2021

Identifying Pathways Affected by the HrpA RNA Helicase

Brandon Simons
University of Central Florida

IDENTIFYING PATHWAYS AFFECTED BY THE HRPA RNA HELICASE

by

BRANDON TAYLOR SIMONS

A thesis submitted in partial fulfillment of the requirements
for the Honors in the Major Program in Biomedical Sciences
in the College of Medicine
and in the Burnett Honors College
at the University of Central Florida
Orlando, FL

Spring Term, 2021

Thesis Chair: Sean Moore, Ph.D.

ABSTRACT

HrpA is a poorly characterized DEAH-box helicase found in many enterobacteria, including *Escherichia coli.* For unknown reasons, HrpA causes sickness in cells lacking Elongation Factor P, a transpeptidation enhancer that prevents ribosome stalling at proline-rich motifs during protein synthesis. HrpA's C-terminal RNA binding motifs and association with the degradosome suggest this protein interferes with an important translational process. Deletion of *hrpA* does not produce an overt phenotype, obscuring its role in cellular physiology. This thesis sought to identify pathways affected by this highly conserved protein. A bioinformatic analysis of HrpA's predicted interactome was conducted using a database called STRING. In an effort to verify HrpA's predicted enzymatic partners, a synthetic lethal screen was performed in order to identify *hrpA*-dependent mutants. Although the screen failed to identify mutant colonies dependent on HrpA for optimal growth, this thesis will serve as a foundation for further investigation of HrpA's role in translation.

DEDICATION

To my beloved grandmother, Gloria, who lives as a Godsent saint to her family and all those touched by her unconditional benevolence. This thesis is in honor of you and your gracious presence in this world.

ACKNOWLEDGMENTS

First and foremost, I want to acknowledge my thesis chair, Dr. Moore, for granting me the resources necessary to complete this thesis. Thank you for your patience while guiding me through my first molecular biology research experience, your adaptability to the challenges a pandemic brought to the thesis, and your mentorship over the past three years. Additional thanks to Dr. Teter and Dr. von Kalm for your assistance as members of the thesis committee. I am grateful to the members of the Moore Lab for their willingness to answer all my questions. Lastly, I want to acknowledge my grandparents, Keith and Cindy Simons, for supplying me with wisdom that will last a lifetime and my parents, Brian and Amal Simons, for always being my number one supporters.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Translation is an essential mechanism underlying the growth and homeostasis of all organisms. A translation cycle uses the coding strand of mRNA to assemble a protein's amino acid sequence through four major steps: initiation, elongation, termination, and ribosome recycling [1]. Within prokaryotes, translation initiation and transcription are coupled through the physical interaction of RNA polymerase with the ribosome [2]. As soon as mRNA emerges from RNA polymerase, the 16S rRNA component of the 30S ribosome subunit interacts with a complementary purine-rich ribosome binding site (Shine-Dalgarno sequence) upstream of the initiation site (usually an AUG codon). Initiation factors IF1, IF2, and IF3 recruit the initiator fMet-tRNA^{fMet} to the P site of the 30S subunit. The ribosome has three active sites: the A site is the entry point of aminoacyl-tRNA, the P site is the location of peptidyl-tRNA formation, and the E site is the exit point of uncharged tRNA. Joining of the 50S subunit promotes the dissociation of initiation factors and the formation of a mature 70S initiation complex with fMet-tRNA^{fMet} housed in the P-site [3]. Translation initiation varies within prokaryotes and remains to be completely characterized.

Elongation starts when the second codon of the open reading frame enters the A site, which undergoes a conformational change due to binding of fMet-tRNA^{fMet} at the P site, allowing an incoming aminoacyl-tRNA anticodon to interact with a mRNA codon. This interaction triggers GTP hydrolysis by elongation factor Tu (EF-Tu), formalizing the binding of aminoacyl-tRNA to the A site [4]. The amino acid in the P site is separated from tRNA and forms a peptide bond between its carboxyl group and the amino group of the aminoacyl-tRNA in the A site through a process catalyzed by the 23S rRNA within the 50S ribosomal subunit. The newly deacylated tRNA in the P site and the dipeptidyl-tRNA in the A site translocate to the E

and P sites, respectively, as elongation factor G (EF-G) catalyzes the movement of the ribosome along mRNA in the 5' to 3' direction [5]. Elongation continues until a stop codon is reached. When a stop codon enters the A site, it is recognized by release factors (RF-1 and RF-2) that promote the hydrolysis of the ester bond between the polypeptide chain and tRNA. Consequently, the polypeptide chain is released from the ribosome. Termination ends as RF-3 initiates the discharge of RF-1 and RF-2. Ribosome recycling factor (RRF) and EF-G promote subunit splitting, releasing mRNA and associated tRNAs from the ribosome. Dissociation of the tRNAs from mRNA is fostered by IF3 [6]. The free 50S and 30S subunits may be reused for future translational cycles.

Many translation factors and pathways are yet to be fully understood. In recent years, numerous studies have characterized the roles of various RNA helicases within the cell. RNA helicases are ubiquitous enzymes found within all domains of life, including some virus families, and share significant structural homology to DNA helicases [7, 8]. The primary function of RNA helicases is RNA duplex unwinding, defined by several highly conserved sequence motifs. Over 60 proteins have the capacity to perform the fundamental RNA helicase function *in vitro* and are engaged in other aspects of RNA metabolism [9]. Many RNA helicases are multifunctional and serve a role in transcription regulation as coactivators or corepressors, pre-mRNA processing, translation initiation, and posttranslational modification [10-12]. RNA helicases coordinate the binding and hydrolysis of nucleoside triphosphates (NTPs) through conserved phosphate binding loops, also known as the Walker A motif [13].

RNA helicases are divided into two categories based on their ability to form oligomeric rings or not. The two categories are broken down into six super families (SFs) according to comparative analysis of structural and sequence homology. Ringed helicases comprise SFs 3

through 6, while non-ringed helicases comprise SFs 1 and 2. Eukaryotic RNA helicases are only found in SFs 1 and 2 [14]. SFs are differed into subfamilies. *HrpA* is a gene that was first identified to share sequence homology with three yeast proteins (Prp2, Prp16, and Prp22) found in the DEAH-box subfamily of SF2 of RNA helicases [15]. The three proteins act as chaperones during pre-mRNA splicing, specifically during the splicing reaction and degradation of the spliceosome [16].

The Moore Lab first became interested with HrpA in a prior study characterizing the function of large ribosomal subunit protein bL9 through inactivation of elongation factor P (EF-P) [17]. EF-P (eIF5A in eukaryotes) is a highly conserved transpeptidation enhancer that was first identified to correctly position the initiator tRNA during translation [18]. EF-P shares a similar structure to tRNA and specifically binds to ribosomes between the P and E sites. More recent studies indicate EF-P prevents stalling at proline-rich motifs by promoting peptide bond synthesis on 70S ribosomes [19]. Proline's structure is comprised of a pyrrolidine ring that contains the nitrogen of an imino group and an α-carbon. The presence of an imino group, instead of an amino group, underlies proline's rigid structure. Consequently, proline is a poor A-site acceptor and P-site donor in peptide bond formation [20, 21]. Ribosomal stalling occurs when ribosomes accumulate at particular positions on mRNAs, slowing down translation. Collisions between actively translating ribosomes induce regulatory pathways to abandon the incomplete polypeptide, significantly impairing cell growth [22].

Loss of EF-P in *Escherichia coli* results in sick cells caused by various pleiotropic effects, but also produces faster-growing suppressor mutants. A suppressor mutation is a second mutation that partially overcomes the effects of the original mutation in order to restore normal growth. Therefore, it is reasonable to assume HrpA's suppressor mutations may help overcome

translational deficiencies. Several independently evolved *Δefp* fast-growing suppressor mutants were isolated and found to contain large deletions (about 200 kb) comprising a domain of unknown function within *hrpA.* The growth of *Δefp ΔhrpA* mutants was improved relative to *Δefp* but was still slower than the *efp*+ parent strain. The growth of *ΔhrpA* showed no difference than the *efp*+ parent strain. Transduction of an *ΔhrpA* allele into *Δefp* cells improved the growth of sick *Δefp* cells. The results initially hinted poly-proline translation may be improved in the absence of HrpA or HrpA impairs an important pathway in *Δefp* cells [23].

Polysome profiling aligns with the observed growth rates of *ΔhrpA*, *Δefp,* and *Δefp ΔhrpA* cells. Deletion of only *hrpA* does not produce an obscured polysome profile. Deletion of only *efp* results in impaired ribosomal subunit maturation, monosome formation, and polysome formation compared to *wild-type E. coli.* Deletion of *hrpA* and *efp* improves subunit maturation, monosome formation, and polysome formation to near *wild-type* levels. Further investigation using a dual luciferase reporter system revealed *Δefp ΔhrpA* cells still experienced obstructed transpeptidation of poly-proline motifs comparable to *Δefp* cells. The dual luciferase reporter findings indicate deletion of *hrpA* does not improve poly-proline translation, pointing the Moore lab in the direction that HrpA interferes with a translational process in the absence of EF-P [23].

