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

Walaa A Hakeem

*Department of Biology, Laurentian University, Canada*

*Corresponding Author*: Mazen Saleh, Associate Professor and Chair, Laurentian University, Sudbury, Ontario, Canada.

**Received**: August 19, 2019; **Published**: November 11, 2019

**Abstract**

In this study we assay that the mRNA encoding secreted proteins; periplasmic protein propyl isomerase chaperone SurA and the phosphatase phoA would preferentially interact with several proteins rather than Ffh secretion factor. As well as the secreted proteins, the mRNAs encoding the cytosolic proteins cytoplasmic protein 3-isopropyl malate dehydrogenase (IsodH) and gMP reductase have the same interactions. We performed the pull-down assays using SurA mRNA as bait that showed that mostly interested proteins corresponding to the *Staphylococcus aureus* SRP protein and *S. aureus* PtsY-like Protein. Therefore, the Nucleoid-associated protein, YbaB and Multifunctional conjugation protein TraI were recognized. Curiously, we were able to identified the following proteins; the fused protein chain initiation factor 2 (IF-2), DEAD/DEAH box helicase protein, multifunctional conjugation protein TraI, Acetyltransferase GNAT family protein I and 3-isopropylmalate dehydratase enzyme that interact with the cytosolic IsodH mRNA. Particularly, it was found that the least enzyme involved in functions associated with mRNA translation in the case of in *E. coli* and thus involved in the regulation of mRNA products. As well as 3-isopropylmalate dehydratase enzyme the ATP-dependent helicase Lhr also has a role in mRNA control. Beside the previous finding, in the case of the PhoA mRNA many proteins such as Poly (A) polymerase I enzyme (PAPI), 30S ribosomal protein S10, 50S ribosomal protein L28, Putative type II secretion system L-type protein YghE and Fumarate and nitrate reduction (FNR) regulatory protein were found. The key factor, Ffh, however was not amongst them.

Ideally, direct binding experiments with a purified Ffh and cloning of the Ffh gene would be conducted for the further direction.

**Keywords**: Signal Recognition Particle (SRP); Secretion of the Protein; Translation; *E. coli*; Fifty Four Homologue (Ffh); SurA; IsodH; RNA; mRNA; tRNA

**Abbreviations**

μJ: Microjoule; μl: Microliter; μg: Microgram; μM: 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; *E. coli*: *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: Secretary; SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis; SRP: Signal Recognition Particle

**Citation**: Walaa A Hakeem. "Identifying *Escherichia coli* Factors that Selectively Bind the mRNA of Secreted Proteins". *EC Microbiology* 15.12 (2019): 01-24.
Identifying *Escherichia coli* Factors that Selectively Bind the mRNA of Secreted Proteins

Introduction

In gram-negative bacteria, up to 40% of cellular proteins are exported to the periplasm 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* [1,2].

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 site [3-5]. 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 [3].

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). This particle is a complex consists of a 4.5S RNA and a 54-kDa Ffh protein (a homologue of eukaryotic SRP) [1]. 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 [1]. 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.

There have been several studies that show that changing the structure of the signal peptide can eliminate or reduce the efficiency of protein secretion. However, other alterations in the signal peptide sequence can enhance protein sorting or have no effect [6].

Additionally, there are many mRNA studies proposed 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 [7,8]. The previous mRNA hypothesis was established because mRNA targeting and subcellular localization have been shown to exist in eukaryotes and prokaryotes [8]. It has been further reported that *fliC* export by the flagellar system in *E. coli* requires 5’ untranslated RNA (UTR) sequences signal [7].

In this study, the periplasmic protein propyl isomerase chaperone SurA, which is involved in the biogenesis of outer membrane proteins was utilized to identify the novel RBP that involved in the secretory pathway in *E. coli*. As well as surA, 3-isopropylmalate dehydratase which involved in the biosynthesis of the amino acid leucine was used as a control for the non-secreted protein [9]. Additionally, we also examined the 5’UTR [10] of alkaline phosphatase phoA and gMP reductase mRNAs that might achieve efficient isolation of such cytoplasmic factors. Biotinylated transcripts of SurA, phoA, IsoH and gMP were produced *in vitro* system. Thus used these as bait for pull down assays with streptavidin-magnetic beads in order to identify a cytoplasmic factor that associate with these mRNAs by SDS-PAGE and MALDI-MS identification.

Our finding found 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.

Materials and Methods

Bacterial cultures

*Escherichia coli* strains B; ATCC 11303 and 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%

Citation: Walaa A Hakeem. "Identifying *Escherichia coli* Factors that Selectively Bind the mRNA of Secreted Proteins". *EC Microbiology* 15.12 (2019): 01-24.
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.

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 (0.1M 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 [0.1M Tris-Cl, pH 8.0, 20 mM EDTA, 0.5M sucrose, 1% (v/v) TritonX-100, 24 mg of lysozyme (30 µl of lysozyme)]. Following an incubation step on ice for 1 hour 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 minutes 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.

Quantitation of nucleic acids

Purified DNA, resuspended in 50 - 250 ml 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.

