A NOVEL METHOD FOR VIRAL PROTEIN TRACKING IN HOST CELLS

by © Jacqueline Patricia Barry

A thesis submitted to the School of Graduate studies in partial fulfillment of the requirements for the degree of

Master of Science / Immunology and Infectious Diseases / Faculty of Medicine

Memorial University

October 2017

St. John’s Newfoundland and Labrador

ABSTRACT

Hepatitis C Virus (HCV) infects approximately 70 million people worldwide and chronic infection can lead to liver cirrhosis and hepatocellular carcinoma. HCV contains a perplexing protein that has a number of proposed functions. This protein, termed p7, is essential for virus infectivity *in vivo*, however its function is a matter of controversy. Research into the function of p7 has been limited because there are no reliable antibodies available for the visualization of this protein. The goal of this project was to establish a system that utilizes fluorescent unnatural amino acids in order to label p7 within the context of a replicating virus. Strategically placed mutations within the viral p7 protein were selected to test their amenability to incorporation of an unnatural amino acid. In order to optimize the incorporation system, plasmids containing mutations in the viral core protein were synthesized for further screening of positions tolerable to substitution. Ultimately, we were successful in incorporating a fluorescent unnatural amino acid into the viral core protein in a single protein expression system. If we can successfully transfer this system back into the context of a replicating virus, this technique could be used to facilitate the study of viral proteins in HCV and other viruses.

ACKNOWLEDGEMENTS

First, I would like to thank my amazing supervisor, Dr. Rodney Russell. Without you, none of this would have been possible, and I am very grateful for all of the support and guidance you have given me over the last two years. You took the role of mentor above and beyond what was to be expected, and I couldn't have done this without you. Thank you so much for helping me throughout this journey and helping pave the way for my future. I promise that I will still be coming to your office for advice long after I have graduated from this program.

I would also like to thank Dr. John Pezacki. As our collaborator at the University of Ottawa, you really helped us get this project off the ground. Thank you for always believing in me and accepting me into your lab. My time with you and the rest of the Pezacki lab team is treasured. With that in mind, I have to send a big thank you to Dr. Megan Powdrill. Megan, you took me under your wing, and I am so grateful for the time I spent working with you. Thank you for always being there for troubleshooting questions and skype meetings to discuss pitfalls and successes!

Next, I must thank my supervisory committee, Dr. Michael Grant and Dr. Mani Larijani. Thank you both for your support and encouragement over the last two years. Your input during committee meetings provided insight into various applications for my project and gave me the confidence I needed to present my results to other audiences.

To my lab family - all the amazing people who I have worked with since starting in Rod's lab. Hassan, Kylie, Ahmed, Bridgette, and Rebecca, we've had some pretty amazing times over the last two years. Thank you all for your friendship and love; I have valued each of our relationships. I must give a special shout-out to Hassan. You were the most amazing role model I could have. Each conversation we had left me in awe of your intelligence and humility. I have said it before and I will say it again, your future students will be very lucky to have you as a professor.

To all my friends in and out of the lab, I couldn't have done this without you. Emilie, thank you for everything you have done for me, I really don't know how I would have survived this journey without you. You have been my best friend inside and out of the lab and have been there for everything; laughter, tears, frustration, excitement and just our everyday adventures. I will miss leaving notes for you, going for treats and casual yoga in the office. I promise I will always keep a spare pair of pants on hand though, because you never know when someone might need them. And as always, Ella… (and maybe me) still need regular visits from "auntie Em".

Last, but certainly not least, a huge thank you to my family. First, to my Newfoundland family, Heather, Vaughan, Sarah, Mike, Andrew, and Keely, thank you for being my home away from home. Adam, you have been my rock. Thank you for always being there to listen to my excitement when things went well, and my annoyance when things

weren't great. Coming home to you every night makes any day better and I am so grateful to have you in my life. Finally, to my amazing family, John, Mom and Dad, I don't know where I would be without you. John, you really are the best baby brother a girl could ask for. Mom and Dad, your unwavering support and encouragement is what got me through this degree. Dad, thank you for always accepting my million phone calls a day, and getting me back on track when I was definitely procrastinating. Your enthusiasm always keeps me motivated. Mom, thank you for coming to visit multiple times throughout the last two years. I always love having you here, even if I might not always show it. I really can't put into words just how amazing you both are. I love you both very much!

Table of Contents

ABSTRACT ..... **ii**

ACKNOWLEDGEMENTS ..... **iii**

LIST OF FIGURES ..... **vii**

LIST OF TABLES ..... **ix**

LIST OF ABBREVIATIONS AND DEFINITIONS ..... **x**

Chapter 1: Introduction **1**

1.1 Overview **1**

1.2 Discovery of the Virus **1**

1.3 Natural History of Infection **3**

1.4 Viral Genome **5**

1.5 Life cycle **5**

1.5.1 Virion Structure, Receptor Binding, Entry and Fusion 5

1.5.2 Translation 10

1.5.3 RNA replication 10

1.5.4 Assembly and Release 11

1.6 History of HCV Therapies **12**

1.6.1 Interferon and Ribavirin 12

1.6.2 Direct-Acting Antivirals 13

1.7 p7 protein **14**

1.7.1 Viral Protein Tagging Methods 19

1.7.2 Unnatural Amino Acids to Visualize Proteins 20

1.8 Translation **22**

1.9 UnAAs **24**

1.10 Project design and hypothesis **27**

1.11 Project aims **28**

I. Screening positions for unAA incorporation in the HCV p7 protein. 28

II. Screening mutants for an amenable position in HCV core protein. 29

Chapter 2: Methodology **29**

2.1 Primer Design **29**

2.2 In-vitro Site-directed mutagenesis **31**

2.3 Transformation **31**

2.4 Miniprep DNA Purification **31**

2.5 *PVUII* Digestion 32
2.6 Sequencing of Plasmids 32
2.7 Maxiprep DNA Purification 32
2.8 Plasmid Linearization 32
2.9 Cell Culture 33
2.10 DNA Transfection 34
2.11 Transcription 34
2.12 RNA Transfection 34
2.13 Determination of Infectious Titre 35
2.14 Immunofluorescence Staining 35
2.15 G418 Treatment 36
2.16 G418 Titration 37
2.17 Transfection of Core Mutants 37

Chapter 3: Results 38
3.1 Mutant Selection: 38
3.2 The Effect of AzF Incorporation on Viral Infection: 39
3.3 The Effect of AzF on Virus Production: 42
3.4 Rationale for Using Anap Instead of AzF 42
3.5 Anap incorporation in Huh-7.5 cells 44
3.6 Rationale for scale down of Anap experiments 46
3.7 Anap was detected in transfected cells stained for core 46
3.8 Treatment of DNA transfected cells with G418 47
3.9 G418 Titration Experiment 49
3.10 Huh-7.5 cells transfected with pAnap and $JFH1_T$ 51
3.11 DNA transfected Huh-7.5 cells can be transfected and infected with $JFH1_T$ 53
3.12 Testing incorporation of Anap into the HCV core protein 55
3.13 Core Mutant Selection 57
3.14 Incorporation of Anap into Core Mutants 58
3.15 Core Positive Cells Detected 63

Chapter 4: Discussion 63
4.1 Future Applications 71

Chapter 5: References 73

LIST OF FIGURES

Figure 1.1	Flow Chart of Viral Infection and Related Outcomes	4
Figure 1.2	HCV Genome and Proteins	6
Figure 1.3	HCV Structure	7
Figure 1.4	Lifecycle of HCV	10
Figure 1.5	HCV p7 protein Monomer and Hexamer	17
Figure 1.6	Structure of p-Azido-$_{L}$-phenylalanine (AzF) (left) and 3-(6acetylnaphthalen-2-ylamino)-2-aminopropanoic acid (Anap) (right)	21
Figure 1.7	Unnatural Amino Acid Incorporation	25
Figure 3.1	DNA gel of maxi prepared samples following *PvuII* digestion	39
Figure 3.2:	The Effect of AzF Incorporation on Virus Infection	41
Figure 3.3	The effect of AzF on Virus Production	42
Figure 3.4	Anap fluorescence was visible in transfected Huh-7.5 cells	44
Figure 3.5	Anap fluorescence was detected in transfected Huh-7.5 cells stained for core	47
Figure 3.6	Treatment of transfected Huh-7.5 cells with G418	49
Figure 3.7	G418 Titration	51

Figure 3.8 DNA transfected Huh-7.5 cells, transfected with WT JFH1T + ANAP 53

Figure 3.9 DNA transfected Huh-7.5 cells transfected and infected with JFH1T 55

Figure 3.10a Incorporation of Anap into core mutants: transfection controls 58

Figure 3.10b Incorporation of Anap into core mutants: tryptophan mutants 60

Figure 3.10c Incorporation of Anap into core mutants: tyrosine mutants 61

Figure 3.11 Relationship between Anap and Detected HCV core staining 63

LIST OF TABLES

Table 2.1	List of all primer sequences within p7	32

LIST OF ABBREVIATIONS AND DEFINITIONS

ΔGDD	Negative cell culture control with NS5B active site removed
A	(Ala) Alanine
aa	Amino acid
AAV	Adeno-associated viral
Anap	3-(6-acetylnapthalen-2-ylamino)-2-aminopropanoic acid
AzF	p-Azido-$_L$-phenylalanine
BSA	Bovine serum albumin
CD81	Cluster of Differentiation 81
cDNA	Complementary deoxyribonucleic acid
CLDN-1	Claudin-1
cLDs	Cytosolic lipid droplets
cm	centimeter
CO_2	Carbon dioxide
DAAs	Direct-acting antivirals
DAPI	4',6-diamidino-2-phenylindole
dH_20	Deionized water
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
dsDNA	Double-stranded Deoxyribonucleic acid
dsRNA	Double-stranded Ribonucleic acid

E1	Envelope Protein 1
E2	Envelope Protein 2
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
EtOH	Ethanol
FBS	Fetal bovine serum
FDA	US Food and Drug Administration
ffu	Focus-forming units
G1a	Genotype 1 a
GFP	Green fluorescent protein
HA	Hemagglutinin antigen
HAV	Hepatitis A Virus
HBV	Hepatitis B Virus
HCV	Hepatitis C virus
HCVcc	HCV cell culture system
HIV	Human Immunodeficiency Virus
Huh-7.5	Human hepatoma 7.5
IF	Immunofluorescence
IFNα	Interferon-α
IRES	Internal ribosome entry site
IU	International units
JFH1	Japanese fulminant hepatitis 1

JFH1T	Japanese fulminant hepatitis 1 triple mutant
Kb	Kilobase
L (Leu)	Leucine
LD	Lipid droplet
LDL	Low density lipoproteins
LVP	Lipovarial
mL	milliliter
mg	milligram
mRNA	Messenger RNA
NaCl	Sodium chloride
NANBH	Non-A, non-B hepatitis
NCR	Noncoding region
ng	Nanogram
NS2	Non-structural protein 2
NS3	Non-structural protein 3
NS4A	Non-structural protein 4A
NS5A	Non-structural protein 5A
NS5B	Non-structural protein 5B
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
peg-IFNα	Pegylated-Interferon-α

PenStrep	Penicillin Streptomycin
PI	Protease inhibitors
PWID	People who inject drugs
RBV	Ribavirin
RdRp	RNA-dependent, RNA-polymerase
rpm	Revolutions per minute
SR-BI	Scavenger receptor B type-I
SVR	Sustained virological response
TAE	Tris base, acetic acid and EDTA buffer
TM1	Transmembrane domain 1
TM2	Transmembrane domain 2
tRNA	Transfer RNA
Trp	Tryptophan
µm	micrometer
µl	microliter
unAA	Unnatural amino acid
UAA	Ochre stop codon
UAG	Amber stop codon
UGA	Opal stop codon
UV	Ultraviolet
Vis	Visible
W (Trp)	Tryptophan

Chapter 1: Introduction

1.1 Overview

Hepatitis C virus (HCV) is a global health concern with an estimated 70 million HCV infected adults or 2.5% of the world population[1–7]. Prevalence ranges between 1-3% of the population in most countries, with a notable difference in Egypt, where there is over 20% prevalence due to negligence of sterility during parenteral antischistosomal therapy[8]. The prevalence of HCV in Canada is estimated to be 1.1% of the general adult population[9,10]. When compared to other infectious diseases, HCV infection causes more years of life lost in Canada due to complications associated with disease progression, [11]. Prevalence rates are highest amongst people who inject drugs (PWID), making intravenous drug use the largest risk factor for HCV transmission[12,13]. Given that HCV is a blood-borne virus, blood transfusions were a leading cause of transmission prior to current screening methods[8]. The role of sexual activity in transmission remains unclear and controversial[1,14].

1.2 Discovery of the Virus

In 1975, what is now known as HCV infection was first described as Non-A, Non-B hepatitis (NANBH); serological tests uncovered that many cases of parenterally-transmitted hepatitis were not due to Hepatitis A virus (HAV) or Hepatitis B virus (HBV)[15]. It would not be referred to as "Type C" until the infectious agent responsible for this form of hepatitis was identified. After many years of extensive research, a breakthrough came in 1989, when a group at the Chiron Corporation isolated the

etiological agent responsible for NANBH as a cDNA clone[15]. The group modified a standard cloning protocol developed to isolate DNA encoding unknown proteins. Modifying the protocol made it possible to isolate and characterize the unknown viral genome. Large volumes of infected chimpanzee plasma showing high infectious titre were ultracentrifuged to extract nucleic acids from the pelleted material. A denaturation step was included to allow both RNA and DNA to act as a template given that the nature of the viral genome was not known at this time. A cDNA library was derived from the isolated nucleic acids and inserted into the bacteriophage λgt11 to be expressed in *Escherichia coli*. Antisera was required for screening the cDNA library, so serum was collected from a NANBH-positive patient as a source of antiviral antibodies. Following immunoscreening of the cDNA library, which contained approximately 1 million clones, a single reactive clone termed 5-1-1, was found to be derived from the HCV genome. Further experiments on the clone confirmed it was not from the host genome and no homologous DNA sequences were found. Cloned cDNA hybridized to RNA extracted from infected chimpanzee sera, but not to RNA extracted from control, uninfected chimpanzees. HCV was confirmed as an RNA virus when hybridization signals were lost following treatment with ribonuclease but not with deoxyribonuclease. Further experiments also confirmed that it was positive sense, single-stranded and large in length (~10,000 nucleotides). The group postulated that HCV belonged to the *Flaviviridae* family of viruses due to sequence similarity. Based on this work, the etiological agent responsible for NANBH was identified as HCV[15].

HCV is divided into seven distinct genotypes (1-7) with each genotypic group containing various subtypes (a, b, c etc.)[2,16]. Despite patients being diagnosed with a specific genotype, multiple HCV quasispecies can be found within each individual patient due to the error-prone nature of the viral polymerase[17]. The estimated viral production rate of 10^{12} virions per day, and the estimated error rate of the polymerase, contribute to the number of mutations that can occur and could exist in every patient every day[18].

1.3 Natural History of Infection

HCV infection varies widely within the infected population, as progression from acute to chronic infection is influenced by many host and viral factors[11,19]. Generally, acute infection presents with little or no symptoms, making it very difficult to diagnose, increasing the likelihood of disease progression[20–23]. In the absence of treatment, approximately 15-25% of cases of acute infection will resolve spontaneously, while 75-85% of infected individuals will develop chronic HCV infection (Fig. 1.1)[24]. Despite the high prevalence of cases that progress to chronic infection recent advances in pharmaceutical interventions for treatment of HCV-infected individuals have made clearance of the virus, or sustained virological response (SVR), more common[20,25,26]. Although treatments are available and effective, progression to chronic HCV infection can still lead to a plethora of complications, including the development of liver cirrhosis, fibrosis and an increased risk of hepatocellular carcinoma[27,28]. The complications associated with chronic infection are what makes HCV infection a leading cause of liver-related deaths worldwide[11].

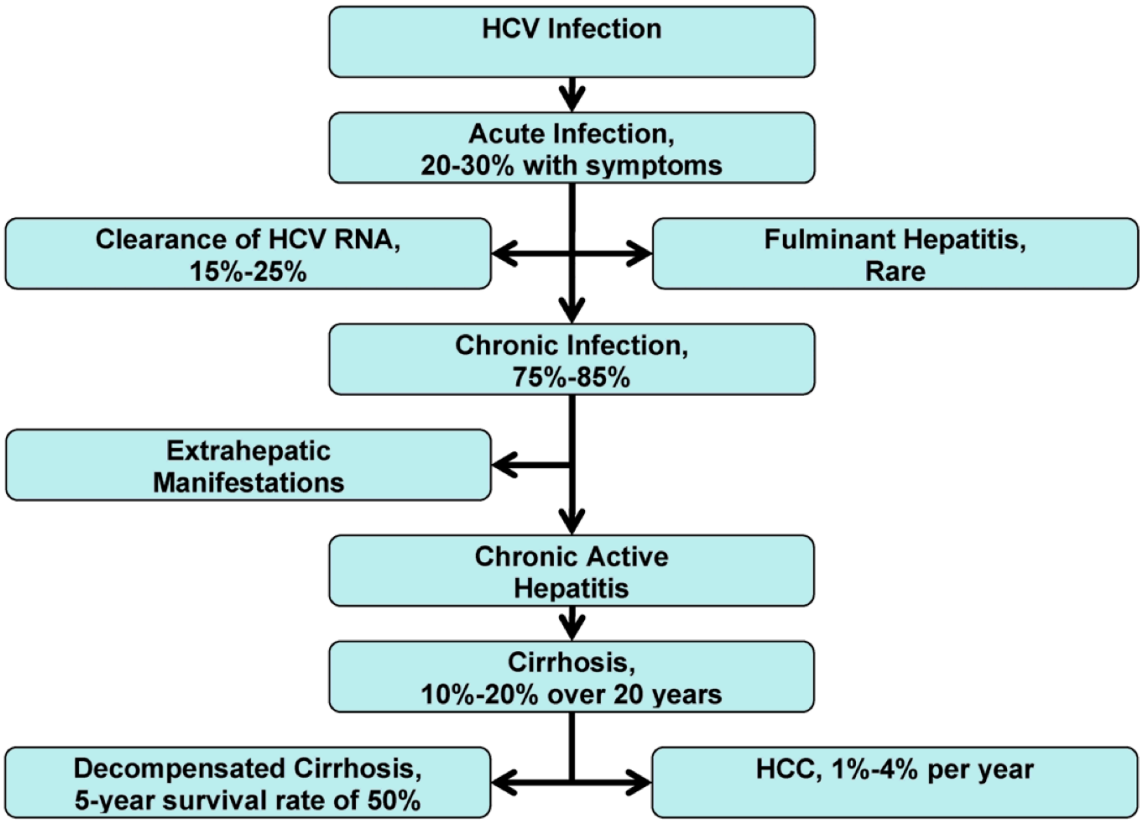

Figure 1.1 Flow Chart of Viral Infection and Related Outcomes (modified from [25]).

Diagnosis of HCV generally occurs during chronic stages when infected individuals begin to present symptoms of liver disease.

1.4 Viral Genome

Since first being cloned in 1989, the viral genome of HCV has now been well characterized. HCV is a 9.6 kilobase (kb), single-stranded, positive-sense RNA virus belonging to the *Flaviviridae* family[15]. HCV has a single, long open-reading frame (ORF) encoding a polyprotein of approximately 3000 amino acids[29]. In addition to the ORF, there is a 5'-noncoding region (NCR), which contains an internal ribosome entry site (IRES) and a conserved region at the 3' end necessary for genome replication[29]. The IRES mediates synthesis of the HCV polyprotein, which is then cleaved by cellular and viral proteases into 10 viral proteins[30,31]. The amino terminal region encodes the three structural proteins: the core (C) protein, and two envelope proteins, E1 and E2. Following the structural proteins is a small, integral membrane protein, termed p7[32]. After p7, the remainder of the genome encodes the six non-structural (NS) proteins: NS2, NS3, NS4A, NS4B, NS5A, and NS5B (Fig. 1.2)[32].