Within the DEAH subfamily of SF2 of RNA helicases, two highly conserved helicase domains connected with a short linker region encompass each protein's characteristic helicase function. Folding of the two core domains forms a pocket biding site for NTPs and RNA. The variable N- and C- extensions are often larger than the helicase core and contain domains specific to the function of each protein [24]. Specifically, the C-terminal domain of HrpA includes an RNA binding motif weakly resembling ribosomal protein S1. Nonetheless, the

function of HrpA's C-terminal domain is not established and is classified as domain of unknown function 3418 (DUF3418) [15]. A mutant version of HrpA composed of HrpA's C-terminal domain linked to a fluorescent protein named mClover 3 was found to co-migrate with polysomes [23]. Previous studies conclude HrpA may associate with the degradosome [25]. Purified HrpA acts as a universal NTPase, hydrolyzing ATP, GTP, CTP, TTP, and UTP. Also, the level of sickness in *Δefp* cells was found to be positively correlated to the quantity of HrpA present [23]. Consequently, the findings suggest HrpA's DUF3418 may guide HrpA into a pathway associated with ribosomal stalling during translation.

RNA helicases are involved in various human diseases and infections including HIV, Chagas disease, Alzheimer disease, Lyme disease, etc [26]. Deletion of a HrpA ortholog found in *Borrelia burgdorferi* results in an inability of Lyme spirochetes to infect mice. The C-terminal domain of the HrpA in *B. burgdorferi* differs considerably from *E. coli* HrpA, suggesting each protein targets a different pathway [27]. Nevertheless, HrpA's high conservation across various domains of life signifies it has a noteworthy role that is yet to be understood.

The research conducted in this thesis serves as a starting point for future characterization of DUF3418 by identifying HrpA's possible molecular interactions through a bioinformatics and wet lab approach. The specific aims of this thesis were as follows: 1) identify HrpA-dependent mutants using a synthetic lethal screen and P1 transduction and 2) identify and analyze HrpA's genetic partners using protein-protein interaction networks.

MATERIALS AND METHODS

Synthetic Lethal Screen

An *N-ethyl-N-nitrosourea* (ENU) mutant library originating from strain KR20 (BW30270, an *fnr*$^+$, *rph*$^+$ derivative of MG1655), comprised of genotype *ΔhrpA lacIZYA::cat* and *pRC-hrpA-T5*, was used for this study. The plasmid includes *lacZ* under control of the *lac* promoter and lacks a partioning locus, allowing the plasmid to be easily lost. Freezer stocked 1 ml aliquots of the mutant library were thawed, resuspended, and diluted 1:1 with LB broth. While held over ice, 50 µl of the diluted library was spread onto an LB plate supplemented with 0.2% glycerol, 0.5 mM IPTG, and 80 µg/mL X-gal. The plates were incubated at 30°C. After 48 to 72 hours of incubation, the plates were examined for potential hits. Isolated, solid dark-blue colonies were restreaked for closer observation. Sometimes the restreaked potential hits displayed solid-white colonies alongside solid-blue colonies. In this case, the solid-white colonies must be smaller than the solid-blue colonies in order to be considered a potential hit. Larger solid-blue colonies proved the solid-white colonies were "sick" or had a growth defect due to lack of HrpA. A mutant comprised of solid-white and solid-blue colonies of similar size was not considered a potential hit. The guidelines were strictly followed to eliminate false-positives [28]. Isolated colonies which fulfilled the guidelines were grown in overnight liquid cultures at 30°C and stored as potential hits at -80°C in 3:1 LB-glycerol.

Diagnostic Polymerase Chain Reaction

In order to confirm the presence of the plasmid and chromosomal knockout in each potential hit, two primer pairs were designed for polymerase chain reaction (PCR). For confirming *hrpA* in plasmid, the first primer pair outlines a 466 base pair DNA segment within

pRC-hrpA-T5. The amplified DNA encompassed the ending of *lacI*, a gene upstream of *hrpA*, and the first 93 base pairs of *hrpA*. For diagnosis of the *hrpA* chromosomal knockout, a second primer pair consisted of a forward primer upstream of *hrpA* and a reverse primer downstream of *hrpA*. If *hrpA* chromosomal knockout was present, PCR produced a 760 DNA base pair product. If a wild-type *hrpA* locus was present, PCR produced a 4,365 base pair product. All primers were designed with a melting temperature (Tm) of 60°C and standard PCR parameters were followed (refer to Appendix A: Diagnostic PCR Primers and Parameters). The DNA of interest was amplified using Taq polymerase (New England Biolabs) and product length was examined using 1% agarose gel electrophoresis.

P1 Transduction

RW220 (genotype: *rna- (scar) hrpA-T5-tet*) was selected as the donor strain and potential hits 1 - 4 (genotype*: rna- ΔhrpA lacIZYA::cat, pRC-hrpA-T5*) were selected as the recipient strains. RW220 was streaked onto a selective LB plate (refer to Appendix B: Reagents for P1 Transduction). Each recipient was streaked on a selective and nonselective LB plate to ensure the recipient forms colonies under nonselective conditions and does not form colonies under selective conditions. An isolated RW220 colony was grown in a 1 mL overnight culture supplemented with 10 μg/mL tetracycline. The culture was used to inoculate a new 1 mL culture lacking tetracycline at a 1:100 dilution. The new culture was placed into a 37ºC shacking incubator. When early log phase was reached, indicated by light turbidity, 40 μL of phage stock was added to the culture. The infection continued for 1 to 3 hours until lysis was evident. The culture was compared to two controls for evidence of lysis: a mock culture (no donor cells and 40 μL of phage) and a mock infection (donor cells with no phage). The culture was saturated with 75 μL of chloroform, transferred to a microfuge tube, and centrifuged for 5 mins at

maximum g. The supernatant was transferred to new microfuge tube along with 50 μL of chloroform and stored at 4ºC as the donor lysate. Each of the four recipient strains were grown in 1 mL overnight LB cultures at 30ºC, centrifuged for 3 minutes at 3,000 g, and the cells were resuspended in 300 μL of P1-LB (refer to Appendix B: Reagents for P1 Transduction). For each of the four transductions, 100 μL of donor lysate stock was transferred to a microfuge tube and placed in a 37ºC shacking incubator with its cap left open for 15-30 minutes in order to evaporate leftover chloroform. 100 μL of recipient cells were added to the 100 μL phage aliquot and quickly mixed with the pipettor. The tubes were transferred to a 37ºC shacking incubator and the infection commenced. After 30 minutes, the tubes were removed from the incubator. 200 μL of 1 M sodium citrate (refer to Appendix B: Reagents for P1 Transduction) and 1 ml of LB broth was added to each tube. After another hour of incubation at 37ºC, the tubes were centrifuged for 5 minutes at 5,000 g. The supernatant was removed and 100 μL of LB-citrate (refer to Appendix B: Reagents for P1 Transduction) was added to each tube. The entire contents of the tubes were spread onto a selective LB plate with 1 mM sodium citrate and incubated at 30ºC overnight [29]. The plates were examined at 24, 48, and 72 hours for the absence of dark blue colonies.

Protein-Protein Interaction Networks

In this study, the STRING online database was used to explore HrpA's alleged protein-protein interactions. HrpA from *E. coli* K12 MG1655 was selected as the model organism and evaluated under the following categories: textmining, experiments, genetic neighborhood, gene fusion, co-occurrence, and co-expression [30]. For every protein-protein interaction on STRING, a confidence score is displayed based upon the supporting evidence. On a scale between zero and one, the confidence score indicates the likelihood a protein-protein interaction is specific, reproducible, and biologically significant [31].

Based upon previous studies of HrpA, three standards were established in order to filter through HrpA's large number of possible enzymatic partners identified in STRING. If the putative protein interaction did not fulfill at least one of the criteria, then it was disregarded. The criteria were as follows: 1) proteins involved in transcription or translation 2) proteins interacting with mRNA, and 3) proteins interacting with ribosomes or involved in ribosome recycling pathways. The criteria were developed based upon the current evidence of HrpA's predicted molecular function.

RESULTS

Synthetic Lethal Screen

The synthetic lethal screen is a method of isolating mutant organisms that are dependent on a non-essential gene of interest. The role of a non-essential gene is uncovered by creating mutations in related pathways, forcing the cell to rely on the non-essential gene for its survival or improved fitness. Two components are necessary for the success of the screen: chromosomal deletion of the gene of interest and an unstable single-copy plasmid containing the gene of interest and phenotypic marker under an inducible promoter. Cells with mutations that necessitate the gene of interest will retain the plasmid and exhibit a phenotypic change, indicating a potential hit. Ideally the mutations of all the potential hits will map to the same region of the chromosome, identifying a mutated gene in a pathway affected by the non-essential gene. Further experimentation may be required to create a model of the gene's exact function. However, the synthetic lethal screen eliminates experimental bias through its reliance on objective results.

