Evaluation of Novel Viral-Host Protein Interactions

By:

Holly Passetti

A Thesis Submitted to the Faculty of Alfred University

In partial fulfillment of the requirements for:

Alfred University Honors

Alfred University Biology Honors

April 2021

Under the Supervision of:

Chair: Dr. Geoffrey Lippa, Assistant Professor of Biology, Alfred University

Committee Members:

Dr. Jean Cardinale, Chair of Biology Department and Professor of Biology, Alfred University

Dr. John D'Angelo, Associate Professor of Chemistry, Alfred University

Acknowledgements

I would like to start by thanking Dr. Lippa for his guidance, time, and help throughout the completion of this thesis. Without his knowledge and support, this would not have been possible. I would also like to thank him for always challenging me to think about things on a deeper level and giving me advice and tools that will be helpful as I continue my academic career in graduate school. The time I have spent in the lab on this thesis has really solidified my desire to pursue a career in research.

I would also like to thank Dr. Cardinale and Dr. D'Angelo for their support of this thesis and my endeavors. Both professors have been mentors to me throughout my academic career and I cannot thank them enough for the role they have played in my time at Alfred University.

I would like to acknowledge the work that Madison Muehl ('19) completed for the foundations of this thesis. Her prior research and analysis made the following experimentation justifiable and helps to explain the purpose of this research.

Lastly, I would like to thank my family, friends, and especially my two roommates for the immense amount of support and encouragement they have given me throughout this process. Whenever I was frustrated or overwhelmed with the processes, I always knew I could count on them. Thank you to everyone who made this thesis possible.

Table of Contents

Table of Figures

Table of Tables

Abstract

Recent studies have shown that some viruses are able to hijack the cell's ubiquitination processes. Ubiquitination is used to target proteins for degradation in the cell; it also has a responsibility in immune signaling. Viruses have been able to manipulate this system to halt immune signaling and to degrade antiviral proteins in the cell. By studying viral and host protein interactions, the virus' mechanism of infection can be revealed, and creation of novel and effective therapeutics can begin.

We predict that the Hepatitis C viruses manipulate the process of ubiquitination. Our lab has been working with the viral protein, NS5A, to evaluate its interactions with host proteins, Elongin B and Elongin C, which are part of the Cullin-RING ligase (CRL) complex of the ubiquitination process. The cell signaling protein suppressor of cytokine signaling protein 2, SOCS2, is used as a control, because it is an endogenous substrate recognition subunit that interacts with Elongin B and C. To test the hypothesis that Hepatitis C virus manipulates the process of ubiquitination by the binding of NS5A to the CRL complex through Elongin B and C, we designed and constructed unique co-expression constructs for Elongin B, Elongin C, and our viral target. Protein expression was conducted in *Escherichia coli*. We isolated the complex while verifying viral-host protein interactions using immobilized metal affinity chromatography.

This research will reveal if this virus infects host cells by hijacking the ubiquitination process. If this is confirmed, continued experimental steps can be taken to start designing unique and effective therapeutics against these viral infections.

Introduction

First identified in 1989, Hepatitis C virus (HCV) is a positive strand single stranded RNA virus of the family Flaviviridae and genus Hepacivirus.[1] Such viruses have a genome comprised of ribonucleic acid that can act as messenger RNA and can be directly translated into proteins using the host cells ribosomes.

HCV causes the disease known as Hepatitis C. Hepatitis C (Hep C) is a liver infection, causing inflammation and damage to the liver, mainly spread through contact with an infected person's blood. The disease can lead to liver cirrhosis, end stage liver disease, and/or hepatocellular carcinoma.[1] Hep C infection can be acute or chronic. The symptoms include fever, fatigue, abdominal pain, loss of appetite, nausea, vomiting, joint pain, and jaundice.[2] Some people may not experience symptoms at all and are accidentally diagnosed through increased alanine aminotransferase levels in routine blood work, indicating hepatic injury and tissue necrosis.[1] HCV has seven genotypes with 67 different subspecies; genotype 1a and 1b are the most common in the United States.[1,2] Genotype 3 is the most common in India, genotype 4 is most common in Africa and the Middle East, and genotype 6 is most common in southeast Asia and South Africa.[1] The genotypes of HCV are area specific, making it that much harder to treat or eradicate on a global level.

The genome of HCV is 9.6 kilobases, which code for a 3,000 amino acid polyprotein. This polyprotein is cleaved into structural and non-structural (NS) proteins. Structural proteins (C, E1, and E2) form the nucleocapsid core (C) and envelope glycoproteins (E1 and E2) as shown in Figure 1. The nonstructural proteins are labeled NS2 to NS5. Some of the functions of these proteins are still unknown. However, NS3 is found to have helicase and protease activities and

NS5 functions in RNA viral replication with RNA dependent RNA polymerase.[1] In this thesis, we specifically focus on NS5A, a phosphorylated metalloprotein. It is said to play roles in virus replication, regulation of cellular pathways, and interferon resistance.[3] These characteristics make it a protein of interest for research.

