiate the polyadenylation process (Kushner et al., 2011). Hence, these results cannot exclude the possibility that sequence in the UTR of

mRNA might be able to bind several proteins that are necessary for the different functions (Table 7).

The S10 protein of *Escherichia coli* is a member of the ribosomal protein (r-proteins) group that is located in the head of the 30 S subunit of the ribosome (Yaguchi et al., 1980). A major function of S10 is its ability to mediate binding of tRNA to the ribosomes (Yaguchi et al., 1980; Zurawski et al., 1985). Studies have also reported that the S10 protein has a role in anti-termination of the transcription process (Das et al., 1985; Friedman et al., 1981). It has been shown that the 50S ribosomal protein L28 cross-links to two distinct RNA regions in *E.coli* (Osswald et al., 1990; Urlaub et al., 1995). Moreover, this protein has a critical role in ribosome assembly and the absence of L28 also prevented the synthesis of the ribosome (Maguire et al., 1997). Further experiments indicated that the function of ribosomal proteins is not only restricted to the ribosome but that it has another role such as mediating the association of ribosomal protein L22 during export of a secreted virulence factor antigen 43 (AG 43) in *E.coli* (Yap et al., 2013). Additionally, it has been verified there is a connection between the physiological roles of ribosomal proteins and RNase since it has been shown that RNase co-migrates with 30S ribosomal subunit (Malecki et al., 2014). Our results might promote the idea that there are additional functions of ribosomal protein in *E .coli*.

The putative type II secretion system L-type protein YghE is another interesting protein identified in the pull-down assays (Figure 17 B, spot N). YghE belongs to the general secretion pathway (GSP L family) that locates in the inner membrane. This protein family is involved in a type II secretion system for the export of proteins (Filloux, 2004). The MALDI MS identification of pulled down proteins further identified the Fumarate and nitrate reduction (FNR) regulatory protein that is recognized as a global transcription factor. FNR regulates a

large family of genes in response to environmental conditions and also appear to bind DNA (Melville et al., 1997; Gostick et al., 1998).

Altogether I was unable to pull down secretion-dependent factors that recognize and bind selectively to the mRNA of secreted proteins. The proteins that were identified are typically associated within mRNA processing. Some of these proteins have been also pulled down using similar methodologies in different laboratories. Maier et al (2008) demonstrated that the SRP - protein KdpD interaction in *Escherichia coli* is required for KdpD protein targeting. The complex was further confirmed using pull down experiments.

4.4 The Streptavidin-Magnetic Beads Capacity

According to the binding capacity for one mg of the Dynabeads streptavidin can bind approximately ~200 pmol single strand-oligonucleotides. Assuming that a single RNA can bind to one molecule of protein (1:1), the beads could occupy the maximum of ~200 pmol of RNA and therefore 200 pmol of protein molecules. The 200 pmol of protein assuming an average mass of 40 kDa this will translate to approximately a mass of 8 microgram of protein (mole of the substance = gram of substance/Molecular Weight (MW)). Therefore, I established that ~200 pmol of beads can bind a theoretical concentration of the 8 microgram (the binding capacity) of biotinylated proteins. This analysis has been done in order to demonstrate that these beads are able to bind enough protein in order to detect a specific target protein on SDS- PAGE gel.

4.5 The Dynabeads Beads Blocking

Blocking of the dynabeads prior to the pull down assay is very important to reduce nonspecific protein binding. In our study, I investigated different blocking solutions including casein, skim

milk, bovine serum albumin (BSA) and gelatin. I have found that casein was in fact the best blocking solution that I can use judging from the weak background that I have seen in Figure 13, 14 and 15.

4.6 Potential Degradation of mRNA in the Procedures

It has been demonstrated that RNase enzymes catalyze the degradation of RNA and therefore affect various processes within the cell (Deutscher, 2006). In order to protect the samples against degradation from environmental RNases and to obtain a more detailed picture of RNA properties, it will be necessary to examine the effectiveness of such inhibitors. I observed that inhibitors such as DEPC (Diethylpyrocarbonate) and ATA (Aurintricarboxylic acid) were ineffective with regard to the pull-down assay as compared with samples that have not been treated with any inhibitors, as shown in lanes 2 and 3 in Figure 16.

Aurintricarboxylic acid has no effect on pull down assays of both *phoA* and *gMP* mRNA's as shown in Figure 16 (lanes 6 and 7) which illustrated the same profile as non-treated pull down assays in lanes 2 and 3. Our previous finding is in agreement with the evidence of that ATA is associated with the inhibition of general nuclease such as DNase I, RNase A, exonuclease III and SI nuclease and several enzymatic reactions of initiation and elongation of protein synthesis (Hallick et al., 1977).

The effect of diethylpyrocarbonate (DEPC) on the pull-down assay was further examined. It has found DEPC to be efficient and a general inhibitor of RNases. In addition, it inhibits RNases through the modification of Cysteine, Tyrosine, and Lysine and primarily with Histidine residues (Zhou et al., 2013). Our results show that there is no detectable effect of DEPC on RNA - protein

interaction. This is explained the similarity between lanes 4 and 5 in comparison to the non-treated pull down assays in lanes 2 and 3 (Figure 16).