In this study, *hrpA* was the non-essential gene of interest. The phenotypic marker, *lacZ*, encodes a β-galactosidase enzyme which hydrolyzes X-gal into an insoluble blue pigment. The plasmid (*pRC-hrpA-T5*) harboring *hrpA* and *lacZ* lacks a partioning locus, allowing the unstable plasmid to be efficiently lost when a cell experiences a lack of selective pressure to retain *hrpA* [32]. The following three phenotypes were observed when the synthetic lethal mutant library was plated with IPTG and X-gal: solid-blue, solid-white, and sectored-blue with an asterisk-like appearance (Figure 1). A solid-blue colony retained the plasmid harboring *hrpA* and *lacZ*, indicating a mutation pressured the colony to keep *hrpA* for a particular advantage. A solid-white colony never possessed *pRC-hrpA-T5*, indicating the colony was derived from a plasmid-free

cell. A sectored-blue colony signified a rapid loss of *pRC-hrpA-T5* as the cells developed [28]. Therefore, the mutations within a solid-white or sectored-blue colony did not demonstrate *hrpA* dependency.

Over several months, about 80,000 colonies were screened for the presence of solid-blue colonies. As expected, the vast majority of colonies were solid-white or sectored-blue (Figure 1). Isolated solid-blue colonies were restreaked for closer observation. Sometimes a restreaked dark blue colony displayed solid-white colonies alongside solid-blue colonies. In order to eliminate false-positives, the solid-white colonies must be smaller than the solid-blue colonies. Larger solid-blue colonies indicated the solid-white colonies were sick, or grew deficiently, due to lack of HrpA. A mutant containing solid-white colonies with equal or improved growth relative to solid-blue colonies was dropped from the potential hit library. In total, 59 colonies were identified as potential hits. Compared to similar synthetic lethal screen studies, the number of potential hits was unusually high [28]. Potential hits 1 to 4 displayed solid-blue and solid-white colonies, but upon close examination the solid-blue colonies displayed improved growth compared to the solid-white colonies (Figure 2). Potential hits 5 to 49 all displayed solid-blue colonies. The intensity, growth, and color of potential hits 5 to 49 were strikingly similar (Figure 3). Given the unusual characteristics and high abundance, it was suspected the highly abundant potential hits represented a contamination of the mutant library. Diagnostic PCR was used to investigate the abnormality.

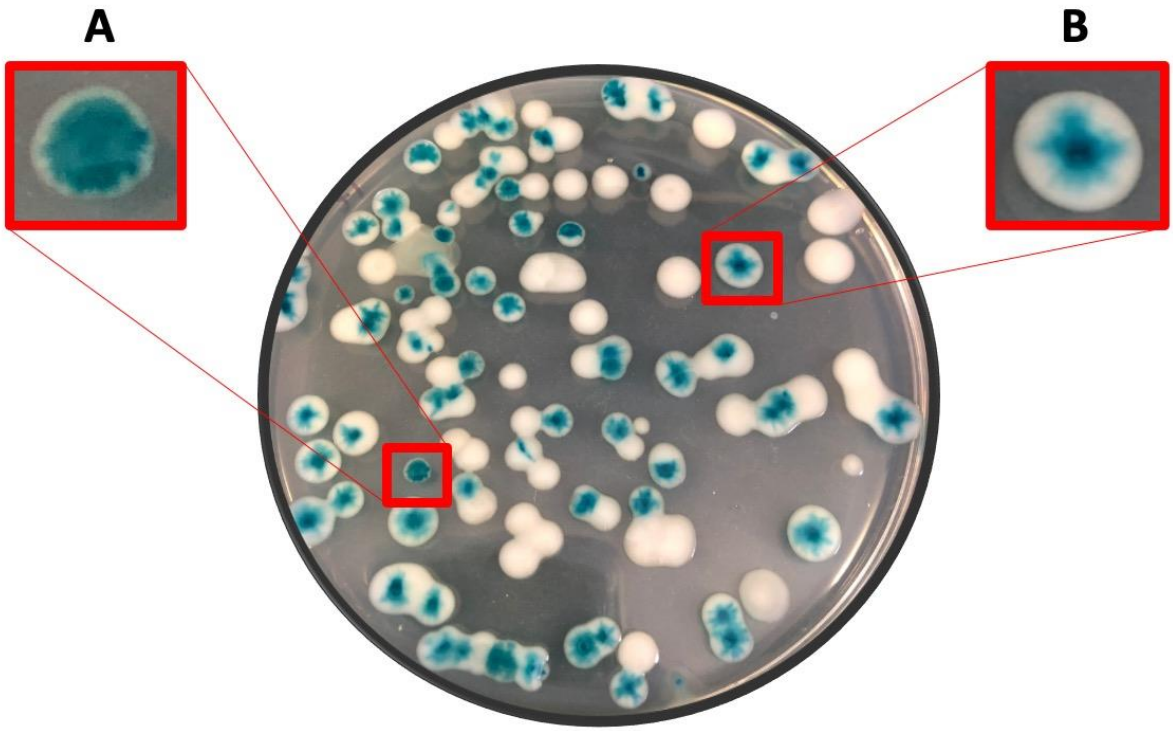

Figure 1: Typical plate from synthetic lethal screen

LB agar Petri plate supplemented with IPTG and X-gal showing ENU treated colonies (genotype: *ΔhrpA lacIZYA::cat* and *pRC-hrpA-T5*) following 72 hours of incubation at 30°C. The solid-blue appearance (A) of potential hit 1 indicates the presence of selective pressure to retain the plasmid. The sectored-blue appearance (B) indicates loss of the plasmid.

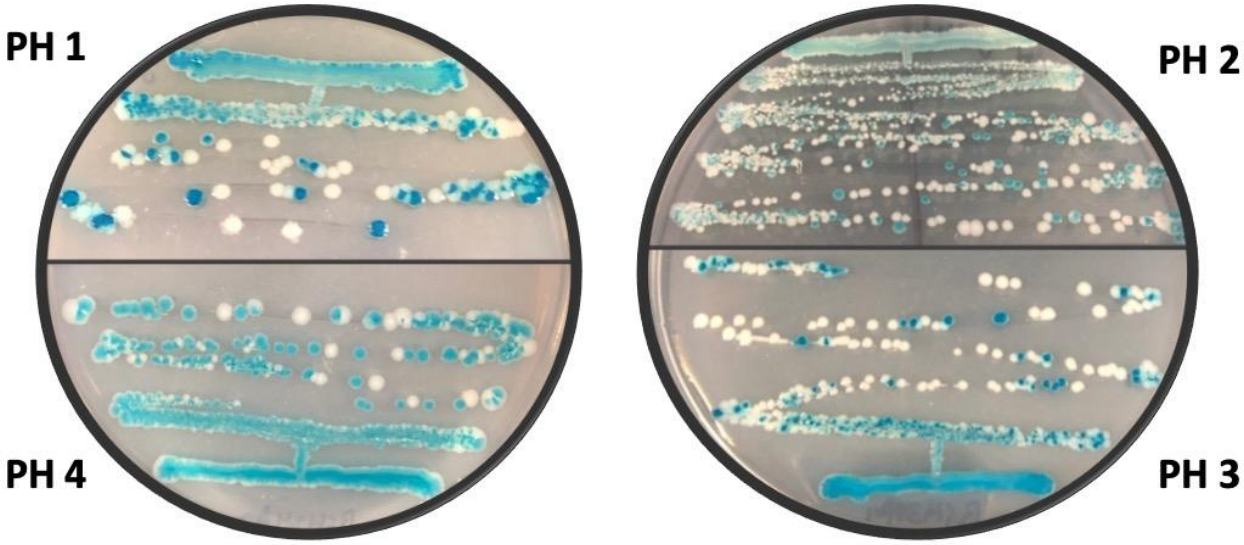

Figure 2: Potential hits 1 through 4

LB agar Petri plates supplemented with IPTG and X-gal showing potential hits following 48 hours of incubation at 30°C.

