

2007

# The Multifunctional Capsid Gene of Hepatitis C Virus

Laura Katherine McMullan

Follow this and additional works at: [http://digitalcommons.rockefeller.edu/student\\_theses\\_and\\_dissertations](http://digitalcommons.rockefeller.edu/student_theses_and_dissertations)



Part of the [Life Sciences Commons](#)

---

## Recommended Citation

McMullan, Laura Katherine, "The Multifunctional Capsid Gene of Hepatitis C Virus" (2007). *Student Theses and Dissertations*. Paper 14.



# THE MULTIFUNCTIONAL CAPSID GENE OF HEPATITIS C VIRUS

A Thesis Presented to the Faculty of  
The Rockefeller University  
in Partial Fulfillment of the Requirements for the degree of  
Doctor of Philosophy

by

Laura Katherine McMullan

June 2007





# THE MULTIFUNCTIONAL CAPSID GENE OF HEPATITIS C VIRUS

Laura McMullan, Ph.D.  
The Rockefeller University, 2007

In populations worldwide, the Hepatitis C Virus displays great diversity. The capsid gene, however, is remarkably conserved and this has led to predictions of RNA secondary structure and an overlapping, alternative reading frame (ARF). We investigated a role of the ARF in HCV infection, by introducing four stop codons into the ARF of a genotype 1a H77 molecular clone. These changes did not alter the capsid protein sequence, but were predicted to disrupt the RNA secondary structures, SLV and SLVI. An infection was launched after inoculation of the mutant HCV RNA into an HCV naïve chimpanzee. The acute infection was attenuated with low peak viremia, minimal ALT elevation, and early virus control by a diverse adaptive immune response.

Sequencing circulating virus revealed progressive reversions at the third and then fourth stop codon mutations. In cell culture, replication of a genome with the four stop codons was severely impaired, but the revertant genomes showed a marked improvement in replicative fitness. Consistent

with the chimpanzee infection, reversions at stop codons 3 and 4 were also selected by passage of mutant, infectious HCV in cell culture.

Genetic evidence for RNA structures that were disrupted by the stop codons was provided by structure-restoring compensating mutations that relieved the defective HCV replication. Further mutagenesis identified bases and secondary structure features critical for function; mutations in the top stem of SLVI severely impaired HCV replication. RNA bearing these mutations was less efficiently translated in cell free extracts. The effect of mutations in SLVI was reduced when the HCV replicase proteins were translated using a heterologous IRES. This data suggests that these RNA structures are important for translation from the HCV IRES and that the defective translation exhibited by the SLVI mutants contributed to their impaired replication.

Thus, RNA structures within the capsid coding region are responsible for modulating translation from the HCV IRES, and do not encode an ARF protein as had been postulated. While dispensable for the viability of subgenomic replicons, strong selective pressure for the integrity of these structures, both *in vivo* and *in vitro*, highlights their importance for the HCV lifecycle.

## ACKNOWLEDGEMENTS

I would like to thank my advisor Charlie Rice for his guidance and support for extensive and expensive experiments. His exacting standards ensured all experiments were completed in triplicate and controls included for each process. His perfecting approach towards science has demonstrated that careful and detailed experiments often reveal subtle, yet repeatable information. I admire his work ethic and his dedication to train a large group of scientists with diverse backgrounds and interests. He was always reliable to give an opinion for the latest results and his insight on how they relate to the work enabled the progress presented in this thesis.

I would also like to thank Andrea Branch for her creativity and enthusiasm. Her willingness to listen and contemplate the minute details of my experiments provided a great sounding board as well as expert guidance and suggestions. She was always available and she thoughtfully considered each aspect of the work seriously.

I owe many thanks to my faculty committee, which was subjected to often longer than planned talks and discussions. John McKinney, the chair, always provided careful insight and interpretation of results up to the final task of the reading of the thesis. Magda Konarska provided positive support and encouragement on the details of RNA structure-function relationships.

Stephen Goff, my external reviewer, for reading carefully the final thesis and offering great suggestions and insights.

The entire rice lab and alumni have contributed to this work. With the professionalism displayed during lab meetings and smaller, subgroup meetings, I have benefited from a collective knowledge of diverse disciplines and approaches of doing science. In particular, Joe Marcotrigiano established the cell free translation reaction using Huh-7 cell extracts and let me use his system for my translation studies. Arash Grakoui who taught how to isolate and identify an immune response. Matt Evans proposed the general method for selection of reversions using the JFH isolate and detected the adaptive mutations allowing a chimera H-JFH to produce virus. Chris Jones introduced the luciferase reporter to the Rice lab, which has greatly increased the speed and ease of doing triplicate, Rice lab experiments. I am grateful for Shihyun You whom always offered support, positive attitude, and patience while listening to various stages of each project. I would also like to thank Mike Flint for his continual support and encouragement. He always believed the work would get finished.

# TABLE OF CONTENTS

Chapter 1: Introduction	1
Identification of hepatitis C virus	2
Epidemiology and disease burden of HCV	3
Diversity of HCV	4
Treatment	6
Transmission	6
Classification and genome organization	7
Lifecycle	14
Model systems to study HCV	18
Cell culture systems	21
Translation of HCV	25
Influence of the capsid gene on translation	29
Conservation of the capsid gene	31
Evidence for an ARF protein in HCV replication and infection	34
Summary of thesis work	42
 Chapter 2: Materials and Methods	 47
Cell culture	47
Plasmid construction	48
Site directed mutagenesis	51
Detection and sequencing of HCV RNA	52

Chimpanzee inoculation and sample collection	53
Isolation of PBMCs	53
Isolation of CD8 <sup>+</sup> and CD4 <sup>+</sup> T cells from liver biopsy	54
Detection of HCV-specific CD8 <sup>+</sup> T cells by ELISPOT	54
HCV recombinant antigen and peptides	55
Detection of HCV-specific antibodies	56
Generation of CTL lines	56
CD4 <sup>+</sup> proliferation assay	57
Flow cytometry	58
<i>In vitro</i> transcription	59
Electroporation of transcribed HCV RNA	60
Quantitation of HCV RNA by Real Time qRT-PCR	61
RNA secondary structure probing	62
Primer extension analysis	63
TransMessenger delivery of HCV RNA to Huh-7.5 cells	64
Immunohistochemistry	64
Luciferase assay	65
Metabolic labeling of proteins and immunoprecipitation	65
Selection of reversions	67
Huh-7 translation cell extracts	67
<i>In vitro</i> translation using Huh-7 cell extracts	68
<i>In vitro</i> translation using rabbit reticulocytes	69
Selection of HCV replicons with G418	69

Chapter 3: The Alternative Reading Frame is not required for infection	74
Introduction	74
Ablation of ARF-encoded gene products by site directed mutagenesis	76
An HCV genome without F and DF expression is infectious <i>in vivo</i>	79
An immune response is detected in both the periphery and liver	81
Unique epitopes are detected by cloning of T cells from liver biopsy	82
The selection of revertants suggests pressure to maintain ARF RNA elements rather than F/DF protein expression	84
Evidence for an RNA structure in the capsid gene	87
Discussion	89
Chapter 4: An RNA structure in the capsid-coding gene affects HCV replication in cell culture	96
Introduction	97
The Stop mutations confer a replication defect in cell culture	98



The Stop mutations do not affect expression or cleavage of the polyprotein	104
The ARF is dispensable for HCV replication in cell culture	105
An RNA structure in the capsid gene is involved in HCV replication	110
RNA structure probing in solution reveals that Stop mutations disrupt an RNA secondary structure	113
RNA structure in the capsid gene influences replication in diverse genotypes	117
The use of luciferase reporter genomes with the JFH-1 replicase	121
The Open mutations impair replication and virus infectivity	126
The mutations in SLVI do not alter RNA stability in Huh-7.5 cells	128
The loops of SLV and SLVI do not have a role in replication and infectivity	128
Mutations in the base of SLVI impair replication	133
Mutations in the base of SLVI additively impair replication with mutations in the top of SLVI	135
Mutations in the middle, c-bulge region, and amino acids in the capsid protein affect replication and infectivity	137

Mutations in the c-bulge region impair replication in the genotype 2a genome	145
Evolution of capsid protein mutants reveals an interaction with p7 and NS2	150
Evolution of RNA structure mutations in Huh-7.5 cells mimic reversions <i>in vivo</i>	163
Discussion	171
Chapter 5: The role of RNA structures in the capsid gene	182
Introduction	182
Initial translation is affected by mutations in SLVI	186
Mutations in SLVI reduced translation <i>in vitro</i>	192
Differences in translation are not due to RNA degradation	196
Visualization of radiolabeled proteins during <i>in vitro</i> translation show differences in translation	197
<i>In vitro</i> translation using extracts from Huh-7 and Huh-7.5 cells	199
Translation using Huh-7.5 cell extracts is less efficient than Huh-7 cell extracts	204
IRES mutations known to affect translation initiation severely reduce HCV translation and replication	206
Does SLVI participate in a long range RNA-RNA interaction with the 5'NTR?	211

Deletions in the 5'NTR reduce translation	219
Deletions in the capsid gene reduce translation	222
Genotype differences in the 5'NTR do not affect replication of the J6-JFH genome	224
Mutations in the 5'NTR eliminate replication	228
Potential interactions of the capsid protein <i>in trans</i>	232
Translation and replication kinetics of HCV in the presence of HCV structural proteins expressed <i>in trans</i>	246
Transcomplementation of point mutations in the capsid gene with the structural proteins increase virus infectivity	253
Mutations alter translation and replication at the single cell	257
The influence of SLV and SLVI in replication are position dependent in the genome	266
The replication defect contributed by mutations in SLVI is abrogated by independent expression of the replicase	267
Discussion	272

Chapter 6: General discussion	289
Predictions of a multifunctional capsid gene	289
Is the ARF essential in an HCV infection?	290
Are the Stop 1,2,3,4 mutations positively selected?	292
Is replication compromised by the ablated ARF or disrupted RNA structures?	294
Does the replication defect in Huh-7.5 cells explain the attenuated infection of the mutant genome Stop 1,2,3,4?	296
What defect in replication is caused by the mutations in SLVI?	297
Can the structural proteins transcomplement mutant HCV genomes in replication?	298
Is there a subpopulation of Huh-7.5 cells capable of supporting mutant SLVI replication?	300
Are SLV and SLVI location dependent in the genome?	302
Model for the role of SLVI in replication	303
Future directions	307
References	310

## LIST OF FIGURES

### Chapter 1:

Figure 1.1	Phylogenetic tree of the principal genotypes of HCV	5
Figure 1.2	Genome organization of HCV	8
Figure 1.3	RNA secondary structures in the HCV genome	10
Figure 1.4	Lifecycle of HCV in the cell	15
Figure 1.5	Molecular clones and bicistronic genomes	23
Figure 1.6	The IRES and translation initiation pathway	26
Figure 1.7	RNA secondary structure of the 3' minus strand	27
Figure 1.8	Sequence conservation of regions of the HCV genome	33
Figure 1.9	The proposed frameshift sites of the F and DF proteins	35
Figure 1.10	Comparison of the HCV-1 and HCV-RH capsid sequences	38

### Chapter 3:

Figure 3.1	Location of Stop codons in the F and DF protein	77
Figure 3.2	Location of mutations for F and DF stop codons in RNA secondary structures SLV and SLVI	78
Figure 3.3	Infection profile of Stop 1,2,3,4 in the chimpanzee	80
Figure 3.4	Reversions detected in the circulation after inoculation with Stop 1,2,3,4 genome RNA	85
Figure 3.5	Reversions restore an RNA structure	88
Figure 3.6	Phylogenetic analysis of conserved nucleotide and covariant base pairs of SLVI	90

### Chapter 4:

Figure 4.1	Replication of the Stop 1,2,3,4 genome is severely impaired in Huh-7.5 cells	100
Figure 4.2	Replication of genomes with combinations of mutations for Stop 1-4 in Huh-7.5 cells	101
Figure 4.3	Flow cytometry of Huh-7.5 cells transfected with Stop mutations	103
Figure 4.4	Radiolabeled HCV proteins expressed from genomes with Stop mutations	106

Figure 4.5	H77 SLV and SLVI with mutations for Stop 1,2,3,4 and Open 1,2,3,4	107
Figure 4.6	Replication of genomes with Open mutations	109
Figure 4.7	Replication of genomes with compensatory mutations for Stop 1,2,3,4	111
Figure 4.8	Mfold of Stop mutations in SLV and SLVI	114
Figure 4.9	RNA structure probing reveals mutations of Stop 1,2,3,4 disrupt the top of SLVI	116
Figure 4.10	Stop 1,2,3,4 and Stop 3,4 mutations impair replication in genotype 1b genomes	118
Figure 4.11	Stop 1,2,3,4 and Stop 3,4 mutations impair replication in genotype 2a and 1a-2a chimeras	120
Figure 4.12	Translation and replication of luciferase reporters in JFH-1 chimeras	122
Figure 4.13	Replication of Stop 1,2,3,4 mutations in the H-JFH/Rluc genome	124
Figure 4.14	Replication of Open 1,2,3,4 in H-JFH/Rluc	127
Figure 4.15	Replication of Stop 3,4 and Open 3,4 mutations in a polymerase defective genome	129
Figure 4.16	Mutations in the loops of SLV and SLVI	130
Figure 4.17	Replication of genomes with mutations in the loops of SLV and SLVI	132
Figure 4.18	Mutations introduced in the base and top of SLVI	134
Figure 4.19	Replication of genomes with mutations in the base of SLVI	136
Figure 4.20	Replication of genomes with mutations in the base and top of SLVI	138
Figure 4.21	Mutations in the middle region of H77 SLVI	140
Figure 4.22	Replication of H77 genomes with mutations in the middle region of SLVI	141
Figure 4.23	Replication of H77 genomes with silent mutations in the middle region of SLVI	144
Figure 4.24	Replication of luciferase reporter genomes of H-JFH and J6-JFH	146
Figure 4.25	Mutations in the middle region of SLVI in genotype 2a	148
Figure 4.26	Replication of genomes with mutants in the middle region of genotype 2a	149
Figure 4.27	Mutation differences between J6 and JFH-1	151

Figure 4.28	Replication of genomes with mutation differences between J6 and JFH	152
Figure 4.29	Selection for reversions increasing virus infectivity from genomes containing a mutant capsid protein	154
Figure 4.30	Alignment of amino acid conservation of the p7 reversion site in all 6 genotypes	157
Figure 4.31	Alignment of amino acid conservation of the NS2 reversion site in all 6 genotypes	158
Figure 4.32	Position of compensatory mutation of a mutant capsid in the predicted membrane topology of p7 and NS2	159
Figure 4.33	IHC of compensatory mutations in p7 and NS2 restoring infectivity	161
Figure 4.34	Selection for reversions improving replication of genomes with disrupted SLVI	165
Figure 4.35	Sequence analysis of reversions during passage of Huh-7.5 cells harboring Stop 1,2,3,4 genomes	167
Figure 4.36	Sequence analysis of reversions during passage of Huh-7.5 cells harboring Stop 3,4 genomes	170

## Chapter 5:

Figure 5.1	Translation of genomes with Stop 1,2,3,4 mutations in electroporated Huh-7.5 cells	187
Figure 5.2	Translation of genomes with Stop mutations in electroporated Huh-7.5 cells	189
Figure 5.3	Translation of genomes with Stop mutations introduced by lipid transfection in Huh-7.5 cells	191
Figure 5.4	<i>In vitro</i> translation of genomes with Stop 1,2,3,4	193
Figure 5.5	End point measurement of <i>In vitro</i> translation of genomes with Stop 1,2,3,4	195
Figure 5.6	Radiolabeled RNA is not differentially degraded during <i>in vitro</i> translation	198
Figure 5.7	Mutant Stop 1,2,3,4 produced less radiolabeled translation products	200
Figure 5.8	Mutant Stop 1,2,3,4 shows reduced translation in Huh-7 and Huh-7.5 cell extracts	202
Figure 5.9	Huh-7.5 cell extracts have a negative factor that reduces translation	205
Figure 5.10	Location of mutations in the IRES known to affect translation initiation	208

Figure 5.11	Translation and replication of initiation mutants	209
Figure 5.12	Diagram of long range RNA-RNA interaction of the 5'NTR and capsid gene	212
Figure 5.13	Mutations in the base of SLVI and 5'NTR	215
Figure 5.14	Translation of genomes with mutants designed to interrupt the interaction between SLVI and 5'NTR	217
Figure 5.15	Translation of genomes with deletions in the 5'NTR	221
Figure 5.16	Translation of genomes with deletions in the capsid gene	225
Figure 5.17	Genotype differences in the 5'NTR do not affect replication	227
Figure 5.18	Mutations in the 5'NTR designed to disrupt miR122 or SLVI interactions	229
Figure 5.19	Replication of genomes with mutations in the 5'NTR	230
Figure 5.20	Translation and replication of genomes with a repeated base SLVI binding site in the 5'NTR	233
Figure 5.21	Constructs used in transcomplementation assay	236
Figure 5.22	Replication of genomes with deletions in the capsid gene in Huh-7.5 cells and VEErepC cell line	238
Figure 5.23	Replication of coelectroporation of VEE replicons with deletion-capsid HCV genomes	241
Figure 5.24	Infection of Huh-7.5 cells and a VEErepC cell line with virus produced after coelectroporation of HCV capsid-deletion genomes and VEErepC	244
Figure 5.25	Translation and replication kinetics of mutant HCV genomes transcomplemented with HCV structural proteins	248
Figure 5.26	Translation and replication kinetics of mutant HCV genomes with point mutations transcomplemented with HCV structural proteins	255
Figure 5.27	IHC and flow cytometry of transfected Huh-7.5 cells with mutant genomes	259
Figure 5.28	Translation and replication of mutant HCV in an end point dilution assay	263
Figure 5.29	Replication at the end point and minimum number of transfected cells replicating HCV	265
Figure 5.30	Replication of genomes with the mutant SLVI positioned behind the EMCV IRES	268
Figure 5.31	G418 selection of cells replicating genomes with the mutant SLVI positioned behind the EMCV IRES	269



Figure 5.32 Replication of bicistronic genomes expressing the  
expressing the replicase independently of the  
HCV IRES and mutant SLVI

271

## LIST OF TABLES

Table 2.1	Mutations introduced into HCV genomes	71
Table 3.1	CD8 <sup>+</sup> T cell lines and epitopes of CD4 <sup>+</sup> T cell clones generated from liver biopsy	83

## **Chapter 1**

## **Chapter 1. Introduction.**

### ***Identification of hepatitis C virus.***

Forty years ago, one in three blood transfusions resulted in the recipient contracting hepatitis <sup>3</sup>. After diagnostic tests identified the hepatitis A and B viruses, the unattributed disease was designated as non-A, non-B hepatitis (NANBH). Harvey Alter and others demonstrated that serum from NANBH patients conveyed hepatitis when injected into a chimpanzee, and the causative agent was recognized as an enveloped virus after transmission was shown to be due to a chloroform-sensitive, filterable agent <sup>1, 139</sup>. The virus responsible for NANBH was identified by Michael Houghton's group at Chiron using a random-primed complementary DNA library prepared from NANBH-infected chimpanzee plasma and screened with serum from a NANBH infected patient <sup>27</sup>. In a 7 year effort, one out of one million cDNAs was recognized by the NANBH patient serum and proved to be derived from the etiological agent by detecting an RNA of approximately 10,000 nucleotides in only NANBH infected patient serum. With the cloned agent sequenced, the newly identified hepatitis C virus (HCV) was classified and its genome organization disclosed. Knowledge of HCV proteins enabled new serological diagnostic assays to be developed. By 1996, the nation's

blood supply could be screened for hepatitis A, B and C viruses reducing the risk of transfusion delivered hepatitis to nearly zero <sup>2,4</sup>.

### ***Epidemiology and disease burden of HCV.***

An estimated 3% of the world population is infected with HCV, and an estimated three to four million persons are newly infected each year <sup>159</sup>. Infection with HCV is rarely cleared by the host immune response. In 65% - 80% of exposures, a persistent, chronic infection is established for the life of the host <sup>51</sup>. Although initial symptoms are often mild, chronic infection can lead to liver damage and hepatocellular carcinoma. Infection with HCV accounts for approximately 20% of cases of acute hepatitis and 70% of cases of chronic hepatitis. HCV is responsible for 50-76% of all liver cancer cases and is the leading cause of liver transplants in the developed world <sup>159</sup>.

The severity of HCV related liver disease is extremely variable between patients. Progression of liver fibrosis to cirrhosis and to hepatocellular carcinoma may take years or decades. Jaundice and fulminant hepatitis are very rare. More commonly, patients show an elevation in the aminotransferase levels and may have mixed cryoglobulinemia <sup>120</sup>.

Disease prevalence is below 1% in Australia, Canada and northern Europe, about 1% in the USA and most of Europe, and greater than 2%

(often 5-10%) in many countries in Africa, Latin America and Central and South-Eastern Asia <sup>159</sup>. Italy has areas of high prevalence rates of over 12.6%. Egypt has an extremely high seroprevalence rate ranging from 20% for those 10-19 years to about 60% for those above 30 years of age <sup>156</sup>.

### ***Diversity of HCV.***

HCV displays remarkable sequence diversity at various levels. There are 6 major genotypes displaying greater than 30% diversity at the nucleotide level <sup>128, 129</sup>. They correlate with geographical location and particular risk groups. Within each genotype are more closely related subtypes, conventionally denoted with a letter, which differ by 20-25% at the nucleotide level. The subtypes 1a, 1b, and 3a are generally found in Western countries but can show a wide geographical range as they have been dispersed through blood transfusions and injecting drug use (Figure 1.1) <sup>134</sup>. Over 90% of HCV patients in Egypt are infected with genotype 4a as it was introduced and spread during a campaign for schistosomiasis treatment reusing nonsterilized needles <sup>32</sup>. Within an infected patient, the virus population exists as a quasispecies differing by 2-5% at the nucleotide level.

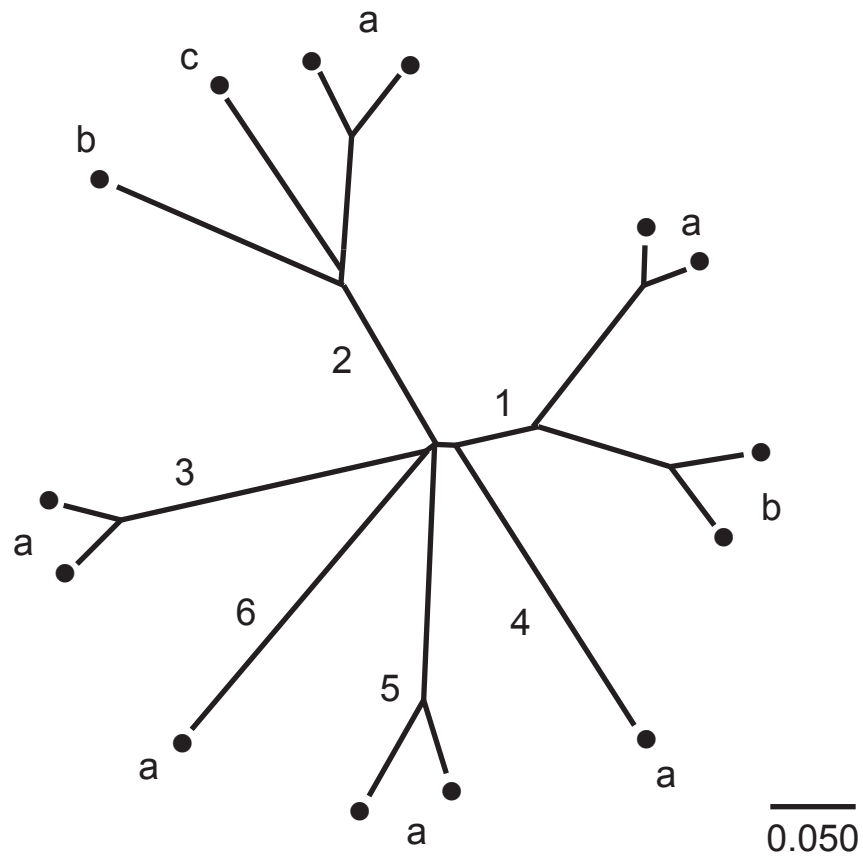


Figure 1.1 Phylogenetic tree of the principal genotypes of HCV. The tree was constructed by using the neighbour-joining method using Jukes–Cantor-corrected distances between complete genome sequences. Adapted from Simmonds 2002.

### ***Treatment.***

The standard of care for HCV is pegylated interferon-alpha (IFN- $\alpha$ ) in combination with ribavirin <sup>28</sup>. Successful treatment is considered a sustained virologic response (SVR) and indicates undetectable HCV RNA in the blood 24 weeks after treatment has ended. Treatment outcomes can vary with 89% of patients achieving SVR if not coinfecting with HIV, relatively healthy, and infected with genotype 2 or 3. The SVR drops to 30% for patients infected with genotype 1 HCV, coinfecting with HIV, or with advanced liver fibrosis. A predictor of responding to treatment is seen within the first 12 weeks in a 2 log drop in virus load <sup>89</sup>.

### ***Transmission.***

Transmission occurs through direct contact of infected blood. Hemophiliacs, hemodialysis patients, drug addicts and people transfused with inadequately screened blood are particularly at risk for infection <sup>4</sup>. In the developed world, intravenous drug use remains the main route of transmission, accounting for almost 90% of new HCV infections <sup>107</sup>. Sexual transmission is thought to be infrequent. Mother-to-child infection ranges from 3-15% and is believed to occur *in utero*.



### ***Classification and genome organization.***

Hepatitis C virus is the sole member of the genus *Hepacivirus* in the family *Flaviviridae*. The virus genome is a single RNA molecule of positive strand polarity<sup>9,74</sup>. The genome organization closely resembles the other genera of the pestiviruses and flaviviruses. Virions are 50 nm in diameter, spherical, and surrounded by a lipid bilayer as measured by filtration and electron microscopy. Virus particles display significant heterogeneity as particles isolated from acutely infected patients have a buoyant density in sucrose of 1.06 g/cm<sup>3</sup> and particles from serum of chronically infected patients have a density of 1.15-1.18 g/cm<sup>3</sup><sup>49</sup>. This lower density is thought to be a result of particle association with very-low-density lipoproteins while the higher density particles are associated with immunoglobulins in the serum. Cell culture derived virus produces a density of 1.14 g/cm<sup>3</sup> but a peak of specific infectivity at 1.10 g/cm<sup>3</sup>. Passage of cell culture derived virus in a chimpanzee resulted in particles having a major peak at 1.10 g/cm<sup>3</sup> and a smaller peak at 1.14 g/cm<sup>3</sup><sup>73</sup>.

The HCV genome is a single-strand, positive polarity RNA molecule of approximately 9,600 nucleotides in length (Figure 1.2). The genome organization closely resembles the other genera of the *Flaviviridae*, the *pestiviruses* and *flaviviruses*. The RNA is uncapped and encodes a

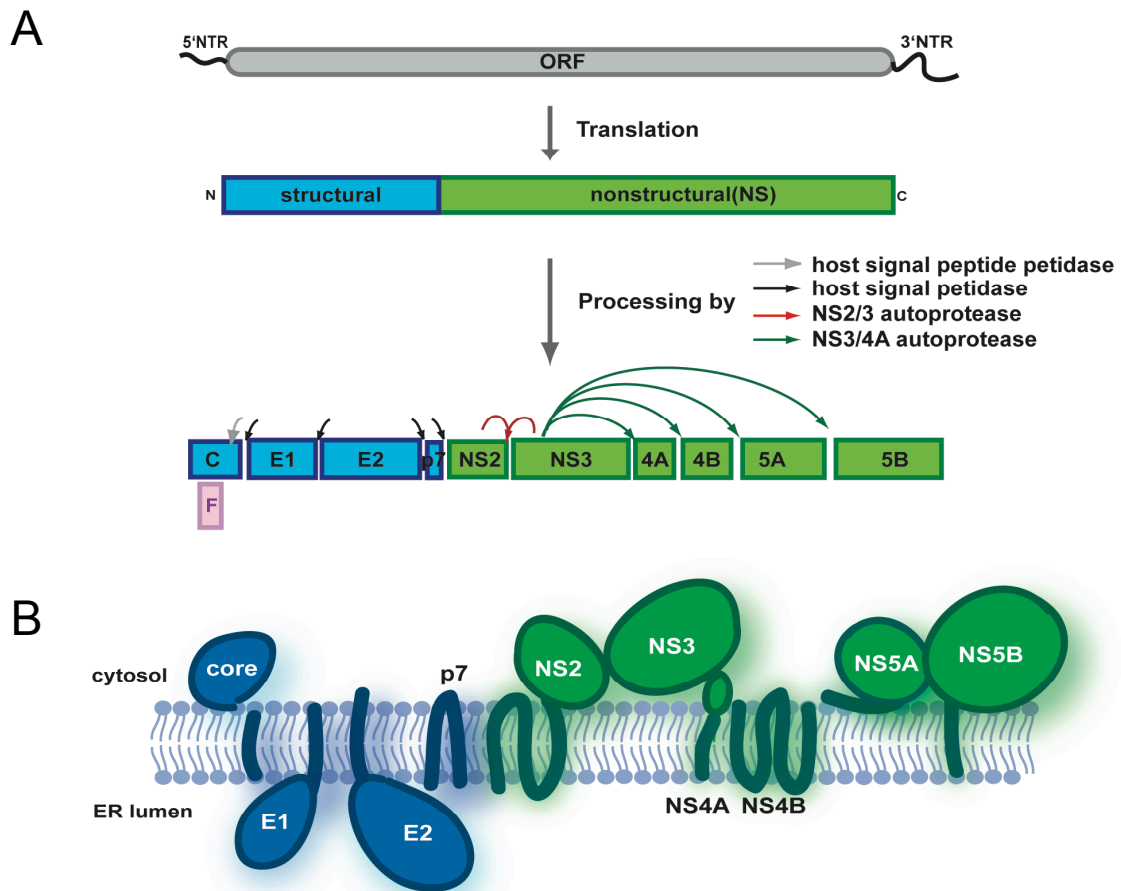


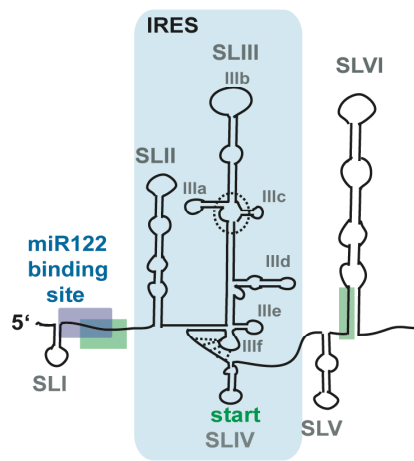
Figure 1.2. A. The genome of HCV is flanked by 5' and 3'NTRs and encodes one, long polyprotein, which is processed by host cell and virus proteases into individual virus structural and nonstructural proteins. B. The proposed membrane topology of each HCV protein.

polyprotein of about 3000 amino acids that is cleaved co- and post-translationally by cellular and virus encoded proteases into at least 10 mature proteins <sup>45</sup>. These proteins act to replicate the virus genome into minus strand intermediates and then plus strand genomes, which are then packaged by virus structural proteins.

### ***NTRs.***

The HCV polyprotein is flanked by non translated regions (NTRs) at the 5' and 3' of the genome which display highly structured RNA elements conducting translation and replication. An internal ribosome entry site (IRES) in the 5' NTR directs translation of the polyprotein in a cap-independent manner and will be discussed in detail <sup>54</sup>. Replication signals are present at the 5' end of the 5'NTR and overlap the IRES <sup>41</sup>. RNA structures in the minus sense are thought to act as promoters for the virus replicase during plus strand synthesis (Figure 1.3). The 3'NTR contains two highly conserved elements, the poly U/UC tract and a three helical structure of 98 bases called the 3'-X at the end of the 3'NTR <sup>15, 162</sup>. Both elements are required for replication and are thought to act as promoters for the replicase during minus strand synthesis. A kissing interaction is required between the SL2 of the 3'-X and a large stem structure in the coding region of NS5B <sup>40</sup>,

A



B

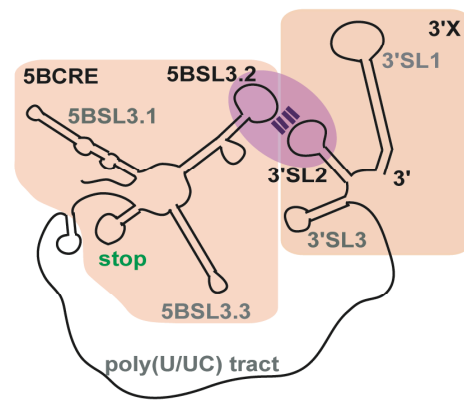


Figure 1.3. RNA secondary structures in the HCV genome. A. The 5'NTR contains RNA structures required for replication, SLI, and signals for translation by the IRES. SLV and SLVI are in the capsid-coding gene. B. RNA secondary structures in the 3'NTR participate in a kissing interaction with RNA structures in the coding-region of NS5B.

<sup>165</sup>. The length of the poly U/UC tract can influence replication and may contribute to the stability of the kissing interaction <sup>40</sup>. Although still controversial, the 3'NTR may contribute to translation efficiency <sup>135</sup>.

### ***Structural proteins.***

The structural proteins are expressed in the amino terminus the polyprotein. The capsid protein is a highly basic RNA binding protein that forms the nucleocapsid <sup>121</sup>. It is the most conserved virus protein and has been described to perform a multitude of functions to modulate the host cell. A hydrophobic signal sequence located at the carboxy terminus of capsid guides the polyprotein into the ER where it is cleaved by the host cell signal peptidase. Another cleavage event by the host cell signal peptide peptidase removes the signal sequence from the anchored capsid protein; nonetheless the capsid remains associated with the ER via an unknown interaction <sup>86, 157</sup>. HCV encodes two envelope glycoproteins, E1 and E2, believed to be type I integral membrane proteins, which are incorporated into the virion and likely interact with cellular receptors and mediate viral entry. Both E1 and E2 contain signals in their transmembrane domains that mediate localization to the endoplasmic reticulum when expressed in cell culture. p7 is a small hydrophobic protein of 63 amino acids and contains two transmembrane

domains and small basic loop in the cytoplasm and is often found uncleaved from E2<sup>71</sup>. It has been shown to have ion channel activity resembling the M2 activity of influenza and other viroporins<sup>100</sup>. It is not essential for replication, but required at an as yet unknown stage of infectious virus function<sup>46</sup>. There are also reports of a protein expressed by ribosome slippage into a -2/+1 frame of the capsid gene<sup>6, 18, 150, 153, 161</sup>. This alternative reading frame (ARF) is a conserved overlapping reading frame present in all 6 genotypes that could express a 125-160 amino acid protein. HCV infected patients have been shown to produce antibodies and express T cell epitopes to peptides from the ARF, but a role for an ARF protein is unknown<sup>66</sup>.

### ***Nonstructural Proteins.***

The remaining proteins are considered nonstructural proteins (NS proteins) and participate in virus protein processing and virus replication. NS2 acts in conjunction with the NS3 protease to autocatalytically cleave itself from NS3<sup>44</sup>. The catalytic domain of NS2 acts as a dimeric cysteine protease with a composite active sites derived from two monomeric chains<sup>80</sup>. NS2 is not required for virus replication, and recent studies have shown adaptive mutations found in NS2 permit virus production<sup>164</sup>. The N terminus of NS3 is a serine protease responsible for four downstream cleavages of the

polyprotein. The carboxy terminus contains RNA helicase and nucleoside triphosphatase (NTP) activity<sup>10, 145</sup>. The 54-residue NS4A is a cofactor of the protease NS3 and is associated with the ER and NS5A. NS4B is a 27kDa hydrophobic protein predicted to span the ER with at least four transmembrane domains with amino and carboxy tails located in the cytoplasm<sup>11</sup>. NS4B also has no currently defined function, but is thought to participate as a membrane scaffolding protein for the virus replicase. Genetic interactions have been described between NS4B and NS5A influencing HCV replication<sup>11</sup>. NS5A is a serine phosphoprotein with hyper and basal phosphorylated states and is the target of numerous cell culture adaptive mutations improving HCV replication in Huh-7 cells<sup>68, 141</sup>. NS5A has also been reported to act with cellular trafficking and antiviral pathways<sup>142</sup>. NS5B is the RNA-dependent RNA polymerase and is anchored to the ER via its 21 amino acid carboxy tail<sup>90</sup>. It has been shown to initiate RNA synthesis in a primer independent manner, but is regulated by interactions with other virus replicase proteins and RNA secondary elements of the genome<sup>10, 110</sup>.

### *Lifecycle.*

Virus attachment to, and entry into, target cells occurs by a multifactorial, multi-step process the details of which are not completely understood. Direct binding does occur between the glycoprotein E2 and the cellular co-receptors CD81 and scavenger receptor class-B member-I (SR-B1)<sup>47, 104</sup>. CD81 is present on numerous cells types and functions via interactions with cell type specific proteins; however, the role of CD81 in hepatocytes is unknown. Claudin-1 is required for infection at a late stage possibly after CD81 binding<sup>35</sup>. Other receptors have also been described such as the liver and lymph node-specific intercellular adhesion molecule-3-grabbing nonintergrin (L-SIGN) and the dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintergrin (DC-SIGN)<sup>43 106</sup>. HCV entry presumably occurs through receptor-mediated endocytosis, that delivers viral particles to endosomes, where low pH initiates fusion between the viral and cellular membranes<sup>146</sup>. Following membrane fusion, although the mechanism has not been studied, the nucleocapsid is disassembled, and the virus genome RNA is released into the cytoplasm (Figure 1.4).

To generate viral proteins required for replication, the virus RNA must first undergo translation. The virus genome associates with the cellular translation machinery via the highly structured IRES of the 5'NTR.



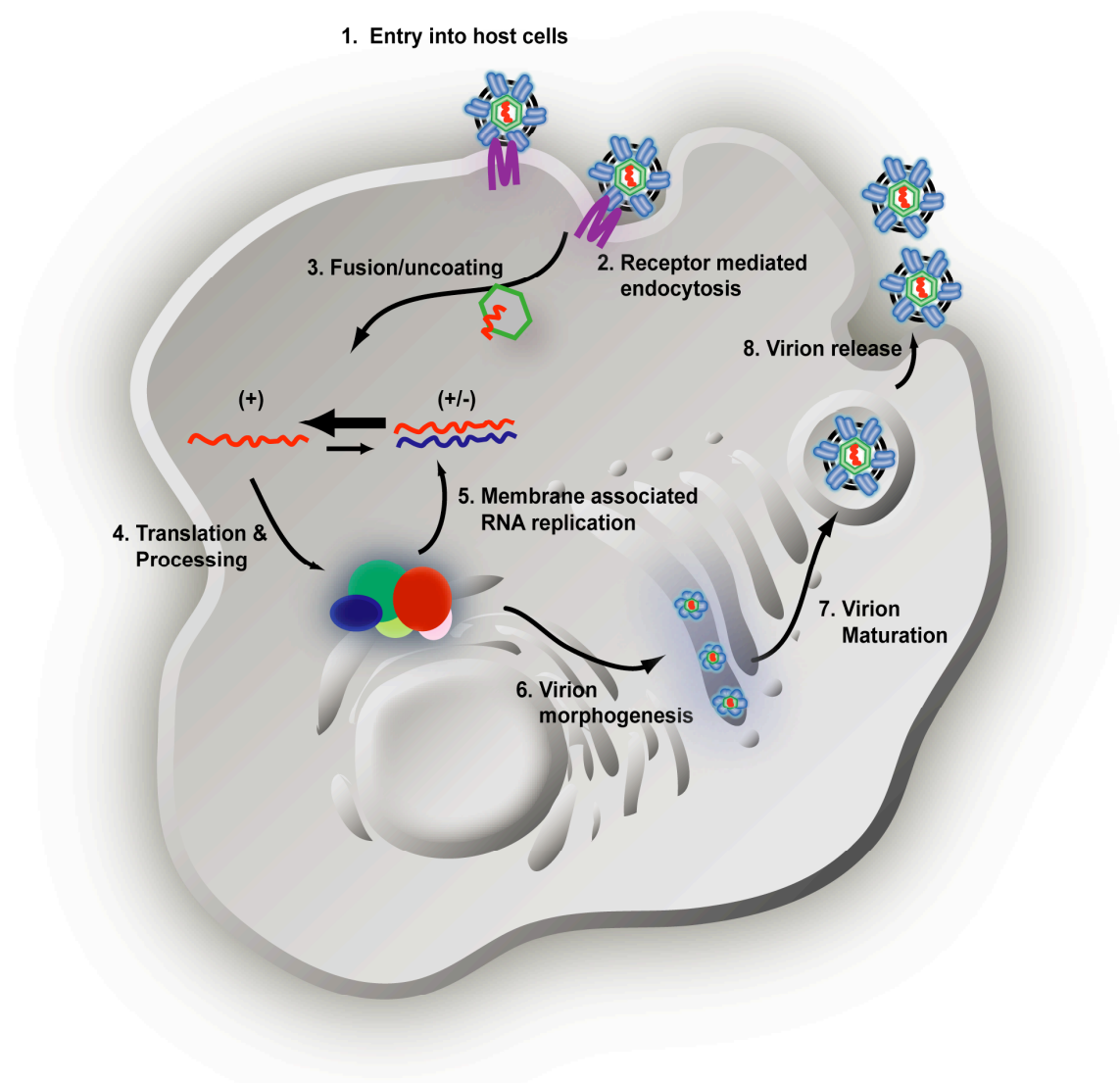


Figure 1.4. Lifecycle of HCV. Virions bind the cell surface and are internalized by interactions with host cell receptors. Virus RNA is released into the cytoplasm where it undergoes translation of the polyprotein. Membrane-associated replicase complexes are formed which synthesize virus RNA via a minus strand intermediate. Genomic RNA is undergoes subsequent rounds of translation and establishing replicase compartments. Eventually genomic RNA is packaged by virus structural proteins and released from the cell.

Translation begins with the assembly of a minimal set of cellular initiation factors to synthesize the entire polyprotein, which is cleaved into the individual virus proteins<sup>58, 98</sup>. The replicase complex is built along the ER membranes, and the translated virus RNA becomes the template for RNA synthesis. Little is known about the transition from translation to replication, or how the virus RNA is recruited into the membrane complex. The RNA polymerase initiates replication at the 3' terminus of the positive sense genome to synthesize a full length complementary minus RNA strand, which is in turn used as a template for synthesis of genomic plus strand RNA. RNA amplification through the minus strand has not been thoroughly described for HCV, but previous work on other classical flaviviruses suggests that the minus strand RNA is base paired with the genomic RNA strand. This is referred to as the double-stranded replicative form, RF. The RF serves as the template to produce multiple positive-strand RNA molecules creating an unequal ratio of 10-100 positive to 1 negative strand RNA molecule in the cell.

Since translation of the polyprotein expresses equal molar ratios of each virus protein, and no subgenomic transcripts are generated, researchers have hypothesized about additional functions for both the structural and nonstructural virus proteins possibly acting as scaffolding or shuttle proteins

coordinating the virus RNA between the two processes of translation and replication. Replication complexes can be isolated from cells harboring replicating HCV. In such complexes, both positive and negative strand virus RNAs were protected from nuclease digestion and, by using protease digestion, only a small fraction of nonstructural proteins were found to contribute to replication <sup>108</sup>. This indicates that the majority of the expressed nonstructural proteins may have a role supplemental to that played in RNA replication.

The mechanism of virus assembly has only recently begun to be studied. The capsid protein functions in virus assembly and domains of the capsid protein required for self assembly have been described <sup>86, 94</sup>. The capsid protein has been observed to associate with cellular lipid droplets to act in virus assembly <sup>117</sup>. The E1 and E2 glycoproteins associate with each other through their transmembrane domains, but their interactions with the capsid protein or other structural proteins during the assembly process is undefined. Compensatory mutations have been found in p7 and NS2 that enable production of infectious virus <sup>164</sup>. Neither protein is thought to be incorporated into virions, and thus act during the assembly process in a yet undetermined manner.

### ***Model systems to study HCV.***

#### ***Animal models.***

The study of HCV has been a challenge due to the inability to isolate virus from infected patients into cell culture. In addition, HCV displays a host range restricted to humans and the chimpanzee. Chimpanzees become infected with HCV after inoculation with serum from infected humans or chimpanzees, direct hepatic inoculation of *in vitro* transcribed HCV RNA, or from tissue culture derived virus<sup>83, 139</sup>. The chimpanzee was the key animal model used to identify the causative agent of NANBH as the RNA virus HCV; inoculation of a chimpanzee with RNA transcripts from an HCV cDNA consensus clone led to infection and hepatitis, proving HCV is the causative agent of disease<sup>65</sup>. Direct intrahepatic inoculation of HCV RNA transcripts, is still the most stringent system to test the *in vivo* importance of HCV genetic elements. This method was used to demonstrate the necessity of p7 and the RNA element 3'X region for HCV viability<sup>119</sup>. Ethical and expense considerations make the chimpanzee model less than ideal, but the similarities between human and higher primate immune systems have enabled categorization of HCV-specific T cell epitopes, the presence of neutralizing antibodies, and study how the diversity of the virus allows escape from the host defenses.

Many studies have followed infections launched with monoclonal HCV genomes generated from cDNA clones and have established that chimpanzee and human HCV infections share similar features. Circulating HCV RNA can generally be detected 7-10 days after infection <sup>123, 126</sup>. During the acute phase of weeks 2-20, HCV RNA levels increase and liver damage becomes apparent within 7-10 weeks as indicated by a rise in serum alanine aminotransferase (ALT) <sup>24</sup>. The chimpanzee liver also displays signs of pathology, with inflammation and focal necrosis, but chimpanzees do not progress as frequently to severe liver disease, such as fibrosis and carcinoma, as do humans. Seroconversion occurs generally between 15-20 weeks <sup>111</sup>. Neutralizing antibodies have been isolated from both human and chimpanzee infections; but appear unable to clear the virus <sup>144</sup>. Levels of circulating HCV RNA show great variability in both species. Chimpanzee infections usually generate  $10^4$  to  $10^7$  HCV RNA copies/ mL during the acute phase. Greater variability is detected in chronically infected humans ranging from  $10^2$  to  $10^8$  HCV genome equivalents/ mL <sup>154</sup>. Like humans, chimpanzees fail to spontaneously clear the virus during the acute phase and progress to chronic infection 60-70% of the time <sup>113</sup>.

Spontaneous clearance has been associated with generation of a diverse adaptive immune response. CD8<sup>+</sup> and CD4<sup>+</sup> T cells have been

isolated from the periphery and the liver and HCV-specific epitopes have been mapped. A correlation has been proposed between the diversity of virus epitopes recognized by the host T cells and the frequency of clearance<sup>30, 112, 158</sup>. HCV specific, interferon gamma producing T cells have also been shown be present in the liver of chimpanzees that have cleared the infection<sup>116, 137, 143</sup>. Development of a response by both CD8<sup>+</sup> and CD4<sup>+</sup> T cells has been shown to contribute to HCV clearance<sup>45, 127</sup>. In both chimpanzee and human infections, HCV displays a similar mutation rate of  $1.44 - 1.92 \times 10^{-3}$  base substitutions per genome per year<sup>95</sup>. The generation of virus sequence diversity is thought to allow epitope escape permitting chronic infection.

### ***Chimeric mouse.***

Establishment of a small animal model would provide a more convenient and economical alternative to study HCV. HCV inoculation of other primate monkey species, such as marmosets and tamarins, and also of tupaia and mice have failed to yield a reliable model for HCV infection. A chimeric mouse, which can express transplanted human liver cells, has been shown to replicate HCV<sup>87</sup>. In this system, an immunodeficient SCID mouse was crossed with a transgenic mouse carrying the urokinase-type plasminogen activator gene controlled by the albumin promoter (Alb-uPA).

The Alb-uPA transgene results in programmed cell death of endogenous hepatocytes in the homozygous mice, thus enabling repopulation by injected human liver cells. Inoculation of serum from human HCV patients resulted in infection of these chimeric mice for 15-17 weeks with titers of  $10^4$  to  $10^6$  HCV RNA copies/ mL in the blood. HCV could also be passed to other chimeric mice. There are numerous limitations for this model as the mice are difficult to breed and maintain because of the precarious state of their immune system. In addition, the mice must undergo a successful transplant of human liver cells, and this procedure often results in death. Studies of the immune response to HCV are also compromised by the SCID status of the chimeric mouse.

### ***Cell culture systems.***

#### ***Cell lines.***

The study of HCV replication in cell culture has been challenging. The parental Huh-7 and subcultured Huh-7.5 cell line can support the most efficient replication for HCV genomes of genotypes 1a and 1b. Nonhepatic cell lines such as HeLa and 293 and the hepatic cell line HepG2 have been reported to replicate genotype 1 HCV, but not to levels equivalent to Huh-7 and the Huh-7.5 cells. The 2a sequence JFH-1 has demonstrated robust

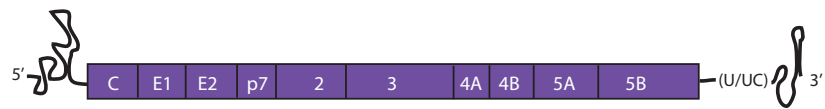
replication in Huh-7 and Huh-7.5 cells as well as HepG2, 293, HeLa, and even primary mouse liver cells <sup>61, 148</sup>. Primary hepatocytes are notoriously difficult to faithfully culture as they often de-differentiate in culture and lose specific functions and markers. The development of stem cell derived liver cells and three dimensional culture systems have not been reproducibly shown to easily propagate HCV replication.

### ***HCV RNA replication in cell culture.***

HCV replication in cell culture was first achieved with the use of a subgenomic replicon that contained only the replicase proteins NS3 to NS5B for the genotype 1b Con1 sequence (Figure 1.5). The replicon included the 5' and 3' NTRs of HCV, and expressed a drug selectable marker by the HCV IRES. A second cistron expressed the HCV replicase proteins behind the encephalomyocarditis virus (EMCV) IRES. Expression of the selectable marker, neomycin phosphotransferase (Neo), conferred resistance to the drug G418 and thus enabled selection for cells capable of harboring HCV replication <sup>79</sup>. Replication of these subgenomic, bicistronic replicons in Huh-7 cells allowed selection of more fit genomes. Adaptive mutations were discovered in NS3, NS4B, NS5A, and NS5B that enabled a substantial increase in the number of Huh-7 cells selected to replicate HCV <sup>12, 78</sup>. These



A



B

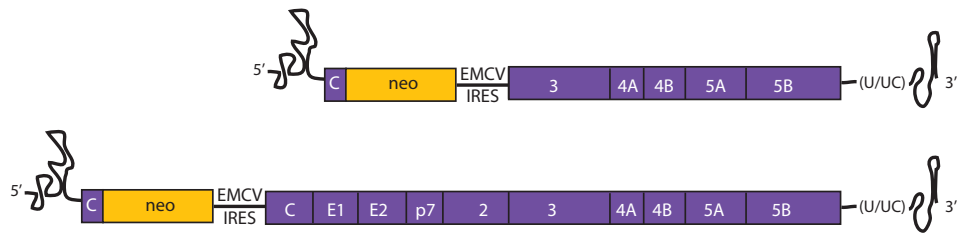


Figure 1.5. A. Diagram of the molecular clone of the full length HCV genome. B. Bicistronic replicons expressing either the subgenomic replicase proteins or the entire polyprotein from the EMCV IRES.

adaptive mutations, however, were shown to be specific only for replication in cell culture, and HCV genomes containing these mutations were not infectious in the chimpanzee <sup>25</sup>. Using this strategy, HCV sequences from genotype 1a and other 1b isolates were shown to replicate in Huh-7 cells in the context of a bicistronic genome <sup>79</sup>.

During the selection of subgenomic replicons, cell colonies were isolated and found to contain no adaptive mutations in the HCV sequence. These cells were allowed to expand, and “cured” of the HCV RNA by treatment with interferon gamma <sup>14</sup>. This Huh-7 subculture line, designated as Huh-7.5, was shown to have an increased capacity to support replication of HCV. It is not entirely clear which cellular factors contribute to HCV replication or are absent to impede replication, but Huh-7.5 cells have been shown to be deficient in the RIG-I antiviral pathway <sup>38</sup>.

Bicistronic RNAs expressing the entire polyprotein of HCV behind the EMCV IRES were shown to replicate in Huh-7 cells, but at a much lower efficiency than the subgenomic RNAs. These full-length, bicistronic genomes did not produce infectious virus despite the inclusion of the structural proteins <sup>8</sup>. Using the highly permissive Huh-7.5 cell line, transient replication could be measured for full-length genomes containing cell adaptive mutations, but lacking a selectable marker.

All HCV genomes shown to replicate in cell culture required adaptive mutations until the isolation of a genotype 2a sequence from a patient displaying fulminant hepatitis. This sequence, JFH-1 (Japanese Fulminant Hepatitis) could replicate without adaptive mutations and was subsequently shown to produce infectious virus in cell culture <sup>151</sup>. The determinants for virus production in cell culture have yet to be revealed, but inter- and intra-genotypic chimeras have been shown to be infectious when structural proteins are fused to the replicase proteins of JFH-1 <sup>102</sup>.

### ***Translation of HCV.***

HCV replication can only begin after the genome has been translated to produce the necessary replicase complex. Host cell machinery initiates polyprotein synthesis from the IRES located in the 5'NTR. The 5' NTR contains 4 stem loop domains and a pseudoknot mapped to positions 44-354 (Figure 1.6). The first stem loop, 5-20, is required for replication, but not for translation and when deleted can actually enhance protein expression <sup>41</sup>. The following region, 21-43, is predicted to be unstructured in the plus sense, but is thought to form an RNA element with SLI in the minus sense and act as a promoter for plus strand synthesis (Figure 1.7). SLII marks the beginning of the IRES between positions 38 and 43 and influences binding of the 40S

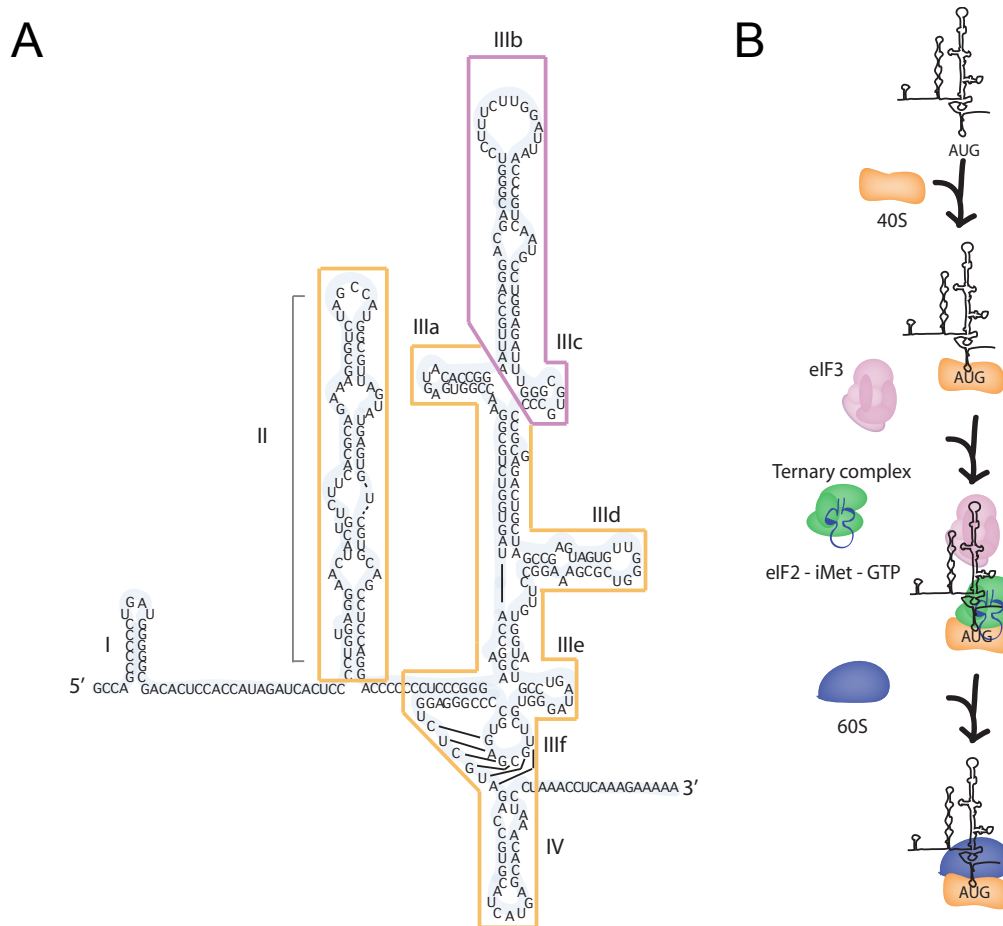
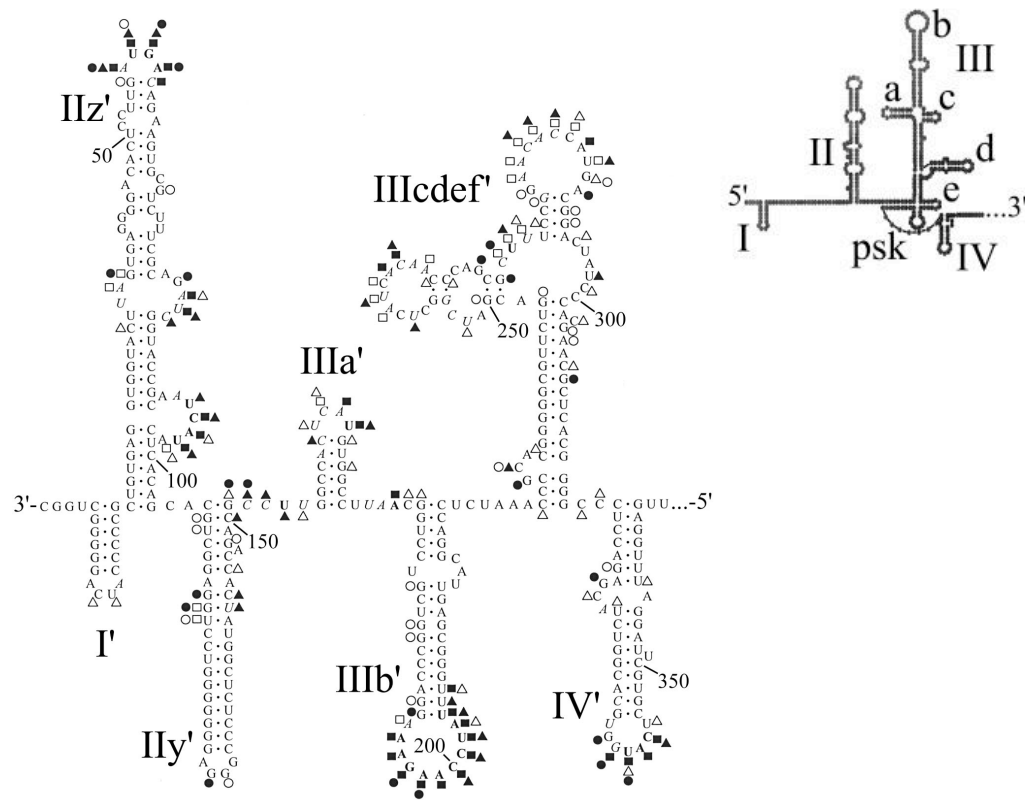


Figure adapted from Fraser et al. 2007

Figure 1.6. The 5'NTR of HCV and the initiation pathway. A. The IRES boundaries are demarcated with a yellow line. eIF3 binds the apical stem of SLIII, and the 40S subunit binds the basal stem of SLIIIId. SLII induces a conformational change of the 40S subunit to position the ribosomal P site at the start codon positioned in the pseudoknot of SLIV. B. The 40S subunit binds the SLIII and is positioned at the AUG start codon. eIF3 and the ternary complex join by contacting the top of SLIII to form the 48S\* complex. After GTP hydrolysis, the 60S subunit joins and forms the 80S complex.



Adapted from Smith et al. 2002

Figure 1.7. Replication elements in the 5'NTR participate as promoters for plus strand synthesis. Predicted RNA secondary structure of the 3' minus strand. The stem structures are numbered based on the RNA structures in the genomic sense (see small IRES structure).

ribosomal subunit and eIF3<sup>77, 136</sup>. The 40S ribosomal subunit and eIF3 bind two separate stem loops of a four-way junction of SLIII<sup>62</sup>. SLIV harbors the initiating AUG codon of the polyprotein and participates in an RNA pseudoknot<sup>52</sup>.

Translation from the HCV IRES has been compared to prokaryotic translation as the 40S ribosomal subunit binds directly to the IRES, which acts in a fashion analogous to a Shine-Dalgarno sequence, without the need for canonical eukaryotic initiation factors such as the eIF4F complex<sup>50, 115</sup>. Initiation proceeds as the 40S subunit binds the basal SLIII d, and eIF3 binds the apical region of SLIII a,b,c. SLII induces a conformational change in the 40S subunit to position the start codon into the ribosomal P site<sup>115, 136</sup>. The ternary complex of eIF2-GTP-Met-tRNA joins to form the intermediate 48S\* complex<sup>98</sup>. After GTP hydrolysis, eIF2 releases the initiator tRNA and disassociates. A second GTP hydrolysis event with eIF5B allows the 60S subunit to join and form the 80S complex. Translation elongation proceeds without any scanning activity.

Single particle cryo-electron microscopy has shown the IRES to mediate conformational changes in the 40S subunit that permit access of the A site to the AUG located in stem loop IV<sup>136</sup>. Mutation of the AUG start codon to AUU or CUG does not affect IRES mediated translation and an

increase in  $Mg^{2+}$  concentration allows translation to proceed without the need for an initiating tRNA<sup>69, 114</sup>. The HCV IRES differs from other virus IRES such as the polio and EMCV IRES, because the boundaries extend downstream into the coding sequence<sup>122</sup>. The first 33 nucleotides of the downstream capsid sequence are required for efficient translation, but providing the full length of the capsid sequence can suppress cap independent translation<sup>37, 124</sup>. It has been demonstrated that the sequence after the AUG must also be free of secondary structure<sup>52</sup>. After docking, the 40S subunit must melt the pseudoknot structure in SLIV to proceed into the polyprotein coding region. Stabilization of the pseudoknot has been shown to inhibit translation, indicating that this RNA structure provides a regulatory role<sup>53</sup>. The pseudoknot has also been predicted in the related classical swine fever and GBV-B viruses. In the pestivirus bovine viral diarrhea virus (BVDV), lack of stable secondary RNA structure rather than a specific RNA sequence was important for optimal translation initiation<sup>93</sup>.

### ***Influence of the capsid gene on translation.***

The initial 33 nucleotides of the capsid gene are required for efficient IRES activity<sup>37</sup>. Inclusion of longer regions of the capsid gene after the HCV IRES has been shown to be repressive for translation of reporter

constructs and has sparked a debate over whether the capsid protein or the capsid gene has a role in regulating translation.

Proponents of the capsid protein being the repressive element have cited the ability of the capsid protein to preferentially bind the HCV 5'NTR over other RNA sequences, specifically SLI, the unstructured region of 23-41, and SLIIId of the IRES<sup>17, 36, 124, 125, 140</sup>. When the capsid protein is expressed by a recombinant baculovirus *in trans*, translation of an HCV IRES driven reporter is reduced in cells<sup>121</sup>. Another group reports capsid expression must occur *in cis* and insertion of a frameshift early in the capsid sequence alleviates the repression of translation.

Proponents of the capsid sequence down regulating HCV translation argue the capsid protein interacts nonspecifically with RNA including other virus RNA, tRNA, and even ribosomal RNA. Other virus capsid proteins can also bind the HCV IRES as efficiently as the HCV capsid protein. Expression of  $\beta$ -gal from a recombinant baculovirus resulted in lower translation activity from the HCV IRES and introduction of a frameshift early in the capsid gene still conveyed less translation than a reporter with minimal capsid gene sequence. The debate continues.



### *Conservation of the capsid gene.*

Prodigious numbers of HCV particles are generated in an infected individual,  $\sim 10^{12}$  per day. Due to error prone replication, HCV has a high rate of mutation, causing each replicated genome to bear at least one change from its template. Over time mutations accumulate, sequence diversity expands, allowing rapid evolution. Viruses must adapt to changing intracellular environments and evade host defenses while maintaining the functions necessary for replication. Compact genomes and multifunctional elements assist in meeting these demands, yet constrain regions from random mutation. In the face of such strong selective pressures, viruses are positively selected where only progeny with an adaptive mutation can survive. By comparing virus sequences from an infected individual, patient H in 1977 and in 1990 (referred to as H77 and H90), the evolutionary rate for HCV was estimated to be  $2 \times 10^{-3}$  mutations per site per year<sup>95</sup>. A slower mutation rate of  $4-7 \times 10^{-4}$  sites/year was observed within a cohort of infected individuals<sup>131</sup>. The error threshold is the mutation rate beyond which viral populations cannot be sustained due to accumulation of deleterious mutations and is inversely proportional to the genome length. Estimating the HCV genome to be  $10^4$  nucleotides, a  $10^{-4}$  error threshold

value is predicted; thus HCV replication occurs at or beyond the critical error limit.

Despite a high mutation rate, there are notably conserved regions of the HCV genome that must possess functions indispensable for the virus life cycle. Ina reported a low level of synonymous substitutions in the capsid gene of HCV and subsequent studies have confirmed this observation with more divergent sequences<sup>57</sup> (Figure 1.8). Pressure to maintain virus protein function would prevent an excessive number of nucleotide changes in the 1<sup>st</sup> and 2<sup>nd</sup> codon positions, but does not explain conservation of 3<sup>rd</sup> base positions that do not alter the amino acid sequence. Ina also noted a release of sequence constraint near the termination of the +1 capsid protein. Smith *et al* confirmed suppression of sequence variability in the capsid gene and in addition noticed sequence constraints located in the NS5B region of the genome<sup>133</sup>. It is interesting that both of these genes are adjacent to areas with highly conserved *cis* acting regulatory elements of RNA structure in the 5' and 3' NTRs. The transition to transversion ratio was also higher in synonymous sites (16:1) than in nonsynonymous sites (1.4:1) of the capsid region, indicating that preservation of the type of nucleotide is important. Walewski *et al* reported that codons conserved at 3<sup>rd</sup> base positions in the capsid gene of all genotypes<sup>153</sup>. Five clusters of conserved codon usage

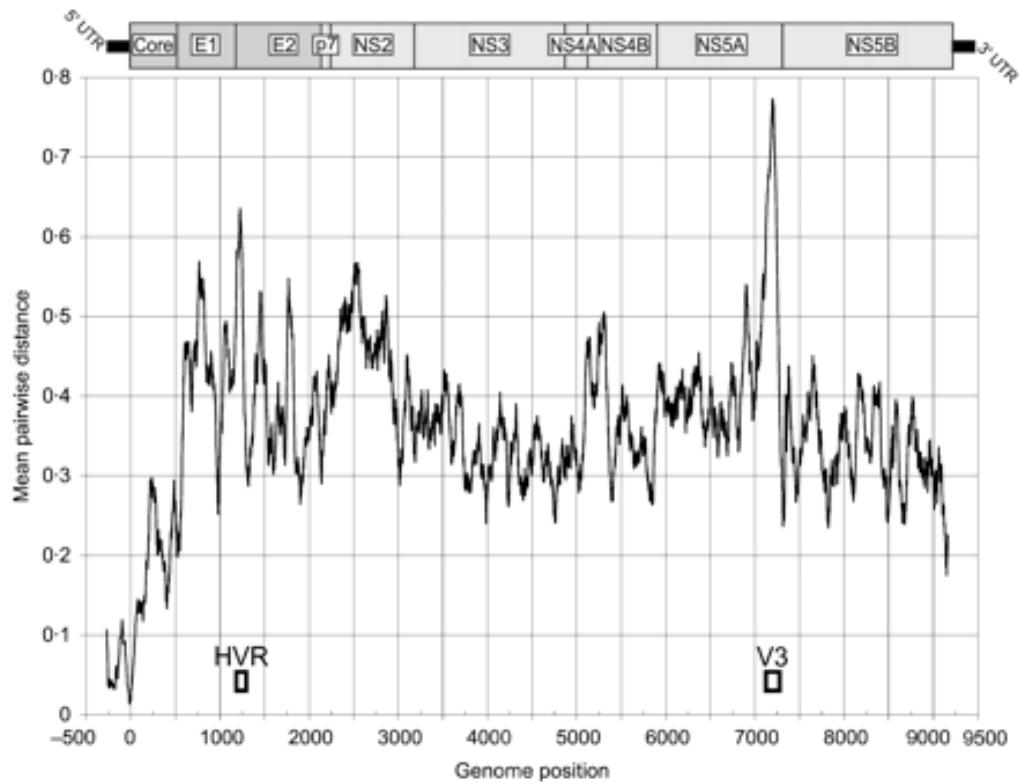


Figure 1.8. Sequence conservation of regions in the HCV genome are shown by measuring the distance with an overlapping 150 bp window of the nonsynonymous/ synonymous mutation rate. This analysis reveals highly conserved regions in the 5' UTR and core regions, and high viral diversity in envelope genes and NS5A. Adapted from Smith et al. 2002.