1.5 Life cycle

1.5.1 Virion Structure, Receptor Binding, Entry and Fusion

The HCV virion is made up of a nucleocapsid comprised of core protein and a single copy of the RNA genome[33,34]. The nucleocapsid is surrounded by a lipid envelope, envelope proteins E1 and E2, and host low-density lipoproteins (LDL) (Fig. 1.3)[33]. The virus is classified as a lipoviral particle (LVP), due to its low buoyant density caused by inclusion of serum lipoproteins in the virion membrane[35–37]. The presence of LDL in the virion is a common characteristic of *Flaviviruses*, and facilitates endocytosis through interactions with LDL receptors on target host cells[38].

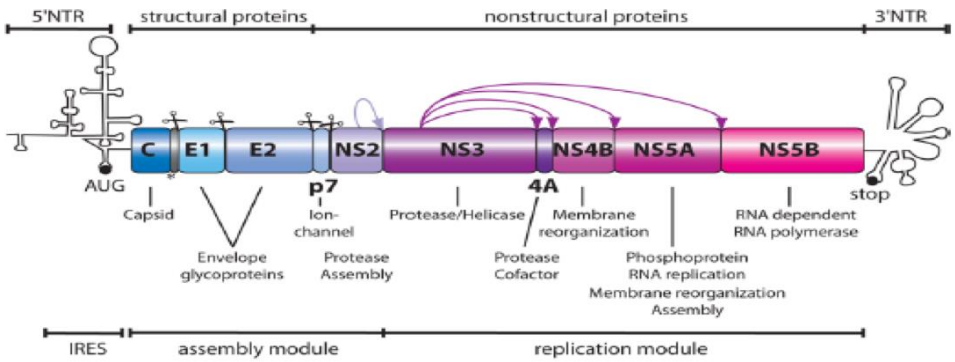

Figure 1.2 HCV Genome and Proteins ([39]). The HCV genome is approximately 9.6 kb and contains a single open-reading frame encoding ten viral proteins. The 5'-noncoding region contains the IRES. HCV contains 3 structural proteins, core, envelope protein 1 and envelope protein 2, and six non-structural proteins, NS2, NS3, NS4A, NS4B, NS5A, and NS5B. For the purposes of this figure, p7 was included in the non-structural proteins, however, it is still unclear whether p7 is structural or non-structural.

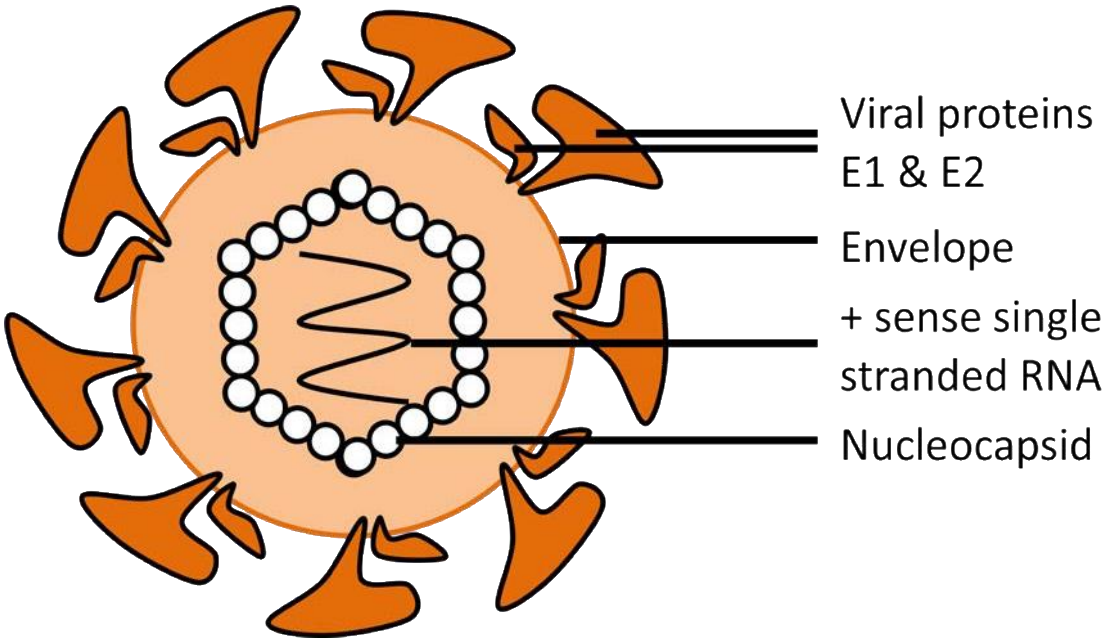

Figure 1.3 HCV Structure (modified from ([33])). The HCV virion is comprised of a single strand of positive-sense RNA contained within a nucleocapsid. The viral envelope is made up of viral proteins E1 and E2.

Many cell surface molecules have been proposed to mediate HCV binding. Cluster of Differentiation 81 (CD81), a member of the tetraspanin superfamily, has been shown to mediate cell binding of HCV E2 and was the first receptor to be identified as necessary in HCV infection[40]. In addition to CD81, scavenger receptor B type I (SR-BI) was also proposed as a surface molecule functioning in binding of HCV through interactions with E2[41]. LDLs present on the LVP play a large role in facilitating HCV entry through the LDL receptor, and a recognition step preceding infection is modulated by glycosaminoglycan receptors[38,42].

The late steps of HCV entry were shown to be mediated by tight junction proteins, claudin-1 (CLDN1) and occludin[43,44]. Evidence that Claudin-1 plays a key role in entry was confirmed as it can be used to make non-hepatic cell lines susceptible to HCV infection[43]. Following the discovery of the role for claudin-1 in entry, another tight junction protein, occludin, was identified, confirming the importance of HCV entering through the tight junction[44].

HCV entry is ph-dependent (acidic conditions) and occurs via endocytosis[33,45,46]. Thus far, the specific mechanism of fusion is poorly understood, however, research suggests that the host receptors identified play key roles in the conformational changes of the envelope glycoproteins that lead to fusion and release of the nucleocapsid[45,47–49]. Both E1 and E2 have been proposed as fusion proteins responsible for the release of the nucleocapsid[50–52]. Fusion of the virion via cellular receptors, endocytosis and entry into hepatocytes is depicted in (Fig. 1.4).

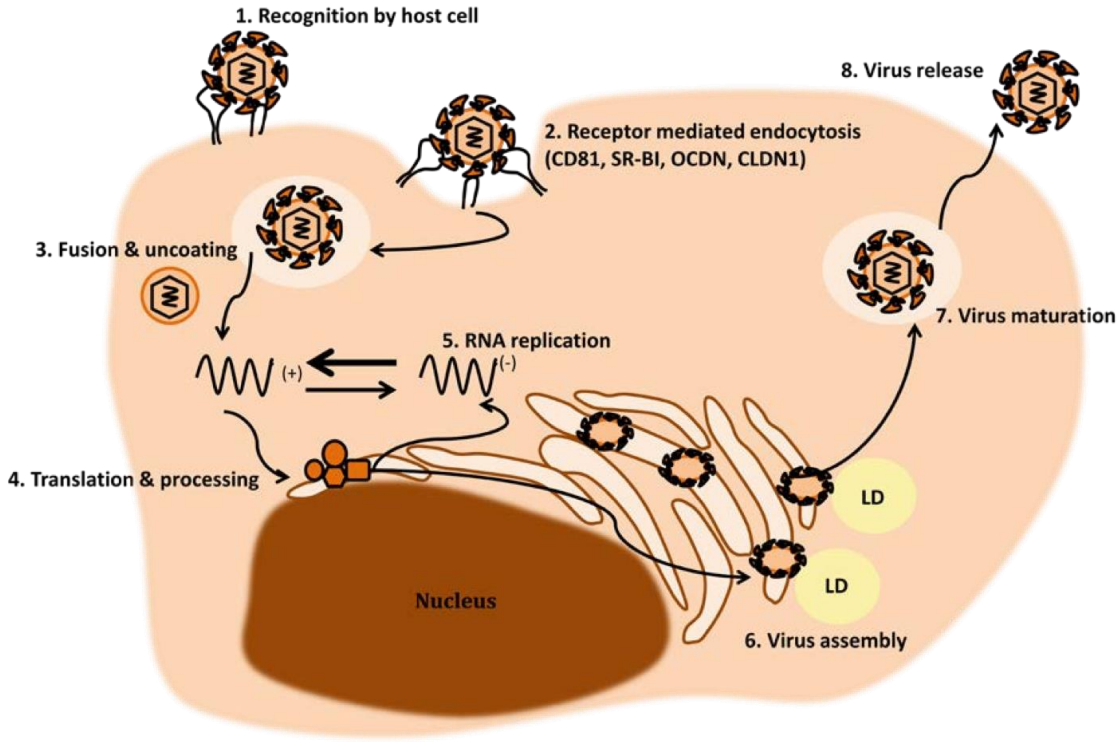

Figure 1.4 HCV Life Cycle (modified from ([33])). HCV enters hepatocytes via receptor-mediated endocytosis, with the proposed cell surface molecules CD81, SR-BI, OCDN, and CLDN1. The positive-strand genomic RNA is released into the cytoplasm and translation produces the polyprotein that will be proteolytically cleaved to release the viral proteins. The replication forms with the non-structural proteins, excluding NS2, and the two step replication process is catalyzed by NS5B. Core protein interacts with lipid droplets and the positive-strand RNA is then assembled into virions. Following assembly, mature virions are released from the hepatocyte.

1.5.2 Translation

Once the nucleocapsid has been released, the positive-strand genomic RNA is liberated into the cytoplasm[32,53]. The 5' terminal IRES initiates translation of the RNA genome generating a polyprotein that is co- and post-translationally cleaved by viral and host proteases[29]. Host proteases act to cleave the structural region of the polyprotein, NS proteins are cleaved through the activity of NS2 cleaving itself from NS3 and NS3/4A acting on the remainder of the polyprotein cleavage sites[53–56].

1.5.3 RNA replication

The non-structural proteins, with the exception of NS2, form a membrane-associated multiprotein complex which acts as the replication complex in HCV[57,58]. These complexes associate with a virally-induced rearrangement of the endoplasmic reticulum (ER), the membranous web[59,60]. This rearrangement of the ER increases the concentration of metabolites and viral proteins[58].

Replication of the HCV genome occurs in two steps[61–64]. The main protein involved in catalyzing replication is NS5B, the RNA-dependent, RNA polymerase (RdRp)[65]. The first step utilizes the positive-strand genomic RNA as a template for synthesis of a negative-strand intermediate[66]. In the second step, the negative-strand intermediate is used as the template to produce large amounts of positive-strand RNA[66]. The positive-strand RNA is utilized for polyprotein translation and synthesizing new replication intermediates and gets packaged into new virus particles[67].

1.5.4 Assembly and Release

Assembly is dependent on merging of the structural proteins, a newly synthesized RNA genome, and several host proteins that impart the low density of the LVP[42]. HCV assembly begins with the viral core protein interacting with cytosolic lipid droplets (cLDs), which are thought to act as the platform for nucleocapsid formation[68,69]. Interactions with core protein at cLDs are necessary for the viral RNA to be included in assembly. Core and NS5A mediate the association of the replication complex with cLDs[70–72]. The envelope glycoproteins E1 and E2 are translated and remain associated with the ER membrane during this process[73]. In addition to host and structural proteins, the non-structural proteins also play key roles in the process of assembly. NS2 has been proposed to coordinate the interaction of the non-structural proteins involved in assembly and the envelope proteins, playing a key role during the early stages of assembly[74–76]. The enigmatic protein p7 has also been proposed to play a role in assembly based on mutational studies, however p7's function has yet to be fully characterized[77–79]. A study utilizing chimeric genomes encoding structural and non-structural proteins showed that interactions between NS2 and NS3 and also between NS2, E1 and p7 are essential for virus assembly and/or release[75]. Mutation, deletion and adaptive studies support the proposal that NS3 helicase activity plays a major role in assembly in concert with the NS4A cofactor[33]. A study demonstrated that an assembly-defective NS4A mutant can be rescued by an NS3 mutation outside of the protease domain supporting the theory that these two non-structural proteins interact during assembly[80].

NS5A plays a multitude of roles during the viral life cycle and interacts with numerous host and viral proteins. It has been suggested that phosphorylation of residue 457 in domain III of NS5A plays and important regulatory step in infectious virus production[81]. In addition, domain III of NS5A functions in the unloading of core protein from lipid droplets, suggesting that interactions between these two proteins is linked with virus production efficiency[82]. Despite the current body of evidenceassembly is a complex and highly organized process that remains to be fully defined.

1.6 History of HCV Therapies

1.6.1 Interferon and Ribavirin

The standard of treatment for HCV infection has changed over the course of the last four decades[83]. The treatment regimen for HCV began in the 1980s as a monotherapy, consisting of treatment with recombinant interferon alpha (IFNα)[84,85]. IFNα was initially used to control the hepatic manifestations caused by NANBH, and liver enzyme tests and histological analysis were used to assess the primary outcomes[86]. In 1998, the standard of care evolved with the addition of RBV to the IFN monotherapy[87]. With the addition of RBV, the response rate to treatment approximately doubled, however, this was also accompanied by additional side effects. The next breakthrough came in 2001 when the classical standard of care evolved one again with the approval of peg-IFN[88]. The addition of these chemical groups increased the half-life of the protein *in vivo* by a substantial amount and peg-IFNα became the recommended treatment in combination with RBV[89,90]. Prior to the discovery of direct-acting antivirals

(DAA) therapies for HCV in 2011, the recommended standard of care for HCV infection was pegylated-interferon α (peg)-IFNα plus Ribavirin (RBV).

1.6.2 Direct-Acting Antivirals

The most promising treatment successes have come with the discovery of various DAAs that target multiple points during the HCV life cycle.[83,91] The first compounds to be approved in 2011, Telaprevir and Boceprevir, quickly became integrated into the standard of care for genotype-1 patients along with the classical peg-IFNα and RBV[92–94]. Both Telaprevir and Boceprevir are NS3/4A protease inhibitors (PI) that act on the proteolytic active site of NS3[95]. Following the success with the first wave of PIs, a second generation of optimized compounds are in clinical development (as reviewed in [96]). The approval of these compounds marked a major breakthrough in the field of antiviral pharmaceutics, since they were the first successful inhibitors specifically designed to cure a chronic infection and result in significantly improved SVR rates in populations infected with HCV genotype-1[96].

The next big breakthrough in HCV treatment came in 2010 with the discovery of Sofosbuvir and subsequent FDA approval in 2013[97,98]. Sofosbuvir differs from the first class of DAAs given that it acts as a nucleotide analogue, inhibiting polymerization after incorporation into a new viral RNA strand[99]. This was also the first inhibitor to be used in combination with RBV in the absence of IFN, indicating the possibility of IFN-free therapies[99]. This is of importance given that the side effects associated with IFN therapies can have an impact on compliance to treatment regimen. As with the PIs,

Sofosbuvir also markedly improves SVR, albeit to a greater extent than is seen with the first-generation PIs and SVR can be achieved in a much shorter time frame[100,101].

In addition to the already-approved compounds, several NS5B inhibitors are also in clinical development but have yet to receive FDA approval[102]. Many of the NS5B inhibitors are genotype-specific and most likely will only be useful in combination DAA therapy. In recent years, several NS5A inhibitors, including Daclatasvir and Ledipasvir have become a focus for developing combination therapies[99,103–107]. Although the mechanism of action is unknown, data suggests these compounds inhibit both replication and assembly. The discovery and subsequent approval of these compounds highlights the drastic changes in the standard of care for HCV treatment[83,108]. Although these successful treatment options are now available, DAA therapy remains extremely expensive. HCV will remain a health concern for North American healthcare systems until treatment is made widely available and affordable, and programs to reduce new infections are in place[13,109]. Furthermore, additional research is still needed to uncover other potential therapeutic targets.

1.7 p7 protein

p7 was first identified through expression of a series of C-terminally truncated HCV polyproteins fused to a human c-myc epitope tag[39,78]. These studies showed p7 to be located between E2 and NS2 within the viral polyprotein[76]. Homologous proteins have been identified in other viruses, such as bovine viral diarrhea virus, classical swine fever virus, and border disease virus, all containing a characteristic protein between E2

and NS2[39]. Studies have shown that cleavage at the E2-p7 and p7-NS2 junctions is delayed, yielding a precursor E2-p7-NS2 polyprotein[39,73]. The function of this precursor polyprotein hasn't been uncovered, however, the most likely hypothesis is that the polyprotein plays a regulatory role in viral kinetics or levels of final product expression.

HCV p7 protein is a small, 63-amino acid residue protein, which spans the endoplasmic reticulum membrane twice, forming transmembrane domains 1 and 2 (TM1 and TM2) connected by a short segment, termed the cytoplasmic loop, with its N- and C-termini oriented toward the cytosol[39]. Localization studies have found p7 in different areas of the cell. Initial subcellular localization studies in HepG2 cells showed a large fraction of p7 in an early compartment of the secretory pathway, which suggests some sort of retention signal maintaining p7 localization in the ER[110]. Conversely, p7 was shown to partially co-localize with mitochondria and adjacent membrane structures when green fluorescent protein (GFP)- or Flag-tagged p7 was visualized in HEK293T cells[111]. In addition, staining native p7 and tagged p7 demonstrated that untagged p7 was exclusively detected in the ER, whereas N-terminally-tagged p7 was detected in the ER or mitochondrial adjacent membranes[111]. Work done with eGFP-p7 or p7 tagged with HA downstream of the potential E2-p7 cleavage site showed that p7 localized only in the ER of Huh-7 cells[39]. This work suggests p7 localizes with multiple organelles indicating it may have a dual-role in HCV assembly and trafficking of nascent virions through cellular pathways. However, the variation in p7 localization dependent on the tagging method used may also suggest that these tags disrupted p7 function and virus production.

Homologous proteins in other viruses, including the p7 protein in BVDV were originally proposed to oligomerize and form ion-channels (Fig 1.5)[39]. Viral ion channels can play important roles in the virus' life cycle, by regulating replication, or aiding in virus entry, assembly or release[112]. Viral ion-channels could also modulate the electrochemical balance in subcellular compartments of host cells. This ion-channel activity was observed when p7 oligomerized to form a hexamer in artificial membranes and functioned as a calcium ion-channel in black lipid membranes[39,78]. Black lipid membranes are a model system that can be used to characterize the physico-chemical properties of lipid membranes and can provide functional information regarding ion-channels[113]. Two drugs, amantadine and rimantadine, are ion-channel blockers that have been used to treat the influenza A virus by blocking the ion channel activity of the M2 protein[78,114]. The role of p7 as an ion-channel was confirmed when the ion-channel activity of p7 could be blocked by amantadine and rimantadine[78,114]. The ion-channel activity of p7 lead to classification of this protein as a viroporin[39]. Viroporins are small hydrophobic proteins with the ability to form pores or channels within membranes for ion and small molecule movement[112].

Studies performed in the chimpanzee model confirmed the role of p7 in infectivity[77]. It was shown that any deletions or mutations in the cytoplasmic loop of p7 within infectious clones of genotype 1a (G1a) failed to cause viremia after intrahepatic transfection of chimpanzees[110]. In addition, substitution of G1a p7 with a p7 derived from a genotype 2a infectious

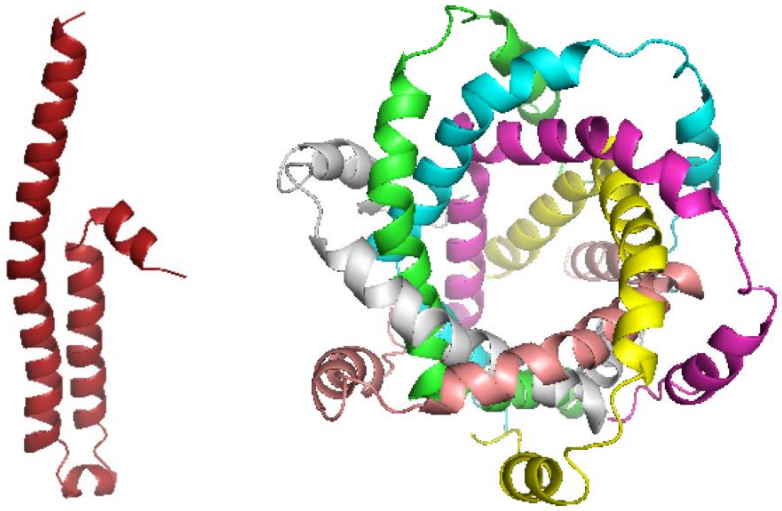

Figure 1.5 HCV p7 protein Monomer and Hexamer (Modified from [39]) HCV p7 protein is a polytopic membrane protein with two trans-membrane domains, TM1 and TM2, connected by a short segment, the cytoplasmic loop, with its N- and C-termini oriented toward the cytosol. HCV p7 can oligomerize to form a hexamer, which has been shown to have ion-channel activity.

clone was also not viable[39]. These results demonstrated, *in vivo*, the importance of p7 in the virus life cycle and virus infectivity, however, the function of p7 remains unknown.