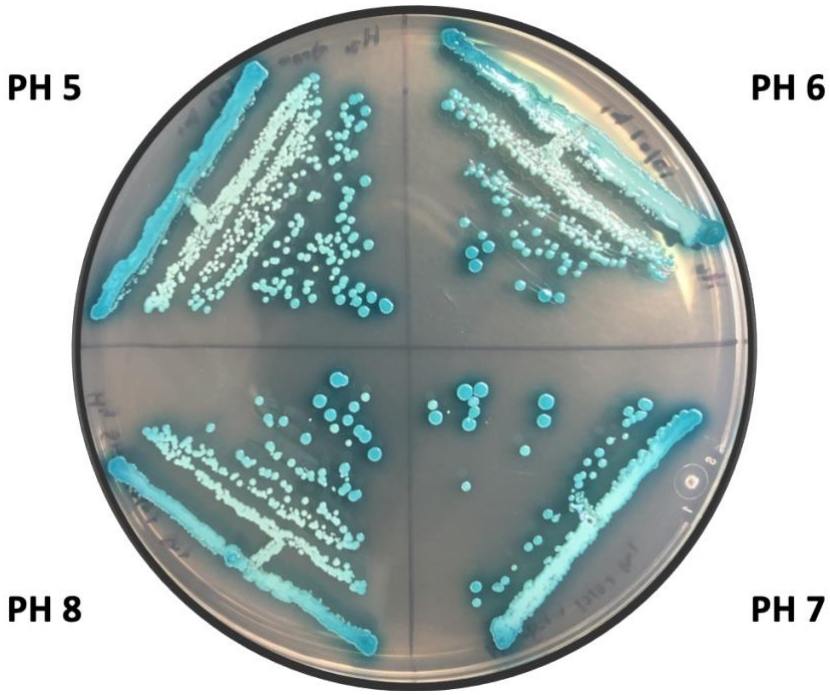

Figure 3: Potential hits 5 through 8

LB agar Petri plates supplemented with IPTG and X-gal showing potential hits following 48 hours of incubation at 30°C. Potential his 9 through 49, not shown, displayed nearly identical phenotypes to potential hits 5 through 8.

Diagnostic Polymerase Chain Reaction

To confirm chromosomal Δ*hrpA* and retention of the plasmid harboring *hrpA* in potential hits 1 through 4, polymerase chain reaction was conducted to amplify specific portions of DNA. For confirming retention of *pRC-hrpA-T5*, a primer pair outlined a 466 bp fragment overlying part of the *lac* operon and the frontal portion of *hrpA* within the plasmid. A different primer pair flanking *hrpA* in wild-type *E. coli* amplified 760 bp if chromosomal Δ*hrpA* or 4,635 bp if wild-type *hrpA*.

Using 1% gel electrophoresis, potentials hits 1 through 4 displayed DNA bands at 760 bp and 466 bp confirming Δ*hrpA* in the chromosome and the presence of *hrpA* in *pRC-hrpA-T5* (Figure 4). Potential hits 5 through 9 did not produce DNA bands in the presence of wild-type *hrpA* primers. Therefore, a second primer pair outlining an approximately 550 bp segment of *rcsF* in *E. coli* was used to determine if the highly abundant potential hits were in fact *E. coli*. Potential hits 5 through 9 did not display DNA bands when PCR amplified with *rcsF* primers, while wild-type *E. coli* displayed a DNA band at 550 bp (Figure 5). The absence of a DNA band indicated the highly abundant potential hits were a contamination of the ENU mutant library. The false positives were removed from further analysis, leaving potential hits 1 through 4 as targets for P1 transduction.

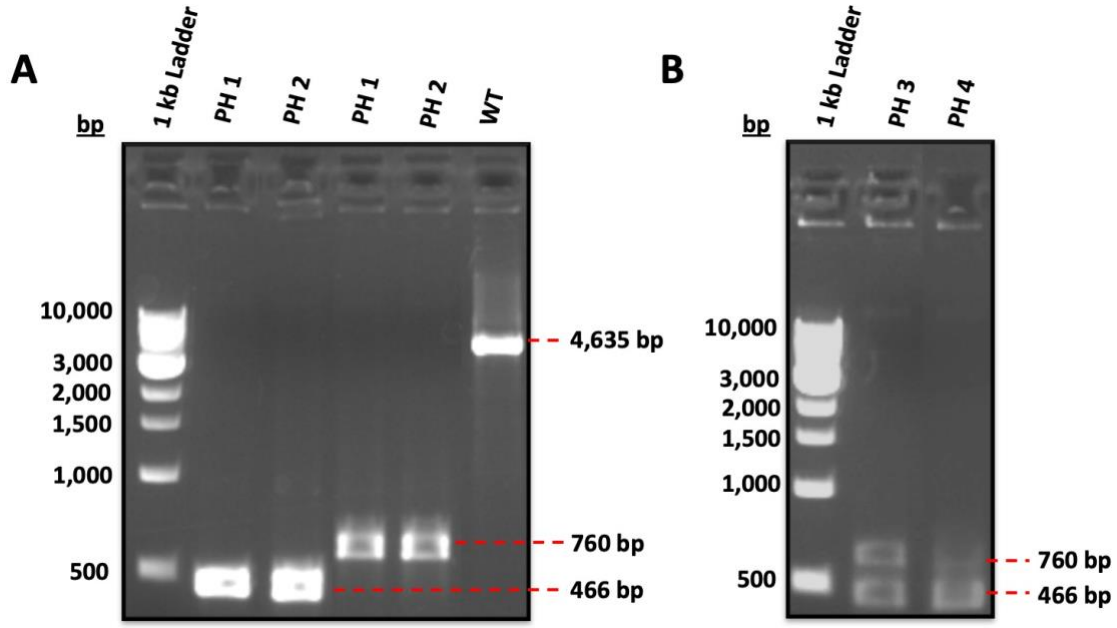

Figure 4: Diagnostic PCR potential hits 1 through 4

Potential hits 1 – 4 displayed DNA bands at 760 bp and 466 bp confirming the presence of chromosomal Δ*hrpA* and plasmid *hrpA*, respectively.

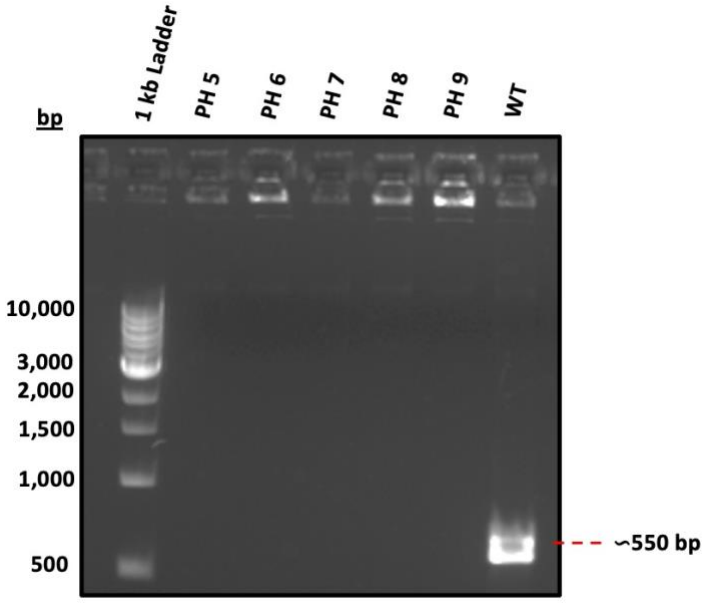

Figure 5: Diagnostic PCR potential hits 5 through 9

The absence of a DNA band at approximately 550 bp indicates potential hits 5 - 9 are contaminants of the *E. coli* mutant library. Similar results were present for potential hits 10 - 49.

P1 Transduction

P1 transduction is a process in *E. coli* by which genes from a donor bacterium are integrated into the genome of a recipient bacterium using a virus known as bacteriophage P1. After infecting a cell, bacteriophage will either enter the lysogenic or lytic cycle. Bacteriophage lay dormant as a prophage in the lysogenic cycle by injecting the phage DNA into the host genome, allowing the prophage genome to be replicated passively as the host cell divides. Stressors to the host cell may cause bacteriophage to enter the lytic cycle, inducting the bacteriophage to hijack the host cell's machinery to build new bacteriophage and eventually release the newly formed bacteriophage by host cell lysis. P1*vir*, a mutated form of phage that has lost its lysogenic ability, is used in P1 transduction to ensure the phage always enters the lytic cycle. Following lytic bacteriophage infection of the donor cell, P1 will sometimes accidently place part of the bacterial genome into a phage head. Release of the newly formed bacteriophage and subsequent infection allows for incorporation of the nonphage donor bacterium DNA into the recipient bacterium's genome through genetic recombination. This process can occur anywhere along the donor bacterium's chromosome or plasmid, allowing any section of the host's DNA to be incorporated into the recipient's DNA. Each phage head can hold about 110 Kb of DNA, allowing the entire *E. coli* genome to be represented in a reasonable number of bacteriophages [29].

Retention of the plasmid harboring *hrpA* and *lacZ* in a solid-blue potential hit did not completely prove the mutant *E. coli* were dependent on *hrpA* for improved fitness. To confirm *hrpA* dependency, P1 transduction was used to restore wild-type *hrpA* into potential hits 1 - 4 containing chromosomal *ΔhrpA.* A cell containing *hrpA* linked to a tetracycline resistant gene was chosen as the donor, allowing verification of successful transduction by plating the

transduced *E.coli* on a tetracycline Petri plate and observing healthy growth. If a potential hit were indeed reliant upon *hrpA*, the mutant *E. coli* would drop the *hrpA* plasmid and phenotypically changed from solid-blue to solid-white, or sectored-blue, upon transduction of chromosomal *hrpA* into the genome of the potential hit. If a potential hit were not actually reliant upon *hrpA*, the mutant *E. coli* retained the *hrpA* plasmid and remained solid-blue upon transduction of chromosomal *hrpA*. P1 transduction allows potential hits that are truly dependent upon *hrpA* to be objectively distinguished from false-positives. Subsequently, P1-mediated transduction mapping is conducted to identify the mutations responsible for *hrpA* dependency in the true potential hits, revealing biological pathways affected by HrpA.