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 seconds 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 -5 8°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-EZD-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.

<table>
<thead>
<tr>
<th>Processes in Translation</th>
<th>Factors</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>IF1</td>
<td>Stabilizes 30S subunit.</td>
</tr>
<tr>
<td></td>
<td>IF2</td>
<td>Binds fmet-tRNA to 30S–mRNA complex: bind to GTP and stimulates hydrolysis.</td>
</tr>
<tr>
<td></td>
<td>IF3</td>
<td>Binds 30S subunit to mRNA</td>
</tr>
<tr>
<td>Elongation</td>
<td>EF-Tu</td>
<td>Binds GTP; brings aminoacyl-tRNA to the A site of ribosome</td>
</tr>
<tr>
<td></td>
<td>EF-Ts</td>
<td>Generates active EF-Tu</td>
</tr>
<tr>
<td></td>
<td>EF-G</td>
<td>Stimulates translation; GTP-dependent</td>
</tr>
<tr>
<td>Termination</td>
<td>RF1</td>
<td>Catalyzes release of the polypeptide chain from TRNA and disassociation of the translocation complex; specific for UAA and UAG termination codons</td>
</tr>
<tr>
<td></td>
<td>RF2</td>
<td>Behaves like RF1; specific for UGA and UAA codons</td>
</tr>
<tr>
<td></td>
<td>RF3</td>
<td>Stimulates RF1and RF2</td>
</tr>
</tbody>
</table>

*Table 1: A summary of the factors involved in protein synthesis in prokaryotes [24].*
Identifying *Escherichia coli* Factors that Selectively Bind the mRNA of Secreted Proteins

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene Fragment Amplified</th>
<th>Primers</th>
<th>Temperature</th>
<th>Fragment Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sura A</td>
<td><em>sura A</em></td>
<td>Forward: 5’TAATACGACTCACTATAGGGATGAAGAACTG-GAAAACGCTG-3’&lt;br&gt;Reverse: 3’CAAGGTCGGTTTCGACAC-5’</td>
<td>55 - 58°C</td>
<td>127</td>
</tr>
<tr>
<td>isoDH</td>
<td><em>isodH</em></td>
<td>Forward: 5’TAATACGACTCACTATAGGGATGCTAAAGCTTATCGAAAAA-3’&lt;br&gt;Reverse: 3’TACCGAAAGTTTTGCCGC-5’</td>
<td>51 - 52°C</td>
<td>178</td>
</tr>
<tr>
<td>pohA</td>
<td><em>pohA</em></td>
<td>Forward: 5’TAATACGACTCACTATAGGGAAAGATCTCTTTCAAGAGCTC-3’&lt;br&gt;Reverse: 3’GCCCGTTTTGCCAGACAG-5’</td>
<td>59 - 60°C</td>
<td>200</td>
</tr>
<tr>
<td>gMP</td>
<td>GMP</td>
<td>Forward: 5’TAATACGACTCACTATAGGGATGCTTGCCGCATGATTAA-3’&lt;br&gt;Reverse: 3’GTAATTGACGTTCCAGTCA-5’</td>
<td>56°C</td>
<td>194</td>
</tr>
</tbody>
</table>

**Table 2: Primers for the PCR reactions.**

**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).

**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 1.5h 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.

**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, 40U 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: isomyl 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 minutes 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 20 µl dH2O.

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 minutes 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 perform 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).

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 0.2 mg/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.

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.

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 254 nM/120,000 µJ/cm^2 for 5 minutes (Microprocessor-Controlled UV Cross linkers). Therefore, the mRNAs and the bound interactome can be efficiently captured using pull down assays.

Pull-down assay

An aliquot of magnetic beads (0.8 mg beads/1 mg 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 2h 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 (3 ml 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].

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 2h (0.8 mg beads/1 mg 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.

**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 [3M 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 5 minutes. The membrane was blocked with 5% milk for 1h at RT followed by an incubation with 1 mg cell extract for an additional 1h at RT in the presence of 200 μl binding buffer (3 ml 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.

**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 separates 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, 4 ml 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). 1 mg 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, 0.1M 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 (144g glycine, 1/10 dilution of 30 g Tris-HCl in 1L 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 minutes 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. 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.

**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, 4 ml 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% 27 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, 0.1M DTT) to final concentration of 1X. The entire contents of the pull down assay were boiled for 10 minutes in water bath and then loaded into the wells. The gels were performed at a constant voltage of 200 V for 45 minutes in Mini -PROTEAN Tetra System (BIO-RAD) with 1x SDS running buffer (144 g glycine, 1/10 dilution of 30g 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.
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 minutes 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 minutes further 10 μl 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 μL of 25 mM ammonium bicarbonate was add to the previous collected liquid for 10 min followed by addition of 20 μL of 5% formic acid for an extra 10 min. This wash was repeated with 20 μL 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.

Results

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 1). 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 2).

Figure 1: 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).
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 3 lane 3) and isodH mRNA (Figure 3 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 4 (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 2: 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).