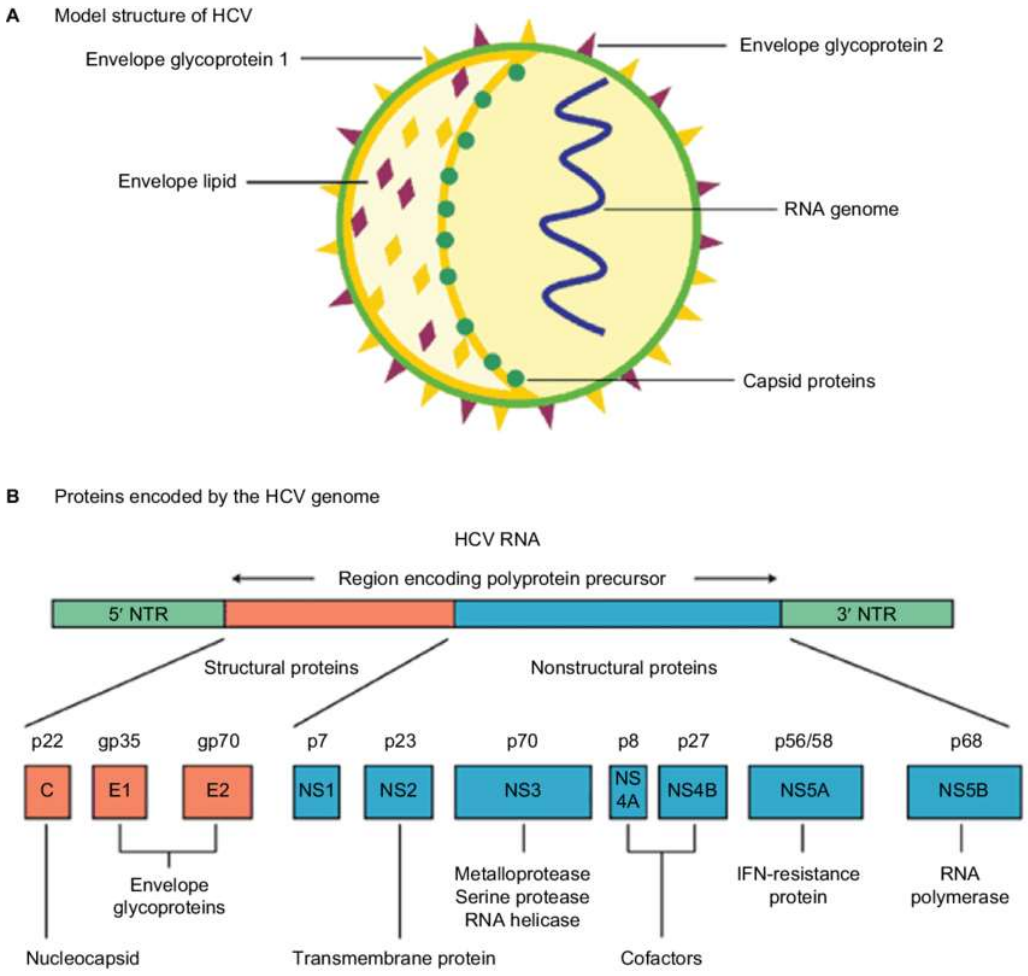


Figure 1. Structure of HCV (A) along with its genome with what proteins it codes for and their function (B).[4]

HCV replicates in the cytoplasm of hepatocytes. It can replicate very quickly, spread cell-to-cell efficiently, and doesn't initiate a T-cell immune response to HCV antigens, making this type of infection persistent. The viral production of HCV is very high, producing 10^{10}-10^{12} virions per cell per day.[1] The genome of the virus also mutates frequently, due to the virus's rapid

replication and lack of error proofing from its RNA polymerase. Therefore, patients positive for HCV may have a plethora of HCV genomes in their body, differing 1%-5% in nucleotide sequence.[1] This heterogeneity within patients and the fact that immune pressure causes the envelop proteins to change, leads to immune escape and chronicity. These reasons, among others, makes HCV very hard to treat. Acute Hep C may be treated with antivirals, depending on the genotype of the virus contributing to the infection. However, chronic Hep C infection may necessitate liver transplantation.[2] Because of the high prevalence of liver transplantations and cancer, it's important to find more effective therapeutics.

There is no vaccine for Hep C, although there is one for Hepatitis A and B, and little progression has been made towards one. This is due to a plethora of reasons. There are no robust tissue cultures of this disease or valuable animal models to use for research. Also, with the multiple genomes of HCV, multiple vaccines would also need to be created, making the task that much more difficult. Even if vaccines were developed, immunity would not be complete, again due to the multiple genotypes of HCV.[1] Therefore, reinfection could easily occur.

The fact that vaccines are improbable leads to research in developing new treatments for Hepatitis C. One method of viral infection that is currently being researched is viral hijacking of host ubiquitination systems. Ubiquitination is a process in which unwanted proteins are tagged with a ubiquitin protein(s) for degradation by the proteosome. These unwanted proteins could be no longer functional or just in surplus in the body. This process occurs in a three step procedure starting with a ubiquitin activating enzyme (E1), then a ubiquitin conjugating enzyme (E2), and finally a ubiquitin protein ligase (E3).[5] The most common ubiquitin protein ligase is the Cullin-RING ligase (CRL) as shown in Figure 3. It is comprised of a

Cullin scaffold protein, RING finger protein, adaptor proteins, and a substrate receptor that binds a substrate. The CRL specifically looked at in this thesis utilizes a Cullin-5 scaffold protein, RING H2 zinc finger protein, Elongin B and C (adaptor proteins), and suppressor of cytokine signaling protein 2 (SOCS2), as a substrate receptor, that binds via a SOCS-box binding domain.[6] In normal function, SOCS2 negatively regulates cytokine signaling.

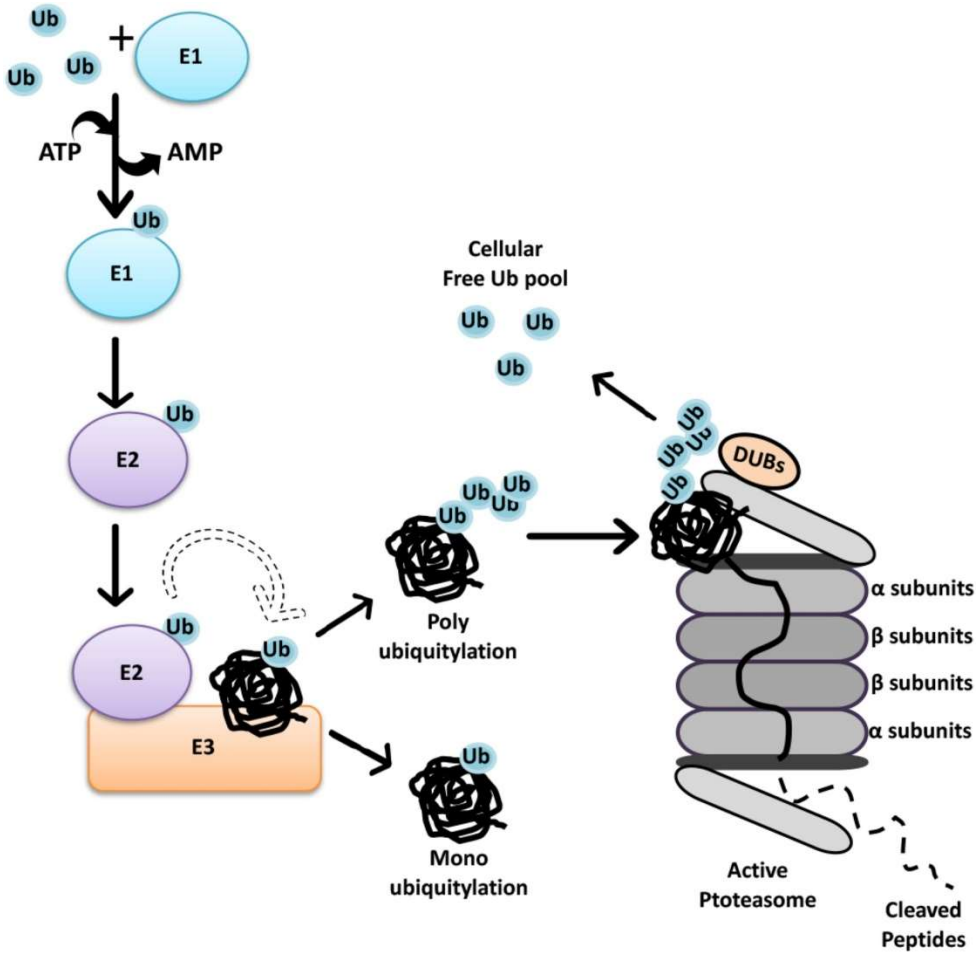


Figure 2. Schematic of how the ubiquitination process in carried out.[5]