Before starting P1 transduction, the recipient and donor were plated on selective and nonselective plates (Figure 6). When plated on a selective plate, the recipient did not exhibit any viable colonies. An absence of colonies confirmed the recipient did not possess the drug resistant marker. When plated on a selective plate, the donor (RW220) produced viable colonies. A single donor strain colony was selected and restreaked on a selective plate to ensure the donor possessed the drug resistance marker.

Upon transduction of chromosomal *hrpA-tet* into the recipient (Δ*hrpA*), blue and white colonies appeared on the selective plate in potential hits 1 - 4. The white and blue colonies exhibited a similar phenotypic appearance (Figure 7). As a result, mutation within a biological pathway affected by HrpA was not responsible for retention of the plasmid harboring *hrpA* in the synthetic lethal screen. Therefore, P1-mediated transduction mapping and Sanger sequencing were not necessary to identify the mutant genes causing dependence on HrpA for optimal growth.

Ideally all the colonies would drop the plasmid and appear solid-white or sectored-blue following successful P1 transduction, indicating restoration of chromosomal wild-type *hrpA* eliminated the cell's dependency on the *hrpA* plasmid. The phenotypic change from solid-blue to solid-white following P1 transduction would verify the *hrpA* dependency within the potential hits was due to mutation necessitating *hrpA*, rather than unforeseen factors.

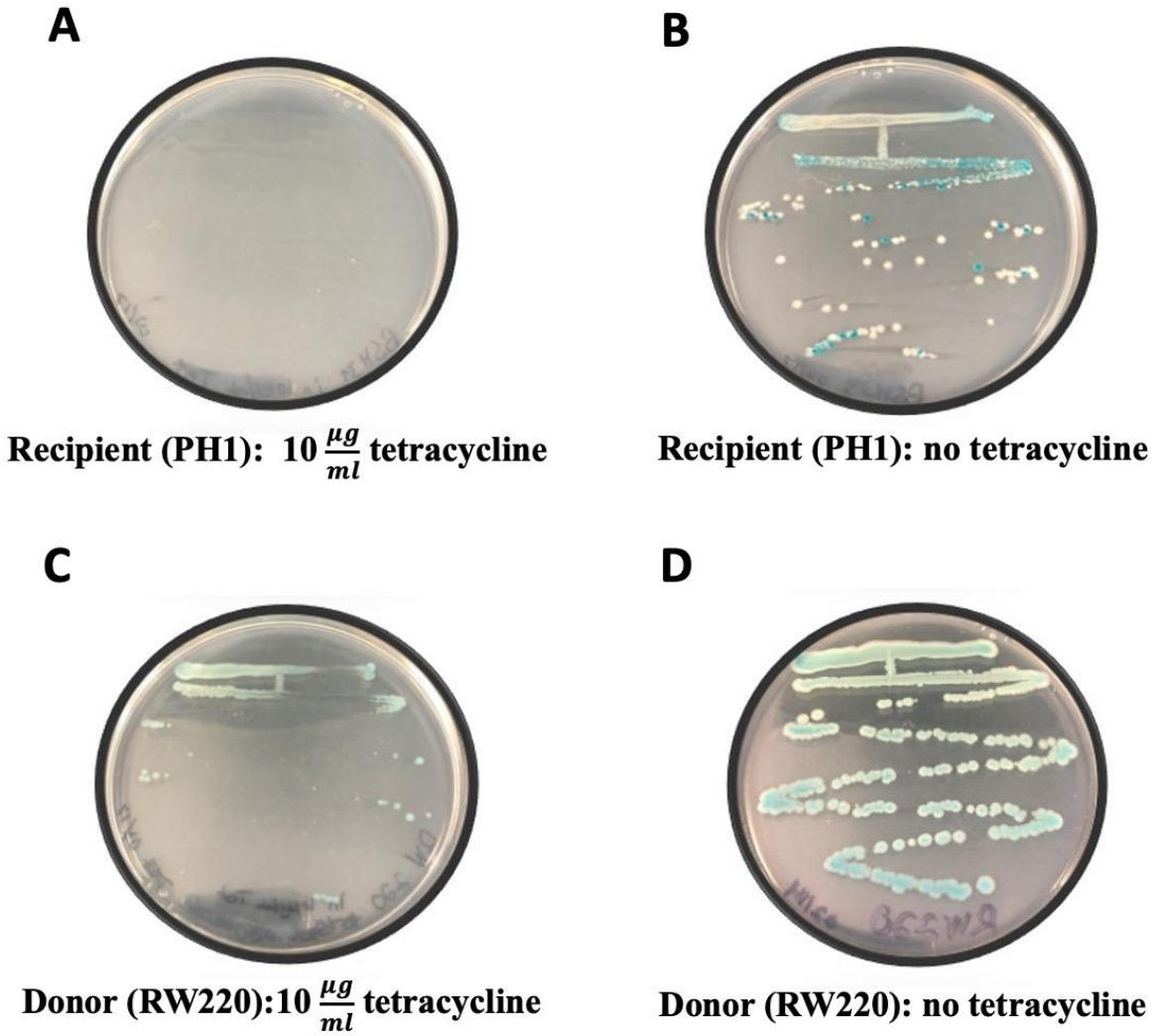

Figure 6: Recipient and donor transduction plates

LB agar Petri plates supplemented with IPTG and X-gal showing potential hits following 48 hours of incubation at 30°C. Recipient PH1 (A) exhibited no viable colonies in the presence of tetracycline. Recipients PH2 – PH4 displayed similar results. Recipient PH1 (B) and Donor RW220 (D) exhibited normal growth in the absence of tetracycline. Recipients PH2 – PH4 displayed similar results. Donor RW220 (C) exhibited growth in the presence of tetracycline.

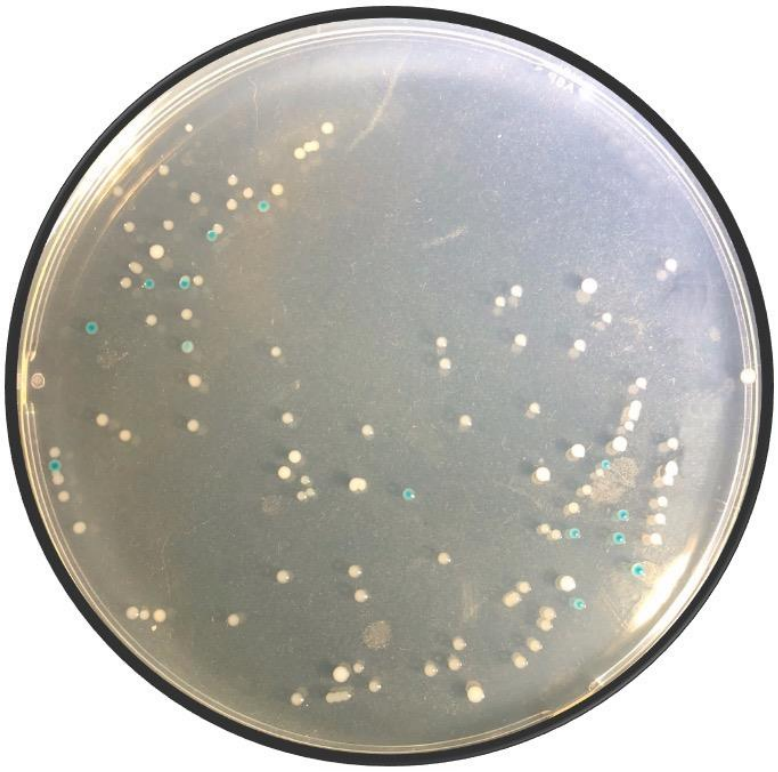

Figure 7: P1 transduced colonies of potential hit 1

LB agar Petri plate supplemented with IPTG, X-gal, sodium citrate, and tetracycline showing transduced colonies following 48 hours of incubation at 30°C. The retention of dark blue colonies following restoration of chromosomal *hrpA* indicates the mutations were not reliant on *hrpA* for optimal growth.

Protein-Protein Interaction Network

STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) is an extensive database of well-established and predicted protein-protein interactions. Comprising over 24 million proteins from more than 5,000 organisms, STRING is one of the most powerful and comprehensive datasets to date [33]. STRING incorporates known functional and physical associations from biochemical, biophysical, and genetic experimental data within primary databases. Predicted interactions are also displayed from computational analysis using

algorithms derived from genomic context and homology across various organisms. The predicated protein-protein interactions are based on the idea that interacting protein orthologs in one organism often interact in another organism. The STRING algorithms take into account the evolutionary distance between two species and the presence of paralogs. Orthologous genes are genes in different species which descended vertically from a common ancestor, while paralogous genes form due to the duplication of one gene into two copies within the same genome. Unlike orthologs, paralogous genes are more likely to diverge and develop new functions due to lack of the original selective pressure on only one copy of the gene. Therefore, the presence of paralogs along the evolutionary tree will lower the predictive association of two proteins [31].