Therefore, although the use of these agents for the protection of the mRNA from degradation during the assays these RNase inhibitors work by inhibiting the interaction between the RNase and the RNA itself. Since the target of these experiments (Ffh) is also postulated to bind these mRNAs, it is conceivable that their use also affected the ability of these baits to pull down the Ffh.

Chapter 5: Conclusion

The model of secreted pre-protein recognition I proposed was based on experimental data from other laboratories working with signal sequences of type III secreted factors of *Yersinia*. In this model, I proposed that pre-protein recognition involves the simultaneous recognition of both the N-terminal signal sequence of the nascent polypeptide and the 5' end of the mRNA. The most direct method was to show that the mRNA of a secreted protein binds to the Ffh secretion factor. The experiments were carried out with *E. coli* proteins. Two secreted proteins (SurA and PhoA) and two cytosolic proteins (IsodH and GMP, as controls) were used to see whether Ffh would preferentially interact with the secreted but not the cytosolic proteins. Ideally, direct binding experiments with a purified Ffh would be conducted but I was not able to get the *E. coli* Ffh protein.

The pull-down assays using these mRNAs as bait did show differences in the types of proteins that both types of mRNAs (of secreted vs cytosolic) were able to pull down as judged by SDS-PAGE and mass spectrometry. The key factor, Ffh, however was not amongst them. The results do not support the proposal that Ffh can interact with the mRNA of secreted proteins and therefore a role for mRNA in targeting proteins for secretion through the general secretory pathway remains to be proven.

5.1 Limitations of Study

One limitation of this study lies in the nature of one of the mRNAs used to detect RNA binding proteins. It is also important to point out that the periplasmic propyl isomerase chaperone SurA is exported to the periplasmic space but is not completely secreted. Potentially it is possible that

other factors will not interact with the mRNA of periplasmic proteins in same way as secreted proteins.

Care needs to be taken in considering the results of blocking experiments as they relate to prevent non-specific protein binding. The backgrounds to each sample of *E.coli* extract incubated with beads from pull down assays were provided in the SDS-PAGE gels, as well as the background related to the alternate membrane binding strategy. The protein pattern from the blocking experiments in fact does not support the idea of preventing a number of proteins that binds non-specifically to RNA.

Other factual issues also the small size of mRNAs. The MS analysis identified proteins that do not interact normally with mRNA however they were isolated in this procedure indicating a need for greater specificity in the pull down assay to target true RNA binding proteins. This limitation also meant that determination the nature of mRNA such as length and size is necessary to consider.

Care was given in explaining each individual method and process as these related to the overall research purposes, as a means of clarifying these limitations.

5.2 Future Directions

- Clone the *ffh* gene and examine the binding either with mRNA alone or with mRNA and signal peptide.
- Isolate the ribosome/mRNA complex followed by purification of any RNA binding factors.
- Determine the size of mRNA that further reveals a number of code protein would bind.

References

- Anderson, D. M., & Schneewind, O. (1997). A mRNA signal for the type III secretion of Yop proteins by *Yersinia enterocolitica*. *Science*, **278**, 1140-3.
- Angelini, S., Boy, D., Schiltz, E., & Koch, H.-G. (2006). Membrane binding of the bacterial signal recognition particle receptor involves two distinct binding sites. *The Journal of Cell Biology*, **174**, 715-24.
- Bae, W., Xia, B., Inouye, M., & Severinov, K. (2000). *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 7784-9.
- Bashirullah, A., Cooperstock, R. L., & Lipshitz, H. D. (1998). RNA localization in development. *Annual Review of Biochemistry*, **67**, 335-94.
- Beckwith, J. (2013). The Sec-dependent pathway. *Research in microbiology*, **164**, 497-504.
- Benelli, D., & Londei, P. (2009). Begin at the beginning: evolution of translational initiation. *Research in microbiology*, **160**, 493-501.
- Berg, L., Lale, R., Bakke, I., Burroughs, N., & Valla, S. (2009). The expression of recombinant genes in *Escherichia coli* can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5'-untranslated part of mRNA. *Microbial biotechnology*, **2**, 379-89.
- Bitto, E., & McKay, D. B. (2003). The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *The Journal of Biological Chemistry*, **278**, 49316-22.
- Blachly-Dyson, E., & Stevens, T. H. (1987). Yeast carboxypeptidase Y can be translocated and glycosylated without its amino-terminal signal sequence. *The Journal of Cell Biology*, **104**, 1183-91.
- Braig, D., Mircheva, M., Sachelaru, I., van der Sluis, E. O., Sturm, L., Beckmann, R., & Koch, H.-G. (2011). Signal sequence-independent SRP-SR complex formation at the membrane suggests an alternative targeting pathway within the SRP cycle. *Molecular Biology of the Cell*, **22**, 2309-23.
- Brescia, C. C., Kaw, M. K., & Sledjeski, D. D. (2004). The DNA binding protein H-NS binds to and alters the stability of RNA *in vitro* and *in vivo*. *Journal of Molecular Biology*, **339**, 505-14.
- Broude, N. E. (2011). Analysis of RNA localization and metabolism in single live bacterial cells: achievements and challenges. *Molecular microbiology*, **80**, 1137-47.