Establishment of the HCV cell culture system allowed for mutational studies to be completed highlighting the importance of p7 for virus production[38]. One study supported a role for p7 at late stages in the viral replication cycle, given that mutants in p7 reduced total infectivity and the ratio of intracellular to released infectious particles. Current research suggests p7 does not play a role in viral entry or viral replication[38]. One study showed that the infectivity of released virions was maintained despite the fact that p7 had been mutated, demonstrating that it is most likely not acting during entry[38]. Despite evidence supporting the importance of p7 for virus production, it is still unclear what role p7 does play[38]. It has been suggested that p7 acts at a late stage in virus assembly, and may be acting as a dual function protein. One possible function of p7, in monomeric form, is assisting NS2 in gathering newly-formed capsids at LDs and glycoprotein complexes on the ER lumen for proper envelopment[38]. Another possible function of p7, in oligomeric form, is protecting glycoproteins from immature degradation during trafficking and release through its ion-channel activity[38]. Although there have been many proposed functions of p7, its role in the life cycle remains unknown. Since it has been shown that p7 plays a role in infectivity, it would be ideal to have a way of visualizing the protein, which would allow for better understanding of its function in the viral life cycle[38]. Despite the body of evidence that currently exists surrounding HCV research, conventional tagging methods have made it difficult to study p7 and its function in the life cycle[38].

1.7.1 Viral Protein Tagging Methods

One traditional method of investigating protein function has been to label or tag the protein of interest with some sort of fluorescent moiety[115]. A commonly used method is epitope tagging, whereby a known epitope is fused to the protein of interest. An epitope with available antibodies is selected, making it possible to detect proteins for which no antibody is available. The epitope-specific antibody will bind the epitope which is fused to the target protein[116]. One of the earliest used epitope tags was an epitope of the c-myc proto-oncogene product. This epitope tag has been used for immunohistochemical analyses, Western blots and subcellular localization studies. The hemagglutinin antigen (HA) tag is one of the most used epitope tags and has been used for immunohistochemistry, immunoprecipitation and Western blot analyses. The FLAG epitope is a synthetic peptide that can be used by placing multiple copies in tandem for enhanced protein detection. Currently, GFP is one of the most widely used tags, given its inherent fluorescence. GFP is a large protein that forms a "drum" structure, with 11β sheets surrounding an α-helix that sits diagonally across the inside of the "drum" and functions as a fluorophore[113,114]. The fluorescent nature of GFP allows for the visualization and localization of proteins without the use of an antibody. Although visualization studies using GFP can be conducted without an antibody, antibodies are available to be used for Western blot analysis and co-immunoprecipitation[113,114].

Although the technique of epitope tagging has facilitated the study of protein structure and function both *in vitro* and *in vivo*, there are some limitations and challenges associated with labelling. Despite numerous successes using the FLAG tag, there have

been complications, including reports of the FLAG tag disrupting activity when attached to the small GTPase H-Ras. The strategy of protein labelling by the fusion of an epitope to the target protein can be limited to either the C- or N-termini based on the epitope being used. Although GFP is extremely versatile, its large size can disrupt the structure and/or function of the target protein. The main limitation in using conventional tagging methods to study p7 is their large size and placement constraints[116]. The limitations associated with conventional tagging methods supports the need for a new way of visualizing proteins that isn't dependent on inserting a relatively large protein tag, which would facilitate the study of p7 and many other viral proteins.

1.7.2 Unnatural Amino Acids to Visualize Proteins

An alternative method for protein labelling involves the co-translational incorporation of a small unnatural amino acid (unAA) directly into the target protein in a site-specific manner. One unAA of interest is p-Azido-$_{L}$-phenylalanine (AzF), which was genetically encoded in *M. jannaschii*[117,118]. AzF is an aryl-azide, one of the most widely used photocrosslinking agents [119] (Fig. 1.6). Photocrosslinkers can be used to photochemically label antibodies with hapten, irreversibly inactivate enzymes and probe protein-peptide and protein-protein interactions[119]. In addition to photocrosslinking assays, AzF can be used in click chemistry reactions, whereby a fluorophore, or another convenient reactive moiety, can be attached to the side chain. Probably the most interesting unAA available is Anap, 3-(6-acetylnapthalen-2-ylamino)-2-aminopropanoic acid, which is particularly interesting because of its fluorescent property (Fig. 1.6)[120–123].

Anap is a derivative of prodan, an environmentally sensitive fluorophore that is commonly used in biochemistry and cell biology[120].

Figure 1.6 Structure of p-Azido-L-phenylalanine (AzF: left) and 3-(6-acetylnaphthalen-2-ylamino)-2-aminopropanoic acid (Anap: right) (Modified from [119,120]). AzF is an aryl-azide; upon irradiation with UV light at wavelengths below 310 nm can form short-lived singlet nitrenes which will rearrange to form dehydroazepines. Dehydroazepines can react with amines to form robust adducts. Aryl-azdes are commonly used as photocrosslinking agents to photochemically label antibodies, irreversibly inactivate enzymes and probe protein-peptide and protein-protein interactions. Anap is an amino acid derivative of prodan, a fluorophore. The absorption and emission maxima for Anap in water are 360 nm and 490 nm, respectively. Anap is inherently fluorescent.

Anap is an ideal candidate for subcellular localization studies of proteins using site-specifically incorporated Anap as a fluorescent reporter[120]. Incorporating Anap into a protein would allow the protein to be visualized without the use of antibodies or fixing, since the Anap incorporated within the protein would be fluorescent itself. Theoretically, any protein could be made fluorescent through the incorporation of Anap within the protein sequence. The fluorescent properties of Anap make it a useful tool for studying protein folding and protein-protein interactions. One group monitored protein misfolding by site-specifically labeling firefly luciferase with Anap[122]. The group was able to track the thermal unfolding and aggregation of luciferase *in vivo* by analyzing the Anap fluorescence emission by confocal imaging. The fluorescent properties of Anap will allow us to label proteins without the restrictions of conventional labeling methods.

1.8 Translation

In order to circumvent the issues mentioned above with the visualization of p7, we chose to employ a relatively novel strategy of protein labelling using unAA incorporation. Proteins play a major role in all biological processes and amino acids are the building blocks required for protein formation.

In order to understand how unAAs can be incorporated into proteins, we must first review basic protein synthesis[124,125]. Protein synthesis occurs through a step-wise process[125]. First, DNA is transcribed yielding mRNA. The process of transcription utilizes the DNA as a template strand and the resulting mRNA is a single-stranded

complementary copy of the gene. This mRNA strand must next undergo a process known as translation, the decoding of mRNA to synthesize a protein. The decoding of mRNA relies on two major components, ribosomes and transfer RNAs (tRNA). tRNAs act as a bridge between mRNA codons and the specific amino acids that they code for and have two specialized ends. One end of the tRNA has a sequence of three nucleotides, the anticodon, which can bind the corresponding mRNA codon. The opposite end of the tRNA carries the amino acid that is specified by the codons. An aminoacyl tRNA synthetase (synthetase) is the enzyme responsible for attaching the specified amino acid onto its tRNA. This process of amino acid binding to a tRNA is often referred to as "charging the tRNA". Ribosomes are the structures where protein synthesis occurs, and where translation machinery is located. Each ribosome has two subunits which will form around the mRNA, and act as a catalyst for amino acid linkage. The process of translation can be broken down into three major steps: initiation, elongation and termination. Once all codons in the mRNA have been read by the tRNA molecules, and their corresponding amino acids linked together, the newly synthesized protein must be released from the mRNA and ribosome. Termination of translation occurs once release factors bind to one of three termination codons and allows for the release of the mRNA from the ribosome. The release of the mRNA from the ribosome also results in the dissociation of the ribosomal subunits. The three termination codons that are present at the end of a protein-coding sequence in mRNA are: UAA, UAG, and UGA. These stop codons are not recognized by any endogenous tRNAs and the amber stop codon (TAG/UAG) is the least frequently used stop codon in the human genome.

1.9 UnAAs

The genetic code consists of 64 triplet codons specific for 20 canonical amino acids and 3 stop signals[126,127]. Although there are some variations of natural amino acids that have been identified, these changes arose through post-translational modifications of the canonical amino acids[115,118,121,126–138]. Two amino acids, selenocysteine and pyrrolysine, have been deemed as natural expansions of the genetic code given that they can be incorporated into proteins co-translationally. In addition to these naturally existing amino acids, there is a group of unAAs that can be incorporated into proteins yielding a variety of applications[126,128]. The applications of this technique are widespread given that there are over 100 unAAs with novel side chains including: bioorthogonal handles, photocross-linking moieties, fluorophores, and *in vitro* or cellular probes[115,118,121,126–138]. The incorporation of unAAs with these novel side chains can facilitate the study of proteins.

Similar to natural amino acids, the genetic encoding of an unAA requires a set of specific components to ensure proper incorporation of the unAA at the desired location (Fig. 1.7)[115,118,121,126–138]. The three requirements for unAA incorporation are a tRNA, a codon, and an aminoacyl-tRNA synthetase (hereafter referred to as synthetase). This orthogonal tRNA/codon/synthetase set must be functionally compatible with other components of the translation apparatus and must not crosstalk with any endogenous tRNA/codon/synthetase sets. The orthogonal tRNA must not be recognized by any

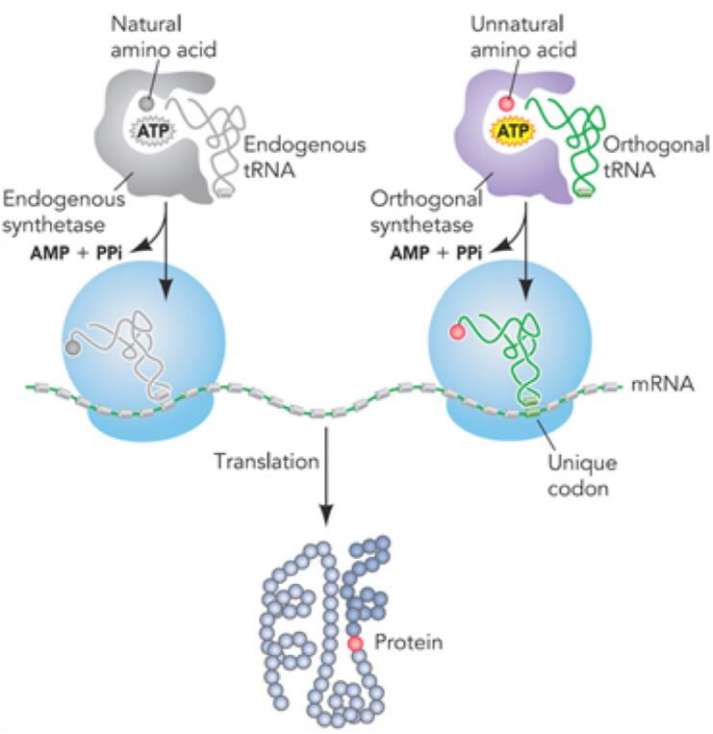

Figure 1.7 Unnatural Amino Acid Incorporation (modified from [136]) UnAA incorporation occurs via the same process as natural amino acids. UnAA incorporation has three requirements for successful incorporation: a unique codon, an unAA that is stable in the cellular environment, and a tRNA/synthetase pair specific for only the unAA. Incorporation of an unAA uses the host cells translational machinery therefore the tRNA/synthetase must not bind any endogenous amino acids.

endogenous synthetase, and should decode only the orthogonal codon, which cannot be assigned to any canonical amino acid. In order for the system to function properly, the orthogonal synthetase should only charge the orthogonal tRNA with the unAA. Once the synthetase is expressed in cells, it can charge its orthogonal tRNA with the desired unAA. This process should result in the unAA being incorporated into proteins in response to the unique codon by utilizing the endogenous translational machinery[115,118,121,126–138].

In order to establish this technique, researchers had to perform very complex positive and negative selection to create tRNA/synthetase pairs that were specific for various unAAs (as reviewed in [136]). Only tRNA/synthetase pairs that could be used in conjunction with the endogenous translational machinery in different organisms would be considered successful[134]. This technique has been applied to expand the genetic code of many cells ranging from bacteria, yeast, mammalian, stem cells and neurons to multicellular organisms including a primitive animal, an insect and a plant[139]. One group sought to expand the genetic code in the mouse (*Mus musculus*) brain using adeno-associated viral (AAV) vectors for unAA incorporation[137]. Establishment of this system in the mouse brain meant that various unAAs could be incorporated and used as tools to probe and control protein functions in a live vertebrate. Despite all of the work done, this technique remains to be applied to visualize viruses[139]. With the development of systems to incorporate unAAs into proteins, researchers have uncovered an innovative technique that can be utilized to both investigate and engineer protein structure and function.

1.10 Project design and hypothesis

As mentioned above, the function of HCV p7 is controversial, with the main reason for uncertainty that Western blots, immunoprecipitation and immunofluorescence cannot routinely be performed on p7, and due to the small size of p7 and inherent complications of protein tags, tagging studies have generated conflicting results. Researchers had to employ conventional tagging methods for p7 studies since there are no reliable antibodies available for the visualization of p7. It has been shown that p7 may function in different ways based on the tagging method that was used. We speculate that the tags were most likely interfering with protein structure or function[39]. This is likely due to the fact that the majority of the protein is buried within membranes in an infected cell. Another obstacle to studying the HCV p7 protein was the absence of a fully infectious cell culture system. Initially, HCV research was conducted using a replicon system, which utilized a neomycin selection gene and an ECMV IRES to mediate the translation of HCV NS proteins and enable replication *in vitro*. Despite its practical uses, the replicon system failed to produce infectious virus and did not express the viral structural proteins. Next, researchers established the HCV pseuoparticle system, which used retrovirus biology to create chimeric viruses containing the HCV glycoproteins on the surface. This system was used for the generation of particles containing E1 and E2 which could then be used to infect permissive hepatocytes to allow for the study of virus entry. Up until 2005, all HCV research was conducted using the replicon and pseudoparticle systems. Although both systems had their specific uses, neither system allowed for the study of p7. A major breakthrough came when the HCV cell culture

(HCVcc) system was discovered and provided the capacity to study the whole HCV life cycle, including assembly and release. The establishment of the HCVcc system allowed researchers to study the viral proteins throughout the life cycle which has provided valuable knowledge into protein function and interactions[39].

The goal of this project was to develop a recombinant strain of HCV containing a fluorescent unnatural amino acid within p7 that would enable us to visualize this protein by immunofluorescence and confocal microscopy. The ability to visualize p7 within the context of a replicating virus would allow us to analyze p7 localization within the cell, as well as co-localization with other viral and cellular proteins, and all of these analyses could theoretically be performed in live cells. The availability of such a tool would allow us and others to study HCV proteins in a way that they've never been analyzed before, and ultimately, this strategy could then be applied to other viruses.

1.11 Project aims

I. **Screening positions for unAA incorporation in the HCV p7 protein.**
Before testing incorporation of an unAA, we needed to select various positions throughout the HCV p7 protein. These selected sites would then have a unique TAG codon inserted using site-directed mutagenesis to allow for incorporation of the unAA. A total of 14 sites were selected to be screened. Once we selected the sites that would be used to screen amenability to unAA incorporation, we tested the system for incorporation of an unAA into the p7 protein in the context of a replicating virus. Following incorporation of the unAA during transfections, cells would be fixed and probed with a specific HCV core antibody to confirm virus

production in the presence of the unAA. Theoretically, with the fluorescent unAA, we should visualize an overlap of green fluorescence from Anap, with the specific HCV core staining.

II. Screening mutants for an amenable position in HCV core protein.

Following challenges with unAA incorporation in the context of a fully replicating virus, we shifted our focus to testing incorporation in select core mutants in the context of a single-protein expression system. We selected 11 potential sites for incorporation of Anap in the HCV core protein. If any sites were found that could successfully incorporate the unAA, Anap, we would re-test those sites in the HCV core protein in a replicating virus. Theoretically, if Anap can be incorporated at a specific position in the HCV core protein in a single-protein expression system, it should also be possible for incorporation of Anap at that site in a replicating virus. In addition, testing incorporation of Anap in a single-protein expression system first allows for optimization without the mutational pressure provided by the replicating virus.

Chapter 2: Methodology

2.1 Primer Design

Individual mutagenic oligonucleotide primers were designed containing the desired mutation. Special consideration had to be given to ensure all primers were approximately 30 bases in length while containing the mutation in the middle of the primer. In addition, high GC content was desired while starting and terminating with at

Table 2.1 List of all primer sequences within p7

Mutant	Primers (5' → 3')
L755TAG	CA GCA TTG GAG AAG TAG GTC GTC TTG CAC G
L758TAG	GAA GTT GGT CGT C TAG CAC GCT GCG AGT GC
A761TAG	CGT CTT GCAC GCT TAG AGT GCG GCT GAC TGC
L769TAG	GCT GAC TGC CAT GGC TAG CTA TAT TTT GCC
Y771TAG	CAT GGC CTC CTA TAT TAG TTT GCC ATC TTC TTC
F772TAG	CAT GGC CTC CTA TAT TAG GCC ATC TTC TTC G
F775TAG	C CTA TAT TTT GCC ATC TAG TTC GTG GCA GC
F776TAG	GCC ATC TTC TAG GTG GCA GCT TGG CAC ATC AGG
W780TAG	C TTC TTC GTG GCA GCT TAG CAC ATC AGG GG
T790TAG	GTG GTC CCC TTG TAG ACC TAT TGC CTC ACT GG
Y792TAG	GTC CCC TTG ACC ACC TAG TGC CTC ACT GGC
W798TAG	CTC ACT GGC CTA TAG CCC TTC TGC CTA CTG
F800TAG	GGC CTA TGG CCC TAG TGC CTA CTG CTC ATG GC
P808TAG	CTC ATG GCA CTG TAG CGG CAG GCT TAT GCC

least one G or C. A complete list of primers can be found in Table 2.1. All primers were purchased from Invitrogen.

2.2 In-vitro Site-directed mutagenesis

Plasmids were constructed using the QuikChange® II XL Site-directed Mutagenesis Kit (Strategene). $JFH1_T$ was used as the dsDNA template (10ng) and was combined with the forward and reverse oligonucleotide primers containing desired mutations (100 ng), in a solution containing the 10X reaction buffer, dNTP mix, QuikSolution, ddH_2O and *PfuUltra* HF DNA polymerase. A three-segment thermal cycle was used to extend primers as per kit protocol. The 68°C temperature in the second segment was held for 13 minutes. Following the thermal cycling *Dpn-1* was added and incubated at 37°C for one hour in order to digest parental dsDNA.

2.3 Transformation

XL 10-gold ultracompetent cells were transformed upon addition of *Dpn-1* treated DNA as per QuikChange® II XL Site-directed Mutagenesis Kit (Strategene) protocol.

2.4 Miniprep DNA Purification

Colonies were selected from transformed XL 10-gold ultracompetent cells and incubated overnight in 10 mL LB broth. Plasmid DNA was isolated from bacteria culture via mini preparation using the QIAprep® Miniprep Kit (Qiagen).

2.5 *PVUII* Digestion

Miniprep DNA samples and control plasmid ($JFH1_T$) were treated with the restriction enzyme *PVUII* at 37°C for two hours, run on a 1.5% agarose gel, and analyzed by UV light to compare digestion patterns for initial screening of mutagenesis success.