**Figure 3: 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).
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 5). 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).

Figure 5: 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/μl), isodH mRNA (71 ng/μl) while surA mRNA (9 ng/μl) and isodH mRNA (3 ng/μl) 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.

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 considered (Figure 6A-6D). This fact explains that the proteins that bind to mRNA are not only based on the actual sequence but can also bind to structural elements. I followed the process posted by the RNA fold Webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) in order to predict secondary structures of mRNA sequences.

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.

<table>
<thead>
<tr>
<th>Sample ID (mRNA)</th>
<th>Label Efficacy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>surA</td>
<td>41</td>
</tr>
<tr>
<td>isodH</td>
<td>36</td>
</tr>
<tr>
<td>phoA</td>
<td>43</td>
</tr>
<tr>
<td>gMP</td>
<td>38.5</td>
</tr>
</tbody>
</table>

Table 3: mRNA labeling efficiency.

To determine the labeling efficiency the following equation was used: labeled RNA/total RNA X 100.
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 7).

**Figure 7:** Detection of biotinylated mRNA on PVDF membrane.
The biotinylation 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.

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.

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

<table>
<thead>
<tr>
<th>mg of Beads/Reaction</th>
<th>%Free Biotin</th>
<th>% Biotin Bound to Beads</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>0.2</td>
<td>93</td>
<td>7</td>
</tr>
<tr>
<td>0.4</td>
<td>54</td>
<td>46</td>
</tr>
<tr>
<td>0.6</td>
<td>28</td>
<td>72</td>
</tr>
<tr>
<td>0.8</td>
<td>18</td>
<td>82</td>
</tr>
</tbody>
</table>

*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.*

**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 8 and 9 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 10).

*Figure 8: 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 9:** 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 10:** 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).
Identifying *Escherichia coli* Factors that Selectively Bind the mRNA of Secreted Proteins

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 11A). 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 11B.

**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 12). 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 12 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).

**Figure 11**: 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).

**Figure 12**: 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).
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.

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 13A) or contain the transcript of phoA (Figure 13B). The gels were stained with Coomassie brilliant blue and a diagonal migration of proteins was generated for both gels. An aggregation in figure 13B (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.

![Figure 13: RNA-protein interaction assay.](image)

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 refers to RNA-protein complex. Electrophoreses of *E. coli* extract in gels contain either control (A) or contain the transcript of phoA (Figure 13B). M = Marker (kDa).

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 8 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*.
Identifying *Escherichia coli* Factors that Selectively Bind the mRNA of Secreted Proteins

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

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

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).

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

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

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 11A and 11B 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 13B was identified as Putative type II secretion system L-type protein YghE (See table 7 below).

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

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

**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. 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 [7]. 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 [7]. 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.

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 [11,12]. 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 (Table 5-7).

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 [13]. 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 [9].

**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. 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 6).

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

secondary structures. These questions remain to be answered and would shed more light onto the specific features of the mRNA needed for targeting.

**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 8 and 9). 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 10). 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. 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. 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 (Table 1). 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). 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.

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 [14]. 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. 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 [15]. 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 [16].

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. Further study proved that this polypeptide also contains an amino region similar to the DEAD family helicases. 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 shown to have a major role in binding and cleaving single-stranded DNA oligonucleotides containing an oriT sequence in plasmid transfer experiments [17]. TraI also functions in the initiation and/or termination of plasmid transfer [17].

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) [10]. Direct evidence supporting the importance of the UTR comes from the recent observation that the ribosomal protein S1 [10,18] which is essential for in vivo translation in E. coli can bind to an 11 nucleotides region in the mRNA immediately upstream of the Shine-Dalgarno (SD) sequence [10]. 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 [10]. The SD sequence-S1 protein interaction further plays a significant role in effecting gene expression [10].

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. 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.

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 11A). 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. 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 it is able to bind first to the RNA and then initiate the polyadenylation process. 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. A major function of S10 is its ability to mediate binding of tRNA to the ribosomes. Studies have also reported that the S10 protein has a role in anti-termination of the transcription process [19,20]. It has been shown that the 50S ribosomal protein L28 cross-links to two distinct RNA regions in E. coli. Moreover, this protein has a critical role in ribosome assembly and the absence of L28 also prevented the synthesis of the ribosome. 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 protein.
virulence factor antigen 43 (AG 43) in *E. coli*. 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. 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 13B, 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 [21]. 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.

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.

**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 Wight (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.

**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 9-11.

**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 [22]. 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 12.

Aurintricarboxylic acid has no effect on pull down assays of both phoA and gMP mRNA's as shown in figure 12 (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 [23].

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. 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 12).

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 [25-39].

Conclusion

In our study a direct interaction between mRNAs encoding secreted proteins and the E. coli SRP equivalent (Ffh) was pursued using pull down assays. 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.

Acknowledgements

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.

Conflict of Interest

There is no any conflict of interest exists.

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

Bibliography

10. Berg L. "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 2.3 (2009): 379-389.


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


Volume 8 Issue 12 December 2019
© All rights reserved by Walaa A Hakeem.