- Christiansen, J. K., Larsen, M. H., Ingmer, H., Søgaaard-andersen, L., Kallipolitis, B. H. (2004). The RNA-Binding Protein Hfq of *Listeria monocytogenes*: Role in Stress Tolerance and Virulence. *Journal of Bacteriology*, **186**, 3355-62.
- Collier, D. N. (1994). *Escherichia coli* Signal Peptides Direct Inefficient Secretion Outer Membrane Protein (OmpA) and Periplasmic Proteins and Alkaline Phosphatase) in *Bacillus subtilis*. *Journal of Bacteriology*, **176**, 3013-20.
- Collinson, I., Breyton, C., Duong, F., Tziatzios, C., Schubert, D., Or, E., Kühlbrandt, W. (2001). Projection structure and oligomeric properties of a bacterial core protein translocase. *The EMBO journal*, **20**, 2462-71.
- Cooley, A. E., Riley, S. P., Kral, K., Miller, M. C., DeMoll, E., Fried, M. G., & Stevenson, B. (2009). DNA-binding by *Haemophilus influenzae* and *Escherichia coli* YbaB, members of a widely-distributed bacterial protein family. *BMC Microbiology*, **9**, 1-8.
- Cuia, J., Huey-Fen Sim, T., Gong, Z., & Shen, H.-M. (2012). Generation of transgenic zebrafish with liver-specific expression of EGFP-Lc3: A new in vivo model for investigation of liver autophagy. *Biochemical and Biophysical Research Communications*, **2**, 268-273.
- Das, a., Ghosh, B., Barik, S., & Wolska, K. (1985). Evidence that ribosomal protein S10 itself is a cellular component necessary for transcription antitermination by phage lambda N protein. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 4070-4.
- Deutscher, M. P. (2006). Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Research*, **34**, 659-66.
- Dna, S., Stern, J. C., & Schildbach, J. F. (2001). DNA Recognition by F Factor TraI36 : Highly Sequence-Specific Binding of Single-Stranded DNA. *Biochemistry*, **40**, 11586-95.
- Du, T. G., Schmid, M., & Jansen, R.-P. (2007). Why cells move messages: the biological functions of mRNA localization. *Seminars in cell & developmental biology*, **18**, 171-7.
- Edman, M., Jarhede, T., Sjöström, M., & Wieslander, A. (1999). Different sequence patterns in signal peptides from mycoplasmas, other gram-positive bacteria, and *Escherichia coli*: a multivariate data analysis. *Proteins*, **35**, 195-205.
- Fekkes, P., & Driessen, a. J. (1999). Protein targeting to the bacterial cytoplasmic membrane. *Microbiology and Molecular Biology Reviews*, **63**, 161-73.
- Feng, Y., Huang, H., Liao, J., & Cohen, S. N. (2001). *Escherichia coli* poly (A)-binding proteins that interact with components of degradosomes or impede RNA decay mediated by polynucleotide phosphorylase and RNase E. *The Journal of biological chemistry*, **276**, 31651-6.
- Filloux, A. (2004). The underlying mechanisms of type II protein secretion. *Biochimica et Biophysica Acta*, **1694**, 163-79.

- Folichon, M. (2003). The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic Acids Research*, **31**, 7302-10.
- Folichon, M., Allemand, F., Régnier, P., & Hajnsdorf, E. (2005). Stimulation of poly(A) synthesis by *Escherichia coli* poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. *The FEBS journal*, **272**, 454-63.
- Friedman, D. I., Schauer, a .T., Baumann, M. R., Baron, L. S., & Adhya, S. L. (1981). Evidence that ribosomal protein S10 participates in control of transcription termination. *Proceedings of the National Academy of Sciences of the United States of America*, **78**, 1115-8.
- Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters*, **582**, 1977-86.
- Goldstein, J., Pollitt, N. S., & Inouye, M. (1990). Major cold shock protein of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 283-7.
- Gouridis, G., Spyridoula, K., Ioannis, G., Charalampos, G., Kalodimos., Anastassios, E. (2009). Signal peptides are allosteric activators of the protein translocase. *Nature*, **462**, 363-67.
- Graumann, P. L., & Marahiel, M. a. (1998). A superfamily of proteins that contain the cold-shock domain. *Trends in biochemical sciences*, **23**, 286-90.
- Habyarimana, F., & Ahmer, B. M. M. (2013). More evidence for secretion signals within the mRNA of type 3 secreted effectors. *Journal of Bacteriology*, **195**, 2117-8.
- Hajnsdorf, E., & Boni, I. V. (2012). Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie*, **94**, 1544-53.
- Hallick, R. B., Chelm, B. K., Gray, P. W., and Orozco, E. M. Jr. (1977). Use of aurintricarboxylic acid as an inhibitor of nucleases during nucleic acid isolation. *Nucleic Acids Research*, **4**, 3055-64.
- He, Y., Chen, B., Pang, Q., Strul, J. M., & Chen, S. (2010). Functional specification of Arabidopsis isopropylmalate isomerases in glucosinolate and leucine biosynthesis. *Plant & cell physiology*, **51**, 1480-7.
- Iino, T., Takahashi, M., & Sako, T. (1987). Role of amino-terminal positive charge on signal peptide in staphylokinase export across the cytoplasmic membrane of *Escherichia coli*. *The Journal of Biological Chemistry*, **262**, 7412-17.
- Iost, I., & Dreyfus, M. (2006). DEAD-box RNA helicases in *Escherichia coli*. *Nucleic acids research*, **34**, 4189-97.
- Jiang, W., Hou, Y., & Inouye, M. (1997). CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. *The Journal of biological chemistry*, **272**, 196-202.