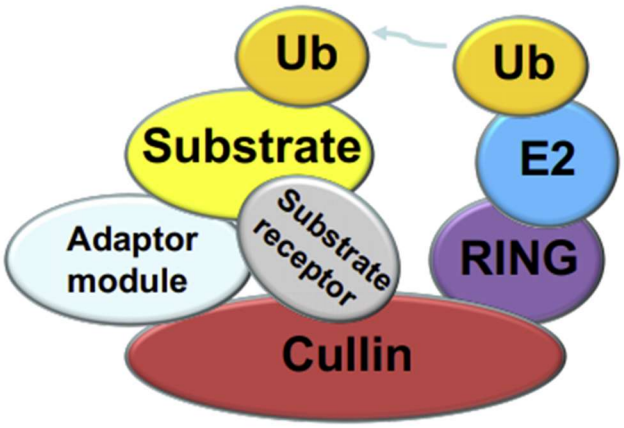


Figure 3. Cullin-RING ligase complex.

Some viruses, such has HIV-1, have taken advantage of hijacking this system. HIV-1 has been able to rewire the CRL complex by coding for a protein, viral infectivity factor (vif), that binds to Elongin B and C, where SOCS2 normally would.[7] This causes human enzyme APOBEC3G to bind to the ligase complex and be degraded by the proteosome. This enzyme is an antiviral protein that triggers mutations in viral mRNA transcripts, damaging the resulting viral proteins, therefore, reducing infection. However, by degrading APOBEC3G though hijacking the ubiquitination process, the virus can enhance its fitness and infectivity. This shows that this is a plausible biochemical pathway for viruses to hijack, making researchers investigate other viruses that may use this pathway to increase their viral fitness. Table 1 shows viruses that use this mechanism of increasing viral fitness.

Table 1. Viral proteins identified to hijack host CRLs.

Virus	Viral Protein	Mechanism
HIV-1	Viral infectivity factor (vif)	Recruits formation of host E3 ubiquitin ligase complex to induce degradation of APOBEC3G
Hepatitis B	HBx	Causes genome instability through unknown mechanism
Respiratory Syncytial Virus (RSV)	NS1	Degrades host STAT proteins, disables host interferon response
Human Papillomavirus	HPV LANA	Generation of E3 ligase complex to degrade p53
Adenovirus	E1B55K/E4orf6	Generation of E3 ligase complex to degrade p53
Epstein-Barr virus	BZLF/BPFL1	Inactivates CRLs preventing the degradation of host cell cycle and DNA damage regulators

Through sequence comparison, it was found that NS5A, a protein found in Hepatitis C virus, has a similar sequence to that of the SOCS-box binding domain (Figure 4). Since, the sequences are conserved in certain areas, this may cause NS5A to also bind to Elongin B and C as SOCS2 or vif does. Using the Protein Data base, the structural alignment of NS5A and the SOCS2-box can be generated (Figure 5). Using this image, the root-mean square deviation (RMSD) can be calculated. The RMSD is the measure of average distances between the atoms of the two superimposed proteins; the lower the RMSD, the better the superposition of the two molecules is. For NS5A and the SOCS box, it was calculated by Madison Muehl, using PyMOL, to be 9.395 Å. For comparison, the RMSD value for Vif and the SOCS box is 7.466 Å. This information leads the hypothesis of this thesis: that the Hepatitis C virus manipulates the process of ubiquitination, like that of the vif protein of HIV-1.

```
SOCS_box  ---------------APSLQHLCR-----------------------------------LT  11
NS5A      MFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIVGPRTCSNTWHGTFPINAYT  60
                          :*  *                                       *
SOCS_box  --------INKCTGAIWGLPLPT-------------------------------------  26
NS5A      TGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVPAPEFFTEVD  120
                    : : *:* :
SOCS_box  ----------------------RLKDYLEE-------------                  34
NS5A      GVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPCEPEPDVAV                  164
                                  *::**
```

Figure 4. NS5A sequence compared to the SOCS-box binding domain. (*=conserved)

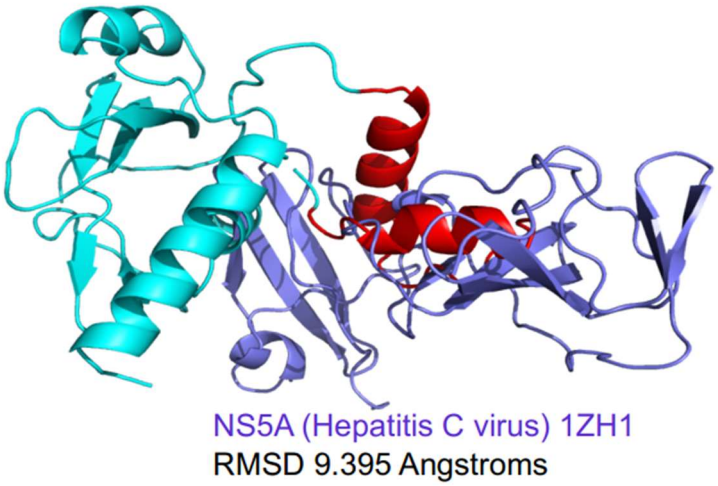


Figure 5. Structural alignment of NS5A (purple) and SOCS2-box (red). The RMSD value was calculated by Madison Muehl using PyMOL.

To test this hypothesis, viral chimeras were made to test protein interactions *in vitro* in *Escherichia coli* model organisms. This was done through co-expression of viral (NS5A) and host (Elongin B and C) proteins. The interactions were tested through immobilized metal affinity chromatography. By doing this, we should be able to see if NS5A is interacting with Elongin B and C, and if this is a viable mechanism for viral infectivity of HCV. If it is, future steps can be taken to evaluate possible therapeutics that inhibit this interaction.

