Title
Combining high-resolution genetics and imaging for the study of hepatitis C virus proteins critical for HCV assembly in infected host cells

Permalink
https://escholarship.org/uc/item/1sq5j3vk

Author
Remenyi, Roland Gilbert

Publication Date
2014

Peer reviewed|Thesis/dissertation
Combining high-resolution genetics and imaging for the study of hepatitis C virus proteins critical for HCV assembly in infected host cells

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular and Medical Pharmacology

by

Roland Gilbert Remenyi

2014
ABSTRACT OF THE DISSERTATION

Combining high-resolution genetics and imaging for the study of hepatitis C virus proteins critical for HCV assembly in infected host cells

by

Roland Gilbert Remenyi
Doctor of Philosophy in Molecular and Medical Pharmacology
University of California, Los Angeles, 2014
Professor Ren Sun, Chair

The establishment of a cell culture system for producing infectious hepatitis C virus (HCV) prompted genetic and functional studies of viral proteins and their roles in the assembly process. Since then, all ten viral proteins have been implicated in HCV assembly. Nonetheless, the exact location of the assembly site within an infected host cell remains unknown. Moreover, an understanding of the chronology of events and individual protein contributions at different stages of the assembly process has been difficult to obtain.

The two studies comprising this dissertation apply high-resolution genetics and high-resolution imaging to the study of HCV in cell culture. Study 1 employs a high-resolution genetics approach to reveal a functional map of the entire HCV genome. Next-generation sequencing of an insertion mutant library following passage in cell culture revealed genetic footprints that reflected known biological functions of the underlying protein regions. We show how these genetic footprints can serve as a resource to identify flexible regions that tolerate insertion of tags useful for a variety of protein detection methods. Moreover, using the genetic footprints, we identify a region in the NS4B protein that plays a role in post-RNA-replication steps. Study 2 examines HCV assembly through imaging, applying electron microscopy,
electron tomography, superresolution light microscopy, multi-color fluorescence microscopy and live-cell imaging to the study of virus assembly. Our results indicate a juxtaposition of LDs, virus-like particles, membrane vesicles, clusters of HCV core protein and areas containing core-E2-NS5A proteins. The high-resolution snapshots underscore the functional compartmentalization of the LD environment, which provides viral proteins with membranous platforms to carry out a complex process such as virion assembly. We also show how our imaging platform can aid in the phenotypic characterization of an assembly-deficient mutant NS5A virus.

The two studies presented in this dissertation further our understanding of the contributions of non-structural proteins such as NS4B and NS5A to the HCV assembly process. We suggest that the combination of the high-resolution genetic platform of study 1 and the high-resolution imaging platform of study 2 facilitates the identification and phenotypic characterization of viral protein regions involved in HCV assembly. A streamlined approach that integrates these two methods has the potential to identify additional targets for therapeutic intervention at post-genome-replication steps.
The dissertation of Roland Gilbert Remenyi is approved.