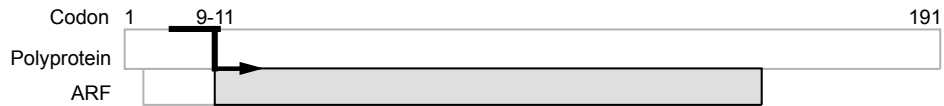
were found in the HCV genome with 22 conserved codons in the capsid region. Another study contradicts the phylogenetic evidence for an ARF-encoded protein citing the distribution of stop codon between genotypes indicates a lack of functional constraints<sup>31</sup>. These reports indicate that the conservation of codons could be attributed to the translation of the overlapping reading frame into a protein, or the presence of RNA secondary structure elements, that are essential for HCV viability.

***Evidence for an ARF protein in HCV replication and infection.***

In 1994, Ina first proposed the possibility of HCV using an overlapping reading frame to encode a protein as an explanation for the conservation of sequence in the capsid region. All HCV genotypes have an open reading frame in the +1 frame of the capsid gene beginning at nucleotide 346, with the polyprotein beginning at 342, but the length of the open frame varies between genotype (Figure 1.9). The +1 frame for genotype 1a is open for 486 bases and could produce a 17 kDa protein, genotype 1b is open for 429 bases potentially encoding a 15 kDa protein, and genotype 2a is open for 375 bases that could express a 13 kDa protein. The +2 frame contains numerous stop codons. Walewski *et al* have termed the +1 frame ARF for alternative reading frame, and the potential encoded

A

Frameshift (F) protein



B

Dual frameshift (DF) protein



Figure 1.9. The proposed frameshift sites of the F and DF proteins.

A. The F protein initiates at the start of the polyprotein and shifts into the ARF at codon 9-11. Translation terminates at a stop codon in the ARF. B. Translation of the DF protein also initiates at the start of the polyprotein, yet shifts further downstream into the ARF at codon 42. Translation continues until reaching a stop codon where expression is terminated or stimulated to shift back into the frame of the polyprotein.

protein, ARFP. The synonymous substitution bias of the capsid gene preserves the amino acid sequence of the ARFP, and there is 65-82% identity between all genotypes. Walewski *et al* aligned and compared 255 HCV sequences to conclude that the potential ARFP is as conserved as other conventional HCV proteins. ARFP has little amino acid identity with the capsid protein, 16%, but both are predicted to be very basic with a pI of approximately 12<sup>153 152</sup>. Proponents of the evolutionary principle of gene overprinting, where a novel protein is created by utilizing an existing coding region, remark that the chemical composition of the overlapping protein is often polarized due to amino acid restrictions imposed by the original protein. The restrictions are due to the codon usage of the 2<sup>nd</sup> base position and could explain the similar amino acid composition.

Experimental evidence for a protein encoded by the +1 frame can be traced to early studies investigating HCV polyprotein processing. There are two forms of capsid protein with sizes of 23/21 kDa and 21/19 kDa. The size discrepancy is due to ambiguous protein markers used between different laboratories. The full-length capsid protein has a hydrophobic region in the carboxy terminus that serves as a signal sequence to deliver the proceeding E1 protein into the lumen of the ER. The host cell signal peptidase cleaves capsid from E1 in the ER and the signal peptide peptidase cleaves the

remaining capsid protein from the ER. The smaller, completely processed form of capsid predominates in cell culture. While investigating these processing patterns, a smaller protein of about 16 kDa was occasionally detected in cell free translation reactions programmed with the capsid gene<sup>76</sup>. This 16 kDa protein was initially believed to be another processed form of the capsid protein, until the laboratory of James Ou began investigating the presence of the 16 kDa protein, named P16, by comparing two HCV genotype 1a sequences, HCV-1 and HCV-RH (Figure 1.10). HCV-1 is the first HCV sequence to be identified by Choo et al; HCV-RH is another name for the H77 sequence<sup>27</sup>. Using *in vitro* translation assays, capsid gene sequences from HCV-1 produced only P16 whereas HCV-RH made only a 21 kDa protein (P21). There are 4 amino acid differences between HCV-1 and HCV-RH capsid protein sequences. Domain swapping experiments and site directed mutagenesis revealed that codon 9 was critical for P21 or P16 production. At codon 9, HCV-1 expresses a lysine (codon AAA) and HCV-RH an arginine (codon AGA). The adenine to guanine difference in HCV-RH interrupts a stretch of 10 consecutive adenines present in HCV-1. HCV-1 is the only isolate with 10 consecutive adenines at this position, in an otherwise conserved A-rich region. Tagging experiments determined that P16 and P21 initiated at the same AUG, but differ at the carboxy termini.

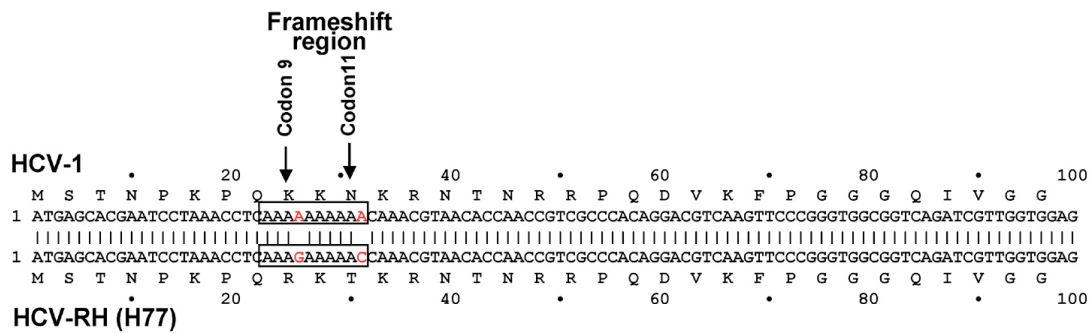


Figure 1.10. Comparison of the HCV-1 and HCV-RH genotype 1a capsid sequences. The sequence HCV-1 contain 10 consecutive adenines at positions 23-33, which are proposed to conduct a ribosomal frameshift event. The sequence of HCV-RH has a guanine at position 26, interrupting the stretch of adenines and is not reported to frameshift.



Xu et al proposed that synthesis of P16 occurs by a ribosomal frameshift into the  $-2/+1$  reading frame and renamed this the Frameshift or F protein (Figure 1.9) <sup>161</sup>. Sequencing of the F protein did not precisely define the shift site, but suggested that it occurred at or near codon 11 and only required codons 8-14. Varaklioti *et al* measured the ability of HCV to frameshift by cloning the luciferase gene after HCV-1 and HCV-H77 capsid genes in all three reading frames <sup>149</sup>. In cell free translation assays programmed with the HCV-1 sequence, ribosomal frameshift into the +1 frame was detected 35% of the time. A  $-1$  frameshift was also detected 6% of the time when only 9 A's were in the slippery sequence. The HCV-H77 sequence, which does not have consecutive adenines, only demonstrated frameshifting less than 1% of the time. Xu *et al* reported the ability of Huh-7 cells to support ribosomal frameshifting 1-2% as measured by a luciferase reporter construct <sup>161</sup>. A triple decoding strategy has also been proposed where both the +1 frame and the +2 frame are expressed. The +2 frame contains eight regularly spaced stop codons before amino acid 120 of the ARF, but using an in vitro translation system, a small 1.5kDa protein was detected <sup>26</sup>.

The downstream E1 gene has been implicated to have a role in P16 production <sup>75</sup>. When the capsid and E1 genes of HCV-1 were used in cell

free translation systems, P21 and P16 were produced, whereas HCV-RH again only expressed P21. Interestingly, both HCV-1 and HCV-RH were able to generate both P21 and P16 in a transient mammalian expression system in HepG2 cells. If the HCV sequence only contained the capsid gene and did not include E1, HCV-1 only made P16 and HCV-RH only made P21. Different cellular localizations were reported. P16 was found in the nucleus and P21 in the cytoplasm, but these systems also lacked the majority of the polyprotein <sup>75</sup>. The HCV capsid protein has been reported to be transported to the nucleus when no other HCV proteins are coexpressed <sup>56</sup>. Studies have investigated the cellular localization of the F protein in transient expression systems of HepG2 cells. The F protein was only detected if the sequence between codons 9 and 11 contained the consecutive adenines <sup>118, 160</sup>.

HCV infected patients also generate antibodies against the F protein. HCV patient sera have been used to detect P16 in the above translation assays and also immunoprecipitate recombinant F protein. Walewski *et al* reported 13/100 chronic HCV patient sera recognized peptides based on the +1 frame by western blot compared to 3 out of 104 healthy controls patients with non-HCV liver disease <sup>153</sup>. Patients infected with different HCV genotypes (1a, 1b, 2b, and 4) have also been reported to produce antibodies

against F by ELISA <sup>66</sup>. Another group sequenced the HCV capsid gene from HCC patients and found 3/10 had non-arginine codons at codon 9, but none had a lysine <sup>96</sup>. When these sequences were expressed in an *in vitro* translation assay, P16 was produced.

Another ARF protein was detected using the +1 overlapping reading frame of the capsid gene. Boulant *et al* isolated a heterogeneous population of *E.coli* expressed capsid proteins using a carboxy terminal affinity tag. Protein sequencing revealed that a +1 frameshift had occurred at codon 42 and upon encountering a stop signal, a –1 shift at codon 144 placed the ribosome back into the capsid reading frame to continue producing the polyprotein <sup>18</sup>. This dual frameshift protein, DF protein, is the first detection of use of the overlapping reading frame usage by genotype 1b (Figure 1.9). Using a transient expression system, the authors showed expression of the DF protein in HepG2 cells.

In summary, there are two reported proteins encoded from the +1 frame of the capsid gene, both involving a translational frameshift. The F protein shares the first 9-11 amino acids of capsid and then moves into the +1 frame. The DF protein uses a frameshift site at codon 42 to enter the +1 frame and shift back into the capsid sequence at codon 144 (Figure 1.9).

### *Summary of thesis work.*

My thesis research began with a desire to learn classical virology with molecular methods to further our understanding the replication of an infectious agent important to public health. Focusing on HCV, I also wanted to examine a function of replication that was not specific for, or limited to, cell culture. Given the early difficulties of studying HCV replication, this attempt could have been very brief and without much success! However, I benefited from joining the field at an exciting time when advances in the methods to study HCV replication were rapidly evolving, and I was in a lab that quickly adopted these new practices and contributed novel systems to the field.

The strong conservation of the capsid gene sequence, based on naturally occurring HCV isolates, was persuasive evidence that the capsid gene, and not simply the capsid protein, had functional significance in the HCV life cycle. Two lines of evidence were in the literature: an expressed, alternative reading frame and RNA secondary structures. I chose to test the requirement for an ARF in HCV infection as the mechanism of translational frameshift allows for many interesting studies of virus-host interactions. Translation from the HCV IRES was already shown to be novel during which a minimal set of initiation factors conduct expression of a ~3000

amino acid protein. The characterization of a newly identified virus protein also promised many questions to solve. The inability to identify a frameshift protein in cell culture hampered these efforts, and the focus shifted to an *in vivo* model.

Chapter 3 describes the investigation for a requirement of an ARF in a natural infection. This approach involved the introduction of silent point mutations conferring multiple stop codons in the ARF. In collaboration with Stephen Feinstone, an HCV genome with these Stop mutations was inoculated into a chimpanzee. Using the conventional sequence of H77, which has been the parent genome for over 20 chimpanzee studies, stop codons were introduced into the ARF without altering the capsid protein sequence. The Rice lab has been involved with numerous studies investigating the pathogenesis and molecular determinants of HCV infection in the chimpanzee model, and has shown this approach to provide definitive evidence for factors that are required for replication.

An infection was launched, yet attenuated. The ARF was therefore not required for infection, but could be responsible for the low virus load and little signs of pathology. We measured the immune response in both the periphery and liver and detected a diverse set of HCV specific T cell epitopes corresponding with the time of virus decline. Sequence analysis of

the circulating virus RNA revealed a reversion at the site of one of the point mutations conferring a stop codon in the ARF. This reversion was selected at the beginning of the acute infection and became fixed in the population. As the infection progressed, additional reversions were identified and indicated selection of the capsid sequence as opposed to an ARF protein.

Chapter 4 describes how mutations conferring the Stop codons in the ARF were shown to be detrimental to HCV replication in Huh-7.5 cells. Mutagenesis studies confirmed the silent mutations of the ARF stop codons actually disrupted an RNA structure. Two RNA structures had been proposed for the capsid gene based on covariant sequence analysis. Mutations in the structure of SLVI were shown to reduce HCV replication, with disruptions in the apical region of the SLVI resulting in the most severe effects. Altering the stability of the base produced a slight reduction in replication, but contributed to a further reduction when combined with mutations in the apical stem. Passage of Huh-7.5 cells transfected with HCV genomes containing a disrupted SLVI identified reversions selected to restore the RNA base pairs of SLVI. These reversions were identical to the reversions selected during the chimpanzee infection and indicated that comparable selective pressures exist to maintain a functional RNA structure during replication in cell culture and *in vivo*.

Chapter 5 describes how the search for a mechanism of SLVI in HCV replication returned the focus on HCV translation. The capsid gene had been shown to have a potential role in regulation of translation, but a specific interaction had never been identified. The mutations in SLVI conveyed a reduction in IRES mediated translation. A 2- 4 fold reduction was observed in transfected Huh-7.5 cells and in cell free assays using rabbit reticulocyte, Huh-7, and Huh-7.5 cell extracts. Mutations in SLVI were shown to reduce translation independent of downstream capsid sequence or of the 5'NTR, contradicting published reports. We showed that the effects of SLVI depend on its position in the HCV genome. We can abrogate the reduced replication due to mutations in SLVI by expression of the replicase proteins *in trans* and out of the context of SLVI and the IRES.

We propose that mutations in SLVI reduce IRES-mediated translation, which consequently lowers the amount of replicase expression leading to a reduced level of synthesized RNA in the cell. The impact of mutations in SLVI is most evident at 24 hours post transfection, a time when translation and replication activity begins to accelerate. The lower efficiency of translation conveyed by mutations in SLVI becomes amplified during this period and if too severe, replication of the mutant cannot be sustained in the cell. Replication of mutant SLVI can proceed in a minor subpopulation of

Huh-7.5 cells, implying either a host factor or an infrequent state of the cell can overcome this translation defect.



## **Chapter 2**

## **Chapter 2. Materials and Methods**

### ***Cell culture.***

Cell monolayers of the human hepatoma cell lines Huh-7 and Huh-7.5 were grown at 37 °C, 5% CO<sub>2</sub> in Dulbecco's modified minimal essential medium (DMEM; Invitrogen, Carlsbad, CA) supplemented with nonessential amino acids, 10% fetal bovine serum (FBS), 100 U of penicillin per ml, and 100ug of streptomycin per ml. Cells were passaged after treatment with 0.05% trypsin-0.02% EDTA and seeded at a dilution of 1:4.

### ***Plasmid construction.***

The HCV genotype 1a H77 infectious clone (GenBank accession number: AF009606) <sup>65</sup> was used to generate RNA for the chimpanzee infection. The HCV genotype 1a H77 full-length genome, pH/FL (L+I) [(P1496L) and (S2204I)] <sup>13</sup>, genotype 1b Con1 (I) [S2204I], genotype 1b Con1 (GIT) [E1202G, T1280I, and K1846T] were used in cell culture replication assays.

The genotype 2a chimera pJ6-JFH <sup>72</sup>, genotype 2a chimera pJ6-JFH/Rluc <sup>146</sup>, genotype 1a-2a chimera pH-JFH/Rluc (V+T) [I348V/S1103T], genotype 1a-2a chimera bicistronic pH-EI-JFH (V+T) (Williams, D.,

unpublished) were used in cell culture replication assays and to produce virus in cell culture.

The plasmid H-JFH/Rluc was created by generating a PCR product containing the Renilla luciferase gene and 2A protease from FMDV using oligos 7041 (5' GGG GGA CGC GTT GGA CAC GAC TTC GAA AGT TTA TGA TCC AG 3') and 7236 (5' CCC CCA CGC GTC GGG CCC TGG GTT GGA CTC G 3'). This product creates MluI sites on either side and inserts the signal sequence for NS2 (amino acids LDT) prior to the Renilla luciferase gene. An extra G was inserted after the 2A protease to maintain the reading frame. PCR was performed by the Expand High Fidelity kit (Roche Applied Science, Indianapolis, IN), the product gel purified, and TA cloned (Invitrogen). The product was verified by sequencing and subcloned as a MluI fragment into plasmid pH-JFH-MluI/V+T, which contains two base changes (A2768G and C2769T) generating a MluI site at the junction of p7 and NS2.

Plasmids containing deletions in the capsid gene were generated by two-step PCR. The first step amplified the flanking regions of the deletion site. A NotI site was included at the junction of the deletion site to allow the two PCR products to anneal. The second PCR step amplified the two PCR products creating a NotI junction at the site of deletion. The deletion of

amino acids 15-171 (pDel 15-171) was generated by PCR using the oligos 6607 (5' GAG TGT CGT GCA GCC TCC AGG ACC C 3'), 7243 (5' ACC TGC GGC CGC GGT GTT ACG TTT GGT TTT TCT TTG AG 3'), 7244 (5' GCG GCC GCA GGT TGC TCT TTC TCT ATC TTC CTT CTG G 3'), and 7245 (5' GCC CAG CGG TGG CCT GGT GTT G 3'). The deletion of amino acids 57-171 (pDel 57-171) was generated by PCR using the oligos 6607, 7241 (5' ACC TGC GGC CGC TTG CGA CCG CTC GGA AGT CTT CC 3'), 7244, and 7245. The plasmid pDel 57-171/ Stop 1,2,3,4 was created using the PCR strategy listed above, but with the template pH-JFH/Stop 1,2,3,4. The fragments of AgeI – MluI were cloned into pH-JFH/Rluc-MluI and their identity verified by sequencing.

The bicistronic plasmids of pH-neo/RRRI and pH-neo/RRRI (Stop 1,2,3,4) were created by two-step PCR. The first reaction used oligos 4016 (5' CCC CAG TGC CAC GTT GTG AGT TGG 3'), 4015 (5' CGT GCT CAT TAT TAT CGT GTT TTT CAA AGG 3') and amplified the EMCV IRES with an extension to the capsid gene. The second PCR amplified the capsid with an extension for the EMCV IRES and used oligos 4014 (5' CAC GAT AAT AAT GAG CAC GAA TCC TAA ACC 3'), and 4017 (5' CGT GCT CAT TAT TAT CGT GTT TTT CAA AGG 3'). The second step joined the two PCR products using oligos 4016 and 4017. For pH-neo/RRRI

the template for the capsid gene was pH/L+I. For pH-neo/RRRI/Stop 1,2,3,4, the template was pH/L+I/Stop 1,2,3,4. The PCR products were purified using QIAquick spin column (Qiagen, Valencia, CA). The KpnI fragment was subcloned into pH-neo/L+I<sup>13</sup>. The adaptive mutations were introduced by subcloning the fragment BsrGI-EcoRI from pH/RRRI [Q1067R, K1691R, K2040R, S2204I]<sup>163</sup>.

VEE replicons expressing the HCV capsid (VEErepC) and C-NS2 (VEErepC-NS2) from genotype 1a H77 sequence were generated by PCR creating SapI sites flanking the fragment using oligos 7043 (5' TAA ATA AAG CTC TTC AAG CAC GAA TCC TAA ACC TCA AAG AAA AAC C 3'), 7044 (5' TTA TTT AGC TCT TCT TTT AGG CTG AAG CGG GCA CgG TCA GGC 3'), and 7240 (5' TTA TTT AGC TCT TCT TTT AAT CAT TGG TGA CAT GGT AAA GCC C 3'). Products were verified by sequencing and the SapI fragment subcloned into pVEErepGFP-Sap<sup>101</sup>.

### ***Site directed mutagenesis.***

Point mutations were introduced using the QuikChange Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA). A small, intermediate plasmid containing EcoR-KpnI of the 5'NTR and capsid gene was cloned into pRS2 and used as the template for the QuikChange PCR. Products were verified

by sequencing and subcloned using either AgeI-KpnI or EcoRI-KpnI fragments into the parent HCV genome. Table 2.1 lists the mutations introduced into HCV genomes.

***Detection and sequencing of HCV RNA.***

Quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR; detection limit 200 RNA GE/ml) was performed with total RNA extracted from 100 ul serum. For nested RT-PCR, total RNA was isolated using QIAamp Virus RNA mini kit (Qiagen). RT-PCR was performed using Superscript III and Platinum High Fidelity One Step system (Invitrogen). RNA was denatured at 60°C for 5 min followed by reverse transcription at 55°C for 40 min. PCR cycling conditions were 94°C 30 sec, 55°C 30 sec, and 68°C 1 min for 25 cycles. For samples requiring nested RT-PCR, 2 ul of the first reaction was used for semi-nested PCR using Pfx polymerase (Invitrogen). The cycling conditions were 94°C 15 sec, 55°C 30 sec, and 68°C 1 min for 20 cycles. Reverse primer 5'-CCG CCT CGT ACA CAA TAC TCG (nt: 970-990) and forward primer 5'-GTG CCC CCG CAA GAC TGC (233-250) were used. Semi-nested PCR combined the forward primer above and an internal primer, 5'-GGT GAC ATG GTA AAG CCC CG (934-

953). PCR products were either sequenced directly or cloned into pCR2.1 (Invitrogen) and individual clones sequenced.

### ***Chimpanzee inoculation and sample collection.***

Chimpanzee 1602 was inoculated with 0.6 mg transcript RNA by direct intrahepatic injection <sup>65</sup>. Peripheral blood was collected in acid citrate dextrose tubes for isolation of PBMCs, expansion of CD8<sup>+</sup> T cells, and plasma. Liver tissue obtained by hepatic needle biopsy was placed in RPMI for T cell studies or flash frozen for later RNA analysis. Peripheral blood and liver needle biopsies were collected weekly for the first 12 weeks and monthly thereafter. Animal housing, maintenance, and care met the NIH requirements for the humane use of animals in scientific research.

### ***Isolation of PBMCs.***

PBMCs were isolated using Ficoll-Paque Plus (GE Healthcare BioSciences, Piscataway, NJ) density gradient and cultured in AIM-V (Invitrogen) and 2% heat inactivated human AB sera (Gemini Biosciences, Calabasas, CA).

### ***Isolation of CD8<sup>+</sup> and CD4<sup>+</sup> T cells from liver bioposy***

To recover intrahepatic lymphocytes (IHL), liver biopsies were gently homogenized in phosphate buffered saline (PBS) containing 1% FCS. CD8<sup>+</sup> T cells were enriched using anti-human CD8<sup>+</sup> dynabeads (Dyna, Oslo, Norway) and expanded in bulk using anti-human CD3 antibody. CD8<sup>+</sup>, CD4<sup>+</sup> and CD8<sup>-</sup>/CD4<sup>-</sup> cells were seeded in one well of a 24 well plate in T cell clone media: RPMI 1640 (Invitrogen) with L-glutamine (Invitrogen), 5% T-stim culture supernatant without PHA (BD Biosciences, San Jose, CA), 50U/mL recombinant human IL-2 (Peprotech, Rocky Hill, NJ), 10% heat inactivated FCS, and antibiotics (penicillin and streptomycin). Cells were expanded using anti-human CD3 monoclonal antibody (clone X35) (Immunotech, Marseille, France) at 0.05 ug/ml in the presence of 2 x 10<sup>6</sup> irradiated (3000 rads) human PBMCs per well as feeder cells. Cultures were fed every 3-4 days by replacing half of the culture media. Populations of CD4<sup>+</sup> cells were confirmed to have 94.5% purity and CD8<sup>+</sup> cells 98.5% purity by flow cytometry.

### ***Detection of HCV-specific CD8<sup>+</sup> T cells by ELISPOT.***

IFN- $\gamma$  ELISPOT kits (U-cytech, Utrecht, The Netherlands) were used to assay PBMCs and expanded CD8<sup>+</sup> cells. 2 x 10<sup>5</sup> freshly isolated PBMCs



cells were incubated with HCV antigen at 2 ug/mL for 48H and then transferred to 96 well plate coated with anti-IFN- $\gamma$ . For expanded CD8<sup>+</sup> cells, 1 x 10<sup>5</sup> cells/well were incubated with autologous irradiated PBMCs as APCs and HCV antigen for 48H before transferring to 96 well plates coated with anti-IFN $\gamma$ . Plates were developed according to the manufacturer's instructions.

***HCV recombinant antigen and peptides.***

HCV recombinant antigens C22-3 (core aa: 2-120), C33c (NS3 aa: 1192-1457), C100 (NS4 aa: 1569-1931), C200 (NS3-NS4 aa: 1192-1931), and NS5A (aa: 2054-2995) were expressed as C-terminal fusion proteins with human superoxide dismutase in either yeast (*Saccharomyces cerevisiae*) or *Escherichia coli* and were a gift from Dr. M. Houghton (Chiron, Emeryville, CA). Peptides were synthesized by Genemed Synthesis (San Francisco, CA). Peptide pools were comprised of 15-20 mer peptides overlapping by 10 residues. Each pool was composed of peptides corresponding to residues, numbered from the beginning polyprotein: 1-214 (pool 1); 205-766 (pool 2); 757-1055 (pool 3); 1046-1365 (pool 4); 1356-1680 (pool 5); 1671-2009 (pool 6); 2000-2443 (pool 7); 2434-2735 (pool 8); 2726- 3011 (pool 9). Peptides used to screen the ARF were composed of

amino acids 1-161 of the +1 frame, beginning with nucleotide 5 of the polyprotein (pool 10).

***Detection of HCV-specific antibodies.***

HCV-specific antibodies were determined using the HCV EIA-2 assay (Abbott, Abbott Park, IL).

***Generation of CTL Lines.***

CD8<sup>+</sup> T cells enriched using anti-human CD8 dynabeads (Dyna) as described above were cloned at a limiting dilution of 10 or 50 cells/well in 96 well plates. CD8<sup>+</sup> T cells were seeded in T cell clone media and expanded using anti-human CD3 monoclonal antibody (clone X35) (Immunotech) at 0.05 ug/ml in the presence of  $5 \times 10^4$  irradiated (3000 rads) human PBMCs per well as feeder cells. Cultures were fed every 3-4 days by replacing half of the culture media. After 2 weeks, growing cell lines were transferred to 24 well plates and subjected to another round of anti-CD3 antibody stimulation in the presence of  $2 \times 10^6$  irradiated human PBMCs. Cultures were fed every 3-4 days as described above. 306 independently derived CD8<sup>+</sup> T cell lines was tested for cytotoxicity against autologous independently derived CD8<sup>+</sup> T cell lines (BLCL) targets infected with

recombinant vaccinia virus expressing the HCV genome (vv1-3011).

Epitope fine mapping was performed by intracellular staining for IFN- $\gamma$  using peptide matrices spanning the antigen of interest. Peptides were 20 amino acids long overlapping by 10 residues.

***CD4<sup>+</sup> proliferation assay.***

Liver derived CD4<sup>+</sup> T cell clones were tested for proliferative activity by incubating  $5 \times 10^5$  CD4<sup>+</sup> T cells with  $5 \times 10^5$  irradiated (10,000 rads) autologous BLCLs in the presence of synthetic HCV peptide in AIM-HS media (AIM-V lymphocyte medium, Invitrogen) containing 2% human AB serum (Gemini Biosciences) in 96 well flat bottom plates. Four replicates were seeded for each CD4<sup>+</sup> cell clone. Cultures were incubated at 37°C in 5% CO<sub>2</sub> for 2 days. Proliferating cells were labeled with 1uCi/well of [<sup>3</sup>H] thymidine (GE Healthcare BioSciences, Piscataway, NJ) for 18 hours. Cells were collected and washed on filters using an automated cell harvester (Tomtec, Hamden, CT) and the amount of incorporated thymidine measured using a scintillation counter (Beckman Coulter, Fullerton, CA).

### ***Flow Cytometry.***

Populations of lymphocytes were confirmed by 4- color flow cytometry using anti- CD3 (UCHT1) and subset-specific antibodies against CD4 (SK3), CD8 (SK1), and CD56 (B159) (all BD-Pharmingen, San Jose, CA). Cells were resuspended at  $2 \times 10^6$  cells/mL in PBS containing 2% FCS and 0.02% sodium azide and incubated with the appropriate antibody for 1 hour. After washing three times, cells were resuspended in 2% paraformaldehyde solution and analyzed on a FACS Calibur instrument using CELLQuest software. 10,000 events were acquired in the lymphocyte gate. The data was analyzed with program FlowJo (Tree Star, Ashland, OR).

To detect HCV protein in transfected Huh-7.5 cells, a single cell suspension was prepared at a density of  $2 \times 10^6$  cells/ml and fixed in a 2% paraformaldehyde solution in PBS for 20 min at room temperature. The cells were washed and permeabilized in 0.1% saponin- PBS for 20 min at room temperature and stained for 1H with an anti-NS5A monoclonal antibody (MAB7094P 1:50) (Maine Biotechnology, Portland, ME) or an isotype control (IgG2a) diluted in 0.1% saponin- PBS and 1% goat serum. Bound monoclonal antibody was detected by incubation for 1 h at room temperature with anti-mouse immunoglobulin G (IgG) conjugated to Alexa 488 (Invitrogen) diluted 1:1,000 in 0.1% saponin- PBS and 1% goat serum. The

stained cells were washed three times with 0.1% saponin- PBS and analyzed as described above.

***In vitro RNA transcription.***

Plasmid DNA for *in vitro* transcription was prepared from large scale bacterial cultures and twice purified by centrifugation in CsCl gradients. For the chimpanzee study, the cDNA plasmid pH/FL Stop1,2,3,4 was linearized by digestion with BsmI and 3' overhangs converted to blunt ends with T4 DNA polymerase. The reaction was terminated by Proteinase K digestion, phenol-chloroform extraction, and ethanol precipitation. RNA transcripts were synthesized at 37°C for 1.5H in a 200 ul reaction volume containing 40 mM Tris-HCl (pH 7.5), 10 mM NaCl, 18 mM MgCl<sub>2</sub>, 2 mM spermidine, 10 mM dithiothreitol (DTT), 3 mM of each NTP, 160 U of Superscript (Ambion, Austin, TX), 400 U of T7 RNA polymerase (Epicentre Technologies, Madison, WI), and 10 ug of linearized DNA. At the end of the reaction, 10 U of DNase I (Ambion) was added and incubated at 37°C for 20 min to remove the DNA template. RNA was purified using the RNeasy Mini kit (Qiagen). An additional DNase treatment was performed on the column using an RNase-free DNase set (Qiagen). The RNA was eluted in water and the integrity of RNA was confirmed by agarose gel electrophoresis.

Parental and mutant derivatives of genotype 1a pH/L+I, pH-neo/RRRI were linearized with HpaI, genotype 1b Con1 (I) and Con1 (GIT) were linearized with ScaI, and genotype 2a and 1a- 2a chimeras pJ6-JFH, pH-JFH (V+T), pH-EI-JFH (V+T) were linearized with XbaI. RNA was synthesized by *in vitro* transcription as above using 2 ug of template DNA in a 50uL reaction or using 5 ug of template DNA in a 20 uL using the T7 MegaScript kit (Ambion). After incubation for 2H at 37°C, 2uL of TurboDNase (Ambion) was added to the reactions and incubated for 20 min at 37°C. The RNA was purified using the RNeasy kit (Qiagen) and eluted twice in a total of 100 uL of water. The yield of RNA was determined by absorbance at 260nm and integrity verified by agarose gel electrophoresis.

### ***Electroporation of transcribed HCV RNA.***

Subconfluent Huh-7.5 cells were trypsinized, collected, and washed twice with ice-cold RNase-free phosphate-buffered saline. The cells were resuspended in PBS at  $1.5 \times 10^7$  cells/ml and placed on ice. One to two micrograms of *in vitro* transcribed RNA was mixed with 0.4 ml of the Huh-7.5 cell suspension, transferred to a 2 mm gap cuvette (BTX, Holliston MA), and pulsed with a BTX ElectroSquarePorator (820 V; five pulses; 99 usec pulse length, 1.1 sec intervals). After 10 min at room temperature, the cells

were transferred to 9.6 ml of DMEM-10% FBS. For Flow cytometry,  $1 \times 10^6$  cells were plated on 100 mm dish. For real time qRT-PCR,  $7.5 \times 10^5$  cells were plated per well of a 6 well plate.

For luciferase assays, a master dilution plate was generated using an 8 x 12 rack of Titer tubes (BioRad, Hercules, CA). Two dilutions were made:  $3 \times 10^5$  cells/ mL and  $1 \times 10^5$  cells/ mL. Transfers of 100 uL from the master plate to a 96 well flat bottom tissue culture plate were performed creating four replicate wells per RNA. Three 96 well plates were seeded at  $3 \times 10^4$  cells/ well to harvest at the early time points and three plates were seeded at  $1 \times 10^4$  cells/ well to harvest at later time points.

#### ***Quantitation of HCV RNA by Real Time qRT-PCR.***

A 100 ng aliquot of total RNA was used to quantify HCV-specific RNA levels using an ABI Prism 7700 sequence detector (Applied Biosystems, Foster City, CA). qRT-PCR amplifications were performed using the Platinum Quantitative RT-PCR ThermoScript One-Step system (Invitrogen) and primers specific for the HCV 5' NTR: 5'-CCT CTA GAG CCA TAG TGG TCT-3' (sense, 10 uM); 5'-CCA AAT CTC CAG GCA TTG AGC-3' (antisense, 10 uM); and 6-carboxyfluorescein-CAC CGG AAT TGC CAG GAC GAC CGG (probe, 10 uM; Applied Biosystems). Reverse

transcription reactions were incubated for 30 min at 50°C, followed by inactivation of the reverse transcriptase coupled with activation of Taq polymerase at 95°C for 5 min. After cooling to 25°C for 2 min, 40 cycles of PCR were performed with cycling conditions of 15 sec at 95°C, 40 sec at 50°C, and 30 sec at 72°C. The reaction mix contained the glyceraldehyde-3-phosphate dehydrogenase detection mix from Applied Biosystems (VIC-MGBNFQ) for normalization. Synthetic HCV RNA standards of known concentration were included with each set of reactions and used to calculate a standard curve. The real-time PCR signals were analyzed in a multiplex format using SDS software (version 1.6.3; Applied Biosystems).

***RNA secondary structure probing.***

One microgram of transcript RNA was treated with cobra venom nuclease V1 (Ambion) or RNase T1 (U.S. Biochemicals, Cleveland, OH) at 0°C for 30 min. The reaction was carried out in a 10 ul reaction volume with 10 mM Tris-HCl (pH 7.0), 0.1 M KCl, 10 mM MgCl<sub>2</sub>, 3 ug of yeast tRNA, and RNases. One or 0.1 U of RNase T1 and 0.1 or 0.01 U of RNase V1 was used. The reaction was extracted with acidic phenol-chloroform followed by ethanol precipitation.



***Primer extension analysis.***

Oligonucleotide used in the primer extension assay: (5'-CCA GGT CCT GCC CTC GGG CCG-3'). The oligonucleotide was gel purified and 5'-end labeled with 10 U of T4 polynucleotide kinase (New England Biolabs, Ipswich, MA) and 100 uCi of [ $^{32}\text{P}$ ]ATP (8,000 Ci/mmol; GE Healthcare) in a reaction mixture containing 70 mM Tris-HCl (pH 7.6), 10 mM  $\text{MgCl}_2$ , 100 mM KCl, and 1 mM  $\beta$ -mercaptoethanol at 37°C for 30 min. The unincorporated radioisotope was removed by passing the reaction mixture twice through QIAquick silica gel (Qiagen). Cleaved RNAs were mixed with the 5'-end-labeled oligonucleotide ( $5 \times 10^4$  cpm as determined by liquid scintillation counting) in a 20  $\mu\text{l}$  reaction volume containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM  $\text{MgCl}_2$ , and 10 mM DTT, prewarmed to 40°C for 2 min, and incubated for 50 min following addition of 100 U of Superscript II RNase H reverse transcriptase (Invitrogen). The cleavage sites on the RNA were determined by analyzing the reverse-transcribed cDNA products separated by denaturing polyacrylamide gel (12% polyacrylamide-7 M urea) electrophoresis. Dideoxy sequencing reactions were carried out on unmodified RNA and run in parallel.

***TransMessenger delivery of RNA to Huh-7.5 cells.***

Huh-7.5 cells were seeded at  $8 \times 10^4$  cells/ well in a 24 well plate the day before. Next, 0.8ug of *in vitro* transcribed RNA was added to the diluted enhancer reagent and incubated for 5 min at room temperature. TransMessenger reagent (Qiagen) was added to the RNA and incubated for 10 min at room temperature. OptiMEM media was mixed with the transfection mix. Cell monolayers were washed with PBS and the TransMessenger-RNA complexes added drop wise to the cells. Cells were either harvested or after 3H, DMEM with 10% FBS added to the wells.

***Immunohistochemistry.***

HCV-positive Huh-7.5 cells were visualized by immunohistochemical staining for NS5A. All washes and diluents consisted of PBS with 0.1% Tween20. Endogenous peroxide was quenched with 3%  $\text{H}_2\text{O}_2$ . Cells were incubated for 1H at 37°C with monoclonal antibody 9E10 diluted 1:200. After washing, bound antibody was detected with ImmPRESS peroxidase-conjugated anti-mouse (Vector laboratories, Burlingame, CA) diluted 1:4 and incubated for 30 min at 37°C. Peroxidase was detected with DAB substrate (Vector labs), and nuclei counterstained with hematoxylin 2.

### ***Luciferase assay.***

At the indicated time point post electroporation, media was removed and cell monolayers were washed twice with PBS. For cells in 96 well plates, 20 uL of lysis reaction was added to each well and the plate was stored at -80°C. To analyze luciferase activity, plate contents were thawed and 15 uL was transferred to a white polystyrene 96 well plate. For a dual luciferase reaction, 25 uL of LARII (Promega, Madison, WI) was added to a well, a 2 sec delay, and firefly luciferase activity read for 2-5 sec. Then 25 uL of Stop and Glo (Promega) was added to the well, a 2 sec delay, and Renilla luciferase activity measured for 5-10 sec. For a Renilla luciferase assay, 25 uL of Renilla substrate was added to each well, a 2 sec delay, 5-10 sec read, and then a 10 uL of 70% ethanol was added to quench the signal.

### ***Metabolic labeling of proteins and immunoprecipitation.***

Cell monolayers in 35 mm wells were incubated for 0.5 to 10H in methionine- and cysteine-deficient minimal essential medium containing 1/40 the normal concentration of methionine, 10% FBS, and Express <sup>35</sup>S-protein labeling mix (140 uCi/ml; NEN). Labeled cells were washed once with cold PBS and harvested in 200 ul of SDS lysis buffer (0.1 M sodium phosphate buffer [pH 7.0], 1% SDS, 1x complete protease inhibitor cocktail

[Roche Applied Science], 80 ug of phenylmethylsulfonyl fluoride [PMSF] per ml), and cellular DNA was sheared by repeated passage through a 27 gauge needle. Equal amounts of protein lysates (50 ul) were heated at 75°C for 10 min and clarified by centrifugation prior to mixing with 200 ul of TNA (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.67% bovine serum albumin, 1 mM EDTA, 0.33% Triton X-100, 80 ug of PMSF per ml). One ul of HCV-positive serum (patient JHF) was added, and immune complexes allowed to form by incubation overnight at 4°C with rocking. Immune complexes were collected by adding 50 ul of prewashed Pansorbin cells (Calbiochem, San Diego, CA) and incubating for 1 to 2H at 4°C with rocking. Immunoprecipitates were collected by centrifugation and washed three times in TNAS (TNA containing 0.125% SDS) and once with TNE (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM EDTA, 80 ug of PMSF per ml), solubilized by heating at 80°C for 20 min in protein sample buffer, and separated on an SDS-10% polyacrylamide gel. Metabolically labeled proteins were visualized by autoradiography.

### ***Selection of reversions:***

Replication was launched following the electroporation protocol as described above with 10 ug of *in vitro* transcribed RNA for  $6 \times 10^6$  cells. Cells were seeded at  $1 \times 10^6$  cells in a 100 mm dish for passage, and  $1 \times 10^4$  cells in a 24 well plate for IHC. Upon reaching about 80% confluency, cells were split 1:5 into a 100 mm dish and one well of a 24 well plate. To test for virus infectivity, media was removed from the transfected cell culture, centrifuged to remove any cell debris, and added to naïve cells in a 24 well plate. HCV positive cells were visualized by IHC and HCV RNA was collected using RNeasy Mini columns (Qiagen) for cloning and sequencing.

### ***Huh-7 Translation Cell Extracts:***

Monolayers of Huh-7 cells were grown in ten P150 flasks to a density of  $1 \times 10^7$  cells per plate. Media was removed and washed twice in ice cold PBS. Cells were dislodged from the dish in ice cold PBS using a flat edge cell scraper and centrifuged at  $2000 \times g$  for 20 min at  $4^{\circ}\text{C}$ . The supernatant was discarded and the cells were gently resuspended in 750 uL (1.5 the volume of cells) of low salt solution (10 mM HEPES [pH 7.4], 10 mM KoAc, 2 mM DTT). Cells were allowed to swell on ice for 10 min. Using a 2 mL Dounce B homogenizer, cells were lysed with 30 strokes on ice and

transferred to two 1.5 mL microfuge tubes. To each tube, 125 uL of 10 x buffer (0.2 M HEPES [pH 7.4], 1.2 M KOAc, 50 mM DTT, 25 mM MgOAc) was added and mixed by pipetting. The lysates were centrifuged at 3000 x g for 5 min at 4°C. The supernatant was transferred to a new microfuge tube and centrifuged at maximum speed for 20 min at 4°C. Avoiding a film of lipids at the surface, supernatant was transferred to a new microfuge tube. Micrococcal nuclease (0.067 U/ul) in 50% glycerol was added, followed by 1 mM CaCl to activate the nuclease. The reaction was incubated at room temperature for 1H. 2 mM EGTA was added to stop the reaction and aliquots were flash frozen in liquid nitrogen before being transferred to -80°C.

***Translation using Huh-7 and Huh-7.5 cell extracts.***

On ice, the reaction was assembled using 12.5 uL of cell extracts, 2.5 uL of 10x Buffer (10 mM ATP, 2.5 mM GTP, 100 mM HEPES, 0.6 M KCl, 12.5 mM MgOAc, 0.2 M creatine phosphate, 10 nM spermadine), 7.5 ug of creatine kinase, 0.04 mM amino acids. Template RNA was added and the reaction was incubated at 30°C.

***Translation using rabbit reticulocyte extracts.***

The Flexi rabbit reticulocyte lysates were obtained from Promega. In a 25 uL reaction volume, 16.5 uL of lysates, 0.04 mM amino acids, 0.2 U superasin, 40 mM KCl, and MgCl<sub>2</sub> added to a final concentration of 3 mM (depending on the endogenous amount in the lysates). Template RNA was added and the reaction was incubated at 30°C.

***Transfection of Huh-7.5 cells and selection with G418.***

Subconfluent Huh-7.5 cells were trypsinized, collected, and washed twice with ice-cold RNase-free phosphate-buffered saline. The cells were resuspended in phosphate-buffered saline at  $2.5 \times 10^7$  cells/ ml. One microgram of *in vitro* transcribed RNA was mixed with 0.4 ml of the Huh-7.5 cell suspension, transferred to a 2 mm gap cuvette (BTX), and pulsed with a BTX ElectroSquarePorator (0.82 kV; five pulses; 99 usec pulse length, 1.1 sec intervals). After 10 min at room temperature, the cells were transferred to 9.6 ml of DMEM with 10% FBS. To measure the efficiency of G418-resistant colony formation, the transfected cells were plated at a series of densities ( $10^6$ ,  $10^5$ , and  $10^4$  cells per plate). To maintain the total number of plated cells at  $10^6$  cells per 100mm dish, cells transfected with

polymerase-defective RNA were used as feeder cells. After 24H, the medium was changed to DMEM with 10% FBS supplemented with G418 (Geneticin; Invitrogen) at 1 mg/ml. The medium was replaced every 4 days. After about 3 weeks, G418-resistant colonies were stained with 1% crystal violet in 50% ethanol.



<b>Table 2.1: List of mutations</b>				
<b>Name</b>	<b>Position</b>	<b>Genotype</b>	<b>AA change</b>	<b>AA ARF change</b>
Stop 1	C407A	1a	None	Stop
Stop 2	T434A	1a	None	Stop
Stop 3	G473A	1a	None	Stop
Stop 4	C480A, C482G	1a	None	Stop
Open 1	C407T	1a	None	L > S
Open 2	T434C	1a	None	S > L
Open 3	T471C, G473C	1a	None	S > W
Open 4	G479C, C480A, C482G	1a	None	S > C G > A
Comp 3	C466T, C407A, T434A, G473A, C480A, C482G	1a	P42L	Stop
Comp 4	G455C, G457T, C407A, T434A, G473A, C480A, C482G	1a	R39L	Stop
Comp 2	A502T, C407A, T434A, G473A, C480A, C482G	1a	E54V	Stop
Stop 1	C407A	1b	None	Stop
Stop 2	C434A	1b	None	Stop
Stop 3	C473T	1b	None	Stop
Stop 4	C480A, C482G	1b	None	Stop
Mut Loop SLV	C407A, A408C, A409T	1a	K23L	NA
Mut Loop SLVI	A468C, A470C, T471C	1a	None	NA
Mid Right	C482A, C484G	1a	A48G	NA
Mid Right Silent	C482A, G485T	1a	None	NA
Mid Left	G452C, G455A	1a	L37F	NA
Mid Right/ Left	C482A, C484G G452C G455A	1a	A48G L37F	NA

**Table 2.1: List of mutations, CONT.**

<b>Name</b>	<b>Position</b>	<b>Genotype</b>	<b>AA change</b>	<b>AA ARF change</b>
Base Left	T434A, T437A, A441T	1a	None	NA
Base Right	T497A, C500A, A502T	1a	E54V	NA
Base Left/ Right	T434A, T437A, A441T, T497A, C500A, A502T	1a	E54V	NA
J6 Mid Right-A	G472A, C481A, C483G	2a	A48G	NA
J6 Mid Left-A	T449C, G451C, G472A	2a	None	NA
J6 Mid Left-B	G451A, G454C, G472A	2a	None	NA
J6 > JFH	G482A	2a	A48T	NA
J6 Mid Right-B	G472A, C481G, G484T	2a	None	NA
J6 Stop 3,4	G472A, C479A, C481G	2a	None	NA
5'NTR mut-A	U26A, A29U, A32U	1a	None	NA
5'NTR mut-B	U33A, G34C, A35G	1a	None	NA
5'NTR mut-C	U26A, A29U	1a	None	NA
5'NTR mut-D	U26A	1a	None	NA
5'NTR mut-E	A32U	1a	None	NA
5'NTR mut-F	A29U, A32U, C38U	1a	None	NA

## **Chapter 3**

### **Chapter 3. The Alternative Reading Frame is not essential for infection.**

#### ***Aim.***

An overlapping open reading frame is present in the capsid gene of all HCV genotypes. Using cell free translation assays as well as transient expression in cell culture, a protein product has been shown to be expressed from this alternative reading frame (ARF)<sup>18, 149, 153, 161</sup>. HCV patients have also been shown to produce antibodies and cellular immune responses to ARF peptides indicating its expression during infection<sup>66, 153, 161</sup>. To test if this ARF is required in an HCV infection, we introduced silent mutations that confer stop codons in the ARF, while preserving the capsid amino acid sequence, in an HCV molecular clone. This ARF-ablated genome was inoculated into a chimpanzee. We monitored the virus load, adaptive and humoral immune response, and virus diversity to determine if the ARF contributes to the HCV lifecycle in a natural infection.

#### ***Introduction.***

Comparison of the capsid gene from divergent HCV isolates has revealed an unusually high level of nucleotide sequence conservation. Synonymous mutations are suppressed, suggesting a functional role for the nucleic acid sequence beyond simply encoding the capsid protein. For all 6

HCV genotypes, an overlapping reading frame in the +1 frame of the capsid gene is present. In the same region, highly conserved RNA secondary structures have been predicted using phylogenetic analyses of covariant base pairs.

The +1 frame of the capsid gene has been termed the alternative reading frame, ARF. Antibodies and cellular immune responses reacting to ARF-encoded peptides or recombinant protein have been detected in HCV-infected patients. These observations provide evidence for expression of the ARF *in vivo*; however, an ARF-encoded product has yet to be detected in infected tissue. Using cell-free translation and transient expression in cell culture, an ARF protein has, however, been reported for the prototype genotype 1a isolate, HCV-1. HCV-1 contains a stretch of 10 consecutive adenine residues where a +1 frameshift event moves translation into the ARF between codons 9-11 of the capsid gene to produce a frameshift product, termed the F protein<sup>160</sup>. Another ARF protein has been reported for a genotype 1b sequence involving two frameshift events. This dual frameshift (DF) protein also begins translation from the AUG of the polyprotein and frameshifts further downstream after codon 42 into the ARF<sup>18</sup>. Translation continues in the ARF until a stop codon at position 144 is

reached, triggering either termination or an additional frameshift that restores translation into the polyprotein reading frame.

***Ablation of ARF-encoded gene products by site directed mutagenesis.***

Despite its high degree of conservation, the region of the capsid gene encoding potential ARF proteins is dispensable for RNA replication in cell culture. Since ARF-encoded proteins may function in other aspects of the HCV lifecycle such as virus assembly, virus-immune system interactions or pathogenesis, we used the chimpanzee model to examine the behavior of a mutant HCV genome with ablated expression of both the ARF-encoded F and DF proteins. The consensus genotype 1a H77 cDNA was chosen as the parent genome for our studies given the wealth of information obtained from chimpanzees infected with H77 RNA transcripts or with the acute-phase virus<sup>82, 83, 84</sup>. Five nucleotide substitutions were introduced to the capsid gene that conferred four stop codons (Stop1,2,3,4) while preserving the amino acid sequence of the capsid protein (Figure 3.1). These mutations would also disrupt RNA secondary structures in the capsid gene (Figure 3.2). To minimize leaky protein expression due to read-through and reversion, two stop codons were introduced downstream from each reported frameshift site. The F protein is reported to use a frameshift signal at codons 9-11. The DF

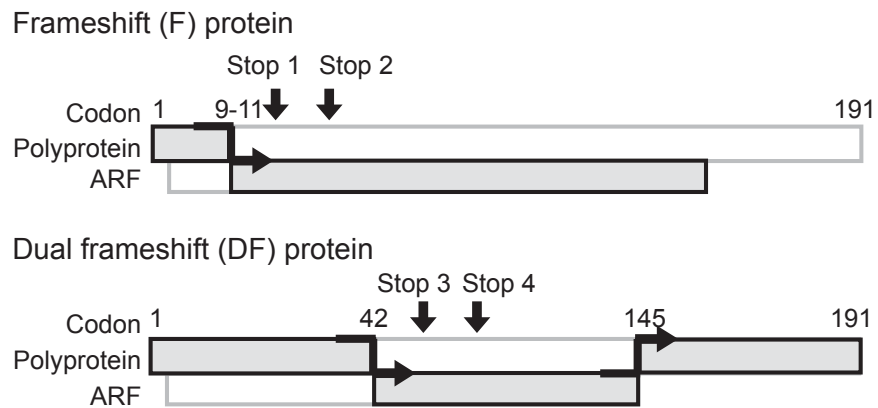


Figure 3.1. The proposed expression of the F and DF proteins are shown. Silent point mutations were introduced to convey stop codons in the ARF. Mutations for Stop 1 and Stop 2 would block expression of the F protein. Stop 3 and Stop 4 would block expression of the F and DF proteins.

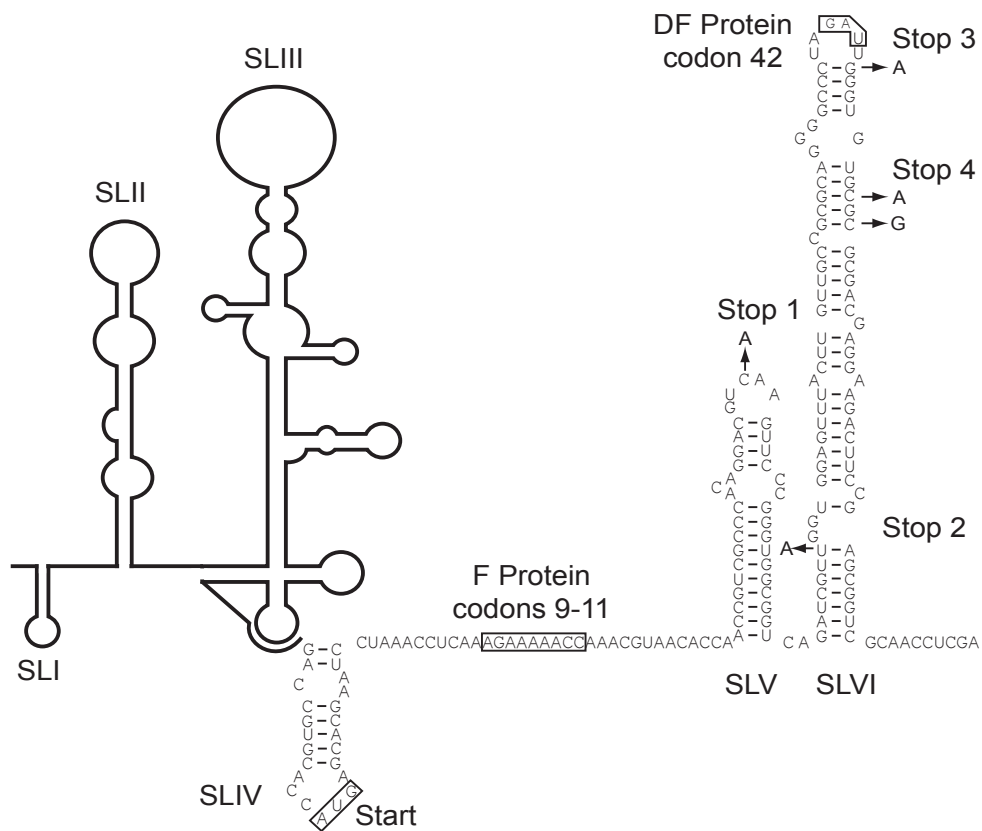


Figure 3.2. The predicted RNA secondary structures SLV and SLVI in the capsid gene. The point mutations conferring Stop 1,2,3,4 in the ARF are shown. Stop 2,3,and 4 would disrupt base paired interactions in SLVI.



protein reportedly shifts into the ARF further downstream at codon 42 (Figure 3.1). Stop 1 would permit expression of only the first 21 amino acids of the F protein, and Stop 2 follows 9 codons downstream. Stops 3 and 4 are located at codon 43 and 45 of the ARF to block expression of the DF protein. Stops 1, 2, and 3 each required one base substitution (C407A, U434A, G473A, respectively), whereas two base changes were required to generate Stop 4 (C480A and C482G). Only Stop 4 created a potentially leaky opal stop codon (UGA), but rather than C residue which promotes translational read-through, this codon was followed by a G<sup>70</sup>.

***An HCV genome without F and DF expression is infectious in vivo.***

An HCV-naïve chimpanzee was inoculated with the HCV genome Stop 1,2,3,4 by direct intrahepatic injection of in vitro transcribed RNA. The animal became infected. Circulating HCV RNA was detected by nested RT-PCR at 1 week post-inoculation. By week 2, HCV RNA levels could be quantified by real time qRT-PCR. HCV RNA levels rapidly increased and peaked at weeks 6 and 7 at  $8 \times 10^4$  GE/ml (Figure 3.3). Circulating HCV RNA then declined rapidly and fell below the limit of detection at week 10. RNA was transiently measurable at week 12 (380 RNA GE/ml), and only detectable by nested RT-PCR for the subsequent weeks analyzed (through

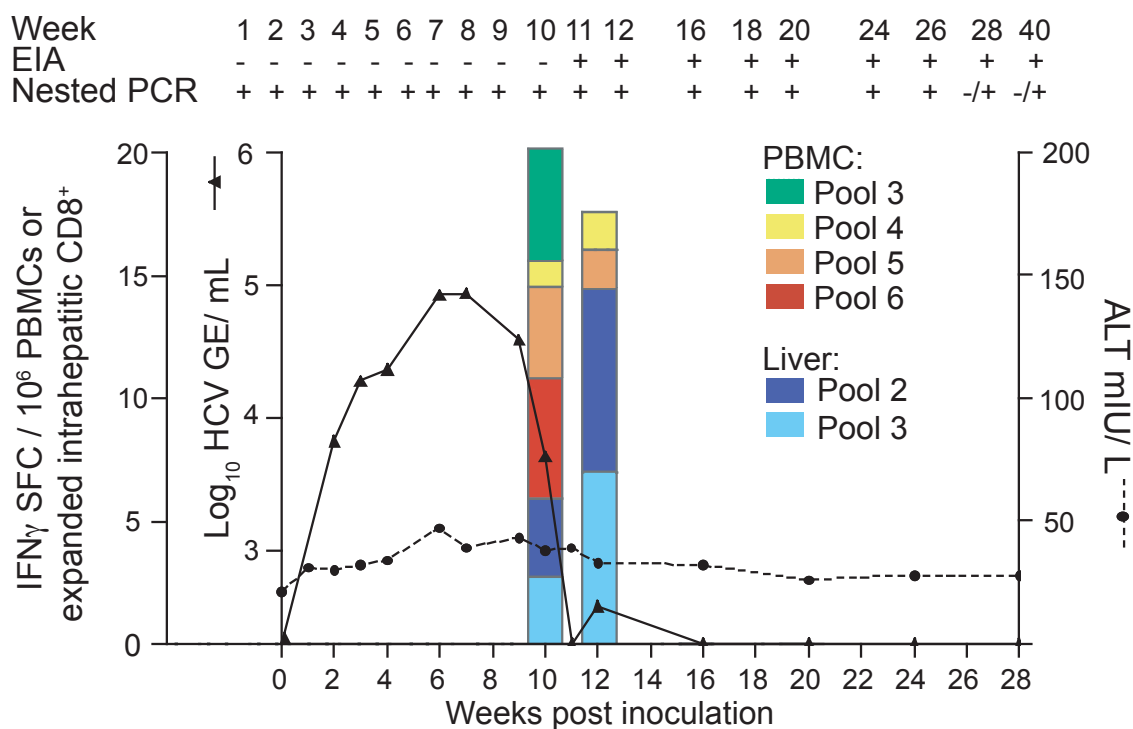


Figure 3.3. The infection profile of the HCV genome Stop 1,2,3,4 in a chimpanzee. The virus RNA levels are measured by Real Time qT-PCR and graphed on the inner left y-axis with the limit of detection as the origin of the x-axis. ALT levels are on the right y-axis. An HCV-specific immune response was measured by PBMCs and liver derived CD8<sup>+</sup> in an ELISPOT assay. The frequency of secreted interferon gamma cells is shown on the outside left y-axis. An antibody response was detected by EIA and positive weeks indicated.

week 40). The animal seroconverted at week 11. No significant elevation in ALT was observed except for a slight rise of 20 U/ml above baseline during the peak of viremia (Figure 3.3). Compared to other monoclonal chimpanzee infections with the parental H77 sequence, Stop 1,2,3,4 infection produced a lower peak viremia by 1-2 logs. The mutant virus was also controlled earlier without apparent liver damage, suggesting that the ARF mutant may be attenuated.

***An immune response is detected in both the periphery and liver.***

Given that ARF-encoded products might function in modulating host immune responses to HCV, HCV-specific T cell responses were monitored both in the periphery and the liver. At weeks 10 and 12, HCV-specific responses were detected in peripheral blood mononuclear cells (PBMCs) and from intrahepatic CD8<sup>+</sup> T cells by ELISPOT (Figure 3.3). PBMCs were stimulated from pools 3, 4, 5 and 6, representing peptides from E2 to NS4B, and HCV-specific CD8<sup>+</sup> cell responses reacted to pools 2 and 3, which include peptides from E1, E2, p7, and NS2. The frequency of HCV-specific CD8<sup>+</sup> cells was lower than observed for other chimpanzee infections launched with a genotype 1a virus, but we were able to detect a diverse response indicating active replication and interchange with a functional

immune system. We monitored the general population of CD4<sup>+</sup>, CD8<sup>+</sup>, CD3<sup>+</sup>, and CD56<sup>+</sup> expressing cells in the PBMCs and found no change, thus indicating no evidence of immunosuppression during the infection.

***Unique epitopes are detected by direct cloning of T cells from liver biopsy.***

Since a response could not be detected after week 12 using the ELISPOT assay, we used a more sensitive approach and cloned intrahepatic CD8<sup>+</sup> and CD4<sup>+</sup> T cells from liver biopsies at week 16 and week 20 post inoculation. Analysis of 308 independently derived CD8<sup>+</sup> T cell clones revealed that 31% were capable of killing target cells presenting HCV proteins. We then determined the epitope of 96 CD8<sup>+</sup> and 7 CD4<sup>+</sup> intrahepatic T cell clones, which revealed previously unseen CD8<sup>+</sup> epitopes spanning proteins E1 to NS5. CD4<sup>+</sup> cells proliferated in response to epitopes in NS4B at week 16 and epitopes in E2, NS3, NS5A by week 20 (Table 3.1). No T cell reactivity was detected against ARF-encoded epitopes. Thus, despite producing a low level of viremia, infection with Stop 1,2,3,4 stimulated a diverse intrahepatic T cell response, first detectable by week 10, and presumably controlling the virus by week 20.

### A. CD8+ T cell lines:

Week 16		<u>No. cell lines</u>
	Pool 2	48
	Pool 3	39
	Pool 6	1
	Pool 7	4
	Pool 8	4
	Tested	308

### B. CD4+ T cell clones:

<u>Clone</u>		<u>Epitope</u>	<u>Protein</u>	<u>AA sequence</u>
Week 16				
5F	Pool 6	LFNILGGWVAAQLAAPGA	NS4B	1808-1825
9F	Pool 6	LFNILGGWVAAQLAAPGA	NS4B	1808-1825
Week 20				
1C	Pool 2	YANGSGLDERPYCWHY	E2	474-489
8C	Pool 2	MVDYPYRLWHYPCTINY	E2	608-624
7E	Pool 5	KLVALGINAVAYYRGLDV	NS3	1406-1423
3F	Pool 5	YRLGAVQNEVTLTHPITK	NS3	1626-1643
7F	Pool 7	EVSVP AEILRKSRRFARA	NS5A	2267-2284

Table 3.1. CD8<sup>+</sup> and CD4<sup>+</sup> cells were extracted from liver biopsy at weeks 16 and 20 and cloned in a limiting dilution assay. A. CD8<sup>+</sup> T cell lines shown to generate a cytotoxic response upon stimulation with HCV peptides from the indicated pools. B. The epitopes of CD4<sup>+</sup> T cell clones were mapped in a proliferation assay upon stimulation with the indicated HCV peptides.

***The selection of revertants suggests pressure to maintain ARF RNA elements rather than F/DF protein expression.***

The low peak viremia and early control of the mutant virus could be due to the Stop 1,2,3,4 substitutions impeding HCV replication. Given the high mutation rate of HCV replication, reversion to the wild-type H77 sequence and/or accumulation of compensating, fitness-restoring mutations would provide clues to the importance of F/DF expression or functional RNA elements in the ARF. At weekly or monthly (after week 12) time points, the 5'-NTR-C/ARF region of circulating HCV RNA was amplified by RT-PCR. Sequences were determined for both the population and multiple clones, to sample sequence heterogeneity.

At week 1, the predominant HCV sequence was identical to the inoculated Stop 1,2,3,4 RNA transcript. Since we have never detected circulating input RNA using mutant transcripts with lethal (pol<sup>-</sup>) mutations<sup>65</sup>, this suggests that the Stop 1,2,3,4 mutant was able to replicate *in vivo*, albeit inefficiently. At week 2, coincident with a rise in circulating virus RNA to nearly 10<sup>4</sup> GE/ml, the parental H77 G473 was found in all HCV genomes sequenced (Figure 3.4). The reversion of Stop 3 (G473) remained fixed in the population for at least 40 weeks, the last time point analyzed.

	Stop 1	Stop 2	Stop 3	Stop 4
Stop 1,2,3,4	GGACGU <u>AA</u> AGU . . . UCGU <u>A</u> GGUG . . . GAUU <u>A</u> GGUGUG <u>A</u> GCGCG			
Week 1	.....	.....	.....	.....
Week 2	.....	.....	G.....	.....
Week 10	{.....	.....	G.....	.....
Week 20	.....	.....	G.....	C.....
Week 40	.....	.....	G.....	C.C...
H77	GGACGU <u>C</u> AAGU . . . UCGU <u>U</u> GGUG . . . GAUU <u>G</u> GGUGUG <u>C</u> GCGCG			
Position	407	434	473	480 482

Figure 3.4. Reversions detected in the circulation after inoculation with Stop 1,2,3,4 genome. At week 2, a reversion at 473 replaced the Stop 3 mutation A with the H77 G. This reversion was the dominant sequence and became fixed in the population. At week 10, a mixed population was detected containing a second reversion of A480C. This second reversion became fixed at week 20. At week 40, a third reversion was detected, G482C.

This reversion removed the Stop 3 ochre codon, but the downstream Stop 4 opal codon was still present to block expression of DF.

At week 10, we detected a second reversion at Stop 4 restoring A480 to the wild-type C480 H77 sequence. This was not the dominant sequence, as virus with both reversions circulated with virus containing only the first reversion at Stop 3. The A480C reversion eliminated the Stop 4 codon, thus restoring possible expression of the DF protein. The capsid protein sequence remained unaltered as Stop 4 preserved capsid Arg47 by creating an alternative arginine codon (AGG, nucleotides 480-482) and the second reversion, A480C, yielded the capsid codon CGG. Virus with one or both reversions existed as a mixed population for the next 10 weeks. At week 20, virus with both first and second reversions was fixed in the population (Figure 3.4).

A third reversion was detected at week 40 involving the second nucleotide substitution of Stop 4, with G482 being replaced by the wild-type H77 nucleotide, C (Figure 3.4). This third reversion was the only sequence detected, but may not necessarily represent the virus population since we succeeded in obtaining an RT-PCR product in only one of three attempts. This third reversion did not change the capsid protein sequence since C482 restores the wild-type Arg47 codon (CGC, nt 480-482). Hence, the selective



force(s) favoring C482 are unclear. The possibility of codon preference for Arg CGC over CGG seems unlikely since they are used at approximately equal frequency in the HCV H77 ORF (38 versus 34 times, respectively). Since week 10, the presence of the first and second reversion could restore possible DF expression. The third reversion would not be expected to affect DF production, although it would alter the amino acid at position 5 of the DF frame from Gly (mutant) to Ala (wild-type). Alternatively, the reversion to C482 may restore an RNA element in the C/ARF region important for HCV replication. The results of the sequence analyses, where Stop 1 and Stop 2 were maintained throughout infection, clearly demonstrate that F protein expression is not required for HCV H77 replication.