Materials and Methods

1. Cell Biology Techniques

1.1 Liquid Media Creation

Luria-Bertani (LB) liquid media was used throughout this experimentation. This was made using tryptone, yeast extract, sodium chloride and water. A standard measurement of 1 g tryptone, 1 g yeast extract and 0.5 g sodium chloride per 100 mL of media was used. Agar was added to this mixture when plate preparation was required. The mixture was divided between baffled flasks. These were put into the autoclave on the "liquid less than 300 mL" cycle and were taken out when finished.

1.2 Bacterial Plating

Bacterial plating was done on LB plates with 50 µg/mL of ampicillin and/or 50 ug/mL streptomycin sulfate. Plating was done through toothpicks that were sterilized via autoclave and Bunsen burner. The plates were put in an incubator overnight at 37°C.

1.3 Overnight Cultures

Overnight cultures were created by pipetting LB media into 15 mL culture tubes. A colony was then added to the tube via toothpick sterilized by the autoclave and Bunsen burner. Ampicillin and/or streptomycin were added at concentrations of 50 µg/mL. The culture tubes were left on the shaker overnight at 37°C. Afterwards, the cell growth was measured through a spectrophotometer reading at 600 nm.

1.4 Glycerol stock creation

Glycerol stocks were made with equal volumes of 50% glycerol and overnight culture. The stocks were then frozen at -80°C to be utilized for bacterial plating.

2. Plasmid Construction

2.1 Test digests of plasmids

For all digestions, a solution of the target plasmid, specific restriction enzymes (BglI, NdeI, and/or Xho1), 10x FastDigest buffer (ThermoFisher), and nuclease free water was created. This solution was spun down in the centrifuge at 6,000 rpm for one minute. It was then put in the thermocycler at 37°C for at least 30 minutes. Digested products were analyzed by agarose gel electrophoresis.

2.2 PCR, Digest, and Purification of Inserts

PCR reactions were done to amplify the Elongin C, NS5A, and SOCS2 genes. The primers needed to include Nde1 and BglII cut sites for NS5A and SOCS2 and BglII and Xho1 cut sites for Elongin C. The primers were diluted to 10 µM with nuclease free water. PCR tubes were prepared with the desired template DNA, each desired primer (forward and backward), 2x PCR Master Mix (ThermoFisher), and nuclease free water, to create a 50 µL solution. These were run on the thermocycler with an annealing temperature of 58°C, denaturing temperature of 95°C, and extension temperature of 72°C, with a 30 second extension on each cycle.

Table 2. Primer sequences that were designed for the chimera constructs. The restriction sites are underlined.

Insert	Forward Primer	Reverse Primer
SOCS2	catccatatgggccatcaccatcaccatcacggc atgaccctgcggtgccttg	gatgagatctgccctgaaaatacaggttttcgcct acctggaatttatattc
NS5A	catccatatgggccatcaccatcaccatcacggc atgccgttcttttcttgtcag	gatgagatctgccctgaaaatacaggttttcgcc aaccgcaacatccggttcg
Elongin C	catcagatctatgggatatgtcaaattg	gatgctcgagttaacaatctaagaagttcgcag

Digestion of the PCR products was also done, however, in this process, the digests were left in the thermocycler for two hours instead of 30 minutes. After this, the products were taken out of the thermocycler and treated with of alkaline phosphatase, Fast AP from ThermoFisher, to cleave the phosphate ends of the inserts. They were put back in the thermocycler at 37°C for 10 minutes and purified using the Promega Wizard SV Gel and PCR clean-up system. The eluted DNA concentration was checked with the Nanodrop at 260 nm.

2.3 Ligation

The volumes of insert and vector needed to obtain a 3:1 ratio of insert to vector were calculated using NEBio Ligation Calculator for ligation. These volumes helped set up control and experimental reactions in PCR tubes. Both reactions were 10 µL and were set up to not exceed 200 ng of vector DNA. The experimental solution consisted of diluted insert, vector, T4 ligase (NEB), 5x buffer, and nuclease free water. The control solution contained everything minus the insert. These reactions were incubated at 16°C overnight.

2.4 Transformation into bacterial cell lines

Transformation was accomplished by thawing a tube of -dam/-dcm cells (NEB). The cell solution was mixed gently and 50 µL of cells were pipetted into a 15 mL culture tube. A concentration of 1 pg to 100 ng of plasmid DNA was added the cell mixture and the tube was flicked to mix the cells and DNA. The mixture was placed on ice for 30 minutes and then heat shocked at exactly 42°C for 30 seconds. The mixture was again placed on ice for five minutes. Next, 950 µL of LB media was pipetted into the mixture, which was placed at 37°C for one hour on the shaker. The cells were mixed then 200 µL were plated on LB media with 50 µg/mL ampicillin and grown overnight. These were used for DNA plasmid amplification and extraction.

Plasmids were also transformed into BL21 DE3 cells (ThermoFisher). This was done by thawing out 50 µL of ThermoFisher One Shot cells and pipetting 2 µL of each plasmid into separate cell tubes. The procedure followed that of transforming plasmids into the -dam/-dcm cells. However, the second ice bath incubation is one minute rather than five. Also, 250 µL of pre-warmed LB media was added to the mixture instead of 950 µL. These were used for protein interaction testing.

2.5 Agarose Gel Electrophoresis

Gels were created using 60 mL of 1x TAE buffer and 1% agar. This mixture was heated in the microwave until the agar was dissolved. Then, after the mixture had cooled a bit, 1x Syber Safe DNA gel stain was added, and the gel was poured into the gel mold. This was let sit until the gel was opaque. The gel was then inserted into the gel apparatus and covered with 1x TAE buffer. The samples were spiked with SDS-containing loading dye and loaded into the gel with a Promega 1kb DNA ladder. The gel

was run at 80 V until the samples were out of the well, then bumped up to 100 V until complete.

3. Protein Interaction Analysis

3.1 Protein Expression

Utilizing a 2-4 mL overnight culture, a final absorbance of 0.05 was desired for optimal protein expression in the targeted growth flask. To reach this optical density, an appropriate amount of the overnight culture was added to LB media in a targeted baffled flask (50-200 mL). Ampicillin and/or streptomycin were again added to the media at a concentration of 50 µg/mL. The baffled flasks were then put in the shaker and the optical density (OD) was periodically observed until it reached 0.5-0.6 OD. When this was achieved, the baffled flasks were taken out of the shaker and cooled on ice to 30°C. An aliquot of 1 mL was transferred to a microcentrifuge tube and put on ice, this represented time zero for expression analysis. After the contents of the baffled flask had cooled, 1mM of IPTG was added to the flask to induce protein expression, by activating the lac-operon. The aliquot for time zero was spun in the centrifuge for five minutes at 6,000 rpm at 18℃ to form a pellet at the bottom of the tube. The supernatant was discarded so that only the cell pellet remained in the tube, which was labeled and put in the -20°C freezer. Other time points were taken periodically throughout the day utilizing the same technique listed above for the pre-induction sample. The optical densities for all the time point samples were also recorded. Expression ran 4-16 hours after induction. Cell pellets from expression were saved at -80°C after recording the mass.