- Kaczanowska, M., & Rydén-Aulin, M. (2007). Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiology and molecular biology reviews*, **71**, 477-94.
- Kaiser, C. A., & Botstein, D. (1986). Secretion-defective mutations in the signal sequence for *Saccharomyces cerevisiae* invertase. *Molecular and Cellular Biology*, **6**, 2382-91.
- Kalapos, M. P., Paulus, H., & Sarkar, N. (1997). Identification of ribosomal protein S 1 as a poly (A) binding protein in *Escherichia coli*. *Biochimie*, **79**,493-502.
- Keiler, K. C. (2011). RNA localization in bacteria. *Current Opinion in Microbiology*, **14**, 155-59.
- Kim, J. S., Park, S. J., Kwak, K. J., Kim, Y. O., Kim, J. Y., Song, J., Jang, B., Jung , C-H., &Kang, H. (2007). Cold shock domain proteins and glycine-rich RNA-binding proteins from *Arabidopsis thaliana* can promote the cold adaptation process in *Escherichia coli*. *Nucleic acids research*, **35**, 506-16.
- Kindler, S., Wang H., Richter, D., & Tiedge, H. (2007). RNA Transport and Local Control of Translation. *Annual Review of Cell and Developmental Biology*, **21**, 223-245.
- Kloc, M., Zearfoss, N. R., & Etkin, L. D. (2002). Mechanisms of Subcellular mRNA Localization. *Cell*, **108**, 533-44.
- Koshland, D. E., JR, Goldbeter, A., & Stock, J. (1982). Amplification and adaptation in regulatory and sensory systems. *Science*, **217**, 220-25.
- Kushner, S .R., and Mohanty , B. K., (2011). Bacterial/archaeal/organelle polyadenylation. *Wiley Interdiscip Rev RNA*, **2**, 256-76.
- Kudva, R., Denks, K., Kuhn, P., Vogt, A., Müller, M., & Koch, H.-G. (2013). Protein translocation across the inner membrane of Gram-negative bacteria: the Sec and Tat dependent protein transport pathways. *Research in microbiology*, **164**, 505-34.
- Laemmli, U.K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
- Lammertyn, E., & Anne, J. (1998). Modifications of Streptomyces signal peptides and their effects on protein production and secretion. *FEMS Microbiology Letters*, **160**, 1-10.
- Lasko, P. (1999). RNA sorting in Drosophila oocytes and embryos. *The FASEB Journal*, **13**, 421-33.
- Laursen, B. S., Sørensen, H. P., Mortensen, K. K., & Sperling-petersen, H. U. (2005). Initiation of Protein Synthesis in Bacteria. *Microbiology and Molecular Biology Reviews*, **69**, 101-23.
- Lazar, S. W., & Kolter, R. (1996). SurA assists the folding of *Escherichia coli* outer membrane proteins. *Journal of bacteriology*, **178**, 1770-3.