***Evidence for an RNA structure in the capsid gene.***

Two RNA helix forming stem structures, SLV and SLVI, have been proposed in the region containing the ARF stop codons and confirmed by enzymatic structure probing <sup>147</sup>. The base change conferring Stop 1 is not predicted to alter a base pair interaction; however, substitutions for Stop 2, Stop 3, and Stop 4 could disrupt the upper stem in a large RNA helix of SLVI (Figure 3.5). A phylogenetic study of covariant base pairing shows high sequence conservation for the structure SLVI for all 6 genotypes

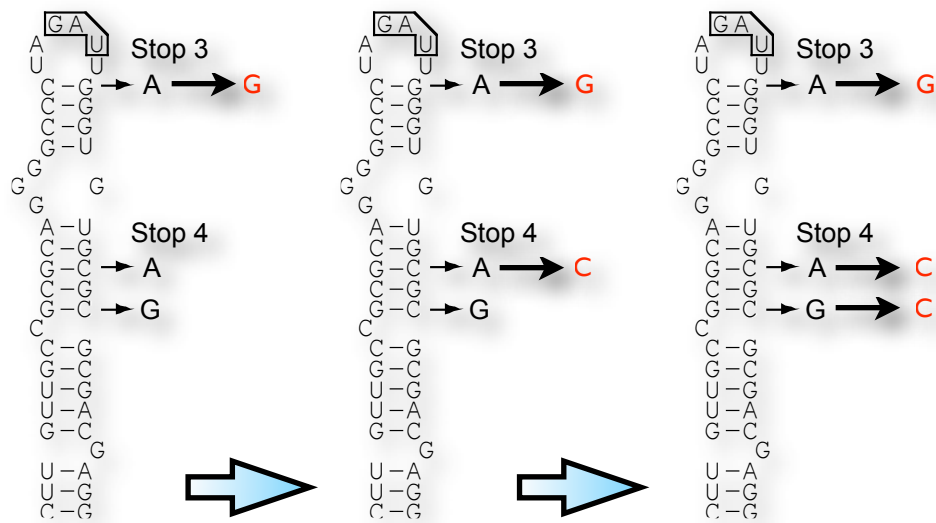


Figure 3.5. Reversions detected in the circulation indicate a selective pressure to improve replication by restoring an RNA structure. The top of SLVI is shown. The first reversion of Stop 3 would re-establish a Watson and Crick base pair below the loop of SLVI. The second and third reversion would restore two base pairs to the top stem of SLVI.

(Figure 3.6). The sequence of the parental H77 was compared with 1274 sequences including all 6 genotypes. The predicted RNA structures were compared by generating M-fold profiles and differences tallied with R-fold analyzer, kindly provided by Andrea Branch. The top of SLVI displays covariant conservation with base pairs likely to exist in 90% or more of the sequences shaded in Figure 3.6. Nucleotides conserved in all but one sequence out of the total 1274 analyzed are circled. The wild-type nucleotide of Stop 3, G473, is conserved in 1272 out of 1274 sequences, and the Stop 4 mutations, 480 and 482, are conserved in 1256 out of 1274 sequences. Covariant base pairs are detected in 18 and 14 sequences for Stop 4 mutant positions 480 and 482, respectively. The first reversion at Stop 3 restores a base pair interaction at the base of the loop of SLVI, and the second and third reversions would repair two base pair disruptions in the stem of SLVI.

## **Discussion.**

We have shown that an HCV genome containing 4 stop codons in the +1 frame overlapping the capsid gene is capable of establishing an infection in a chimpanzee. The infection was, however, atypical with low peak viremia, short duration of the acute phase, and no appreciable liver pathology. A multi-specific adaptive immune response was elicited in both the liver and

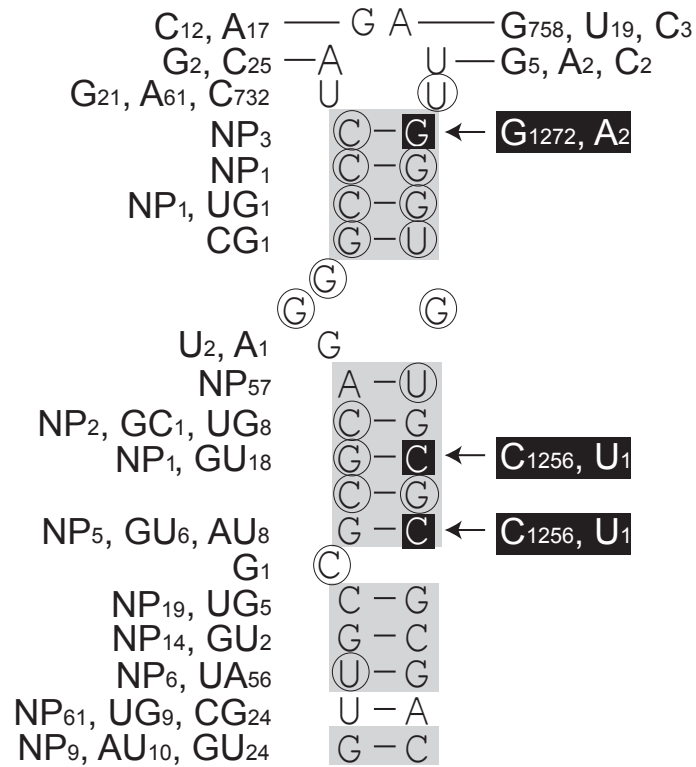


Figure 3.6. Phylogenetic analysis of the conserved nucleotides and covariant base pairs of SLVI. The sequence of the parental H77 was compared with 1274 sequences and include all 6 genotypes. Regions shaded indicate Watson-Crick or wobble base pairs likely to exist in 90% or more of sequences in all 6 genotypes. The nucleotides conserved in all but one sequence out of 1274 are circled. Covariant bases are counted adjacent to the structure, with the number of sequences with bases unable to pair (NP, not paired) indicated. The base mutations in the chimpanzee study are indicated by a square with the number of sequences with identical or different bases at each position indicated

the periphery that coincided with a decline in circulating HCV RNA. During infection, a series of reversions emerged indicating a selective pressure for RNA sequence and predicted RNA structure rather than a frameshift protein. Base substitutions creating the two downstream stop codons were found to revert during the infection. These substitutions disrupt base pairing of a predicted RNA secondary structure in the capsid gene, implicating this RNA element in the modulation of HCV infection.

Features of this infection differed from previous young naïve chimpanzee infections launched with this consensus H77 genome. The titer reached only  $8 \times 10^4$  GE/mL, 1-2 logs lower than previous studies<sup>154</sup>. Also, circulating HCV RNA peaked at week 6 and was undetectable by week 12, several weeks earlier than a typical acute infection. The low level of circulating HCV RNA during acute infection may indicate that few hepatocytes became infected, low replication fitness of the mutant virus on a per cell basis, or both. Virus levels declined without elevated liver transaminases in the serum, signifying little liver damage. Again, this is atypical where an increase in serum ALT (usually to 100-200 IU/mL) coincides with immune-mediated hepatocellular injury and a rapid decrease in circulating HCV RNA. We did, however, find a diverse adaptive immune response in both the liver and periphery at the time of virus decline and persisted as virus

disappeared from circulation. HCV specific CD8<sup>+</sup> and CD4<sup>+</sup> T cells detectable at low frequencies were capable of recognizing HCV peptides in 5 of 10 pools representing the entire H77 polyprotein. No T cells targeting epitopes from the ARF were found. Thus, HCV lacking expression of an ARF protein established an infection, albeit with low levels of circulating virus and of short viremia. Infection with the mutant virus elicited a diverse adaptive immune response, though virus-specific CD4<sup>+</sup> and CD8<sup>+</sup> T cells were detected at a lower frequency.

During infection, we detected a series of sequence reversions in the virus population. The predominant sequence in the first week of infection was the input mutant Stop 1,2,3,4. Since we have never previously detected carry over of RNA from inoculation in the circulation, this suggests that the input mutant was able to replicate. The dominant virus during the acute phase harbored a reversion that eliminated the Stop 3 codon. Stops 1, 2 and 4 remained intact, however, blocking the expression of both the F and DF proteins. A second reversion restored one of two nucleotide substitutions in Stop 4 and was selected at a time of immune stimulation. Finally, a reversion of the second substitution creating Stop 4 was selected as the dominant virus. Since the DF protein expression was possible after the second reversion, and at no time was there a change in the capsid amino acid

sequence, these results are consistent with selection to maintain the capsid nucleotide sequence rather than DF expression. Taken together, these data indicate that neither the F or DF proteins are required for HCV H77 replication *in vivo*.

RNA elements in both the non-coding and coding regions regulate translation and replication, but the functions of SLV and SLVI are unknown. The downstream boundary of the IRES does extend into the capsid-coding region, and IRES function can be influenced by both the capsid protein and RNA structures immediately downstream of the AUG start codon<sup>52, 124, 155</sup>. The 5' base of the SLVI stem in the capsid gene has also been shown to base pair with a complementary region in the 5'NTR between SLI and the IRES, down modulating IRES function<sup>55, 64</sup>. The base substitutions we used to create stop codons in the ARF do not affect the sequence proposed to base pair with the 5'NTR, but would disrupt the stem of SLVI itself. Also, a liver specific microRNA, miR122, can base pair with the same region in the 5'NTR and enhance HCV replication<sup>59</sup>. This interplay between cellular miR122, the 5'NTR, and SLVI may indicate that HCV translation and replication are regulated by competing higher order RNA structures.

Our results clearly show that the F and DF proteins are dispensable for HCV replication *in vivo* and *in vitro*. We cannot exclude that functionally

important ARF products are expressed by internal initiation from regions downstream of the Stop 4 codon (ARF codon 46). In this regard, internal initiation has been proposed to occur between ARF codons 80 and 86, and the importance of such ARF-encoded products remains to be examined <sup>6</sup>.

In summary, a study designed to test the importance of the ARF-encoded F and DF proteins, instead revealed a new, functionally important RNA element in the HCV protein-coding region. This stem-loop structure (SLVI) resides in the ARF and capsid-protein coding region and may be part of an assemblage of higher order RNA structures that regulate HCV translation and replication.



## **Chapter 4**

## **Chapter 4. An RNA structure in the capsid-coding region affects HCV replication in cell culture.**

### ***Aim:***

The course of a chimpanzee infection with an H77 genome containing mutations for Stop 1,2,3,4 differed from that with the wild type, parental genome. Low circulating RNA levels and a series of reversions that were rapidly selected, as early as the second week, imply that the mutant was attenuated. To determine if silent “Stop” mutations in the ARF affected HCV replication in cell culture, we engineered these mutations into an HCV genome that is replication competent in Huh-7.5 cells. The mutations were studied in a functional assay measuring HCV RNA levels and protein expression in cell culture. The regions of the capsid gene contributing to a replication defect were mapped and compared with HCV sequences from diverse genotypes. Using an HCV isolate capable of efficient replication and virus production in cell culture, we evaluated the selective pressure to restore the original RNA sequence.

## ***Introduction.***

The region of the core gene encoding these Stop mutations is dispensable for replication of subgenomic replicons. These replicons undergo HCV replication with a minimum of genetic elements. Both promoter-containing NTRs are present, and the replicase is expressed from the heterologous EMCV IRES. It was thus surprising to detect reversions for this region of capsid sequence during the infection. The first reversion was selected at the launch of the acute stage of the infection and was rapidly fixed in the population. Additional reversions were soon selected and dominated the circulating virus population. The appearance and shift in the population indicate a strong selective pressure, and the order of the reversions implies selection for an RNA element in the capsid gene.

RNA structures are known to conduct translation and replication in positive strand viruses. Virus families such as the *Picornaviridae*, in addition to the *pestivirus* and *hepacivirus*, use IRES-driven translation to express the virus proteins as well as RNA structures at the both ends of the genome to conduct replication via virus polyprotein precursors and cellular proteins<sup>109</sup>. RNA structures at the carboxy terminus of the genome can often influence translation as well as enhance RNA stability. The tRNA-like elements of the family *Bromoviridae* are aminoacylated and require binding

of the capsid protein for replication <sup>16</sup>. *Cis*- acting elements have been identified in plus and minus strands in almost all RNA viruses. RNA signals in the coding regions have been shown to act as frameshifting signals in genera as diverse as the *retrovirs*, *coronaviruses*, and *astrovirus* <sup>22, 85</sup>.

Minimal RNA structures as well as additional enhancer elements have been described as acting as packaging signals for the virus nucleocapsid in assembly. A role for the RNA structures SLV and SLVI of the capsid gene in HCV are unknown, therefore we measured the ability of HCV genomes with mutations in SLV and SLVI to replicate in cell culture.

***The Stop mutations confer a replication defect in cell culture.***

The point mutations conferring Stop 1,2,3,4 in the ARF were introduced into an HCV genome capable of replicating in cell culture. The parent genome is identical to the genotype 1a H77 sequence used in the chimpanzee infection, with the exception of two amino acid differences in the replicase that confer replication in cell culture; one adaptive mutation in the helicase domain of NS3 (P1496L) and another in NS5A (S2204I) <sup>13</sup>. All experiments were performed using an adapted cell line capable of increased HCV replication, Huh-7.5, which was derived from the human hepatoma cell line Huh-7 <sup>12</sup>. Replication was measured by electroporating *in vitro*

transcribed HCV RNA into Huh-7.5 cells and HCV RNA levels determined at successive time points by quantitative real time RT-PCR. RNA levels for the mutant constructs were compared to the wild-type (H/L+I) parent or a replication defective (pol<sup>-</sup>) control. For the H/L+I parent, RNA levels declined slightly as input RNA was degraded, then rose to plateau by day 4 as new RNA was generated (Figure 4.1). By day 3, pol<sup>-</sup> RNA remained detectable, but at levels more than 1000-fold lower than the H/L+I parent. The Stop 1,2,3,4 mutant was highly impaired with transient, low-level replication observed at 3 days, which then decreased to pol<sup>-</sup> levels by day 5. The defect in replication caused by five silent point mutations in the capsid gene was striking, as this region is dispensable for replication in subgenomic replicons, and therefore, thought not to participate in replication. To determine which mutations contributed to the defect, we made a further series of constructs and evaluated their capacity for replication. Mutations for Stop 1 and Stop 2 had no deleterious effect on replication; however, mutations for Stop 3 and 4 did impact replication levels (Figure 4.2). Individually, Stop 3 and Stop 4 reduced replication to 47% and 38% of wild type levels, respectively, and together they crippled replication to just 2% of wild type (Figure 4.2). Combining the mutations of Stop 3 with Stops 1 and 2 also attenuated replication by 76%. Stop 4 yielded an even more dramatic

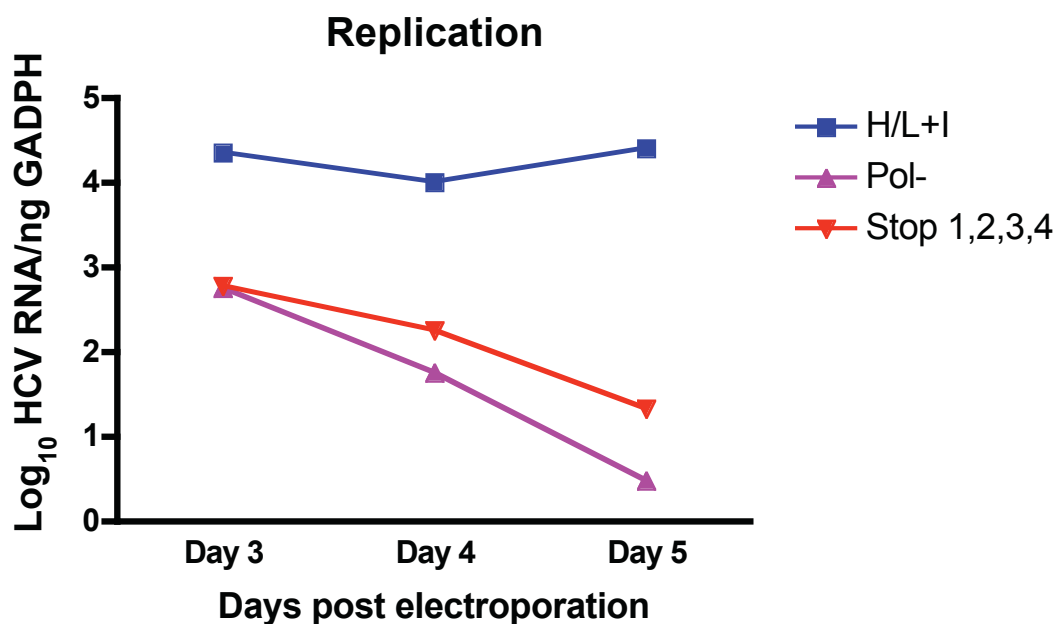


Figure 4.1. Replication of the Stop 1,2,3,4 genome is severely impaired. HCV RNA was measured by Real Time qRT-PCR and normalized with cellular GADPH RNA. Replication is indicated where HCV RNA levels are present above a polymerase defective control.

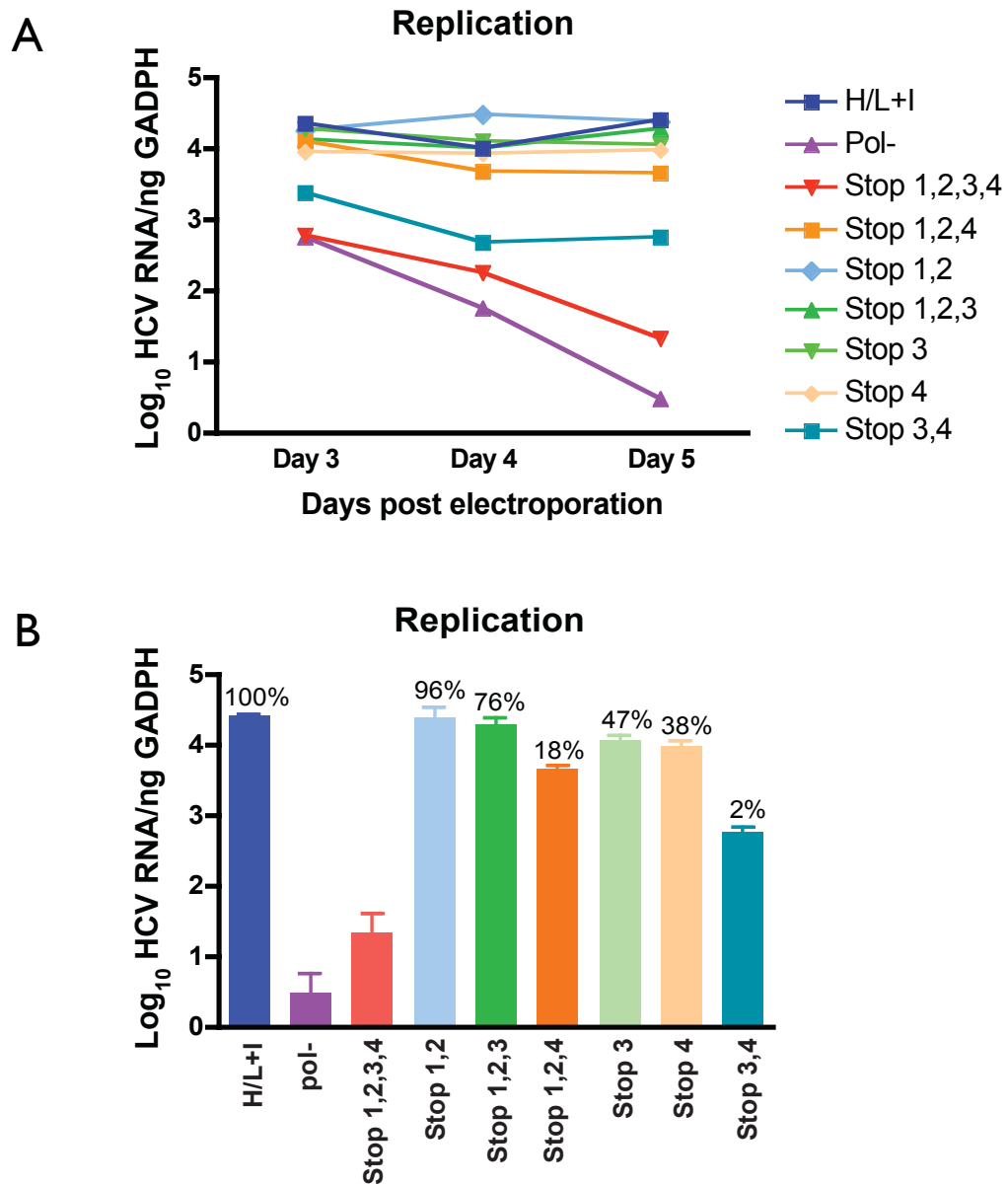


Figure 4.2. Replication of HCV genomes with point mutations for Stop codons 1-4. A. HCV RNA measured by Real Time qRT-PCR at days 3, 4, 5 post electroporation. B. HCV RNA levels at day 5 post electroporation. The percent difference from wild type is shown above each bar. Data represent averages of two independent experiments with two RNA preparations for each genome.

affect in combination with Stops 1 and 2 with 18% replication compared to wild type (Figure 4.2). These results imply that the Stop mutations compromised replication in the chimpanzee infection. Consistent with the results from the chimpanzee infection, the genome with all mutations was severely compromised, almost to the level of the pol- control, yet the genome mimicking the first reversion, Stop 1,2,4, replicated more efficiently although to only 18% of wild type. Furthermore, the second and third reversions in the chimpanzee removed both Stop 3 and Stop 4 mutations, and we found that this genome could replicate to wild-type levels.

The replication defect was confirmed by detecting HCV antigen in transfected Huh-7.5 cells by flow cytometry. This allowed a measurement of HCV protein expression on a per cell basis, while the quantitative real time RT-PCR results were a study of RNA levels within the transfected cell population. At day 5 post electroporation, 19% of the cells transfected with the wild type H/L+I were positive for NS5A, whereas only 2% of cells were positive when transfected with the pol- control (Figure 4.3). Genomes with the mutations for Stop 3 (Stop 1,2,3 and Stop 3 alone) conveyed a 15% and 14% antigen positive cell population, respectively. Fewer cells were positive when transfected with genomes with Stop 4 mutations. A four-fold reduction, 5% of the population, in the number of positive cells resulted



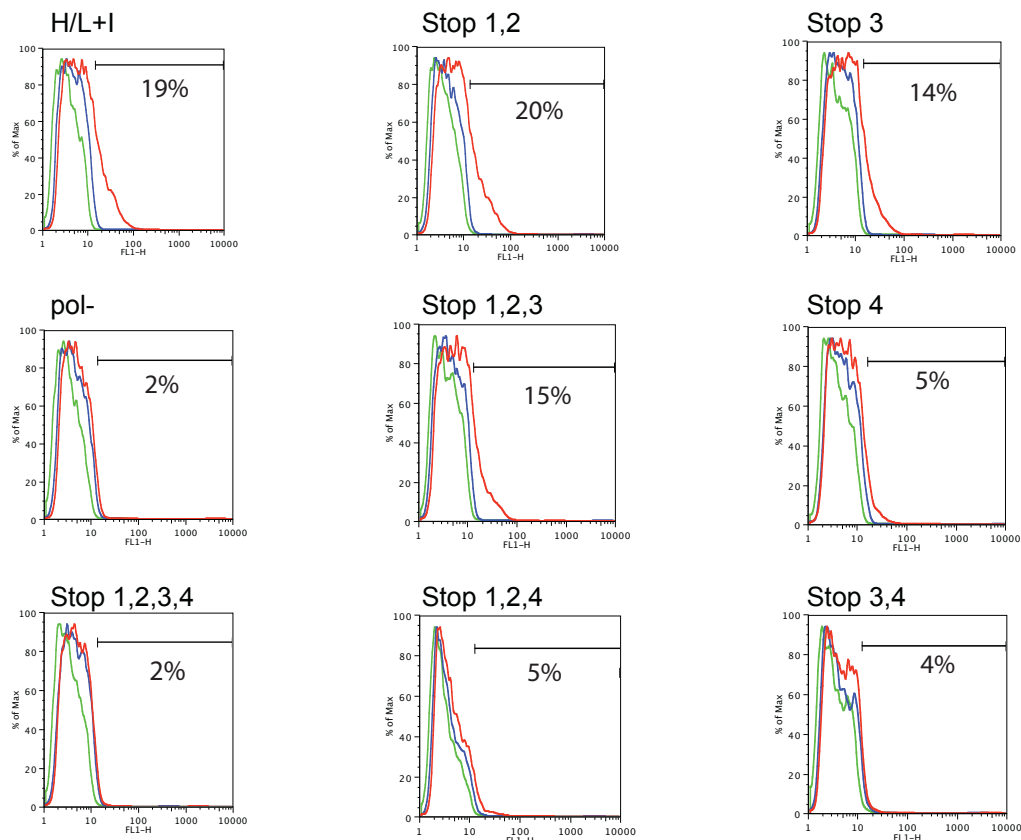


Figure 4.3. Flow cytometry of Huh-7.5 cells positive for HCV 4 days post electroporation. Histograms show percent cells stained with anti-NS5A (red), IgG2a (blue), and cells unstained (green). The percent positive for NS5A is indicated.

from transfection with genomes with Stop 1,2,4 and Stop 4 alone (Figure 4.3). Stops 3 and 4 together resulted in only 4% of the population being HCV positive. These results suggest that mutations for Stops 3 and 4 are deleterious for replication. This indicates that the reversions observed *in vivo*, which replaced the Stop 3 and 4 mutations with the original H77 sequence, were selected because they improved the replication ability, and thus the fitness of the virus.

***The Stop mutations do not affect expression or cleavage of the polyprotein.***

The mutations of Stop 3 and Stop 4 impair replication. We examined the ability of HCV RNA containing the Stop mutations to express and process the polyprotein. DNA plasmids with the HCV genome behind a T7 promoter were transfected into Huh-7.5 cells. Subsequently, cells were infected with a recombinant vaccinia virus expressing the T7 RNA polymerase, which transcribed the HCV RNA in the cytoplasm of the cell. Translation was mediated by the HCV IRES, and the polyprotein was processed into the separate virus proteins. Newly generated proteins were <sup>35</sup>S-Met labeled, HCV proteins immunoprecipitated with serum from an HCV infected patient, and analyzed by SDS-PAGE and autoradiography. We compared the expression and processing of HCV proteins from genomes

containing the Stop mutations (Figure 4.4). No difference in the expression levels of NS3, NS4B, and NS5A could be detected except in the lane for Stop 3,4. The lower amount of radiolabeled HCV proteins correlated with a lower level of background contaminants therefore it appears the low level is due to a loading error. In this system, the Stop mutations had no effect on HCV expression and polyprotein processing.

***The ARF is dispensable for HCV replication in cell culture.***

The replication defect in genomes with silent mutations in the capsid gene, yet conveying Stop codons in the ARF, could be due to lack of expression of an ARF protein product. To address this possibility, HCV genomes were constructed with point mutations in the same location as the Stop mutations; however, these genomes were engineered with both the capsid and ARF reading frames open. These “open” substitutions did not alter the capsid amino acid sequence, but did change the coding of the alternative reading frame. The genome with Open 3 mutations contains two substitutions (T471C, G473C) and Open 4 consists of three substitutions (G479C, C480A, C482G). Open 1 and Open 2 both contain only one substitution each (C407T and T434C) (Table 2.1 and Figure 4.5). These substitution sets were introduced into the H/L+I parent, *in vitro* transcribed

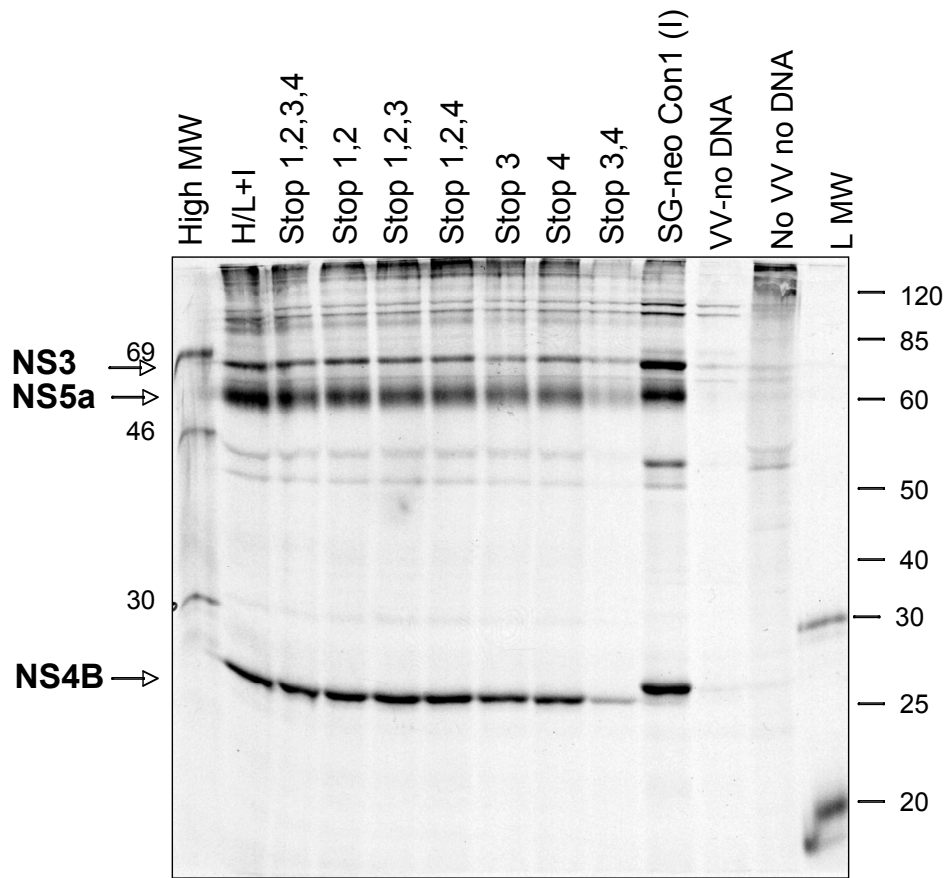


Figure 4.4. The Stop mutations do not alter polyprotein processing. Immunoprecipitation of radiolabeled HCV proteins expressed from genomes containing the Stop mutations. cDNA clones containing the HCV genome behind the T7 promoter were transfected in Huh-7.5 cells, and HCV RNA transcribed from the T7 polymerase provided by infection with the recombinant vaccinia virus vTF7-3. Proteins were radiolabeled with  $^{35}\text{S}$ -Met, and HCV proteins immunoprecipitated with serum from an HCV patient. A subgenomic replicon was included as a control.

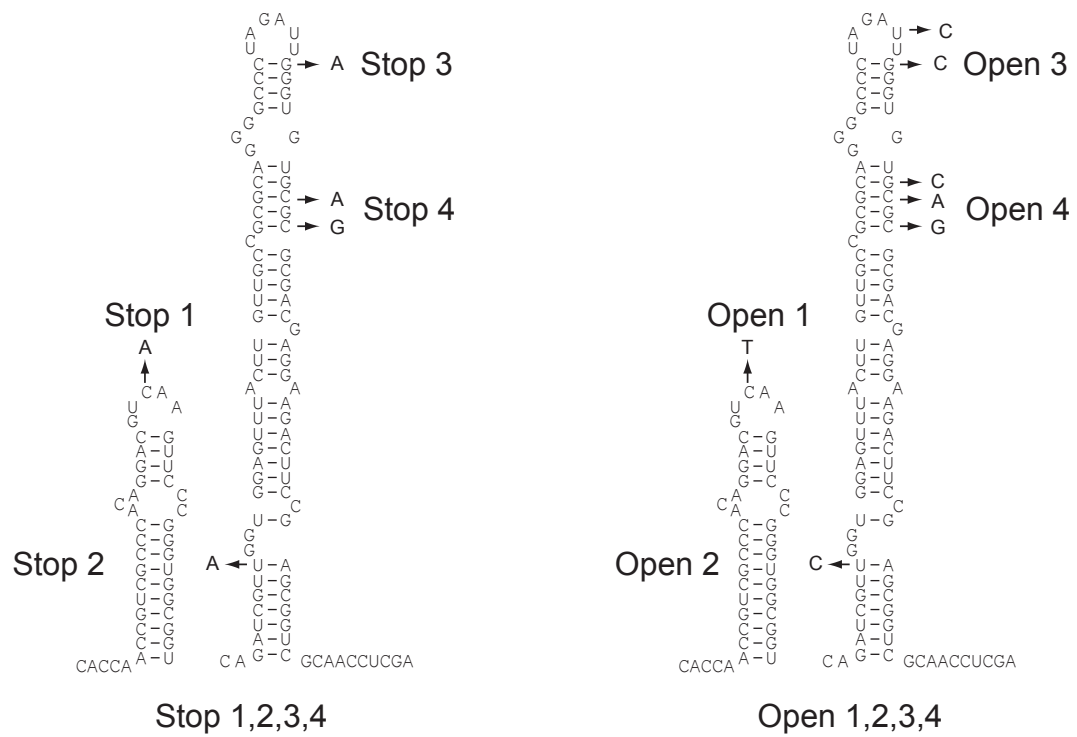


Figure 4.5. H77 SLV and SLVI with mutations for Stop and Open sets indicated. The mutations do not change the capsid protein.

RNA transfected into Huh-7.5 cells, and replication measured as described previously. At day 5 post electroporation, the pol- genome measured RNA levels to 3% of the parental genome. Similar to the genomes with the Stop mutations, the substitutions of Open 3 and Open 4 were detrimental to replication. Genomes with Open 3, 4 alone, or in combination with Open 1 or Open 2, replicated to just 1% of wild type (Figure 4.6); the same level as the genome with all five of the Stop mutations, Stop 1,2,3,4. Although both genomes were severely compromised for replication, Open 1,2,3,4 replicated to a two fold lower level than Stop 1,2,3,4. This could be due to the additional five substitutions in the Open 3,4 mutation set as opposed to the three in Stop 3,4. Interestingly, the mutations of region 3 and region 4, either Stop 3,4 or Open 3,4, had a dominant replication defect. The mutations of regions 1 and 2 had no effect independently; yet contributed a further reduction in the replication capacity when combined with the regions 3 and 4 mutations. The genome with Open 3,4 substitutions was 100 fold reduced from wild type, yet Open 1,2,3,4 showed a 400 fold reduction. An identical pattern was observed with the Stop mutation set. Thus, mutations in regions 1 and 2 exert a synergistic effect on the replication defect caused by mutation in regions 3 and 4. The replication defect in cell culture contributed by the Stop and Open mutations in the capsid gene indicated that either an

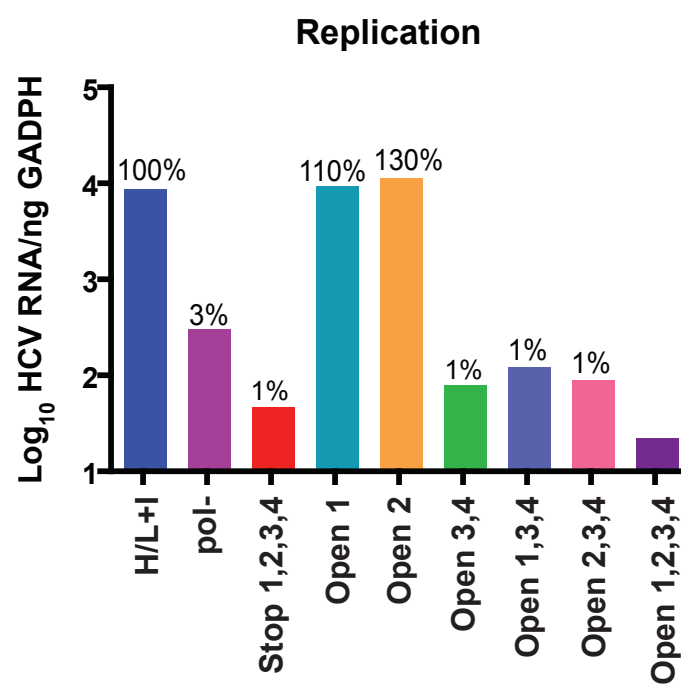


Figure 4.6. HCV replication of Open set of mutations. RNA levels were measured by Real Time RT-PCR day 5 post electroporation.

ARF protein product does not participate in HCV replication or that the amino acid alterations introduced in the ARF were disruptive for replication.

***An RNA structure in the capsid gene sequence is involved in HCV replication.***

As discussed in chapter 3, the conservation of the capsid gene may indicate the use of an ARF or of functional RNA secondary structures. Two helical stems, named SLV and SLVI, have been predicted by phylogenetic studies of codon usage and covariant base changes<sup>133, 147</sup>. The substitutions conferring Stop codons designed to block expression of the ARF would also disrupt these predicted RNA secondary structures. The mutations of the Stop and Open substitutions would disrupt the large helix of SLVI, but, as mentioned above, we could not discriminate between an absent or dysfunctional ARF protein and an alteration of the RNA structure as the cause of the replication defect. To address this, using the genome containing all five substitutions of Stop 1,2,3,4, we introduced compensatory mutations that would be predicted to restore the base pair interactions of SLVI (Figure 4.7). Since the mutation for Stop 1 is located in the loop of SLV, no compensatory change was possible. The mutation for Stop 2 substituted an adenosine for a uridine at the base of SLVI, so we introduced the



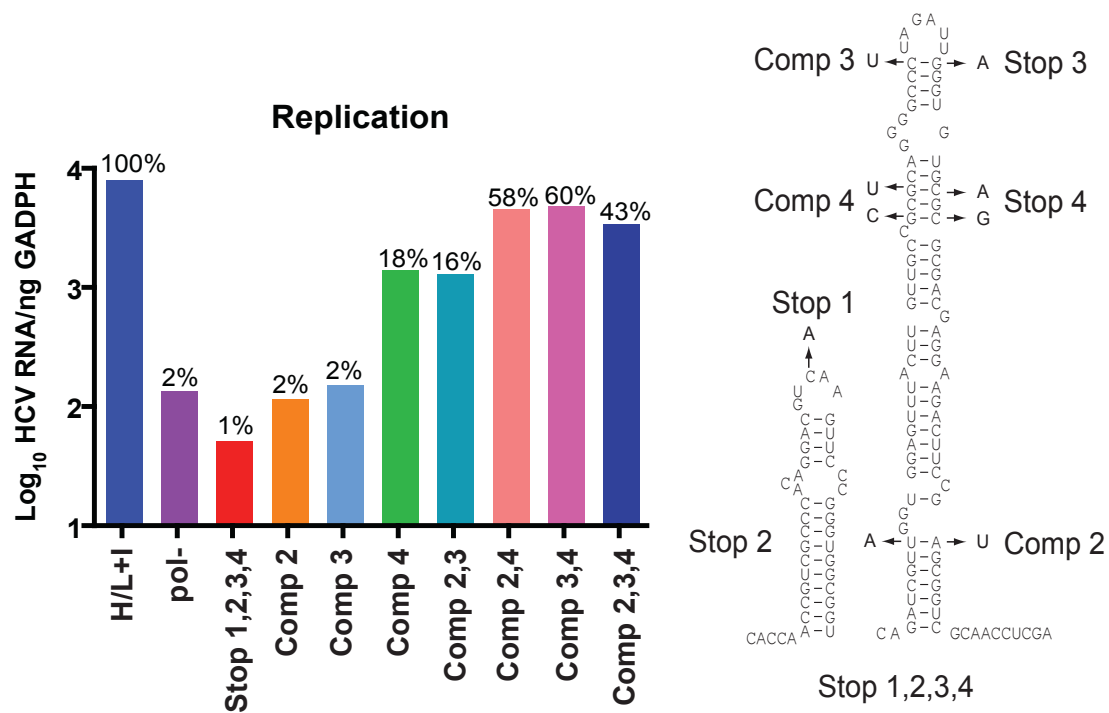


Figure 4.7. HCV replication of compensatory set of mutations. RNA levels were measured by Real Time qRT-PCR 5 days post electroporation. Compensatory mutations of H77 SLV and SLVI are shown on the predicted RNA structure.

compensatory substitution A502T. This change alters the amino acid of the capsid protein from a glutamate to a valine at position 54 (Table 2.1). The mutations for Stop 3 and Stop 4 disrupt three base pairs at the top helix of SLVI, therefore three mutations were introduced to restore these interactions. These mutations also change the capsid protein at P42L and R39L. The ARF remains blocked by the four original Stop codons. Genomes containing the Stop 1,2,3,4 mutations with individual or combinations of compensatory mutations were assayed for their ability replicate in cell culture. The replication level of the genome of Stop 1,2,3,4 was severely debilitated to 1% of wild type. On their own, the compensatory mutations for region 2 and region 3 (Comp 2 and Comp 3), did not restore replication. However, providing the compensatory mutation for Stops 2 and 3 together, Comp 2,3, replication did increase 25 fold over the parent Stop 1,2,3,4 (Figure 4.7).

Recall that replication for the genome with the Stop 1,2,4 mutations was debilitated to 18% of wild type, indicating that mutation of Stop 4 has the most severe replication defect of the tested mutations. The genome with Comp 3 alone could not restore replication of Stop 1,2,3,4 to the level of Stop 1,2,4, and required the addition of the Comp 2 mutation to replicate to levels 16% of the H/L+I sequence. In contrast, compensating the Stop 4

region allowed for more robust replication. Genomes with only the Comp 4 mutations improved replication 27 fold above the parent. The benefit did not reach the level of Stop 1,2,3 mutations, which was 76% of wild type, but replication was improved to 18% of wild type. The addition of compensating mutations for Stop 2, Stop 3, and Stop 4 further improved replication: Comp 2,4 showed an 88 fold increase and Comp 3,4 measured levels 93 fold above the Stop 1,2,3,4 parent. Genomes with the compensating mutations for Stop 3, 4, Comp 3,4, could replicate to 60% of the H/L+I genome; however, did not restore replication to level of Stop 1,2 (which was equivalent to the wild type). These results imply that it is the defect of an RNA structure, and not the absence of an ARF protein, that debilitates replication.

***RNA structure probing in solution reveals that Stop 3 and 4 mutations disrupt RNA secondary structure.***

Genetic analysis revealed HCV replication is influenced by RNA secondary structures in the capsid region. Disruption of SLVI by the introduction of base pair breaking mutations, led to a decrease in the replication capacity of HCV genomes in cell culture. Computer modeling programs of the thermodynamic folding of RNA, MFOLD, predicted the top of SLVI to be disrupted by the Stop 3 and Stop 4 mutations (Figure 4.8). We

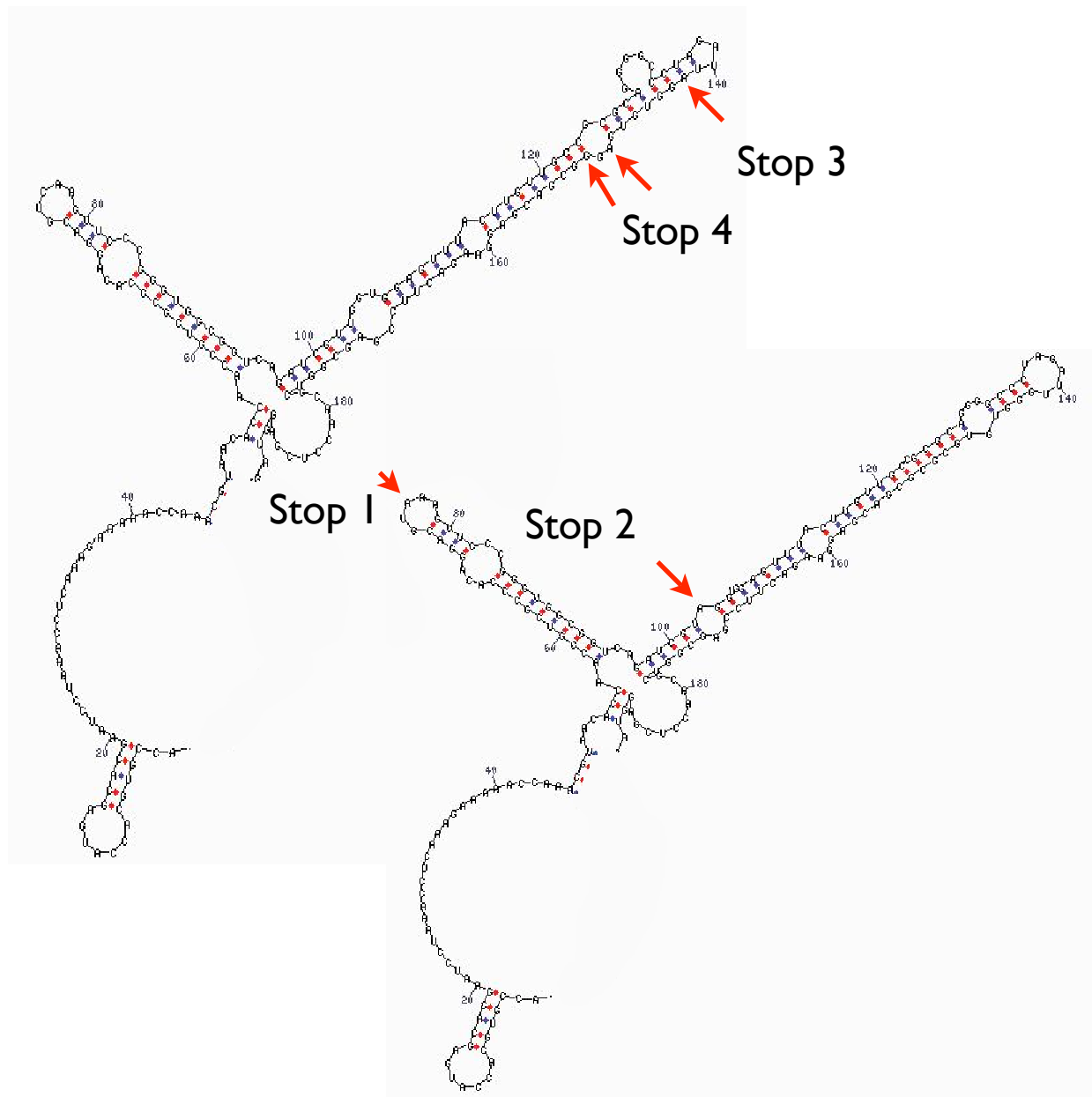


Figure 4.8. The mutations for Stop 2,3, and 4 would predictably disrupt SLVI. The predicted RNA structures for Stop 1,2 and Stop 3,4 generated from Mfold are shown.

wanted to confirm these predictions using the biochemical method of detecting paired and single stranded RNA in solution. Using the RNase T1, which cleaves the 3' phosphate bond of single stranded guanine and the cobra venom nuclease V1, which cleaves double stranded RNA, we probed the RNA structure of the capsid region using primer extension analysis (Figure 4.9). The RNA containing Stop 1,2,3,4 mutations displayed different sites of T1 cleavage than the wild type. In the mutant RNA, T1 cleavage was detected at 481A and 482G, which are the locations of the Stop 4 mutations. T1 cleavage sites could also be detected on the other side of the stem bearing the Stop 4 mutations at positions 452G and 455G. This data demonstrates that the Stop 4 mutations disrupt RNA base paired interactions. Cleavages were also detected in the location of Stop 3 at 474G and 475G in both the wild type and mutant RNA indicating potential "breathing" of unpaired bases near the loop of the stem. V1 nuclease cleavage sites indicating helical RNA were infrequent, but we did detect cleavage in the wild type RNA 464C and 465C, which is the base below the loop and at 478U, which lies below the G bulge in the top of SLVI. In the mutant RNA, a V1 cleavage site was detected in the loop at 476. Thus, the Stop 3 and Stop 4 mutations disrupted base pairing of the top of SLVI and may have induced base paired interactions within the loop of SLVI.

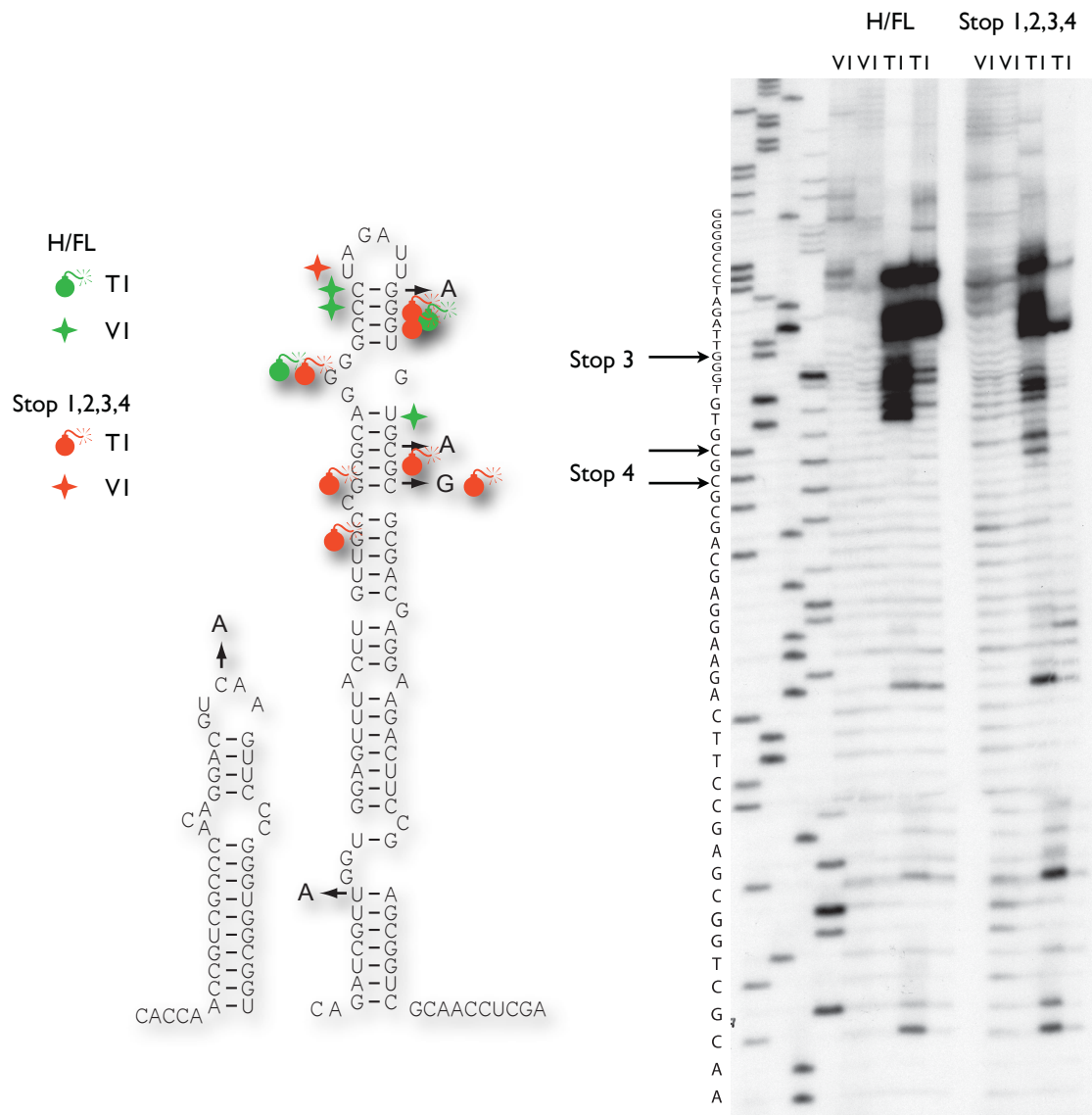


Figure 4.9. RNA structure probing reveals mutations in SLVI disrupt the top of SLVI. Cleavage with nucleases V1 (0.1U and 0.02U) and T1 (1U and 0.2U) demonstrate that Stop 3 and Stop 4 mutations disrupt base paired interactions at the top stem of SLVI.

***RNA structure in the capsid gene influences replication in diverse genotypes.***

Replication studies of HCV in cell culture have been challenging. Currently, few sequences are capable of replication in Huh-7 cells and all but one requires cell culture adaptive mutations. The replication studies described so far have involved the full length H77 genome of genotype 1a containing the adaptive mutations L+I. To test if the mutations in SLVI impaired replication in sequences from other genotypes, we introduced the Stop mutations into the genotype 1b sequence of Con1 and the genotype 2a chimera J6-JFH. First, two Con1 sequences with different adaptive mutations were tested. The Con1/ I sequence contains the S2204I mutation, one of the same adaptive mutations in the H L+I genome, and the Con1/ GIT contains three adaptive mutations: two in NS3 (E1202G and T1280I) and one in NS4B (K1846T). As observed for the genotype 1a genome, the mutations for Stop 1,2 had no effect on, or actually improved, replication, as opposed to genomes containing mutations for Stop 3 and 4 which reduced replication to 20% of wild type (Figure 4.10). A dramatic reduction in replication, to 3 - 9% of wild type, again distinguished the genomes containing all five substitutions for Stop 1,2,3,4 (Figure 4.10). Recently, an

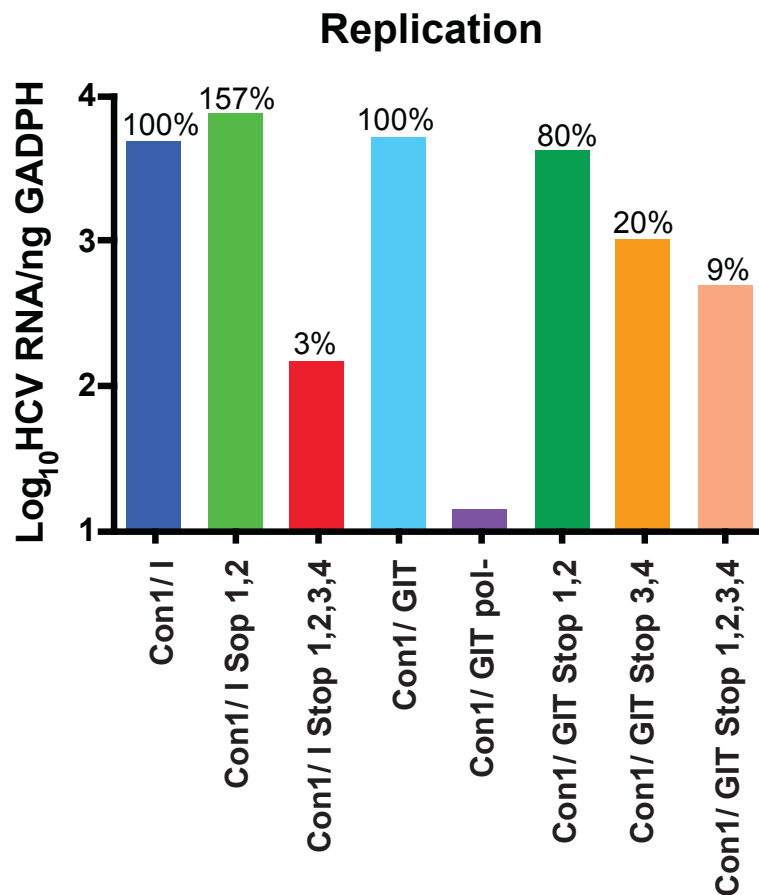


Figure 4.10. Stop 1,2,3,4 and Stop 3,4 convey a replication defect in Con1 sequence genotype 1b with adaptive mutations I and GIT. HCV RNA was measured by Real Time qRT-PCR after 5 days post electroporation with percent differences from wild type indicated above each bar. Two independent RNAs were included for each mutation and the average shown.



HCV isolate, JFH-1, has proven to efficiently replicate without any adaptive mutations and produce infectious virus in Huh-7.5 cells. These remarkable features appear to stem from the replicase proteins of JFH-1, therefore chimeric genomes can be constructed where the 5' NTR and capsid-NS2 portions of the genome can originate from other genotype 2a sequences, such as the J6 sequence. Other genotype sequences may also be fused with the JFH-1 replicase proteins and 3'NTR and efficiently replicate; however, adaptive mutations are required for virus production.

We wanted to test if the influence of SLVI is conserved in a genome with greater than 30% difference from the H77 sequence and if the mutations of SLVI also conveyed a phenotype in the more robust replication environment of JFH-1<sup>60</sup>. The mutations for Stop 3 and 4 were introduced into the genotype 2a J6-JFH genome and replication was measured by quantitative real time qRT-PCR at 3 days post electroporation. Replication was compromised to 18% of wild type levels (Figure 4.11). The replication ability of the mutants was also tested in a genotype 1a-JFH-1 chimera where the 5'NTR to NS2 coding region from H77 is fused to the NS3 gene to 3'NTR of JFH-1. While this chimera can replicate efficiently, it requires two adaptive mutations for virus production, (I348V/S1103T), hence is named H-JFH/ V+T. Consistent with previous observations, the mutations for Stop

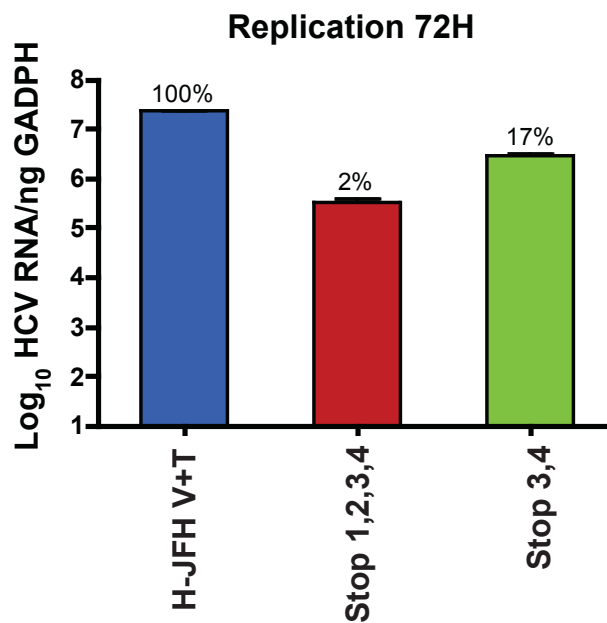
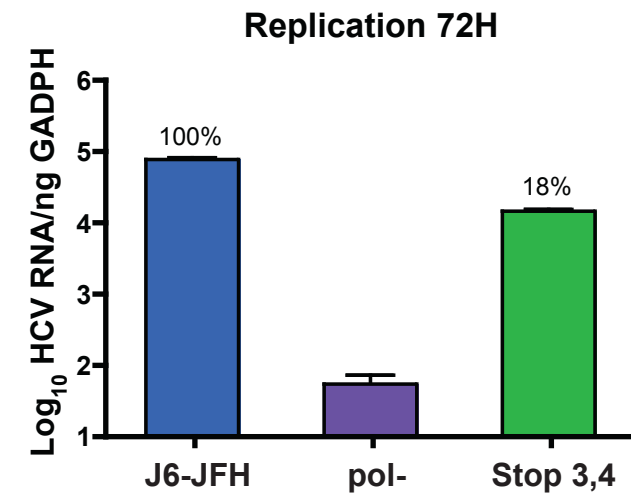


Figure 4.11. Stop 1,2,3,4 and Stop 3,4 convey a replication defect in genotype 2a and 1a-2a chimeras. HCV RNA was measured by Real Time qRT-PCT at 72H post electroporation. Two RNAs were tested for each genome and averages shown. The percent differences from wild type are listed above each bar.

1,2,3,4 contribute to a severe replication defect whereas mutations for Stop 3,4 conveyed only a 17% reduction (Figure 4.11).

***The use of luciferase reporter genomes with the JFH-1 replicase.***

Genomes with the JFH-1 replicase have increased our ability to easily detect replicating HCV RNA and allow us to test for virus assembly and infection in cell culture. A reporter genome was created that expresses the Renilla luciferase protein in the context of the HCV. The Renilla gene was cloned into the junction of the p7 and NS2 genes and included the 2A protease from Foot and Mouth Disease virus (FMDV) at the carboxyl terminus to allow proper processing (Figure 4.12). The Renilla gene and 2A protease were introduced into the parent genome of the H77- JFH chimera with the V+T adaptive mutations, H-JFH/ V+T, to create H-JFH/ Rluc. Insertion of the reporter and protease allowed efficient replication, albeit one log below the parent (Figure 4.12). Virus production was reduced from  $6 \times 10^3$  TCID<sub>50</sub> /ml of the parent to 30 TCID<sub>50</sub>, but due to the sensitivity of the luciferase assay, virus infectivity was easily measured. Analysis of the infected monolayer by IHC revealed that about 5% of the cells were positive for NS5A antigen, therefore cells were not simultaneously infected and virus may spread during the analysis. Since expression of the luciferase gene was

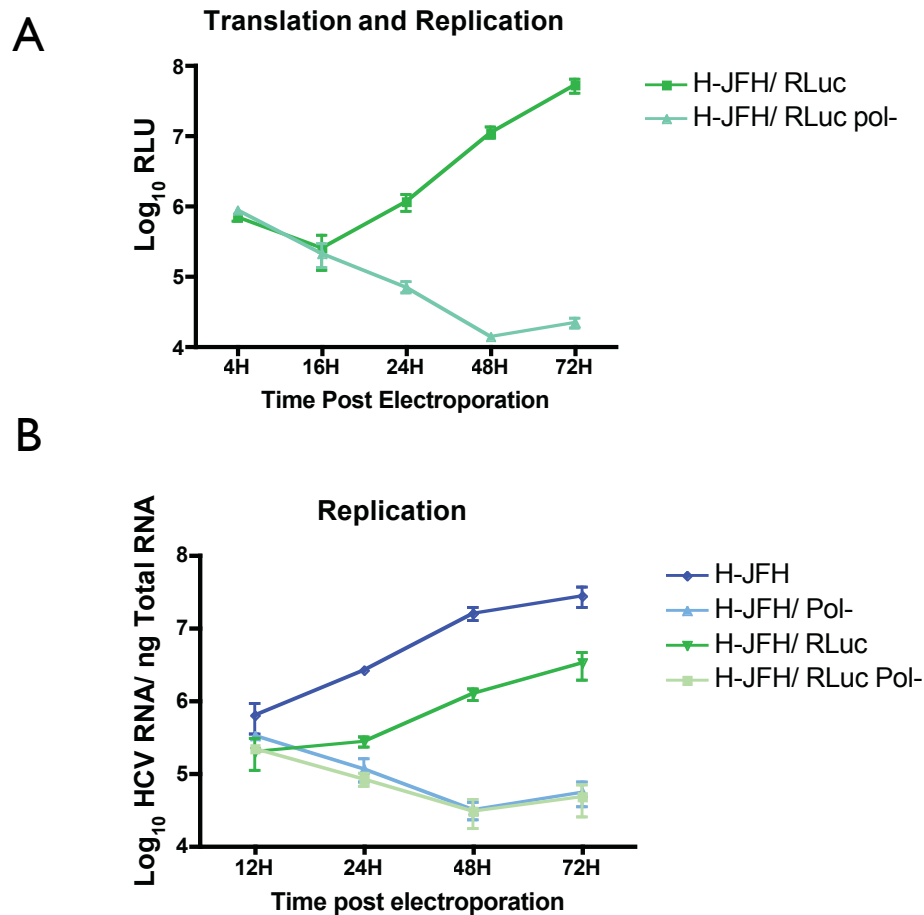
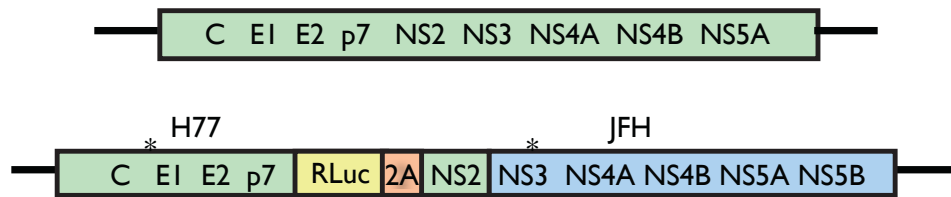


Figure 4.12. Genotype 1a genomes with a luciferase reporter between p7 and NS2 show similar replication kinetics, but with levels one log below the parental genome. A. Time course of luciferase activity of the reporter H-JFH/ RLuc. B. HCV RNA levels measured by Real Time qRT-PCR for the reporter and parental genomes.

dependent on translation of the polyprotein, the luciferase signal increased as replication generated nascent RNA that was itself translated. This demonstrated the link between translation and replication. As the RNA entered the cell, either via electroporation or infection, translation occurred and expression of the polyprotein and luciferase reporter was evident for all genomes, including the pol- genome at 4H (Figure 4.12). As the input RNA degraded, the luciferase levels declined was monitored for the pol- genome. The logarithmic phase of replication is detected at 12H by measuring HCV RNA by real time qRT-PCR (Figure 4.12). At 16H, luciferase expression rapidly increased as HCV RNA was generated by the replicase and entered into subsequent rounds of translation. By 48H the pol- genome was degraded to background levels.

Using these reporter constructs, we were able to monitor how replication and virus production were affected by mutations of RNA structures in the capsid gene. The mutations for Stop 1,2,3,4 and Stop 3,4 were introduced into the H-JFH/ Rluc genome and replication measured. Initially, luciferase levels increased between 4H and 8H as the input RNA was translated and luciferase protein was accumulated (Figure 4.13). After 8H, luciferase levels began to decline as luciferase degraded, with a half-life of approximately 6H. Replication could be detected at 16H as levels for the

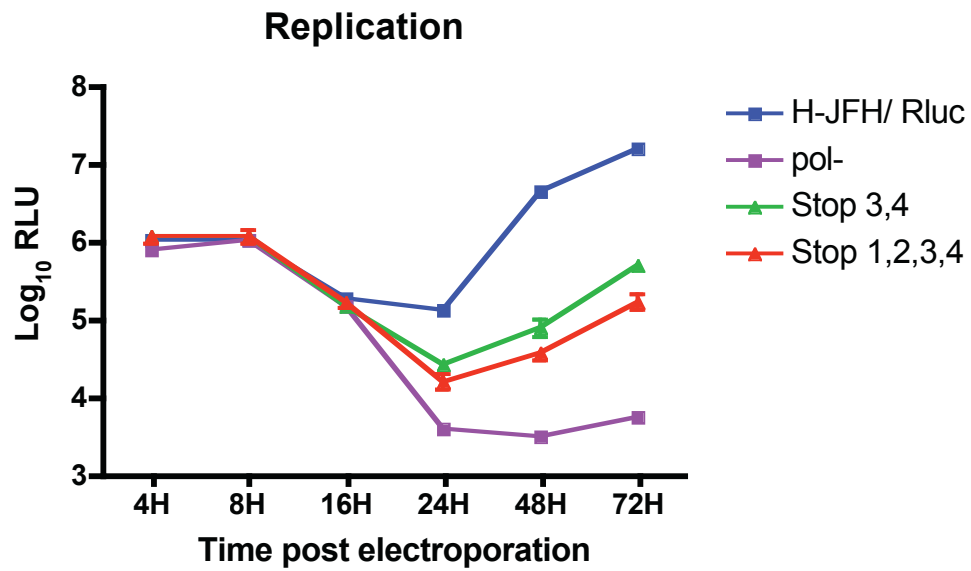


Figure 4.13. Mutations for Stop 1,2,3,4 and Stop 3,4 show a reduction for replication at 16H post electroporation in the reporter H-JFH/ Rluc. Two RNAs were transfected and duplicate wells measured for luciferase activity for each genome. The averages are shown.

pol- genome continued to decline yet the wild type genome began to rise. Following the initial translation from input RNA, the signal from the Stop mutants also declined, but not to pol- levels. After 24H, replication accelerated for the wild type genome. Levels of the Stop mutant genomes also rose, but did not reach wild type levels as the increase began from a lower point than the wild type. This lag in the exponential phase of replication could be due to numerous factors as the translation and replication processes overlap and interactions with the host cell or movement of RNA between replication and translation could alter the replication kinetics. Apparently, the mutations in the RNA structures contributed to a less than optimal transition between one or multiple components required for efficient replication. The result was a lower total level of HCV RNA and reporter expression. The reporter levels for the mutant genome Stop 3,4 was 1.7 logs below wild type, and Stop 1,2,3,4 replicated 2.1 logs lower at 48H post electroporation.

Using the luciferase reporter, we are able to detect virus infection of naïve cells. The sensitivity of the luciferase assay allows for consistent detection and quantitation of mutant HCV genomes that could not easily be studied using IHC or Real Time qRT-PCR. However, the requirement of the reporter to be translated and the ability of the genome to undergo replication

in infected cells, potentially complicates the result. If a mutant conveys a defect in translation or replication, differences in the luciferase activity in the infected cells may be a result of reduced translation, replication and virus production. In this regard, we refer to the luciferase activity measured from infected cells as “virus infectivity.” This general term is used to describe measurement of virus production, entry, and subsequent translation and replication in the infected cell. There was no measurable virus infectivity for the mutant Stop 1,2,3,4 and the Stop 3,4 genome, likely a consequence of the low levels of replication.