3.2 Immobilized Metal Affinity Chromatography

10 mL of lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 10 mM imidazole, 5 mM B-mercaptoethanol, and protease inhibitor tablet) were added per gram of bacterial cell pellet. Lysozyme was added at a concentration of 2 mg/mL and the mixture was incubated for 20 minutes. The mixture was sonicated to lyse the cells at 30% amplitude for five 10 second intervals with 50 seconds in between each pulse. The cells were then spun at 6,000 rpm for 20 minutes and filtered using a 0.45 µL filter and syringe. This filtrate was added to a Ni-NTA spin column (ThermoFisher) in three 5 mL intervals, each being incubated for 10 minutes before spinning for two minutes at 700 x g. The supernatant flow through was collected each time and then the column was washed three times with 2 mL of lysis buffer, spinning for two minutes at 700 x g. The column was then eluted 3 times with 2 mL of elution buffer (each time) (50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 250 mM imidazole, and 5 mM B-mercaptoethanol), spinning for two minutes at 700 x g. The concentrations of the elution fractions were measured by nanodrop at 280 nm.

3.3 SDS-PAGE Analysis

Samples from expression or purification were prepared with 2x SDS buffer (BioRad). These samples were boiled in a hot water bath for four minutes before loading into the 4-20% Mini-PROTEAN TGX pre-cast gel (BioRad). The gel was loaded into the Mini-PROTEAN Tetra Vertical Electrophoresis gel apparatus with, or without, a buffer dam. Tris-Glycine-SDS Running Buffer (1x) was added. The wells were washed out with loading buffer and the samples were loaded into the wells along with a protein ladder (Precision Plus Dual Color Protein Standard-BioRad). The gel was run at 200 V until the

bands were a few mm from the bottom (approximately 20 to 30 minutes). The gel was removed from the apparatus and plastic casing and put in Coomassie stain. This was put on the shaker for 1 to 2 hours at a speed of 84. After, the gel was put into destain overnight, or until there was enough contrast between the bands and the gel. The gel was then placed in water to let the stains set.

Results

Table 3. All the plasmids and strains created throughout the experimentation and what each is composed of.

Name	Composition	Key
Plasmid AN	PCDFDuet-1_(1)NS5A	(1) = MCS1
Plasmid AS	PCDFDuet-1_(1)SOCS2	(2) = MCS2
Plasmid B	PETDuet-1_(1)EB_(2)EC	EB = Elongin B
Plasmid CN	PETDuet-1_(1)EB_(2)6His-NS5A	EC = Elongin C
Plasmid CS	PET-Duet-1_(1)EB_(2)6His-SOCS2	
Plasmid DN	PETDuet-1_(1)EB_(2)6His-NS5A-TEV-EC	
Plasmid DS	PETDuet-1_(1)EB_(2)6His-SOCS2-TEV-EC	
Strain ANB	Glycerol stocks containing Plasmids AN and B	
Strain ASB	Glycerol stock containing Plasmids AS and B	
Strain CN	-dam/-dcm cells containing Plasmid CN	
Strain CS	-dam/-dcm cells containing Plasmid CS	
Strain DN1	-dam/-dcm cells containing Plasmid DN	
Strain DS1	-dam/-dcm cells containing Plasmid DS	
Strain DN2	BL21 DE3 cells containing Plasmid DN	
Strain DS2	BL21 DE3 cell containing Plasmid DS	

This table was created to help simplify the different strains and plasmids eluted to in the results section of this thesis.

1. Plasmid Construction

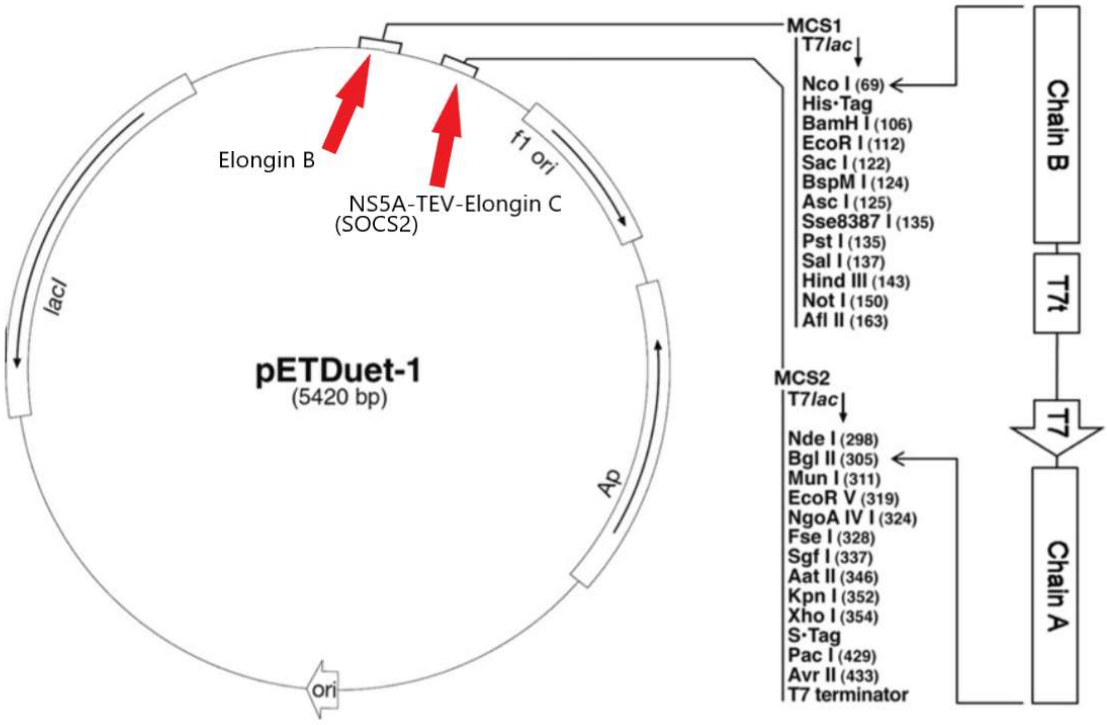


Figure 6. pETDuet-1 plasmids showing insertion of NS5A (SOCS2), Elongin B, Elongin C, and TEV. This represents the template for Plasmid DN and DS.