The extensive amount of information within STRING has successfully assisted researchers in an array of endeavors, including identifying biomarkers in various cancers and creating a method to accurately predict operons within *E. coli*. [34, 35]. Although using data obtained from the protein interaction network to prove a protein's molecular function is a difficult task, HrpA's previous experimental data provided a strong foundation for evaluating HrpA's alleged protein interactions within STRING. To date, an in-depth protein interaction network analysis had not been conducted for HrpA. STRING was analyzed in order to identify HrpA's possible enzymatic partners, serving as a steppingstone for further investigation of this protein. STRING yielded twenty-five putative protein interactions with HrpA.

Textmining

Searching for the mention of an uncharacterized protein, such as HrpA, in peer-reviewed publications is often a tedious and time-consuming task. STRING simplified the process by displaying statistically relevant co-incidence of gene names within PubMed abstracts and a wide array of other scientific texts. The database highlighted a study in which an evolved *E. coli*

lineage possessed several mutations resulting in improved fitness when in the presence of isobutanol-induced stress. The temporal order of the mutations revealed *hrpA* was mutated within the first 50 generations. During this period, the growth rate of the evolved *E.coli* more than doubled. Around the 125th generation a mutation in ribosomal protein S2, encoded by *rpsB*, occurred [36]. Given the relationship between *hrpA* and ribosomal stalling in suppressor mutant Δ*efp* cells, the occurrence of mutated *hrpA* and *rpsB* in an evolved isobutanol-tolerant *E. coli* lineage may serve as a starting point for investigating HrpA's specific interaction with ribosomes.

Gene Fusion

A gene fusion is the process by which a hybrid gene is formed from the juxtaposition of two previously independent genes. Gene fusion may occur through structural rearrangements such as translocation, deletion, and inversion or from transcription readthrough of adjacent genes. The resulting multi-domain bacterial proteins may be advantageous through the increased efficiency of coupled biochemical reactions or signal transduction events and the co-regulation of fused domains. A gene fusion becomes fixed during evolution when the fused genes cooperate functionally. Analyzing gene fusions comprising a conserved protein of interest across various species helps identify putative protein-protein interactions [37].

The STRING gene fusion view presented instances of hybrid genes, comprising HrpA (or an HrpA ortholog) and an independent gene, across organisms. STRING revealed HrpA and EndA undergo a gene fusion in *Bombus impatiens* within phylum Arthropoda.

Genetic Neighborhood

Interacting genes are often located in close proximity to each other within the genome for efficient co-regulation and physical interaction. Therefore, analyzing a gene's genetic neighborhood may be used to identify possible interacting or functionally related proteins. A high level of genetic neighborhood conservation across many species may indicate an evolutionary advantage obtained by a particular genomic organization [38]. The STRING genetic neighborhood view displayed gene pairs with a conserved topological neighborhood across organisms.

Upstream of *hrpA* is *azoR*, which encodes a quinone reductase used to reduce thiol-specific stress inflicted by electrophilic quinones. Specifically, AzoR helps process azo dyes that spread into the environment. Azo dyes are commonly found in textiles, cosmetics, and pharmaceuticals. The breakdown products of azo compounds are often toxic or mutagenic [39]. Directly downstream of *hrpA* is *ydcF*, which encodes for a plasma membrane protein containing domain of unknown function 218. Information on YdcF is limited. YdcF is thought to interact with ribosomal protein L2 and may be involved in the anaerobic respiratory pathway [40]. Downstream of *ydcf* is *aldA*, which encodes a lactaldehyde dehydrogenase. This oxidoreductase is involved in several metabolic pathways [41].

HrpB

HrpB is a DEAH-box helicase comprised of multiple motifs sharing strong homology to HrpA. Instead of HrpA's large domain of unknown function 3418, HrpB contains a much smaller ATP-dependent helicase C-terminal domain. A comparison of HrpA's and HrpB's

STRING networks reveals drastically different protein-protein interactions across all active interaction sources (Figure 8). Based on HrpA and HrpB's domain structure analysis, the C-terminal domain of each protein is responsible for the two different protein interaction networks.

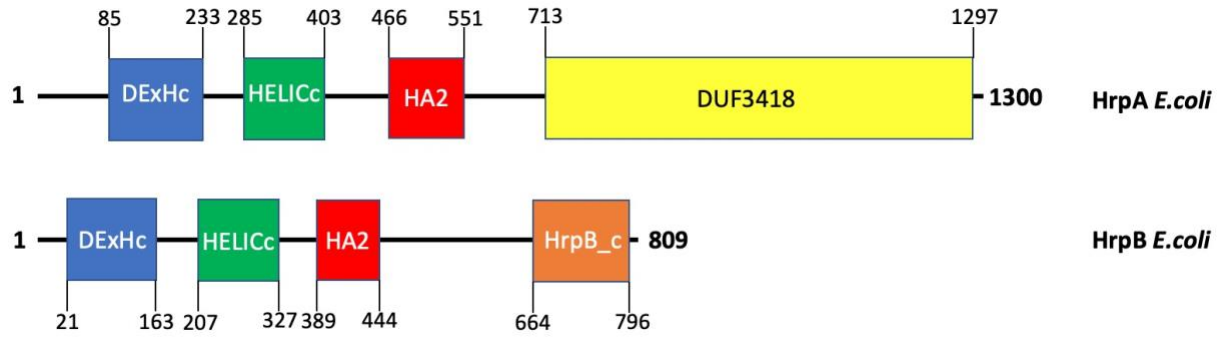

Figure 8: Domains of HrpA and HrpB

Domain structure analysis of *hrpA and hrpB* in *E.coli.* The domain structures were identified using the Pfam protein families database. Domain symbols: DExHc Box, DEAH-like helicase superfamily; HELICc, Helicase superfamily C-terminal domain; HA2, Helicase-associated domain; DUF3418, Domain of unknown function 3418; HrpB_c, ATP-dependent helicase C-terminal domain.

DISCUSSION

Synthetic Lethal Screen

In order to identify mutations causing *hrpA* dependency in *E. coli*, a synthetic lethal screen was conducted. The screen was anticipated to reveal novel genes and biological pathways affected by the HrpA helicase. Synthetic lethal screens of similar size have successfully identified mutations causing dependence on the gene of interest [28]. Therefore, the failure of the synthetic lethal screen in this study was likely due to user-error or perhaps Δ*hrpA*'s inability to visually confer a synthetically lethal phenotype. The solid-blue appearance of the potential hits varied in intensity from a heavier dark blue to a lighter aqua blue. The varying phenotypes suggest some of the potential hits may have exhibited only slight levels of *hrpA*-dependency. Nevertheless, the complications involved in the synthetic lethal screen should not rule out another attempt of this powerful genetic tool.

Diagnostic Polymerase Chain Reaction

Unfortunately, diagnostic PCR revealed a widespread contamination within the mutant library. The false-positives were eliminated from further investigation but may have acted as a source of distraction during the process of scanning for potential hits. Due to the contamination's similar appearance to *E. coli*, the species of the contamination was not identified and was difficult to distinguish by visual means. The contamination was found in all 1 mL aliquots of the mutant library, indicating the contamination was wide-spread and infection likely occurred during ENU treatment.

P1 Transduction

P1 transduction is a powerful tool to confirm *hrpA* dependence within the context of a synthetic lethal screen, especially in restreaked bilateral potential hits displaying both solid-blue and solid-white colonies. Since individual cells in a bilateral colony retaining the *hrpA* plasmid cannot be feasible separated from cells without the *hrpA* plasmid, conventional methods of comparing solid-blue and solid-white colony fitness with bacterial growth curves could not be used in this study. Therefore, visual observation was used to ensure the solid-blue colonies displayed improved fitness compared to solid-white colonies in the bilateral potential hits. In order to eliminate potential inaccuracies while comparing growth of sold-white and solid-blue colonies within a single potential hit by visual means alone, P1 transduction distinguished true potential hits from false-positives by observing phenotypic change upon transduction of wild-type *hrpA* into each potential hit.

The transduced colonies remained solid-blue and solid-white, indicating retention of the *hrpA* plasmid in the solid-blue colonies following integration of chromosomal *hrpA* into potential hits 1 - 4. Therefore, mutation of an HrpA related biological pathway was not responsible for retention of the *hrpA* plasmid. Perhaps random mutation caused overexpression of *lacZ*, resulting in increased conversion of X-gal to a blue precipitate. Additionally, visual observation may have falsely concluded the solid-blue colonies displayed improved fitness relative to the solid-white colonies within a bilateral potential hit. Nevertheless, the framework of this study remains a powerful tool by providing a three-step process of confirming *hrpA* dependency in the potential hits and eliminating all possible subjective experimental bias through screening a synthetic lethal library, diagnostic PCR, and P1 transduction.