***The Open mutations impair replication and virus infectivity.***

As shown previously, the mutations consisting of the Open set were severely detrimental to HCV replication if located in the top of SLVI. These mutations were engineered into the H-JFH/ Rluc parent and luciferase activity measured. All genomes with mutations in the top of SLVI, Open 3 and Open 4, had impaired replication activity, even when present with the more robust JFH replicase (Figure 4.14). Again, there was a lag in the transition to the logarithmic replication phase and levels remained 1-2 logs below wild type. As in the H/ L+I parent, genomes with mutations Open 1 and Open 2 alone did reach wild type levels. The failure of the mutants to



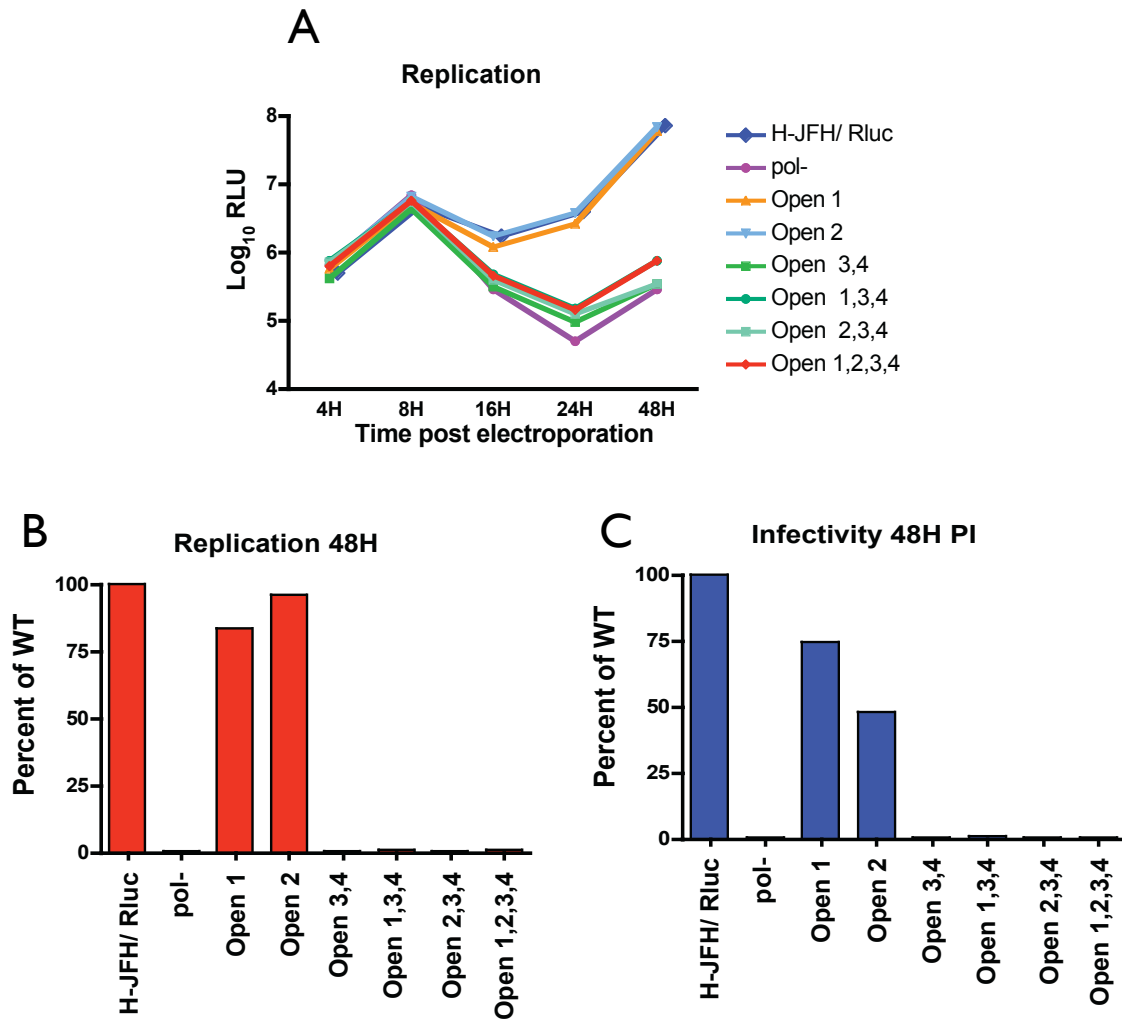


Figure 4.14. Mutations of Open 3 and 4 reduce replication and eliminate virus infectivity when both the capsid and ARF can be expressed. A. Time course of luciferase expression. B. Luciferase expression displayed as percent of wild type at 48H post transfection. C. Virus infectivity graphed as percent of wild type at 48H post infection. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

accumulate a substantial amount of HCV RNA correlated well with virus production. Genomes capable of efficient replication were also able to infect naïve cells (Figure 4.14). These results show that the mutations of Open 1 and Open 2 did not impede translation, replication, or virus infectivity.

***The mutations in SLVI do not alter RNA stability in Huh-7.5 cells.***

The reduced replication of genomes harboring mutations in SLVI could be due to RNA instability and preferential degradation in transfected cells. The mutations of SLVI were introduced into a replication-incompetent genome. Luciferase expression and degradation were compared between the mutant and parental genomes and found to be equivalent indicating similar RNA stabilities in transfected cells (Figure 4.15).

***The loops of SLV and SLVI slightly influence replication and virus infectivity.***

RNA structures are known to interact with virus and host proteins to coordinate replication and virus production. The predicted size of SLVI, 81 bases, leads one to hypothesize that it may participate in these functions. We introduced mutations to change the nucleotide sequence of three bases in the loop of SLV and three bases in the loop of SLVI (Figure 4.16). The

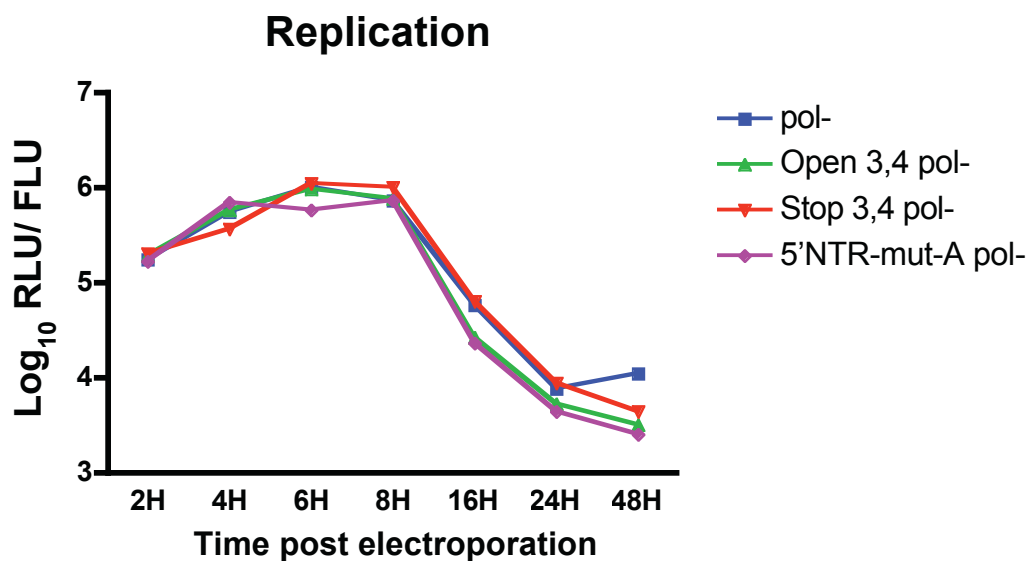


Figure 4.15. The replication defects contributed by mutations in SLVI and the 5'NTR are not due to RNA stability. Luciferase expression of mutant genomes with a defective polymerase show equivalent levels as a wild type genome with a defective polymerase.

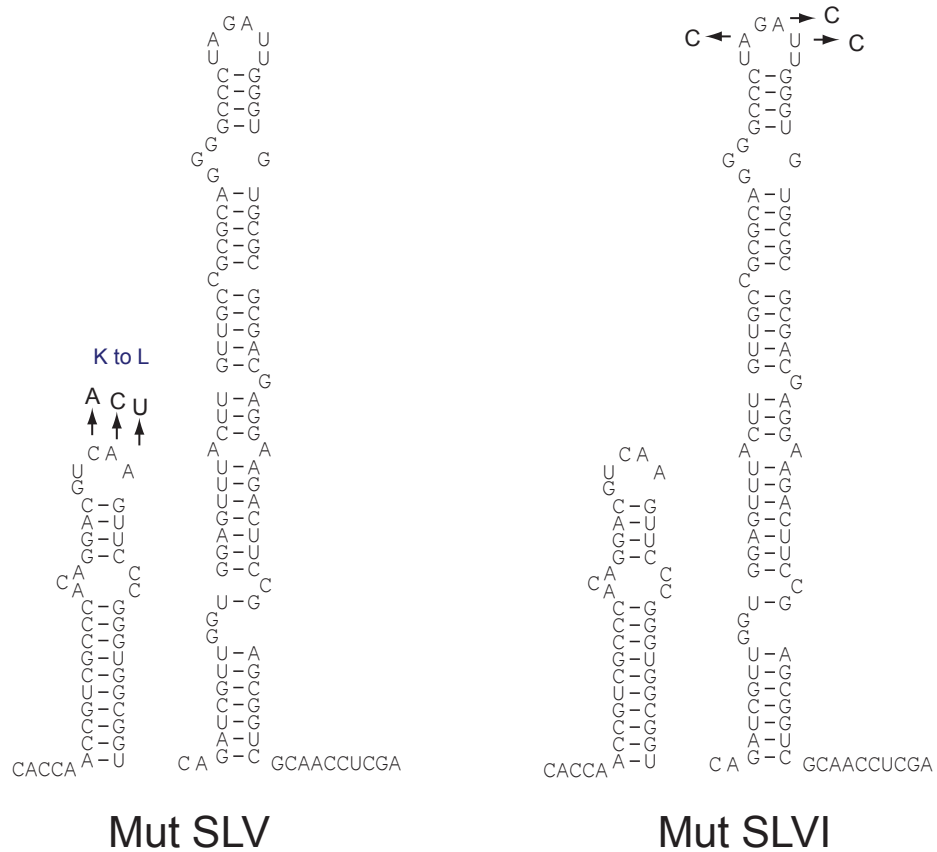


Figure 4.16. Mutations of the loop in SLV and SLVI in the H77 sequence. The mutation in the loop of SLV would change a lysine to a leucine in the capsid protein. The mutations in the loop of SLVI do not change the capsid protein.

mutations for the SLVI loop were silent for the capsid protein, but the mutations in SLV involved one amino acid change, a lysine to a leucine at position 23. Replication and infectivity were measured by luciferase assay. Mutations in the loops reduced replication to 44% and 42% of wild type at 48H post electroporation (Figure 4.17), but this reduction was not as severe as the replication defect observed for the mutations in the helices of SLVI. The difference between genomes with the mutant loops and wild type was apparent at an earlier time, 8H, when translation of the input RNA was occurring (Figure 4.17). This difference was not due to a transfection or preparation in RNA discrepancy as a control firefly luciferase reporter was always used to equalize transfection efficiency and two independent RNA preparations were used for each construct in each experiment. The kinetics followed a similar pattern, yet failed to accumulate to wild type levels within two days. Infectivity of naïve cells was also reduced to 54% for the mutant loop SLV which may be a result of the lower level of accumulated HCV RNA, the change in the amino acid of the capsid protein, or an effect of the RNA structure (Figure 4.17). Infectivity was further reduced to 24% for the genome with the mutated SLVI loop despite replicating to a similar level as the mutated SLV. This difference was not due to changes in the capsid protein and may indicate a defect in virus assembly. The three mutations in

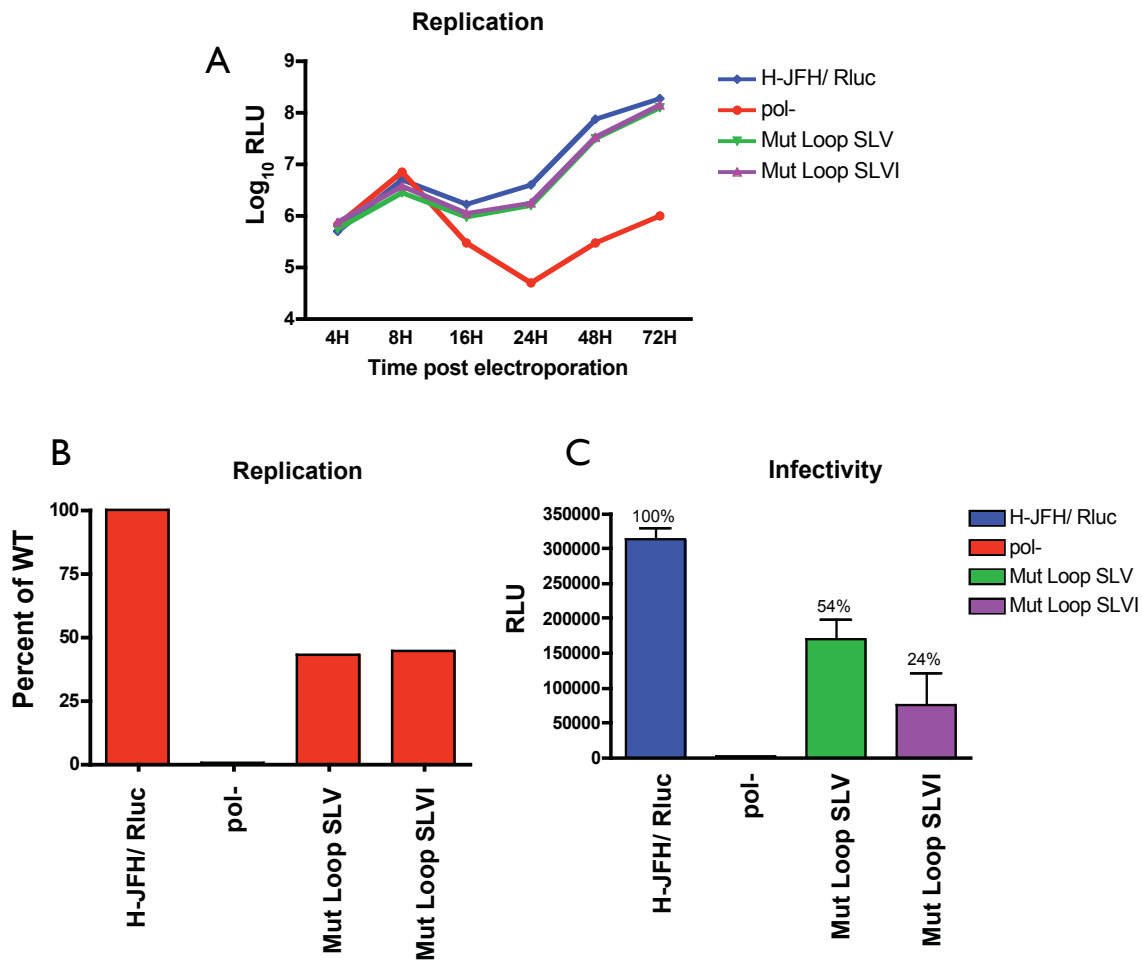


Figure 4.17. Mutations in SLV and SLVI slightly reduce replication and moderately reduce virus infectivity. A. Time course of luciferase expression. B. Luciferase expression displayed as percent of wild type at 48H post transfection. C. Virus infectivity graphed as percent of wild type at 48H post infection. Percent differences from wild type are noted above each bar. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

the loop of SLV conveyed a more detrimental replication effect than the single base change of Stop 1 or Open 1, both of which were located in the loop of SLV. This data indicates a role for the SLV loop in replication, but this region may not be as critical as the helix of SLVI.

***Mutations in the base of SLVI impair replication.***

Structure predictions of the helix of SLVI reveal a large stalk of 81 bases. We have demonstrated that the top helix of SLVI has a critical role in replication, while the loops of SLV and SLVI are also important, but to a lesser degree. To test the importance of the base of SLVI, we introduced three base mutations at the left side of the base of SLVI, called Base Left (Figure 4.18). All were silent for the capsid protein. We also created three mutations at the right side of the base to compensate for the left sided mutations (Figure 4.18). This mutant genome, Base Right, also generates an amino acid change at position 54 of the capsid protein with a glutamate to a valine substitution. The genome with both sets of mutations was named Base Left/ Right. Each set of mutations was tested in the parental H-JFH/ Rluc genome. Disruption of the base of SLVI reduced replication to 53% of wild type for Base Left and to 16% for Base Right at 48H post electroporation. Again the difference became apparent at the 16H time-point; the window for

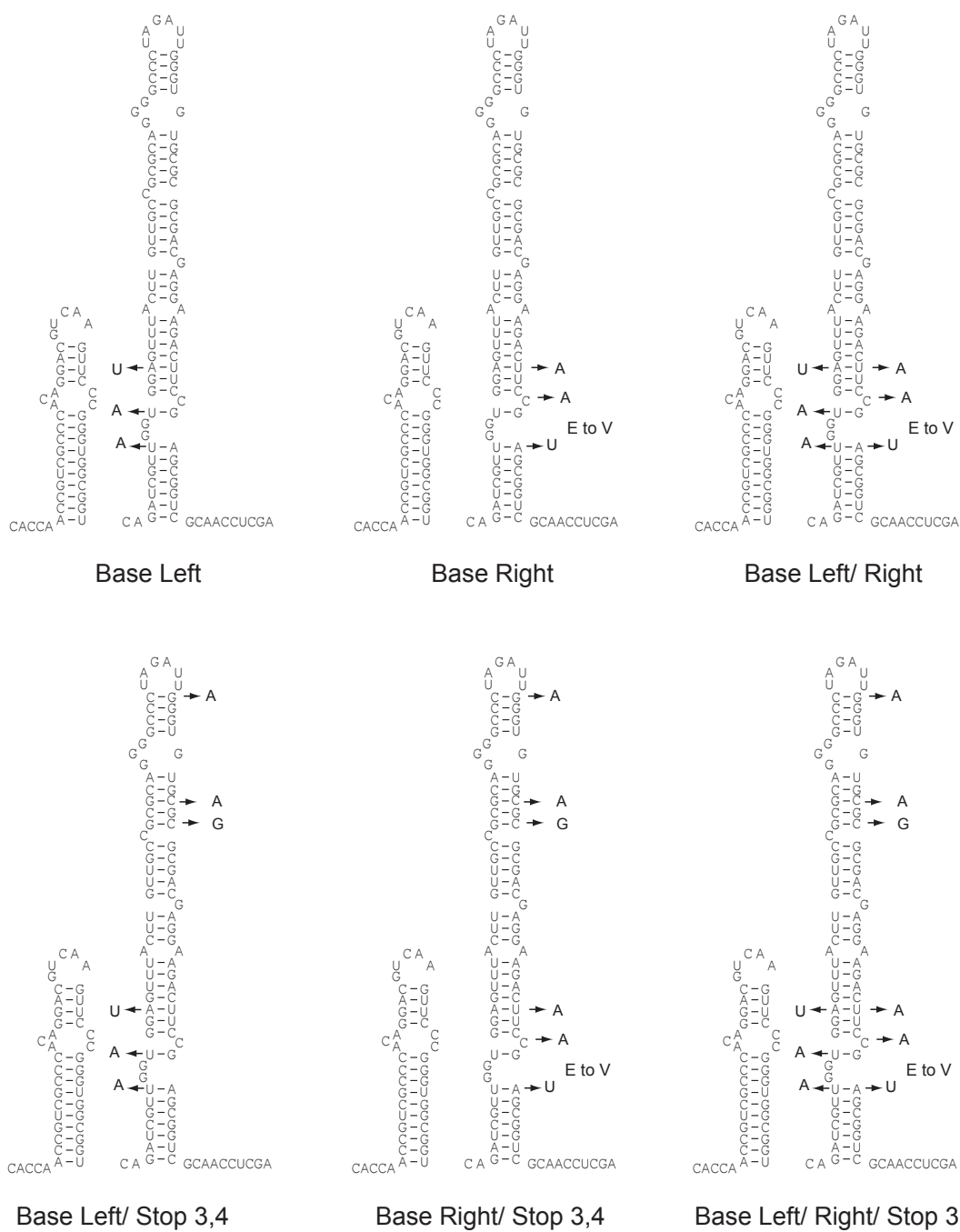


Figure 4.18. Mutations introduced at the base and top of SLVI in the H77 sequence.



transition to efficient replication (Figure 4.19). The compensatory mutant, Base Left/ Right, was reduced to 45% of wild type, an improvement from the defect with the Base Right mutations alone. There was also reduced virus infectivity from genomes where the base of SLVI was disrupted. With the level of replication reduced to just half of wild type, virus infectivity was down to 32% for Base Left. The Base Right mutations eliminated virus infectivity (Figure 4.19). This could be a result from the amino acid change introduced at residue 54 in the capsid protein; infectivity was not restored with the compensation of Base Left/ Right. This mutant genome was also examined in a later study concerning its ability to produce infectious virus (Chapter 5).

***Mutations in the base of SLVI additively impair replication with mutations in the top of SLVI.***

The base of the SLVI does influence replication, but not as drastically as mutations in the top stem of SLVI. To determine if the replication phenotype observed for the top of SLVI is influenced by the base structure of the stem, we combined the mutations of the base with those at the top of the helix, mutants Stop 3,4 (Figure 4.18). We found that mutations for Stop 3,4 in the apical stem contributed to a more severe reduction in replication

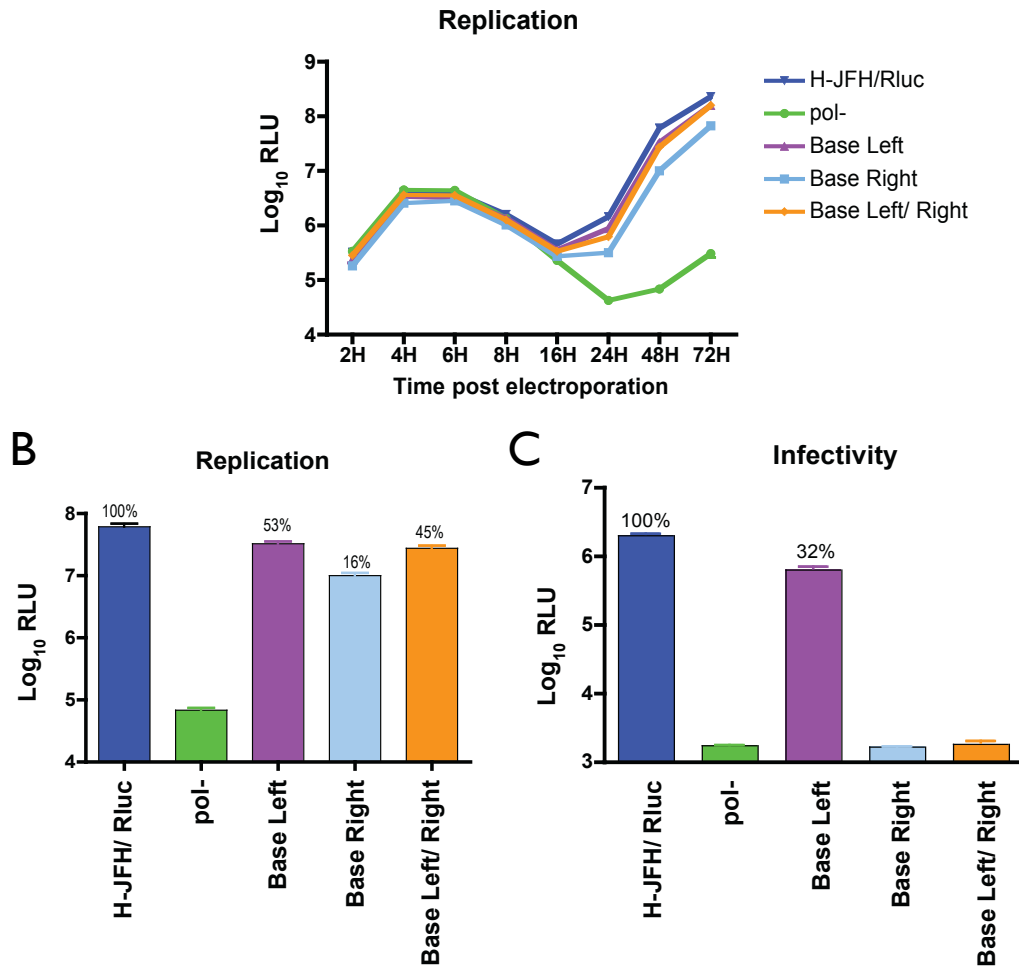


Figure 4.19. Mutations in the base of SLVI reduce replication. The Base Right mutations, which change an amino acid of the capsid protein, eliminate virus infectivity. A. Time course of luciferase expression. B. Luciferase expression at 48H post transfection. C. Virus infectivity at 48H post infection. Percent differences are noted above each bar for B and C. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

than mutations in the base of the stem. Genomes with both mutations in the top and the base of the stem were additively suppressed for replication (Figure 4.20). The mutations for Base Right were more deleterious than the mutations for Base Left, 12% versus 36%, and the replication of each was further reduced when combined with the Stop 3,4 mutations. The genome with Stop 3,4 / Base Left was 199- fold below wild type levels, and the combination with the Base Right mutations further reduced replication (876- fold below the parent). When base pairing was restored by the introduction of compensatory mutations, Stop 3,4/ Base Left/ Base Right, the replication defect was less pronounced (296- fold below wild type). Genomes with the compensatory mutations were still impaired, but not to the same levels as the Stop 3,4/ Base Right mutations (Figure 4.20). These results imply that while the top of SLVI has a critical role in replication, it is important to maintain the overall structure of SLVI for efficient replication.

***Mutations in the middle, c-bulge region and amino acids in the capsid protein that affect replication and infectivity.***

The central region of SLVI shows remarkable conservation at both the nucleotide and amino acid level. Extending down from the top loop and just below the Stop 4 mutations, the region around the c-bulge at nucleotide 454

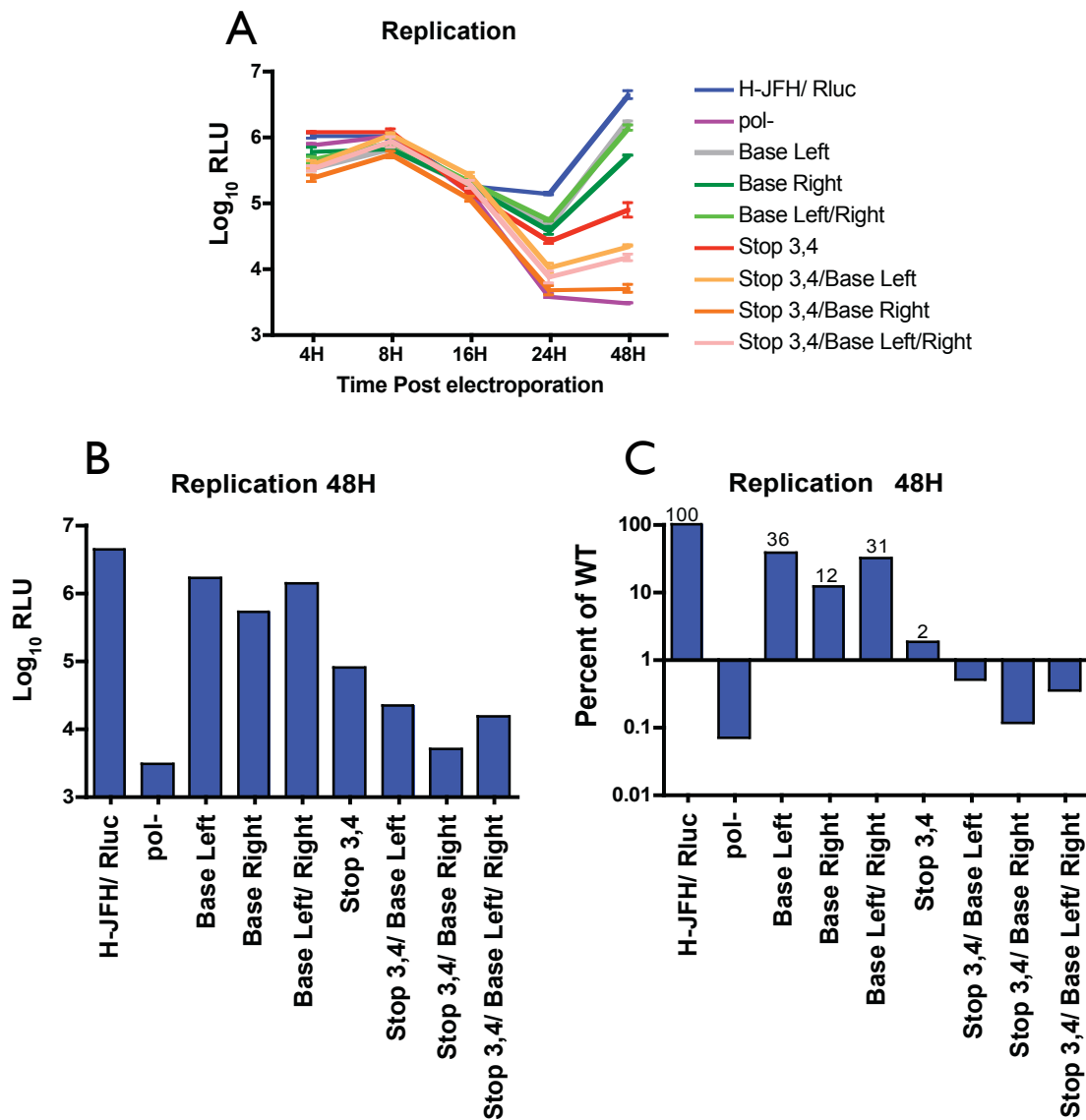


Figure 4.20. Mutations in the base of SLVI additively reduce replication in combination with mutations in the top of SLVI. A. Time course of luciferase expression. B. Luciferase expression at 48H post transfection. C. Percent differences from wild type at 48H. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

is conserved in over 90% of sequences from all six genotypes. This region is predicted to be a foundation for the overall structure of SLVI. The program RNAmute predicted the area surrounding the c-bulge and specifically nucleotide 484 to be critical to maintain the structure of SLVI<sup>29</sup>. To test if this region participates in replication, two mutations were introduced on either side of SLVI in the H-JFH/Rluc parent (Figure 4.21). The mutations Mid Right, C482A and C484G, disrupt two base paired interactions and test the predicted critical base at 484. An amino acid is also changed at residue 48 from an alanine to a glycine. The first mutation, C482A, was also one of the mutations in the Stop 4 mutant. To alter the left side of SLVI, the mutations Mid Left, G452C and G455A, were introduced, and change residue L37F of the capsid protein (Figure 4.21). Genomes with both mutation sets, Mid Right/ Left would restore the base pairing of SLVI in this region according to the prediction models. Replication was measured in three separate experiments, each using two independent preparations of *in vitro* transcribed RNAs. At 48H post electroporation, genomes with the mutations Mid Left or Mid Right were reduced for replication to 34% and 28% of wild type, respectively (Figure 4.22). Consistent with this, in a separate experiment, the genome with mutations for Mid Right gave lower levels of replication, 44% of wild type, than the genome with Mid Left,

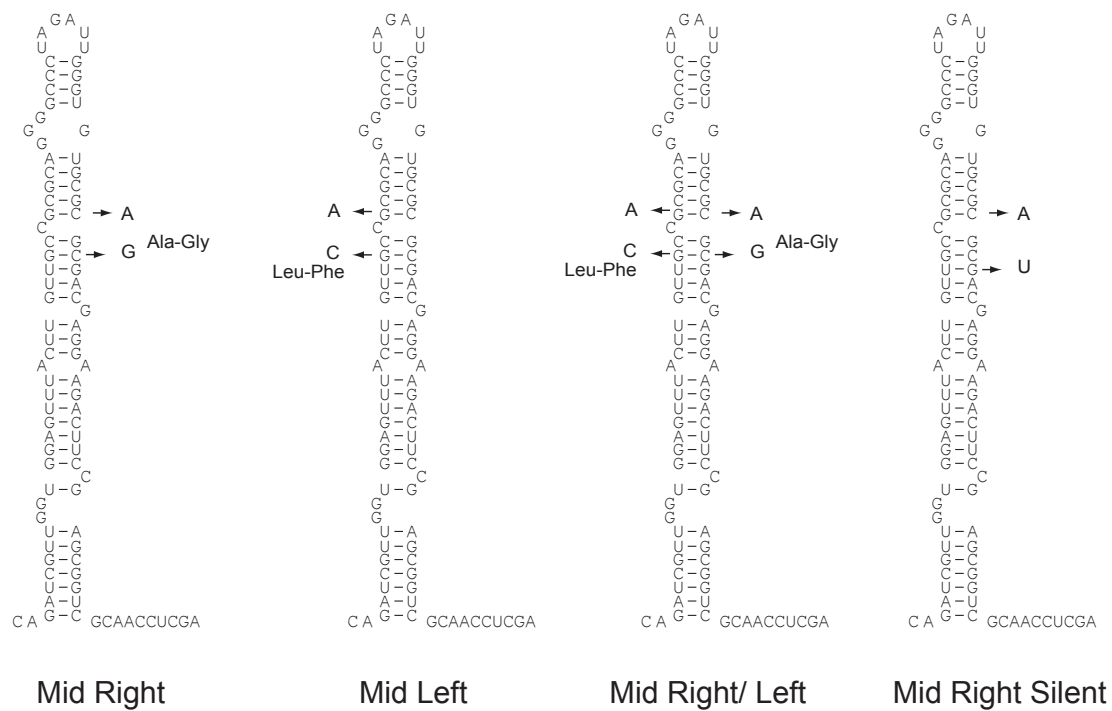


Figure 4.21. Mutations in the middle region of SLVI in the H77 sequence. Amino acid changes are noted.

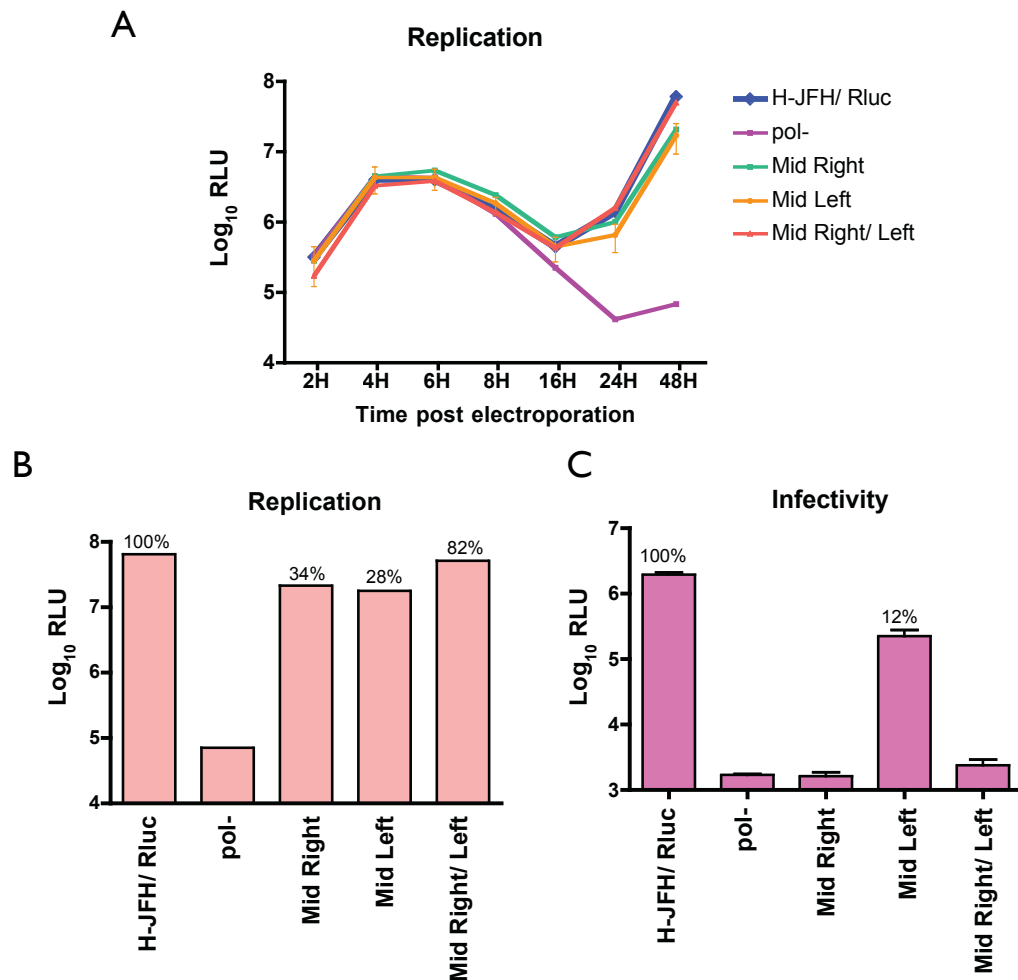


Figure 4.22. Mutations in the middle region of SLVI reduce HCV replication. Mid Right failed to make infectious virus. A. Time course of luciferase expression. B. Luciferase expression at 48H post transfection with percent differences from wild type noted above each bar. C. Virus infectivity 48H post infection with percent differences noted. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

which was 30% of wild type (data not shown). By combining both sets of mutations in the genome Mid Right/ Left, replication was improved to 82% of wild type. This improved replication implied a restoration of a functional RNA structure, as the compensatory genome efficiently replicated, while harboring the two amino acid changes in the capsid protein. Virus infectivity, however, was reduced by the amino acid changes. While both genomes, Mid Right and Mid Left, replicated to similar levels, only the genome with Mid Left mutations produced infectious virus (Figure 4.22). This indicate that either the original residue, alanine, at position 48 is required for virus production/ infection or that the structure of SLVI that facilitates infection was not sufficiently restored. To address these possibilities, the genome with Mid Right mutations was examined in additional assays discussed later (Chapter 5).

To distinguish if the RNA structure or amino acid sequence contributed to the deficiency in infection, we created a genome with mutations in the Mid Right region, but without changing any amino acids (Figure 4.21). The first mutation was the identical to Mid Right, C482A, but the second mutation, G485T, was one base downstream of the “critical base” predicted by RNAmute. Unexpectedly, this mutant genome, Mid Right Silent, displayed a more dramatic reduction in replication than did Mid



Right. The genome Mid Right Silent could only replicate to 6% of wild type levels whereas Mid Right was moderately compromised to 34% of wild type (Figure 4.23). This defect was barely apparent at 4H post electroporation, but was significantly delayed from the wild type level at the 16H to 24H transition window. Infectivity was reduced to 5% of wild type levels. This is probably a consequence of the reduced genomic RNA available for virus production due to the compromised replication. The replication defect by Mid Right Silent was observed repeatedly where the replication level was 6% of wild type and infection was reduced to 1.2% (data not shown). This result is in contrast to the Mid Right mutant genome, which could efficiently replicate, but not produce infectious virus (Figure 4.22). The larger replication defect by Mid Right Silent, which differs from Mid Right by one adjacent nucleotide, may indicate a sensitive area for the RNA structure. The abrogation of infectious virus production from Mid Right while virus replication is still evident, implicates A48 to be critical for production of infectious virus.

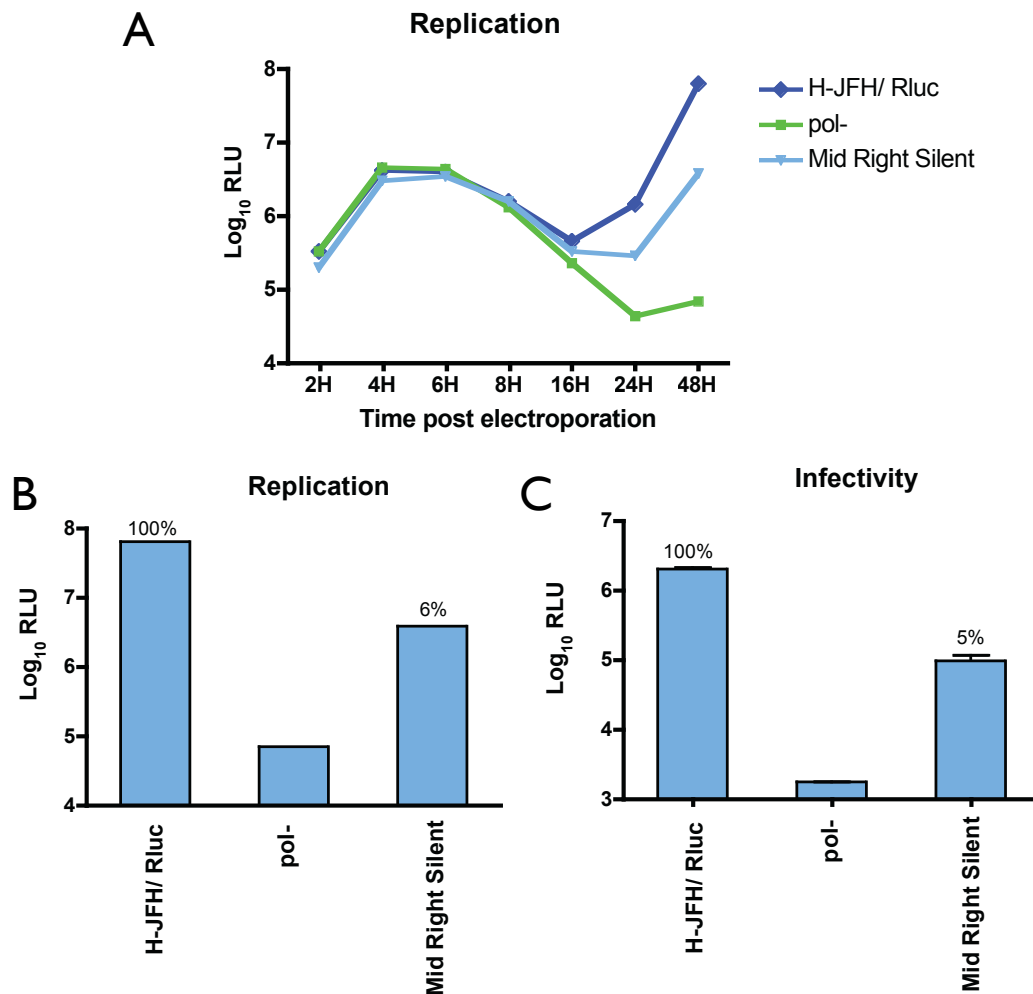


Figure 4.23. Mutations in the middle region of SLVI reduce HCV replication. A. Time course of luciferase expression. B. Luciferase expression at 48H post transfection with percent differences from wild type noted above each bar. C. Virus infectivity 48H post infection with percent differences noted. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

***Mutations in the c-bulge region impair replication in the genotype 2a genome.***

Mutations in the Mid region of SLVI were also examined in the genotype 2a J6-JFH/Rluc genome. This genome has similar levels of replication and virus infectivity as the H-JFH/ Rluc chimera but with a greater rate of replication from 24H to 48H (Figure 4.24). Replication was equivalent between the two genotypes at 72H post electroporation as were virus titers.

Mutations at the c- bulge region were introduced into a genome with the Stop 3 mutation (Figure 4.25). We have previously shown the Stop 3,4 mutations in J6-JFH to drastically reduce replication to 16% of wild type, and hypothesize the defect to be due to disrupting the RNA structure at the top of SLVI (Figure 4.11). To test the c-bulge region of SLVI in genotype 2a, we combined the Stop 3 mutation (473A) and with mutations 481G and 484T, located downstream from the Stop 4 mutations (479A and 481G). This genome, J6 Mid Right-B, was impaired for replication, although with some variation, at 37% and 5% of wild type in a two separate experiments (Figure 4.26). Virus infectivity was detected, but reduced to 8% and 4% of the wild type levels. These mutations are in the same location as the H-JFH Mid Right Silent genome, but change an adenine rather than a guanine at

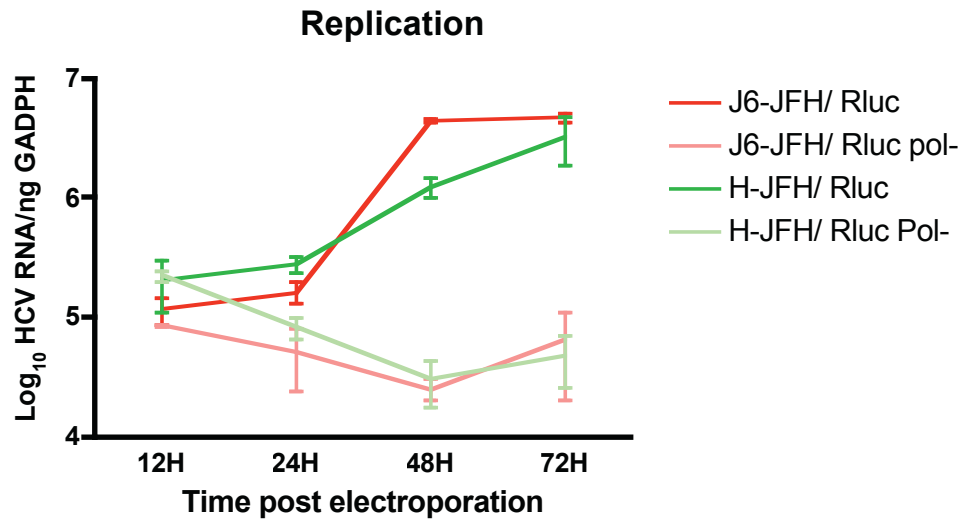


Figure 4.24. Comparison of replication kinetics of the genotype 1a-2a chimera reporter H-JFH/ Rluc and genotype 2a chimera reporter J6-JFH/ Rluc. Replication is indicted by HCV RNA levels measured by Real Time qRT-PCR.

484 (485 in H-JFH numbering) to a uridine due to the covariant base usage between genotypes. Recall that H-JFH Mid Right Silent also showed a dramatic reduction in replication, 6% of wild type (Figure 4.23).

We introduced mutations in the left side of the c-bulge region of SLVI. Silent mutations at 451 and 454 are on either side of the highly conserved C453 and in the same location as the H-JFH Mid Left genome (Figure 4.25) (see phylogenetic structure, Figure 3.6). Replication was also compromised, with some variability, with a range of being reduced 82-98% of wild type (Figure 4.26). Introduction of silent mutations in the downstream codon (449 and 451) in the genome J6 Mid Left-A, resulted in a minimal effect on replication (82% of wild type) and a modest reduction of virus infectivity (60% of wild type) (Figure 4.26). Mutations at the complementary side of SLVI, J6 Mid Right-A, also had a modest effect on replication, to 62 - 65% of wild type. However, virus infectivity was not detectable (Figure 4.26). The mutations in J6 Mid Right-A change the amino acid of an alanine to a glycine at codon 48. This could account for the elimination of virus infectivity, but also could be due to the reduced replication. This codon is conserved in all genotypes of HCV, except in the JFH-1 sequence where it is a threonine.

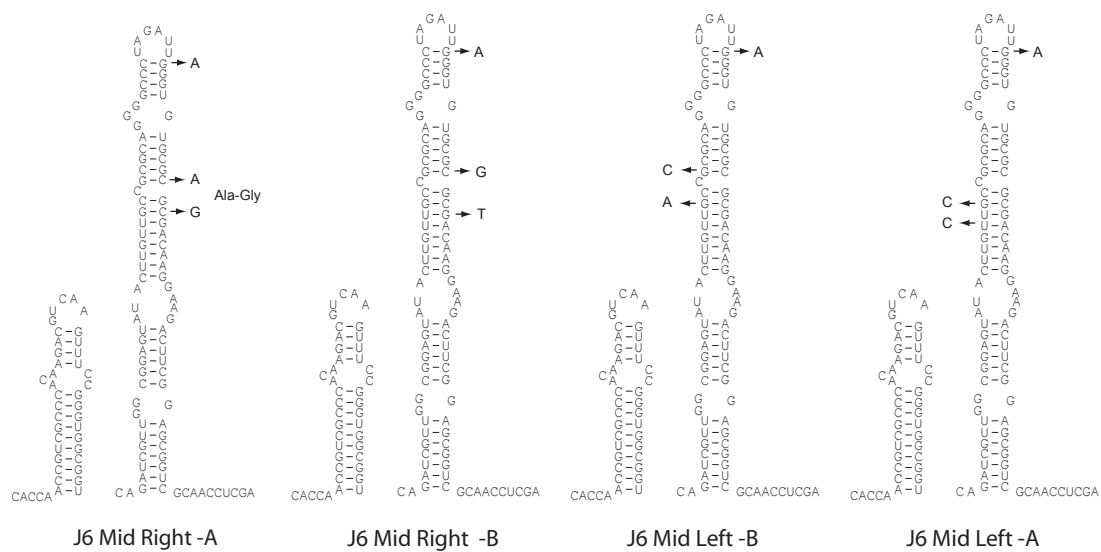


Figure 4.25. Mutations in the middle region of SLVI in genotype 2a J6 sequence. Mid Right-A changes an amino acid of the capsid protein, A48G.

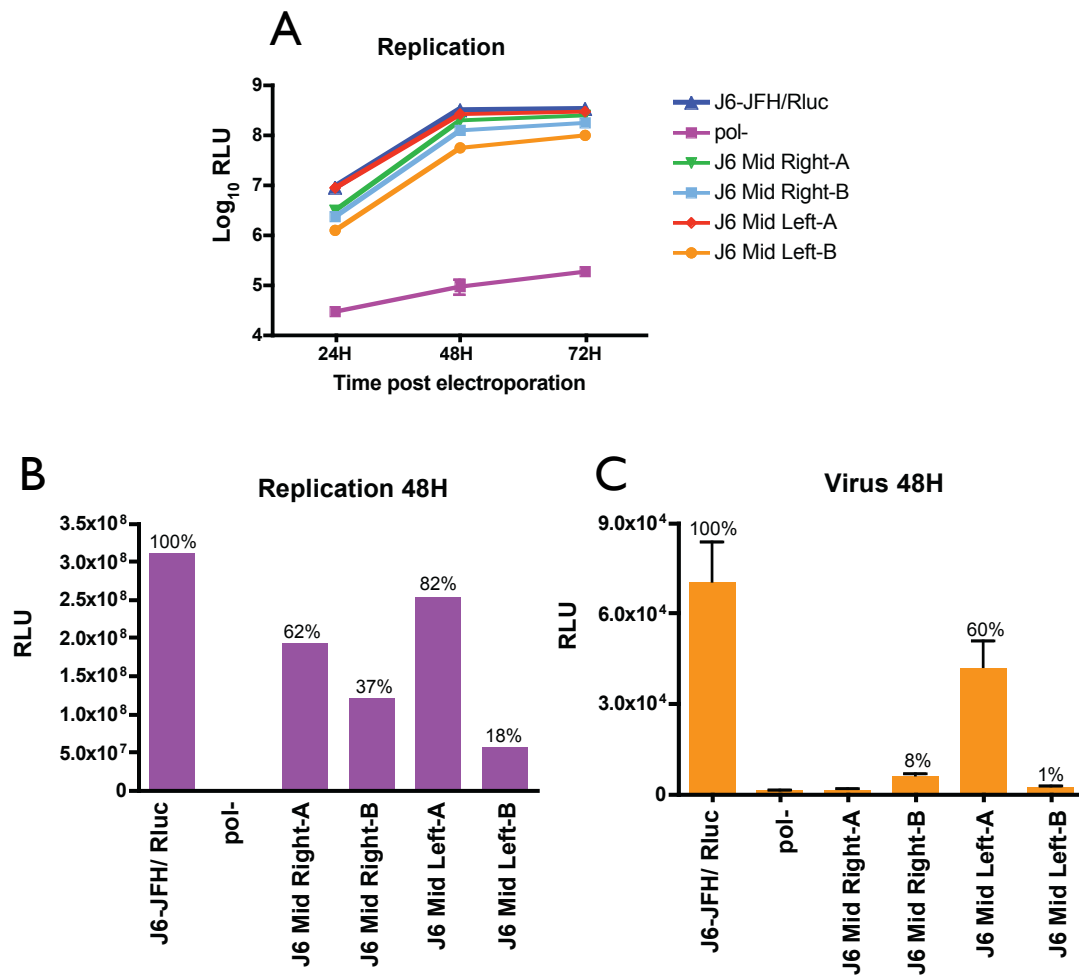


Figure 4.26. Mutations in the middle region of SLVI in genotype 2a J6 sequence reduce HCV replication and virus infectivity. A. Time course of luciferase expression. B. Replication at 48H post electroporation with percent differences noted above each bar. C. Virus infectivity 48H post infection with percent differences noted. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

To evaluate the impact of this amino acid difference at residue 48, we introduced a single nucleotide change at position 482 that disrupts one predicted base pairing of SLVI and changed the J6 amino acid alanine at codon 48 to the JFH-1 threonine (Figure 4.27). This genome, J6 > JFH, replicated efficiently, 82% of wild type, but virus infectivity was compromised to 45% of wild type (Figure 4.28). This result indicates that the amino acid at position 48 does contribute to virus production and/ or infection.

***Evolution of capsid protein mutants reveal a compensatory mutation at p7 and NS2.***

The genome with three mutations J6 Mid Right-A was capable of efficient replication, but failed to produce infectious virus in cell culture (Figure 4.26). In order to understand this defect, we subjected the virus genome of J6 Mid Right to continuous replication while monitoring the sequence for the selection of reversions or pseudo-reversions that enable virus production. It has been well established that the HCV polymerase has an estimated mutation rate of  $10^{-3}$  to  $10^{-4}$  errors per generation<sup>128</sup>. With a genome size of 9700 nucleotides, approximately one error will be incorporated in each genome generated. In addition, the T7 polymerase





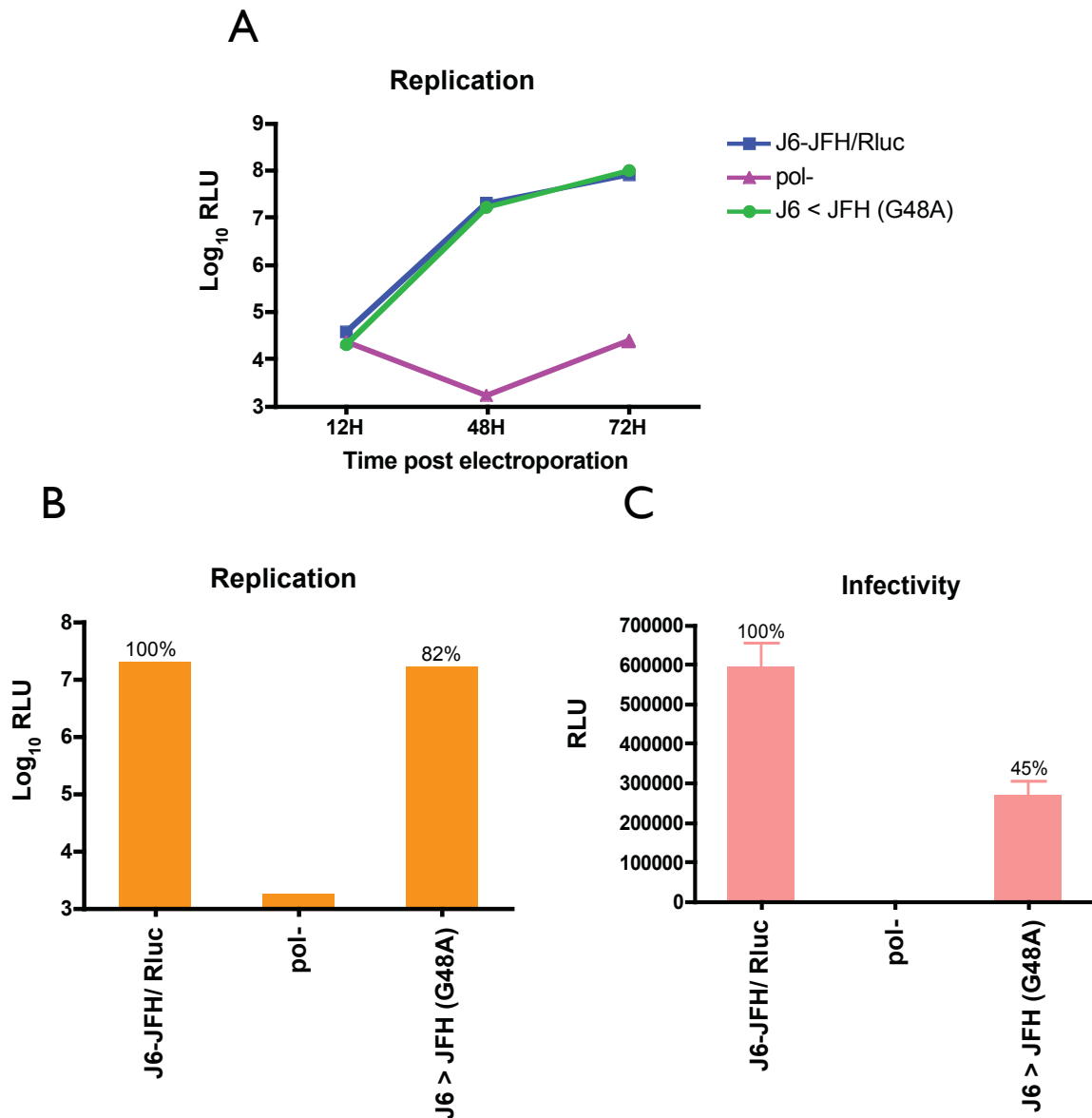


Figure 4.28. Mutation differences between J6 and JFH in SLVI slightly reduce HCV replication and moderately reduce virus infectivity. A. Time course of luciferase expression. B. Replication at 48H post electroporation with percent differences noted above each bar. C. Virus infectivity 72H post infection with percent differences noted. Each genome was measured from duplicate wells of two transfected RNAs. Averages are shown.

generates mutations with equal frequency during *in vitro* transcription, the source of HCV RNA in our replication studies. To discover adaptive mutations that can overcome the block in virus production caused by the mutations of J6 Mid Right-A, we electroporated *in vitro* transcribed RNA into Huh-7.5 cells and allowed the error prone replication to create a genomic diversity. Two preparations of each genome were tested in parallel. HCV replication diminishes as cell growth slows; therefore cells were seeded at low density in 100mm dishes and passaged as the monolayers approached 80% confluency. Replication was monitored by seeding cells in a separate 24 well plates and staining for NS5A expression by IHC (Figure 4.29). Initially, 80% of the cell monolayers transfected with the parental J6-JFH genome were positive for NS5A, while 50% of cells stained positive after transfection with the J6 Mid Right-A genome (Figure 4.29). Cells transfected with the J6-JFH genome maintained an approximately 80% positive culture while the percentage of J6 Mid Right-A positive cells was reduced after the first split. It has been observed that cells harboring HCV replication grow at a slower rate than naïve cells. Since the J6-JFH genome is capable of producing virus, the faster growing naïve cells become infected and thus maintain an HCV positive culture. Cells transfected with the virus-impaired Mid Right-A genome were out competed by the faster growing

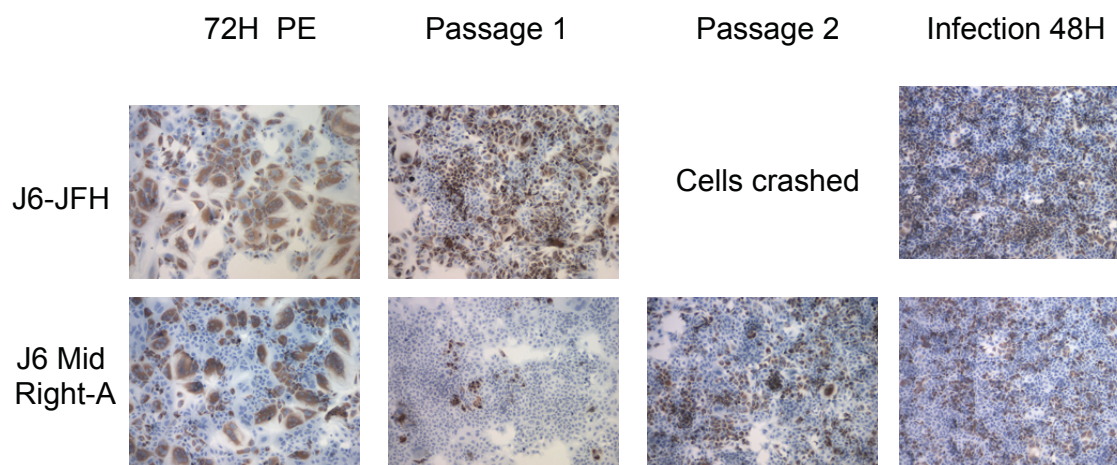


Figure 4.29. Selection for reversions increasing virus infectivity.

Transfected cells were passaged upon reaching 80% confluency and monitored for HCV by immunohistochemistry staining for NS5A. After 2 passages, cells with the wild type crashed, but cells transfected with the mutant showed an increase in the number of HCV positive cells. Infection of naive cells with media from the second passage revealed virus was now being produced by the mutant.

naïve cells and the percentage of positive cells was reduced to approximately 5% in the monolayer (Figure 4.29). Within three days after the first split, cells harboring wild type J6-JFH became cytopathic and could no longer be cultured, but media was cleared of cell debris and used to measure virus. The culture of cells transfected with J6 Mid Right-A continued to grow without appreciable cell death and after the second split an increase in the number of positive cells was apparent, to about 30% of the population, is detected (Figure 4.29).

To determine if virus was being released, the media was removed, cleared of any debris, and transferred onto naïve cells for two days. Staining for NS5A revealed virus-infected cells in both the wild type and the J6 Mid Right-A genome after the second split. HCV RNA was isolated from both the transfected and infected cells and sequenced. From the J6 Mid Right-A cell population, no change was evident in the capsid gene, including the mutations in SLVI, but sequence analysis of the entire genome revealed an amino acid change in the p7 gene. This change was a transition from a thymidine to a cytidine at nucleotide 2667 in the polyprotein, which resulted in a phenylalanine to a serine at amino acid 776 in the polyprotein, the 26th residue of p7. One of the two cultures launched from independent RNA preparations contained a mixed population of both wild type p7 and the

776S change while the second culture contained only the 776S change. Media was passaged a second time for the J6 Mid Right-A culture onto naïve cells, again, HCV RNA was isolated and sequenced. Population sequencing of both J6 Mid Right-A cultures again found the p7 776S mutation; however one of the cultures contained an additional change in NS2. This mutation in NS2 involved a transition of a cytidine to thymidine at nucleotide 2937 and recoded a proline to leucine at the 53<sup>rd</sup> residue of NS2, which is amino acid 866 in the polyprotein. Both cultures were passaged a third time and HCV RNA sequenced. One culture retained the 776S mutation in p7 and no other change in the polyprotein was detected. By the third passage, the second culture had completely reverted the 776S mutation in p7 back to the wild type codon and now only harbored the 866L mutation in NS2. The location of the pseudo-reversions occurred in very conserved regions for both p7 and NS2 (Figure 4.30 and 4.31). The F776 is conserved in 4 out of 6 genotypes while the proline at 866 in NS2 is conserved in the reference sequences for all 6 genotypes. No change was ever found in the capsid gene at any time during passage of either culture. Also, there were no changes detected at the population level in the wild type J6-JFH cultures.

Continual replication of a virus-defective genome, apparently allowed amino acid changes in p7 and NS2 to become selected during passage. In

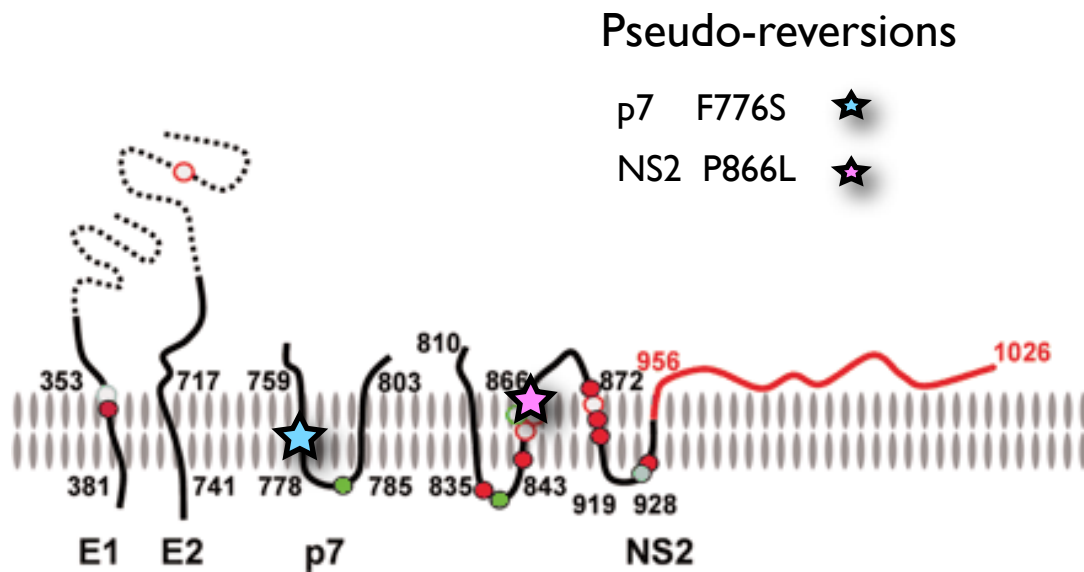
	10	20	30
1a.JOC10-C12B	A L E N L V L L N A A S L A G T H G L	V S S L V F F C F A W	
1a.NR1/W12/C6	A L E N L V I L N A A S L A G A H G F	V S F L V F F C F A W	
1a.H77	A L E N L V I L N A A S L A G T H G L	V S F L V F F C F A W	
1b.HCV-S	A L E N L A V L N A A S V A G E H G V	L S F L M F F C A A W	
1b.AD22c06	A L E N L V V L N A A S V A G T H G I	L S F L V F F C A A W	
1b.274933RU	A L E N L V V L N A A S V A G A H G I	L S F L V F F C A A W	
1c.HC-G9	A L E N L I V L N A A S L V G T H G I	V P F F I F F C A A W	
1c.AY051292	A L E N L I V L N A A S L A G T H G I	V P F F I F F C A A W	
2a.HC-J6	A L E K L V V L H A A S A A S C N G F	L Y F V I F F V A A W	
2a.JCH-6	A L E K L V V L H A A S T A S C N G F	L Y F V I F F V A A W	
2a.G2AK1	A X E K L V V L H A A S A A S C N G F	V X F V I F F V A A W	
2b.JPUT971017	A L E K L I I L H S A A S A A S A N G P	L V F F I F F T A A W	
2b.MD2B-1.	A L E K L I I L H S A A S A A S A N G P	L V F F I F F V A A W	
2b.Q2B	A L E K L I I L H S A A S A A S A N G P	L V F F I F F V A A W	
2c.BEBE1	A L E K L V I L H A A S A A S S N G L	L Y F I L F F V A A W	
2i.D54	A L E K L V V L H A A S A A S S H G L	L Y F I L F F V A A W	
2k.VAT96	A L E K L V I L H A A S A A S S H G M	L C F I I F F I A A W	
3a.HCVCENS1	A L E N L V T L K A V A A A G T H G I	G V Y L V A F C A A W	
3a.CB	A L E N L V T L N A V A A A G T H G I	G V Y L V A F C A A W	
3a.NZL1	A L E N L V T L N A V A A A G T H G I	G V Y L V A F C A A W	
3b.HCV-Tr	A M E N L V M L N A L S A A G Q Q G Y	V V Y L V A F C A A W	
3k.JK049	A L E N L I V L N A I S A A G T H G I	W V S L V A F C V A W	
4a.ED43	A L S N L I N I N A A S A A G A Q G F	W Y A I L F I C I V W	
5a.EUH1480	T C K N V I V L N A A A A A G N H G F	F W G L L V V C L A W	
5a.SA13	A L E N V I V L N A A A A A G T H G F	F W G L L V I C F A W	
6a.HK.6a33	A V E R L V V L N A A S A A G T A G W	W V A V L F L C C V W	
6a.HK.EUHK2	A V E R L V V L N A A S A A G T A G W	W V A V L F L C C V W	
6b.TH.TH580	A L E R L V V L N A A S A A G T A G W	C W T L I F L C C V W	
6d.VN.VN235	A L E N L I V L N A A S A A S Q G W	I Y C L V F I C C A W	
6g.ID.JK046	A L E N L I V L N A A S A A A T R G W	E C F L L F M C W A W	
6h.VN.VN004	A L E N V I V L N A A S A A S C Q G L	L W G L I F I C C A W	
6k.CN.KM45	A L E N L I I L N A A S A A S N Q G W	V C G I A F I C A A W	
6k.VN.VN405	A L E N L I V L N A T S A A G S Q G W	V V G V V F I C A A W	
	A L E N L V V L N A A S A A G : : G	: : F L : F F C A A W	

Figure 4.30. The amino acid in p7 involved in a pseudo-reversion restoring virus infection in an HCV genome with the mutation A48T in the capsid protein is conserved in 4 out of 6 genotypes. The p7 pseudo-reversion changes P866L. An alignment created by ClustalX of HCV reference sequences obtained from the HCV database (Los Alamos).