The figure above (Figure 8) shows the viral chimera that was created. Chimeras were created through the molecular cloning techniques, such as digestion, purification of inserts, and ligation. Several successive digestions were done to end with the result of Plasmids DN and DS. Plasmid B was the first plasmid digested to insert the genes for NS5A or SOCS2 into multiple cloning site 2 (MCS2). Both restriction enzymes used to insert these genes (Table 3) includes a 6-His tag for labelling the proteins of interest. This process created Plasmids CN and CS.

Plasmids CS and CN were further digested to insert the gene for Elongin C in MCS2. The Elongin C gene was linked to the NS5A and SOCS2 genes using a Tobacco Etch Virus (TEV) linker. This was necessary so both proteins would be expressed in the second cloning sight.

The TEV linker sequence was included in the created primers (Table 3). This insertion and ligation resulted in Plasmids DN and DS.

2. Protein Expression

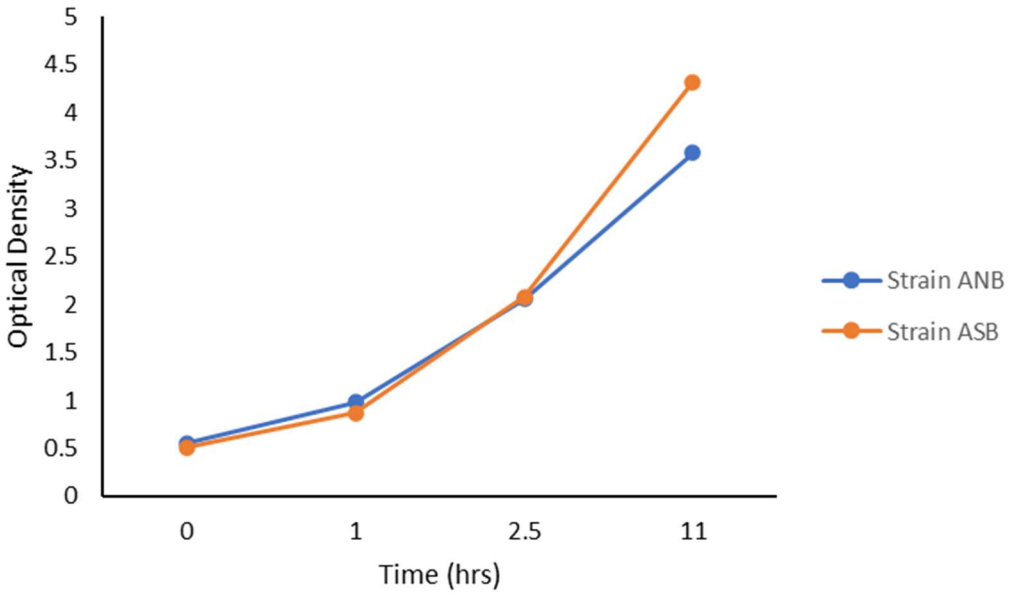
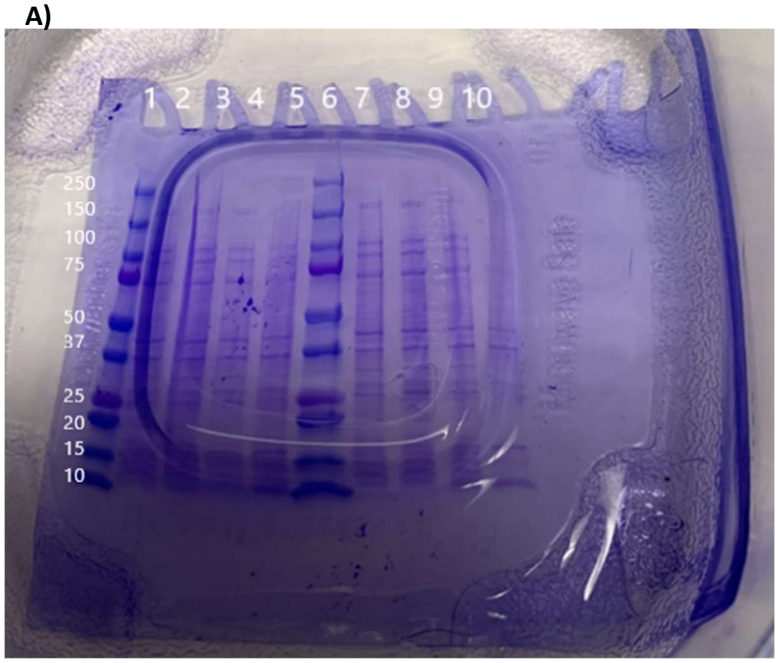


Figure 7. Graph of times versus optical density of Strain ASB (orange) and Strain ANB (blue) for culture growths.

Strains ANB and ASB were plated on LB media plates overnight. Overnight cultures were then grown overnight from the cell cultures on the plates. The protein expression protocol from section 3.1 was then followed. Figure 6 shows Strains ANB and ASB were able to grow and express proteins. The time points that were obtained from ASB were analyzed through SDS-PAGE, shown in Figure 7. Here, bands occur at around 11 and 13 kDa for Elongin C and B respectively. Therefore, the cells are expressing these proteins. However, in this figure, which looks at the proteins from Strain ASB, there is no band at or around 22 kDa, meaning SOCS2 was not being expressed in these cells. This was also true for NS5A.

A)



B)

1	MW ladder
2	T0 (10 µL)
3	T1 (10 µL)
4	T2.5 (10 µL)
5	T11 (10 µL)
6	MW ladder
7	T0 (5 µL)
8	T1 (5 µL)
9	T2.5 (5 µL)
10	T11 (5 µL)

Figure 8. SDS PAGE gel of proteins expression time points from Strain ASB (A), with corresponding lane assignments (B).

After construction of Plasmids DN and DS, these plasmids were transformed into -dam/-dcm cells for expression trials. Figure 9 shows that the bacteria were able to selectively grow in overnight cultures before and after induction with IPTG. This allowed for subsequent analysis of the proteins being expressed.