Protein-Protein Interaction Network

Proteins rarely function alone and are often regulated by other cellular components. Protein-protein interactions govern almost every process within a cell and their characterization is critical to fully outlining pathways of cellular physiology. HrpA's obscured physiological role made investigating HrpA under wet experimental procedures difficult.

Textmining

Using STRING's textmining feature, a study was found in which E. coli were subjected to isobutanol-induced stress. Isobutanol, along with other alcohols, intercalate the cell's lipid bilayer disrupting the cell's electrochemical gradients and physiological processes. Microbial tolerance to toxic solvents is complex and involves various cellular pathways [42]. Therefore, the mutations involved in conferring isobutanol tolerance in the evolved *E. coli* lineages often involve epistatic interactions among distal genes. Epistasis occurs when the effect of a certain mutation is dependent on the genomic context. In other words, a mutation may only be effective in the presence or absence of other mutations [43]. Genomic resequencing of evolved isobutanol tolerant colonies revealed multiple mutations, which were identified temporally. Within the first 50 generations of the *E. coli* lineage, a mutation within *hrpA* was identified and resulted in improved fitness [36]. Similar to *hrpA*'s deletions found in *Δefp* fast-growing suppressor mutants, the study suggests disrupting HrpA's function may play a role in isobutanol tolerance or other multifaceted stress tolerance responses. Interestingly, a mutation in *rpsB* was identified around the 125th generation in the same *E. coli* lineage. RpsB, also known as ribosomal protein S2, allows ribosomal protein S1 to bind to the 30S subunit [44]. S1 is essential for translation initiation and may be involved in other mRNA transactions [45]. HrpA's C-terminal domain includes an RNA binding motif weakly resembling ribosomal protein S1. Further evidence must

be conducted to evaluate if *hrpA* and *rspB* comprise an epistatic protein-protein interaction, but their presence in the same evolved *E. coli* lineage ought to be noted given HrpA's association with ribosomal stalling. To date, there is little evidence on isobutanol's impact on translation. However similar aliphatic alcohols, such as ethanol, are known to decrease translational processivity and exhibit ribosomal stalling [46].

Gene Fusion

Gene fusion is a powerful tool because the formation of hybrid genes fixed during evolution indicates a functional advantage obtained by the gene fusion event of two independent genes. Therefore, a functional link can be predicted to exist between two highly preserved fusion-linked genes [37]. HrpA's fusion with EndA only occurs in *Bombus impatiens* of phylum Arthropoda. This single gene fusion event does not demonstrate evolutionary fixation, thus a selective advantage to HrpA and EndA's fusion cannot be outright inferred. EndA is an endonuclease that cleaves duplex DNA to form 7 base pair oligonucleotides. EndA's cellular role is unknown and *⊿endA* does not produce a phenotypic effect [47]. EndA's biological activity does not align with any pathways of HrpA's proposed cellular function. As a result, the gene fusion analysis failed to add a significant contribution to the aims of this thesis.

Genetic Neighborhood

Careful analysis of the evolutionary tree showcasing *hrpA*'s genetic neighborhood with *azoR*, *ydcF*, and *aldA* reveals there is a lack of evidence to infer that the location of *hrpA's* gene neighbors resulted from a significant evolutionary advantage of localizing interacting proteins. *HrpA*, *ydcF*, and *aldaA* are only proximal to each other in a handful of families under the Proteobacteria phylum. Likewise, *hrpA* and *azoR* exsist proximal in the genomes of only two

families: Enterobacteriaceae and Vibrionaceae. Due to the lack of genetic conservation, the genetic neighborhood analysis failed to add a significant contribution to the aims of this thesis.

HrpB

HrpB and HrpA's vastly different STRING networks adds to the evidence that HrpA's DUF3418 is responsible for HrpA's biological role. HrpA's C-terminal domain contains an RNA binding motif similar to ribosomal protein S1. Also when fused with a fluorescent protein, HrpA's C-terminal co-migrates in polysomes nearly identical to wild-type HrpA. The findings indicate DUF3418 likely serves as a guide as it recruits HrpA's RNA helicase activity to a specific elongational event in the cell.

Future Directions

The discussed protein-protein interactions are based solely on STRING's algorithms and published experimental evidence. Previous studies demonstrate HrpA's obscured physiological role makes it difficult to investigate through wet experimental procedures. However, the predicted protein interactions analyzed in this thesis provide a source of further investigation into HrpA's function and related pathways. Additionally, the synthetic lethal screen remains a powerful tool for identifying *E. coli* mutants reliant on HrpA for optimal growth. The widespread presence of false-positive contaminants may have drawn attention away from finding true potential hits. The creation of a brand-new, larger, and contamination-free mutant library is optimal for uncovering HrpA's enzymatic partners. Before conducting the synthetic lethal screen, the possible presence of contamination should be examined by plating small samples from several freezer stocked 1 ml aliquots of the mutant library. If non-*E.coli* highly abundant solid-blue colonies are present, then the sources of the mutant library need to be tested for

contamination and the methodology behind building the mutant library ought to be examined for possible routes of contamination.

Polysome profiling and examining growth on a Petri plate indicated Δ*efp* Δ*hrpA* colonies were healthier than Δ*efp* colonies [23]. In order to gain further insight on how deletion of *hrpA* can affect the global physiology of *E. coli*, total RNA sequencing could be conducted to identify how the transcriptome differs between *wild-type*, Δ*hrpA*, Δ*efp*, and Δ*efp* Δ*hrpA* colonies. Ribosome profiling, also known as ribosome foot printing, will indicate how the ribosomes are arranged throughout the transcriptome in each of the genotypes. Additionally, wild-type HrpA and a mutant version of HrpA composed of HrpA's C-terminal domain linked to a fluorescent protein were both found to co-migrate with polysomes [23]. The protein interaction network uncovered a study in which *hrpA* and *rpsB* were both found mutated in an evolved isobutanol-tolerant *E.coli* lineage [36]. Inactivating *rpsB* and observing how *rpsB*'s loss of functionality affects the comigration of HrpA and DUF3418 with polysomes would provide more guidance in evaluating the alleged interaction between *hrpA* and *rpsB*. Also, the author did not indicate the specific base pair location of the mutations in the evolved isobutanol-tolerant lineage. Contacting the author will shed light into whether the mutation occurred in HrpA's N-terminal helicase core domains or DUF3418. Comparing the location and type of mutation in fast-growing suppressor mutant Δ*efp* cells and the mutant isobutanol-tolerant cells may help decipher which portions of *hrpA* are vital to functionality.

APPENDIX A: DIAGNOSTIC PCR PRIMERS AND PARAMETERS

Table 1: HrpA Primers

Gene	Location	T_m	Primers		Product Length (bp)	
			Forward	Reverse		
hrpA	pRC-hrpA-T5	60 °C	5’ – CAG CTC ATG TTA TAT CCC GCC G – 3’	5’- CAG ACG GCG AGA AAA ACG CAG – 3’	466	
hrpA	Chromosome	60 °C	5’ – CCC TGC CAG GAT GCT GG – 3’	5’ – GCG GAG CAC TCT GAA AGG AG – 3’	**WT** 4365	**Δ*hrpA*** 760

Table 2: Polymerase Chain Reaction Parameters

Step	PCR Temp. (°C)	Time (min)	Cycles
Initial denaturation	95	2	1x
Denaturation	95	0.5	30x
Annealing	60	0.75	
Extension	72	1	
Final Extension	72	5	1x
Storage	4	∞	-

APPENDIX B: REAGENTS FOR P1 TRANSDUCTION

LB Plates (Formula per Liter)

- 5.0 g NaCl
- 5.0 g Yeast
- 10.0 g Tryptone
- 2 pellets KOH
- 12.0 g Agar
- 0.5 mM IPTG
- 80 $\frac{\mu g}{mL}$ X-gal
- 0.2% glucose
- Dissolve in deionized water and autoclave at 121ºC for 35 minutes
- **Selective plate:** 10 $\frac{\mu g}{mL}$ tetracycline
    - (spread on surface of plate and let diffuse for several hours)
- **Selective plate with 1 mM Sodium Citrate:** Create mixture of 1 mM sodium citrate stock, 10 $\frac{\mu g}{mL}$ tetracycline, and add diH_2O up to a final volume of 100 μL
    - (spread on surface of plate and let diffuse for several hours)

P1-LB

- LB broth
- 12.5 mM $MgCl_2$
- 5 mM $CaCl_2$
- 0.1% glucose

LB-Citrate

- LB Broth
- 100 mM Na-citrate
- Final pH 5.5 +/- 0.2

1M Sodium Citrate

- 1 M Citric Acid
- Adjust to Final pH 5.5 +/- .2 with NaOH

Chloroform

- Stabilized with 1% ethanol or 0.1% isoamyl alcohol

Bacteriophage P1*vir* stock

- Preexisting phage stock with a confirmed transduction activity

REFERENCES

1. Rodnina, M.V., *Translation in Prokaryotes.* Cold Spring Harb Perspect Biol, 2018. **10**(9).

2. Kohler, R., et al., *Architecture of a transcribing-translating expressome.* Science, 2017. **356**(6334): p. 194-197.