2.6 Sequencing of Plasmids

Plasmids were sent for sequencing to the Centre of Applied Genomics. For sequencing, the plasmid concentration was approximately 300 ng/μL. Primers were diluted to a stock concentration of 5 ρmol in 0.7 μL and the diluted primer was used in sequencing reactions.

2.7 Maxiprep DNA Purification

Miniprep DNA samples containing the correct sequence were re-transformed in E. coli Dh5-α competent cells and incubated overnight in 100 mL LB broth. Plasmid DNA was isolated from bacteria culture via maxi preparation using QIAprep® Maxiprep Kit (Qiagen).

2.8 Plasmid Linearization

Plasmid DNA (25μg) was incubated at 37°C for 1 hour with 5 μL *XbaI* (Invitrogen), 5 μL BSA, 5 μL buffer and dH_2O up to 50 μL. Following the digestion, dH_2O was added to increase volume to 100 μL followed by 100μL of phenol/chloroform/isoamyl alcohol (Invitrogen) and the sample was then mixed by vortex for 5 seconds. Samples were spun in a table top microcentrifuge for 3 minutes at 14,000xg. Next, 100 μL of upper phase was transferred to a new tube. Then, 3 μL of 5 M

NaCl was added followed by 230 μL of 100% EtOH. Samples were incubated at -20°C for 1 hour followed by a 5 minute spin at 14,000xg. Supernatant was removed and the pellet was dried then resuspended in 30 μL of dH_2O and DNA concentration was determined using a Nanodrop. Linearized DNA was either stored at -20C or used directly for transcription.

2.9 Cell Culture

A human-hepatoma derived cell line (Huh-7.5) was cultured at 37°C in a 5% CO_2 incubator in 15 cm tissue culture dishes (Corning Incorporated). Cells were grown in Gibco® Dulbecco's Modified Eagle Medium (DMEM; Invitrogen) supplemented with 10% heat-inactivated fetal bovine serum (FBS; Invitrogen) and antibiotics penicillin and streptomycin (PenStrep; Sigma) yielding complete medium ($DMEM_{(comp)}$).
Cells were passaged when cell growth reached approximately 75-90% confluency. To passage cells, medium was removed from the culture dish and cells were treated with 0.5% trypsin-EDTA (Invitrogen) and incubated at 37°C in 5% CO_2 for 6 minutes. The plate was washed with 10 mL $DMEM_{(comp)}$ twice, and suspended cells were transferred to a 50 mL sterile polypropylene tube. Cells were then centrifuged at 400xg for 5 minutes. Cell pellets were resuspended in 20 mL of $DMEM_{(comp)}$ Cells were counted using a hemocytometer and $2x10^6$ cells were replated into culture dishes containing 25 mL of $DMEM_{(comp)}$.

2.10 DNA Transfection

Huh-7.5 cells were seeded in 10 cm dishes at a density of $1x10^6$ per dish 24 hours before DNA transfection. Cells were washed twice and replaced with 2 mL of serum-free (SF) media/plate. Next, 500 μL of SF media and 20 μL of Lipofectamine®2000 (Thermo Fisher Scientific) were combined with 500 μL of SF media and 1 μg of either pMah or pAnap and mixed gently. pMah is the plasmid that makes the tRNA and synthetase for incorporation of the unAA AzF; pAnap is the plasmid that makes the tRNA and synthetase for incorporation of the unAA Anap. Transfection mixes were then added to cells and incubated for 3 hours. Following incubation, cells were washed twice and 7 mL of $DMEM_{(comp)}$ was added to each plate. Transfected cells were incubated for 24 hours to allow expression of the plasmid (pMah or pAnap) before transfection with viral RNA.

2.11 Transcription

On the day of transfection, 1 μg of linearized DNA was transcribed using the T7 RiboMAX™ Express Large Scale RNA Production System (Promega) following the kit protocol.

2.12 RNA Transfection

Huh-7.5 cells were transfected with the pMah or pAnap DNA plasmid 24 hours before the RNA transfection. Cells were washed twice and 2 mL of serum-free (SF) medium was added to each plate. Next, 50 μL of DMRIE-C transfection reagent (Invitrogen) and 4 μL of RNA transcripts were mixed lightly. Transfection mixes were

then added to cells along with either 1 or 10 μM of the unAA (1 μM Anap; 10 μM AzF) and incubated for 3 hours. Following incubation, cells were washed 2 times and 7 mL of $DMEM_{(comp)}$ supplemented with 10 μM of the respective unAA was added to each plate. Cells were incubated for 72 hours before supernatants were collected and clarified.

2.13 Determination of Infectious Titre

At 48 hours pre-infection, Huh-7.5 cells were transfected with the DNA plasmid coding for the tRNA/synthetase. At 24 hours pre-infection, 50,000 transfected Huh-7.5 cells were plated in 8-well chamber slides (ThermoFisher) in 400 μl $DMEM_{comp}$ and incubated overnight. To determine virus titre, clarified transfection supernatants were serially diluted 10-fold and 100 μL of each dilution were inoculated in triplicate in chamber slides containing pMah-transfected Huh-7.5 cells and incubated for 4 hours before $DMEM_{comp}$ was added and cells were incubated for an additional 72 hours. Focus-forming units were counted after 72 hours using a Zeiss Axio Imager M2 microscope after immunofluorescence staining. Infectious titres were determined in triplicate and averaged to yield the supernatant titre.

2.14 Immunofluorescence Staining

At 72 hours post-infection supernatants and plastic chambers were removed. Slides were dipped in 10% phosphate buffered saline (PBS; Invitrogen) for 2 minutes, fixed in acetone for 1½ minutes and rubber gaskets removed. The primary antibody, mouse anti-HCV core monoclonal antibody (B2; Anogen), was diluted 1:200 in 5% BSA in PBS solution and 25 μL was placed on each well. Slides were covered with a glass

cover slip (ThermoFisher) and incubated for 20 minutes. Slides were dipped in PBS for 1 minute to remove the cover slip and then washed in PBS for 5 minutes. Secondary antibody, Alexa Fluor ® 488 (Invitrogen) or Alexa Fluor ® 594 (Invitrogen), was diluted 1:100 in PBS and 25 μL was placed on each well. Slides were covered with a glass cover slip and incubated for 20 minutes. Slides were dipped in PBS for 30 seconds to remove cover slip and then washed in PBS for 5 minutes. Slides were stained with Vectashield Hardset mounting medium containing DAPI (Vector Labs) and covered with a glass cover slip. Once the Vectashield had dried, slides were examined using a Zeiss Axio Imager M2 microscope and observed using the 20X objective.

2.15 G418 Treatment

To determine the effect of G418 on pAnap-transfected Huh-7.5 cells, 200,000 Huh-7.5 cells were plated in 2-well chamber slides (ThermoFisher) in 2 ml $DMEM_{comp}$ and incubated overnight. After 24-hours Huh-7.5 cells were transfected with pAnap and incubated for 24 hours before treatment with G418. pAnap-transfected cells were then treated with 75 μg/mL G418 and incubated overnight. After 24-hours G418 treated, pAnap-transfected Huh-7.5 cells were transfected with L755X, L758X, and A761X encoded RNA. Anap was added to the cells during the RNA transfection to ensure the possibility of incorporation during translation. In addition, cells were maintained in G418 during the RNA transfection. Following the 48-hour incubation after the RNA transfection, cells were fixed and visualized by indirect immunofluorescence.

2.16 G418 Titration

100,000 Huh-7.5 cells were plated in a 24-well plate in 1 mL $DMEM_{comp}$ and incubated overnight. On the following day, cells were treated, in triplicate, with increasing concentrations of G418. The concentrations used for the titration were 0 μg/mL, 10 μg/mL, 20 μg/mL, 40 μg/mL, 60 μg/mL, 80 μg/mL, and 100 μg/mL. After 48 hours, the supernatant was removed and the cells were washed with 300 μL of trypsin. Cells were then incubated for 6 minutes in 300 μl of trypsin. Following the 6 minute incubation, 500 μl $DMEM_{(comp)}$ was added to the wells to inactivate the trypsin and the total volume was transferred to 12 mL round bottom tubes. Cells were then counted using a hemocytometer.

2.17 Transfection of Core Mutants

At 24-hours before the DNA transfection, 300,000 Huh-7.5 cells were plated in 2-well chamber slides and incubated overnight. Transfection mix A was made from 50 μL of Optimem medium (Gibco) and 1.5 μL of Lipofectamine®3000 (Thermo Fisher Scientific). Transfection mix B was made from 50 μL of Optimem medium, 2 μL of P®3000, and 1 μg of pAnap was added and mixed gently. Both mixes were aliquoted for the number of mutants being tested, and respective DNA (wild-type core, core mutants and GFP) was added to mix B. Transfection mixes were then combined and incubated for 15 minutes at room temperature. Culture fluids were replaced by 1 mL of Optimem medium/well. The combined transfection mix was then added to cells along with 10 μM Anap and incubated for 3 hours. Following incubation, medium was replaced with 2 mL

of $DMEM_{(comp)}$. Anap was added at a concentration of 10 μM and transfected cells were incubated for 48 hours before being fixed and stained for immunofluorescence.

Chapter 3: Results

3.1 Mutant Selection:

The first step that was necessary for unAA incorporation into p7 was the selection of potential sites for substitution. We wanted to select sites that could potentially tolerate the substitution of a variety of unAAs. Given that different unAAs have varied side chains, making them useful for a number of different applications, we wanted to select a wide variety of amino acid residues to account for various structural substitutions. This selection process yielded a panel of 14 mutants that could potentially be used for different unAAs. One unAA that we would be working with was AzF and as such, a select number of our mutants had phenylalanine → TAG substitutions. The other unAA that we would be working with, for the purpose of protein tracking, was Anap. We selected tryptophan and tyrosine residues to potentially accommodate the substitution of Anap based on their moderate structural similarity. In addition to those amino acids, we also selected various others to test amenability to incorporation. To maximize the potential for successful incorporation, we selected residues spanning both transmembrane domains. We avoided making any potential substitutions in the cytoplasmic loop, since our previous mutational studies had shown those residues to be essential for activity[77]. The selected mutations were: L755X, L758X, A761X, L769X, Y771X, F772X, F775X, F776X, W780X, Y792X, T790X, W798X, F800X, and P808X. Once the mutants were

designed, site-directed mutagenesis using custom primers was performed to introduce the mutations into the viral genome. The mutated HCV plasmids were digested using a restriction enzyme to confirm sequence similarity with the $JFH1_T$ template (Fig. 3.1). All successfully mutated plasmid constructs were validated by Sanger sequencing at The Center for Applied Genomics. Cloning of some mutants turned out to be problematic and those will be revisited in the future using more elaborate cloning procedures.

3.2 The Effect of AzF Incorporation on Viral Infection:

AzF is a chemically active unAA that can be used for various functions, including binding a fluorophore. Attaching a fluorophore directly to a residue within the p7 protein would allow us to visualize the protein without the use of a tag. To determine whether an unAA could be successfully incorporated into the HCV genome, we transfected pMah, the DNA plasmid coding for the tRNA and synthetase specific for AzF, followed by transfection of mutated RNA 24-hours post-pMah transfection. Cells were cultured with AzF during the RNA transfection to ensure the possibility of unAA incorporation during translation. Transfected cells were harvested three days post-transfection and seeded into 8-well chamber slides for 3 days. The HCV core protein was detected in L758X + AzF virus, and A761X + AzF virus but not in L755X + AzF or L769X + AzF (Fig. 3.2). We postulated that the unnatural amino acid was successfully incorporated into L758X + AzF virus, and A761X + AzF virus given that the HCV core protein was detected. If the unAA wasn't incorporated, translation should terminate at the stop codon located in the p7 sequence resulting in no core staining. Although core is upstream of p7, it won't be produced in the absence of AzF given that HCV is translated as a polyprotein.

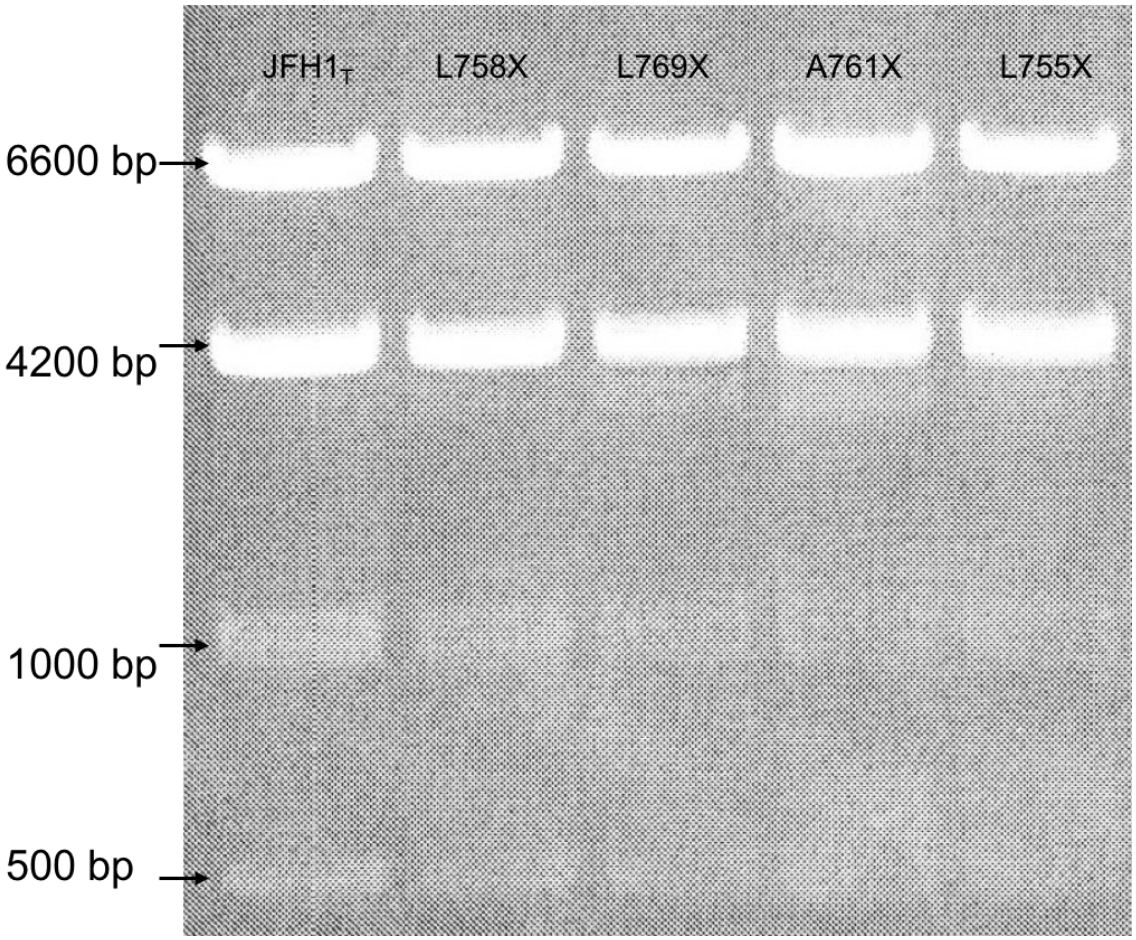

Figure 3.1 DNA gel of maxi prepared samples following *PvuII* digestion. The presence of bands at the same places when compared to the template $JFH1_T$ confirms that all mutated versions of the plasmid were intact.

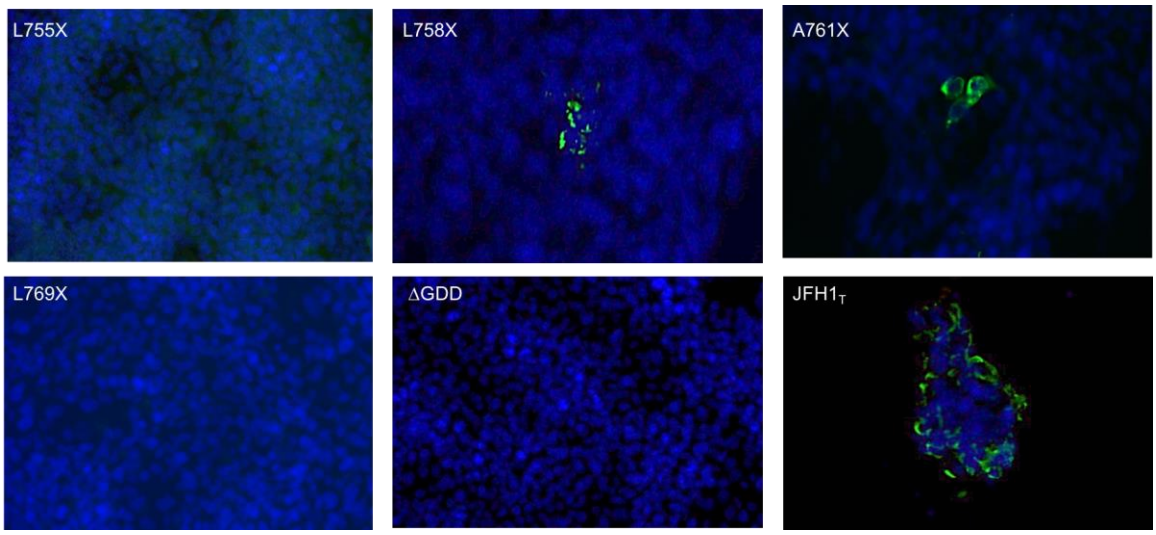

Figure 3.2: The Effect of AzF Incorporation on Virus Infection. $1x10^6$ Huh-7.5 cells were transfected with $JFH1_T$, L755X, L758X, A761X, L769X or ΔGDD encoded RNA. Cells were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (green) after 72 hours, observed using the 20X objective. The pictures shown are representative of the entire area of the well. Results shown are representative of 2 independent experiments.

These results were compared to the wild-type $JFH1_T$, which showed expected levels of viral spread, and to ΔGDD, containing a deletion in the active site of NS5B, representing the replication-defective negative control for the experiment.

3.3 The Effect of AzF on Virus Production:

The effect of AzF incorporation on virus production was determined by measuring the levels of infectious virus titre using a focus-forming unit assay. Three days post-infection viral titres were analyzed. $JFH1_T$- L755X -AzF, $JFH1_T$-L758X - AzF, $JFH1_T$-A761X - AzF, $JFH1_T$-L769X - AzF, represent the system negative controls, i.e., without the unAA present, translation should terminate at the amber codon. Variants with or without AzF produced very low levels of infectious titre when compared to the wild-type $JFH1_T$ (Fig. 3.3). No infectious virus was detected in ΔGDD, A761TAG + AzF, L769X + AzF, and L769X - AzF. Infectious virus was detectable in L755X - AzF, L758X - Anap, and A761X - Anap; this was unexpected given that translation should terminate at the amber stop codon in the absence of the unAA. As we weren't able to visualize the unAA at this point, we couldn't confirm that the successful virus production was due to unAA incorporation as opposed to read-through or reversion of mutations.