- Lazar, S. W., Almirón, M., Tormo, A., Almiro, M., & Kolter, R. (1998). Role of the *Escherichia coli* SurA Protein in Stationary-Phase Survival. *Journal of Bacteriology*, **180**, 5704-11.
- Lee, M.-H., & Schedl, T. (2006). RNA-binding proteins. *WormBook: the online review of C. elegans biology*, **1**, 1-13.
- Lee, V. T., & Schneewind, O. (2001). Protein secretion and the pathogenesis of bacterial infections. *Genes & development*, **15**, 1725-52.
- Lehnhardt, S., Pollitt, N. S., Goldstein, J., & Inouye, M. (1988). Modulation of the effects of mutations in the basic region of the OmpA signal peptide by the mature portion of the protein. *The Journal of Biological Chemistry*, **263**, 10300-3.
- Link, T. M., Valentin-Hansen, P., & Brennan, R. G. (2009). Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19292-7.
- Liu, D.-Q., Liu, H., Shen, H.-B., Yang, J., & Chou, K.-C. (2007). Predicting secretory protein signal sequence cleavage sites by fusing the marks of global alignments. *Amino Acids*, **32**, 493-6.
- Low, K. O., Muhammad Mahadi, N., & Md Illias, R. (2013). Optimisation of signal peptide for recombinant protein secretion in bacterial hosts. *Applied Microbiology and Biotechnology*, **97**, 3811-26.
- Maier, K. S., Hubich, S., Liebhart, H., Krauss, S., Kuhn, A., & Facey, S. J. (2008). An amphiphilic region in the cytoplasmic domain of KdpD is recognized by the signal recognition particle and targeted to the *Escherichia coli* membrane. *Molecular Microbiology*, **68**, 1471-84.
- Maguire, B. a, & Wild, D. G. (1997). The roles of proteins L28 and L33 in the assembly and function of *Escherichia coli* ribosomes in vivo. *Molecular Microbiology*, **23**, 237-45.
- Malecki, M., Bárria, C., & Arraiano, C. M. (2014). Characterization of the RNase R association with ribosomes. *BMC Microbiology*, **14**, 1-9.
- Malys, N. (2012). Shine-Dalgarno sequence of bacteriophage T4: GAGG prevails in early genes. *Molecular Biology Reports*, **39**, 33-9.
- Malys, N., & McCarthy, J. E. G. (2011). Translation initiation: variations in the mechanism can be anticipated. *Cellular and molecular life sciences*, **68**, 991-1003.
- Marintchev, A., & Wagner, G. (2004). Translation initiation: structures, mechanisms and evolution. *Quarterly reviews of biophysics*, **37**, 197-284.

- Maris, C., Dominguez, C., & Allain, F. H.-T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS journal*, **272**, 2118-31.
- Martin, K. C., & Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell*, **136**, 719-30.
- Mao, C., Cheadle, C. E., Hardy, S. J. S., Lilly, A. a, Suo, Y., Sanganna Gari, R. R., Randall, L. L. (2013). Stoichiometry of SecYEG in the active translocase of *Escherichia coli* varies with precursor species. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 11815-20.
- McGinness, K. E., & Sauer, R. T. (2004). Ribosomal protein S1 binds mRNA and tmRNA similarly but plays distinct roles in translation of these molecules. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 13454-9.
- Michaelis, S., Hunt, J., & Beckwith, J. (1986). Effects of signal sequence mutations on the kinetics of alkaline phosphatase export to the periplasm in *Escherichia coli*. *Journal Bacteriology*, **167**, 160-167.
- Milon, P., Carotti, M., Konevega, A. L., Wintermeyer, W., Rodnina, M. V., & Gualerzi, C. O. (2010). The ribosome-bound initiation factor 2 recruits initiator tRNA to the 30S initiation complex. *EMBO Reports*, **11**, 312-6.
- Mori, H., & Ito, K. (2001). The Sec protein-translocation pathway. *Trends in Microbiology*, **9**, 494-500.
- Nakashima, K., Kanamaru, K., Mizuno, T., & Horikoshi, K. (1996). A novel member of the cspA family of genes that is induced by cold shock in *Escherichia coli*. *Journal of Bacteriology*, **178**, 2994-7.
- Nakaminami, K., Karlson, D. T., & Imai, R. (2006). Functional conservation of cold shock domains in bacteria and higher plants. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 10122-7.
- Nesmeyanova, M. a, Karamyshev, A. L., Karamysheva, Z. N., Kalinin, A. E., Ksenzenko, V. N., & Kajava, A. V. (1997). Positively charged lysine at the N-terminus of the signal peptide of the *Escherichia coli* alkaline phosphatase provides the secretion efficiency and is involved in the interaction with anionic phospholipids. *FEBS Letters*, **403**, 203-207.
- Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S., & Amster-Choder, O. (2011). Translation-independent localization of mRNA in *E. coli*. *Science*, **331**, 1081-84.
- Nevo-Dinur, K., Govindarajan, S., & Amster-Choder, O. (2012). Subcellular localization of RNA and proteins in prokaryotes. *Trends in genetics*, **28**, 314-22.