				40				50		60																				
1a.JP.HC-J1	C	I	S	W	C	L	V	V	L	Q	Y	F	L	T	R	V	E	A	Q	L	H	V	V	V	P	P	L	N	V	R
1a.H77	Y	I	S	W	C	M	V	V	L	Q	Y	F	L	T	R	V	E	A	Q	L	H	V	V	V	P	P	L	N	V	R
1a.LTD1	Y	I	S	W	C	L	V	V	L	Q	Y	F	L	T	R	V	E	A	Q	L	H	V	V	I	P	P	L	N	V	R
1b.CN.HC-C	F	L	A	R	L	I	V	V	L	Q	Y	F	I	T	R	A	E	A	H	L	Q	V	V	I	P	P	L	N	V	R
1b.JP.HCV-J4	F	L	A	R	L	I	V	V	L	Q	Y	F	I	T	R	A	E	A	H	L	Q	V	V	I	P	P	L	N	V	R
1b.US.11v07c07	F	L	A	K	L	I	V	V	L	Q	Y	F	I	T	R	A	E	A	Q	L	Q	V	V	I	P	P	L	N	V	R
1c.ID.HC-G9	Y	M	A	R	G	I	V	V	L	Q	Y	M	L	T	R	A	E	A	L	L	Q	V	V	V	P	P	L	N	A	R
1c.IN.AY051292	Y	M	A	R	G	I	V	V	L	Q	Y	M	L	T	R	A	E	A	L	L	H	V	V	V	P	S	L	N	A	R
1c.IN.Khaja1	Y	M	A	R	G	I	V	V	L	Q	Y	M	L	T	R	A	E	A	L	L	Q	V	V	V	P	P	L	N	A	R
2a.JP.AY746460	L	L	S	R	F	L	V	V	L	C	Y	L	L	T	L	G	E	V	M	V	Q	E	V	V	P	P	M	Q	A	R
2a.G2AK3	L	L	S	R	C	L	V	V	L	C	Y	L	L	T	L	G	E	A	M	V	Q	E	V	A	P	S	M	Q	A	R
2a.NDM59	L	L	S	R	S	L	V	V	L	C	Y	L	L	T	L	G	E	A	M	I	Q	E	V	V	P	P	M	Q	A	R
2a.pJ6CF	L	L	S	R	F	L	V	V	L	C	Y	L	L	T	L	G	E	A	M	V	Q	E	V	A	P	P	M	Q	V	R
2b.MD2b10-2	L	L	S	R	S	V	V	V	L	S	Y	M	L	V	L	A	E	A	Q	I	Q	Q	V	V	P	P	L	E	V	R
2b.JPUT971017	L	L	S	R	S	V	V	V	L	S	Y	M	L	V	L	A	E	A	Q	V	Q	Q	V	V	P	P	L	E	A	R
2b.MD2B-1	L	L	S	R	S	V	V	V	L	S	Y	M	L	V	L	A	E	A	Q	I	Q	Q	V	V	P	P	L	E	A	R
2c.BEBE1	L	L	S	R	C	L	V	V	T	C	Y	M	L	V	L	A	E	A	L	I	Q	D	V	I	P	P	L	Q	A	R
2i.VN.D54	L	L	T	R	C	L	V	V	L	C	Y	M	L	V	L	A	E	A	Q	I	Q	Q	V	V	P	P	L	E	A	R
2k.MD.YAT96	L	L	C	R	S	L	V	V	L	S	Y	L	L	V	R	A	E	A	L	I	Q	D	V	V	P	P	W	Q	A	R
3a.DE.HCVCENS1	W	I	G	R	L	I	V	V	N	Q	Y	T	I	C	R	C	E	S	A	L	Q	V	V	V	P	P	L	L	A	R
3a.CB	W	I	G	R	L	I	V	V	N	Q	Y	T	I	C	R	C	E	S	A	L	Q	V	V	V	P	P	L	L	A	R
3a.NZL1	W	I	G	R	L	M	V	V	N	Q	Y	T	I	C	R	C	E	S	A	L	H	V	V	V	P	P	L	L	A	R
3b.JP.HCV-Tr.	W	I	G	R	L	I	V	V	N	Q	Y	A	I	C	R	G	E	A	A	L	Q	V	V	V	P	P	L	L	V	R
3k.ID.JK049	W	I	A	R	L	I	V	V	N	Q	Y	T	I	A	R	C	E	A	A	L	Q	I	V	V	P	P	L	L	A	R
4a.EG.ED43	W	L	A	R	G	L	V	V	I	Q	Y	F	I	A	R	T	E	A	V	L	H	V	Y	I	P	S	F	N	V	R
5a.GB.EUH1480	K	V	V	L	V	M	V	V	L	Q	Y	F	I	A	R	V	E	A	I	I	H	V	V	V	P	P	L	Q	V	K
5a.ZA.SA13	M	V	V	L	V	I	V	V	L	Q	Y	F	I	A	R	V	E	A	F	I	H	V	V	V	P	P	L	Q	V	R
6a.HK.6a33	L	L	T	C	F	I	V	V	N	Q	Y	F	L	A	R	A	E	A	M	V	H	E	V	V	P	D	L	R	V	R
6a.HK.EUHK2	L	L	N	C	F	I	V	V	N	Q	Y	F	L	A	R	A	E	A	M	V	H	E	V	V	P	D	L	R	V	R
6b.TH.TH580	L	L	T	H	A	I	V	V	N	Q	Y	M	L	A	R	A	E	A	M	I	H	D	V	V	P	D	L	R	V	R
6d.VN.VN235	M	L	V	H	F	L	V	V	N	Q	Y	F	I	A	R	S	E	A	L	I	Q	Q	V	V	P	S	L	R	V	R
6g.ID.JK046	L	L	A	H	F	L	V	V	T	Q	Y	F	I	A	H	I	E	A	K	L	H	V	V	V	P	F	L	R	V	R
6h.VN.VN004	L	L	I	S	M	L	V	V	I	Q	Y	F	I	A	L	T	E	A	Q	L	H	Q	V	V	P	S	L	L	V	R
6k.CN.KM45	L	L	I	S	T	L	V	V	I	Q	Y	F	I	A	R	A	E	A	I	L	H	V	V	V	P	P	L	Q	V	R
6k.VN.VN405	L	L	I	S	T	L	V	V	I	Q	Y	Y	I	A	R	A	E	A	M	L	Y	V	V	V	P	S	L	Q	V	R
	L	:	R	:	W	W	L	Q	Y				L	:	R	:	E	A		L	Q	V	W	V	P	P	L	:	V	R

Figure 4.31. The amino acid in NS2 involved in the second pseudo-reversion restoring virus infection in an HCV genome with a mutant capsid (A48T) is absolutely conserved in all genotypes. An alignment created by ClustalX of HCV reference sequences obtained from the HCV database (Los Alamos).



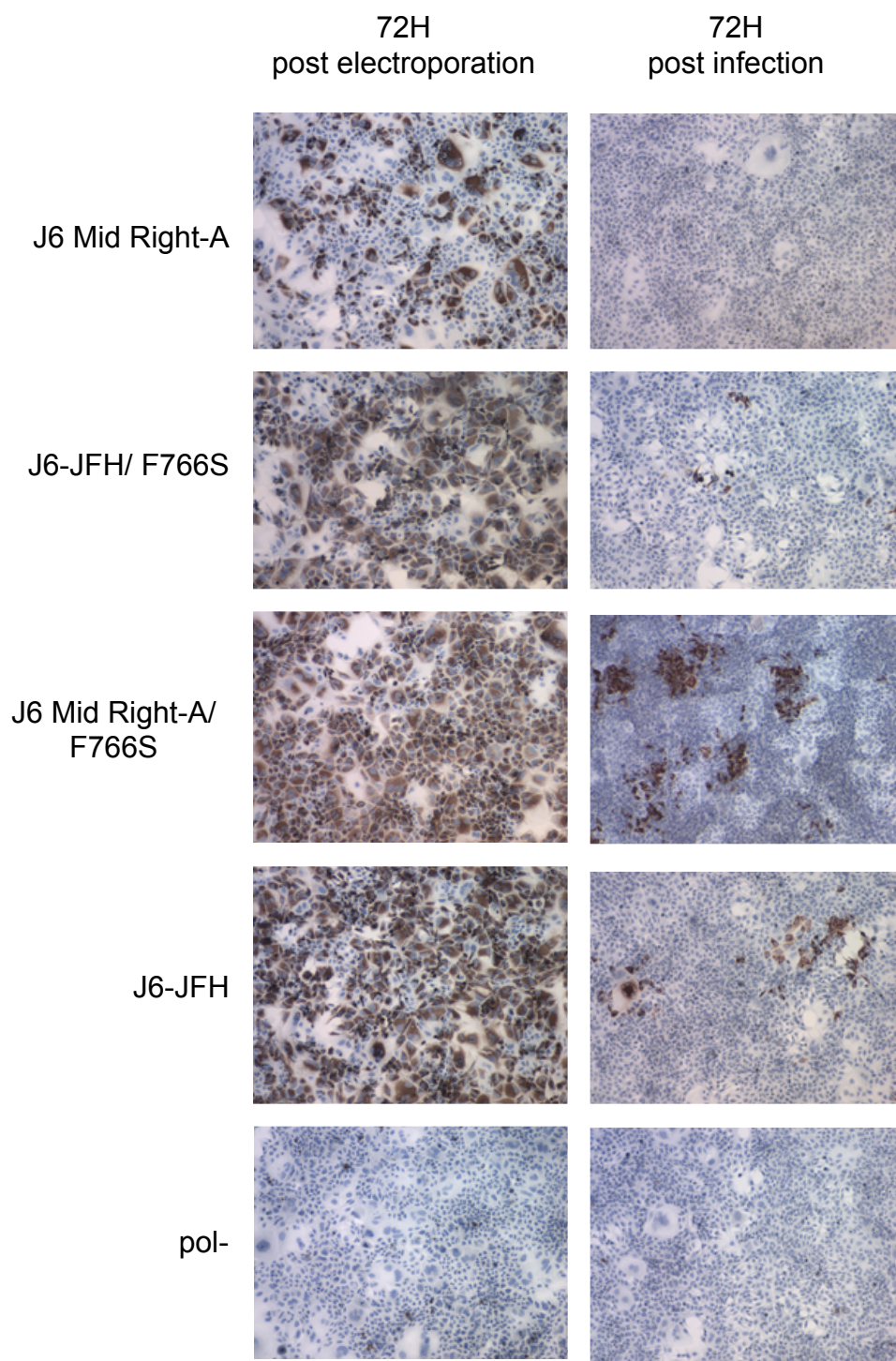


Adapted from Yi et al. 2007

Figure 4.32. Position of compensatory mutations found in p7 (F776S) and NS2 (P866L) that restore virus infectivity of the mutant J6 Mid Right-A containing an amino acid change in the capsid protein at residue 48. Compensatory mutations in p7 and NS2 found by Yi et al. 2007 are also shown.

order to confirm that these mutations were responsible restoring virus production, the 776S mutation for p7 was cloned into the wild type J6-JFH genome and the J6 Mid Right-A genome harboring the 48G capsid mutation. Replication of the J6-JFH/ 776S genome was comparable to the wild type J6-JFH with over 90% of cells positive for NS5A staining, whereas the genome J6 Mid Right-A (J6-JFH/ 48G) was only 50% positive. Combining both capsid and p7 mutations in J6-JFH/ 48G+776S enabled wild type level replication (Figure 4.33). To prove that the p7 mutation was selected as a compensatory mutation that repaired the virus defect conveyed by the capsid mutation of 48G, we measured virus infectivity by IHC. Using two independent RNA preparations for each construct, media was collected from transfected cells after 96H, transferred onto naïve cells, and stained for NS5A expression. The J6 Mid Right-A genome with mutation 48G failed to yield infectious virus (Figure 4.33). The J6-JFH/ 776S genome with the p7 mutation was infectious, but with few cells in two infection foci. In contrast, the genome J6-JFH/ 48G +776S was capable of virus infection and infected larger number of cells with multiple foci (Figure 4.33). This confirmed that the p7 mutation at 776S was able to rescue the virus defect contributed by a mutation in the capsid protein. We were thus able to identify compensatory

Figure 4.33. (next page) The compensatory mutation in p7, F776S, can restore virus infectivity in a mutant genome with an amino acid change in the capsid protein at residue 48. Introduction of the amino acid changes found in p7 into the molecular clones J6-JFH and J6-JFH/ Mid Right-A confirm that sequence changes found during the reversion analysis can improve virus infectivity and imply an association between residue 48 of the capsid protein and 766 of p7. IHC of cells either transfected with HCV genomes after 72H or naive cells infected with media from 72H transfected cells transfected after 3 days.



mutations for virus infection by increasing the pool of diverse genomes and selecting for mutations that restored virus infection.

***Evolution of RNA structure mutations in Huh-7.5 cells mimic reversions in vivo.***

Given the successful selection of reversions that restored virus production in a genome with an amino acid mutation, we wanted to test if reversions that improve virus replication could be selected in cell culture. We have shown that mutations for Stop 1,2,3,4 severely compromise HCV replication by disrupting the RNA structure SLVI. During the infection of the chimpanzee inoculated with the Stop 1,2,3,4 mutant, reversions were selected that repaired the disrupted SLVI. Given the high mutation rate of HCV replication, a spectrum of diverse genomes is likely to be generated, but we did not know if reversions in SLVI would be selected in cell culture and how frequently reversions could arise when replication was drastically impaired. There appeared to be little cytopathic effect on cells harboring HCV genomes with attenuated replication, and therefore our strategy was to launch replication with a large amount of HCV RNA and monitor replication while passaging cells in culture. Replication is highly impaired by the Stop 1,2,3,4 mutations and since the generation of diversity depends on the rate of

replication, either adaptive mutations would be selected, improving replication, and the number of HCV positive cells increase, or HCV RNA would be lost from the population. The SLVI mutations were tested in the parent H-JFH genome that has a more robust replication and allows virus production. We have shown that mutations in SLVI initially affect replication, but that virus infection is possible if the defect on replication is not too severe. Using a genome that can produce virus, if replication occurs at a sustainable rate, virus spread and an increase in the number of cells harboring HCV replication should occur — increasing the likelihood of propagating a fitter genome as an adaptive mutation is selected.

To improve our chances of selecting an adaptive genome, we followed replication of two mutant genomes: H-JFH/ Stop 1,2,3,4 conferring impaired replication levels of 2% of wild type, and H-JFH/ Stop 3,4 which is moderately reduced to 16%. At four days post electroporation, over 90% of the population was positive for HCV antigen in the parental H-JFH/ V+T genome (Figure 4.34), but the mutant genomes were both reduced to less than 10% of the population. Cells were allowed to grow and passaged upon confluency. By day 5, the culture of the wild type genome displayed significant cytopathic effect, due to efficient replication and virus

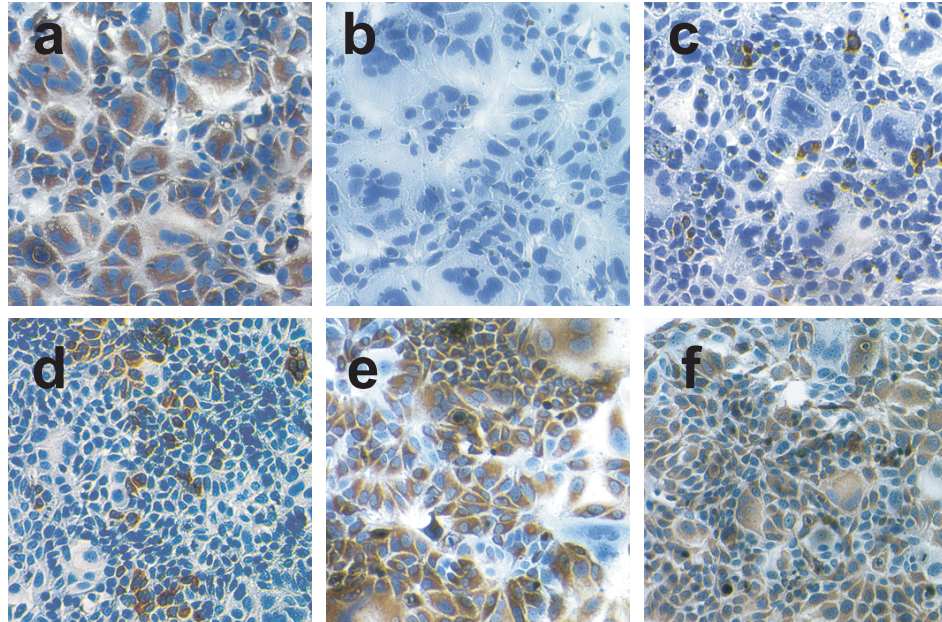


Figure 4.34. Selection of reversions conferring an increase in replication of mutant H-JFH/ Stop 1,2,3,4 in Huh-7.5 cells. (a–c) IHC staining of NS5A showing replication of H-JFH (a), a pol- control (b), and H-JFH/ Stop 1,2,3,4 (c) at day 2 post electroporation. (d–f) Cells harboring H-JFH/ Stop 1,2,3,4 were passaged and stained for NS5A (brown) and nuclei (blue) at days 7 (d), 21 (e), and 41 (f).

production, and could no longer be passaged. The cultures with the mutant genomes remained healthy.

Initially, we sequenced HCV RNA at the population level. For the culture transfected with H-JFH/ Stop 1,2,3,4, there was no change in the capsid gene or the rest of the genome for the first 7 days post electroporation (Figure 4.35). By day 11, the chromatogram revealed a mixed population at 473, the site of the Stop 3 mutation. The majority of the sequences contained the input mutation of adenine at 473, yet a minor population could also be detected with a reversion to the original H77 nucleotide of guanine (Figure 4.35). No other sequence changes could be detected at the population level. The appearance of the reversion corresponded with an increase in replication as the number of NS5A positive cells increased (Figure 4.34). Sequence analysis at day 15 again showed a mixed population at the site of Stop 3, with the input 473A and revertant 473G, and another change in the population at 480, which was the first base of the Stop 4 mutations. This second change, A480C, was also a reversion to the H77 sequence and was present as a minor species in the population (Figure 4.35). By day 21, the number of HCV positive cells had increased to over 60% indicating a marked improvement in replication and/or virus infection. Sequence analysis revealed an equal mixture of both reversions, 473G and 480C, present in the



**A**

	Stop 1	Stop 2	Stop 3	Stop 4
Stop 1,2,3,4	GGACGU <u>AA</u> AGUUC	GGGUGCGGUCAGAU <u>CGU</u> AGGUGGAGUUUACUUGUUGCCGCGCAGGGGCCCUAGAU <u>U</u> AGGUGUG <u>Ag</u> GCGG		
				C
				G
H77	GGACGU <u>CA</u> AGUUC	GGGUGCGGUCAGAU <u>CGU</u> UGGUGGAGUUUACUUGUUGCCGCGCAGGGGCCCUAGAU <u>U</u> CGGUGUG <u>Cg</u> CgCGG		
Position	407	434	473	480 482

**B**

Day post electroporation	Reversion			
	A473G	A480C	A473G + A480C	none
Day 1-7				27/27
Day 10	1/6	1/6		4/6
Day 11	1/8	1/8		6/8
Day 15	1/10	3/10		6/10
Day 21	3/13	9/13		1/13
Day 23	9/29	19/29	1/29	
Day 38	3/9	6/9		
Day 41	1/10	9/10		
Day 50	1/10	9/10		

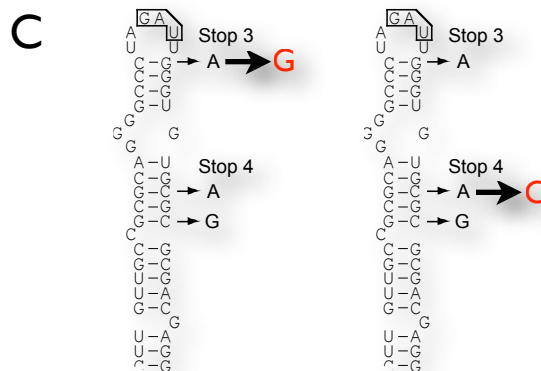


Figure 4.35. Sequence analysis of HCV RNA during passage of cells transfected with H-JFH/ Stop 1,2,3,4. A. Sequence changes found in the passaged cells. B. The frequency of reversions found at the times indicted. C. The reversions would restore the RNA structure of the top of SLVI.

population. To measure the frequency of revertants in the population, we sequenced individual clones. This also enabled us to detect transient changes that may be present as minor species and therefore missed in the bulk sequence analysis. Initially, we only detected the input sequence containing the mutations Stop 1,2,3,4. In approximately half of the clones, we detected additional mutations in the amplicon region of the 5'NTR and capsid gene. These changes were not concentrated in one area and appeared to be transient changes in the population (Figure 4.35). By day 10, two out of six clones contained a reversion. One clone contained the A473G change and the other contained the A480C reversion indicating that these reversions were not on the same molecule, but arose independently. The input sequence was still the dominant genome of the population. By the third week, one-third of the clones contained the 480C reversion while only one-tenth had the 473G reversion. The population shifted where the majority of clones possess a revertant, either 473G or 480C by week 4. In the following weeks, the revertant 480C steadily became the dominant genome. Replication showed a marked improvement as over 90% of the monolayer was positive for HCV antigen (Figure 4.34).

Following replication launched with the H-JFH/ Stop 3,4 genome showed the identical revertants emerging as for the genome with Stop

1,2,3,4, but with a different pattern. There was no sequence change detected in the population for two weeks post electroporation (Figure 4.36). At day 14, all clones sequenced contained the reversion of Stop 3 to the H77 sequence, 473G. This sequence change of the population corresponded with an increase in the number of HCV positive cells (Figure 4.36). During the following week, the second reversion of 480C was selected. This reversion was present on the same genome containing the reversion 473G. The frequency of genomes containing this second reversion remained only one out of ten clones for all time points after day 21, suggesting that a genome with Stop 1,2,3 and mutation A482 is more fit than the genome with Stop 1,2,4. Indeed the mutant Stop 1,2,3 can replicate 58% more efficiently than mutant Stop 1,2,3,4 in both the H/L+I parent as well as in the H-JFH/V+T parent (Figure 4.2).

HCV genomes with Stop 1,2,3,4 and Stop 3,4 mutations were both capable of at least low levels of persistent replication in Huh-7.5 cells. Reversions to the original H77 sequence were selected at the mutations of Stop 3 and 480 of Stop 4, indicating pressure to maintain a functional RNA structure of SLVI. The reversions selected in Huh-7.5 cells were identical to the reversions selected in the chimpanzee infection, demonstrating similar selective pressures in both the *in vivo* and cell culture systems. We conclude

A

	Stop 3,4	Stop 3	Stop 4	clones
Day 1-12	GGACGUCAAGUUC	CGGGUGGCGGUCAGAU	CGU	
Day 14-21			G	10/10
Day 23-41			G	9/10
			G	1/10
H77	GGACGUCAAGUUC	CGGGUGGCGGUCAGAU	CGU	
Position		473	480 482	

B

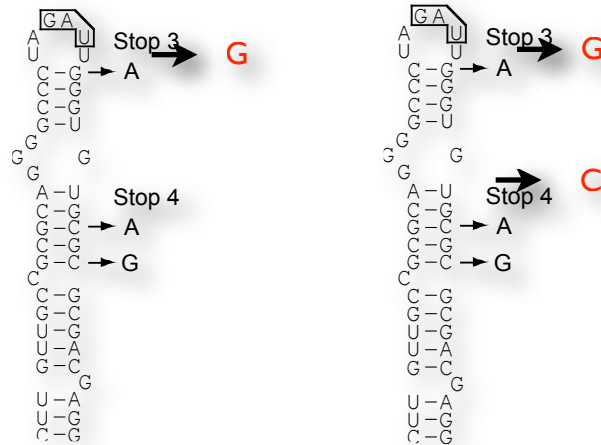


Figure 4.36. Reversions found after passaging Huh-7.5 cells transfected with H-JFH/ Stop 3,4. HCV RNA was extracted and individual clones sequenced. A. Sequence changes found at times post electroporation and frequency of reversions detected in the cloned sequences. B. Reversions would restore the top of SVI.

that the genetic analysis defining the structure of SLVI in HCV replication in Huh-7.5 cells also describes the function of SLVI in a natural infection.

### ***Discussion.***

#### ***The Stop mutations convey a replication defect in Huh-7.5 cells.***

A genome with mutations Stop 1,2,3,4 led to an attenuated infection in a chimpanzee. Within 2 weeks of the infection, a reversion at Stop 3 was selected and fixed in the population. Additional reversions of Stop 4 were soon selected and fixed in the circulating virus population. The order of the reversions and the speed of their selection indicate pressure to restore an RNA structure as opposed to open the ARF. As shown in the phylogenetic analysis, the mutations for Stop 3 and 4 disrupt three base pair interactions (Figure 3.6). The reversions restore the RNA structure for SLVI and indicate a role of this structure in the virus lifecycle.

We found that HCV genomes harboring the Stop mutations had a reduced ability to replicate in Huh-7.5 cells. The genome with Stop 1,2,3,4 displayed a severe replication defect, with RNA levels greater than 2 logs below wild type. Comparing replication for genomes with different combinations of Stop mutations, we found that mutations for Stop 3 and Stop 4 conveyed the decrease in replication. Individually, the mutations

slightly reduced replication, but together a synergistic decrease in replication was observed. Genomes with mutations for Stop 1 and 2 showed no defect in replication; however, the genomes displaying the most impaired ability to replicate contain Stop 1,2,3,4. Even though the mutations of Stop 3 and 4 disrupt the top of SLVI as predicted by the secondary RNA structure, tertiary interactions could also be disrupted and account for the additional replication defect with the inclusion of Stop 1 and 2. The defect in replication contributed by the Stop 1,2,3,4 mutations could be responsible for the selective pressure during the chimpanzee infection. Mutants of Stop 1,2,4, mimicking the first reversion in the chimpanzee, showed an improvement of 205-fold compared to the Stop 1,2,3,4 genome. Reversions replaced the Stop 3 mutations with the H77 sequence and restored the top stem structure of SLVI. Since replication was reduced in the presence of the Stop 3 and 4 mutations, the reversions occurring in the natural infection indicate selection for a functional RNA element *in vivo*.

***Disruption of an RNA element impairs HCV replication.***

By introducing mutations that allow both reading frames to be expressed, we confirmed that the Stop 3 and 4 mutations affect replication by disrupting an RNA structure rather than blocking expression of an ARF

protein. These Open mutations also conveyed a replication defect when the helical stem of SLVI was disrupted. Since the Open mutations also introduced amino acid differences in the ARF protein, although silent for the capsid protein, the question if the ARF protein affects replication remained. We ultimately demonstrate that replication was affected by an RNA structure by introducing compensatory mutations in the genome with Stop 1,2,3,4 that restore the structure of SLVI. Compensatory mutations restoring the base pairings of Stop 3 and 4 improved replication by 60%. The restoration of the top helix of SLVI and consequent improvement of replication strongly argues for a functional RNA element and not the use of an ARF protein in HCV replication in Huh-7.5 cells. Enzymatic structure probing confirmed that mutations for Stop 3 and 4 disrupt the helix base pair for an RNA structure.

***SLVI is conserved amongst diverse genotypes of HCV.***

The importance of SLVI in diverse genotypes was demonstrated by comparing replication of genomes with Stop mutations in genotypes 1a, 1b, and 2a. Similar effects on replication were observed for each genotype. Mutations for Stop 1,2,3,4 reduced replication to below 8% of wild type levels and mutations for Stop 3 and 4 conveyed a replication decrease to 17-

20% of wild type for all three HCV genotypes. This indicates that SLVI is a functional RNA element in sequences as diverse as 19% at the nucleotide level.

***Different regions of SLVI have varying importance in HCV replication.***

To define the regions of SLVI that affect replication, we used a luciferase reporter genome that expressed the Renilla luciferase gene and the protease 2A from FMDV between the p7 and NS2 coding regions. The reporter was introduced into a chimeric genome of H-JFH allowing early measurements of translation, replication and virus infectivity. Point mutations in the loops of SLV and SLVI had a moderate effect on HCV replication; however mutations in the loop of SLVI dramatically reduced virus infectivity by 75%. The mutations did not alter the amino acid sequence of the capsid protein, thus implying the loop of SLVI has a role in virus assembly or infection.

Mutations along the basal region of SLVI contributed to a minor reduction in RNA replication. When combined with mutations at the apical stem (Stop 3 and 4 mutations), however, an additional decrease in replication was observed. This implies the overall structure of SLVI influences replication, but it is the integrity of the top stem structure that is



critical. Mutations for Base Right in the base of SLVI, had a more significant impact on virus infectivity than on replication. These mutations were introduced to compensate for the Base Left mutation set, therefore restricting the position for the mutations. The Base Right mutations introduced an amino acid change into the capsid protein. This change was later found to be important for virus infectivity as it could be trans complemented by expression of the capsid, E1, E2, p7 and NS2 proteins (See chapter 5).

Mutations in the middle region of SLVI had an effect on replication and indicated overlapping features of replication and virus infectivity. In the H-JFH/Rluc genome, mutations introduced along either side of the cytidine bulge of position 453, decreased replication to approximately 30% of wild type. By combining the left and right side mutations, replication was restored to 82% of wild type, by re-establishing the base pairing of SLVI. The mutations for Mid Right did alter the capsid protein and dramatically reduced virus infectivity. This mutant genome was also shown to be trans complemented for virus production by expression of the virus structural proteins and NS2 (Chapter 5). We found a critical base in the structure of SLVI that affects replication. The mutations for Mid Right Silent shared one substitution with Mid Right, 482A but also contained a mutation at 485T.

There was no amino acid change in the capsid protein. The genome Mid Right Silent was severely impaired for replication, to 6% of wild type whereas the similar mutant genome of Mid Right was only reduced to 34% of wild type. This difference in replication implied that the structure of SLVI may be considerably altered by the mutation at position 485 and may also indicate the existence of additional tertiary conformations not yet predicted.

We also compared mutations of the middle region of SLVI in the J6-JFH/Rluc parental genome. Again, we found mutations of position 485 to contribute to a more severe reduction in replication than mutations of the adjacent positions. A comparable mutant to Mid Right in the H-JFH genome, Mid Right-A, was found to slightly reduce replication, but had a severe impact on virus infectivity. Mid Right-A involves three nucleotide substitutions at the top and middle region of SLVI and an amino acid change at residue 48. This genome could also be transcomplemented for virus infectivity by expression of the proteins capsid, E1, E2, p7 and NS2 (Chapter 5). In addition, this genome was used in an analysis to detect reversions in Huh-7.5 cells.

The mutations for Mid Right-A changed the alanine at position 48 to a glycine. This position varies between J6 and JFH. In J6, as in all other six genotypes, residue 48 encodes an alanine, but in JFH, residue 48 is a

threonine. A genome with one base substitution conferring the JFH amino acid in the J6-JFH/Rluc parent only slightly reduced replication, but significantly reduced virus infectivity. This result describes another capsid amino acid critical for virus infectivity.

***Passage of a genome with a mutation in the capsid protein selects for second-site reversions in p7 and NS2.***

The mutant genome Mid Right-A displayed a minor reduction in replication, but was significantly impaired in virus infectivity. After passage of Huh-7.5 cells transfected with this mutant RNA, we detected reversions in p7 and NS2 that coincided with an increase in virus infectivity. The first reversion was found in p7 and changed a phenylalanine to a serine at residue 776 of the polyprotein. The F776 codon is not specific for genotype 2a, but is conserved in all 6 genotypes. This residue is predicted to lie in the first transmembrane helix of p7. Introduction of the F776S pseudoreversion alone into the parental genome did not alter replication or virus infectivity, but when combined with the mutation A48G in the capsid protein, increased virus infectivity. A second pseudoreversion was selected after, and eventually replaced, the F776S mutation in the population. The second pseudoreversion is located within the first transmembrane helix of NS2 and

changes a proline to a leucine at residue 866. In a study investigating virus determinants for an H-JFH chimera, Yi et al. found similar compensatory mutations in p7 and NS2 <sup>164</sup>. Using an analogous approach of passaging cells transfected with replication competent chimeric genomes, virus infectivity was detected within 2 days, but increased by day 15. The compensatory mutations were found to also cluster in the transmembrane domains of p7 and NS2, and the authors concluded that these mutations increased the specific infectivity of the particles rather than increase the number of released particles. In this study, we found genetic evidence for an interaction between residue 48 of the capsid protein and p7 at 776 and NS2 at 866 to restore virus infectivity. The residue of 48 of the capsid protein has not been specifically studied, but lies in a domain known to be involved with RNA binding activity and homo-oligomerization <sup>86</sup>. Although not required for replication, p7 is essential for virus production. It has been shown to homo-oligomerize and form an ion channel akin to other virus viroporins <sup>100, 119</sup>. p7 displays genotype compatibility indicating necessary interactions with other virus proteins, but there has been no description of a p7-capsid interaction <sup>164</sup>. The role of an ion channel activity or another mechanism of p7 in virus assembly has yet to be defined. Less is known for a role of NS2 in virus assembly, but an increase in virus infectivity was mapped in Con1-JFH and

J6-JFH chimeras, that implicates an interaction between the structural proteins of capsid, E1, and E2 and p7 and NS2<sup>102</sup>. Shifting the junction of a J6-JFH chimera between the first and second predicted transmembrane segment of NS2 increased virus release and infectivity indicating sequence specificity for NS2. The second site reversions detected in p7 and NS2 could be a result of a general increase in fitness that compensates for the injury conveyed by the mutations in the capsid.

***Comparable selective pressures for the structure of SLVI are present in Huh-7.5 cells and in a natural infection.***

The mutations in SLV and SLVI of Stop 1,2,3,4 severely reduced HCV replication, but low levels of replication persisted in transfected Huh-7.5 cells. Reversions were identified after 7 days that restored the base pairing of SLVI and corresponded with an improvement of HCV replication in cell culture. Two reversions were found in the population, yet were not present on the same HCV genome. The reversion of 473G removed the Stop 3 mutation and was also the first reversion found in the chimpanzee infection. The reversion of 480C involved one out of two mutations for Stop 4 and was the second reversion selected during the chimpanzee infection. Both revertants arose simultaneously and were initially present at equal

frequencies. However the 480C revertant became the dominant genome in the population after 38 days in culture. This indicated that a genome with the Stop 1,2,3 mutations and G482 had an increase in replicative fitness over a genome with just Stop 1,2,4. This was indeed the case when replication was measured after electroporation where genomes with Stop 1,2,3 were compromised to just 76% of wild type and Stop 1,2,4 conferred a defect to 18% of wild type (Figure 4.2).

Continued replication of genomes with mutants Stop 1,2,3,4 selected for reversions at either Stop 3 or the mutation 480 of Stop 4 in Huh-7.5 cells. A similar reversion pattern was observed in the chimpanzee infection where reversions for both Stop 3 and Stop 4 were selected. This demonstrates that there are parallel selective pressures to restore the structure of SLVI in both a chimpanzee infection and in a cell culture system. Replication in cell culture was impaired by the Stop mutations, which disrupt the RNA structure of SLVI. Given both systems selected reversions of mutants located at the top of SLVI, it is likely that reversions were selected to improve replication fitness by repairing the base pair interactions of SLVI.

## **Chapter 5**

## **Chapter 5. The Role of RNA Structures in the Capsid Gene**

### ***Aim:***

We have shown that mutations in the RNA structures SLV and SLVI of the capsid gene reduce HCV replication, yet do not know what stage of replication is compromised. HCV, like many positive-stranded RNA viruses, conducts replication through interactions between virus RNA elements and host cell components. In this chapter, we demonstrate how mutations in SLVI affect translation and thus replication, and how virus proteins and cellular factors may participate.

### ***Introduction.***

Replication is dependent on the process and efficiency of translation, creation of membrane associated replicase complexes, and movement of the virus RNA between both of these sites. It is believed that translation of an HCV RNA molecule establishes a replicase complex that can only synthesize copies of the originally translated genome.

Transcomplementation of a heterologous RNA that did not undergo translation for the replicase complex cannot occur<sup>5 34</sup>. Nascent HCV RNA is released from the replicase complex and is translated to create new, independent replicase complexes that generate copies of their originating



genomes. This interdependence of translation and replication ensures that the fittest HCV genomes must possess elements that enable both of these processes to occur efficiently.

HCV translation requires a minimum sequence of bases 43-375 from the 5'NTR and capsid gene to form SLII, SLIII, and SLIV of the IRES<sup>37, 41</sup>. The stability of SLIV requires the downstream sequence and also an absence of RNA structure immediately downstream<sup>52</sup>. The secondary structure of SLIV, cofactors and contact points have been determined<sup>39, 97, 105</sup>. The 3'NTR is also thought to influence translation, potentially through interactions between proteins bound to RNA elements, but the exact mechanism of action is unknown and the IRES can conduct translation in the absence of the 3'NTR<sup>67, 135</sup>. Translation initiation was monitored using purified 40S ribosomal subunits and eIF3 and comparing the kinetics of cofactors binding to mutant IRES RNA<sup>58, 69, 98</sup>. The basal stem of SLIII<sub>d</sub> binds the 40S subunit and the apical SLIII<sub>b</sub> binds eIF3. The ternary complex of eIF2-GTP-initiator tRNA joins eIF3 to form 48S\*. In contrast to the canonical 48S complex, eIF4F does not bind the HCV IRES and this complex is therefore referred to as 48S\*. SLII induces a conformational change in the 40S subunit, GTP hydrolysis releases eIF2 and the 60S subunit joins to form the 80S complex. In all of the above mentioned translation

studies, none have included the RNA structures of SLV and SLVI, therefore it is unknown how these RNA elements may influence IRES formation, the initiation of HCV translation, and thus replication.

Because HCV infection is persistent *in vivo* and noncytopathic *in vitro*, it is hypothesized that the virus must control the level of virus protein and RNA in the infected cell. Both the capsid protein and sequence have been described as modulating translation. The capsid protein reportedly down regulates translation by binding the IRES. Several studies claim the interaction with the IRES can occur *in trans*, yet another report claims the protein must be expressed *in cis*<sup>17, 36, 124, 125, 140</sup>. Other studies suggest that it is the capsid gene sequence and RNA structures therein that decrease translation. The HCV capsid efficiently binds RNA, in a non-sequence specific fashion, and other virus capsid proteins, in addition to the HCV capsid protein, have been shown to bind the HCV IRES sequence<sup>121, 155</sup>. In two contradictory studies, an identical frameshift was introduced early in the capsid sequence to conclude that it is the RNA sequence, and perhaps RNA structures of SLV and SLVI, that reduce IRES translation or that it is the capsid protein that is expressed *in cis* that reduces translation<sup>124, 155</sup>. It should be noted that the frameshift would express the ARF protein. The specific elements of SLV and SLVI were never formally addressed in any of the

studies. Furthermore, the question of capsid protein or gene sequence acting as a modulating mechanism was never addressed in the context of HCV replication, with the exception of a study including the HCV IRES and capsid gene to drive translation of a chimeric poliovirus. In this context, the capsid gene was found to be essential for larger plaque size and titer<sup>81, 166, 167</sup>.

In this chapter, we defined when and how the RNA structures of the capsid gene affect HCV replication. Mutations in SLVI contribute to a defect in replication resulting in a lower steady-state level of HCV RNA compared to wild type genomes. The effects of mutations in SLVI are observed at a time when input RNA has decayed and initial replication activity accelerates. During this time, replicase complexes are formed, nascent HCV RNA synthesized, and translation of nascent HCV RNA establishes additional replicase complexes. We examined the possible effects of SLVI on HCV replication by measuring translation and assessing the effect of expression of virus proteins *in trans*, exploring the contribution from host cell factors, and probing the effects of position within the genome.

***Initial translation is affected by mutations in SLVI.***

To determine if translation was compromised by the mutations in SLVI, a construct was created that expresses the Renilla luciferase gene near the C terminus of NS5A. This position within NS5A was chosen based on another construct where a Flag epitope was randomly inserted into the replicase of genotype 1b, Con1 sequence with the adaptive mutations GIT<sup>91</sup>. With the Flag epitope, replication was not affected, but when it was replaced with the Renilla luciferase gene, replication no longer occurred. We could however use this construct as a tool to measure translation in the absence of replication as the luciferase gene was expressed from the input RNA and declining levels indicate RNA degradation in the cell (Figure 5.1). We introduced the mutations in the capsid gene that were shown to significantly impair HCV replication by altering the RNA structure of SLV and SLVI. These mutations, Stop 1,2,3,4 and Stop 3,4, were introduced into the parental genome of 1bGIT/RLuc (5A) as well as the Stop 1,2 mutations that do not confer a replication defect. *In vitro* transcribed RNA was electroporated into Huh-7.5 cells and luciferase expression at early time points was measured. At 4H, the wild type, pol-, and genome with Stop 1,2 showed equivalent levels of luciferase expression, but a significantly lower (about 9 -fold less) luciferase activity was observed from the Stop 1,2,3,4

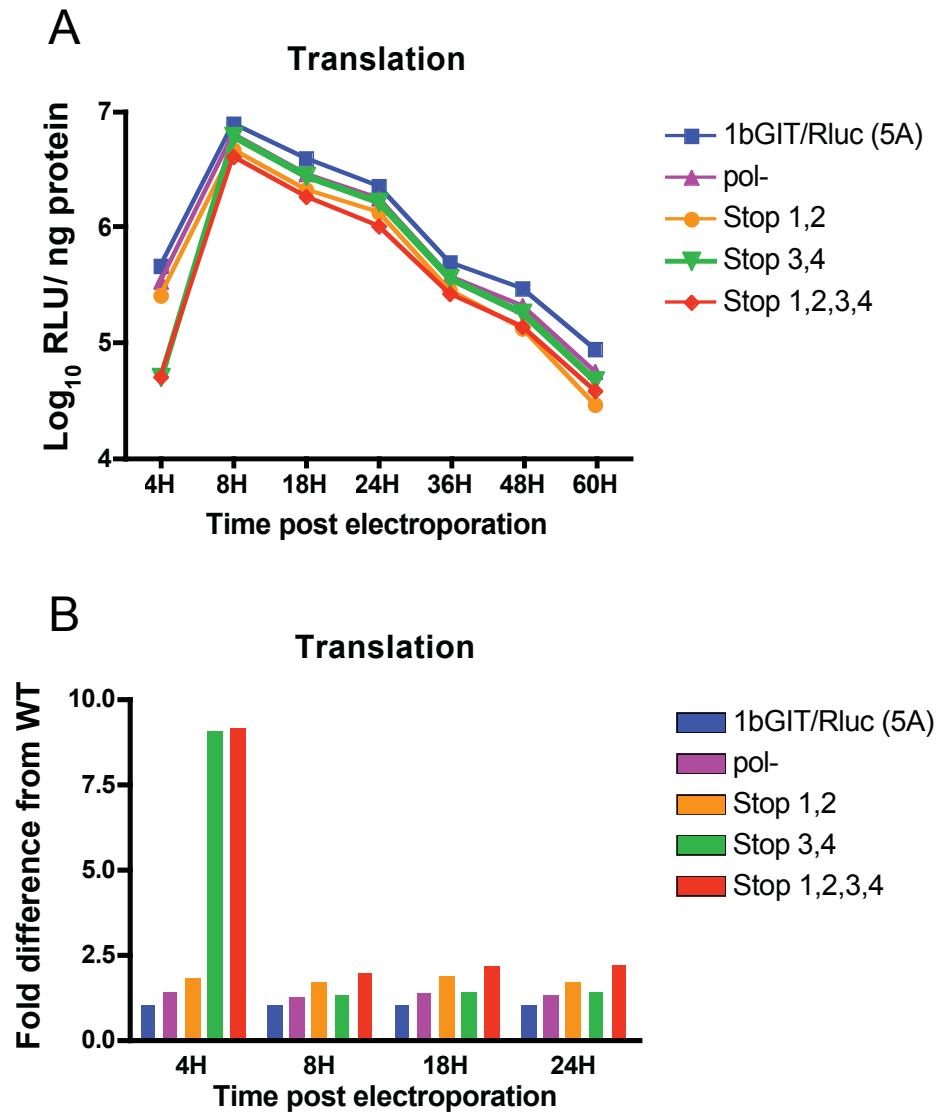


Figure 5.1 Mutations for Stop 1,2,3,4 reduce translation 4H after electroporation in a replication-defective genome. A. Time course of luciferase activity. B. Differences of mutant genomes from wild type at each time point.

and Stop 3,4 genomes (Figure 5.1). At subsequent time points, all genomes were capable of comparable levels of translation of the polyprotein and luciferase reporter gene. These results indicate a dysfunction at an early stage of translation, but not a complete inability to translate the input HCV RNA. Since replication does not occur, only translation of the input RNA was measured and levels declined as the RNA undergoes degradation in the cell.

We repeated this experiment and included additional time points early after electroporation. Any transfection differences were controlled by coelectroporating a capped, poly-adenylated transcript expressing the firefly luciferase gene with the HCV genomes containing the Renilla luciferase reporter and expressing the results as a ratio. No significant differences in transfection efficiency were observed. Again, the genomes with mutations Stop 1,2,3,4 and Stop 3,4 showed a markedly lower translation compared to wild type and the Stop 1,2 genomes (Figure 5.2). The genome with the Stop 1,2,3,4 mutations showed a 3-4 fold reduced level of luciferase expression immediately after electroporation (0H) as well as the early times 2H and 4H post electroporation. Reduced translation was also observed for the genome with Stop 3,4 mutations at levels 2-3 fold below wild type (Figure 5.2).

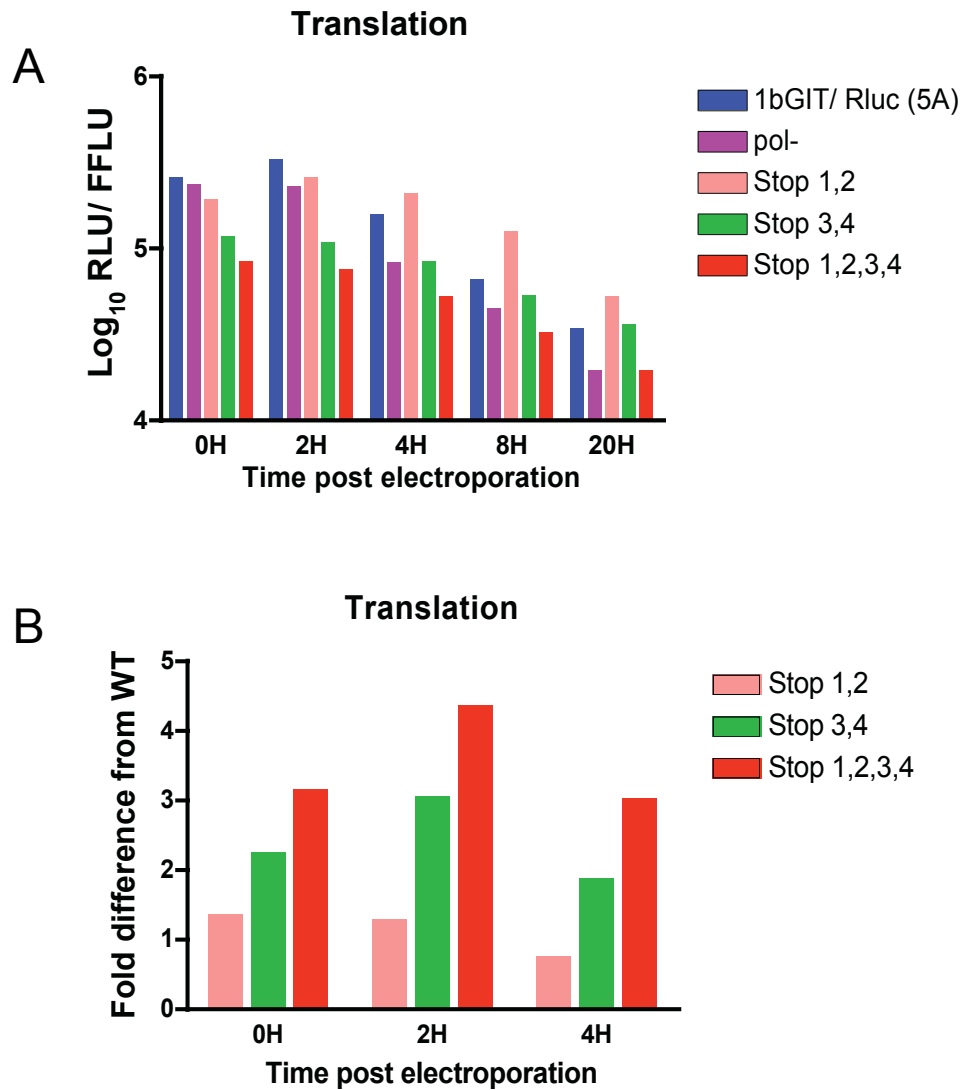


Figure 5.2 Mutations for Stop 1,2,3,4 reduce translation 4H after electroporation in a replication-defective genome. A. Time course of luciferase activity. B. Differences of mutant genomes from wild type at each time point.

Electroporation may alter the state of the cell and may accentuate the mutant genomes' defect in translation. To study the early translation events and differences contributed by mutations of SLV and SLVI, we tested another delivery method of HCV RNA to Huh-7.5 cells. Using a lipid-based RNA delivery, the TransMessenger system, HCV RNA was transfected to Huh-7.5 cells and luciferase expression measured. A capped firefly transcript was included to normalize any transfection efficiencies between RNAs. The firefly luciferase and HCV Renilla luciferase RNA were mixed and incubated with the TransMessenger reagent, and then applied to the Huh-7.5 cell monolayer. Without an additional incubation period, cells were washed and lysed at the time points indicated in Figure 5.3. After 75 minutes, translation was comparable for all genomes, with wild type slightly higher. As translation continued, and more RNA taken up by the cells, luciferase levels increased for the wild type, pol- and genome with the Stop 1,2 mutations. Despite being transfected with equal amounts of RNA and normalized to the firefly luciferase control RNA, luciferase expression barely increased for the genomes with Stop 1,2,3,4 and Stop 3,4 mutations (Figure 5.3) over the 6H time frame. Luciferase expression was above background indicating a low level of translation did occur. We conclude that while the wild type, pol-, and Stop 1,2 genomes could efficiently launch



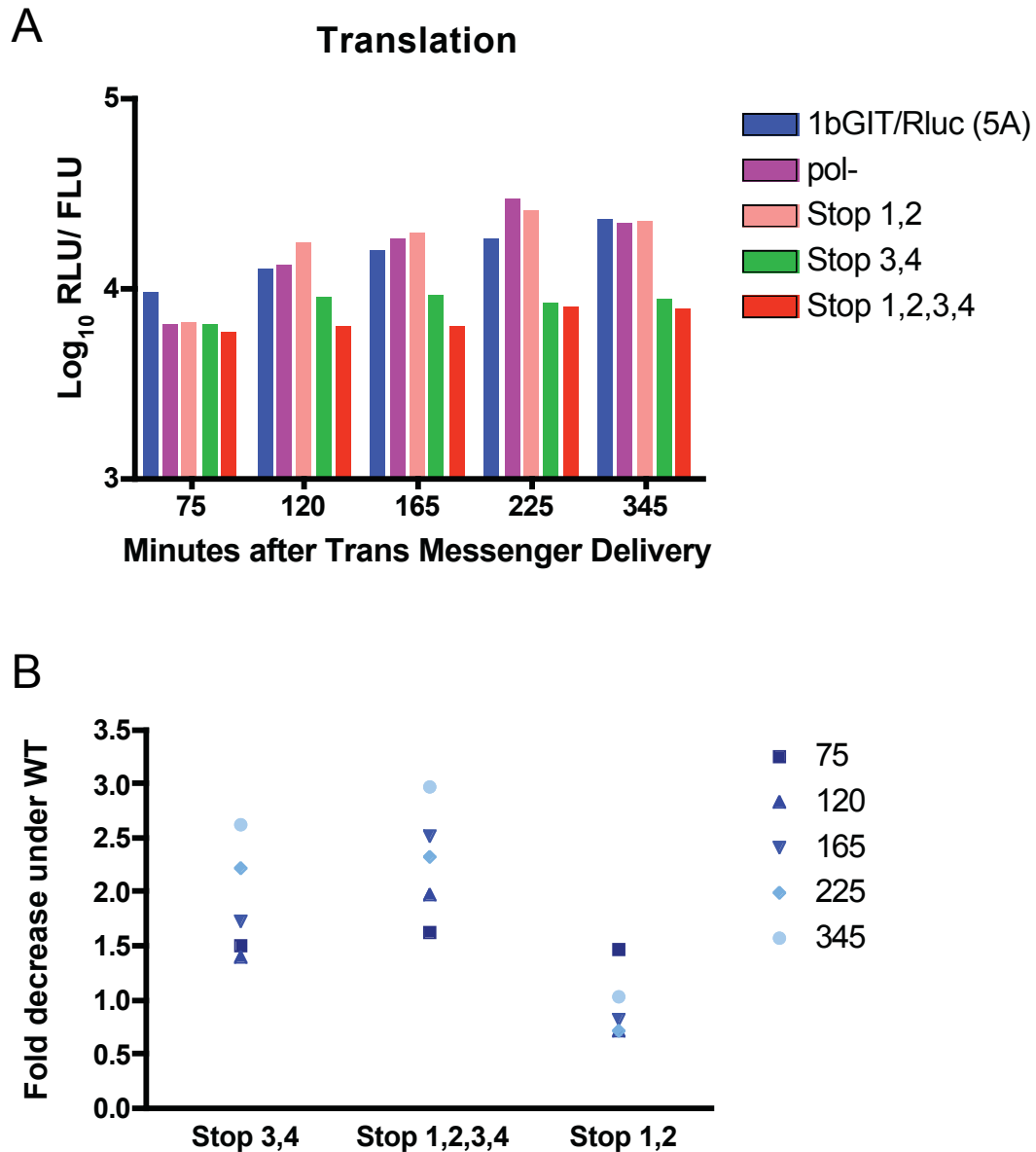


Figure 5.3 Mutations for Stop 1,2,3,4 reduce translation after RNA delivery by lipid mediated TransMessenger delivery. A. Time course of luciferase activity after RNA delivery. B. Differences of mutant genomes from wild type at each time point.

translation and luciferase expression increased with time, this process was defective in the genomes with Stop 1,2,3,4 and Stop 3,4 mutations. The disrupted RNA structure could cause this phenotype in many ways, by altering cellular localization of the incoming RNA, association with inhibitory proteins, or perturbing the normal HCV IRES.

***Mutations in SLV and SLVI reduce translation in vitro.***

The low level of luciferase expression at early time points of RNA delivery indicates the RNA structure of SLVI may participate at the stage of translation initiation or early events localizing the virus RNA with the translation machinery. To ask the question is translation *per se* altered by mutations in SLVI, and to remove the potentially complicating events occurring in cell culture, we compared our mutant genome containing the Stop 1,2,3,4 mutations with a wild type genome in a cell free translation reaction using rabbit reticulocytes extracts. Both genomes contain the Renilla luciferase gene near the carboxy terminus of NS5A to facilitate measurement of translation. Initial translation at 15 minutes was equivalent between the wild type and mutant genomes indicating no defect with translation initiation. The mutant genome however did not maintain wild-type translation levels (Figure 5.4). Between 45 and 90 minutes, translation

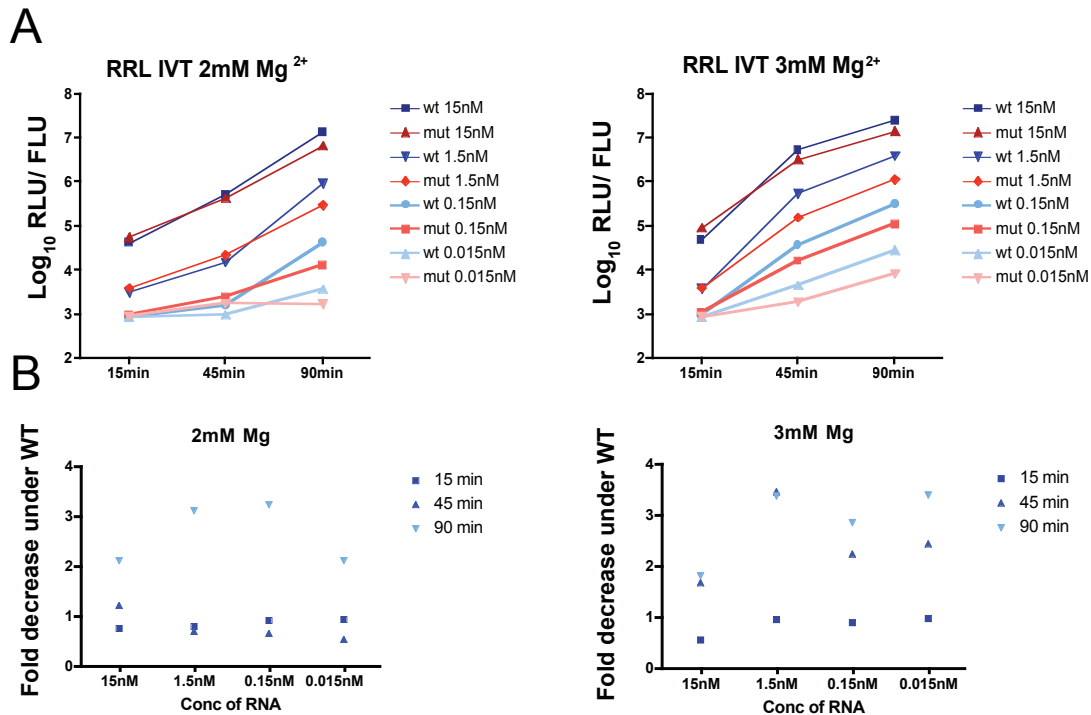


Figure 5.4. Mutations for Stop 1,2,3,4 reduce translation in rabbit reticulocyte lysates. A. The wild type and mutant genome with Stop 1,2,3,4 mutations were measured for luciferase activity with three concentrations of input RNA and two concentrations of magnesium. The amount of Renilla luciferase activity indicating IRES dependent translation of the HCV genomes was normalized to activity from a capped, polyadenylated firefly luciferase message RNA B. Difference of luciferase activity measured from the mutant RNA from the wild type over time.

of wild type had outpaced the mutant. The reaction was programmed with a standard 15nM of HCV RNA as well as ten-fold dilutions to 0.015nM to ensure that we did not saturate the translation system. Compared to the wild-type, the mutant displayed a reduced translation in all four dilutions of input RNA indicating that the phenotype was independent of RNA levels in the system. Previous studies of HCV IRES dependent translation have shown  $Mg^{2+}$  to have a beneficial effect in the rabbit reticulocytes system. Increasing the  $Mg^{2+}$  concentration from 2mM to 3mM increased translation for both genomes (Figure 5.4). Again the mutant and wild type RNA were translated equally at 15 minutes, yet in this system, a difference was observed earlier, at 45 minutes, as the mutant genome failed to increase translation to wild type levels. Again, this was true for all dilutions of input RNA.

The mutant genome exhibited a 2 -3 fold decrease in translation after 90 minutes in the standard reaction, and a 2 -3.5-fold difference when the system was enhanced with the addition of  $Mg^{2+}$  at 45 min (Figure 5.4). This difference was consistent as we extended the time measuring translation up to 20H to determine the end point of the reaction. The maximum level of translation was detected at 4H followed by a decline in luciferase activity presumably as the input RNA and luciferase protein degraded, and necessary translation factors are exhausted (Figure 5.5). The mutant and the wild type

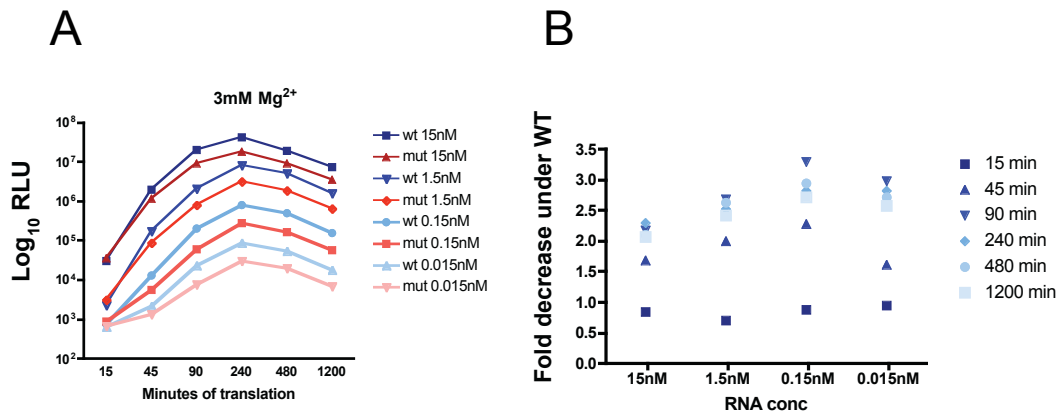


Figure 5.5. End point measurement of translation of wild type and mutant HCV RNA. A. Translation of the mutant RNA shows a reduction after 45 minutes and remains below wild type as the reaction ends for all concentrations of input RNA. B. The difference in translation between mutant and wild type is 2 - 3 fold for all RNA concentrations and are similar for each time point.

levels declined at a similar rate indicating RNA degradation does not account for the differences in luciferase expression. This will be addressed more formally in the next sections. The mutant genome with SLVI disrupted can engage in translation, yet lags as compared to wild type. This could be due to a defect either in the efficiency in initiation, elongation, or termination and thus recycling of the ribosomes. The defect in translation found in the cell free system (2-3 fold of wild type) was similar to that measured in transfected cells (2-4 fold of wild type).

***Differences in translation are not due to RNA degradation.***

While HCV genomes with mutations in SLVI can initiate translation, they fail to reach the activity of genomes with unaltered RNA structures. This disparity could be a result of different rates of RNA degradation, either due to the RNA being lethally misfolded in the cell or the chance of carrying a contaminant that degrades the RNA during the reaction. To address this concern, HCV RNA was transcribed in the presence of  $^{32}\text{P}$ -CTP for the wild type, the genome containing the Stop 1,2,3,4 mutations, and a genome containing a frameshift in E2, which prevents expression of the luciferase reporter from the start of the polyprotein. The radiolabeled RNA was used to program a standard cell free translation reaction in rabbit reticulocyte lysates

and RNA extracted. A short, 6KB radiolabeled RNA was added to the reaction prior to extraction as a control for any loss during the phenol chloroform extraction process. Radiolabeled products were separated by denaturing agarose gel electrophoresis and visualized by autoradiography. No differences in the amount or length of RNA were detected between any of the HCV genomes at all time points (Figure 5.6). Degradation does occur as less HCV RNA is present at later times, but was consistent for all three genomes. Thus, the observed defects in translation are not due to different rates of RNA degradation in the reaction.

***Visualization of radiolabeled protein products reveals a difference in translation at late time point.***

Translation was quantified by expression of a luciferase gene positioned within NS5A, the next-to-last gene of the HCV genome. It was previously shown that the mutations in SLVI do not affect polyprotein processing (Figure 4.4, Chapter 4). To verify that the mutations in SLVI contributed to IRES-dependent translation differences and were not due to faulty expression of the polyprotein, we radiolabeled the protein products during *in vitro* translation using rabbit reticulocytes extracts. Processing of the polyprotein was inefficient in rabbit reticulocyte extracts, despite the

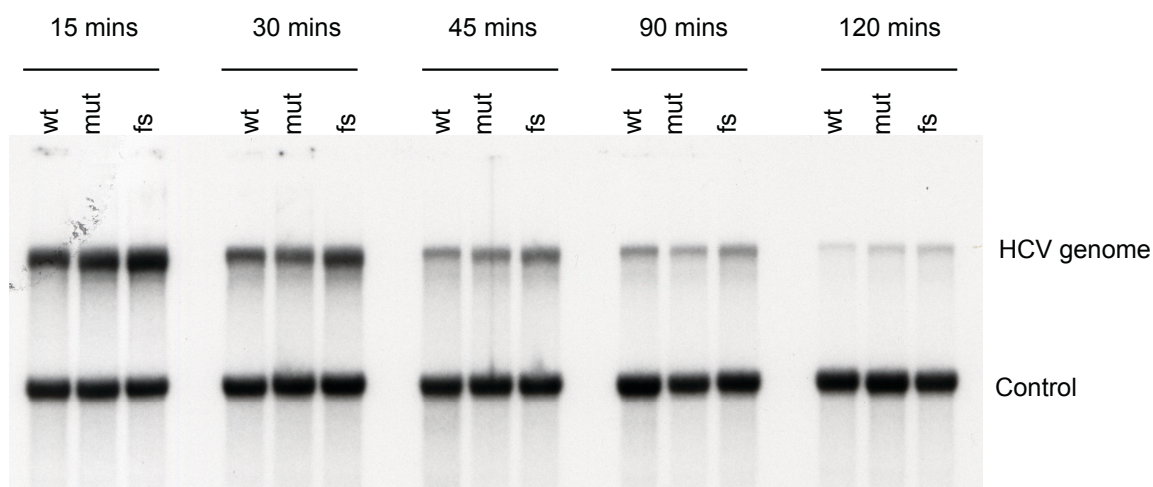


Figure 5.6. The reduced translation observed from HCV genomes with mutations in SLVI is not due to RNA degradation. Labeled HCV RNA for the wild type (wt), mutant (mut) and frameshift (fs) genomes was extracted from rabbit reticulocyte lysates at time points indicated and integrity visualized on a denaturing agarose gel. A control RNA was added to the cell free lysates before extraction to detect any loss during the extraction procedure.



addition of microsomal membranes, creating difficulties when visualizing the individual products. To simplify the system, we truncated a genome by removing 7580 nucleotides between E1 and NS5B. This construct, Delta7580, is a genotype 1a H77 sequence and encodes a protein product of approximately 54kDa. The mutations Stop 1,2 and Stop 1,2,3,4 were introduced into the capsid gene of Delta7580 and RNA generated by *in vitro* transcription. HCV RNA was added to the rabbit reticulocytes extracts and protein products labeled with <sup>35</sup>S-methionine. The truncated HCV protein could be visualized faintly by 15 minutes. The protein product also appeared to be identical in size for all genomes. By the 90 minutes, it was apparent there was less translated protein encoded by the genome harboring the Stop 1,2,3,4 mutations (Figure 5.7). Again, this demonstrates the mutations in SLVI cause a translation defect evident at a later time point in translation.