3.4 Rationale for Using Anap Instead of AzF

Although it appeared as though we had incorporation of AzF in our initial trials, without the addition of a fluorophore, we couldn't easily confirm that the unAA had been successfully incorporated. We couldn't eliminate the possibility of read-through or reversion events in our first trial with AzF without some way to distinguish where in the

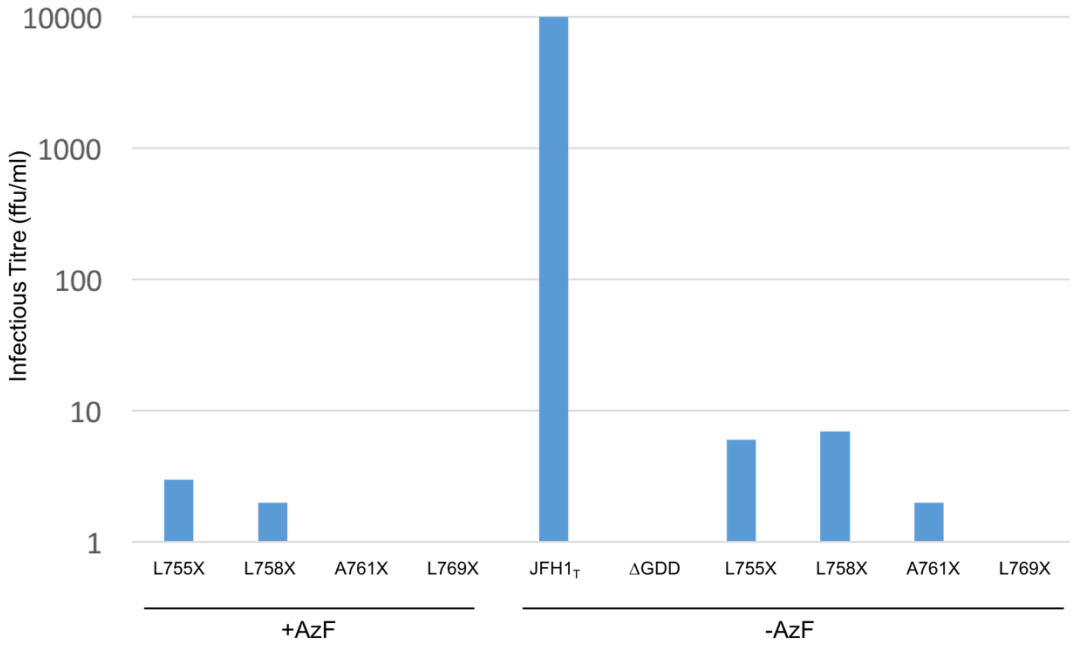

Figure 3.3 The effect of AzF on Virus Production. Culture fluids from cells transfected with $JFH1_T$, L755X, L758X, A761X, L769X and ΔGDD were collected 72 hours post-transfection, serially-diluted, and used to infect naïve Huh-7.5 cells that had been pre-transfected with pMah. $JFH1_T$ and $JFH1_T$-ΔGDD-transfected cells were used to infect naïve Huh-7.5 cells that had not been pre-transfected with pMah. Seventy-two hours post-infection cells were fixed and probed for the presence of core protein in order to determine the number of focus-forming units (ffu) per mL of supernatant. Viral titres were measured in triplicate. Results shown are representative of 1 experiment.

pMah-transfected Huh-7.5 cells the unAA was being incorporated. We surmised that we could overcome this issue by testing the amenability of specific amino acid positions to incorporation of an unAA using a fluorescent unAA. Based on the literature, we decided that switching to Anap, a fluorescent unAA, would also make conducting localization studies significantly easier, since it would eliminate the need for a p7-specific antibody.

3.5 Anap incorporation in Huh-7.5 cells

To test whether Anap could be incorporated into the HCV genome we transfected pAnap, followed by transfection of RNA 24 hours post-transfection of pAnap. Anap was added during the RNA transfection to ensure the possibility of incorporation during translation. Transfected cells were harvested three days post-transfection and seeded into 8-well chamber slides and cultured for 3 days. Anap fluorescence was visible in L755X +Anap, L758X +Anap, A761X +Anap, and L769X +Anap (Fig. 3.4). The presence of fluorescence confirmed incorporation of the unAA, however, since these slides were not stained for HCV core, it cannot be confirmed whether incorporation was into the HCV genome or into cellular proteins. We did not stain for core during this experiment given that the secondary antibody we normally used in our lab, Alexa Fluor ® 488, would fluoresce at the same wavelength as Anap. Any green staining observed with this secondary antibody would result in undistinguishable results, since we wouldn't be able to discriminate between Anap incorporation versus core fluorescence. Therefore, in subsequent experiments, HCV core protein was visualized using a different secondary antibody that fluoresced red, i.e., Alexa Fluor ® 594.

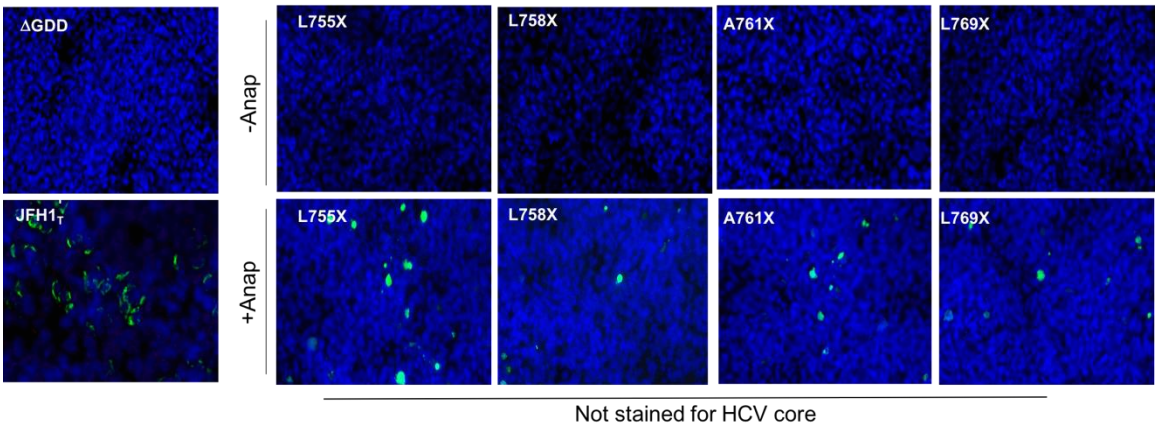

Figure 3.4 Anap fluorescence was visible in transfected Huh-7.5 cells. At 48 hours pre-RNA transfection, Huh-7.5 cells were transfected with pAnap. At 24 hours pre-RNA transfection, 50,000 pAnap-transfected Huh-7.5 cells were plated in 8-well chamber slides in 400 µl $DMEM_{comp}$ and incubated overnight. The pAnap-transfected Huh-7.5 cells were then transfected with $JFH1_T$, L755X, L758X, A761X, L769X and ΔGDD encoded RNA. Seventy-two hours post-transfection, cells were fixed and probed with DAPI (blue). ΔGDD and $JFH1_T$ were probed with a specific HCV core antibody, by indirect immunofluorescence (green), observed using the 20X objective. The green visible in well containing L755X + Anap, L758X + Anap, A761X + Anap, and L769X + Anap was the fluorescence emitted by the unAA itself. The pictures shown are representative of the entire area of the well.

3.6 Rationale for scale down of Anap experiments

One potential issue that we wanted to address was the transfection efficiency of Huh-7.5 cells. Huh-7.5 cells generally have a low transfection efficiency, and up to this point, our experimental setup required the initial DNA transfection as well as an additional DNA transfection when setting up chamber slides for immunofluorescence. We were transfecting cells with the DNA plasmid to then be seeded into 8-well chamber slides for immunofluorescence. In addition to the issues with transfection efficiency, we wondered if it would be possible to cut down the time it takes between starting an experiment and staining the slides for immunofluorescence. Given that we were now working with a fluorescent unAA, our results are based on the images that we get following visualization on the microscope. At this point, we decided that it would be optimal to perform the entire experiment in 2-well chamber slides to eliminate the need for the additional transfection step. Scaling down the experiment into these 2-well chamber slides would allow us to do both the DNA and RNA transfections directly into the slides, followed by immunofluorescence staining directly on the slides. This scale down took the experimental set up from an 8-day protocol to a 5-day protocol. We continued using the 2-well chamber slides for the duration of our experiments.

3.7 Anap was detected in transfected cells stained for core

To test incorporation of Anap into the HCV genome, Huh-7.5 cells were transfected with pAnap, followed by transfection of mutated RNA 24 hours post-pAnap transfection into 2-well chamber slides. Anap was added during the RNA transfection to allow incorporation during translation. Transfected cells were fixed and stained 2 days after the

RNA transfection and visualized by indirect immunofluorescence. ΔGDD and $JFH1_T$ represent our experimental controls, with high levels of HCV core in cells transfected with $JFH1_T$, and no HCV core detected in cells transfected with ΔGDD. We included an additional control in this experiment, the 'transfected cells' well to determine how much background Anap fluorescence is detectable when cells are transfected with pAnap in the absence of viral RNA. As seen in the panel 'transfected cells' well, it is evident that there is background expression which may be attributed to incorporation into cellular proteins. No HCV core was detected in L769X in the presence or absence of Anap. HCV core was detected in L755X, A761X, and L758X in the presence and absence of Anap (Fig. 3.5). The fact that there was HCV core detected in the absence of Anap confirms that reversion or read-through was occurring. We expected we would be able to distinguish successful incorporation from any read-through or reversion events since incorporation would be characterized by the presence of green from Anap in the same location as red from HCV core staining. We were unable to isolate any HCV core positive cells containing green fluorescence from Anap.

3.8 Treatment of DNA transfected cells with G418

Since we encountered the issue of HCV core being detected in cells that did not contain Anap, we wanted a way to isolate the DNA transfected cells to maximize the chance of DNA transfected cells then being transfected with RNA. The DNA plasmid for the tRNA/synthetase (pAnap) also encodes the gene for G418 resistance so treatment of transfected cells with G418 should result in cell death of any cells that weren't

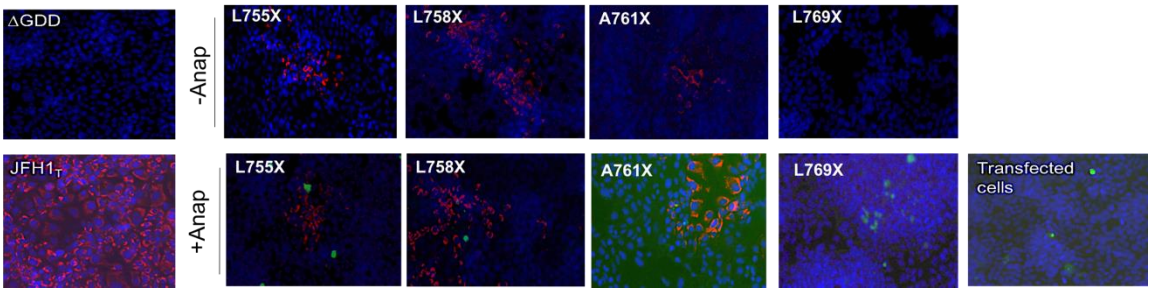

Figure 3.5 Anap fluorescence was detected in transfected Huh-7.5 cells stained for core. To determine if Anap fluorescence is detectable, 200,000 Huh-7.5 cells were plated in 2-well chamber slides in 2 ml $DMEM_{comp}$ and incubated overnight. After 24 hours, Huh-7.5 cells were transfected with pAnap. After 48 hours, pAnap-transfected Huh-7.5 cells were transfected with RNA representing $JFH1_T$, L755X, L758X, A761X, L769X, and ΔGDD. An additional control was included to determine if background Anap was detectable, the 'transfected cells' well was pAnap-transfected cells in the absence of any viral RNA. Forty-eight hours post-RNA transfection, cells were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (red), observed using the 20X objective. The green that was visible was the fluorescence emitted by the unAA, Anap. The pictures shown are representative of the entire area of the well. Results shown are representative of 3 independent experiments.

transfected with the plasmid. To test the impact of G418 on our transfected Huh-7.5 cells, 24 hours post-pAnap transfection, we treated our cells with 75 μg/ml G418 (Thermo Fisher Scientific). We selected this concentration of G418 based on previous work done on Huh-7.5 cells. Twenty-four hours post-treatment with G418 we transfected Huh-7.5 cells with RNA representing L755X, L758X and A761X. Anap was added during the RNA transfection to allow incorporation during translation. Transfected cells were fixed and stained 2 days after the RNA transfection and visualized by indirect immunofluorescence. Low levels of HCV core were detected in L755X and A761X + Anap, without the addition of G418 (Fig 3.6). No Anap fluorescence was detected in either L755X or A761X. Due to an experimental error, we were unable to get an image of L758X + Anap, without the addition of G418. There was an evident reduction in the number of cells present following treatment with G418, however, the treated cells looked unhealthy. Anap fluorescence was detected in L755X, L758X and A761X; however, there was no HCV core detected in any of the samples following the G418 treatment. This result confirmed more work was needed to determine the ideal concentration of G418 to work with.

3.9 G418 Titration Experiment

Despite seeing a reduction in total cells, the G418 treatment seemed to have an impact on the quality of cells. A titration experiment was performed to determine what

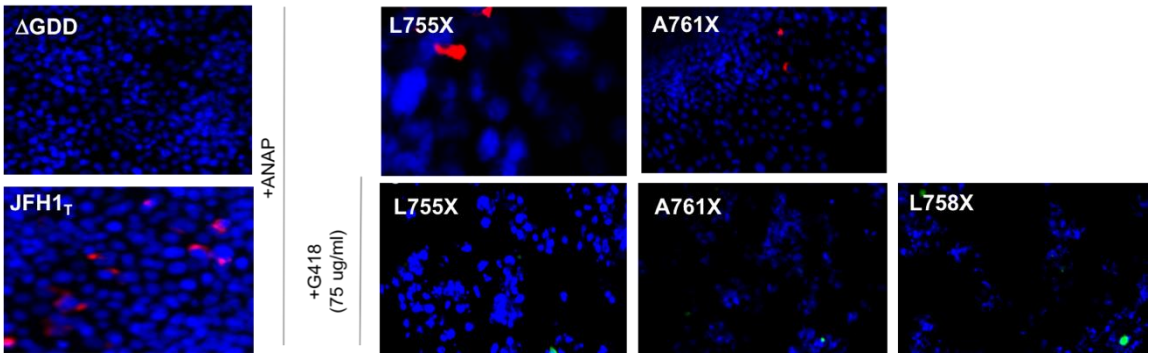

Figure 3.6 Treatment of transfected Huh-7.5 cells with G418. At 24 hours pre-pAnap transfection, 200,000 Huh-7.5 cells were plated in 2-well chamber slides in 2 mL $DMEM_{comp}$ and incubated overnight. After 24 hours, Huh-7.5 cells were transfected with pAnap. At 48 hours, pAnap-transfected Huh-7.5 cells were treated with 75 μg/mL G418 and incubated. After 24-hours, pAnap-transfected, G418 treated cells were transfected with RNA representing L755X, L758X, and A761X. Untransfected, and untreated Huh-7.5 cells were transfected with RNA representing $JFH1_T$ and ΔGDD. Forty-eight hours post-RNA transfection, cells were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (red) and observed using the 20X objective. The green that is visible is the fluorescence emitted by the unAA, Anap. The pictures shown are representative of the entire area of the well. Results shown are indicative of 1 experiment.

concentration of G418 results in 50% Huh-7.5 cell death, in the interest of potentially establishing a cell line stably-transfected with pAnap. To determine what the ideal concentration of G418 was, 100,000 Huh-7.5 cells were plated into each individual well in a 24-well plate. The cells were then treated, in triplicate, with increasing concentrations of G418: 0 μg/mL, 10 μg/mL, 20 μg/mL, 40 μg/mL, 60 μg/mL, 80 μg/mL and 100 μg/mL (Fig. 3.7). The well containing 0 μg/mL is representative of 100% cell growth given that there is no inhibition. Based on the results obtained, we estimated that 50% cell death would be obtained at a G418 concentration of 65 μg/mL. Any future experiments using G418 for selection of Huh-7.5 cells will be done using a concentration of 65 μg/ml.

3.10 Huh-7.5 cells transfected with pAnap and JFH1$_T$

During the treatment of Huh-7.5 cells with G418 experiment, we noted that there was no HCV core detected in any of the samples following the G418 treatment. Given that we had detected HCV core in previous experiments, we considered the possibility that cells that are successfully transfected with pAnap may not be amenable to transfection with RNA representing JFH1$_T$. To determine if it was possible to have pAnap-transfected cells also get transfected with virus, we transfected Huh-7.5 cells with pAnap, and then after 24 hours we transfected those cells with RNA representing wild-type JFH1$_T$. Anap was added during the RNA transfection to demonstrate which cells were successfully transfected with pAnap. The 'transfected cells' well represents a virus negative control, and is representative of typical levels of background incorporation of

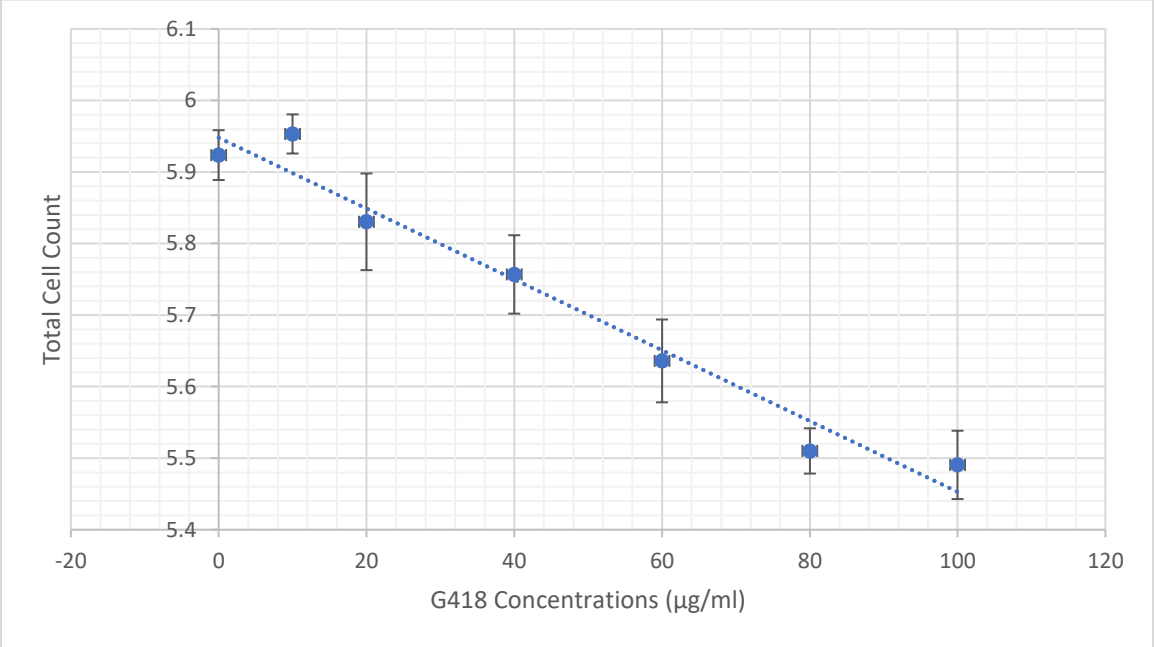

Figure 3.7 G418 Titration. 100,000 Huh-7.5 cells were plated in 1 mL $DMEM_{comp}$ and incubated overnight. After 24 hours the cells were treated, in triplicate, with increasing concentrations of G418: 0 µg/mL, 10 µg/mL, 20 µg/mL, 40 µg/mLl, 60 µg/mL, 80 µg/mL, and 100 µg/mL. After 48 hours, cells were trypsinized and counted.

Anap into cellular proteins. There were comparable levels of RNA transfection when comparing transfection of $JFH1_T$ into Huh-7.5 cells and pAnap-transfected Huh-7.5 cells (Fig. 3.8). Despite the moderate level of RNA transfection, there was no clear overlap between HCV core staining and Anap fluorescence. Based on this result, we concluded that the experiment should be repeated with an additional infection of pAnap transfected cells with $JFH1_T$.