- Nijtmans, L. G. J., Henderson, N. S., & Holt, I. J. (2002). Blue Native electrophoresis to study mitochondrial and other protein complexes. *Methods (San Diego, Calif.)*, **26**, 327-34.
- Ordonez, H., & Shuman, S. (2013). *Mycobacterium smegmatis* Lhr Is a DNA-dependent ATPase and a 3'-to-5' DNA translocase and helicase that prefers to unwind 3'-tailed RNA:DNA hybrids. *The Journal of Biological Chemistry*, **288**, 14125-34.
- Osswald, M., Greuer, B., & Brimacombe, R. (1990). Localization of a series of RNA-protein cross-link sites in the 23S and 5S ribosomal RNA from *Escherichia coli*, induced by treatment of 50S subunits with three different bifunctional reagents. *Nucleic Acids Research*, **18**, 6755-60.
- Palacios, I. M., & Johnston, D. S. (2001). GETTING THE MESSAGE ACROSS: The Intracellular Localization of mRNAs in Higher Eukaryotes. *Annual Review of Cell and Developmental Biology*, **17**, 569-614.
- Palacios, I. M. (2007). How does an mRNA find its way? Intracellular localisation of transcripts. *Seminars in cell & developmental biology*, **18**, 163-70.
- Papanikou, E., Karamanou, S., & Economou, A. (2007). Bacterial protein secretion through the translocase nanomachine. *Nature reviews. Microbiology*, **5**, 839-51.
- Perlman, P., Jarrell, K., Dietrich, R., Peebles, C., Romiti, S., & Benatan, E. (1986). Mitochondrial gene expression in yeast: further studies of a self-splicing group II intron. *Basic Life Science*, **40**, 39-55.
- Peterson, J. H., Szabady, R. L., & Bernstein, H. D. (2006). An unusual signal peptide extension inhibits the binding of bacterial presecretory proteins to the signal recognition particle, trigger factor, and the SecYEG complex. *The Journal of Biological Chemistry*, **281**, 9038-48.
- Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, **8**, 785-6.
- Phadtare, S., & Inouye, M. (1999). Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of *Escherichia coli*. *Molecular microbiology*, **33**, 1004-14.
- Pugsley, a P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiological reviews*, **57**, 50-108.
- Puziss, J. W., Fikes, J., & Bassford, P. (1989). Analysis of mutational alterations in the hydrophilic segment of the maltose-binding protein signal peptide. *Journal of Bacteriology*, **171**, 2303-11.
- Raynal, L. C., & Carpousis, a J. (1999). Poly(A) polymerase I of *Escherichia coli*: characterization of the catalytic domain, an RNA binding site and regions for the interaction with proteins involved in mRNA degradation. *Molecular microbiology*, **32**, 765-75.

- Reuven, N. B., Koonin, E. V., & Rudd, K. E. (1995). The Gene for the Longest Known *Escherichia coli* Protein Is a Member of Helicase Superfamily II. *Journal of Bacteriology*, **177**, 5393-5400.
- Roegiers, F., & Jan, Y. N. (2000). Staufen: a common component of mRNA transport in oocytes and neurons? *Trends in cell biology*, **10**, 220-4.
- Rosch, J. W., Vega, L. A., Beyer, J. M., Lin, A., & Caparon, M. G. (2008). The signal recognition particle pathway is required for virulence in *Streptococcus pyogenes*. *Infection and Immunity*, **76**, 2612-19.
- Saier, M. H. (2006). Protein Secretion Systems in Gram-Negative Bacteria. *Microbe*, **1**, 414-419.
- Samander, H., Passi, K., Saleh, M., & Science, C. (2013). Prediction of Sec-dependent Secreted Proteins Based on mRNA Structure of Signal Peptides. *World Congress in Computer Science, Computer Engineering, and Applied Computing*, 1-7.
- Sarkar, N. (1997). Polyadenylation of mRNA in prokaryotes. *Annual review of biochemistry*, **66**, 173-97.
- Sengupta, J., Agrawal, R. K., & Frank, J. (2001). Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 11991-6.
- Schneewind, O., & Missiakas, D. M. (2012). Protein secretion and surface display in Gram-positive bacteria. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **367**, 1123-39.
- Sevo, M., Buratti, E., & Venturi, V. (2004). Ribosomal Protein S1 Specifically Binds to the 5' Untranslated Region of the *Pseudomonas aeruginosa* Stationary-Phase Sigma Factor *rpoS* mRNA in the Logarithmic Phase of Growth. *Journal of Bacteriology*, **186**, 4903-09.
- Sibbald, M. J. J. B., Ziebandt, a K., Engelmann, S., Hecker, M., de Jong, a., Harmsen, H. J. M., van Dijl, J. M. (2006). Mapping the pathways to staphylococcal pathogenesis by comparative secretomics. *Microbiology and Molecular Biology Reviews*, **70**, 755-88.
- Silhavy, T. J., Benson, S. A., & Emr, S. D. (1983). Mechanisms of Protein Localization. *American Socitey of Microbiology*, **47**, 313-344.
- Sklar, J. G., Wu, T., Kahne, D., & Silhavy, T. J. (2007). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes & Development*, **21**, 2473-84.
- Slobodin, B., & Gerst, J. E. (2010). A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. *RNA*, **16**, 2277-90.