3. Gualerzi, C.O. and C.L. Pon, *Initiation of mRNA translation in bacteria: structural and dynamic aspects.* Cell Mol Life Sci, 2015. **72**(22): p. 4341-67.

4. Rodnina, M.V., et al., *Ribosome dynamics during decoding.* Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1716).

5. Yamamoto, H., et al., *EF-G and EF4: translocation and back-translocation on the bacterial ribosome.* Nature Reviews Microbiology, 2014. **12**(2): p. 89-100.

6. Janosi, L., et al., *Ribosome recycling by ribosome recycling factor (RRF) — An important but overlooked step of protein biosynthesis.* Advances in Biophysics, 1996. **32**: p. 121-201.

7. Ranji, A. and K. Boris-Lawrie, *RNA helicases: emerging roles in viral replication and the host innate response.* RNA biology, 2010. **7**(6): p. 775-787.

8. Anantharaman, V., E.V. Koonin, and L. Aravind, *Comparative genomics and evolution of proteins involved in RNA metabolism.* Nucleic Acids Research, 2002. **30**(7): p. 1427-1464.

9. Jankowsky, E. and M.E. Fairman, *RNA helicases--one fold for many functions.* Curr Opin Struct Biol, 2007. **17**(3): p. 316-24.

10. Fuller-Pace, F.V., *DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation.* Nucleic Acids Research, 2006. **34**(15): p. 4206-4215.

11. Gustafson, E.A. and G.M. Wessel, *DEAD-box helicases: Posttranslational regulation and function.* Biochemical and Biophysical Research Communications, 2010. **395**(1): p. 1-6.

12. Parsyan, A., et al., *mRNA helicases: the tacticians of translational control.* Nature Reviews Molecular Cell Biology, 2011. **12**(4): p. 235-245.

13. Hanson, P.I. and S.W. Whiteheart, *AAA+ proteins: have engine, will work.* Nat Rev Mol Cell Biol, 2005. **6**(7): p. 519-29.

14. Jankowsky, E., *RNA helicases at work: binding and rearranging.* Trends in Biochemical Sciences, 2011. **36**(1): p. 19-29.

15. Moriya, H., H. Kasai, and K. Isono, *Cloning and characterization of the hrpA gene in the terC region of Escherichia coli that is highly similar to the DEAH family RNA helicase genes of Saccharomyces cerevisiae.* Nucleic Acids Res, 1995. **23**(4): p. 595-8.

16. Liu, Y.C. and S.C. Cheng, *Functional roles of DExD/H-box RNA helicases in Pre-mRNA splicing.* J Biomed Sci, 2015. **22**: p. 54.

17. Naganathan, A., M.P. Wood, and S.D. Moore, *The large ribosomal subunit protein L9 enables the growth of EF-P deficient cells and enhances small subunit maturation.* PLoS One, 2015. **10**(4): p. e0120060.

18. Blaha, G., R.E. Stanley, and T.A. Steitz, *Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome.* Science, 2009. **325**(5943): p. 966-70.

19. Huter, P., et al., *Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P.* Mol Cell, 2017. **68**(3): p. 515-527.e6.

20. Pavlov, M.Y., et al., *Slow peptide bond formation by proline and other N-alkylamino acids in translation.* Proc Natl Acad Sci U S A, 2009. **106**(1): p. 50-4.

21. Doerfel, L.K., et al., *EF-P is essential for rapid synthesis of proteins containing consecutive proline residues.* Science, 2013. **339**(6115): p. 85-8.

22. Buchan, J.R. and I. Stansfield, *Halting a cellular production line: responses to ribosomal pausing during translation.* Biology of the Cell, 2007. **99**(9): p. 475-487.

23. Wingo, R., *Exacerbation of ΔEFP Sickness in Escherichia coli By an Uncharacterized RNA Helicase.* STARS, 2018: p. 82.

24. Jankowsky, E. and M.E. Fairman-Williams, *Chapter 1 An Introduction to RNA Helicases: Superfamilies, Families, and Major Themes*, in *RNA Helicases*. 2010, The Royal Society of Chemistry. p. 1-31.

25. Carabetta, V.J., T.J. Silhavy, and I.M. Cristea, *The Response Regulator SprE (RssB) Is Required for Maintaining Poly(A) Polymerase I-Degradosome Association during Stationary Phase.* Journal of Bacteriology, 2010. **192**(14): p. 3713.

26. Steimer, L. and D. Klostermeier, *RNA helicases in infection and disease.* RNA Biol, 2012. **9**(6): p. 751-71.

27. Salman-Dilgimen, A., et al., *HrpA, a DEAH-Box RNA Helicase, Is Involved in Global Gene Regulation in the Lyme Disease Spirochete.* PLOS ONE, 2011. **6**(7): p. e22168.

28. Bernhardt, T.G. and P.A. de Boer, *Screening for synthetic lethal mutants in Escherichia coli and identification of EnvC (YibP) as a periplasmic septal ring factor with murein hydrolase activity.* Mol Microbiol, 2004. **52**(5): p. 1255-69.

29. Moore, S.D., *Assembling new Escherichia coli strains by transduction using phage P1.* Methods Mol Biol, 2011. **765**: p. 155-69.

30. Szklarczyk, D., et al., *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.* Nucleic Acids Res, 2019. **47**(D1): p. D607-d613.

31. Szklarczyk, D., et al., *The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible.* Nucleic Acids Research, 2016. **45**(D1): p. D362-D368.

32. de Boer, P.A., R.E. Crossley, and L.I. Rothfield, *A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli.* Cell, 1989. **56**(4): p. 641-9.

33. Rao, V.S., et al., *Protein-protein interaction detection: methods and analysis.* International journal of proteomics, 2014. **2014**: p. 147648-147648.

34. Li, L., et al., *Screening and identification of key biomarkers in hepatocellular carcinoma: Evidence from bioinformatic analysis.* Oncol Rep, 2017. **38**(5): p. 2607-2618.

35. Taboada, B., C. Verde, and E. Merino, *High accuracy operon prediction method based on STRING database scores.* Nucleic Acids Res, 2010. **38**(12): p. e130.

36. Minty, J.J., et al., *Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in Escherichia coli.* Microbial Cell Factories, 2011. **10**(1): p. 18.

37. Yanai, I., Y.I. Wolf, and E.V. Koonin, *Evolution of gene fusions: horizontal transfer versus independent events.* Genome Biology, 2002. **3**(5): p. research0024.1.

38. Dandekar, T., et al., *Conservation of gene order: a fingerprint of proteins that physically interact.* Trends Biochem Sci, 1998. **23**(9): p. 324-8.

39. Liu, G., et al., *The Escherichia coli Azoreductase AzoR Is Involved in Resistance to Thiol-Specific Stress Caused by Electrophilic Quinones.* Journal of Bacteriology, 2009. **191**(20): p. 6394-6400.

40. Chao, K.L., et al., *The Escherichia coli YdcF binds S-adenosyl-L-methionine and adopts an α/β-fold characteristic of nucleotide-utilizing enzymes.* Proteins: Structure, Function, and Bioinformatics, 2008. **72**(1): p. 506-509.

41. Limón, A., E. Hidalgo, and J. Aguilar, *The aldA gene of Escherichia coli is under the control of at least three transcriptional regulators.* Microbiology (Reading), 1997. **143 (Pt 6)**: p. 2085-95.

42. Nicolaou, S.A., S.M. Gaida, and E.T. Papoutsakis, *A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation.* Metab Eng, 2010. **12**(4): p. 307-31.

43. Domingo, J., P. Baeza-Centurion, and B. Lehner, *The Causes and Consequences of Genetic Interactions (Epistasis).* Annu Rev Genomics Hum Genet, 2019. **20**: p. 433-460.

44. Moll, I., et al., *Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in Escherichia coli.* Mol Microbiol, 2002. **44**(5): p. 1387-96.

45. Delvillani, F., et al., *S1 ribosomal protein and the interplay between translation and mRNA decay.* Nucleic Acids Res, 2011. **39**(17): p. 7702-15.

46. Haft, R.J.F., et al., *Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria.* Proceedings of the National Academy of Sciences, 2014. **111**(25): p. E2576-E2585.

47. Lehman, I.R., G.G. Roussos, and E.A. Pratt, *The deoxyribonucleases of Escherichia coli. II. Purification and properties of a ribonucleic acid-inhibitable endonuclease.* J Biol Chem, 1962. **237**: p. 819-28.