***In vitro translation using extracts from Huh-7 and Huh-7.5 cells.***

HCV efficiently replicates only in Huh-7 cells and more efficiently in Huh-7.5 cells. Although there are several reports noting the differences of Huh-7 and Huh-7.5 cells it is still unclear what causes these lines to support HCV replication. Given the ability of Huh-7 cells to support HCV

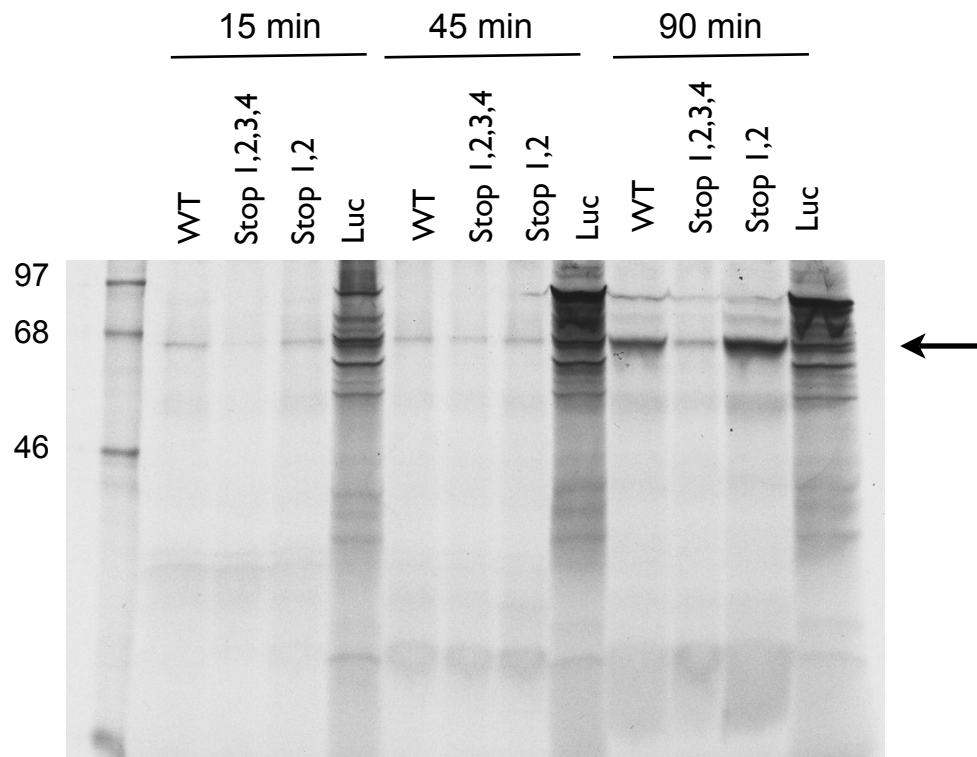
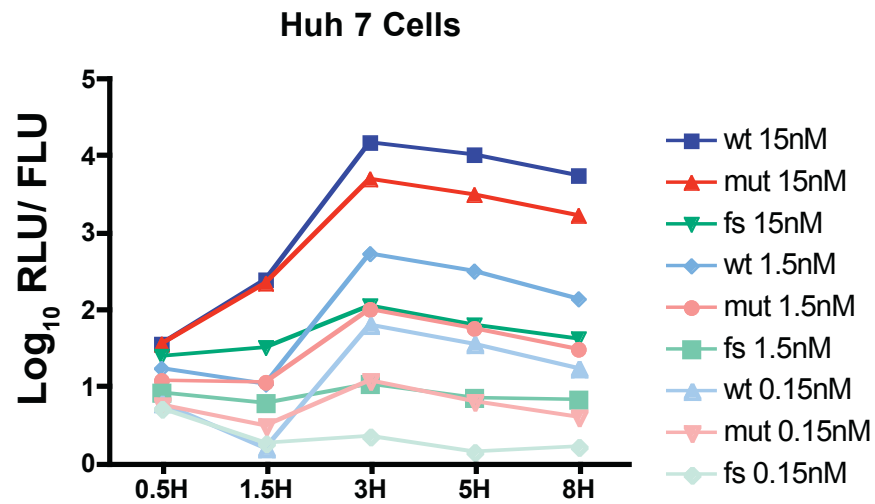


Figure 5.7. Mutant Stop 1,2,3,4 produces less radiolabeled translation products in an *in vitro* translation assay. Rabbit reticulocyte lysates were programmed with HCV RNA templates and translation products labeled with  $^{35}\text{S}$ -Met for the time indicated. A capped, polyadenylated firefly luciferase control RNA was included.

replication, our lab and others developed a cell free translation system using Huh-7 and Huh-7.5 cell lysates (Marcotrigiano, J., unpublished, <sup>138</sup>).

Since Huh-7 cell lines harbor a beneficial environment for HCV replication and we have shown that mutations in SLV and SLVI cause deficient HCV replication in these cells, we asked if the SLV and SLVI mutants have a translation defect in these cells. Using the genotype 1b genome with the Renilla luciferase reporter in NS5A, we compared genomes with the Stop 1,2,3,4 mutations to the parental genome. To detect any aberrant translation that allowed expression of the downstream luciferase gene in a non-IRES dependent manner, a frameshift was introduced in the E2 gene, which should terminate translation of the polyprotein at residue 571. Translation was measured from the three genomes, the wild type (wt), mutant with Stop 1,2,3,4 mutations (mut), and the frameshift genome (fs) using three concentrations of input RNA. Initially, translation occurred with a lower level of luciferase expression and at a slower rate in Huh-7 extracts than in rabbit reticulocytes; therefore, we monitored translation over a longer period of time (Figure 5.8). We included a capped, poly-adenylated RNA expressing the firefly luciferase gene as an internal control. By 3H a substantial difference in translation was detected between the mutant and wild type RNAs. This difference was observed for all three concentrations of

A



B

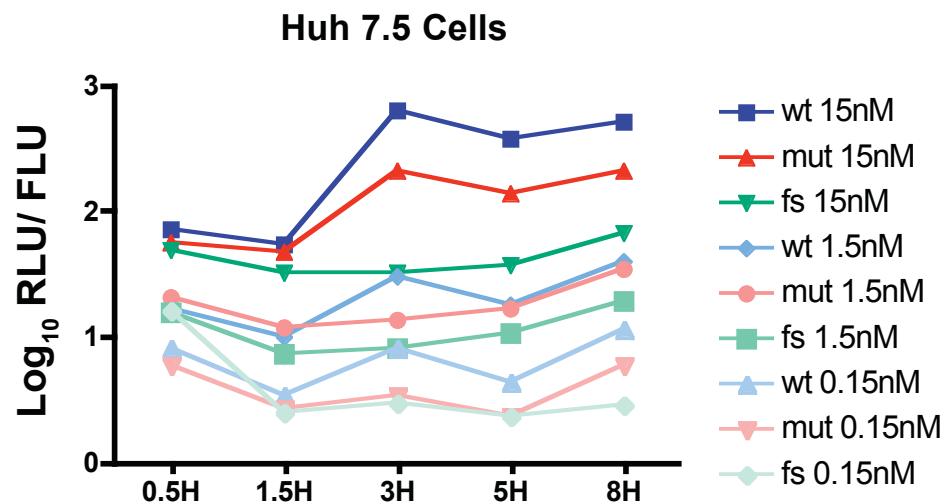


Figure 5.8. An HCV genome with mutation in SLV and SLVI is translated less efficiently in Huh-7 and Huh-7.5 cell free extracts. A. Translation in Huh-7 cell extracts of the wild type (wt), mutant SLVI (mut) and frameshift control (fs) B. Translation in Huh-7.5 cell extracts shows less activity.

template RNAs with a 3- 4 fold reduction at the 15 nM concentration and 4.5 -6 fold reduction at the lower concentrations of 1.5 nM and 0.15 nM.

Luciferase expression was measured for the frameshift genome, but at 130 - 170- fold lower level than wild type. This indicates that either IRES-independent translation can occur, but at a very low frequency or there is a read through of the stop codon at residue 572.

We also compared translation in extracts generated from Huh-7.5 cells. Following a similar preparation protocol, the Huh-7.5 cell lysates showed a reduced ability to translate even the capped firefly transcript. In Huh-7.5 cells, the peak of wild type genome translation was 1.5 logs below that observed in the Huh-7 cell extracts. Despite the lower translation capacity of the Huh-7.5 extracts, there was a difference in luciferase expression between the wild type RNA and the genome harboring the mutations in SLVI (Figure 5.8). At the 3H time point, the mutant genome was reduced 2- fold at the 15 nM concentration and 4-fold reduced for at the lower concentrations of input RNA. The genome with the frameshift did express luciferase, but again at a level 160- fold below wild type. Consistently, the mutant genome with an altered SLVI demonstrated a defect in translation of 2- 4 fold below that of wild type.

***Huh-7.5 cells contribute to an inhibitory factor versus Huh-7 cells.***

The lower level of *in vitro* translation from Huh-7.5 cell extracts is in contrast to the replication ability of these cells. To understand the differences in translation between Huh-7 and Huh-7.5 cells, translation was measured when the extracts were combined. This would indicate if one of the cell lines contained a beneficial or inhibitory factor. In the same reaction, we measured cap- dependent translation of a firefly transcript, and IRES- dependent translation of a wild type or SLVI mutant genome, HCV genome containing the Renilla luciferase gene or a mutant HCV genome with an altered SLVI also expressing the Renilla luciferase gene. The lysates were programmed with 15 nM of HCV RNA and 1.6 nM of capped firefly luciferase RNA. As previously, the Huh-7 cell lysates allowed for more efficient translation than the Huh-7.5 extracts. Huh-7 extracts permitted enhanced translation for both IRES- and cap-dependent templates. We again observed that IRES dependent translation was reduced in the presence of mutations altering SLV and SLVI (Figure 5.9). The capped RNA transcript was included as a control when measuring translation of the wild type and mutant HCV genomes, and therefore measured in two independent experiments (Figure 5.9). When the extracts from Huh-7 and Huh-7.5 cells were mixed, translation only reached the level observed for the Huh-7.5 cell

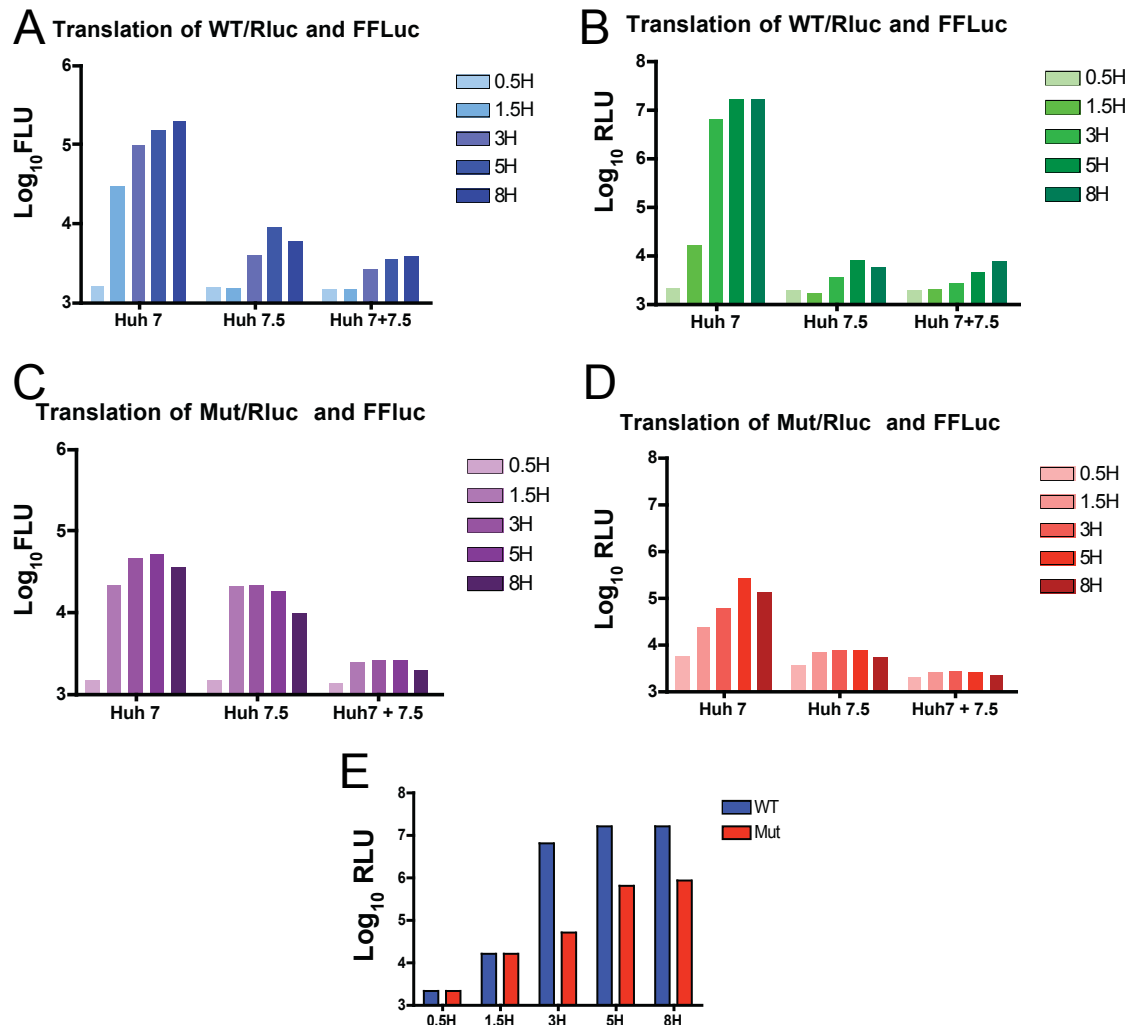


Figure 5.9. Huh-7.5 cells have a negative factor that reduces translation. Luciferase activity of Huh-7, Huh-7.5, and mixtures of Huh-7 and Huh-7.5 cell extracts programmed with a capped, polyadenylated firefly mRNA and HCV RNA expressing Renilla luciferase. Firefly luciferase activity is measured with A. wild type HCV RNA and C. mutant HCV RNA. Renilla luciferase is measured with B. wild type HCV RNA and D. mutant HCV RNA. E. Renilla luciferase of wild type and mutant HCV RNA over time.

lysates or was even further compromised. The reduced ability of translation was evident in both IRES and cap dependent translation. This indicates a potential inhibitory factor is present in the Huh-7.5 cell lysates and not a positive factor in Huh-7 cells. Since both lysates were prepared under the same conditions, i.e. magnesium and potassium concentrations, and translation was conducted using identical reaction conditions, it is unlikely that the reduction in translation is due to unbalancing the optimized reaction.

In all, the reduced capacity of Huh-7.5 cells to conduct *in vitro* translation of both IRES and capped RNAs, is in contrast to the high level of HCV replication supported by these cells.

***IRES mutations known to affect translation have a severe consequence on replication.***

We have shown that translation of HCV genomes with the Stop 1,2,3,4 mutations disrupting SLVI is reduced 25-50% (2- 3- fold reduction) of wild type, yet replication is reduced to an even greater extent to 1-2% of wild type (Chapter 4). The HCV IRES has been the focus of many research groups who have characterized mutations in the RNA sequence and structure of the IRES and their effect on translation. We compared the effects of these IRES mutants with the SLV and SLVI mutants for their ability to translate



and replicate. Two types of IRES mutations were engineered into the H-JFH/Rluc genome. The mutant G266-268A substitutes three guanidines located on the loop of stem III<sub>d</sub> with adenines (Figure 5.10). This mutant was studied for its ability to translate an HCV IRES driven firefly luciferase reporter gene and found to severely reduce translation to 8% of wild type in rabbit reticulocytes and to less than 5% of wild type in HeLa S10 lysates<sup>98</sup>.<sup>58</sup> Despite being able to bind purified 40S subunits, studies of translation initiation kinetics map this defect to a complete absence of 40S formation, the initial complex that must form for translation to begin. The mutant III<sub>d</sub>-delE replaces loop E in stem III of the IRES with a helical segment (Figure 5.10). Translation is only slightly reduced to 46% in rabbit reticulocytes lysates, but shows no translation defect in HeLa S10 lysates. This mutant IRES is highly active and can bind 40S as well as form the intermediate 48S\* and 80S complex, although the latter is formed slightly less frequently than wild type.

We measured these mutations in the context of H-JFH/Rluc genome for translation in rabbit reticulocyte lysates and for replication in Huh-7.5 cells to determine the extent a defect in translation, however mild, would contribute to a reduction in replication. The mutant G266-268A was severely compromised for translation (Figure 5.11). After only 15 minutes,



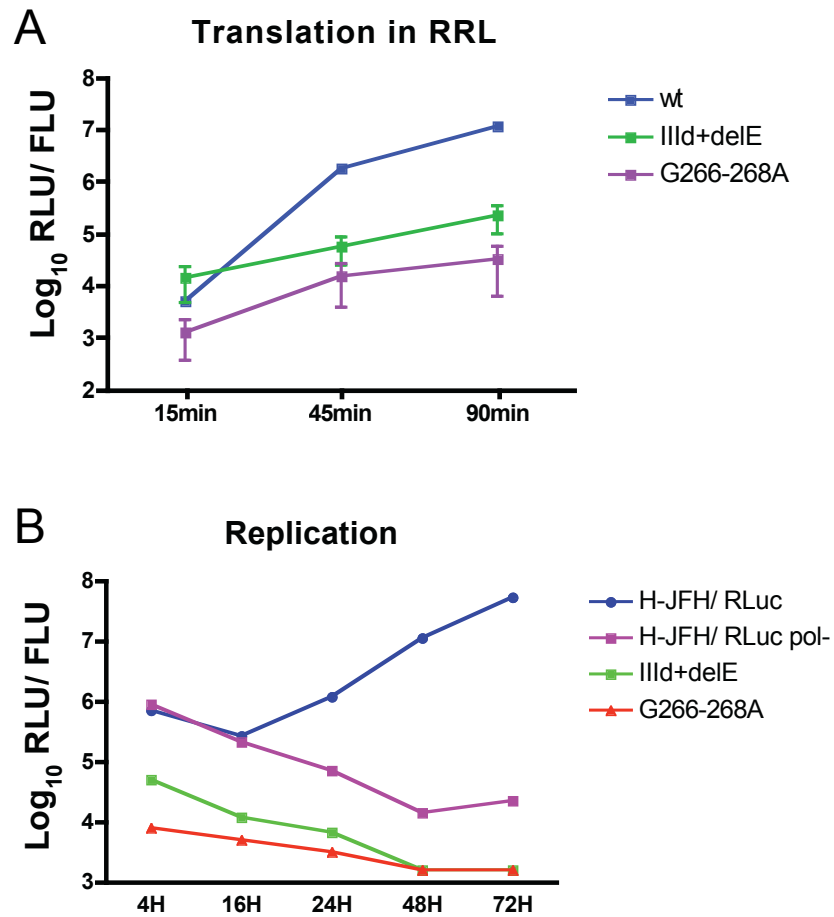


Figure 5.11. Mutations in the IRES that impair initiation confer severe defects in translation and replication assays. A. Luciferase expression in a translation assay using rabbit reticulocyte lysates and B. Replication in Huh-7.5 cells as measured by luciferase.

translation was one-half logs below wild type, which is consistent with a defect in formation of the initial 40S complex. The mutant IIIId-delE showed initiated translation equivalently to wild type, but luciferase levels failed to increase and only reached 2 logs below wild type levels (Figure 5.11). This mutant displayed a much greater defect in translation than in previous reports and may be due to the lower concentration of input RNA (15 nM as opposed to 100 nM) or reflect the further distance of translation required to express the luciferase gene located 2773 nucleotides away from the start site. Both IRES mutations caused a reduction in translation although not complete elimination of all activity. Even the most active IRES mutant able to form all three initiation complexes, IIIId-delE, was translated an order of magnitude less than SLVI mutant genomes. The SLVI mutations may influence IRES driven translation, but at a less critical step in translation.

The IRES mutants displayed an obvious translation and replication defect in Huh-7.5 cells (Figure 5.11). After just 4H post transfection, both IRES mutants were 1-2 logs below both the wild type and pol- control genomes. Luciferase levels rapidly declined and never reached pol- levels. Even for the most active IRES mutant, replication could not be sustained, or perhaps launched in the face of such debilitating translation. The SLVI mutants did not show such a dramatic inability to initiate translation at the

4H time point, but did show a defect as replication was launched and nascent RNA underwent translation. This indicates the role of SLVI may act at the level of replication as well as influencing subsequent translation of newly made RNA.

***Does SLVI participate in a long range RNA-RNA interaction with the 5'NTR?***

The conservation of the 5'NTR and capsid sequence has led to predictions of numerous functions for this region, and in addition, a long range RNA-RNA interaction between the 5'NTR and the base of SLVI has been hypothesized. This interaction involves the unstructured region between SLI and SLII in the 5'NTR (24-38) and the left side basal region of SLVI (428-442) to form a 12 base complementary sequence (Figure 5.12). There is covariant support for this interaction; where most genotypes have an adenine-uridine base pair between bases 29 and 437, genotype 2a sequences contain a compensatory guanine-cytosine base pair at this location. The proposed role of this interaction was to negatively regulate translation. In the reported study, three substitutions were designed to disrupt the complementary interaction with the 5'NTR and base of SLVI and were introduced in a bicistronic, dual reporter construct. The first reporter, a



truncated CAT gene, was expressed by cap-dependent translation and the second cistron driven by the HCV IRES with either 85 or 101 nucleotides of the capsid gene fused to the downstream reporter, a full length CAT gene. Inclusion of the 101 nucleotides of the capsid gene was found to partially reduce *in vitro* translation in rabbit reticulocytes as measured by radiolabeled products visualized by autoradiography. Substitutions in the 5'NTR designed to disrupt three base pairings of the proposed interaction with the capsid gene improved translation slightly above the parental genome, and the corresponding three mutations in the capsid gene improved translation to a higher level of 3 units above the parental. Translation was reduced, to levels below that of wild type, when all six mutations were included to compensate the 5'NTR and capsid gene interaction. The authors proposed that this interaction between the capsid sequence and 5'NTR is a regulatory mechanism that down regulates translation and allows replication to proceed.

The proposed region of the capsid gene that may interact with the 5'NTR is located at the left base of SLVI. We have shown the apical stem of SLVI to influence translation and reduce replication. Previous replication studies have shown the base of SLVI to influence replication, but not to the same extent as the top region of SLVI (Chapter 4). To test if SLVI is acting

in concert with the 5'NTR to regulate translation, we introduced mutations identical to the published study in the 5'NTR and base of SLVI that would disrupt the proposed interaction in the parental genome of H-JFH/Rluc (Figure 5.13). The genome “5'NTR mut” contains mutations in the 5'NTR that would disrupt three predicted base pair interactions with the left side of SLVI in the capsid gene. Three silent mutations were introduced in the capsid gene to create genome “Base Left” that would also disrupt the interaction with the 5'NTR as well as base paired interactions within the base of SLVI. Together, “5'NTR mut/ Base Left”, these mutations should compensate the interaction with the 5'NTR and capsid sequence; however, the base of SLVI would remain disrupted. Previously, translation was measured in the context of only the first 101 nucleotides of the capsid gene, which contain only the left side of the basal stem of SLVI and not the entire stem structure present in the first 166 nucleotides of the capsid gene. Using a full length, monocistronic genome allowed a complete study of the entire stem of SLVI and any additional RNA elements that may also participate in translation. By including the entire HCV genome, we can determine if the competing right side of the base of SLVI has an effect as well as measure any consequences on replication. Mutations along the right side of SLVI, “Base Right”, were therefore introduced that would disrupt the base of SLVI



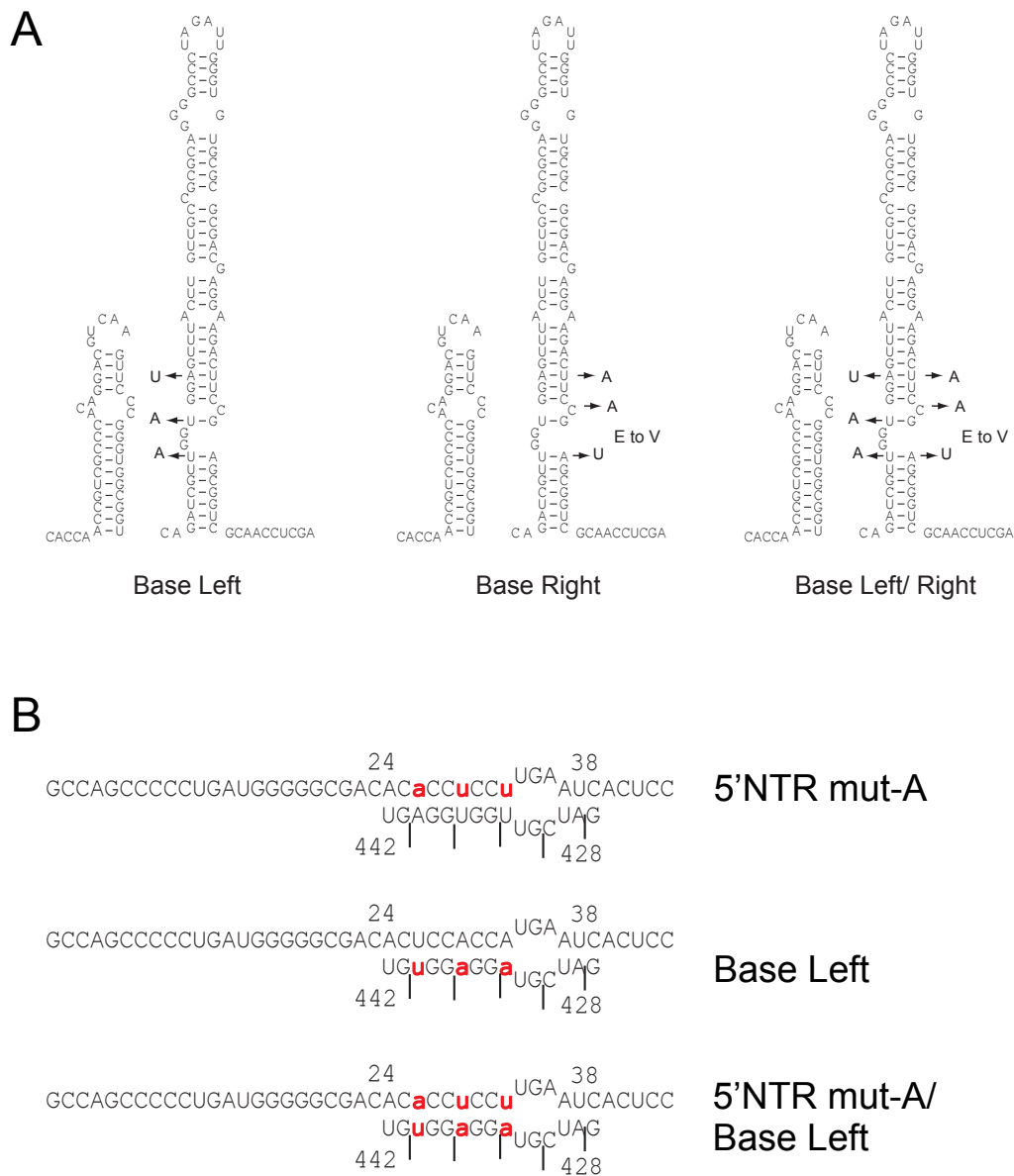
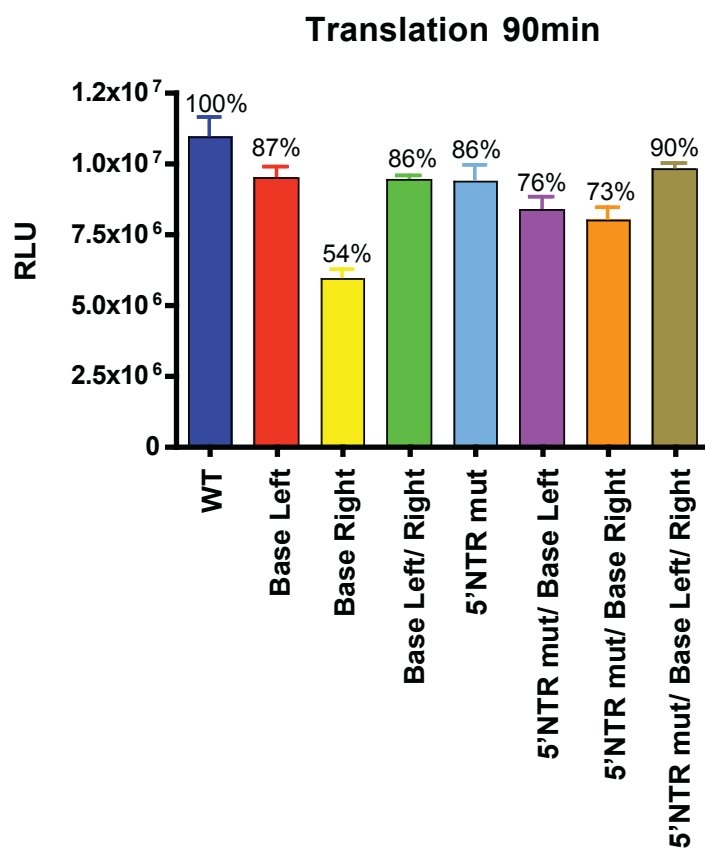


Figure 5.13. Mutations in the base of SLVI and 5'NTR designed to test the proposed interactions. A. Mutations in the base of SLVI of the H77 sequence. B. Mutations in the 5'NTR designed to interrupt the complementary sequence with the base of SLVI.

and in combination with mutations on the left side, Base Left, would predictably compensate the structure of SLVI, “Base Left/ Right”.

Translation was measured in rabbit reticulocyte lysates by quantifying Renilla luciferase expression from two independent preparations of RNA for each HCV genome (Figure 5.14). Reaction conditions were similar to the published study in which lysates were programmed with a 10 uM concentration of RNA in a 20uL reaction for 1H. We measured translation of two concentrations of input RNA at 7.5 nM and 0.75 nM over a time course of from 15 to 180 minutes. Data from the 7.5 nM reaction at 90 minutes is presented, yet both dilutions of RNA yield the same relative results. The proposed interaction between the 5'NTR and capsid gene was reported to negatively regulate translation with disruptive mutations in either location leading to an increase in translation. In the H-JFH/Rluc parent, mutations in the 5'NTR (5'NTR mut) and in the interacting partner of the capsid gene (Base Left) did not increase translation, but reduced translation to 87% and 86% of wild type, respectively (Figure 5.14). Compensating the mutation sets in the 5'NTR and Base Left of the capsid (5'NTR mut/ Base Left) to restore the proposed interaction also reduced translation to 76% of wild type. Furthermore, mutations in the 5'NTR and the right side of the base of SLVI, the competing element for the Base Left, (5'NTR mut/ Base Right) also



5'NTR interaction:	+	-	+	-	-	+ <sup>c</sup>	-	+ <sup>c</sup>
SLVI:	+	-	-	+ <sup>c</sup>	+	-	-	+ <sup>c</sup>

Figure 5.14. Translation of mutants designed to test the 5'NTR and Base of SLVI interaction in rabbit reticulocyte lysates. Translation was measured by luciferase activity at 90 minutes. Mutants permitting an interaction are designated with a "+" and mutants potentially disrupting an interaction is listed with a "-". Compensatory mutation sets include a "c".

reduced translation to 73% of wild type. In any context, disrupting the interaction between the 5'NTR and Base Left region of the capsid gene actually impaired translation, although only slightly, as opposed to improving translation as reported. This suggests there is not an interaction between the 5'NTR and left base of SLVI, which acts as a negative influence on translation.

Translation could, however, be influenced by the base of the SLVI structure itself. Where mutations along the left side (Base Left) minimally reduced translation to 87%, mutations on the right side of the basal structure, Base Right, impaired translation to 54%. Genomes harboring both sets of mutations, restoring the base of SLVI, compensated for the effect of Base Right alone to 86%. Disrupting the base of SLVI in the context of the 5'NTR mutations also demonstrated that SLVI influences translation. The genome with the 5'NTR and Base Left mutations was impaired for translation to 76%. This result favors the SLVI structure having a role in translation, since compensating the 5'NTR and Base Left did not restore translation to wild type levels. Similarly, genomes with 5'NTR and Base Right mutations were reduced to 73%, again indicating that it is the disruption of SLVI that compromises translation. Restoring the base pairing

of SLVI, even in the presence of the 5'NTR mutations (5'NTR mut/ Base Left/ Right), improved translation to 90% of wild type.

These data indicate SLVI participates in translation and does not indicate a role for the complementary sequences of the 5'NTR and capsid gene. The translation effects of the mutations in the basal region of SLVI results are consistent with the effects observed during replication. Mutations at the left base of SLVI (Left Base) reduced replication slightly, yet the mutations for the right side (Right Base) were more severe (see Chapter 4). Replication improved in the presence of both mutation sets that restored the base of SLVI. This replication defect could be a reflection of the reduced capacity for translation as well as indicate a role of the capsid protein as the mutations in Right Base also change an amino acid. The possible role of the capsid protein will be addressed in an upcoming section.

***Deletions of the 5'NTR reduce translation.***

Point mutations introduced into the 5'NTR designed to disrupt a long range RNA-RNA interaction with the capsid gene proved to not increase translation as previously reported. Since individual base changes may have unanticipated effects on RNA structure and the predicted participating bases may be incorrect, we deleted the first 40 nucleotides in the parent genome of

H-JFH/Rluc to determine if translation was affected. This region of the 5'NTR contains SLI at position 5-20 and the unstructured area bordering the 5' end of the IRES. Both regions have been shown to be required for replication and while deletion of SLI leads to an increase in translation, little is known about the unstructured area predicted to interact with the capsid gene <sup>41</sup>. It is also unknown if the regions of SLVI involved with HCV replication are influenced by this region in the 5'NTR. We have shown that the base of SLVI can affect translation independently of the 5'NTR. It has also been shown that the basal stem of SLVI influences HCV replication, but that the apical stem is more critical for replication (Chapter 4). We were curious if the detrimental effects observed by mutations in the apical stem occur in the presence of the 5'NTR. The mutations conferring Stop 1,2,3,4 and Stop 3,4 were introduced into the Delta40 construct with bases 1-40 removed and translation compared in rabbit reticulocyte lysates at two different RNA concentrations. Enhanced translation was found in the Delta40 genome, confirming the previous findings for SLI. The mutations in SLVI suppress translation to 66% and 68% of wild type again indicating a role of this RNA structure in translation as well as replication (Figure 5.15). In the context of the Delta40 genome, the mutations of SLVI also showed a similar reduction in translation, despite the enhanced activity due to the

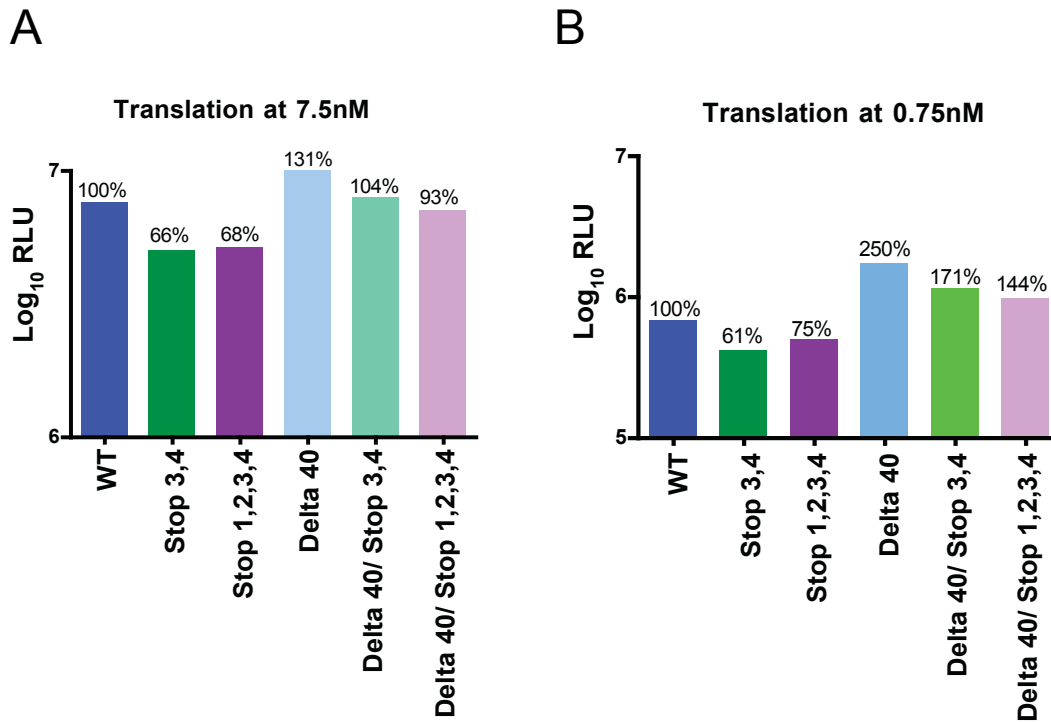


Figure 5.15. Translation of mutants designed to test the 5'NTR and Base of SLI interaction in rabbit reticulocytes. A. Translation at 7.5nM and B. 0.75nM of HCV RNA as measured by luciferase activity.

removal of the first 40 nucleotides. While SLI and the region 21-40 of the 5'NTR did display an effect on translation, the mutations on SLVI in the capsid gene alter translation in an independent manner.

***Deletion of the capsid gene reduces translation.***

The influence of the capsid gene on translation has been the subject of numerous studies predicting interactions with the 5'NTR, altering the stability of the IRES, and debate between the capsid protein or capsid sequence as being the component responsible. The HCV IRES requires sequence in the capsid-coding gene to function; the minimum length being 33 nucleotides, though generally the first 16 amino acids of capsid are included in replicon and bicistronic constructs. It has been demonstrated that the region immediately downstream of the IRES must be devoid of RNA structure, presumably to not compete with the integrity of the IRES structure. Early studies of the IRES found that inclusion of additional sequences of the capsid gene led to a decrease in translation activity. Also the IRES varied in its ability to translate certain downstream reporter genes. Although the IRES does not require sequences beyond 33 nucleotides of the ORF to function, these results suggest that the efficiency of HCV IRES-dependent translation can be influenced by sequence further downstream.



The RNA structures of SLV and SLVI are located just 46 nucleotides from the AUG start codon and 13 nucleotides downstream of the minimal 3' boundary of the IRES. Since the entire capsid gene has been shown to down regulate IRES activity, we compared translation in the presence of only SLV and SLVI and in the absence of both structures. In the parental construct of H-JFH/Rluc, we deleted the capsid gene either from amino acid 15 (position 387) or amino acid 57 (position 513) to amino acid 171 (position 851), the start of the signal sequence of E1. A nine nucleotide linker that encodes three alanines was included at the deletion junction. The mutations of Stop 1,2,3,4 were introduced to determine if SLVI influences translation in the absence of downstream capsid sequence. In an *in vitro* translation reaction, we found that removal of the capsid gene, Core Del 15-171, actually slightly decreased translation to 88% of wild type (Figure 5.16). This effect on translation is the opposite of what has been previously reported. Other translation studies of the capsid gene have utilized bicistronic constructs where the extended regions of the capsid gene were fused directly to a reporter and frequently lack other elements of the HCV genome such as the rest of the polyprotein and the 3'NTR. In this system, all elements of the HCV genome are present as a monocistronic construct, but with the partial deletion of the capsid gene. RNA structures originally downstream in the

genome are brought closer to the 5'NTR and may interfere with IRES function. Extending the region of the capsid gene to include the RNA structures of SLV and SLVI leads to a further reduction in translation, to 80% of wild type (Figure 5.16). This is consistent with previous studies where RNA structures have been proposed to contribute a negative effect on translation. Inclusion of the silent mutations of SLV and SLVI compromised translation additionally to 65% of the parent genome. We have shown mutations in SLVI to reduce translation, and this study is consistent with those results, yet in addition, reveals that the role of SLVI in translation does not involve downstream elements.

***Genotype differences in the 5'NTR do not affect replication of the J6-JFH genome.***

The 5'NTR sequence shows great conservation, yet there are several nucleotide differences that are genotype specific. It has been shown that the genotype 1b sequence conducts translation twofold less efficiently than genotype 1a sequence<sup>55</sup>. The genotype 1b sequence contains two nucleotide differences, an AG at positions 34 and 35 (numbering based on H77 sequence) whereas genotype 1a contains the dinucleotide GA. Exchanging the AG of genotype 1b with the GA of genotype 1a removed the two-fold

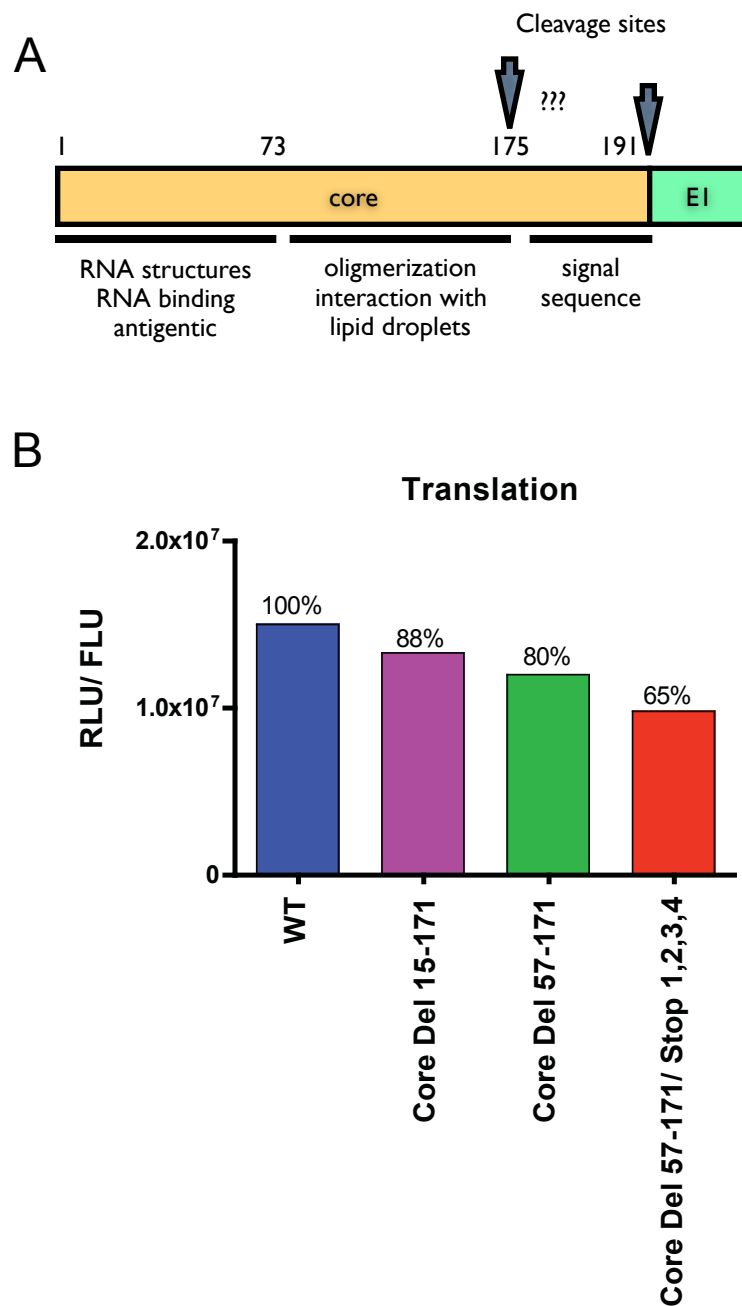


Figure 5.16. Translation of HCV genomes with deletions in the capsid gene is reduced. A. Domains of the capsid protein. B. Translation in rabbit reticulocyte lysates after 90 minutes as measured by luciferase expression.

defect in translation. It was also found that the increased genotype 1b translation was dependent on the presence of the capsid gene, therefore predicting a long range RNA-RNA interaction between the unstructured region of the 5'NTR and the capsid gene or the capsid protein <sup>64</sup>. The 5' NTR sequence of genotype 2a sequence is identical to the genotype 1a sequence at positions 34 and 35, yet there is a difference at position 29 (H77 numbering) where most genotype 2a sequences contain a guanine and most other genotypes contain an adenine (Figure 5.17). This position was described in the previous section as being phylogenetic evidence for interacting with a partner sequence in the capsid gene. To address if these genotype differences contributed any effect on HCV replication, we introduced substitutions at position 29 and at positions 34 and 35 in a genotype 2a J6-JFH/Rluc genome. In this chimera, the 5'NTR sequence is from the JFH genome except for the J6-substitution of G301U. *In vitro* transcribed RNA for the genomes with the G29A mutation (JFH 5'NTR 1a) and the dinucleotide mutations AG to GA at 34/35 (JFH 5'NTR 1b) were electroporated into Huh-7.5 cells and replication measured by luciferase expression. There was no difference at any time point between any of the genomes, indicating no effect of the genotype differences between the 5'NTRs (Figure 5.17).

**A**

	H77	GCCAGCCCCCUGAUGGGGGCGACACUCCACCAUGAAUCACUCC
	Con1	GCCAGCCCCCGAUUGGGGGCGACACUCCACCAUAGAUCACUCC
	J6	-ACCCGCCCCUAAUAGGGGGCGACACUCCGCCAUGAAUCACUCC
	JFH	-ACCUGCCCCUAAUAGGGGGCGACACUCCGCCAUGAAUCACUCC
Mutants:	JFH (1a)	-ACCUGCCCCU <b>G</b> AUAGGGGGCGACACUCCGCCAUGAAUCACUCC
	JFH (1b)	-ACCUGCCCCUAAUAGGGGGCGACACUCCGCCAU <b>A</b> GAUCACUCC
	JFH (J6)	-ACC <b>C</b> GCCCCUAAUAGGGGGCGACACUCCGCCAUGAAUCACUCC

**B**

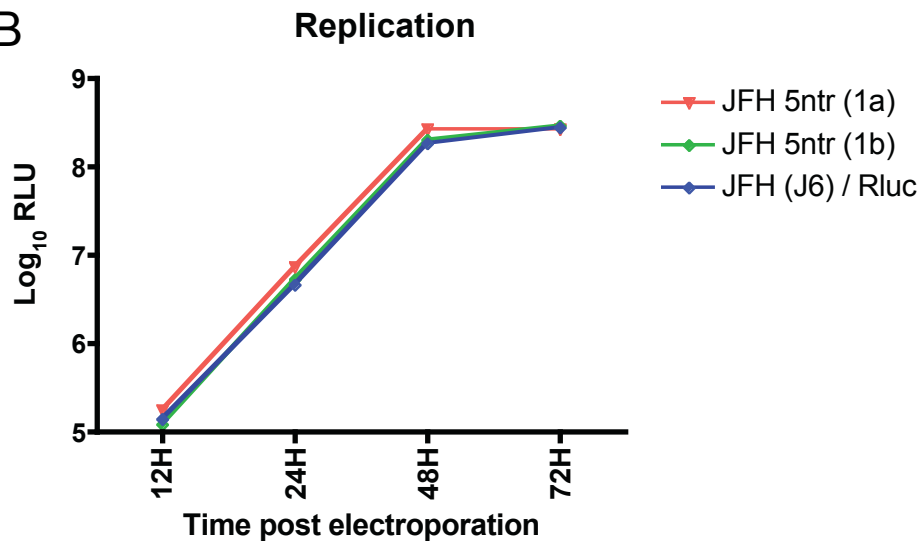


Figure 5.17. Genotypic differences in the 5'NTR do not affect replication. A. Alignment of genotype 1a, 1b, and 2a sequenced of the 5'NTR from position 1-43. B. HCV RNA measured by Real Time qRT-PCR shows no differences in replication.

***Mutations in the 5'NTR eliminate HCV replication.***

We have shown that mutations in the 5'NTR that would disrupt the predicted Watson-Crick base pair interaction with the capsid gene do not lead to an increase in translation. However we did discover that mutants in the predicted interacting partner, the left base of SLVI, did have both a translation and replication defect. To determine if there was any replication consequence of these changes in the 5'NTR in Huh-7.5 cells, the mutations, 5'NTR mut-A, were engineered into the parent genome of H-JFH/Rluc and replication measured by luciferase expression (Figure 5.18). Although input RNA was translated, there was no replication as luciferase levels quickly fell below even levels for the pol- genome (Figure 5.19). This region of the 5'NTR is known to be an essential element for replication; possibly because it forms the promoter for plus strand synthesis in the minus sense.

The long range RNA-RNA interaction of the capsid gene and 5'NTR involves 9 base pairs at the 5' end of the 5'NTR, a bulge of three bases each side, and 3 additional complementary nucleotides (Figure 5.18) <sup>64</sup>. It has also been published that increasing the interaction between the 5'NTR and capsid gene by increasing the base pairing of the bulge region, further reduces translation. These three mutations, 5'NTR mut-B, involved the changes

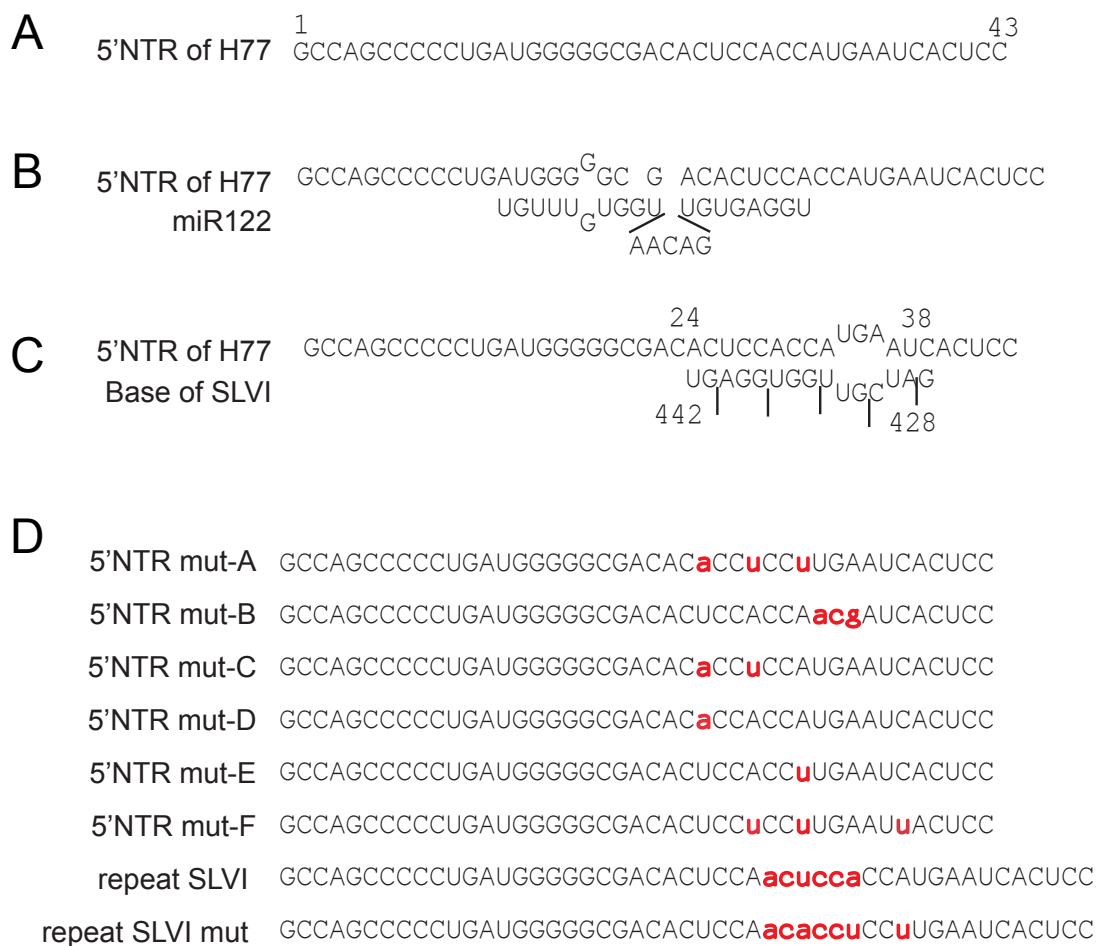


Figure 5.18. The 5'NTR has been shown to interact with a microRNA and RNA sequences from the capsid gene. A. Positions 1-43 of the 5'NTR of H77. B. The seed region for the interaction of miR122 and the 5'NTR. C. The interaction of the 5'NTR and the base of SLVI. D. Mutations introduced in the 5'NTR designed to disrupt the proposed interactions and test in HCV replication.

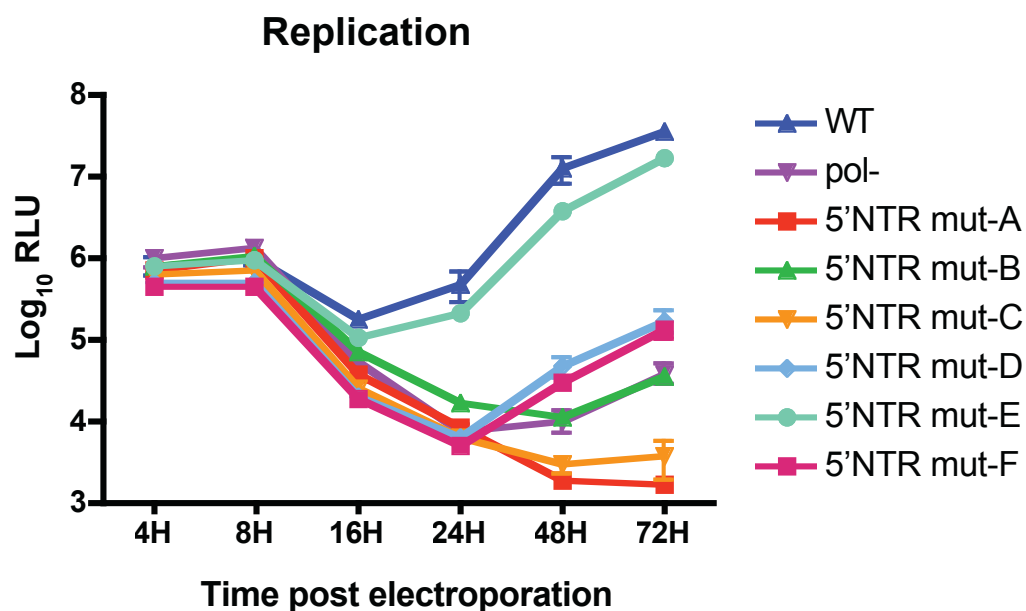


Figure 5.19. Mutations in the 5'NTR affect HCV replication. All mutations introduced in the 5'NTR debilitate HCV replication with the exception of the single base change of mut-E, which lies outside of the miR122 seed region and would disrupt 1 out of 12 base pairings with the capsid gene. Two RNAs were tested for each genome and averages shown.



U33A, G34C, A35G and were tested for their ability to replicate in the H-JFH/Rluc genome (Figure 5.19). Replication was eliminated in the presence of these mutations implying a lethal alteration of the 5'NTR or the minus strand promoter.

Another feature of this region in the 5'NTR is that it is a target for a microRNA and the recognition of this sequence has the unusual outcome of enhancing HCV replication <sup>59</sup>. The microRNA, miR122, is very abundant in liver cells and is present in Huh-7 and Huh-7.5 cells. In fact Huh-7 cells and cell lines derived thereof are the only transformed liver cell line to have maintained their expression of miR122, leading to the hypothesis that miR122 expression is responsible for the ability of these cells to support HCV replication. It is still unknown how or at what stage miR122 is beneficial for replication. The seed region for miR122 shown to be critical for HCV replication is located between bases 22 and 29 of the 5'NTR (Figure 5.18). In addition to potentially disrupting the minus strand promoter mutations, the mutations of 5'NTR mut-A would also alter the target site of miR122. To disrupt the interaction between the 5'NTR and the capsid gene, without interfering with the miR122 seed region, a set of mutations, 5'NTR mut-C, D, E, F were generated in the 5'NTR and tested for their ability to replicate (Figure 5.19). All constructs were unable to replicate except 5'NTR

mut-E. This mutant contained only one substitution, A32U, which is outside of the miR122 seed area and could replicate efficiently. Unfortunately, this one base change may not completely abrogate an interaction with the capsid gene sequence, leaving that interaction untestable in a replication context.

We did attempt to separate these two features by duplicating this overlapping region in the 5'NTR. We introduced a 6 base repeat at position 29, which moves the interacting capsid region downstream from the miR122 seed region. This genome, surprisingly, could efficiently translate and also could replicate, but to a much attenuated degree (Figure 5.20).

### ***Potential interactions of the capsid protein in trans.***

The reduction of IRES directed translation in the presence of extended lengths of capsid gene has been the subject of numerous studies and debates as to whether the capsid sequence or the capsid protein acting in a regulatory fashion. Early biochemical studies of the capsid protein identified the amino terminus, residues 1-75, as capable of binding RNA <sup>121</sup>. Given the highly basic amino acid composition and location in the genome, the capsid protein was predicted to be a structural component of the virion and was shown to bind RNA, although without specificity as it could bind HBV and ribosomal RNA as well as HCV RNA. Subsequently, the capsid protein was reported

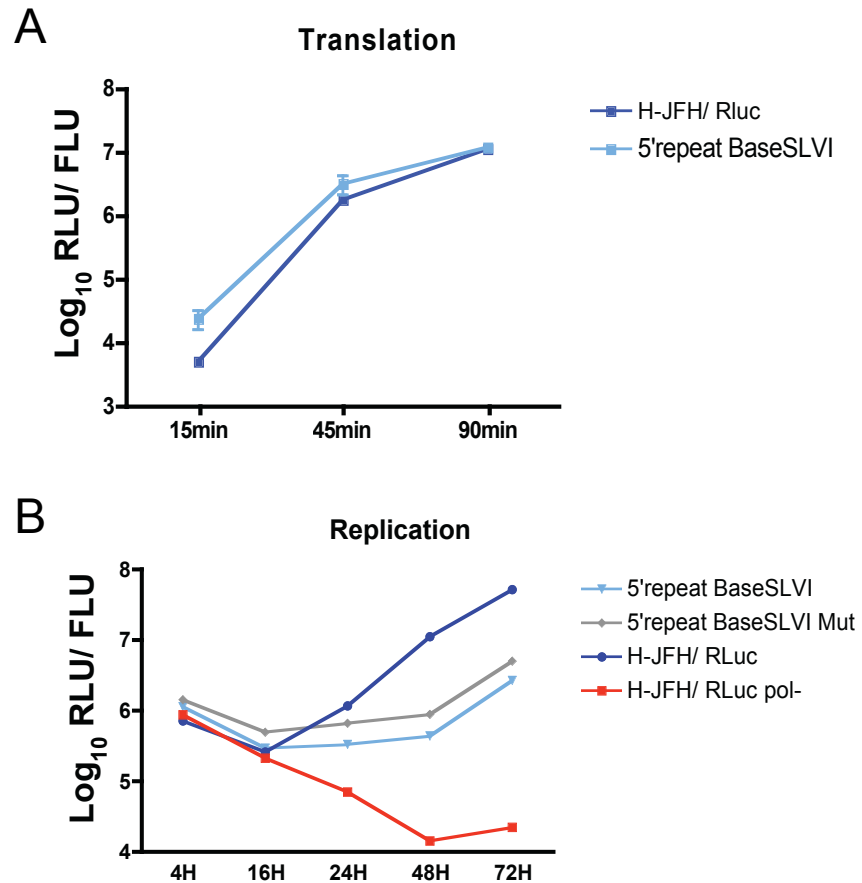


Figure 5.20. Introduction of a repeated sequence predicted to base pair with the SLVI in the 5'NTR. A. *In vitro* translation of HCV with the repeated base of SLVI element shows no reduction in translation. B. Replication, however, is impaired. Replication is measured by luciferase expression.

to bind the 5'NTR and repress translation. Binding the HCV IRES has been described for the capsid protein expressed both *in cis* and *in trans* and using either purified recombinant protein in UV crosslinking and gel shift assays or recombinant baculovirus and mammalian systems expressing capsid protein in cell culture. Studies including other basic, virus capsid proteins, such as that from VEE have argued against the HCV capsid preferentially binding the IRES and interpret the reduced translation a consequence of downstream RNA structures, potentially SLV and SLVI. Increasing the length of capsid gene was shown to have a positive effect on replication of an HCV-poliovirus chimeric virus. Chimeric viruses with the longer regions of capsid gene were shown to have an improved replication and virus producing capacity than chimeric viruses harboring shorter capsid regions. It is unknown if the capsid sequence is used in virus assembly and packaging. To date, there have been no studies on the specific RNA structures of SLV and SLVI in the capsid gene and how they affect IRES activity and influence virus replication and infectivity.

To study the effects of the capsid protein and capsid RNA structures on virus replication and infection, we compared replication and virus infectivity of genomes with deletions in the capsid gene. Using the parent H-JFH/Rluc, the capsid gene was deleted from amino acids 15-171 which

includes the sequence necessary for IRES stability but removes SLV and SLVI as well as the majority of the capsid protein up to the hydrophobic E1 signal sequence (Figure 5.21). To compare the effects on the presence of SLV and SLVI we extended the amount of capsid gene to amino acid 57 and deleted downstream sequence to residue 171. We also introduced the mutations Stop 1,2,3,4 into the capsid sequence to investigate if the structure of SLV and SLVI contributed to an effect. Previously we have shown that the deletion of the capsid gene slightly reduced the level of translation *in vitro* and the presence of mutated structures SLV and SLVI further compromised translation.

To test if the capsid protein or sequence can act *in trans* to influence replication and virus production, we placed the HCV capsid gene in a recombinant Venezuelan equine encephalitis (VEE) virus expression system. Replication of the alphavirus, VEE, occurs entirely in the cytoplasm and independently of its structural proteins which are encoded from a subgenomic RNA transcribed from the minus strand replication intermediate. Due to the early and abundant expression of the structural proteins, expression vectors were developed that replace the alphavirus structural genes with heterologous genes for a protein of interest. Duplication of the subgenomic promoter allows expression of two

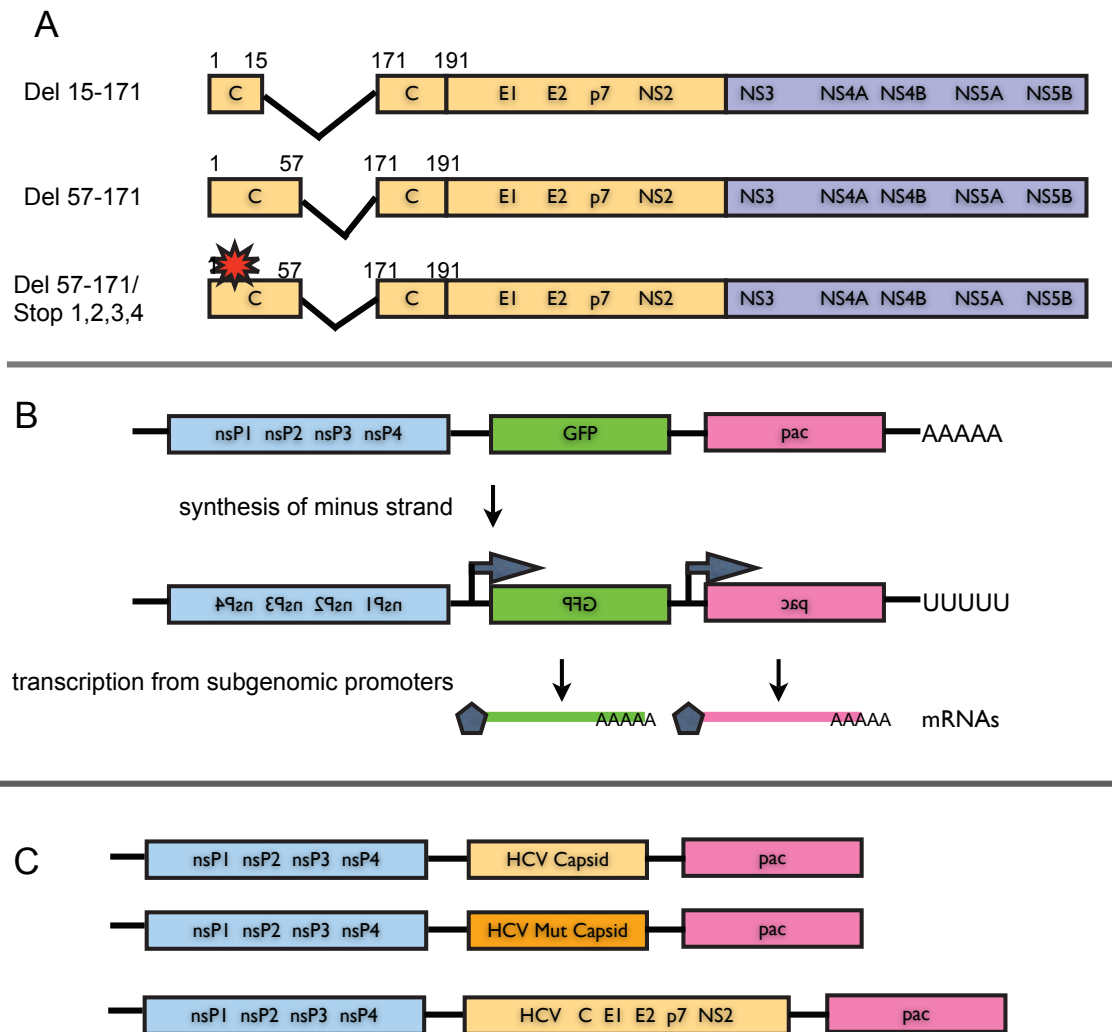


Figure 5.21. Transcomplementation assay. A. The HCV genomes with deletions in the capsid gene and/ or mutations in SLV and SLVI. B. Expression from VEE dual subgenomic promoters occurs after synthesis of minus sense genome and produces capped mRNA transcripts. C. HCV proteins expressed from a VEE subgenomic promoter.

heterologous proteins from independent mRNAs. Petrakova et al. selected for adaptive mutations that enable replication of a dual genomic VEE replicon without cytopathic effect in Huh-7 cells <sup>101</sup>. We cloned the HCV capsid protein (residues 1-191) behind one subgenomic promoter of the adapted VEE replicon. The puromycin N-acetyl transferase gene, *pac*, was also included behind a second subgenomic promoter to allow for selection of cell lines harboring the VEE replicon.

We measured replication of HCV genomes containing the capsid deletions of: Del 15-171, Del 57-171, and Del 57-171/Stop 1,2,3,4 in Huh-7.5 cells. After 72H, replication was severely impaired for the Del 15-171 genome to levels of the pol- control (Figure 5.22). Inclusion of the capsid structures SLV and SLVI in the genome Del 57-171, improved replication, but was still 59-fold lower than wild type. Disrupting SLV and SLVI furthermore limited replication and levels decreased to 807 fold below that of wild type. This impairment of replication is more dramatic than that observed for translation. While genomes with Del 15-171 showed only a slight reduction in translation (88% of wild type), replication was obviously compromised (Figure 5.22). This result was unexpected, as the capsid protein is not required for replication of subgenomic replicons. The minimal capsid sequence required for IRES activity was included; however,

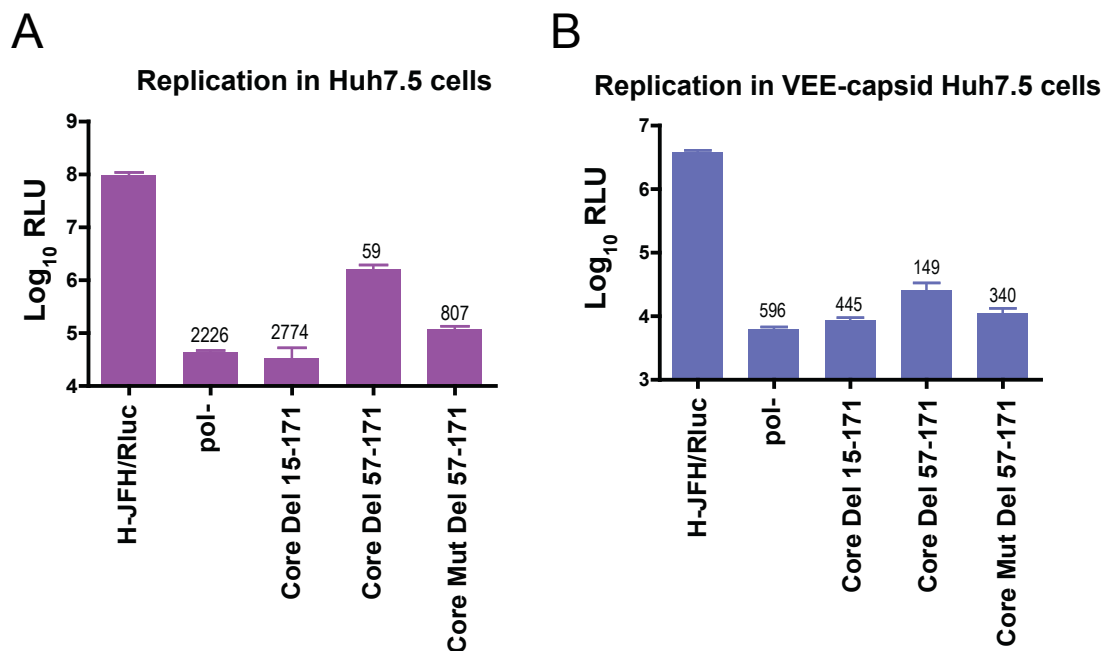


Figure 5.22. Deletions in the capsid gene and mutations in SLV and SLVI confer a defect in replication. Replication was measured by luciferase expression at 72H post electroporation. A. Replication in Huh-7.5 cells B. Replication in a VEE cell line expressing the HCV capsid protein. Values are averaged measurements from two electroporations of two independent preparations of RNAs each plated into two wells. The fold difference of luciferase levels below wild type is indicated above each bar.



relocating downstream capsid sequence in closer proximity to the IRES may have affected as yet undefined replication elements in the 5'NTR. The presence of the capsid sequence containing SLV and SLVI was able to replicate more efficiently than the Del 15-171 genome despite showing a more defect in translation. Again it is surprising that replication was impaired, as the capsid gene was not believed to participate in replication. The reduced replication in the presence of the Stop 1,2,3,4 silent mutations indicate that it is the RNA structure and not protein that are affecting HCV replication. However, the genome Del 57-171 showed a significant decrease in replication ability, which might be due to capsid protein or additional RNA elements in the capsid gene.