- Sørensen, M. a., Fricke, J., & Pedersen, S. (1998). Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli* in vivo. *Journal of molecular biology*, **280**, 561-9.
- Sorg, J. a., Miller, N. C., & Schneewind, O. (2005). Substrate recognition of type III secretion machines--testing the RNA signal hypothesis. *Cellular Microbiology*, **7**, 1217-25.
- St Johnston, D. (1995). The intracellular localization of messenger RNAs. *Cell*, **81**, 161-70.
- Street, L. M., Harley, M. J., Stern, J. C., Larkin, C., Williams, S. L., Miller, D. L., Schildbach, J. F. (2003). Subdomain organization and catalytic residues of the F factor TraI relaxase domain. *Biochimica et Biophysica Acta - Proteins and Proteomics*, **1646**, 86-99.
- Szostak, E., & Gebauer, F. (2013). Translational control by 3'-UTR-binding proteins. *Briefings in functional genomics*, **12**, 58-65.
- Tseng, T.-T., Tyler, B. M., & Setubal, J. C. (2009). Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC microbiology*, **9**, 1-9.
- Urlaub, H., Kruft, V., Bischof, O., Müller, E. C., & Wittmann-Liebold, B. (1995). Protein-rRNA binding features and their structural and functional implications in ribosomes as determined by cross-linking studies. *The EMBO Journal*, **14**, 4578-88.
- Van der Wolk, J. P., Fekkes, P., Boorsma, a., Huie, J. L., Silhavy, T. J., & Driessen, a J. (1998). PrlA4 prevents the rejection of signal sequence defective preproteins by stabilizing the SecA-SecY interaction during the initiation of translocation. *The EMBO Journal*, **17**, 3631-9.
- Vetting, M. W., S de Carvalho, L. P., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L., & Blanchard, J. S. (2005). Structure and functions of the GNAT superfamily of acetyltransferases. *Archives of biochemistry and biophysics*, **433**, 212-26.
- Velíšek, J. A. N., & Cejpek, K. (2006). Biosynthesis of Food Constituents : Amino Acids : 2. The Alanine-Valine-Leucine, Serine-Cysteine-Glycine, and Aromatic and Heterocyclic Amino Acids Groups-a Review. *Czech Journal of Food. Science*, **24**, 45-58.
- Wilhelm, J. E., & Vale , R. (1993). RNA on the Move: The mRNA Localization Pathway. *Journal of Cell Biology*, **123**, 269-74.
- Windbichler , N., Pelchrzim , F., Mayer ,O., Csaszar,E., & Schroeder, R. (2008)., Isolation of small RNA-binding proteins from *E. coli* : evidence for frequent interaction of RNAs with RNA polymerase. *RNA Biology* , **5**, 30-40.
- Yaguchi, M., Roy, C., & Wittmann, H. G. (1980). The primary structure of protein S10 from the small ribosomal subunit of *Escherichia coli*. *FEBS Letters*, **121**, 113-6.

Yamanaka, K., Fang, L., & Inouye, M. (1998). The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation. *Molecular microbiology*, **27**, 247-55.

Yamanaka, K., & Inouye, M. (2001). Induction of CspA, an *E. coli* major cold-shock protein, upon nutritional upshift at 37°C degrees. *Genes to cells: devoted to molecular & cellular mechanisms*, **6**, 279-90.

Yap, M.-N. F., & Bernstein, H. D. (2013). Mutations in the *Escherichia coli* ribosomal protein L22 selectively suppress the expression of a secreted bacterial virulence factor. *Journal of Bacteriology*, **195**, 2991-9.

Yosef, I., Bochkareva, E. S., & Bibi, E. (2010). *Escherichia coli* SRP, Its Protein Subunit Ffh, and the Ffh M Domain Are Able To Selectively Limit Membrane Protein Expression When Overexpressed. *mBio*, **1**,1-9.

Zhou, Y., & Vachet, R . (2013). Increased Protein Structural Resolution from Diethylpyrocarbonate-based Covalent Labeling and Mass Spectrometric Detection. *Journal of American Society for Mass Spectrometry* , **23**, 708-17.

Zurawski, G., Zurawski, S. M., Watsn, J., & Limited, I. R. L. P. (1985). Structure of the *Escherichia coli* S10 ribosomal protein operon. *Nucliec Acids Research*, **13**, 4521-26.

Appendix

MALDI-MS identification of various bands from pull down assay of surA mRNA.

Sample /•	Proteomics Analysis	Identify Protein	Calculate Mass	Protein sequence coverage	Score	Index
surA-A 45-50*	MASCOT	<i>S.aureus</i> SRP protein	-	-	-	gi 291466947 gi 293510003
		<i>S. aureus</i> FtsY-like protein	-	-	-	gi 341858158 gi 417892191
		Treponema succinifaciens DSM 2489 flagellar export protein FliJ	-	-	-	gi 328947998 YP004365335.1
	Protein Prospector	N/A	-	-	-	-
	PROFOUND	DEAD/DEAH box helicase	-	29	-	gi 300925835
surA-B 30-36*	MASCOT	Response regulatory protein ypdB	28	39	-	gi 222157114
		Hypothetical Protein		41	-	gi 428962431
		ECF RNA polymerase sigma-E factor	-	59	34	-
		Resolvase	24	74	40	-
		Putative protein RhsE	-	55	39	-
	Protein Prospector	Protein HflK	45	14.3	540	166080
		Putative protein RhsE	33	17.9	8471	336398
		Nucleoid-associated protein YbaB	12	50.5	2453	521358
	PROFOUND	Efa1/LifA protein	-	28	-	gi 261226294
	surA-C 15-20*	MASCOT	Hypothetical Protein	52	81	49
Excl1 protein			15	58	52	gi 168998552
Fumarate and nitrate reduction regulatory protein			28	44	24	-
Urease subunit beta			11	71	34	-