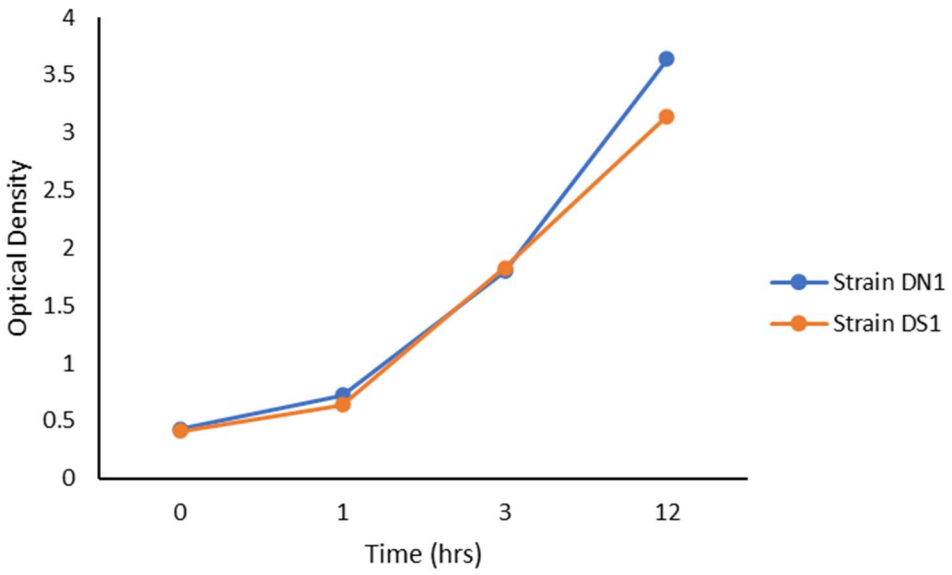


Figure 9. Graph of times versus optical density of Strain DS1 (orange) and Strain DN1 (blue) for culture growths.

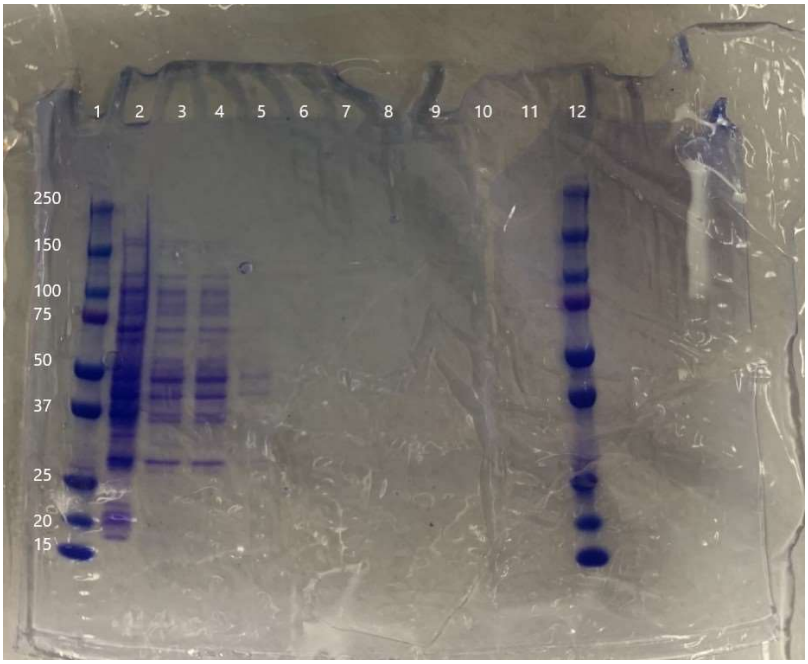
3. Immobilized Metal Affinity Chromatography (IMAC)

After the protein expression trials, Plasmids DN and DS were transformed into BL21 DE3 cells to test protein interactions, creating Strains DN2 and DS2. This was done through immobilized metal affinity chromatography (IMAC). It is expected that the elution fractions would contain the proteins in question (SOCS2/NS5A, Elongin B, and Elongin C). These proteins would then show up at on an SDS-PAGE gel at their corresponding molecular weights, seen in Tables 4 and 5. The combined molecular weights of SOCS2 and Elongin C and NS5A and Elongin C are included due to their linkage through the TEV linker. However, no proteins were visualized in the elution wells, meaning to proteins were present in our elution fractions, for both our control (Strain DS2) and our experimental (Strain DN2). This makes it impossible to draw a conclusion regarding viral-host protein interactions of NS5A, Elongin B, and Elongin C.

Table 4. Predicted molecular weights (MW) in kilodaltons (kDa) for SOCS2, Elongin B, and Elongin C. The combined molecular weights show where the band for SOCS2 and Elongin C would occur due to the TEV linker.

Protein	MW (kDa)	Combined MW (kDa)
SOCS2	22.1	33
Elongin C	10.9	
Elongin B	13.2	13.2

A)



B)

1	MW ladder
2	Insoluble pellet
3	Flow through 1
4	Flow through 2
5	Flow through 3
6	Wash 1
7	Wash 2
8	Wash 3
9	Elution 1
10	Elution 2
11	Elution 3
12	MW ladder

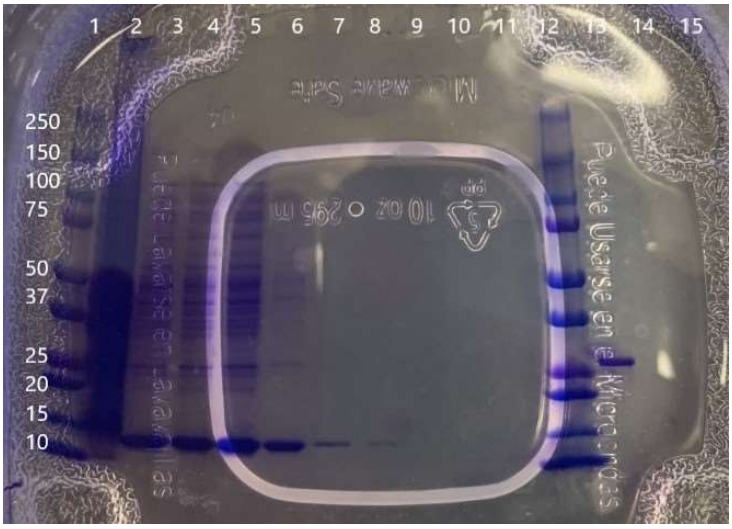
Figure 10. SDS-PAGE gel of proteins from Strain DS2 (A) with corresponding lane assignments (B).

In the experimental, TEV protease was introduced to the second elution fraction. This was done to test the effects of removing the TEV linker on the appearance of protein in the SDS-PAGE analysis. However, even with the cut TEV linker, no bands showed up on the SDS-PAGE gel.

Table 5. Predicted molecular weights (MW) in kilodaltons (kDa) for NS5A, Elongin B, and Elongin C. The combined molecular weights show where the band for NS5A and Elongin C would occur due to the TEV linker.

Protein	MW (kDa)	Combined MW (kDa)
NS5A	18.1	29
Elongin C	10.9	
Elongin B	13.2	13.2

A)



B)

1	MW ladder
2	Insoluble pellet
3	Flow through 1
4	Flow through 2
5	Flow through 3
6	Wash 1
7	Wash 2
8	Wash 3
9	Elution 1
10	Elution 2
11	Elution 3
12	MW ladder
13	TEV
14	Elution 2 + TEV protease

Figure 11. SDS-PAGE gel of proteins from Strain DN2 (A) with corresponding lane assignments (B).