To determine if the capsid protein can restore the replication defect found in the deletion genomes, we established an Huh-7.5 cell line selected to harbor the VEE replicon expressing the capsid protein. We generated the cell line using a VEE replicon with the pac gene and the HCV capsid gene behind two subgenomic promoters. After the selection process, notable cell death was found in cells transfected with the VEE replicon containing the pac and HCV capsid gene. Cells transfected and selected for a VEE replicon expressing the pac gene and gene for GFP did not undergo such dramatic cell death and cytopathic effect. This indicates that over expression of the

HCV capsid conveys toxic effects on Huh-7.5 cells. We confirmed that the HCV capsid protein was expressed and was the appropriate size by immunofluorescence and western blot analysis (data not shown). At 72H post electroporation, replication levels of the parent construct, H-JFH/Rluc in the Huh-7.5-VEErepC cell line was found to be several logs below replication levels in Huh-7.5 cells, possibly due to adaptations in the cellular environment allowing replication of the VEE replicon or perhaps due to a down regulation of replication in the presence of large amounts of capsid protein. Expression of capsid protein *in trans* did not restore replication for the genomes harboring deletions in the capsid gene. The Del 15-171 genome again was severely debilitated for replication and showed no benefit from the capsid expressed in the VEErepC cell line (Figure 5.23). As observed in the Huh-7.5 cells, replication was depressed for the Del 57-171 and Del 57-171/Stop 1,2,3,4 genomes in the Huh-7.5-VEErepC cell lines. Again, the silent mutations in SLV and SLVI contributed to a further reduction in replication even in the presence of capsid protein *in trans*.

Given the altered state of the Huh-7.5 cell selected with the VEErepC replicon and the reduced level of replication detected even in the parental genome, we decided to test for capsid complementation by delivering the HCV genome and VEE replicon simultaneously. We combined 1 µg of *in*

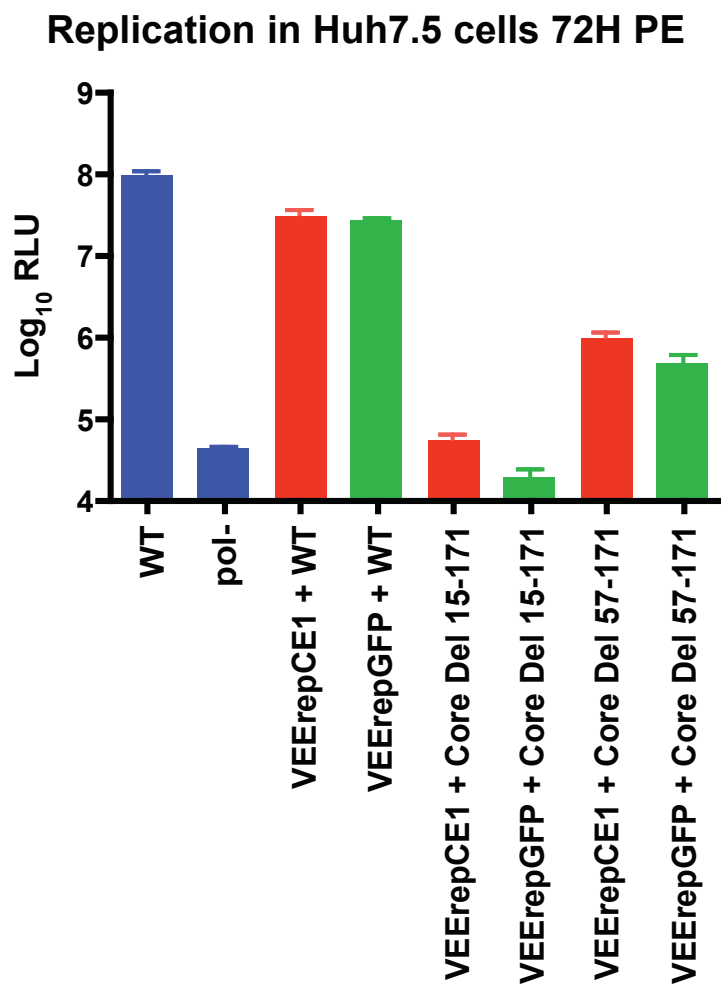


Figure 5.23. Coelectroporation of VEE replicon expressing HCV capsid and 15 amino acids of E1 can slightly improve replication of HCV genomes harboring deletions in the capsid gene. Replication is measure by luciferase expression at 72H post electroporation for two RNAs plated in duplicate wells. Averages are shown.

*vitro* transcribed RNA for a VEE replicon expressing the capsid protein and the first 15 amino acids of E1 with 2 µg of HCV genomic RNA and coelectroporated RNAs into Huh-7.5 cells. To control for replication effects contributed by the VEE replicon, we included a VEE replicon expressing GFP instead of the capsid protein. Replication was measured by luciferase expression 72H post electroporation. Replication in the presence of a VEE replicon either expressing the capsid protein or GFP, resulted in a reduction of the wild type HCV genome to levels 33% and 29% as compared to replication of HCV alone (Figure 5.23). The genomes Del 15-171 and Del 57-171 showed impaired levels of replication as previously observed; however replication was improved in the presence of expression of the capsid protein by the VEE replicon. For the Del 57-171 genome, a 200% increase in luciferase signal was observed in the presence of capsid protein. Replication of the Del 15-171 genome improved 277% with the capsid protein expressed by the VEE replicon, although the overall replication levels were very low. As compared to levels with the VEErepGFP, the wild type genome showed slightly improved replication by 113% when the capsid protein was provided *in trans*. This result indicates that the capsid protein acted in a stimulatory manner to improve HCV replication.

We next asked whether providing the capsid protein *in trans* could complement virus production by HCV genomes with the capsid gene deleted. After 72H post electroporation, media was collected and transferred onto naïve Huh-7.5 cell or the Huh-7.5 cell line containing the VEErepC replicon. At 48H, we measured luciferase to indicate infection, translation, and replication of packaged RNA produced from cells coelectroporated with the VEE replicon expressing the capsid protein or GFP and HCV genomes with the capsid protein removed. The Huh-7.5 VEErepC cell line was included to determine if infection, translation or replication is affected when the capsid protein is already abundantly present in the cell. None of the HCV genomes with deletions in the capsid gene could be *trans* complemented for virus infectivity (Figure 5.24). No virus infectivity was detected in either Huh-7.5 or the Huh-7.5 VEErepC cell line. The lack of detectable virus infectivity could be due to the inability of the capsid protein to package HCV RNA *in trans* or simply reflect the low level of virus replication in both the virus producer cells and the acceptor cells.

We did discover that the capsid protein expressed *in trans* can increase virus infectivity of the replication uncompromised genome, H-JFH/Rluc. The parental genome, H-JFH/Rluc was able to release virus and infect both Huh-7.5 cells and the Huh-7.5-VEErepC cell line, although infection of

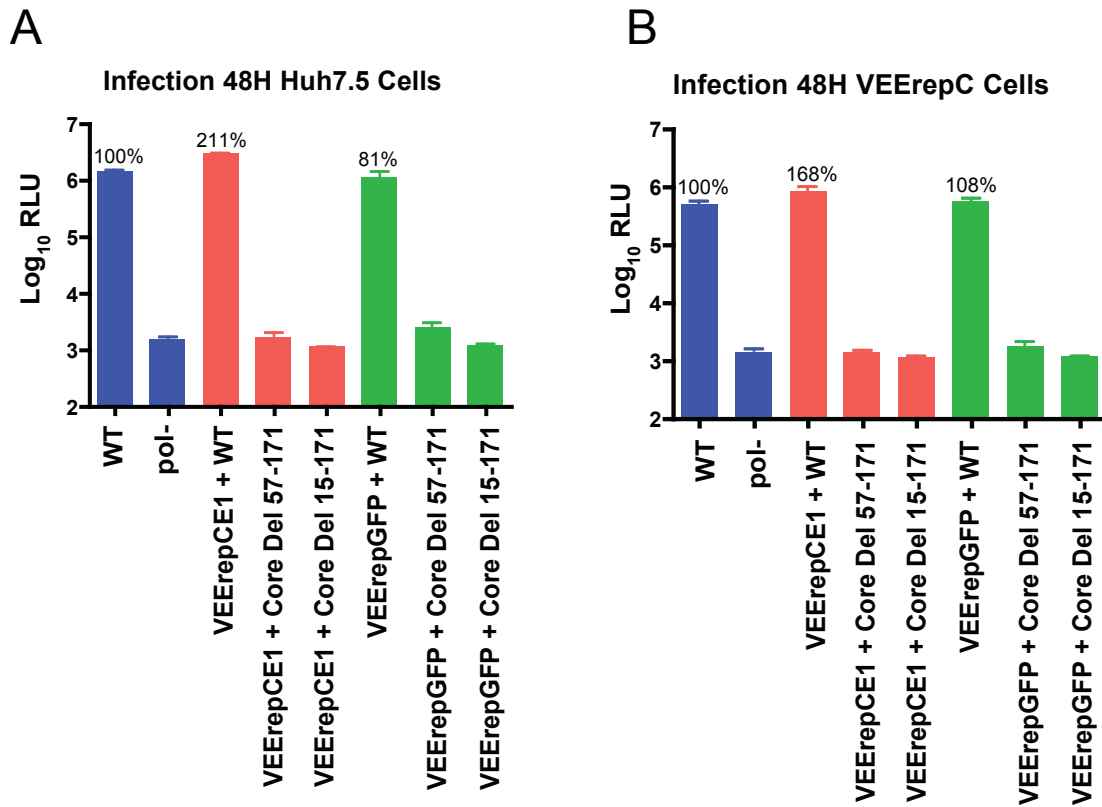


Figure 5.24. Coelectroporation of VEE replicon expressing HCV capsid and 15 amino acids of E1 increases virus infectivity in A. naive Huh-7.5 cells and in B. a Huh-7.5 cell line selected to harbor the VEE replicon expressing HCV capsid protein. Infectivity is measure by luciferase expression 48H post infection.

the latter was reduced by 63% (Figure 5.24). Cells originally cotransfected with the wild type genome and VEErepGFP RNA could also produce virus that infected both cell lines, and similarly was reduced by 60% in the Huh-7.5-VEErepC receptor cells. Infectivity of Huh-7.5 cells was 81% of that observed from cells transfected with the wild type RNA alone and may be a consequence of the lower amount of accumulated RNA in the producer cells due to the lower replication level (29% of wild type). An increase in infectivity was detected in both receiving cell lines when infected with virus produced from the cotransfection of wild type HCV RNA and VEErepC RNA. Cells harboring both wild type HCV RNA and the VEE replicon expressing HCV capsid produced 211% more virus infectivity in Huh-7.5 cells than cells expressing HCV RNA alone. An increase of 168% over wild type alone was also measured in the VEErepC infected cells. This increase in infectivity is a direct result from an increase in the amount of capsid protein provided *in trans* in the virus producing cells. Replication levels of HCV RNA in the presence of VEE-expressed HCV capsid protein were reduced to 33% of HCV RNA replication alone. This reduced replication level is comparable to the level measured for HCV RNA replicating with VEE expressed GFP, yet infectivity was below that of wild type. An increase in infectivity of the VEErepC receiving cells was only observed for

virus produced in the presence of excess capsid, indicating no benefit from the capsid protein in the receiving cells.

***Translation and replication kinetics of HCV in the presence of HCV structural proteins expressed in trans.***

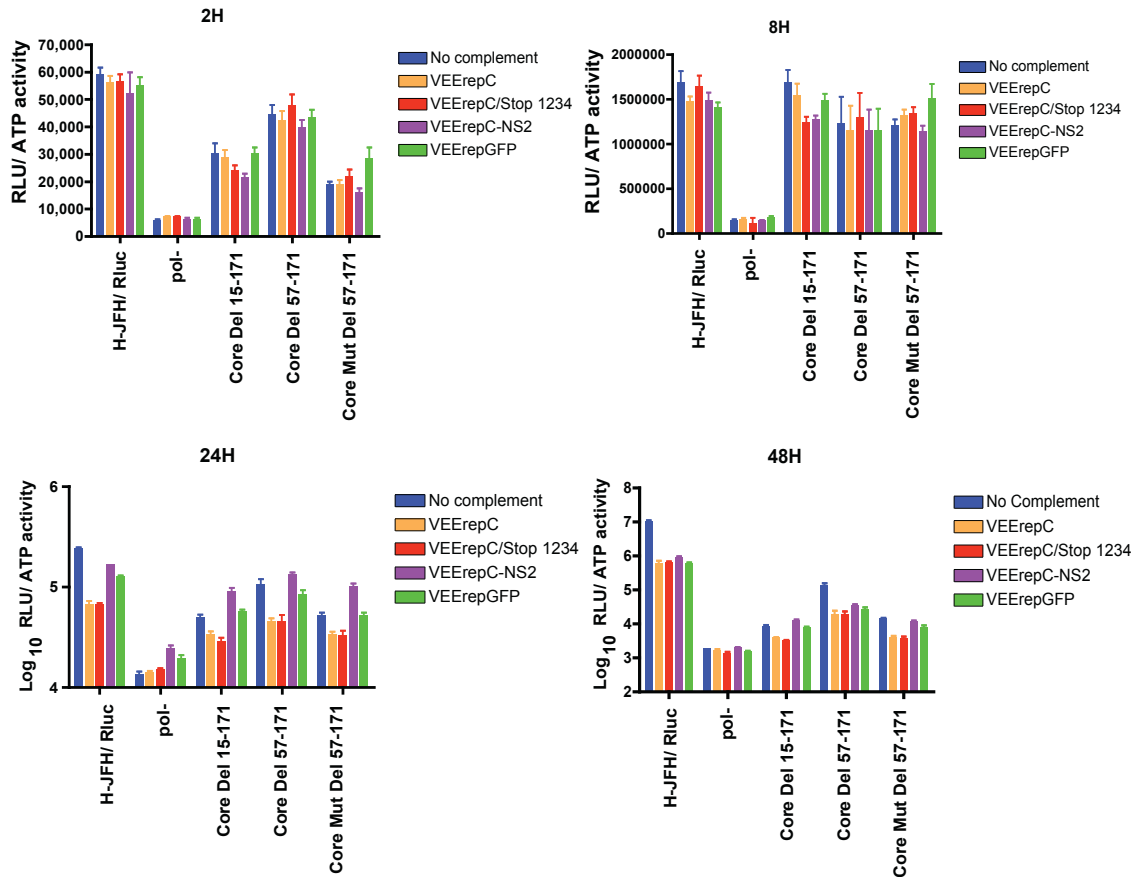
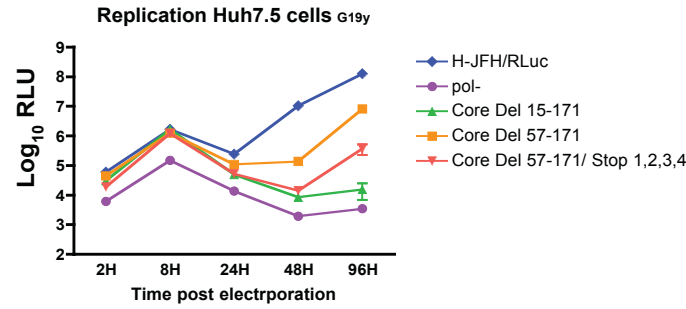
Previous studies describing reduced translation in the presence of the capsid gene or protein were conducted using *in vitro* translation assays or transient expression systems in cell culture in the absence of a replication competent HCV genome. We have found the capsid protein to actually stimulate replication for both wild type and genomes harboring deletions in the capsid gene when expressed *in trans*. In addition, cells expressing the capsid protein from a VEE replicon increased virus infectivity from cells transfected with the replication competent, wild type genome. To further define the enhanced effect by the capsid protein, we measured translation and replication of HCV genomes containing deletions in the capsid gene (Del 15-171, Del 57-171, and Del 57-171/ Stop 1,2,3,4) in the H-JFH/ Rluc parent when cotransfected with VEE replicons encoding the capsid protein, structural proteins capsid-NS2, or a GFP control. Because the VEE replicon generates an independent mRNA transcript from each of its subgenomic promoters, the RNA sequence of the capsid gene is expressed in great



abundance *in trans*. Since we have shown the capsid structures of SLV and SLVI affect HCV replication *in cis*, we wanted to determine if the RNA structures of the capsid gene or disrupted RNA structures contribute an effect *in trans*. We, therefore, introduced the silent mutations Stop 1,2,3,4 into the VEErepC replicon.

We coelectroporated 3µg of VEE replicon RNA with 1 µg of HCV RNA, both generated by *in vitro* transcription, into Huh-7.5 cells and measured luciferase expression over time (Figure 5.25). Due to cell toxicity associated with expression of the VEE replicons in Huh-7.5 cells, we normalized Renilla luciferase levels to cellular ATP activity indicating cell viability. At 2H post electroporation, a difference was observed in translation between HCV genomes with deletions in the capsid gene (Figure 5.25). This difference did not significantly vary with *trans* expression by the VEE replicon. The genome Del 15-171 was, on average, 48% of wild type and below the level of Del 57-171, which was reduced to 78% of wild type (Figure 5.25). The reduced translation of Del 15-171 is not consistent with translation measured in the *in vitro* translation assay showing only a slight inhibition of 88%, but does mirror the repeated replication defect observed for this genome. The reduced level of translation by genome Del 57-171/ Stop 1,2,3,4 (37% of wild type) is very consistent; again demonstrating a

Figure 5.25. (Next page). Transcomplementation of HCV genomes containing deletions in the capsid gene and/ or mutations in the RNA structures for replication over time. Coelectroporation of VEE replicon expressing HCV capsid, HCV capsid with Stop 1,2,3,4, HCV capsid to NS2, GFP, or no complement. Translation and replication is measured by luciferase expression normalized to cellular ATP activity.



functional significance of SLV and SLVI. The pol- genome displayed especially low levels of luciferase expression. Although each transfection was performed using two RNA preparations and luciferase measured for two independent wells for each combination, the low levels of pol- indicate a technical difficulty with these RNA preparations.

By 8H, all HCV genomes had reached similar levels of luciferase expression (Figure 5.25) indicating the initial differences noted for translation was not due to variations in transfection. Notably, the Del 15-171 genome peaked to levels within 0.5% difference of the wild type genome, and the genomes of Del 57-171 with and without mutations for SLV and SLVI were within a 1.6% difference. All combinations expressed luciferase within 38% of each other.

Replication differences became distinct by 24H. For the wild type genome, replication with any VEE replicon decreased luciferase expression, but a significant inhibition was caused by the VEErepC replicons, with and without RNA mutations (Figure 5.25). For the Del 15-171 genome, a marked improvement, above wild type, was found in the presence of C-NS2. Trans expression of the structural proteins also improved replication for the Del 57-171 and Del 57-171/ Stop 1,2,3,4 genomes. Replicon expression of the capsid protein alone hindered replication of all HCV genomes.

For the wild type HCV genome, replication differences became less apparent in the presence of all VEE replicons, yet remained one log below replication alone at 48H. Expression of the C-NS2 proteins led to an increase in replication by 153% above replication with expression of the capsid alone or GFP by the VEE replicon. A similar pattern was measured for the HCV genomes containing deletions in the capsid gene. Replication was suppressed when occurring with the capsid protein expressed *in trans*, yet an increase was observed for all HCV genomes when the C-NS2 proteins were present. For Del 15-171, expression of C-NS2 led to enhanced replication of 164% above levels measured for replication in the presence of GFP. HCV genomes Del 57-171 and Del 57-171/ Stop 1,2,3,4 improved replication by 129% and 147% when replicating with C-NS2 as opposed to GFP.

In this study, we monitored translation and replication of HCV genomes harboring deletions in the capsid gene and compared the ability of the capsid protein or the capsid protein with structural proteins to influence replication *in trans*. A clear translation defect of input RNA was detected at 2H for genomes with deletions in the capsid gene and was more pronounced when the RNA structures of SLV and SLVI were disrupted. We did not find a significance difference in translation when the capsid protein was expressed, alone or with C-NS2, in either the wild type or capsid deletion

genomes indicating the capsid protein did not influence translation *in trans*. Since luciferase levels for all HCV genomes reached similar levels by 8H, the initial differences in translation were not lethal, but a compromise in the efficiency of an early event in translation or RNA localization to the host cell's translation machinery. As replication accelerated for the wild type genome, a lag was observed for genomes with the capsid deletions. We detected additional repression in the presence of the capsid protein *in trans* for all HCV genomes. The capsid protein suppressed luciferase activity when expressed by both VEErepC and VEErepC/Stop 1,2,3,4 indicating no effect *in trans* by RNA structures of the capsid gene. However, HCV replication in the presence of the capsid protein in the context of E1-NS2 was significantly more efficient for all HCV genomes. Either the capsid protein could not negatively interfere when expressed with the downstream structural proteins or E1-NS2 can positively influence HCV replication *in trans*. Expression of the C-NS2 proteins *in trans* improved luciferase activity to or above wild type levels for Del 15-171, Del 57-171, and Del 57-171/ Stop 1,2,3,4. While expression of the capsid protein alone continued to suppress HCV replication and/ or translation, the presence of C-NS2 significantly increased HCV activity.

***Transcomplementation of point mutations in the capsid gene with the structural proteins can increase virus infectivity.***

HCV genomes with deletions in the capsid were unable to be trans complemented for virus infectivity, probably a consequence of the significant reduced replication. To determine if the HCV structural proteins could trans complement virus infectivity, we tested HCV genomes harboring point mutations that convey no or little compromise in replication. We chose mutations previously generated to study the RNA structure of SLVI and their effects on HCV replication (Chapter 4) in the context of the H-JFH/Rluc genome. The genome Base Right contains three point mutations that would disrupt the basal stem and shows a modest replication defect. The Open 3,4 mutations convey a more severe effect on replication and are predicted to disrupt the apical stem. The Mid Right Silent, Mid Right, and Mid Left mutation sets disrupt the middle stem region, while the genome with Mid Right/ Left mutations attempts to restore base pairing of the middle region. The Mut Loop SLVI introduces three base changes at in the loop of SLVI. Some of the mutations changed the amino acid sequence of the capsid protein, and the residue substitution is listed next to the mutant genome. Therefore, we compared replication in the presence of the VEE replicon expressing the C-NS2 structural proteins, VEErepC-NS2, and a

VEE replicon encoding GFP as a control, VEErepGFP to determine if the effect on replication was due to the amino acid difference or the altered RNA structure and if replication could be influenced *in trans*.

To reduce the toxicity conveyed by the VEE replicon, we coelectroporated a lower amount of VEE replicon RNA, 0.5 ug, with 2 ug of HCV RNA. At 48H post electroporation, we again observed a reduced ability of all HCV genomes to replicate in the presence of the VEE replicon (Figure 5.26). The genomes with mutations for Open 3,4 and Mid Right Silent showed the most pronounced replication defect, as previously, while the other mutations permitted replication within one log of wild type. An enhancement of replication was again observed with the expression *in trans* of C-NS2. In the presence of VEErepC-NS2, genomes with mutations for Open 3,4, Mid Left and Mid Right/Left showed an increase in luciferase activity 145% above levels measured for VEErepGFP (Figure 5.26).

Genomes with mutations Mid Right Silent and Mid Right improved replication by 115% with the expression of C-NS2 as opposed to GFP. The increase in activity as a result of C-NS2 expression *in trans* did not correlate with a change in the amino acids of the capsid protein from the HCV genome.



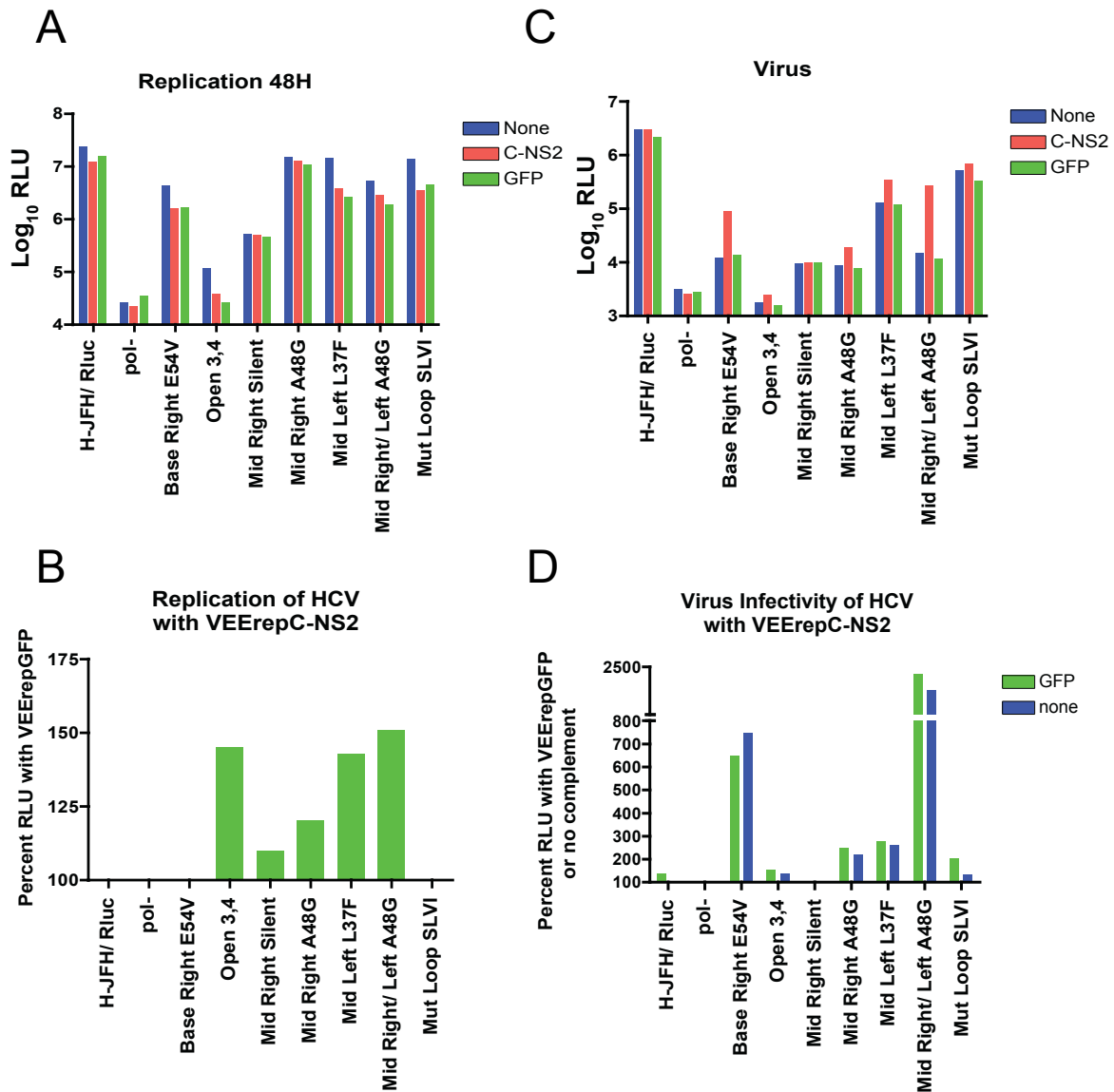


Figure 5.26. Transcomplementation of HCV genomes containing point mutations in SLVI and amino acid changes in the capsid protein. A. Replication and C. virus infectivity are measured by luciferase at 48H. B. The percent of luciferase activity of activity measured when cotransfected with a GFP control. D. The percent of luciferase activity of activity when cotransfected either with C-NS2 or nothing.

We measured the ability of C-NS2 to trans complement virus infectivity of the mutant HCV genomes by incubating naïve Huh-7.5 cells with media harvested from transfected cells after 48H. After 48H incubation, the wild type genome produced the highest level of infectivity (Figure 5.26). Coexpression of the wild type genome with C-NS2 led to a moderate increase of virus infectivity, 140% of the infectivity produced when coexpressed with GFP (Figure 5.26). The Open 3,4 genome failed to establish an infection and could not be trans complemented, potentially a result of the low replication capacity of this genome. Despite modest levels of replication, the Mid Right genome could launch an infection, yet without any benefit when coexpressed with C-NS2. We detected a dramatic increase in infectivity for the genomes Base Right and Mid Right/ Left when both replicated in presence of trans expressed C-NS2. In the presence of VEErepC-NS2, the genome Base Right did not show an increase in replication; however, virus infectivity was significantly increased, to over 600% of levels when replicating with C-NS2 versus GFP. Replication of the genome Mid Right/ Left was complemented by C-NS2 with a 150% increase over GFP, and yet we detected the largest improvement in infectivity by coexpression of this genome with C-NS2 in the virus producing cell. Replication of Mid Right/Left with VEErepC-NS2 resulted in an >2000%

increase in virus infectivity as compared to cells replicating the genome alone or with GFP. Both genomes Base Right and Mid Right/ Left harbor amino acid changes in their capsid proteins. The ability of both of these genomes to be trans complemented with the wild type capsid and structural proteins indicates the reduced level of infectivity was due to the amino acid mutation present in the mutant genomes. The genome Mid Left and genome Mid Right also harbor amino acid changes, L37F and A48G respectfully, in their capsid protein. Both show an increase in virus infectivity when replication and virus production occurred in the presence of replicon expressed C-NS2. The genome Mid Right/ Left contains both L37F and A48G mutations in the capsid protein and is the genome with the highest success of transcomplementation by authentic C-NS2.

***Mutations alter translation and replication at the single cell level.***

We have examined the effects of mutations in SLVI by measuring the replication capacity at a population level. Replication was launched by electroporating equal amounts of HCV RNA into identical number of Huh-7.5 cells; however the level of replication from HCV genomes containing mutations in SLVI was consistently found to be lower at early and late times after transfection. It was unclear if the lower level of replication was due to

fewer cells capable of replicating the mutant HCV genome or if replication was occurring at a reduced level in the population as a whole. Visualizing individual HCV positive cells by immunohistochemistry (IHC) and flow cytometry are routine assays used to quantitate differences between cells, but may not have the sensitivity to measure low levels of HCV antigen. After two days post electroporation of the mutant HCV RNA, it is apparent that, compared to the wild type, fewer cells are positive for HCV antigen by IHC (Figure 5.27). The presence of darker and light shades of antigen staining indicates variation of HCV protein expression in the population, but it is difficult to quantitate the degree of difference. Flow cytometry can measure a range of positive signal for individual cells, over a background of nonspecific immunoglobulin staining. Transfection with a wild type genome leads to a 70% positive population of Huh-7.5 cells (Figure 5.27). The wild type genome produces a distinct population of positive cells with a narrow range of antigen expression level. Analysis of the histogram for cells transfected with a SLVI mutant HCV RNA shows fewer positive cells and weaker antigen expression, indicated by a shoulder of positive cells near the background of nonspecific IgG2a controls. Flow cytometry of Huh-7.5 cell lines selectively harboring HCV replicons and bicistronic genomes expressing a drug selectable marker where, by definition, all cells should be

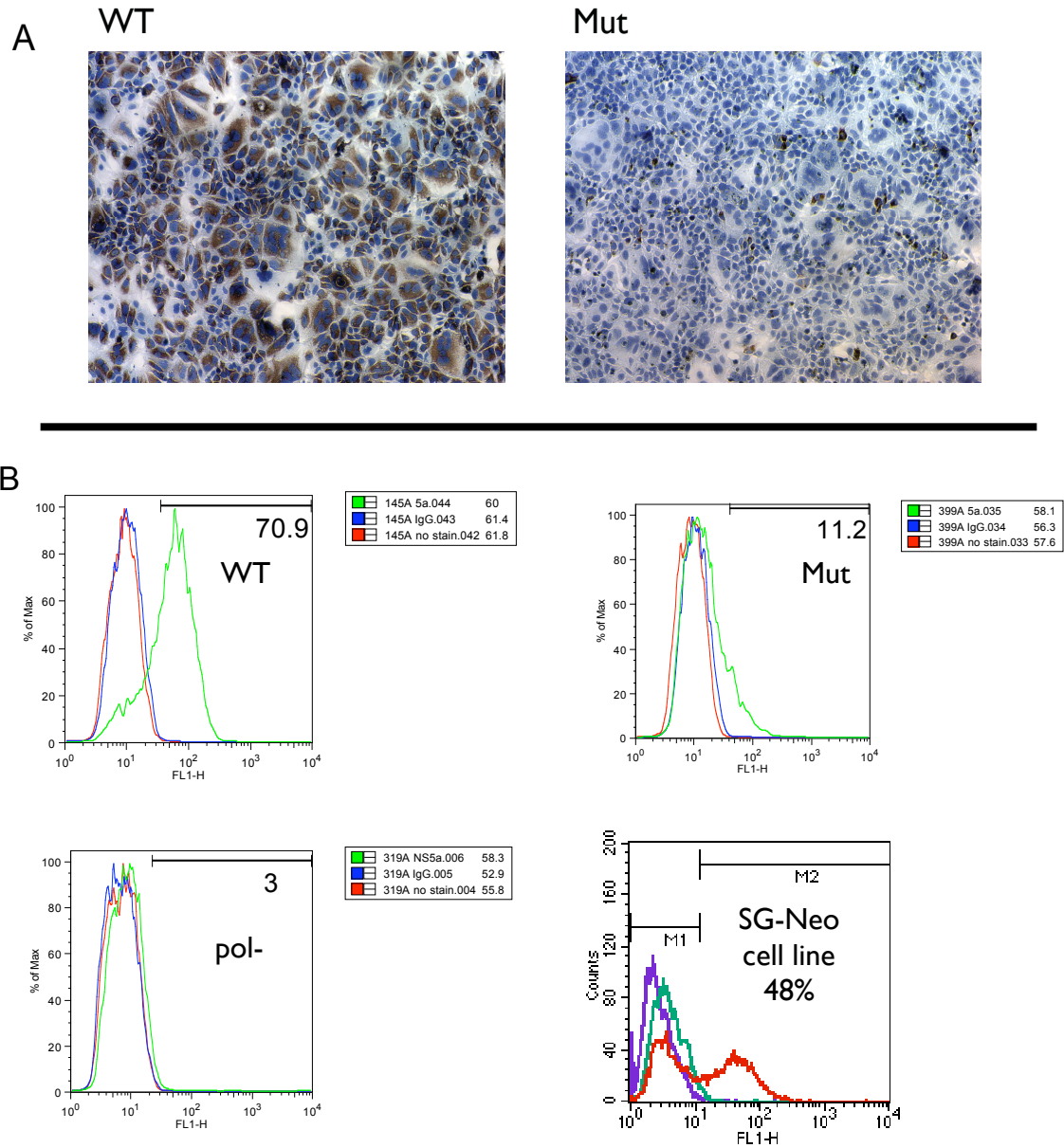


Figure 5.27. Fewer transfected Huh-7.5 cells with mutant genome are positive for HCV replication. A. IHC of NS5A positive Huh-7.5 cells after 2 days post electroporation. B. Flow cytometry of NS5A positive Huh-7.5 cells

positive for HCV antigen, only yield a 50-70% positive population. This demonstrates the low level of sensitivity offered by flow cytometry to measure HCV protein expression. Both IHC and flow cytometry require a sufficient level of antigen expression for analysis, preventing early measurements of translation and replication differences.

Using a luciferase reporter, with high sensitivity and wide dynamic range, early differences of translation and replication can be monitored. We compared a wild type, pol-, and genome harboring the Stop 1,2,3,4 mutations in SLV and SLVI in the parent genome of H-JFH/ Rluc. Huh-7.5 cells were electroporated with HCV RNA and serially diluted with multiple replicates to allow comparisons of HCV activity to a single cell level and evaluate the heterogeneity of the population. We hypothesized that if the entire cell population could replicate the mutant HCV, yet at a low level, the luciferase signal would decline with the number of cells in a linear relationship. If the population were homogenous with respect to HCV replication, no significant differences between replicates at each dilution would be observed. If only a minor population of cells was capable of replicating the mutant genome, the luciferase signal would abruptly reach the background level as the small, positive cell population was diluted. Differences among the replicates would be apparent at the end point

dilution, as some wells received this minor population of cells capable of mutant replication while other did not.

In a master plate, electroporated cells were diluted two fold and 8 wells of a 96 well plate seeded at a beginning concentration of 40,000 cells per well. To ensure positive luciferase activity can be detected at a single cell at later time points, additional dilutions were performed for cells transfected with the wild type genome, seeding with starting dilutions of 4,000 cells, 800 cells, and 400 cells per well. Cells electroporated with no RNA were diluted and seeded in the reverse order so that each well contained the same number of total cells.

At 4H a linear association of luciferase expression and number of cells was observed for the wild type, mutant, and pol- genomes indicating equivalent levels of transfection and translation of the input RNA (Figure 5.28). For the mutant and pol- genomes, luciferase activity could be detected to the last dilution of 39 cells/well. The wild type genome showed slightly less luciferase activity in wells seeded with fewer cells and the final positive detection of translation occurred at 125 cells/well. The theoretical line shows a two-fold dilution starting at the 40,000 value and is included as an orientation guide. By 12H, luciferase levels fell one log as input RNA and luciferase protein degrades. The reduction of luciferase activity is equivalent

for all three HCV genomes as well as the final positive signal at 313 cells/well (Figure 5.28). At 24H, replication increased for the wild type genome. Dilutions of cells previously below the limit of detection for luciferase activity are now positive for replication. The mutant genome shows a reduced level of luciferase expression in both the cells positive for luciferase expression and in the number of positive cells. Luciferase activity indicating cells replicating the mutant RNA is positive for only 5000 transfected cells/well while the wild type RNA could replicate with a minimum of 125 cells (Figure 5.28). This evidence suggests that despite being able to launch translation with equivalent amounts of HCV RNA, only a small proportion of cells were able to establish replication of a mutant genome. At the end point of dilution, the lower luciferase activity observed for the mutant genome indicates that replication was attenuated.

At 72H, the wild type genome increased in luciferase expression. In all replicates of the first three wells seeded with a higher number of transfected cells, an additional increase is observed in luciferase expression. This could be due to a cooperative effect for HCV replication exuded from cells replicating at a particular level. Cells harboring the mutant genome also showed an increase in luciferase activity, but an average 2 logs below wild type. Where replication could be detected to 8 transfected cells in the wild



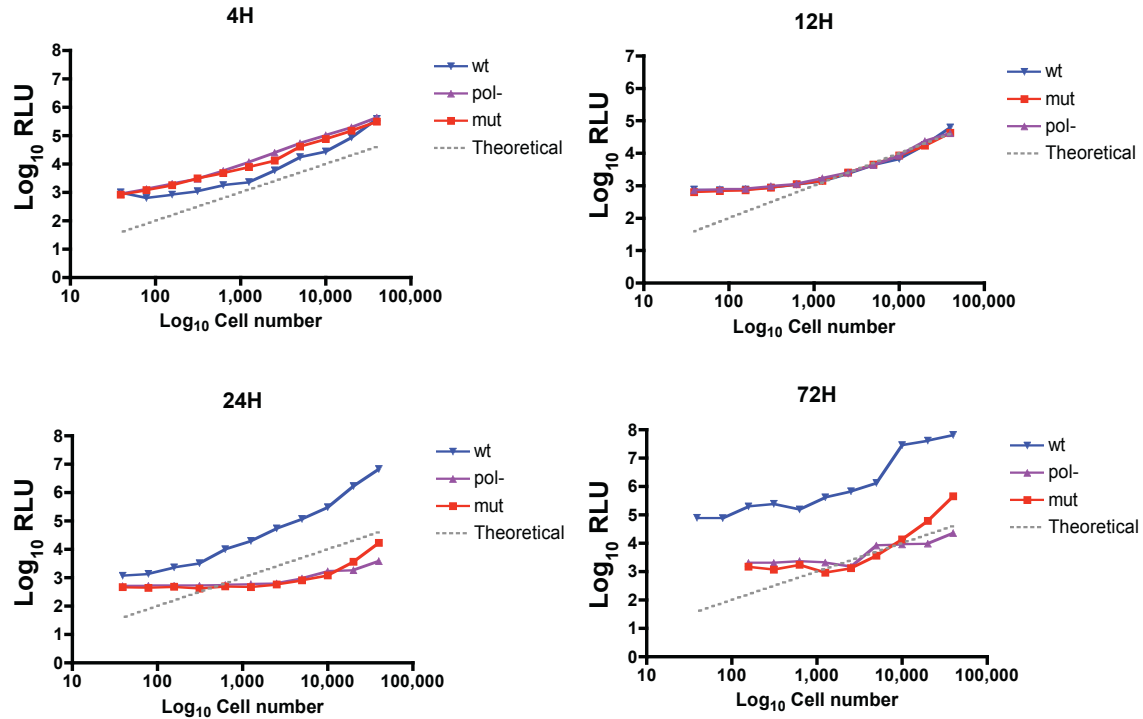


Figure 5.28. Translation and replication kinetics as measured in an end point dilution assay. Luciferase activity was measured in 8 replicate wells for dilutions of transfected cells electroporated with wild type, mutant SLV and SVLI, or pol- HCV genomes over time. A theoretical line is provided for orientation.

type, the last point of detecting a positive signal for cells harboring the mutant genome is 313 cells. At this time point, the increase in number of positive cells could be due to virus spread increasing the amount of luciferase expression in the population.

Comparing the number of cells capable of HCV replication over time, it is observed that more transfected cells are required to launch replication of a genome harboring mutations in SLV and SLVI. The end point was determined by the dilution having 50% of the wells positive for luciferase activity. Averaging the luciferase activity for the positive wells determines the RLU at the end point dilution. At 24H, the wild type genome required 125 cells for an average of 2005 RLU (Figure 5.29). The mutant genome required 5000 transfected cells for an average luciferase signal of 916 RLU. The average background luciferase was 700 RLU. At 72H, the wild type genome required only 50 transfected cells to express an average luciferase unit of 5714 RLU. The mutant genome required 5000 transfected cells to express an average luciferase unit of 2316 RLU.

This result implies either that a particular type or state of host cell is required to replicate a mutant HCV RNA or the mutant genome establishes replication very infrequently. For a cell to harbor HCV replication at 72H, the mutant genome required 100 times more transfected cells than the wild

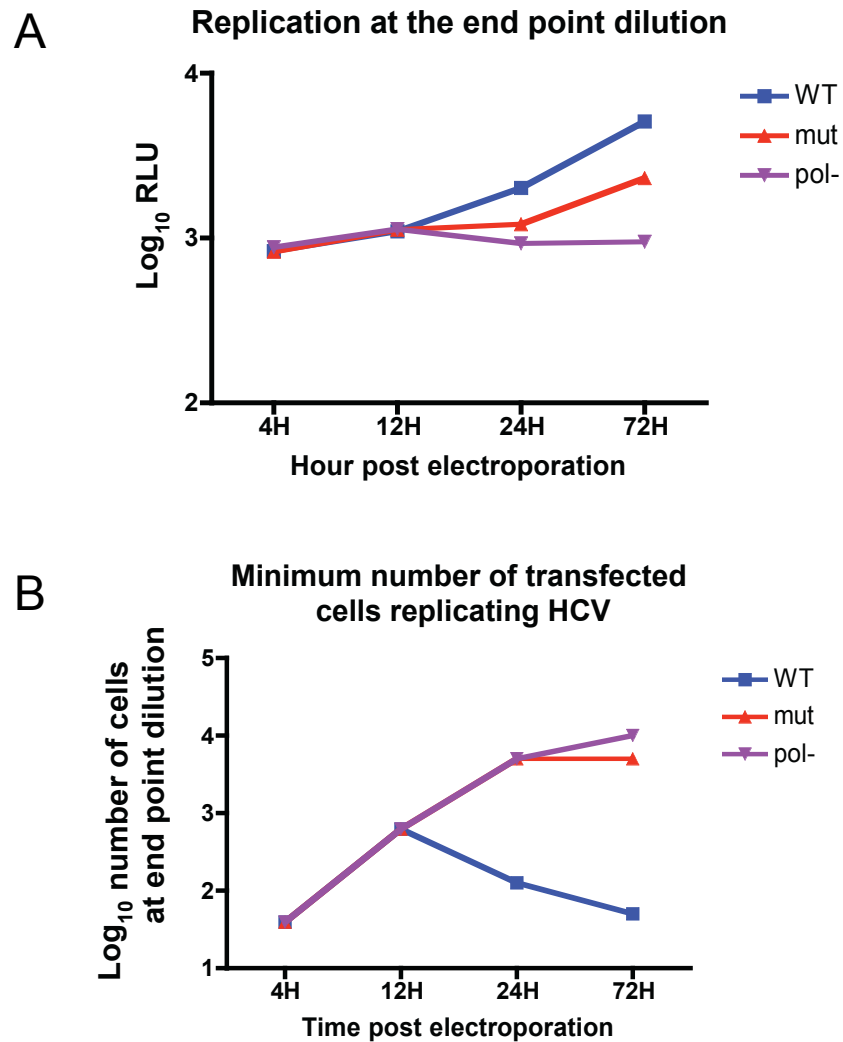


Figure 5.29. Replication of the mutant HCV genome occurs less efficiently and in fewer transfected cells. A. Luciferase activity at the end point dilution. B. The number of transfected cells required to give a positive HCV signal at the end point dilution.

type. Either a particular host cell capable of mutant HCV replication is present in 1% of the population or mutant replication is launched 100 times less frequently than the wild type. The contribution of the host cell to mutant HCV replication is unknown, but whatever the function, it does not completely restore replication of the mutant RNA to wild type levels, as the average luciferase expression is reduced by half.

***The functions of SLV and SLVI are position-dependent.***

We have shown mutations in the RNA structures of SLV and SLVI result in a severe reduction in replication. These mutations also contribute to lower efficiency of translation as measured in transfection of Huh-7.5 cells and using *in vitro* translation assays of rabbit reticulocyte and Huh-7 cell extracts. These RNA structures are located in close proximity to the IRES and 5'NTR replication elements, and therefore we wanted to determine if the action of SLV and SLVI was dependent on the location within the genome. Previous translation and replication studies have compared the effects of mutations on SLVI and SLVI in the context of a monocistronic genome. To test the role of these structures outside of the context of the 5'NTR, we used a bicistronic genome of the genotype 1a H77 sequence with four adaptive mutations: Q1067R, K1691R, K2040R, S2204I. This bicistronic genome, H-

neo/RRRI, contains the H77 5'NTR and the first 14 amino acids of the capsid gene fused to the neomycin phosphotransferase gene followed by the encephalomyocarditis virus (EMCV) IRES which drives translation of the complete H77 genome, capsid-NS5B. The mutations of Stop 1,2,3,4 were introduced into the capsid gene positioned behind the EMCV IRES. *In vitro* transcribed RNA was electroporated into Huh-7.5 cells and replication measured by Real Time qRT-PCR and selection of G418 resistant colonies. We found no defect in the ability of the H-neo/RRRI Stop 1,2,3,4 genome to replicate (Figure 5.30). Equivalent numbers of G418-resistant colonies were selected for both the parental genome and the genome with Stop 1,2,3,4 (Figure 5.31). In addition, HCV RNA levels in transfected Huh-7.5 cells were slightly above levels of the parent genome at time points indicating active replication (Figure 5.30). These results demonstrate that the mutations of Stop 1,2,3,4 do not convey a replication defect when located apart from the 5'NTR and translation is dependent on the EMVC IRES.

***Independent expression of the replicase restores the replication for genomes with mutant SLVI.***

The mutations in SLV and SLVI had no effect on replication when expression of the HCV polyprotein was driven by a heterologous IRES. We

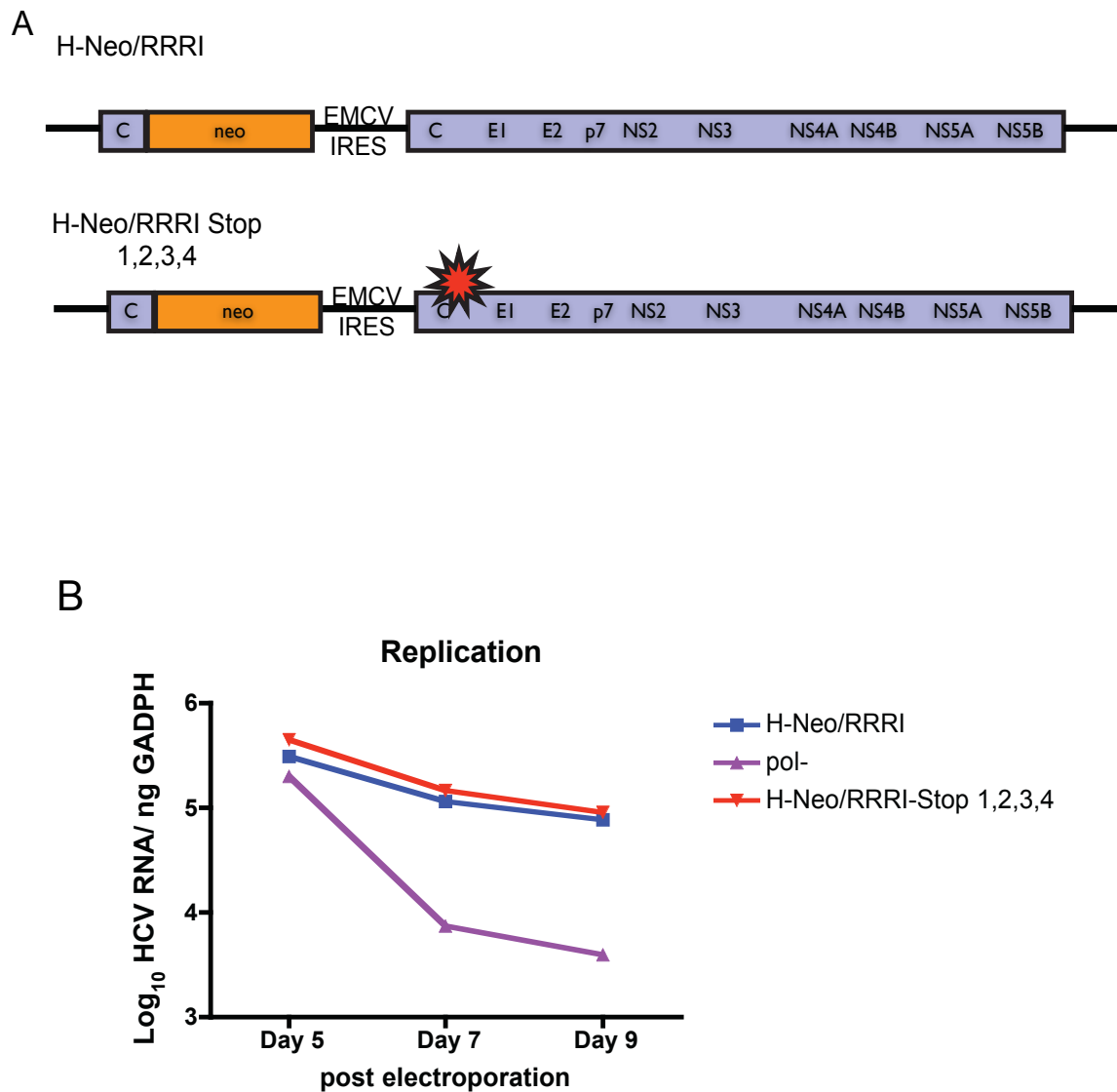


Figure 5.30. Mutations in SLV and SLVI do not confer a replication defect when positioned behind the EMCV IRES in a bicistronic genome. A. Diagram of the bicistronic genome B. HCV RNA levels measured by Real Time qRT-PCR and normalized to cellular GADPH levels.

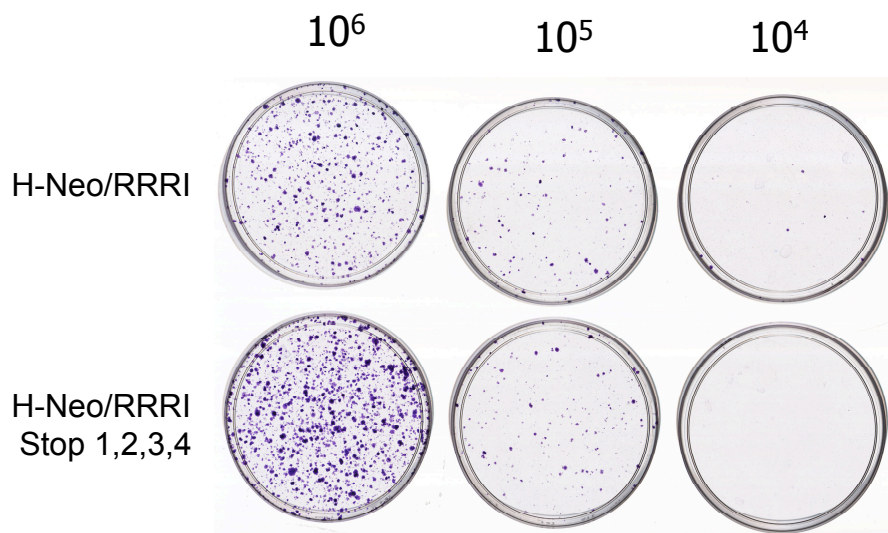


Figure 5.31. Mutations in SLV and SLVI do not reduce replication when positioned behind the EMCV IRES. G418 resistant colonies indicating HCV replication. The number of transfected cells is indicated above each plate.

next wanted to examine if the mutations for SLV and SLVI could alter replication when located in their original position, but separate any effects of translation by expressing the replicase proteins from a heterologous IRES. We used a bicistronic genome consisting of the H77 sequence from the 5'NTR thru NS2 followed by the EMCV IRES and NS3-3'NTR sequence from JFH-1. We introduced the mutation sets of Stop 3,4 and Open 3,4 into this parental genome, H-EI-JFH. Both mutation groups reduced replication to 2% of wild type in the context of the monocistronic genomes of genotype 1a, H/L+I and in the genotype 1a/2a chimera, H-JFH/ Rluc. Replication was measured by Real Time qRT-PCR of HCV RNA over time. We distinguished between replicating HCV RNA and input RNA by comparing RNA levels with that of the pol- control. Given the robust replication ability of the JFH replicase, differences become noticeable after 20H, where the pol- RNA levels begin to decline (Figure 5.32). HCV RNA was quantitated and normalized as a percent of RNA at 4H to control for any differences in transfection. RNA levels for the parent genome began to rise after 14H and increased ~ 2 logs by 72H. RNA levels for the Stop 3,4 and Open 3,4 bicistronic genomes did not initially increase and were at least 50% below wild type. After 30H, RNA levels increased, and finally approaching wild type levels at 72H. The restoration of replication indicates that the mutations



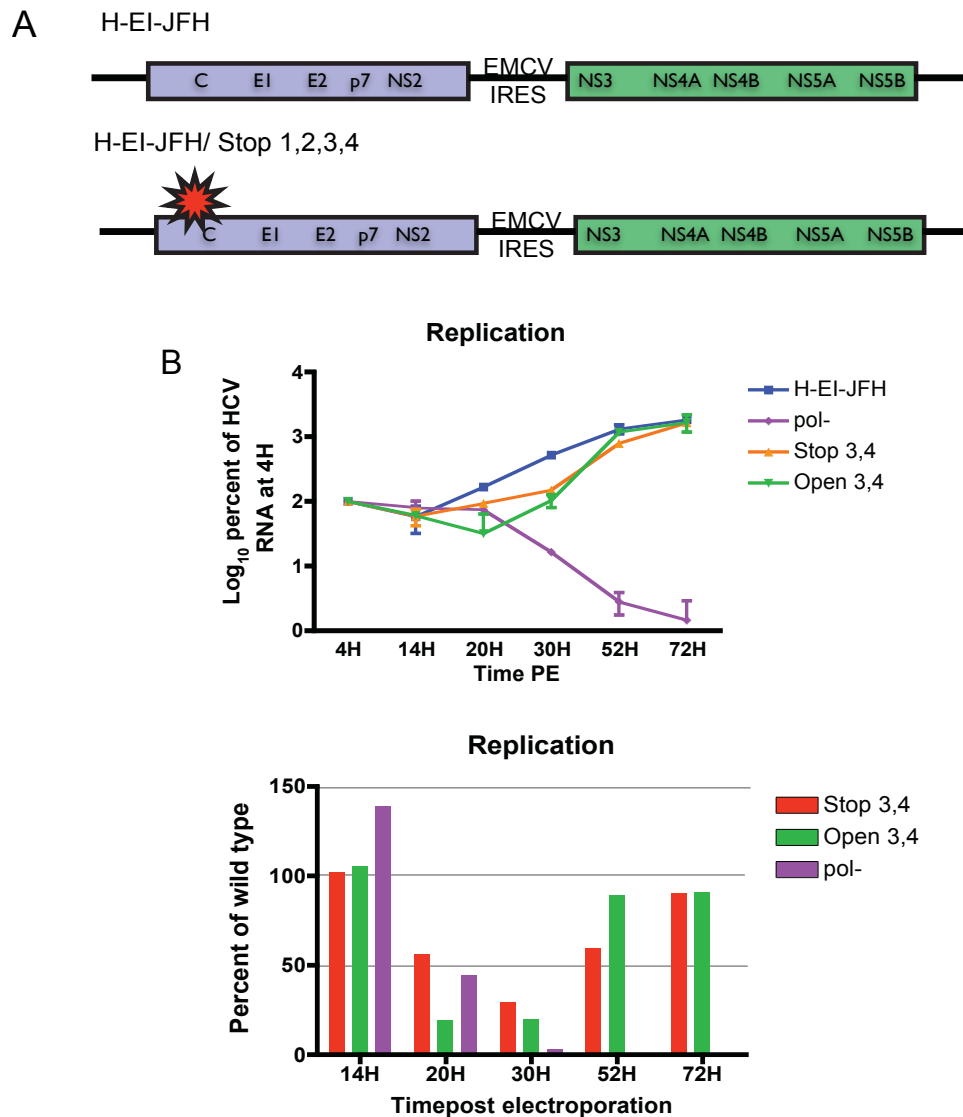


Figure 5.32. Replication is not reduced with mutations in SLV and SLVI when translation of the replicase is independent of the HCV IRES. A. Diagram of the bicistronic HCV genomes separating the replicase from the HCV IRES. B. HCV RNA measured by Real Time qRT-PCR. Values are normalized to the RNA level at 4H to control for transfection differences. C. Percent differences than wild type.

of SLVI are most likely acting at the level of translation leading to an insufficient amount of replicase expression. By separating the expression of the replicase proteins from the HCV IRES and downstream structures SLV and SLVI, replication could proceed with only a slight delay. This delay in replication detected between 20H and 52H could be a consequence of additional functions of the RNA structures in replication, such as affecting RNA association with the replicase or influencing the minus strand promoter. Although the structural proteins of C-NS2 are not required for replicase activity, a reduction in their expression by mutations in SLVI could also possibly reduce the efficiency of the replicase. By independently translating the replicase, the mutations in SLVI have little effect and replication can proceed to almost wild type levels.

### ***Discussion.***

#### ***Translation is compromised by mutations in SLV and SLVI.***

The capsid gene was previously thought to be unnecessary for HCV replication, yet we have shown mutations in RNA structures of SLV and SLVI contribute to a replication defect. Given the close proximity of the RNA structures to the HCV IRES, we tested if mutations in SLV and SLVI affected translation and thus led to reduced replication. Using a luciferase

reporter embedded within the NS5A gene of the polyprotein, we found that mutation of SLV and SLVI caused a reduction in translation, either in electroporated or lipid-mediated transfected Huh-7.5 cells. Since the luciferase insertion at NS5A does not allow replication, we were able to measure translation of only input HCV RNA.

Translation was reduced 2 -4 fold in the presence of mutations previously shown to confer a severe replication defect, Stop 3,4 and Stop 1,2,3,4 while mutations having no effect on replication, Stop 1,2, did not reduce translation. Despite the lower level of luciferase expression found at the early time points in the mutant genomes, translation could occur and luciferase expression eventually reached wild type levels. This lower efficiency of translation was also detected in cell free systems, where translation was launched by mutant genomes, but to levels 2-3 fold below wild type. The differences in translation were not due to preferential degradation of the mutant RNA.

Translation of the HCV genomes containing mutations in SLV and SLVI was also assessed using cell lysates derived from Huh-7 and specialized Huh-7.5 cells. Both types of cell lysates were able to initially translate the mutant HCV RNA to similar levels of wild type, but mutant luciferase levels failed to reach wild type activity. We found that translation

extracts from Huh-7.5 cells had a reduced capacity to translate even wild type genomes. Combining extracts from Huh-7 and Huh-7.5 cells, translation remained below levels found for Huh-7 cell extracts indicating an inhibitory factor present in the Huh-7.5 cells. The reduced capacity for Huh-7.5 cell extracts to conduct translation is in contrast to the enhanced ability of these cells to replicate HCV. This lower level of translation, although in a cell free system, may indicate a state where HCV replication benefits from less competition of the host cell translation machinery.

The secondary structure of the HCV IRES has been characterized and the required cofactors and contacts points defined using the minimal sequence containing nucleotides 40-372 <sup>39</sup>. The downstream sequence of the capsid gene including SLV and SLVI was not included in these studies and is also absent in studies of the assembly of the initiation complex. We tested mutations in the IRES used to define the domains responsible for binding the initiation complex of 40S, ternary complex of eIF2-Met-tRNA<sub>i</sub><sup>Met</sup>-GTP, and 60S in our translation and replication assays. We found the mutations of G266-268A and IIIId-deIE to contribute to more severe defect in both translation and replication than that observed for mutations in SLVI. The mutant G266-268A is unable to efficiently form the 40S complex and an early reduction in translation and severe replication defect was observed. In

contrast, the mutant IIIId-delE was reported to have high IRES activity with only a minimal reduction in translation. Despite being able to form the initiation complex, translation and replication were dramatically reduced. These results indicate that SLVI is not required for translation initiation, nor does it restrict assembly of the initiation complex, but may subtly influence the efficiency of initiation or affect subsequent steps in elongation or termination.

Translation of SLV and SLVI mutant genomes was equivalently reduced, 2- 4 fold relative to wild type, in rabbit reticulocyte, Huh-7 and Huh-7.5 extracts, and in transfected Huh-7.5 cells. Translation was initiated in the mutant genomes, but proceeded with lower efficiency.

***The potential long range RNA-RNA interaction with the 5'NTR and base of SLVI does not affect translation.***

The region between SLI and SLII of the 5'NTR is essential for HCV replication and is not required for IRES activity, but has been the focus of several research groups studying translation. Kim et al reported a complementary sequence of 12 nucleotides between the 5'NTR and the base of SLVI in the capsid gene that results in a decrease in translation <sup>64</sup>. Mutations in either the 5'NTR or capsid gene designed to disrupt this

interaction led to an increase in translation of a bicistronic reporter. The capsid sequence thought to participate in this interaction involves the left side of the base of SLVI. In our studies, we tested the same mutations in the 5'NTR and base of SLVI in translation and replication assays. We found mutations designed to interrupt the 5'NTR and capsid gene interaction actually slightly impaired translation. We used similar translation conditions as Kim et al., but included the sequence for the competing right side of SLVI in our analysis. In this context, we found the structure of SLVI to influence translation and not an interaction with the 5'NTR. Indeed, mutations in the basal region of SLVI impaired translation and replication of full length HCV genomes, and both activities were improved in genomes harboring the compensatory mutations restoring the base pairing of SLVI.

Previous reports suggested that differences between genotypes in the unstructured region of the 5'NTR contribute to a discrepancy in translation<sup>55</sup>. We compared the 5'NTR sequences between genotype 1a, 1b and 2a and found no effect on translation or replication. In fact a genotype 1a/2a chimera genome that contains the 5'NTR from JFH-1, the capsid through NS2 genes from H77 and replicase to the 3'NTR from JFH-1 can replicate equally well as the H-JFH chimera where the entire 5'NTR is from the H77 sequence (data not shown). This is in contrast to a recent report of a

genotype 1b/ 2a chimera displaying genotype specificity between the replicase proteins and the 5'NTR.

Deletions were introduced into the 5'NTR and capsid gene to determine if the long-range interaction between these two regions controls translation. Deletion of the first 40 nucleotides of the 5'NTR resulted in an increase in translation, previously shown to be due to the removal of SLI. Mutations in SLVI conferred a decrease in translation in the presence of the deletion in the 5'NTR indicating the influence of SLVI on translation is independent of the 5'NTR. Deletions in the capsid gene reduced translation, with a more severe reduction observed when more capsid gene was deleted. Despite the low levels of translation, mutations in SLVI confer an additional translation defect in genomes with the downstream capsid gene deleted. This signifies that the role of SLVI in translation does not involved downstream elements in the capsid gene between amino acids 57 and 171. Branch et al. has proposed an RNA structure in the “terminal” region of the capsid gene. Since this region also overlaps the signal sequence for E1, it is present in these deletion constructs and therefore may have an effect. A role for the capsid terminal RNA element is unknown <sup>21</sup>.

***Transcomplementation with the capsid, E1, E2, p7 and NS2 in replication and virus infectivity.***

Previous studies have shown the capsid gene and/or protein to negatively influence translation by acting either *in cis* or *in trans*. The modulation of translation by the capsid protein is reported to occur by directly binding the IRES. We found that deletions in the capsid gene slightly reduce translation, but considerably impair replication. We have made similar observations with mutants in SLVI that convey subtle effects in translation, yet dramatic effects on replication. This result is in contrast to previous assumptions that the capsid protein did not have a role in replication, at least in the context of the subgenomic replicon. The impact of the capsid protein or sequence acting as a translation repressor on replication has not yet been addressed. The adenine-rich region of the capsid sequence has been shown to interact with a cellular protein, NSAP1, to increase translation, and this sequence is present in all of the mutant genomes<sup>63</sup>. If the capsid protein acts, as reported, to down regulate translation, this may consequently stimulate replication. Expression of the capsid protein *in trans* by a VEE replicon did moderately increase replication of the genomes harboring deletions in the capsid gene. Furthermore, virus production was increased in the wild type genome when replication occurred in the presence



of expression *in trans* of the capsid protein. The genomes with deletions in the capsid gene could not be trans complemented for virus production

The effects of expressing the capsid protein and polyprotein of capsid, E1, E2, p7 and NS2 on translation and replication of deleted capsid genomes were closely analyzed in a time course. Translation of the input RNA was compromised for Del 15-171 and Del 57-171 harboring mutations in SLVI, but no specific effect from trans expression of HCV proteins from the VEE replicon was observed. The capsid protein expressed alone or with the HCV structural proteins did not modulate translation *in trans* as previously reported. The reduced translation observed for Del 57-171/ Stop 1,2,3,4 is consistent with the findings in the *in vitro* translation study. For all HCV genomes, the differences in translation diminished by 8H indicating the defect in translation was not as severe as seen in the IRES initiation mutants. Coexpression of the VEE replicon led to a reduction in HCV replication.

At 24H, luciferase differences become apparent between mutant and wild type genomes. At this time point, wild type luciferase levels begin their rapid increase indicating the launch of replication and translation of nascent HCV RNA. Luciferase activity was enhanced in the presence of expression *in trans* of the capsid-NS2 polyprotein; however, luciferase activity was suppressed by expression *in trans* of the capsid protein alone. Expression of

the capsid protein in the context of the structural proteins may coordinate the localization of the capsid protein in the cell. This may lead to a lower amount of capsid protein to act *in trans* on HCV translation and replication. Also, the structural proteins of E1-NS2 may influence translation and replication by acting *in trans*. The expression of the capsid protein alone did reduce luciferase expression of even the wild type genome. The capsid protein could be repressing translation of the nascent HCV RNA and thus indirectly lower RNA synthesis or act directly against a component of replication. The overlapping functions make a distinction of the defect impossible. Initial translation of input RNA was not affected by capsid *in trans* and luciferase levels did reach wild type, but perhaps translation after replication involves different factors than those used in translation of the incoming virus RNA.