3.11 DNA transfected Huh-7.5 cells can be transfected and infected with $JFH1_T$

Although there was no clear overlap between the Anap fluorescence and HCV core staining during our first trial with $JFH1_T$ transfection, we wanted to determine if pAnap-transfected cells could be infected with wild-type $JFH1_T$. To test this, 200,000 Huh-7.5 cells were plated in 2-well chamber slides. After 24-hours, Huh-7.5 cells were transfected with pAnap and incubated overnight. After 24 hours, pAnap-transfected cells were either transfected with RNA representing $JFH1_T$ or infected with $JFH1_T$. Anap was added during the RNA transfection or infection to demonstrate which cells had been successfully transfected with pAnap. Forty-eight hours later, cells were fixed and stained to be visualized by indirect immunofluorescence. As a control, non-pAnap-transfected Huh-7.5 cells were transfected with ΔGDD and $JFH1_T$ and also infected with $JFH1_T$. There was evident overlap between Anap fluorescence and HCV core staining when pAnap-transfected cells were transfected or infected with $JFH1_T$ (Fig. 3.9). This overlap indicates that the cells that were successfully transfected with pAnap, could also be

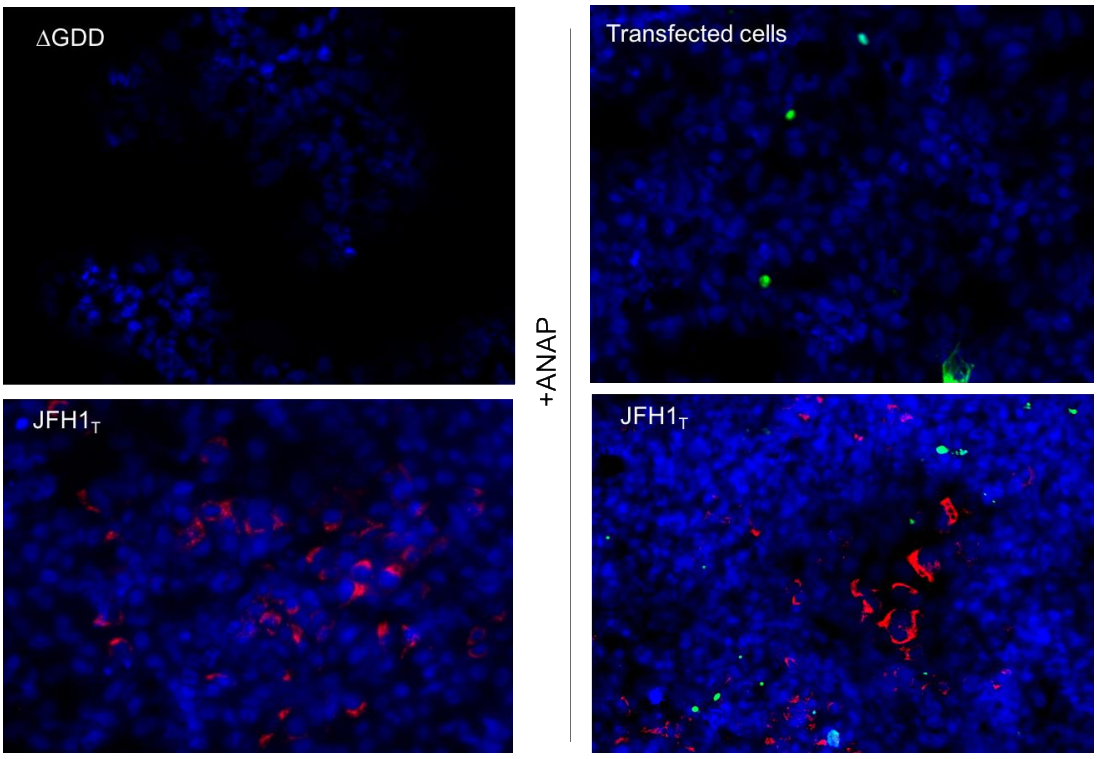

Figure 3.8 DNA transfected Huh-7.5 cells, transfected with WT JFH1T + ANAP. 200,000 Huh-7.5 cells were plated in 2 well chamber slides in 2 mL $DMEM_{comp}$ and incubated overnight. After 24 hours the cells were transfected with pAnap. After 48 hours, we transfected the cells with RNA representing wild-type $JFH1_T$ or ΔGDD. Anap was added during the RNA transfection to identify which cells were successfully transfected with the pAnap. Forty-eight hours post-RNA transfection, cells were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (red) and observed using the 20X objective. The green that is visible is the fluorescence emitted by the unAA, Anap. The pictures shown are representative of the entire well. The results are representative of one experiment.

infected or transfected with $JFH1_T$. This excluded the possibility that there was any mutual exclusion between the virus and pAnap.

3.12 Testing incorporation of Anap into the HCV core protein

One obstacle we encountered when testing incorporation of unAA into the p7 protein in the context of the full virus was that even if unAA was successfully incorporated into p7, we did not know what to expect regarding the localization of p7. It was possible that unAA could get incorporated, but since the normal localization pattern of p7 is not known, we had no way of knowing whether what we were seeing would have been correct. In addition, we were testing incorporation in the context of a full replicating virus, and were thus limited by the fact that mutations could arise eliminating our inserted UAG codon in the viral RNA. Despite this limitation, we saw that an individual cell that was transfected with pAnap could also be transfected or infected with $JFH1_T$. Since mutual exclusion didn't appear to be the main problem, we decided to test Anap incorporation in a single protein expression system. We concluded that using a single-protein expression system would facilitate the selection of potential amenable sites to unAA incorporation since the potential for mutations would be eliminated. Since there is a reliable antibody for core staining, following incorporation of Anap, cells could be stained for core and images overlaid to show that

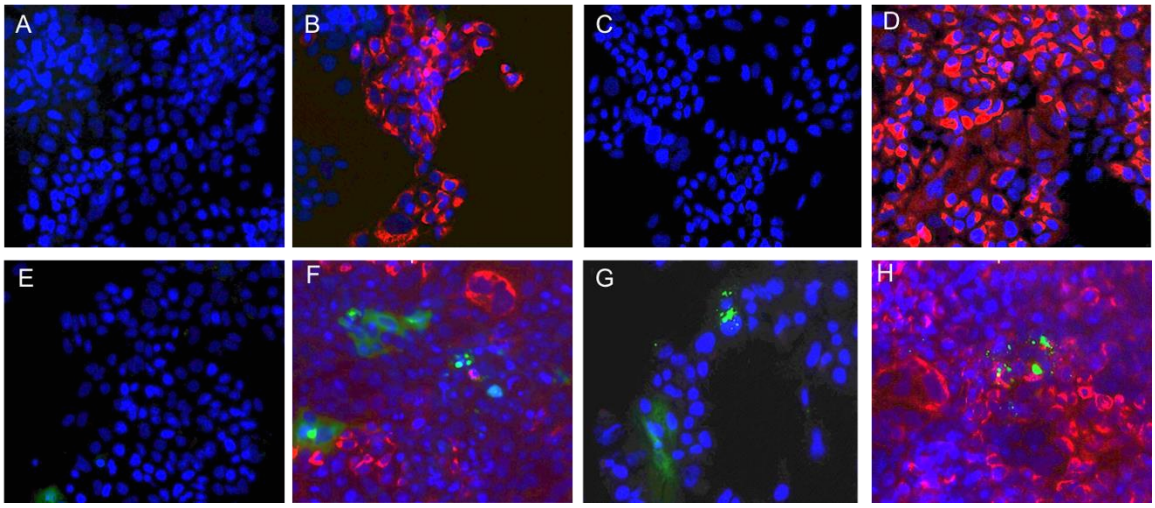

Figure 3.9 DNA transfected Huh-7.5 cells transfected and infected with JFH1$_T$. To set-up, 200,000 Huh-7.5 cells were plated in 2 well chamber slides in 2 mL DMEM$_{comp}$ and incubated overnight. After 24 hours, the cells were transfected with pAnap. After 48 hours various samples of cells were transfected with RNA representing either wild-type JFH1$_T$ or ΔGDD and also infected with JFH1$_T$. Forty-eight hours post-RNA transfection or infection, cells were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (red) and observed using the 20X objective. The green that is visible is the fluorescence emitted by the unAA, Anap. The top row of panels represents the controls for the experiment: panel A is untransfected Huh-7.5 cells; panel B is Huh-7.5 cells infected with JFH1T; panel C is Huh-7.5 cells transfected with ΔGDD and panel D is Huh-7.5 cells transfected with JFH1$_T$. The bottom row of panels are the experimental wells: panel E is pAnap-transfected Huh-7.5 cells; panel F is pAnap-transfected Huh-7.5 cells infected with JFH1$_T$; panel G is pAnap-transfected Huh-7.5 cells transfected with ΔGDD and panel H is pAnap-transfected Huh-7.5 cells transfected with JFH1$_T$. The pictures shown are representative of the entire well.

core staining and Anap fluorescence were in the same location. Given that p7 cannot be easily visualized, trying to incorporate an unAA into the HCV core protein would make the most sense. A major advantage to incorporating Anap into core is that it can be easily confirmed whether Anap disrupts core localization based on where core staining appears since we know where core should be. We postulate that if Anap does not disrupt core localization, it may not affect function either.

3.13 Core Mutant Selection

The first step was the selection of potential sites for incorporation within the HCV core protein. As with p7, we wanted to select sites that could potentially tolerate the substitution of a variety of unAAs. This selection process yielded a panel of 11 mutants that could be used for different unAAs. As previously mentioned, some work was done with AzF and as such, a select number of our mutants had phenylalanine → UAG substitutions. We selected any tryptophan and tyrosine residues to account for the substitution with Anap, given that they are the most structurally similar. To maximize the potential for successful incorporation, we selected residues throughout the core protein. The selected mutations were: F24X, F130X, W76X, W83X, W93X, W96X, W107X, Y35X, Y86X, Y136X, and Y164X. Once the mutants were designed, plasmids were ordered from Genscript.

As part of an Honour's project, Bridgette Green performed some preliminary work with the F24X, F130X, W76X and W86X mutants to optimize the concentration of Anap to be used. It was found that using a concentration of 10 μM was ideal, and all subsequent work using Anap was performed at this optimal concentration.

3.14 Incorporation of Anap into Core Mutants

Since we had shifted to using a single protein expression system, we were able to condense the transfection protocol to a one-day procedure because both pAnap and the core mutant plasmids are DNA plasmids. To test Anap incorporation, 300,000 Huh-7.5 cells were plated in 2 well chamber slides in 2 mL $DMEM_{comp}$ and incubated overnight. After 24 hours, the cells were transfected with both pAnap and individual core mutant plasmids. Wild-type core was used as a positive control for successful protein expression and comparison of core protein localization during this experiment. Two transfection controls were also included during the experimental setup. A negative control, cells only, had the transfection reagents in the absence of any DNA. A positive control, GFP, was included to confirm successful transfection by the presence of green fluorescence. The transfection controls were consistent; the untransfected cells showed no fluorescence detected, and the GFP control shows green fluorescence indicating a successful transfection (Fig 3.10a). The wild-type core sample had high levels of core staining detected in the absence and presence of Anap. In order to ensure all samples were treated the same way, the Huh-7.5 cells to be transfected with wild-type core were also transfected with pAnap and incubated in the presence and absence of Anap. We expected this high expression in both absence and presence of Anap given that translation of the wild-type core plasmid is independent of Anap incorporation since there was no specific mutation site included in the protein sequence. Images of the tryptophan mutants that we screened for any amenable positions to unAA incorporation indicate successful incorporation (Fig 3.10a). All of the tryptophan mutants that we tested for Anap

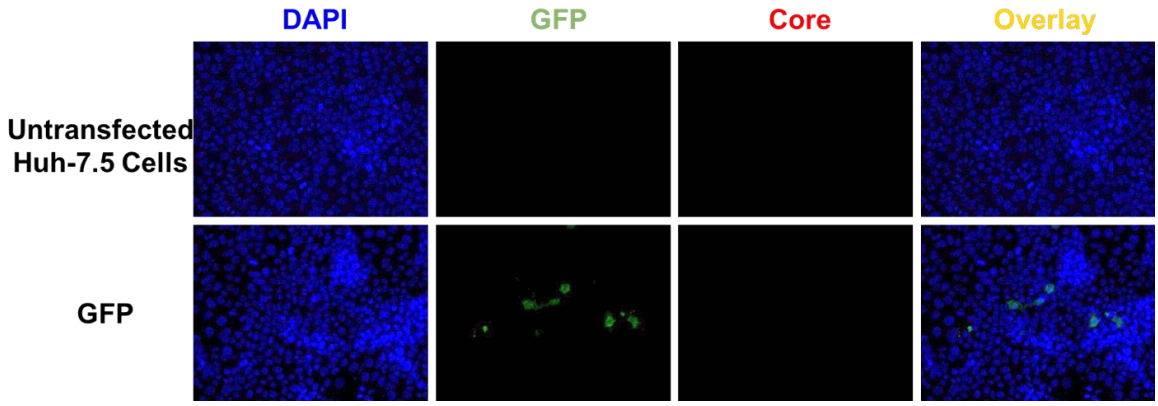

Figure 3.10a Incorporation of Anap into core mutants: transfection controls. To set-up, 300,000 Huh-7.5 cells were plated in 2 well chamber slides in 2 mL $DMEM_{comp}$ and incubated at 37°C in 5% CO_2. After 24 hours, as controls, cells were either transfected with the plasmid containing GFP or mock transfected in the absence of DNA plasmid. Forty-eight hours post-transfection, slides were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (red) and observed using the 20X objective. The green that is visible is the fluorescence emitted by GFP.

incorporation were successful in expressing core (Fig. 3.10b). In W83X, W93X and W107X, the HCV core protein was only detected in the presence of Anap. The red staining shows the typical perinuclear staining of core and the Anap fluorescence co-localizes with it indicating that the unAA is actually being incorporated into the core protein. There was very clear overlap of the green fluorescence from Anap and the core staining in W93X and W107X in the presence of Anap. An important point to note here is that, although the emission is faint in W83X and W96X, Anap fluorescence was detected in each of the tryptophan mutants, and overlapped with the HCV core staining, indicating successful Anap incorporation. W96X was positive in the presence and absence of Anap. This was not expected, as expression of the protein should not occur in the absence of Anap as translation should terminate at the unique amber stop codon. We postulate that this might have been an instance of contamination. As seen with the tryptophan mutants, Y35X, Y136X and Y164X in the presence of Anap showed successful incorporation of Anap into the core protein (Fig. 3.10c). There is very clear co-localization of the green fluorescence from Anap and the core staining in Y35X and Y136X in the presence of Anap. The fluorescence of Anap is not evident in Y164X in the overlay image, however it is evident in the panel containing only Anap that it is present. The images shown in Fig. 3.10b and c are all from the same experiment, but results were presented as groups based on the original amino acid substituted. The images generated from the wild-type core-

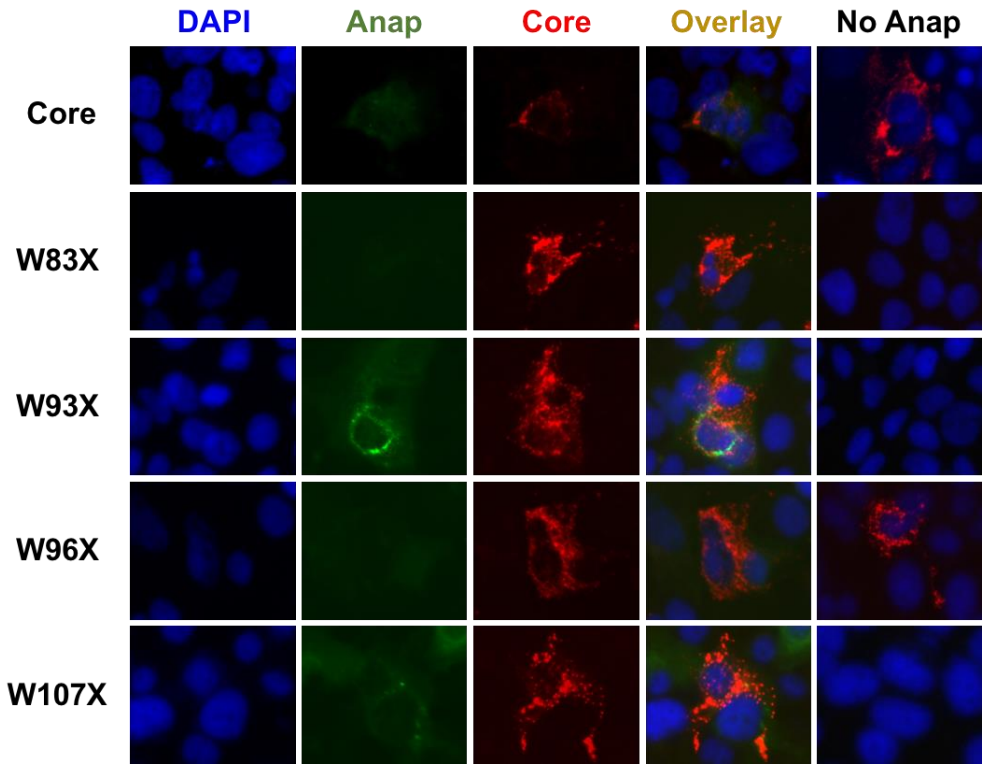

Figure 3.10b Incorporation of Anap into core mutants: tryptophan mutants. To test incorporation into the HCV core proteins, 300,000 Huh-7.5 cells were plated in 2 well chamber slides in 2 mL $DMEM_{comp}$ and incubated overnight. After 24 hours, cells were transfected with pAnap, along with the plasmid encoding W83X, W93X, W96X, and W107X core mutants. In addition, we also transfected cells with wild-type core as a positive control. As a negative control, each mutant was transfected in the absence of Anap, which is shown on the far right as a 'no Anap' panel; this panel is an overlay image of DAPI, Anap and core staining. Forty-eight hours after the transfection slides were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (red). The green that is visible is the fluorescence emitted by Anap. Results shown are representative of 2 independent experiments.

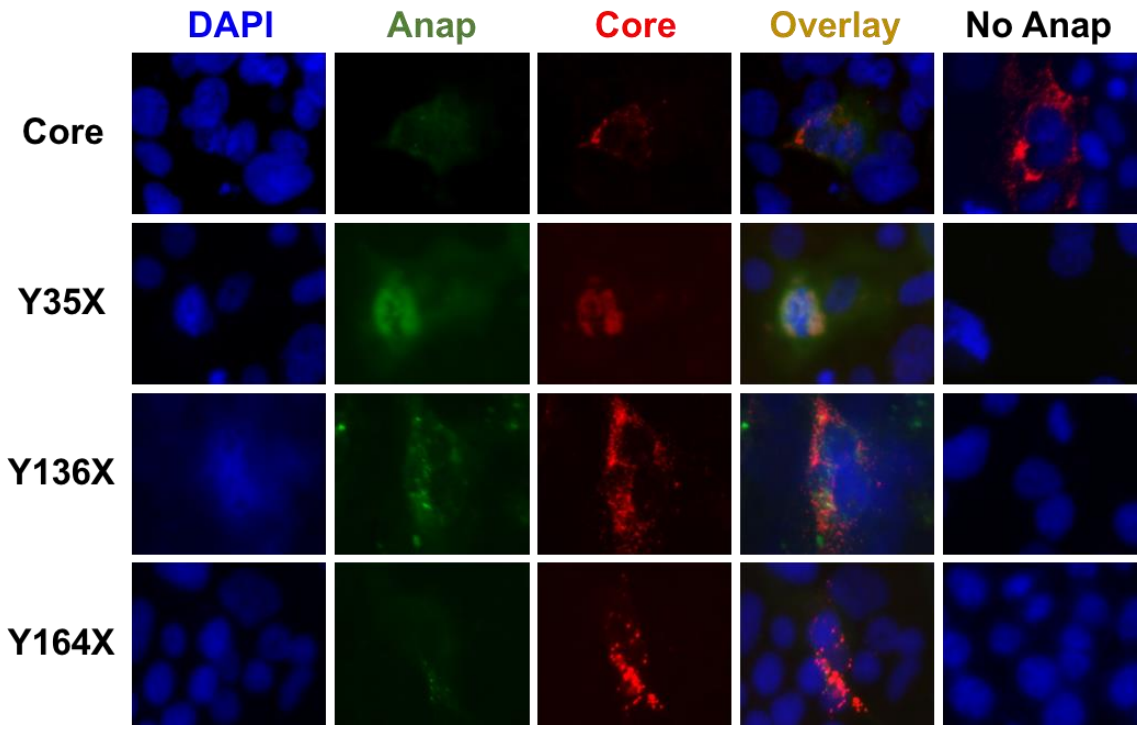

Figure 3.10c Incorporation of Anap into core mutants: tyrosine mutants. To test Anap incorporation into the HCV core protein, 300,000 Huh-7.5 cells were plated in 2 well chamber slides in 2 mL $DMEM_{comp}$ and incubated overnight. After 24 hours, cells were transfected with pAnap, along with the plasmids encoding Y35X, Y136X, and Y164X core mutants. Cells were transfected with wild-type core as a positive control. As a negative control, each mutant was transfected in the absence of Anap, which is shown on the far right as a 'no Anap' panel; this panel is an overlay image of DAPI, Anap and core staining. Forty-eight post-transfection slides were fixed and probed with DAPI (blue) and a specific HCV core antibody, by indirect immunofluorescence (red). The green that is visible is the fluorescence emitted by Anap. Results shown are representative of 2 independent experiments.

transfected cells were duplicated in the figures. Based on these results, we concluded that it was possible to incorporate Anap into the HCV core protein.