Discussion

Through this process, several new plasmids were successfully created through molecular cloning. The pETDuet-1 plasmid was selected to make use of its two cloning sites. This allowed for more equal protein expression between the proteins coded for at each site. Elongin C was linked to NS5A or SOCS2 through a TEV linker, so that both proteins could be coded for in one cloning site.

Several new bacterial strains were also created and successfully grown in cultures. Cell growth and protein expression were able to be optimized, showing the strains were able to express the proteins coded for through the genes inserted in their plasmids. The data seen in Figures 7 and 9, shows that the proteins were not toxic to the cells during expression. This is because we do not see any cell death. The plasmids also weren't cytostatic, meaning they didn't limit or inhibit growth of the cells. This is because over time we see an increased about of growth at an exponential rate. There wasn't a problem with protein expression, and this could be ruled out as something that potentially went wrong in the research. The graphs show that over time, the optical density increased, meaning the cells were growing and expressing proteins.

It was hypothesized that, due to its similar sequence to the SOCS-box binding domain, NS5A would bind to host proteins Elongin B and Elongin C (like SOCS2 does), disrupting the endogenous ubiquitination process. Through immobilized metal affinity chromatography (IMAC), the 6-His tag on NS5A should bind to the nickel resin of the Ni-NTA columns, causing Elongin B and Elongin C to stay on the resin as well by binding to NS5A. This would cause all three proteins to show up in the elution fractions when analyzed by SDS-PAGE. However, no

bands showed up in the elution samples that were loaded into the SDS-PAGE in both the control and the experimental.

Upon analysis of the SDS-PAGE results, no bands were visual in any of the elution or wash fractions, making the evidence inconclusive. There was found to be proteins in the insoluble pellet and the flow through, characteristic of contaminant proteins without the 6-His tag. This was true for both Strain DS2 and DN2. It is known that SOCS2 binds to Elongin B and Elongin C, so there should have been bands showing up at least on this gel.[8] The fact that there wasn't, poses three possible conclusions: 1) this is not a successful method in showing interaction between these proteins, 2) our proteins are not soluble in our elution buffer, or 3) the proteins were too dilute to show up on the gel.

The fact that this methodology didn't show interactions between the expressed proteins, or any proteins at all, leads us to believe that there is a problem within the methodology. In theory, the control should work. Therefore, maybe interactions between these proteins cannot be shown using IMAC. There is also a possibility that the TEV linker may be restricting the protein from binding to the resin or interacting with the target proteins. The TEV linker could possibly bind the 6-His tag, preventing it from binding to the resin in the column, causing our target proteins to be washed through the column during the washes or the flow through.

However, there may be other smaller problems within the experimental designs, like listed above. If our proteins are not soluble in the elution buffer, it will not pull them down into the elution fractions, meaning the proteins may still be stuck to the resin. This would explain

why we don't see any proteins, even in the control. This could be further evaluated by boiling the resin with loading buffer for SDS-PAGE analysis.

Lastly, the concentration of proteins may be too dilute to show up on the gel. This could also explain why there were no proteins shown in the control. The concentrations may be too low to be picked up by the Coomassie stain, but still high enough to induce activity in expressions. This could be fixed by using a stain that can pick up lower concentrations, such as silver nitrate. This could also be resolved by Western blotting to isolate the 6-His tag on NS5A/SOCS2. However, this would not show any other proteins.

Conclusion

In conclusion of these experimental procedures, it cannot be said for certain whether or not this is a plausible biomechanistic pathway used to increase viral fitness of HCV. Further studies should be done to correct problems that may have occurred in the studies above. The fact that the methodology didn't work with the control, prevents ruling out these viral-host protein interactions as possible mechanisms of viral infectivity for HCV. In order to rule it out, clear evidence needs to show interaction within the control and not the experimental variable.

Further studies may try a protein assay, rather than IMAC to test for viral-host protein interactions. This could include *in vivo* assays, such as yeast two-hybrid assays. However, adjustments can be made to what was done here to further test possible interactions, before completely starting over with new methodology. To adjust this study, new elution buffers may be constructed and used to target the possible insolubility problem with the elution buffer. Also, here, only approximately 15 mL of cells were analyzed. This may need to be done on a bigger scale to have a higher concentration of proteins assessed, or overexpression of proteins may need to be triggered.

In light of these events, new plasmids and bacterial strains were successfully constructed and expressed, meaning that we have useful genomes to continue the study of NS5A protein interactions. This is an important first step in the research that can be done to further test the hypothesis that NS5A binds to Elongin B and C to hijack host cell ubiquitination.

References

1. Hoofnagle JH. 2002. Course and outcome of hepatitis C. Hepatology. 36(S1):S21–S29. doi:10.1002/hep.1840360704

2. Hepatitis C - FAQs, Statistics, Data, & Guidelines | CDC. [accessed 2021 Apr 26]. https://www.cdc.gov/hepatitis/hcv/index.htm

3. Chevaliez S, Pawlotsky JM. 2006. Hepatitis C Viruses: Genomes and Molecular Biology. 1st ed. Norfolk, (UK): Horizons Bioscience. Chapter 1. HCV Genome and Life Cycle; p. 5-48. ISBN: 1904933203

4. Elgharably A, Gomaa AI, Crossey MME, Norsworthy PJ, Waked I, Taylor-Robinson SD. 2017. Hepatitis C in Egypt - past, present, and future. Int J Med. 10:1–6. doi:10.2147/IJGM.S119301

5. Kocaturk NM, Gozuacik D. 2018. Crosstalk between mammalian autophagy and the ubiquitin-proteasome system. Front Cell Dev Bio. 6(OCT):128. doi:10.3389/fcell.2018.00128

6. Genschik P, Sumara I, Lechner E. 2013. The emerging family of CULLIN3-RING ubiquitin ligases (CRL3s): Cellular functions and disease implications. EMBO Journal. 32(17):2307–2320. doi:10.1038/emboj.2013.173

7. Rajsbaum R, García-Sastre A. 2013. Viral evasion mechanisms of early antiviral responses involving regulation of ubiquitin pathways. Trends Microbiol. 21(8):421–429. doi:10.1016/j.tim.2013.06.006

8. Bulatov E, Ciulli A. Targeting Cullin-RING E3 ubiquitin ligases for drug discovery: structure, assembly and small-molecule modulation. Biochem J. 2015;467:365–386.

doi:10.1042/BJ20141450