In our initial study, we found the capsid protein to slightly enhance luciferase expression at 72H. In this study, the capsid protein was found to suppress luciferase activity, notably at 24H, but also at 48H. The disparity between the studies could be due to the amount of transfected VEE RNA. When the capsid protein was found to enhance replication, it was expressed at a low ratio of VEerepC to HCV RNA, 0.5 ug to 2 ug. When the capsid protein was found to suppress luciferase expression, the amount of

transfected VEE RNA had been increased to 3 ug and cotransfected with 1 ug of HCV RNA. It could be that a lower amount of HCV capsid protein results in a slight increase in translation and/or replication. When the capsid protein is expressed in the context of HCV structural proteins less is available to repress translation/ replication and therefore acts in a stimulatory manner. HCV RNA naturally enters the cell encapsidated with the structural proteins of the capsid, E1, and E2. The presence of small amounts of capsid protein could act as a stimulant to initial translation. As replication proceeds and translation continues to express more HCV polyproteins, an increase in the localized amount of capsid protein could act as a repressor for translation of nascent RNA.

***Transcomplementation of silent and noisy point mutations in the capsid gene.***

Expression of C-NS2 *in trans* was found to increase luciferase expression of HCV genomes harboring mutations in SLVI. Genomes with mutations that disrupted the top stem of SLVI, but not the loop increased their luciferase expression in the presence of C-NS2 *in trans*. Some of the mutations were silent; some changed the amino acid of the capsid protein. Expression of C-NS2 was found to transcomplement virus infectivity.

Genomes with mutations in the capsid protein sequence showed an increase in virus infectivity when packaged in cells coexpressing VEErepC-NS2. The most dramatic complementation was found in genomes with mutations in the base of SLVI, Base Right, and genomes with mutations in the middle of SLVI. All of the successful complemented genomes contained an amino acid change in the capsid protein. The transcomplementation for the mutant genome Base Right again shows no benefit at the level of replication, yet almost 1 log increase in infectivity when packaged in the presence of exogenous C-NS2. A slight increase in infectivity is detected again for the wild type genome.

The contribution of C-NS2 in virus production is unknown. We have detected compensatory mutations p7 and NS2 that enable virus infectivity in the presence of a mutant capsid protein. Two second site reversions in p7 and NS2 increased virus infectivity for a genome with an amino acid change in the capsid protein indicating either a cooperative interaction or a general improvement of replication or virus fitness. The mechanism of transcomplementation by the capsid, E1, E2, p7, and NS2 proteins in virus infectivity is unknown. The observation that these proteins cannot act *in cis* in the presence of a mutated capsid is intriguing. It could be the mutant capsid protein acts as a dominant-negative in virus production, and

overexpression of the structural proteins can overcome this effect. It would be interesting to know if individual expression of E1, E2, p7 or NS2 could restore virus infectivity. Individual expression of the capsid protein led to an inhibition of replication whereas an improvement in replication was detected when coexpressed with the structural proteins. The difference could be due to localization of the capsid protein. When expressed in the context of E1-NS2, *cis*-acting interactions could prevent or reduce the amount of capsid protein interacting with HCV replication *in trans*.

#### ***Analysis of HCV translation and replication by end point dilution.***

Few cells transfected with HCV RNA with mutations in SLVI are positive for HCV antigen expression as measured by flow cytometry and IHC. Measuring replication at the population level has shown mutant HCV to replicate 1-2% of wild type. We questioned whether the mutant HCV is replicating at a low level throughout the population or if a subpopulation of cells can efficiently replicate the mutant RNA. Transfected Huh-7.5 cells were serially diluted to analyze replication at the single cell level.

Replication at each dilution was monitored in multiple wells. Two types of events were expected to describe the population of cells replicating the mutant genome. In one case, luciferase values would decline in a linear

association with the cell numbers if all cells in a homogenous population could replicate the mutant genome. Alternatively, luciferase activity would not be observed to steadily decline with cell number, but rapidly fall as the small number of cells positive for HCV replication were seeded only in the higher dilution wells. For cells transfected with the genome containing mutations in SLV and SLVI, we observed the latter situation. Translation of the wild type, pol-, and mutant RNA was detected in a linear relationship consistent with the number of cells. A difference between the HCV genomes was detected as replication was launched. A larger proportion of the cell population could replicate the wild type genome than the mutant. At 24H, the mutant genome required 40 times the number of cells as the wild type to achieve half of the average luciferase units. At the height of replication, the mutant genome requires 100 times the number of transfected cells to achieve a replication level half of wild type.

This result indicates there is a small population of Huh-7.5 cells capable of replicating the mutant genome. One explanation for this observation is that this population is genetically distinct and contributes to HCV replication by an unknown host cell factor. Another explanation is that the mutant HCV can only launch replication at an infrequent rate. To

distinguish between these two possibilities, the following experiments could be performed.

To determine if there is a beneficial host factor, cells harboring mutant HCV replication could be isolated and “cured” of replicating HCV by treatment with a replication inhibitor or interferon, and retransfect the cells with mutant HCV RNA. If replication of the mutant RNA was near wild type levels, it could be concluded that a host cell factor contributes to replication of a HCV with disrupted SLVI. If the small population of cells is a result of an infrequent event allowing mutant replication to launch, then transfecting more mutant RNA into Huh-7.5 cells may increase the number of replication-success events. The infrequency of launching replication could be due any numerous intrinsic/extrinsic variables of the cell state such as cell cycle, stress, growth factors, a state of a pathway, etc. Treatment of the cells to a variety of stimulants may trigger more cells in the population to exhibit the beneficial condition. The use of a fluorescent marker in the genome would also allow a visual distinction between HCV positive and negative cells and gradients in between.

***The role of SLVI is dependent on the position within the genome and mutants of SLVI act to reduce translation.***

We have shown mutations in SLVI to subtly reduce translation and severely impair replication. We tested if the effects of SLVI are dependent on its location within the genome. No change in replication was detected in a bicistronic genome where mutations of SLV and SLVI were introduced into the capsid gene of a polyprotein located behind the EMCV IRES. This indicates that the mutations in SLVI affect replication when located near the 5'NTR and behind the HCV IRES. We tested if the mutations of SLVI reduce replication when in their original position, but expression of the replicase proteins independent of the HCV IRES and any effects on translation from the mutant SLVI. In a bicistronic genome where the replicase is expressed by the EMCV IRES, we found mutations for SLVI impair replication slightly between 20H and 52H; however, replication could reach wild type levels at 72H. This indicates the mutations of SLVI contribute to a reduced level of translation of the polyprotein and thus replicase proteins. The lower expression of replicase proteins leads to a drastic decline in the level of replication. The reduced replication at 20H and 30H may be due to an influence of the mutant SLVI on an additional component of replication or may act to reduce expression of C-NS2, and



thus lower the efficiency of replication. We have shown the mutant genome of Open 3,4 to be enhanced by expression of C-NS2 *in trans*. Even though C-NS2 are thought not to contribute to replication, they are expressed with equal ratio as the replicase proteins, and our data indicate they do have a positive effect on replication. In summary, mutations in SLV and SLVI were found to slightly reduce translation, yet dramatically reduce replication. The replication defect contributed by mutations in SLVI was restored by separating expression of the replicase proteins from the HCV IRES and mutant SLV and SLVI. These observations indicate that mutations in SLVI reduce translation of the polyprotein and thus lead to a reduction in replication.

## **Chapter 6**

## **Chapter 6. Discussion.**

### **Predictions of a multifunctional capsid gene.**

Phylogenetic studies have shown HCV to display great diversity; although the variation is not evenly distributed through the HCV genome, comparison of sequences from all 6 genotypes shows overall nucleotide differences of 30%<sup>130</sup>. Regions of conservation indicate elements indispensable for the virus life cycle that cannot tolerate mutation without fitness being compromised. The 5'NTR and 3'NTR are the most conserved areas of the genome as they encode essential RNA structures. Studies of the synonymous/ nonsynonymous mutation rate and codon usage of the polyprotein coding region revealed exceptional sequence conservation for the capsid and polymerase genes<sup>132, 153</sup>. These genes flank the polyprotein and are adjacent to the NTRs. The significance of this conservation was demonstrated by the presence of an RNA secondary structure in the coding region of NS5B that acts in a long range RNA-RNA kissing interaction with the 3'NTR to conduct HCV replication<sup>40, 165</sup>. The sequence of the NS5B coding region is thus constrained to maintain both an essential protein and an essential RNA structure.

The capsid gene also displays great sequence conservation, which has led to predictions that this gene sequence possesses functions in addition to

simply encoding the capsid protein. All 6 HCV genotypes have a substantial open reading frame that overlaps the amino terminus of the capsid gene and could encode a 9-16kDa protein<sup>153, 161</sup>. HCV patients have been shown to produce antibodies and cellular immune responses to peptides from this alternative reading frame indicating its expression during infection<sup>66, 153, 161</sup>. Using phylogenetic analysis of covariant base change, RNA secondary structures have also been predicted to form in this region<sup>133, 147</sup>. These structures, SLV and SVI, are positioned in the amino terminus of the capsid and ARF gene. The requirement for, and functions of, the ARF protein and the SLV and SLVI RNA elements in HCV infection and replication are the subject of this dissertation.

### **Is the ARF essential in an HCV infection?**

The chimpanzee model remains the gold standard for defining the essential genetic elements of the HCV genome, and describing the interplay between virus and the immune system. We therefore decided to test if the ARF was required for an HCV infection and if so, how the host's immune system would respond to an ARF-ablated virus. We did detect an infection following inoculation of an H77 genome with mutations conferring Stop 1,2,3,4 in the ARF. The peak of circulating HCV RNA was several logs

lower than previous chimpanzee infections launched with the parental H77 sequence, and the duration of the acute infection was shorter. This suggested an attenuated infection, and we hypothesized that the absence of an ARF protein might contribute to a more robust immune response. We isolated CD8<sup>+</sup> and CD4<sup>+</sup> cells from both the periphery and the liver and found in both compartments, only a low frequency of T cells that recognized HCV-specific epitopes. Since it is thought that virus infection leads to recruitment of immunocytes that contribute to liver damage <sup>19,48</sup>, the absence of an elevated ALT also indicated low immune activity. The absolute number of CD4<sup>+</sup>/CD3<sup>+</sup>, CD8<sup>+</sup>/CD3<sup>+</sup>, CD56<sup>+</sup>/CD8<sup>+</sup> and CD56<sup>+</sup>/CD4<sup>+</sup> cell populations in the circulation did not change during the HCV infection indicating a functional immune system. However, when we mapped the HCV-specific CD8<sup>+</sup> and CD4<sup>+</sup> epitopes, we found a diverse repertoire from a range of peptides from E2 to NS4B. In previous studies, this level of diversity was correlated with clearance of the virus infection while a limited repertoire of one or two HCV-specific epitopes was predictive of virus persistence and chronic infection <sup>88</sup>. The appearance of HCV-specific lymphocytes at the cellular and humoral level coincided with virus decline in the circulation.

The launch of an infection with the Stop 1,2,3,4 genome does prove that the ARF is not absolutely required for virus production. A diverse

immune response was stimulated indicating that virus replication did occur but, given the low ALT and low frequency of HCV-specific T cells, the infection was attenuated. As we later determined in cell culture studies, the replication of the Stop 1,2,3,4 genome was compromised and resulted in the attenuation.

### **Are the Stop 1,2,3,4 mutations positively selected?**

In a natural infection, HCV is present in a quasispecies. Launching the infection with a monoclonal genome allowed us to study the fitness of the specific input virus RNA sequence. Within 2 weeks of the infection we detected a reversion that removed the Stop 3 mutation. The appearance of this reversion corresponded with the launch of the acute phase and was the reversion quickly became fixed in the population. Using a sensitive nested RT-PCR, we could detect the original input sequence at week 1 in the circulation indicating replication and therefore creation of diverse HCV progeny. The fact that we could not detect the input sequence after week 1 indicates a high selective pressure for the reversion. Numerous studies have investigated the evolution of HCV in natural and experimental infections. The dynamics of selective pressures on HCV have been described for the hypervariable region of E2 where amino acid mutation potentially allows for

antibody escape, CD8<sup>+</sup> T cell escape, and resistance to interferon therapy<sup>20, 23, 143</sup>. Most of these studies have concluded selection to be positive selection where the rate of fixation exceeds the rate of mutation. This is the case for the selection of the reversion at Stop 3 where a bottleneck was apparent and only those genomes with the reversion survived in the population. Since protein expression from the ARF was still blocked in this revertent, this indicates a positive selection for the RNA secondary structure, as opposed to a restoration of the ARF. During the acute infection, the F and DF proteins remained blocked by the downstream Stop 4. A second reversion was positively selected within the first mutation conveying Stop 4. This second reversion occurred late, at week 10, and did remove the fourth stop codon, allowing the DF protein to be expressed, but did not alter the capsid protein. A third reversion (of the second mutation within Stop 4) was subsequently detected, but since this occurred in the context of a revertent that could already express the DF protein, this is indicative of selective pressure to restore the parental H77 RNA sequence.

## **Is replication compromised by the ablated ARF or disrupted RNA secondary structures in cell culture?**

HCV genomes containing the mutations for Stop 1,2,3,4 were severely compromised for replication in Huh-7.5 cells. The replication defect was mapped to the mutations for Stop 3 and 4. Alone, the mutations of Stop 1 and 2 had no effect on replication; however, a synergistic reduction of replication was observed when combined with Stop 3 and 4 mutations. Phylogenetic analysis of covariant base changes predicted this region to form the RNA structure SLVI<sup>147</sup>. The sequence of top region of SLVI is highly conserved; the 473A mutation of Stop 3 is only present in 2 of 1274 HCV sequences from all 6 genotypes, while the Stop 4 mutations 480A and 482G are completely absent. The mutations of Stop 3 and 4 disrupt three base pair interactions in a very conserved helical stem.

The mutations of Stop 1,2,3,4 are silent for the capsid, but do alter the codon usage. Mueller et al. found that recoding the capsid gene to suboptimal codons in poliovirus resulted in a complete loss of virus production<sup>92</sup>. The defect was due to inefficient translation and consequently compromised replication. Recoding only portions of the capsid gene resulted in virus particles with a reduced infectivity. These mutant polioviruses were, however, capable of launching infection and causing pathology in the



poliovirus mouse model, but were found to produce fewer particles. The authors contend that since RNA viruses exist close to the error threshold, creating one suboptimal step in the virus pathway can lead to a dramatic attenuation of the virus fitness. The mutations for Stop 4 do not introduce a rare codon, but in fact exchange a CGC arginine codon for a more frequently used AGG. The mutation for Stop 3, however, does replace a leucine codon for a less frequent codon. So in addition to disrupting SLVI, translation could be attenuated by the rare leucine codon.

We were able to measure replication of genomes containing other mutations in the same location of SLVI. The Open mutations allow expression of both the capsid and ARF proteins. Open 3 contains a mutation at the same nucleotide position as Stop 3 (473) and another substitution at 471. The Open 3 genome was reduced in replication, yet the leucine codon was not the rare TTA as in the Stop 3, but a more frequently used CTC. This argues that the defect in replication is due to disruptions in SLVI as opposed to slow translation induced by usage of a rare codon.

Mutagenesis experiments showed that replication was most severely affected by mutations in the top stem of SLVI, the regions of 3 and 4 mutations. Disruptions in the base of SLVI also reduced replication, but not with the same severity as with the apical stem. There was an additive

reduction in replication when mutations at the base and the top were combined. This suggests that the overall structure is important for SLVI function.

**Does the replication defect in Huh-7.5 cells explain the attenuated infection of the mutant genome Stop 1,2,3,4?**

Passaging Huh-7.5 cells transfected with genomes containing Stop 1,2,3,4 and Stop 3,4 resulted in selection of reversions that led to an increase in replication. The reversions at Stop 3 (473G) and the first base of Stop 4 (480C) detected in Huh-7.5 cells were the identical to the reversions found during the chimpanzee infection. This suggests that the replication defect contributed by the Stop 3 and 4 mutations have similar deleterious effects *in vivo* and *in vitro*. Thus, the attenuation of the chimpanzee infection was consistent with the replication defect observed in cell culture. Interestingly, the reversion of Stop 3 and 480C of Stop 4 were found to have occurred on independent genomes in Huh-7.5 cells. The 480C revertant became more frequent in the population with passage. Replication of the Stop 1,2,3 genome was more efficient than that of Stop 1,2,4, suggesting that the Stop 4 mutation of 480A had an impact on fitness. A combination of both of the reversions observed in cell culture and the third reversion of Stop 4 (482C)

were required to increase fitness in the chimpanzee infection; a likely consequence of the substantial selective pressures during a natural infection. We did not observed any replication effect in genomes harboring the Stop 1,2 mutations, the genome found at the conclusion of the chimpanzee infection.

### **What defect in replication is caused by mutations in SLVI?**

We found that mutations of SLVI reduced translation in both transfected Huh-7.5 cells and using cell free lysates. The base of SLV is just 13 bases downstream of the proposed 3' boundary of the IRES (position 375) while the base of SLVI lies 53 nucleotides downstream of the IRES. We compared translation and replication of HCV genomes containing mutations in the IRES that impair translation initiation <sup>98</sup>. Genomes with mutations that prevent 40S binding or reduce the transition from 48S\* to 80S were severely compromised in both assays. Measuring translation and replication with a luciferase reporter in Huh-7.5 cells revealed that the IRES mutant genomes were severely compromised for initial translation and levels of luciferase failed to reached even the polymerase defective control at all time points. The mutations in SLVI did not convey such a dramatic effect. In the context of replication defective genomes, a slight reduction of 2- 4 fold

below the wild type signal was measured for mutant SLVI genomes, both in cell free assays and after transfection. This indicates that mutant SLVI does not impede the association of the 40S, eIF3, ternary complex or 60S with the IRES. Instead, the mutant SLVI could influence IRES function, by forming a tertiary interaction with the IRES sequence. In previous reports, IRES structure, and the determination of initiation factor binding, was always studied in the absence of SLV and SLVI. Although the RNA complex is very large, NMR analysis of the IRES in the presence of SLV and SLVI could determine if an altered conformation is formed. Similarly, an assessment of the binding of initiation factors to the IRES, in the presence of SLV and SLVI, could help define a role for these RNA elements in the initiation of HCV IRES- mediated translation.

**Can the HCV structural proteins and NS2 act *in trans* to assist replication of mutant SLVI genomes?**

Translation of the polyprotein is assumed to express each virus protein to equal amounts. The structural proteins are first to be expressed, yet cannot assemble a virus until replication has synthesized a sufficient level of progeny virus RNA. The capsid protein has been reported to act as a translational repressor in addition to its role as a structural protein, and may

regulate the transition between translation and replication. Deletion of the capsid gene from amino acids 15- 171 reduced replication. In genomes where the capsid sequence of SLV and SVI was included, but downstream amino acids 57-171 were deleted, replication was improved, but not to wild type levels. Replication was further reduced when mutations were introduced in SLV and SLVI. To determine if the capsid protein or RNA elements in the capsid sequence could transcomplement these replication defects, we expressed the capsid protein alone or in the context of E1, E2, p7, and NS2. We found that expression of the capsid protein alone reduced translation and replication of the wild type and all mutant genomes. However, expression of C-NS2 led to an increase in replication of wild type and all mutants, after 24H post transfection. Despite the replication increase in the presence of C-NS2, replication of the mutant SLV and SLVI structures was still below wild type. The increase in translation and/or replication contributed by expression of C-NS2 was transitory; by 48H only a slight improvement was observed. Expression of the capsid protein alone however, continued to result in a lower level of luciferase expression.

The influence on translation and/ or replication by C-NS2 at the 24H time point implies a role for these proteins beyond virus assembly. Genomes containing deletions of the capsid gene, either with or without mutations in

SLV and SLVI, were shown to have a reduced translation in cell free extracts. This could indicate a threshold amount of C-NS2 proteins is required for efficient replication. Expression of the capsid protein in the context of the structural proteins and NS2 could tether the capsid location in the cell and thus reduce the amount of protein able to interact with the HCV genome. In addition, the C-NS2 complex could have a function in replication. Although there are many studies claiming the capsid protein affects cellular and virus functions when expressed alone, little is known of a function for C-NS2 beyond virus assembly.

### **Is there a subpopulation of Huh-7.5 cells capable of supporting replication of SLVI mutants?**

Observations of Huh-7.5 cells transfected with mutant SLVI genomes revealed a small number of cells positive for HCV antigen. We have shown that replication of SLVI mutants is impaired by measuring differences at the population level. It was unclear, however, if the low levels we detected were due to a few cells replicating the mutant to high levels or if most cells in the population were weakly replicating the mutant. We therefore measured the level of translation and replication of mutant genomes in a limiting dilution assay. We found that 100 times more transfected cells were required to

detect mutant replication than wild type. At the lowest points of dilution, the wild type luciferase level was twice that of the mutant. This indicated that a small population of Huh-7.5 cells was able to replicate the mutant genome, and that this subset of cells was able to replicate the mutant genome only half as well as wild type.

Clonal cell populations are known to display heterogeneity that results in “noise” when measuring biological pathways <sup>33</sup>. Huh-7 cells are no exception and subclones have been isolated that differ from the parental line with respect to HCV permissivity. Subpopulations of Huh-7 cells capable of increased HCV replication were selected to establish the Huh-7.5 cell line <sup>14</sup>. We observed that mutant HCV RNA was able to replicate in either a subpopulation of cells genetically distinct from the rest of the population or a small population of cells that were in a particular “state” that enabled the launch of mutant HCV replication. Numerous cellular factors have been described in HCV replication ranging in function from RNA binding proteins to interferon antagonists to lipid metabolism to most known cellular pathways. Cellular factors required for HCV replication have been shown to be limiting in Huh-7 and Huh-7.5 cell populations, although the identity of these factors and the extent of their cell-to-cell variation is unknown <sup>7</sup>.

### **Are SLV and SLVI location dependent?**

We found that the mutations in SLV and SLVI do not affect replication when located behind the EMCV IRES in a bicistronic genome. This implies that mutations in SLVI were acting in cis and interfere with either the IRES or a replication element in the 5'NTR. Finally, we discovered that we could rescue the defect in replication conferred by mutant SLVI by expressing the replicase complex independently of the HCV IRES. This is consistent with the hypothesis that mutation of SLVI reduces translation of the replicase proteins, leading to a reduction in HCV replication.

A similar reduction in translation and replication was shown for the flavivirus, BVDV. Replication of a subgenomic replicon of BVDV was found to be initially impaired, but improved upon passage in cell culture<sup>93</sup>. Adaptive mutations were identified that removed RNA secondary structure from sequence downstream of the BVDV IRES. The authors concluded that the presence of RNA structure partially inhibited 40S subunit binding. The result reduced translation slightly (8-10-fold), but dramatically reduced replication (10-100- fold). The BVDV IRES shares many similarities to the HCV IRES; in fact, the BVDV IRES can be functionally replaced by the HCV IRES. The selection for adaptive mutations that increase the efficiency



of translation relate again to the coupled processes of translation and replication and how virus fitness can be dramatically reduced if one stage is suboptimal <sup>42</sup>.

### **Model for the role of SLVI in replication**

The following model is proposed for how SLVI influences HCV replication. Using the luciferase reporter genome as a guide of translation and replication, the following differences can be determined between wild type and mutant HCV RNA. Translation of transfected RNA is initiated in both genomes; however, the mutant genome is slightly impaired. For both the mutant and wild-type, the polyprotein is expressed and processed into the mature viral proteins. Translation continues and in the presence of a saturating amount of transfected RNA, little difference between wild type and mutant can be detected. Luciferase levels steadily increase and peak at 8H. Quink et al. quantitated the nonstructural proteins and plus and minus strand RNAs in a replicase complex <sup>108</sup>. They determined that the molar ratio of nonstructural proteins to negative strand RNAs is 1000:1 and concluded that “excessive translation” must first occur before replication begins. In our assays, we cannot distinguish between input RNA and replicating RNA by qRT-PCR until 12H post electroporation. If a substantial amount of

nonstructural proteins must be expressed before replication can proceed, replication may not begin until hours after transfection. Upon reaching this threshold of protein concentration, the input RNA then enters the replication complex. Due to the defect in the efficiency of translation, the mutant may lag behind the wild type as it takes longer to achieve the critical amount of replicase complex expression.

A decline in luciferase expression is detected at 12H post transfection in the wild type, mutant, and polymerase defective genomes. The polymerase defective genome can express and assemble into a replicase complex; it is just nonfunctional. This decline in luciferase is then due, either to degradation, or to the input RNA entering the replicase complex and being sequestered from the translation machinery. This consistent decline in translation indicates a level of synchronicity. An obvious dip in luciferase activity was not detected at later time points, where RNA is either being replicated or translated. Perhaps additional measurements between 24H and 72H would detect a trend or the decrease in luciferase is so apparent because of the large amount of translation that initially occurred without the noise of replication. The half-life of the Renilla luciferase protein is approximately 6H. Using a destabilized version might give a more accurate measure of the level of translation ongoing at particular time point.

The replicase complex is thought form a membrane-associated compartment, which restricts entry of proteins from the cellular environment, yet can exchange nucleotides. Replication proceeds through synthesis of a minus strand intermediate. By qRT-PCR, an increase of wild type RNA levels above the input value can be detected at 12H post transfection. Using the luciferase reporter, activity begins to increase at 16H, presumably through translation of nascent HCV RNA. It is unknown if newly synthesized virus RNA is actively exported from the replicase complex or only released when the complex is turned over. The life of a replicase complex is though to be 11H to 16H, and the increase in translation could be due to a sudden decay of the initial replicase complexes<sup>99, 103</sup>. A decrease in translation is detected for the polymerase defective genome as nascent RNA was not generated. Luciferase activity is lower for the mutant genome than the wild type at this time point. This lower level could be due to numerous factors. For example, although sufficient nonstructural proteins were expressed to initiate replication, the mutant RNA could have built fewer replicase complexes than the wild type, resulting in less synthesized RNA. The lag in building the replicase complexes could result in lower amounts of RNA generated at this time point. The mutant genome could be trapped in the replicase complex, perhaps because of an ineffective

interaction between a transport protein and the mutant RNA structure, or perhaps the replicase complex was built later and therefore will not degrade until later. Finally, the lower activity could simply be because translation of the newly synthesized RNA is attenuated.

Translation of newly synthesized RNA builds more replicase complexes. It is possible that, due to protein degradation, an “excessive amount” of translation must again occur to reach the amount of required nonstructural proteins. It has been observed that replicase complexes cannot be trans complemented. That is, a complex cannot replicate an RNA genome that it did not translate. Therefore, the nascent RNA must express and build its own replicase complex. Since the mutant genome is inefficiently translated, replicase complexes take longer to build, and a lag in the amount of mutant RNA is always detected. Alternatively, the mutant SLVI may also contribute to inefficient replication initiation and therefore reduce the amount of synthesized virus RNA.

Measuring replication using end point dilution, it was observed that fewer cells could replicate the mutant genome. If the mutant genome took too long in translation, a sufficient amount of nonstructural proteins would never be achieved, the replicase complex would never be built and virus RNA could never replicate. The cell would soon lose the HCV RNA. In a

minority cell population, however, this did not occur. Perhaps these cells express an abundance of translation machinery, are defective in protein or RNA degradation pathways, or express a factor that enables the mutant RNA to exit the replicase and undergo translation. The minor cell population does not, however, support replication of the mutant to wild type levels. It is possible that the mutant is also defective in an additional replication step. In the bicistronic construct where the replicase was expressed independently of the HCV IRES and mutant SLVI, RNA levels were restored to wild type levels, but only by 72H post transfection.

### **Future Directions:**

Mutations in SLVI attenuate translation and thus compromise replication. This reduction is most apparent between 16H and 24H using the luciferase system. During this time, initial replication and translation of nascent HCV RNA is occurring and the greatest differences between wild type and mutant are observed. Since we do not detect such a marked difference between the wild type and mutant after initial RNA transfection, I am curious if there are differences in translation between the translated RNA and recently synthesized virus RNA. To test if translation of the mutant is different from wild type after replication, translation could be compared, by

luciferase expression, after replication has been stopped by the addition of replication inhibitors. If performed at a time post transfection where input RNA has decayed, we would only detect translation of replicated HCV RNA. A reduction in translation could be a result of disruption of the IRES by the mutant SLVI or indicate an inability of the mutant RNA to associate with the translation machinery. The mutant RNA could be trapped in the replication complex. The replicase complex can be isolated by sucrose gradients and the amount of recovered HCV RNA compared.

Another remaining question concerns the observation that the liver abundant microRNA, miR122 can enhance HCV replication by an as yet undefined mechanism <sup>59</sup>. In parallel to published results, we found mutations of the miR122 seed region in the 5'NTR eliminate HCV replication and yet could be restored by compensatory mutations of a lentivirus-expressed mutant miR122. The effect of miR122 on HCV replication is a novel observation for microRNA, which normally act to down-regulate translation of cellular RNA. MicroRNAs are thought to control mRNA expression by targeting the cellular RNAs to either processing bodies or stress granules. One hypothesis for a mechanism of miR122 in HCV replication is to test if HCV is localized to either stress granules or processing bodies, and if this action is required for replication. Studies of processing bodies (P bodies)

have identified cellular scaffolding proteins required for their formation.

Knocking down key proteins by siRNA disassembles the P bodies.

Comparing HCV replication and virus production in Huh-7.5 cells with knocked down P body components would determine if these processes contribute to the HCV lifecycle and a role for miR122.

Although the structure of SLV and SLVI has been probed by chemical and enzymatic analysis, a structure with higher resolution would benefit the genetic studies and potential interactions with other RNA or protein molecules. The size of the complete SLVI is rather large, 86 bases, but analysis with NMR, Cryo-EM or X-ray crystallography would provide greater detail of the stability of the structure and perhaps a similarity of other RNA structures with known function. Determination of the structure and interacting proteins of the IRES has not been studied in the context of SLV and SLVI. Repeating previous studies of structure and protein binding with additional capsid sequence including SLVI may alter the results and provide insight into how mutations in SLVI could reduce virus translation and replication.

## References.

1. Alter, H. 1999. Discovery of non-A, non-B hepatitis and identification of its etiology. *Am J Med* 107:16S-20S.
2. Alter, H. J. 2005. HCV natural history: the retrospective and prospective in perspective. *J Hepatol* 43:550-2.
3. Alter, H. J., and M. Houghton. 2000. Clinical Medical Research Award. Hepatitis C virus and eliminating post-transfusion hepatitis. *Nat Med* 6:1082-6.
4. Alter, H. J., S. L. Stramer, and R. Y. Dodd. 2007. Emerging infectious diseases that threaten the blood supply. *Semin Hematol* 44:32-41.
5. Appel, N., U. Herian, and R. Bartenschlager. 2005. Efficient rescue of hepatitis C virus RNA replication by trans-complementation with nonstructural protein 5A. *J Virol* 79:896-909.
6. Baril, M., and L. Brakier-Gingras. 2005. Translation of the F protein of hepatitis C virus is initiated at a non-AUG codon in a +1 reading frame relative to the polyprotein. *Nucleic Acids Res* 33:1474-86.
7. Bartenschlager, R., M. Frese, and T. Pietschmann. 2004. Novel insights into hepatitis C virus replication and persistence. *Adv Virus Res* 63:71-180.
8. Bartenschlager, R., and V. Lohmann. 2001. Novel cell culture systems for the hepatitis C virus. *Antiviral Res* 52:1-17.
9. Bartenschlager, R., and V. Lohmann. 2000. Replication of the hepatitis C virus. *Baillieres Best Pract Res Clin Gastroenterol* 14:241-54.
10. Binder, M., D. Quinkert, O. Bochkarova, R. Klein, N. Kezmic, R. Bartenschlager, and V. Lohmann. 2007. Identification of determinants involved in initiation of hepatitis C virus RNA synthesis by using intergenotypic replicase chimeras. *J Virol*.
11. Blight, K. J. 2007. Allelic variation in the hepatitis C virus NS4B protein dramatically influences RNA replication. *J Virol*.
12. Blight, K. J., A. A. Kolykhalov, and C. M. Rice. 2000. Efficient initiation of HCV RNA replication in cell culture. *Science* 290:1972-4.
13. Blight, K. J., J. A. McKeating, J. Marcotrigiano, and C. M. Rice. 2003. Efficient replication of hepatitis C virus genotype 1a RNAs in cell culture. *J Virol* 77:3181-90.
14. Blight, K. J., J. A. McKeating, and C. M. Rice. 2002. Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *J Virol* 76:13001-14.



15. Blight, K. J., and C. M. Rice. 1997. Secondary structure determination of the conserved 98-base sequence at the 3' terminus of hepatitis C virus genome RNA. *J Virol* 71:7345-52.
16. Bol, J. F. 2005. Replication of alfamo- and ilarviruses: role of the coat protein. *Annu Rev Phytopathol* 43:39-62.
17. Boni, S., J. P. Lavergne, S. Boulant, and A. Cahour. 2005. Hepatitis C virus core protein acts as a trans-modulating factor on internal translation initiation of the viral RNA. *J Biol Chem* 280:17737-48.
18. Boulant, S., M. Becchi, F. Penin, and J. P. Lavergne. 2003. Unusual multiple recoding events leading to alternative forms of hepatitis C virus core protein from genotype 1b. *J Biol Chem* 278:45785-92.
19. Bowen, D. G., and C. M. Walker. 2005. Adaptive immune responses in acute and chronic hepatitis C virus infection. *Nature* 436:946-52.
20. Bowen, D. G., and C. M. Walker. 2005. Mutational escape from CD8+ T cell immunity: HCV evolution, from chimpanzees to man. *J Exp Med* 201:1709-14.
21. Branch, A. D., D. D. Stump, J. A. Gutierrez, F. Eng, and J. L. Walewski. 2005. The hepatitis C virus alternate reading frame (ARF) and its family of novel products: the alternate reading frame protein/F-protein, the double-frameshift protein, and others. *Semin Liver Dis* 25:105-17.
22. Brierley, I., and F. J. Dos Ramos. 2006. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* 119:29-42.
23. Brown, R. J., V. S. Juttla, A. W. Tarr, R. Finnis, W. L. Irving, S. Hemsley, D. R. Flower, P. Borrow, and J. K. Ball. 2005. Evolutionary dynamics of hepatitis C virus envelope genes during chronic infection. *J Gen Virol* 86:1931-42.
24. Bukh, J., X. Forns, S. U. Emerson, and R. H. Purcell. 2001. Studies of hepatitis C virus in chimpanzees and their importance for vaccine development. *Intervirology* 44:132-42.
25. Bukh, J., T. Pietschmann, V. Lohmann, N. Krieger, K. Faulk, R. E. Engle, S. Govindarajan, M. Shapiro, M. St Claire, and R. Bartenschlager. 2002. Mutations that permit efficient replication of hepatitis C virus RNA in Huh-7 cells prevent productive replication in chimpanzees. *Proc Natl Acad Sci U S A* 99:14416-21.
26. Choi, J., Z. Xu, and J. H. Ou. 2003. Triple decoding of hepatitis C virus RNA by programmed translational frameshifting. *Mol Cell Biol* 23:1489-97.

27. Choo, Q. L., G. Kuo, A. J. Weiner, L. R. Overby, D. W. Bradley, and M. Houghton. 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244:359-62.
28. Chung, R. T. 2006. Assessment of efficacy of treatment in HCV: infection and disease. *J Hepatol* 44:S56-9.
29. Churkin, A., and D. Barash. 2006. RNAmute: RNA secondary structure mutation analysis tool. *BMC Bioinformatics* 7:221.
30. Cooper, S., A. L. Erickson, E. J. Adams, J. Kansopon, A. J. Weiner, D. Y. Chien, M. Houghton, P. Parham, and C. M. Walker. 1999. Analysis of a successful immune response against hepatitis C virus. *Immunity* 10:439-49.
31. Cristina, J., F. Lopez, G. Moratorio, L. Lopez, S. Vasquez, L. Garcia-Aguirre, and A. Chunga. 2005. Hepatitis C virus F protein sequence reveals a lack of functional constraints and a variable pattern of amino acid substitution. *J Gen Virol* 86:115-20.
32. El-Zayadi, A., P. Simmonds, H. Dabbous, and O. Selim. 1994. Hepatitis C virus genotypes among HCV-chronic liver disease patients in Egypt: a leading trial. *J Egypt Public Health Assoc* 69:327-34.
33. Elowitz, M. B., A. J. Levine, E. D. Siggia, and P. S. Swain. 2002. Stochastic gene expression in a single cell. *Science* 297:1183-6.
34. Evans, M. J., C. M. Rice, and S. P. Goff. 2004. Genetic interactions between hepatitis C virus replicons. *J Virol* 78:12085-9.
35. Evans, M. J., T. von Hahn, D. M. Tscherne, A. J. Syder, M. Panis, B. Wolk, T. Hatzioannou, J. A. McKeating, P. D. Bieniasz, and C. M. Rice. 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature*.
36. Fan, Z., Q. R. Yang, J. S. Twu, and A. H. Sherker. 1999. Specific in vitro association between the hepatitis C viral genome and core protein. *J Med Virol* 59:131-4.
37. Fletcher, S. P., I. K. Ali, A. Kaminski, P. Digard, and R. J. Jackson. 2002. The influence of viral coding sequences on pestivirus IRES activity reveals further parallels with translation initiation in prokaryotes. *Rna* 8:1558-71.
38. Foy, E., K. Li, R. Sumpter, Jr., Y. M. Loo, C. L. Johnson, C. Wang, P. M. Fish, M. Yoneyama, T. Fujita, S. M. Lemon, and M. Gale, Jr. 2005. Control of antiviral defenses through hepatitis C virus disruption of retinoic acid-inducible gene-I signaling. *Proc Natl Acad Sci U S A* 102:2986-91.

39. Fraser, C. S., and J. A. Doudna. 2007. Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat Rev Microbiol* 5:29-38.
40. Friebe, P., J. Boudet, J. P. Simorre, and R. Bartenschlager. 2005. Kissing-loop interaction in the 3' end of the hepatitis C virus genome essential for RNA replication. *J Virol* 79:380-92.
41. Friebe, P., V. Lohmann, N. Krieger, and R. Bartenschlager. 2001. Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication. *J Virol* 75:12047-57.
42. Frolov, I., M. S. McBride, and C. M. Rice. 1998. cis-acting RNA elements required for replication of bovine viral diarrhea virus-hepatitis C virus 5' nontranslated region chimeras. *Rna* 4:1418-35.
43. Gardner, J. P., R. J. Durso, R. R. Arrigale, G. P. Donovan, P. J. Maddon, T. Dragic, and W. C. Olson. 2003. L-SIGN (CD 209L) is a liver-specific capture receptor for hepatitis C virus. *Proc Natl Acad Sci U S A* 100:4498-503.
44. Grakoui, A., D. W. McCourt, C. Wychowski, S. M. Feinstone, and C. M. Rice. 1993. A second hepatitis C virus-encoded proteinase. *Proc Natl Acad Sci U S A* 90:10583-7.
45. Grakoui, A., C. Wychowski, C. Lin, S. M. Feinstone, and C. M. Rice. 1993. Expression and identification of hepatitis C virus polyprotein cleavage products. *J Virol* 67:1385-95.
46. Griffin, S. D., L. P. Beales, D. S. Clarke, O. Worsfold, S. D. Evans, J. Jaeger, M. P. Harris, and D. J. Rowlands. 2003. The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett* 535:34-8.
47. Grove, J., T. Huby, Z. Stamataki, T. Vanwolleghem, P. Meuleman, M. Farquhar, A. Schwarz, M. Moreau, J. S. Owen, G. Leroux-Roels, P. Balfe, and J. A. McKeating. 2007. Scavenger receptor BI and BII expression levels modulate hepatitis C virus infectivity. *J Virol* 81:3162-9.
48. Hanada, K., Y. Tanaka, M. Mizokami, T. Gojobori, and H. J. Alter. 2006. A reduction in selective immune pressure during the course of chronic hepatitis C correlates with diminished biochemical evidence of hepatic inflammation. *Virology*.
49. Heinz, F. X., Collett, M. S., Purcell, R. H., Gould, E. A., Howard, C. R., Houghton, M., Moormann, R. J. M., Rice, C. M. & Thiel, H. J. . 2000. Family Flaviviridae., p. 859–878. In C. M. F. M. H. V. van Regenmortel, D. H. L. Bishop, E. Carstens, M. K. Estes, S. Lemon, J. Maniloff, M. A. Mayo, D. McGeogh, C. R. Pringle & R. B. Wickner

- (ed.), *Virus Taxonomy*, 7th Report of the International Committee for the Taxonomy of Viruses. Academic Press, San Diego.
50. Hellen, C. U., and T. V. Pestova. 1999. Translation of hepatitis C virus RNA. *J Viral Hepat* 6:79-87.
  51. Heller, T., and B. Rehmann. 2005. Acute hepatitis C: a multifaceted disease. *Semin Liver Dis* 25:7-17.
  52. Honda, M., M. R. Beard, L. H. Ping, and S. M. Lemon. 1999. A phylogenetically conserved stem-loop structure at the 5' border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *J Virol* 73:1165-74.
  53. Honda, M., E. A. Brown, and S. M. Lemon. 1996. Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. *Rna* 2:955-68.
  54. Honda, M., L. H. Ping, R. C. Rijnbrand, E. Amphlett, B. Clarke, D. Rowlands, and S. M. Lemon. 1996. Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA. *Virology* 222:31-42.
  55. Honda, M., R. Rijnbrand, G. Abell, D. Kim, and S. M. Lemon. 1999. Natural variation in translational activities of the 5' nontranslated RNAs of hepatitis C virus genotypes 1a and 1b: evidence for a long-range RNA-RNA interaction outside of the internal ribosomal entry site. *J Virol* 73:4941-51.
  56. Hwang, S. B., S. Y. Lo, J. H. Ou, and M. M. Lai. 1995. Detection of Cellular Proteins and Viral Core Protein Interacting with the 5' Untranslated Region of Hepatitis C Virus RNA. *J Biomed Sci* 2:227-236.
  57. Ina, Y., M. Mizokami, K. Ohba, and T. Gojobori. 1994. Reduction of synonymous substitutions in the core protein gene of hepatitis C virus. *J Mol Evol* 38:50-6.
  58. Ji, H., C. S. Fraser, Y. Yu, J. Leary, and J. A. Doudna. 2004. Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc Natl Acad Sci U S A* 101:16990-5.
  59. Jopling, C. L., M. Yi, A. M. Lancaster, S. M. Lemon, and P. Sarnow. 2005. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* 309:1577-81.
  60. Kato, T., T. Date, M. Miyamoto, A. Furusaka, K. Tokushige, M. Mizokami, and T. Wakita. 2003. Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon. *Gastroenterology* 125:1808-17.

61. Kato, T., T. Matsumura, T. Heller, S. Saito, R. K. Sapp, K. Murthy, T. Wakita, and T. J. Liang. 2007. Production of Infectious Hepatitis C Virus of Various Genotypes in Cell Culture. *J Virol*.
62. Kieft, J. S., K. Zhou, R. Jubin, M. G. Murray, J. Y. Lau, and J. A. Doudna. 1999. The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J Mol Biol* 292:513-29.
63. Kim, J. H., K. Y. Paek, S. H. Ha, S. Cho, K. Choi, C. S. Kim, S. H. Ryu, and S. K. Jang. 2004. A cellular RNA-binding protein enhances internal ribosomal entry site-dependent translation through an interaction downstream of the hepatitis C virus polyprotein initiation codon. *Mol Cell Biol* 24:7878-90.
64. Kim, Y. K., S. H. Lee, C. S. Kim, S. K. Seol, and S. K. Jang. 2003. Long-range RNA-RNA interaction between the 5' nontranslated region and the core-coding sequences of hepatitis C virus modulates the IRES-dependent translation. *Rna* 9:599-606.
65. Kolykhalov, A. A., E. V. Agapov, K. J. Blight, K. Mihalik, S. M. Feinstone, and C. M. Rice. 1997. Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science* 277:570-4.
66. Komurian-Pradel, F., A. Rajoharison, J. L. Berland, V. Khouri, M. Perret, M. Van Roosmalen, S. Pol, F. Negro, and G. Paranhos-Baccala. 2004. Antigenic relevance of F protein in chronic hepatitis C virus infection. *Hepatology* 40:900-9.
67. Kong, L. K., and P. Sarnow. 2002. Cytoplasmic expression of mRNAs containing the internal ribosome entry site and 3' noncoding region of hepatitis C virus: effects of the 3' leader on mRNA translation and mRNA stability. *J Virol* 76:12457-62.
68. Krieger, N., V. Lohmann, and R. Bartenschlager. 2001. Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations. *J Virol* 75:4614-24.
69. Lancaster, A. M., E. Jan, and P. Sarnow. 2006. Initiation factor-independent translation mediated by the hepatitis C virus internal ribosome entry site. *Rna* 12:894-902.
70. Li, G., and C. M. Rice. 1993. The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. *J Virol* 67:5062-7.
71. Lin, C., B. D. Lindenbach, B. M. Pragai, D. W. McCourt, and C. M. Rice. 1994. Processing in the hepatitis C virus E2-NS2 region: identification of p7 and two distinct E2-specific products with different C termini. *J Virol* 68:5063-73.

72. Lindenbach, B. D., M. J. Evans, A. J. Syder, B. Wolk, T. L. Tellinghuisen, C. C. Liu, T. Maruyama, R. O. Hynes, D. R. Burton, J. A. McKeating, and C. M. Rice. 2005. Complete replication of hepatitis C virus in cell culture. *Science* 309:623-6.
73. Lindenbach, B. D., P. Meuleman, A. Ploss, T. Vanwolleghem, A. J. Syder, J. A. McKeating, R. E. Lanford, S. M. Feinstone, M. E. Major, G. Leroux-Roels, and C. M. Rice. 2006. Cell culture-grown hepatitis C virus is infectious in vivo and can be recultured in vitro. *Proc Natl Acad Sci U S A* 103:3805-9.
74. Lindenbach, B. D. a. R., C.M. 2001. Flaviviridae: the viruses and their replication., p. 991-1041. In D. M. K. B.N. Fields, and P.M. Howley (ed.), *Field's Virology*, 4th ed, vol. 1. Lippincott-Raven Press, Philadelphia.
75. Lo, S. Y., F. Masiarz, S. B. Hwang, M. M. Lai, and J. H. Ou. 1995. Differential subcellular localization of hepatitis C virus core gene products. *Virology* 213:455-61.
76. Lo, S. Y., M. Selby, M. Tong, and J. H. Ou. 1994. Comparative studies of the core gene products of two different hepatitis C virus isolates: two alternative forms determined by a single amino acid substitution. *Virology* 199:124-31.
77. Locker, N., L. E. Easton, and P. J. Lukavsky. 2007. HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *Embo J* 26:795-805.
78. Lohmann, V., F. Korner, A. Dobierzewska, and R. Bartenschlager. 2001. Mutations in hepatitis C virus RNAs conferring cell culture adaptation. *J Virol* 75:1437-49.
79. Lohmann, V., F. Korner, J. Koch, U. Herian, L. Theilmann, and R. Bartenschlager. 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* 285:110-3.
80. Lorenz, I. C., J. Marcotrigiano, T. G. Dentzer, and C. M. Rice. 2006. Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature* 442:831-5.
81. Lu, H. H., and E. Wimmer. 1996. Poliovirus chimeras replicating under the translational control of genetic elements of hepatitis C virus reveal unusual properties of the internal ribosomal entry site of hepatitis C virus. *Proc Natl Acad Sci U S A* 93:1412-7.
82. Major, M. E., H. Dahari, K. Mihalik, M. Puig, C. M. Rice, A. U. Neumann, and S. M. Feinstone. 2004. Hepatitis C virus kinetics and host responses associated with disease and outcome of infection in chimpanzees. *Hepatology* 39:1709-20.

83. Major, M. E., K. Mihalik, J. Fernandez, J. Seidman, D. Kleiner, A. A. Kolykhalov, C. M. Rice, and S. M. Feinstone. 1999. Long-term follow-up of chimpanzees inoculated with the first infectious clone for hepatitis C virus. *J Virol* 73:3317-25.
84. Major, M. E., K. Mihalik, M. Puig, B. Rehmann, M. Nascimbeni, C. M. Rice, and S. M. Feinstone. 2002. Previously infected and recovered chimpanzees exhibit rapid responses that control hepatitis C virus replication upon rechallenge. *J Virol* 76:6586-95.
85. Marczinke, B., R. Fisher, M. Vidakovic, A. J. Bloys, and I. Brierley. 1998. Secondary structure and mutational analysis of the ribosomal frameshift signal of rous sarcoma virus. *J Mol Biol* 284:205-25.
86. McLauchlan, J. 2000. Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. *J Viral Hepat* 7:2-14.
87. Mercer, D. F., D. E. Schiller, J. F. Elliott, D. N. Douglas, C. Hao, A. Rinfret, W. R. Addison, K. P. Fischer, T. A. Churchill, J. R. Lakey, D. L. Tyrrell, and N. M. Kneteman. 2001. Hepatitis C virus replication in mice with chimeric human livers. *Nat Med* 7:927-33.
88. Meyer-Olson, D., N. H. Shoukry, K. W. Brady, H. Kim, D. P. Olson, K. Hartman, A. K. Shintani, C. M. Walker, and S. A. Kalams. 2004. Limited T cell receptor diversity of HCV-specific T cell responses is associated with CTL escape. *J Exp Med* 200:307-19.
89. Moradpour, D., and H. E. Blum. 2005. Pathogenesis of hepatocellular carcinoma. *Eur J Gastroenterol Hepatol* 17:477-83.
90. Moradpour, D., V. Brass, E. Bieck, P. Friebe, R. Gosert, H. E. Blum, R. Bartenschlager, F. Penin, and V. Lohmann. 2004. Membrane association of the RNA-dependent RNA polymerase is essential for hepatitis C virus RNA replication. *J Virol* 78:13278-84.
91. Moradpour, D., M. J. Evans, R. Gosert, Z. Yuan, H. E. Blum, S. P. Goff, B. D. Lindenbach, and C. M. Rice. 2004. Insertion of green fluorescent protein into nonstructural protein 5A allows direct visualization of functional hepatitis C virus replication complexes. *J Virol* 78:7400-9.
92. Mueller, S., D. Papamichail, J. R. Coleman, S. Skiena, and E. Wimmer. 2006. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. *J Virol* 80:9687-96.
93. Myers, T. M., V. G. Kolupaeva, E. Mendez, S. G. Baginski, I. Frolov, C. U. Hellen, and C. M. Rice. 2001. Efficient translation initiation is

- required for replication of bovine viral diarrhea virus subgenomic replicons. *J Virol* 75:4226-38.
94. Nakai, K., T. Okamoto, T. Kimura-Someya, K. Ishii, C. K. Lim, H. Tani, E. Matsuo, T. Abe, Y. Mori, T. Suzuki, T. Miyamura, J. H. Nunberg, K. Moriishi, and Y. Matsuura. 2006. Oligomerization of hepatitis C virus core protein is crucial for interaction with the cytoplasmic domain of E1 envelope protein. *J Virol* 80:11265-73.
  95. Ogata, N., H. J. Alter, R. H. Miller, and R. H. Purcell. 1991. Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proc Natl Acad Sci U S A* 88:3392-6.
  96. Ogata, S., M. Nagano-Fujii, Y. Ku, S. Yoon, and H. Hotta. 2002. Comparative sequence analysis of the core protein and its frameshift product, the F protein, of hepatitis C virus subtype 1b strains obtained from patients with and without hepatocellular carcinoma. *J Clin Microbiol* 40:3625-30.
  97. Otto, G. A., P. J. Lukavsky, A. M. Lancaster, P. Sarnow, and J. D. Puglisi. 2002. Ribosomal proteins mediate the hepatitis C virus IRES-HeLa 40S interaction. *Rna* 8:913-23.
  98. Otto, G. A., and J. D. Puglisi. 2004. The pathway of HCV IRES-mediated translation initiation. *Cell* 119:369-80.
  99. Pause, A., G. Kukolj, M. Bailey, M. Brault, F. Do, T. Halmos, L. Lagace, R. Maurice, M. Marquis, G. McKercher, C. Pellerin, L. Pilote, D. Thibeault, and D. Lamarre. 2003. An NS3 serine protease inhibitor abrogates replication of subgenomic hepatitis C virus RNA. *J Biol Chem* 278:20374-80.
  100. Pavlovic, D., D. C. Neville, O. Argaud, B. Blumberg, R. A. Dwek, W. B. Fischer, and N. Zitzmann. 2003. 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.
  101. Petrakova, O., E. Volkova, R. Gorchakov, S. Paessler, R. M. Kinney, and I. Frolov. 2005. Noncytopathic replication of Venezuelan equine encephalitis virus and eastern equine encephalitis virus replicons in Mammalian cells. *J Virol* 79:7597-608.
  102. Pietschmann, T., A. Kaul, G. Koutsoudakis, A. Shavinskaya, S. Kallis, E. Steinmann, K. Abid, F. Negro, M. Dreux, F. L. Cosset, and R. Bartenschlager. 2006. Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. *Proc Natl Acad Sci U S A* 103:7408-13.



103. Pietschmann, T., V. Lohmann, G. Rutter, K. Kurpanek, and R. Bartenschlager. 2001. Characterization of cell lines carrying self-replicating hepatitis C virus RNAs. *J Virol* 75:1252-64.
104. Pileri, P., Y. Uematsu, S. Campagnoli, G. Galli, F. Falugi, R. Petracca, A. J. Weiner, M. Houghton, D. Rosa, G. Grandi, and S. Abrignani. 1998. Binding of hepatitis C virus to CD81. *Science* 282:938-41.
105. Pisarev, A. V., N. E. Shirokikh, and C. U. Hellen. 2005. Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites. *C R Biol* 328:589-605.
106. Pohlmann, S., J. Zhang, F. Baribaud, Z. Chen, G. J. Leslie, G. Lin, A. Granelli-Piperno, R. W. Doms, C. M. Rice, and J. A. McKeating. 2003. Hepatitis C virus glycoproteins interact with DC-SIGN and DC-SIGNR. *J Virol* 77:4070-80.
107. Prati, D. 2006. Transmission of hepatitis C virus by blood transfusions and other medical procedures: a global review. *J Hepatol* 45:607-16.
108. Quinkert, D., R. Bartenschlager, and V. Lohmann. 2005. Quantitative analysis of the hepatitis C virus replication complex. *J Virol* 79:13594-605.
109. Racaniello, V. R., and R. Ren. 1996. Poliovirus biology and pathogenesis. *Curr Top Microbiol Immunol* 206:305-25.
110. Ranjith-Kumar, C. T., and C. C. Kao. 2006. Recombinant viral RdRps can initiate RNA synthesis from circular templates. *Rna* 12:303-12.
111. Rehmann, B. 2000. Interaction between the hepatitis C virus and the immune system. *Semin Liver Dis* 20:127-41.
112. Rehmann, B., K. M. Chang, J. McHutchinson, R. Kokka, M. Houghton, C. M. Rice, and F. V. Chisari. 1996. Differential cytotoxic T-lymphocyte responsiveness to the hepatitis B and C viruses in chronically infected patients. *J Virol* 70:7092-102.
113. Rehmann, B., and M. Nascimbeni. 2005. Immunology of hepatitis B virus and hepatitis C virus infection. *Nat Rev Immunol* 5:215-29.
114. Reynolds, J. E., A. Kaminski, A. R. Carroll, B. E. Clarke, D. J. Rowlands, and R. J. Jackson. 1996. Internal initiation of translation of hepatitis C virus RNA: the ribosome entry site is at the authentic initiation codon. *Rna* 2:867-78.
115. Reynolds, J. E., A. Kaminski, H. J. Kettinen, K. Grace, B. E. Clarke, A. R. Carroll, D. J. Rowlands, and R. J. Jackson. 1995. Unique features of internal initiation of hepatitis C virus RNA translation. *Embo J* 14:6010-20.

116. Robek, M. D., B. S. Boyd, and F. V. Chisari. 2005. Lambda interferon inhibits hepatitis B and C virus replication. *J Virol* 79:3851-4.
117. Rouille, Y., F. Helle, D. Delgrange, P. Roingeard, C. Voisset, E. Blanchard, S. Belouzard, J. McKeating, A. H. Patel, G. Maertens, T. Wakita, C. Wychowski, and J. Dubuisson. 2006. Subcellular localization of hepatitis C virus structural proteins in a cell culture system that efficiently replicates the virus. *J Virol* 80:2832-41.
118. Roussel, J., A. Pillez, C. Montpellier, G. Duverlie, A. Cahour, J. Dubuisson, and C. Wychowski. 2003. Characterization of the expression of the hepatitis C virus F protein. *J Gen Virol* 84:1751-9.
119. Sakai, A., M. S. Claire, K. Faulk, S. Govindarajan, S. U. Emerson, R. H. Purcell, and J. Bukh. 2003. The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proc Natl Acad Sci U S A* 100:11646-51.
120. Sansonno, D., and F. Dammacco. 2005. Hepatitis C virus, cryoglobulinaemia, and vasculitis: immune complex relations. *Lancet Infect Dis* 5:227-36.
121. Santolini, E., G. Migliaccio, and N. La Monica. 1994. Biosynthesis and biochemical properties of the hepatitis C virus core protein. *J Virol* 68:3631-41.
122. Sarnow, P., R. C. Cevallos, and E. Jan. 2005. Takeover of host ribosomes by divergent IRES elements. *Biochem Soc Trans* 33:1479-82.
123. Shimizu, Y. K., A. J. Weiner, J. Rosenblatt, D. C. Wong, M. Shapiro, T. Popkin, M. Houghton, H. J. Alter, and R. H. Purcell. 1990. Early events in hepatitis C virus infection of chimpanzees. *Proc Natl Acad Sci U S A* 87:6441-4.
124. Shimoike, T., C. Koyama, K. Murakami, R. Suzuki, Y. Matsuura, T. Miyamura, and T. Suzuki. 2006. Down-regulation of the internal ribosome entry site (IRES)-mediated translation of the hepatitis C virus: critical role of binding of the stem-loop IIIId domain of IRES and the viral core protein. *Virology* 345:434-45.
125. Shimoike, T., S. Mimori, H. Tani, Y. Matsuura, and T. Miyamura. 1999. Interaction of hepatitis C virus core protein with viral sense RNA and suppression of its translation. *J Virol* 73:9718-25.
126. Shindo, M., A. M. Di Bisceglie, R. Biswas, K. Mihalik, and S. M. Feinstone. 1992. Hepatitis C virus replication during acute infection in the chimpanzee. *J Infect Dis* 166:424-7.
127. Shoukry, N. H., A. Grakoui, M. Houghton, D. Y. Chien, J. Ghrayeb, K. A. Reimann, and C. M. Walker. 2003. Memory CD8+ T cells are

- required for protection from persistent hepatitis C virus infection. *J Exp Med* 197:1645-55.
128. Simmonds, P. 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. *J Gen Virol* 85:3173-88.
  129. Simmonds, P., J. Bukh, C. Combet, G. Deleage, N. Enomoto, S. Feinstone, P. Halfon, G. Inchauspe, C. Kuiken, G. Maertens, M. Mizokami, D. G. Murphy, H. Okamoto, J. M. Pawlotsky, F. Penin, E. Sablon, I. T. Shin, L. J. Stuyver, H. J. Thiel, S. Viazov, A. J. Weiner, and A. Widell. 2005. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* 42:962-73.
  130. Simmonds, P., E. C. Holmes, T. A. Cha, S. W. Chan, F. McOmish, B. Irvine, E. Beall, P. L. Yap, J. Kolberg, and M. S. Urdea. 1993. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J Gen Virol* 74 ( Pt 11):2391-9.
  131. Simmonds, P., and D. B. Smith. 1997. Investigation of the pattern of diversity of hepatitis C virus in relation to times of transmission. *J Viral Hepat* 4 Suppl 1:69-74.
  132. Smith, D. B., J. Mellor, L. M. Jarvis, F. Davidson, J. Kolberg, M. Urdea, P. L. Yap, and P. Simmonds. 1995. Variation of the hepatitis C virus 5' non-coding region: implications for secondary structure, virus detection and typing. The International HCV Collaborative Study Group. *J Gen Virol* 76 ( Pt 7):1749-61.
  133. Smith, D. B., and P. Simmonds. 1997. Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both ends of the genome. *J Mol Evol* 45:238-46.
  134. Smith, D. B., and P. Simmonds. 1997. Review: molecular epidemiology of hepatitis C virus. *J Gastroenterol Hepatol* 12:522-7.
  135. Song, Y., P. Friebe, E. Tzima, C. Junemann, R. Bartenschlager, and M. Niepmann. 2006. The hepatitis C virus RNA 3'-untranslated region strongly enhances translation directed by the internal ribosome entry site. *J Virol* 80:11579-88.
  136. Spahn, C. M., J. S. Kieft, R. A. Grassucci, P. A. Penczek, K. Zhou, J. A. Doudna, and J. Frank. 2001. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science* 291:1959-62.
  137. Su, A. I., J. P. Pezacki, L. Wodicka, A. D. Brideau, L. Supekova, R. Thimme, S. Wieland, J. Bukh, R. H. Purcell, P. G. Schultz, and F. V.

- Chisari. 2002. Genomic analysis of the host response to hepatitis C virus infection. *Proc Natl Acad Sci U S A* 99:15669-74.
138. Svitkin, Y. V., A. Pause, M. Lopez-Lastra, S. Perreault, and N. Sonenberg. 2005. Complete translation of the hepatitis C virus genome in vitro: membranes play a critical role in the maturation of all virus proteins except for NS3. *J Virol* 79:6868-81.
  139. Tabor, E., R. J. Gerety, J. A. Drucker, L. B. Seeff, J. H. Hoofnagle, D. R. Jackson, M. April, L. F. Barker, and G. Pineda-Tamondong. 1978. Transmission of non-A, non-B hepatitis from man to chimpanzee. *Lancet* 1:463-6.
  140. Tanaka, Y., T. Shimoike, K. Ishii, R. Suzuki, T. Suzuki, H. Ushijima, Y. Matsuura, and T. Miyamura. 2000. Selective binding of hepatitis C virus core protein to synthetic oligonucleotides corresponding to the 5' untranslated region of the viral genome. *Virology* 270:229-36.
  141. Tellinghuisen, T. L., J. Marcotrigiano, and C. M. Rice. 2005. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. *Nature* 435:374-9.
  142. Tellinghuisen, T. L., and C. M. Rice. 2002. Interaction between hepatitis C virus proteins and host cell factors. *Curr Opin Microbiol* 5:419-27.
  143. Thimme, R., J. Bukh, H. C. Spangenberg, S. Wieland, J. Pemberton, C. Steiger, S. Govindarajan, R. H. Purcell, and F. V. Chisari. 2002. Viral and immunological determinants of hepatitis C virus clearance, persistence, and disease. *Proc Natl Acad Sci U S A* 99:15661-8.
  144. Thomson, M., M. Nascimbeni, M. B. Havert, M. Major, S. Gonzales, H. Alter, S. M. Feinstone, K. K. Murthy, B. Rehermann, and T. J. Liang. 2003. The clearance of hepatitis C virus infection in chimpanzees may not necessarily correlate with the appearance of acquired immunity. *J Virol* 77:862-70.
  145. Tomei, L., C. Failla, E. Santolini, R. De Francesco, and N. La Monica. 1993. NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *J Virol* 67:4017-26.
  146. Tscherne, D. M., C. T. Jones, M. J. Evans, B. D. Lindenbach, J. A. McKeating, and C. M. Rice. 2006. Time- and temperature-dependent activation of hepatitis C virus for low-pH-triggered entry. *J Virol* 80:1734-41.
  147. Tuplin, A., D. J. Evans, and P. Simmonds. 2004. Detailed mapping of RNA secondary structures in core and NS5B-encoding region sequences of hepatitis C virus by RNase cleavage and novel bioinformatic prediction methods. *J Gen Virol* 85:3037-47.

148. Uprichard, S. L., J. Chung, F. V. Chisari, and T. Wakita. 2006. Replication of a hepatitis C virus replicon clone in mouse cells. *Virology* 3:89.
149. Varaklioti, A., N. Vassilaki, U. Georgopoulou, and P. Mavromara. 2002. Alternate translation occurs within the core coding region of the hepatitis C viral genome. *J Biol Chem* 277:17713-21.
150. Vassilaki, N., and P. Mavromara. 2003. Two alternative translation mechanisms are responsible for the expression of the HCV ARFP/F/core+1 coding open reading frame. *J Biol Chem* 278:40503-13.
151. Wakita, T., T. Pietschmann, T. Kato, T. Date, M. Miyamoto, Z. Zhao, K. Murthy, A. Habermann, H. G. Krausslich, M. Mizokami, R. Bartenschlager, and T. J. Liang. 2005. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med* 11:791-6.
152. Walewski, J. L., J. A. Gutierrez, W. Branch-Elliman, D. D. Stump, T. R. Keller, A. Rodriguez, G. Benson, and A. D. Branch. 2002. Mutation Master: profiles of substitutions in hepatitis C virus RNA of the core, alternate reading frame, and NS2 coding regions. *Rna* 8:557-71.
153. Walewski, J. L., T. R. Keller, D. D. Stump, and A. D. Branch. 2001. Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *Rna* 7:710-21.
154. Walker, C. M. 1997. Comparative features of hepatitis C virus infection in humans and chimpanzees. *Springer Semin Immunopathol* 19:85-98.
155. Wang, T. H., R. C. Rijnbrand, and S. M. Lemon. 2000. Core protein-coding sequence, but not core protein, modulates the efficiency of cap-independent translation directed by the internal ribosome entry site of hepatitis C virus. *J Virol* 74:11347-58.
156. Wasley, A., and M. J. Alter. 2000. Epidemiology of hepatitis C: geographic differences and temporal trends. *Semin Liver Dis* 20:1-16.
157. Weihofen, A., K. Binns, M. K. Lemberg, K. Ashman, and B. Martoglio. 2002. Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science* 296:2215-8.
158. Weiner, A., A. L. Erickson, J. Kansopon, K. Crawford, E. Muchmore, A. L. Hughes, M. Houghton, and C. M. Walker. 1995. Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a cytotoxic T lymphocyte escape variant. *Proc Natl Acad Sci U S A* 92:2755-9.

159. WHO. 2000. Fact Sheet Hepatitis C Virus.
160. Xu, Z., J. Choi, W. Lu, and J. H. Ou. 2003. Hepatitis C virus f protein is a short-lived protein associated with the endoplasmic reticulum. *J Virol* 77:1578-83.
161. Xu, Z., J. Choi, T. S. Yen, W. Lu, A. Strohecker, S. Govindarajan, D. Chien, M. J. Selby, and J. Ou. 2001. Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *Embo J* 20:3840-8.
162. Yi, M., and S. M. Lemon. 2003. 3' nontranslated RNA signals required for replication of hepatitis C virus RNA. *J Virol* 77:3557-68.
163. Yi, M., and S. M. Lemon. 2004. Adaptive mutations producing efficient replication of genotype 1a hepatitis C virus RNA in normal Huh7 cells. *J Virol* 78:7904-15.
164. Yi, M., Y. Ma, J. Yates, and S. M. Lemon. 2007. Compensatory mutations in E1, p7, NS2, and NS3 enhance yields of cell culture-infectious intergenotypic chimeric hepatitis C virus. *J Virol* 81:629-38.
165. You, S., D. D. Stump, A. D. Branch, and C. M. Rice. 2004. A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for hepatitis C virus RNA replication. *J Virol* 78:1352-66.
166. Zhao, W. D., and E. Wimmer. 2001. Genetic analysis of a poliovirus/hepatitis C virus chimera: new structure for domain II of the internal ribosomal entry site of hepatitis C virus. *J Virol* 75:3719-30.
167. Zhao, W. D., E. Wimmer, and F. C. Lahser. 1999. Poliovirus/Hepatitis C virus (internal ribosomal entry site-core) chimeric viruses: improved growth properties through modification of a proteolytic cleavage site and requirement for core RNA sequences but not for core-related polypeptides. *J Virol* 73:1546-54.