3.15 Core Positive Cells Detected

In order to return to the context of the full virus with a potential position that is likely to work, we needed to develop a way to determine which mutants were most successful. In order to accomplish this, we counted each cell that was positive for core and checked to see if that cell also exhibited green fluorescence indicating the presence of Anap to note any discrepancies (Fig. 3.11). We believe there might be some cases where the core staining was blocking or obstructing the Anap fluorescence. With the exception of the wild-type core, W96X + Anap had the highest number of core positive cells. The lowest count of core positive cells was detected in Y35X in the presence of Anap. W83X, W93X, W107X, Y136X and Y164X, in the presence of Anap, all had comparable levels of core positive cells. During the counting process, the W96X -Anap core positive cells could not be located, potentially due to obstruction from older Vectashield or DAPI stain diffusing out. Based on the results, W83X, W96X, and Y164X might be most likely to work in the context of the full virus.

Chapter 4: Discussion

The role of the HCV p7 protein remains controversial due in part to the limitations of current protein labelling methods. A novel method of protein labelling could be achieved through the incorporation of an unAA into the target protein. As mentioned above,

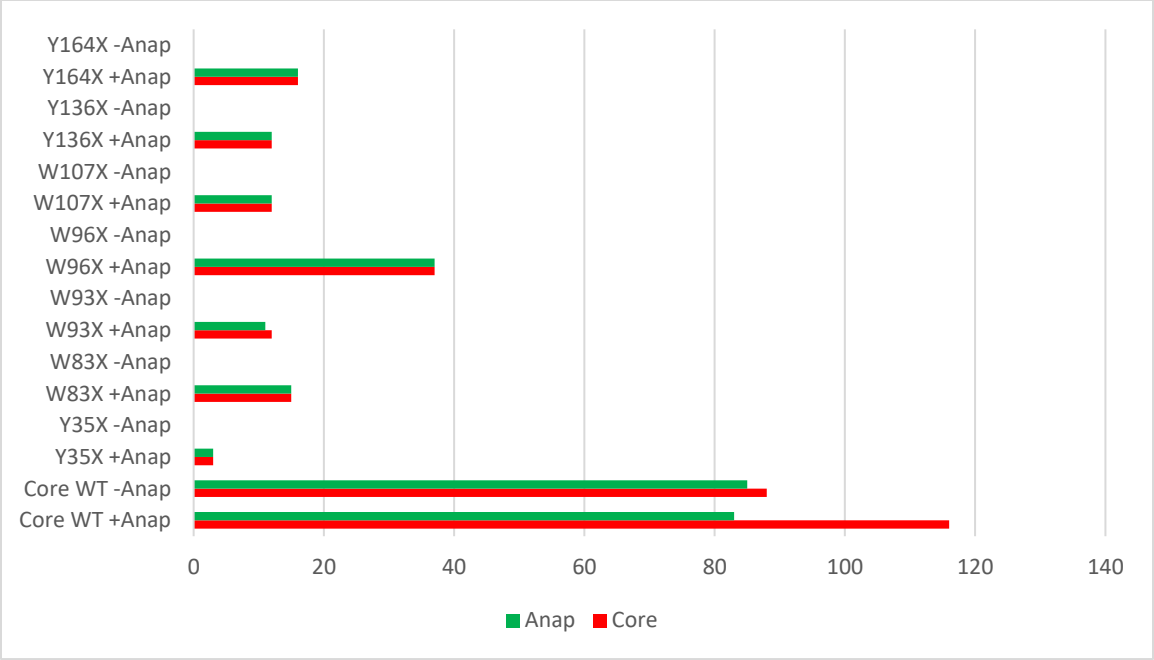

Figure 3.11 Relationship between Anap and Detected HCV core staining.

Immunofluorescence slides from the previous experiment were scored for the number of HCV core-positive cells (red bar), as well as core-positive cells with corresponding Anap fluorescence (green bar) representing potential contamination or read-through.

unAAs have been used to expand the genetic code in a number of organisms[137]. Different groups have utilized this technique in a variety of applications including probing protein-protein interactions, visualization of target proteins, and controlling protein functions in live vertebrates[137]. Although this technique has been extensively used, it has not yet been applied for visualization of viral proteins[137]. The rationale behind our experiments was that incorporation of an unAA that could have a fluorophore attached, or that is inherently fluorescent, would allow visualization of the protein of interest.

Although we selected a panel of 14 mutants to screen for their amenability to unAA incorporation, site-directed mutagenesis only yielded 4 successfully cloned mutants. We performed numerous attempts to clone the additional 10 mutants, however, we were unsuccessful. Despite this setback we attempted incorporation of an unAA using the 4 mutants we had cloned, given that a variety of residues have been substituted with unAAs in the literature. The 4 mutants that were successfully cloned were L755X, L758X, A761X, and L769X. We used these 4 mutants to attempt unAA incorporation of AzF into the p7 protein in the context of a replicating virus. We selected AzF for our preliminary trials given that it has a chemically active side chain that can be utilized in a number of different reactions. One potential use for AzF is probing protein-protein interactions using the photocrosslinking capabilities of AzF, an aryl-azide. Another potential application of AzF is visualizing a target protein; a click-chemistry reaction to attach a fluorophore directly onto the unAA could be used for localization. This click-chemistry reaction would be accomplished via an azide-alkyne cycloaddition; the azide is provided by the unAA, and the alkyne would be from an alkyne-dye, such as a

fluorophore. Before we could utilize AzF for a click-chemistry reaction, we had to first successfully incorporate AzF into p7. Our first trial suggested that we had successful incorporation of the AzF given that we saw core-positive cells in L758X, and A761X (Fig. 3.2). The first trial had one major drawback which we addressed in all subsequent experiments; we hadn't included a control for the unAA incorporation system. For future experiments, we realized we would need to include a control per mutant, whereby the cells would be transfected with the specific DNA plasmid (pMah or pAnap), followed by the transfection of the viral RNA, in the absence of the unAA. This would provide a system negative control, since translation of the polyprotein should terminate when the unique amber stop codon in p7 is reached. Core-positive cells should not be detected in the absence of the unAA.

Based on our preliminary screen with AzF in p7, we wanted to determine what effect AzF had on virus production (Fig. 3.3). Although virus was produced in L755X +AzF and L758X +AzF, it was evident that AzF was negatively impacting the degree of virus production when compared with $JFH1_T$. In addition, virus was produced in L755X, L758X, and A761X in the absence of AzF. These results were not expected and unfortunate, given that virus was being produced in the mutants without AzF added, which should not occur. As previously mentioned, translation should terminate at the amber stop codon in the absence of the unAA resulting in no virus production. At this point, we realized that reversion or translational read-through might be occurring. HCV is a highly mutatable virus, and it was possible that the virus was mutating in order to eliminate the stop codon. If mutations were occurring to alter the unique codon to be

specific for one of the 20 natural amino acids, translation would no longer be dependent on incorporation of the unAA. Another possible explanation for the core-positive cells detected would be read-through. Read-through is when the ribosomes misread the termination signal provided by a stop codon and continue translation past the stop codon. Although read-through events are possible, they are less likely to occur than reversion. Also, the degree of read-through necessary to account for the discrepancies seen is higher than we would expect to occur.

To overcome this obstacle, we decided to further screen the available positions using a fluorescent unAA, Anap[118]. Utilizing a fluorescent unAA would allow us to distinguish between successful incorporation and reversion/read-through events, since we would expect to see overlap of fluorescence from Anap and staining of core-positive cells when successful incorporation occurred. We demonstrated that Anap fluorescence could be detected in Huh-7.5 cells following transfections with the DNA plasmid for incorporation and the virus encoded RNA (Fig. 3.4). During the first screening trial with Anap, we did not stain for HCV core given that the secondary antibody normally used in our lab would fluoresce at the same wavelength as Anap. During the next trial, we screened the same four mutants and used a HCV core specific secondary antibody that would be detectable at a different wavelength than Anap (Fig. 3.5). We demonstrated that core-positive cells were detected, however there was no clear overlap between the core-positive cells and Anap fluorescence. L755X, L758X, and A761X were all positive for virus production in both the presence and absence of Anap, indicating that replication was not dependent on incorporation of the unAA. There was no virus detected in L769X

in any of our trials, so we ceased testing incorporation with this mutant in subsequent experiments given its consistent negative results.

In order for replication of the mutants to occur, the transfected viral RNA must enter a cell that was successfully transfected with pAnap. In order to increase the chances of finding an event where Anap fluorescence is detected in a core-positive cell, we selected for cells that were successfully transfected with pAnap using G418. The pAnap plasmid encodes for Geneticin® (G418) resistance, therefore it should be possible to eliminate any nontransfected cells using G418 treatment. We hypothesized that treating the cells with G418 would result in cell death of any cells not transfected with pAnap. In theory, treating the cells with G418 would be beneficial since the only cells available to be transfected with the viral RNA would already be transfected with pAnap. Unfortunately, no core-positive cells were detected following treatment with G418 and the cells did not appear to be healthy (Fig. 3.6). Before continuing with G418 treatment, we determined what concentration was optimal for use with Huh-7.5 cells (Fig. 3.7). We carried a titration experiment to establish that for the Huh-7.5 cells used in our lab, the ideal concentration of G418 to use is 65 μg/mL (Fig 3.7).

Since we did not detect any core-positive cells following treatment with G418, we considered the possibility that there was mutual exclusion between pAnap and the viral-encoded RNA. We postulated that it was possible that once a cell was transfected with the plasmid for Anap incorporation, that cell was no longer susceptible to HCV infection. To refute this possibility, we transfected cells with pAnap, followed by transfection with RNA representing wild-type $JFH1_T$ (Fig. 3.8). In the first trial with transfected cells

subsequently transfected with $JFH1_T$, there were no cells that stained positive for core which also had detectable fluorescence from Anap. We attributed this absence of overlap between core-staining and Anap fluorescence to a relatively low transfection efficiency with the virus and repeated the experiment. In the second trial, we also included an infection of the transfected cells to eliminate the possibility of mutual exclusion. We demonstrated that pAnap transfected Huh-7.5 cells could also be transfected or infected with $JFH1_T$ (Fig 3.9). This confirmed that the virus could enter and replicate within a cell that had been transfected with the pAnap.

Given that our issues with incorporation were not due to mutual exclusion between the virus and the DNA plasmid, we decided we would optimize the system using a viral protein in a single protein expression system before returning to the full virus. This technique would allow us to test incorporation into a viral protein without the mutational pressure of the virus. Since there is no reliable antibody for immunofluorescence studies with p7, we decided that core would be an ideal target for optimization in a single protein expression system. Successful incorporation of the unAA would be indicated by an overlap of the core-staining and the fluorescence emitted by Anap. We selected a panel of 11 potential sites for incorporation and had the DNA plasmids synthesized. Some preliminary work was done by another student in the lab using 4 of the 11 core mutants. Following success with the first 4 mutants tested (F24X, W76X, Y86X, and F130X), we screened the remaining 7 mutants for their amenability to incorporate Anap at the selected positions. We demonstrated that Anap could be successfully incorporated into Y35X, W83X, W93X, W96X, W107X, Y136X, and

Y164X in the presence of Anap (Fig. 3.10a, b and c). No core-positive cells were detected in the absence of Anap, except in the case of W96X. Two core-positive cells were detected in W96X; which may have been due to contamination or read-through. No core-positive cells were detected in W96X in the absence of Anap in subsequent experiments (not shown in this thesis) supporting the belief that contamination may have occurred in the first trial. In addition to the 7 mutants that we screened, core-positive cells were also detected in the presence of Anap in the remaining 4 mutants: F24X, W76X, Y86X and F130X (Fig. 3.10 a, b, and c).

Since we demonstrated successful incorporation of the unAA, Anap, into the HCV core protein, we evaluated the mutants based on their transfection efficiency to elucidate which unAAs might be ideal candidates for trials in the context of the full virus. We plotted the relationship between detection of Anap and the presence of a core-positive cell (Fig. 3.11). We postulated that the mutants with the highest number of core-positive cells might be ideal target positions for attempting incorporation of Anap into core in the context of the replicating virus. Based on these criteria, W83X, W96X, and Y164X might be more likely to work when we return to the full virus based on the number of core-positive cells that were detected. Although the number of core-positive cells may provide some indication as to which mutants will be successful in the virus, it is more likely that structure will have the greatest impact. Since we can't be certain that the number of core-positive cells is reflective of successful incorporation in the virus, we will attempt to incorporate Anap into all 11 core mutants in the full virus.

If our attempts at incorporation of Anap into HCV core protein in the full virus are successful, we will be able to contribute to the growing body of research about the practical applications of unAA incorporation. This would also provide a novel way to label viral proteins in host cells. Thus far, unAAs have been used in many different ways[139]. One study utilized Anap incorporation to monitor the voltage-dependent motion of the catalytic region of voltage-sensing phosphatase in Xenopus oocytes[123]. Although this example of Anap incorporation was not done in the context of a virus, it supports the fact that Anap can be very useful in localization studies.

4.1 Future Applications

Once we have optimized the system so that Anap can be incorporated into HCV proteins in the full virus, we can expand this method to include other viruses as well. Theoretically, this technique could be applied to any viral protein and could be used for localization and co-localization studies. One interesting application using fluorescent unAAs would be incorporating multiple different fluorescent unAAs into different viral proteins. One group has proved that unAA incorporation into viruses is possible through the establishment of a system to control multicycle replication of live-attenuated HIV-1 by incorporating AzF[138]. The Guo group sought to improve the safety of live-attenuated vaccines through the incorporation of AzF into selected sites in HIV-1. They constructed all-in-one live-attenuated HIV-1 mutants that contained genomic copies of the amber suppression system necessary for unAA incorporation to impart the ability to "turn on" or "turn off" the virus. They were successful in establishing a system that incorporated AzF at various locations. Although, their system and ours vary in the specific aim of study,

this research paper provides proof of concept that unAAs can be incorporated into a virus. Another major difference between the two applications is that their success was based on cessation of virus production in the absence of AzF, whereas we are aiming to use the unAAs as a potential tool for viral protein labelling. This group was not interested in the localization or protein function applications of unAA incorporation.

Although we have faced some challenges in developing our system of unAA incorporation, we remain optimistic that the system will ultimately be used successfully. Some of our main concerns have been based on the known low transfection efficiency of Huh-7.5 cells. The Guo group had one major advantage when working with HIV since they were able to utilize HEK 293T cells, which are known to have high transfection efficiency[136]. This would facilitate successful incorporation of an unAA since more cells would be transfected with pMah. Despite this success with unAA incorporation in HIV, unAA incorporation has yet to be used as a tool for labelling viral proteins in host cells. We believe that using a fluorescent unAA for visualiziation of viral proteins would be a major breakthrough in the field of virology since it would eliminate the need for a protein-specific antibody.

Chapter 5: References

1. Lavanchy, D. Evolving epidemiology of hepatitis C virus. *Clin. Microbiol. Infect.* **17,** 107–115 (2011).
2. Gower, E., Estes, C., Blach, S., Razavi-Shearer, K. & Razavi, H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J. Hepatol.* **61,** S45–S57 (2014).
3. Petruzziello, A., Marigliano, S., Loquercio, G., Cozzolino, A. & Cacciapuoti, C. Global epidemiology of hepatitis C virus infection: An up-date of the distribution and circulation of hepatitis C virus genotypes. *World J. Gastroenterol.* **22,** 7824–40 (2016).
4. Gower, E., Estes, C., Blach, S., Razavi-Shearer, K. & Razavi, H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J. Hepatol.* **61,** S45–S57 (2014).
5. Mohd Hanafiah, K., Groeger, J., Flaxman, A. D. & Wiersma, S. T. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology* **57,** 1333–42 (2013).
6. Wedemeyer, H., Dore, G. J. & Ward, J. W. Estimates on HCV disease burden worldwide - filling the gaps. *J. Viral Hepat.* **22 Suppl 1,** 1–5 (2015).
7. Vieyres, G. *et al.* Subcellular localization and function of an epitope-tagged p7 viroporin in hepatitis C virus-producing cells. *J. Virol.* **87,** 1664–78 (2013).
8. Frank, C. *et al.* The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet (London, England)* **355,** 887–91 (2000).
9. Zou, S., Tepper, M. & El Saadany, S. Prediction of hepatitis C burden in Canada. *Can. J. Gastroenterol.* **14,** 575–80
10. van Buuren, N. *et al.* The 5th Canadian Symposium on Hepatitis C Virus: We Are Not Done Yet-Remaining Challenges in Hepatitis C. *Can. J. Gastroenterol. Hepatol.* **2016,** 7603526 (2016).
11. van Buuren, N. *et al.* The 5th Canadian Symposium on Hepatitis C Virus: We Are Not Done Yet-Remaining Challenges in Hepatitis C. *Can. J. Gastroenterol. Hepatol.* **2016,** 7603526 (2016).
12. Pouget, E. R., Hagan, H. & Des Jarlais, D. C. Meta-analysis of hepatitis C seroconversion in relation to shared syringes and drug preparation equipment. *Addiction* **107,** 1057–65 (2012).
13. Werb, D. *et al.* Treatment costs of hepatitis C infection among injection drug users in Canada, 2006-2026. *Int. J. Drug Policy* **22,** 70–6 (2011).
14. Shepard, C. W., Finelli, L. & Alter, M. J. Global epidemiology of hepatitis C virus infection. *Lancet. Infect. Dis.* **5,** 558–67 (2005).
15. Choo, Q. L. *et al.* Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244,** 359–62 (1989).
16. Kalita, M. M., Griffin, S., Chou, J. J. & Fischer, W. B. Genotype-specific differences in structural features of hepatitis C virus (HCV) p7 membrane protein. *Biochim. Biophys. Acta* **1848,** 1383–92 (2015).
17. Murphy, D. G. *et al.* Hepatitis C virus genotype 7, a new genotype originating from central Africa. *J. Clin. Microbiol.* **53,** 967–72 (2015).
18. Martell, M. *et al.* Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J. Virol.* **66,** 3225–9 (1992).

19. Grebely, J. *et al.* Hepatitis C virus clearance, reinfection, and persistence, with insights from studies of injecting drug users: towards a vaccine. *Lancet Infect. Dis.* **12,** 408–414 (2012).
20. Cho, Y.-K., Kim, Y. N. & Song, B.-C. Predictors of spontaneous viral clearance and outcomes of acute hepatitis C infection. *Clin. Mol. Hepatol.* **20,** 368 (2014).
21. Thimme, R. *et al.* Determinants of viral clearance and persistence during acute hepatitis C virus infection. *J. Exp. Med.* **194,** 1395–406 (2001).
22. Orland, J. R., Wright, T. L. & Cooper, S. Acute hepatitis C. *Hepatology* **33,** 321–7 (2001).
23. Micallef, J. M., Kaldor, J. M. & Dore, G. J. Spontaneous viral clearance following acute hepatitis C infection: a systematic review of longitudinal studies. *J. Viral Hepat.* **13,** 34–41 (2006).
24. Hoofnagle, J. H. Hepatitis C: The clinical spectrum of disease. *Hepatology* **26,** 15S–20S (1997).
25. Chen, S. L. & Morgan, T. R. The natural history of hepatitis C virus (HCV) infection. *Int. J. Med. Sci.* **3,** 47–52 (2006).
26. Beames, B., Chavez, D. & Lanford, R. E. GB virus B as a model for hepatitis C virus. *ILAR J.* **42,** 152–60 (2001).
27. Perz, J. F., Armstrong, G. L., Farrington, L. A., Hutin, Y. J. F. & Bell, B. P. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J. Hepatol.* **45,** 529–38 (2006).
28. Donato, F., Boffetta, P. & Puoti, M. A meta-analysis of epidemiological studies on the combined effect of hepatitis B and C virus infections in causing hepatocellular carcinoma. *Int. J. cancer* **75,** 347–54 (1998).
29. Wang, C., Sarnow, P. & Siddiqui, A. Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *J. Virol.* **67,** 3338–44 (1993).
30. Friebe, P. & Bartenschlager, R. Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. *J. Virol.* **76,** 5326–38 (2002).
31. Yi, M. & Lemon, S. M. 3' nontranslated RNA signals required for replication of hepatitis C virus RNA. *J. Virol.* **77,** 3557–68 (2003).
32. Reed, K. E. & Rice, C. M. Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties. *Curr. Top. Microbiol. Immunol.* **242,** 55–84 (2000).
33. Lindenbach, B. D. & Rice, C. M. The ins and outs of hepatitis C virus entry and assembly. *Nat. Rev. Microbiol.* **11,** 688–700 (2013).
34. Catanese, M. T. *et al.* Ultrastructural analysis of hepatitis C virus particles. *Proc. Natl. Acad. Sci.* **110,** 9505–9510 (2013).
35. Carrick, R. J., Schlauder, G. G., Peterson, D. A. & Mushahwar, I. K. Examination of the buoyant density of hepatitis C virus by the polymerase chain reaction. *J. Virol. Methods* **39,** 279–89 (1992).
36. Miyamoto, H., Okamoto, H., Sato, K., Tanaka, T. & Mishiro, S. Extraordinarily low density of hepatitis C virus estimated by sucrose density gradient centrifugation and the polymerase chain reaction. *J. Gen. Virol.* **73 (Pt 3),** 715–8 (1992).
37. Kanto, T. *et al.* Buoyant density of hepatitis C virus recovered from infected hosts: two different features in sucrose equilibrium density-gradient centrifugation related to degree of liver inflammation. *Hepatology* **19,** 296–302 (1994).