Sample / •	Proteomics Analysis	Identify Protein	Calculate Mass	Protein sequence coverage	Score	Index
		HTH-type transcriptional regulator MatA	23	64	30	-
	Protein Prospector	Transposon gamma-delta resolvase	20	24	848	460624
		Chaperone protein FocC	25	27.5	179	130712
	PROFOUND	N/	-	-	-	-
surA-I *25	MASCOT	DNA-binding transcriptional or f kdgK, kdgT, eda	29	23	-	gi 16129781
		ABC transporter ATP-binding protein	25	26	-	gi 26248011
	Protein Prospector	N/A	-	-	-	-
	PROFOUND	N/A	-	-	-	-
surA-J *40-45	MASCOT	Multifunctional conjugation protein TraI	191	41	43	-
	Protein Prospector	HTH-type transcriptional regulator ZntR	16	27	494	540073

- Apparent Mass from SDS gel.
- Score: the degree of difference between the structure of the unknown protein and a protein in the database.
- Protein sequence coverage: the percent of sequence from the unknown protein (peptides) that is matched to the database sequence.

MALDI-MS identification of various bands from pull down assay of isodH mRNA.

Sample /•	Proteomics Analysis	Identify Protein	Calculate Mass	Protein sequence coverage	Score	Index
isodH-D 40-45*	MASCOT	Hypothetical Protein	-	30	-	gi 423037776
		3-isopropylmalate dehydrogenase	39	-	-	-
		Fused protein chain initiation factor 2(IF-2)	97	-	-	-
		Protein K1aB	42	64	39	-
		ATP-dependent RNA helicase RhlE	50	55	31	-
	Protein Prospector	50S ribosomal protein L33	63	56.4	442	361087
	PROFOUND	ATP-dependent RNA helicase HrpA	100	27	-	gi 419936745
isodH-E 25-30*	MASCOT	Resolvase, N-terminal domain protein	19	34	-	gi 417246748
		Transcriptional regulatory protein YehT	27	57	32	-
		Fumarate and nitrate reduction regulatory protein	28	44	41	-
	Protein Prospector	N/A	-	-	-	-
	PROFOUND	N/A	-	-	-	-
isodH-F 15-20*	MASCOT	Resolvase domain-containing protein	19	44	-	gi 366161248
	Protein Prospector	N/A	-	-	-	-
	PROFOUND	N/A	-	-	-	-
isodH(G)/ 20-25*	MASCOT	Endopeptidase	16	34	-	gi 415778122
		Bacteriophage lysis protein	17	45	-	gi 194430045
	Protein Prospector	N/A	-	-	-	-
	PROFOUND	N/A	-	-	-	-
isodH(H)/ 15-20*	MASCOT	DEAD/DEAH box helicase	16	13	-	gi 301026837
		ATP-dependent helicase Lhr	16	14	-	gi 15802067
		Acetyltransferase, GNAT family	18	-	-	-
isodH(K)/ 40-45*		Multifunctional conjugation protein TraI	191	41	48	

- Apparent Mass from SDS PAGE gel.
- Score: the degree of difference between the structure of the unknown protein and a protein in the database.
- Protein sequence coverage: the percent of sequence from the unknown protein (peptides) that is matched to the database sequence.

MALDI-MS identification of various bands from pull down assay of *phoA* mRNA.

Sample /•	Proteomics Analysis	Identify Protein	Calculate Mass	Protein sequence coverage	Score	Index
phoA-L 45-50*	MASCOT	Poly (A) polymerase I		50	28	-
	Protein Prospector	N/A	-	-	-	-
	PROFOUND	N/A	-	-	-	-
phoA-M 15-20*	MASCOT	30S ribosomal protein S10	11	46	68	gi 487397367
		Fumarate and nitrate reduction regulatory protein	28	44	43	-
		Ribosomal protein S10		34	69	gi 487397367
		Nickel import ATP-binding protein Nike	29	30	48	-
	Protein Prospector	30S ribosomal protein S10		44.7	2003	387038
		50S ribosomal protein L28	9007	59	1090	356052
	PROFOUND	N/A	-	-	-	-
phoA -N 15-20*	MASCOT	YHBP_ECO7I	16	74	31	-
	Protein Prospector	Putative type II secretion system L-type protein YghE	32058	19.2	159	527985
	PROFOUND	N/A	-	-	-	-

- Apparent Mass from SDS PAGE gel.
- Score: the degree of difference between the structure of the unknown protein and a protein in the database.
- Protein sequence coverage: the percent of sequence from the unknown protein (peptides) that is matched to the database sequence.