38. Agnello, V., Abel, G., Elfahal, M., Knight, G. B. & Zhang, Q. X. Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 12766–71 (1999).
39. Atoom, A. M., Taylor, N. G. A. & Russell, R. S. The elusive function of the hepatitis C virus p7 protein. *Virology* **462–463,** 377–387 (2014).
40. Pileri, P. *et al.* Binding of hepatitis C virus to CD81. *Science* **282,** 938–41 (1998).
41. Scarselli, E. *et al.* The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J.* **21,** 5017–25 (2002).
42. Germi, R. *et al.* Cellular glycosaminoglycans and low density lipoprotein receptor are involved in hepatitis C virus adsorption. *J. Med. Virol.* **68,** 206–15 (2002).
43. Evans, M. J. *et al.* Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* **446,** 801–5 (2007).
44. Liu, S. *et al.* Tight junction proteins claudin-1 and occludin control hepatitis C virus entry and are downregulated during infection to prevent superinfection. *J. Virol.* **83,** 2011–4 (2009).
45. Meertens, L., Bertaux, C. & Dragic, T. Hepatitis C virus entry requires a critical postinternalization step and delivery to early endosomes via clathrin-coated vesicles. *J. Virol.* **80,** 11571–8 (2006).
46. Tscherne, D. M. *et al.* Time- and temperature-dependent activation of hepatitis C virus for low-pH-triggered entry. *J. Virol.* **80,** 1734–41 (2006).
47. Codran, A. *et al.* Entry of hepatitis C virus pseudotypes into primary human hepatocytes by clathrin-dependent endocytosis. *J. Gen. Virol.* **87,** 2583–93 (2006).
48. Blanchard, E. *et al.* Hepatitis C Virus Entry Depends on Clathrin-Mediated Endocytosis. *J. Virol.* **80,** 6964–6972 (2006).
49. Sharma, N. R. *et al.* Hepatitis C virus is primed by CD81 protein for low pH-dependent fusion. *J. Biol. Chem.* **286,** 30361–76 (2011).
50. Krey, T. *et al.* The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule. *PLoS Pathog.* **6,** e1000762 (2010).
51. Li, Y., Wang, J., Kanai, R. & Modis, Y. Crystal structure of glycoprotein E2 from bovine viral diarrhea virus. *Proc. Natl. Acad. Sci.* **110,** 6805–6810 (2013).
52. El Omari, K., Iourin, O., Harlos, K., Grimes, J. M. & Stuart, D. I. Structure of a pestivirus envelope glycoprotein E2 clarifies its role in cell entry. *Cell Rep.* **3,** 30–5 (2013).
53. Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. & Rice, C. M. Expression and identification of hepatitis C virus polyprotein cleavage products. *J. Virol.* **67,** 1385–95 (1993).
54. Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. & Shimotohno, K. Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 5547–51 (1991).
55. Kim, J. L. *et al.* Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **87,** 343–55 (1996).
56. Hijikata, M. *et al.* Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. *J. Virol.* **67,** 4665–75 (1993).
57. Romero-Brey, I. *et al.* Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. *PLoS Pathog.* **8,** e1003056 (2012).

58. Paul, D. & Bartenschlager, R. Architecture and biogenesis of plus-strand RNA virus replication factories. *World J. Virol.* **2,** 32–48 (2013).
59. Romero-Brey, I. & Bartenschlager, R. Membranous replication factories induced by plus-strand RNA viruses. *Viruses* **6,** 2826–57 (2014).
60. Egger, D. *et al.* Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J. Virol.* **76,** 5974–84 (2002).
61. Lohmann, V., Körner, F., Herian, U. & Bartenschlager, R. Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J. Virol.* **71,** 8416–28 (1997).
62. Lohmann, V., Overton, H. & Bartenschlager, R. Selective stimulation of hepatitis C virus and pestivirus NS5B RNA polymerase activity by GTP. *J. Biol. Chem.* **274,** 10807–15 (1999).
63. Lohmann, V., Roos, A., Körner, F., Koch, J. O. & Bartenschlager, R. Biochemical and kinetic analyses of NS5B RNA-dependent RNA polymerase of the hepatitis C virus. *Virology* **249,** 108–18 (1998).
64. Behrens, S. E., Tomei, L. & De Francesco, R. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* **15,** 12–22 (1996).
65. Gu, M. & Rice, C. M. Structures of hepatitis C virus nonstructural proteins required for replicase assembly and function. *Curr. Opin. Virol.* **3,** 129–136 (2013).
66. Bartenschlager, R., Frese, M. & Pietschmann, T. Novel Insights into Hepatitis C Virus Replication and Persistence. in *Advances in virus research* **63,** 71–180 (2004).
67. Appel, N., Schaller, T., Penin, F. & Bartenschlager, R. From Structure to Function: New Insights into Hepatitis C Virus RNA Replication. *J. Biol. Chem.* **281,** 9833–9836 (2006).
68. Barba, G. *et al.* Hepatitis C virus core protein shows a cytoplasmic localization and associates to cellular lipid storage droplets. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 1200–5 (1997).
69. Boulant, S., Targett-Adams, P. & McLauchlan, J. Disrupting the association of hepatitis C virus core protein with lipid droplets correlates with a loss in production of infectious virus. *J. Gen. Virol.* **88,** 2204–13 (2007).
70. Boulant, S., Vanbelle, C., Ebel, C., Penin, F. & Lavergne, J.-P. Hepatitis C virus core protein is a dimeric alpha-helical protein exhibiting membrane protein features. *J. Virol.* **79,** 11353–65 (2005).
71. Boulant, S. *et al.* Structural determinants that target the hepatitis C virus core protein to lipid droplets. *J. Biol. Chem.* **281,** 22236–47 (2006).
72. Miyanari, Y. *et al.* The lipid droplet is an important organelle for hepatitis C virus production. *Nat. Cell Biol.* **9,** 1089–1097 (2007).
73. Dubuisson, J. *et al.* Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and Sindbis viruses. *J. Virol.* **68,** 6147–60 (1994).
74. Phan, T., Beran, R. K. F., Peters, C., Lorenz, I. C. & Lindenbach, B. D. Hepatitis C virus NS2 protein contributes to virus particle assembly via opposing epistatic interactions with the E1-E2 glycoprotein and NS3-NS4A enzyme complexes. *J. Virol.* **83,** 8379–95 (2009).
75. Yi, M., Ma, Y., Yates, J. & Lemon, S. M. Compensatory mutations in E1, p7, NS2,

and NS3 enhance yields of cell culture-infectious intergenotypic chimeric hepatitis C virus. *J. Virol.* **81,** 629–38 (2007).
76. Popescu, C.-I. *et al.* NS2 protein of hepatitis C virus interacts with structural and non-structural proteins towards virus assembly. *PLoS Pathog.* **7,** e1001278 (2011).
77. Atoom, A. M., Jones, D. M. & Russell, R. S. Evidence suggesting that HCV p7 protects E2 glycoprotein from premature degradation during virus production. *Virus Res.* **176,** 199–210 (2013).
78. Griffin, S. D. C. *et al.* The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett.* **535,** 34–8 (2003).
79. Jones, C. T., Murray, C. L., Eastman, D. K., Tassello, J. & Rice, C. M. Hepatitis C Virus p7 and NS2 Proteins Are Essential for Production of Infectious Virus. *J. Virol.* **81,** 8374–8383 (2007).
80. Phan, T., Kohlway, A., Dimberu, P., Pyle, A. M. & Lindenbach, B. D. The acidic domain of hepatitis C virus NS4A contributes to RNA replication and virus particle assembly. *J. Virol.* **85,** 1193–204 (2011).
81. Appel, N. *et al.* Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. *PLoS Pathog.* **4,** e1000035 (2008).
82. Masaki, T. *et al.* Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. *J. Virol.* **82,** 7964–76 (2008).
83. Scheel, T. K. H. & Rice, C. M. Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nat. Med.* **19,** 837–49 (2013).
84. Hoofnagle, J. H. *et al.* Transmission of non-A, non-B hepatitis. *Ann. Intern. Med.* **87,** 14–20 (1977).
85. Hoofnagle, J. H. *et al.* Treatment of Chronic Non-A, Non-B Hepatitis with Recombinant Human Alpha Interferon. *N. Engl. J. Med.* **315,** 1575–1578 (1986).
86. Davis, G. L. Recombinant alpha-interferon treatment of non-A, and non-B (type C) hepatitis: review of studies and recommendations for treatment. *J. Hepatol.* **11 Suppl 1,** S72-7 (1990).
87. Poynard, T. *et al.* Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. International Hepatitis Interventional Therapy Group (IHIT). *Lancet (London, England)* **352,** 1426–32 (1998).
88. Kozlowski, A., Charles, S. A. & Harris, J. M. Development of pegylated interferons for the treatment of chronic hepatitis C. *BioDrugs* **15,** 419–29 (2001).
89. Bailon, P. *et al.* Rational Design of a Potent, Long-Lasting Form of Interferon: A 40 kDa Branched Polyethylene Glycol-Conjugated Interferon α-2a for the Treatment of Hepatitis C. *Bioconjug. Chem.* **12,** 195–202 (2001).
90. Glue, P. *et al.* Pegylated interferon-α2b: Pharmacokinetics, pharmacodynamics, safety, and preliminary efficacy data. *Clin. Pharmacol. Ther.* **68,** 556–567 (2000).
91. Hofmann, W. P. & Zeuzem, S. A new standard of care for the treatment of chronic HCV infection. *Nat. Rev. Gastroenterol. Hepatol.* **8,** 257–64 (2011).
92. Hofmann, W. P. & Zeuzem, S. A new standard of care for the treatment of chronic HCV infection. *Nat. Rev. Gastroenterol. Hepatol.* (2011). doi:10.1038/nrgastro.2011.49
93. Pawlotsky, J.-M. The results of Phase III clinical trials with telaprevir and

boceprevir presented at the Liver Meeting 2010: a new standard of care for hepatitis C virus genotype 1 infection, but with issues still pending. *Gastroenterology* **140,** 746–54 (2011).
94. Butt, A. A. & Kanwal, F. Boceprevir and telaprevir in the management of hepatitis C virus-infected patients. *Clin. Infect. Dis.* **54,** 96–104 (2012).
95. Halfon, P. & Locarnini, S. Hepatitis C virus resistance to protease inhibitors. *J. Hepatol.* **55,** 192–206 (2011).
96. De Luca, A., Bianco, C. & Rossetti, B. Treatment of HCV infection with the novel NS3/4A protease inhibitors. *Curr. Opin. Pharmacol.* **18,** 9–17 (2014).
97. Rodríguez-Torres, M. Sofosbuvir (GS-7977), a pan-genotype, direct-acting antiviral for hepatitis C virus infection. *Expert Rev. Anti. Infect. Ther.* **11,** 1269–79 (2013).
98. Sofia, M. J. *et al.* Discovery of a β-d-2'-deoxy-2'-α-fluoro-2'-β-C-methyluridine nucleotide prodrug (PSI-7977) for the treatment of hepatitis C virus. *J. Med. Chem.* **53,** 7202–18 (2010).
99. Gentile, I. *et al.* Efficacy and Safety of Sofosbuvir in the Treatment of Chronic Hepatitis C: The Dawn of a New Era. *Rev. Recent Clin. Trials* **9,** 1–7 (2014).
100. Lam, B., Henry, L. & Younossi, Z. Sofosbuvir (Sovaldi) for the treatment of hepatitis C. *Expert Rev. Clin. Pharmacol.* **7,** 555–566 (2014).
101. Lau, G. *et al.* Efficacy and safety of 3-week response-guided triple direct-acting antiviral therapy for chronic hepatitis C infection: a phase 2, open-label, proof-of-concept study. *lancet. Gastroenterol. Hepatol.* **1,** 97–104 (2016).
102. Sofia, M. J., Chang, W., Furman, P. A., Mosley, R. T. & Ross, B. S. Nucleoside, Nucleotide, and Non-Nucleoside Inhibitors of Hepatitis C Virus NS5B RNA-Dependent RNA-Polymerase. *J. Med. Chem.* **55,** 2481–2531 (2012).
103. Fontana, R. J. *et al.* Sofosbuvir and daclatasvir combination therapy in a liver transplant recipient with severe recurrent cholestatic hepatitis C. *Am. J. Transplant* **13,** 1601–5 (2013).
104. Link, J. O. *et al.* Discovery of Ledipasvir (GS-5885): A Potent, Once-Daily Oral NS5A Inhibitor for the Treatment of Hepatitis C Virus Infection. *J. Med. Chem.* **57,** 2033–2046 (2014).
105. Bari, K. & Sharma, P. Combination of Daclatasvir and Sofosbuvir for Hepatitis C Genotypes 1, 2, and 3. *Gastroenterology* **147,** 534–536 (2014).
106. Lemm, J. A. *et al.* Identification of Hepatitis C Virus NS5A Inhibitors. *J. Virol.* **84,** 482–491 (2010).
107. Gao, M. *et al.* Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nature* **465,** 96–100 (2010).
108. Smith-Palmer, J., Cerri, K. & Valentine, W. Achieving sustained virologic response in hepatitis C: a systematic review of the clinical, economic and quality of life benefits. *BMC Infect. Dis.* **15,** 19 (2015).
109. Brennan, T. & Shrank, W. New Expensive Treatments for Hepatitis C Infection. *JAMA* **312,** 593 (2014).
110. Griffin, S. D. C. *et al.* A conserved basic loop in hepatitis C virus p7 protein is required for amantadine-sensitive ion channel activity in mammalian cells but is dispensable for localization to mitochondria. *J. Gen. Virol.* **85,** 451–461 (2004).
111. Griffin, S., Clarke, D., McCormick, C., Rowlands, D. & Harris, M. Signal Peptide Cleavage and Internal Targeting Signals Direct the Hepatitis C Virus p7 Protein to Distinct Intracellular Membranes. *J. Virol.* **79,** 15525–15536 (2005).

112. Wang, K., Xie, S. & Sun, B. Viral proteins function as ion channels. *Biochim. Biophys. Acta - Biomembr.* **1808,** 510–515 (2011).
113. Winterhalter, M. Black lipid membranes. *Curr. Opin. Colloid Interface Sci.* **5,** 250–255 (2000).
114. Pavlović, D. *et al.* The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 6104–8 (2003).
115. Lotze, J. *et al.* Peptide-tags for site-specific protein labelling in vitro and in vivo. *Mol. BioSyst.* **12,** 1731–1745 (2016).
116. Maue, R. A. Understanding ion channel biology using epitope tags: Progress, pitfalls, and promise. *J. Cell. Physiol.* **213,** 618–625 (2007).
117. Chin, J. W. *et al.* Addition of *p* -Azido- l -phenylalanine to the Genetic Code of *Escherichia c oli*. *J. Am. Chem. Soc.* **124,** 9026–9027 (2002).
118. Alexander Deiters *et al.* Adding Amino Acids with Novel Reactivity to the Genetic Code of Saccharomyces Cerevisiae. (2003). doi:10.1021/JA0370037
119. Jason W. Chin, † *et al.* Addition of p-Azido-l-phenylalanine to the Genetic Code of Escherichia coli. (2002). doi:10.1021/JA027007W
120. Chatterjee, A., Guo, J., Lee, H. S. & Schultz, P. G. A Genetically Encoded Fluorescent Probe in Mammalian Cells. *J. Am. Chem. Soc.* **135,** 12540–12543 (2013).
121. Lee, H. S., Guo, J., Lemke, E. A., Dimla, R. D. & Schultz, P. G. Genetic incorporation of a small, environmentally sensitive, fluorescent probe into proteins in Saccharomyces cerevisiae. *J. Am. Chem. Soc.* **131,** 12921–3 (2009).
122. Hsieh, T. *et al.* Monitoring Protein Misfolding by Site-Specific Labeling of Proteins In Vivo. *PLoS One* **9,** e99395 (2014).
123. Sakata, S., Jinno, Y., Kawanabe, A. & Okamura, Y. Voltage-dependent motion of the catalytic region of voltage-sensing phosphatase monitored by a fluorescent amino acid. *Proc. Natl. Acad. Sci.* **113,** 7521–7526 (2016).
124. Ryoji, M. *et al.* Read-through translation. *Trends Biochem. Sci.* **8,** 88–90 (1983).
125. Kaeberlein, M. & Kennedy, B. K. Protein translation, 2007. *Aging Cell* **6,** 731–734 (2007).
126. Young, T. S. & Schultz, P. G. Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J. Biol. Chem.* **285,** 11039–44 (2010).
127. Liu, W., Brock, A., Chen, S., Chen, S. & Schultz, P. G. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat. Methods* **4,** 239–244 (2007).
128. Liu, W., Brock, A., Chen, S., Chen, S. & Schultz, P. G. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat. Methods* **4,** 239–44 (2007).
129. Kalstrup, T. & Blunck, R. Reinitiation at non-canonical start codons leads to leak expression when incorporating unnatural amino acids. *Sci. Rep.* **5,** 11866 (2015).
130. Wang, W. *et al.* Genetically encoding unnatural amino acids for cellular and neuronal studies. *Nat. Neurosci.* **10,** 1063–1072 (2007).
131. Xiang, L. *et al.* Crucial optimization of translational components towards efficient incorporation of unnatural amino acids into proteins in mammalian cells. *PLoS One* **8,** e67333 (2013).
132. Kelemen, R. E. *et al.* A Precise Chemical Strategy To Alter the Receptor Specificity of the Adeno-Associated Virus. *Angew. Chem. Int. Ed. Engl.* **55,**

10645–9 (2016).
133. Young, D. D. *et al.* An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. *Biochemistry* **50,** 1894–900 (2011).
134. Greiss, S. & Chin, J. W. Expanding the genetic code of an animal. *J. Am. Chem. Soc.* **133,** 14196–9 (2011).
135. Liu, C. C. & Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* **79,** 413–444 (2010).
136. Wang, Q., Parrish, A. R. & Wang, L. Expanding the Genetic Code for Biological Studies. *Chem. Biol.* **16,** 323–336 (2009).
137. Li, F. *et al.* Expanding the Genetic Code for Photoclick Chemistry in *E. coli* , Mammalian Cells, and *A. thaliana*. *Angew. Chemie Int. Ed.* **52,** 9700–9704 (2013).
138. Yuan, Z. *et al.* Controlling Multicycle Replication of Live-Attenuated HIV-1 Using an Unnatural Genetic Switch. *ACS Synth. Biol.* **6,** 721–731 (2017).
139. Han, S. *et al.* Expanding the genetic code of Mus musculus. *Nat. Commun.* **8,** 14568 (2017).