Samson Chow
Asim Dasgupta
Thomas Graeber
Otto Yang
Ren Sun, Committee Chair

University of California, Los Angeles
2014
This dissertation is dedicated to my parents Elizabeth and Eugen Remenyi who are the reason I stand where I stand today and my grandparents who have left us too soon but are watching from above.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT OF THE DISSERTATION</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATIONS</td>
<td>v</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>xiii</td>
</tr>
<tr>
<td>VITA</td>
<td>xvi</td>
</tr>
<tr>
<td>PUBLICATIONS AND PRESENTATIONS</td>
<td>xvi</td>
</tr>
<tr>
<td>CHAPTER 1</td>
<td>1</td>
</tr>
<tr>
<td>Overview and specific aims</td>
<td></td>
</tr>
<tr>
<td>SPECIFIC AIMS FOR DISSERTATION</td>
<td>2</td>
</tr>
<tr>
<td>Aim 1: To systematically identify essential and non-essential areas within the entire HCV genome at high resolution</td>
<td>3</td>
</tr>
<tr>
<td>Aim 2: To acquire high-resolution snapshots of viral assembly sites in HCV-infected cells</td>
<td>4</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td>7</td>
</tr>
<tr>
<td>Background</td>
<td></td>
</tr>
<tr>
<td>Hepatitis C infection, a modern-day pandemic</td>
<td>8</td>
</tr>
<tr>
<td>The molecular virology and life cycle of HCV</td>
<td>9</td>
</tr>
<tr>
<td>Surrogate systems enable the study of HCV infection in cell culture</td>
<td>9</td>
</tr>
<tr>
<td>Cultured Huh-7 cells have an inefficient VLDL assembly pathway</td>
<td>11</td>
</tr>
<tr>
<td>Multiple viral proteins appear to play a role in the HCV assembly process occurring close to lipid droplets (LDs)</td>
<td>12</td>
</tr>
<tr>
<td>The potential of advanced microscopy techniques for imaging of HCV assembly</td>
<td>12</td>
</tr>
<tr>
<td>References</td>
<td>15</td>
</tr>
<tr>
<td>CHAPTER 3</td>
<td>18</td>
</tr>
<tr>
<td>Profiling of 15-nucleotide insertion mutations by next-generation sequencing reveals a comprehensive functional map of the hepatitis C virus genome</td>
<td>18</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>19</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>20</td>
</tr>
<tr>
<td>RESULTS</td>
<td>23</td>
</tr>
<tr>
<td>Transposon insertion sequencing precisely maps complex HCV mutant libraries across different selection pools</td>
<td>24</td>
</tr>
</tbody>
</table>
Selection profiles within individual HCV genes vary according to gene’s encoded function ................................................................. 26
HCV profile reveals bias for particular amino-acid-insertions within special functional areas of the genome ................................................................................................................................. 28
Functional annotation of the C-terminal region of the core protein supports biological relevance of protein structure ................................................................................................................................. 29
Functional annotation of transmembrane protein structures supports their biological relevance ................................................................................................................................. 31
Mapping of tolerated insertion mutants to protein structures of p7 viroporin and the viral polymerase NS5B reveals structural context for insertion toleration ............................................. 33
Identification of nonessential regions aids in the design of viable epitope-tagged viruses ............................................................................................................................................... 34
Further filtering of transposon mutant database uncovers new functional regions encoded by HCV genes ................................................................................................................................. 36

DISCUSSION ................................................................................................................................................................................................. 37
Genetic footprints reflect HCV biology .............................................................................................................................................. 40
Genetic footprints complement structural studies of HCV proteins ................................................................................................. 42
Genetic footprint resource enables extraction of new information regarding HCV biology ................................................................................................................................. 45
Perspectives for transposon insertion sequencing in HCV biology .................................................................................................................. 48
Input library contains at least one insertion mutant between nearly every amino acid encoded by the HCV genome ....................................................................................................................................... 48
Profiling method improves on previous approaches to determine functional maps of entire viral genomes .............................................................................................................................................. 38

METHODS ................................................................................................................................................................................................. 49
TABLES AND FIGURES ................................................................................................................................................................. 55
REFERENCES ................................................................................................................................................................................................. 74

CHAPTER 4 ................................................................................................................................................................................................. 79
High-resolution imaging of the LD microenvironment in HCV-infected cells ................................................................................................................................. 79
ABSTRACT ................................................................................................................................................................................................. 80
INTRODUCTION ................................................................................................................................................................................................. 81
RESULTS ................................................................................................................................................................................................. 85
Electron microscopy analysis of infected cells reveals the juxtaposition of lipid droplets, membrane vesicles and endoplasmic reticulum .................................................................................. 85
Electron tomography provides three-dimensional views of membrane alterations surrounding LDs .............................................................................................................................................. 86
Particles with sizes of 50-75nm are found at various subcellular localizations, including areas enriched in membrane vesicles and ER ......................................................... 88
Free VLPs as well as membrane-vesicle-associated VLPs locate close to LD interfaces................................................................................................................................. 89
Electron tomography provides three-dimensional views of spherical particles close to LDs ................................................................................................................................. 90
Immunoelectron microscopy confirms presence of HCV core and NS5A proteins within LD microenvironment........................................................................................................ 91
Clusters of HCV core protein are close to LDs but show variability in size, shape, distribution patterns and E2 co-localization ............................................................ 93
Development of a panel of infectious reporter viruses containing NS5A tagged with fluorescent proteins ........................................................................................................ 95
STED microscopy of FNX\textsuperscript{NS5A-YPet}-infected cells reveals presence of 110-150nm-sized NS5A clusters .................................................................................................. 98
Infectious reporter viruses containing NS5A tagged with YPet facilitates analysis of NS5A, core and E2 localization around LDs ............................................................... 99
Phenotypic characterization of a point mutant in NS5A domain IA using the FNX\textsuperscript{NS5A-YPet}-based fluorescent imaging system .................................................................. 101
DISCUSSION.................................................................................................................. 104
Infected cells contain an interconnected microenvironment of LDs, membrane vesicles and ER ......................................................................................................................... 105
LD microenvironment contains virus-like particles of varying sizes and densities close to LD surface ............................................................................................................. 107
Core and NS5A proteins localize in clusters of varying sizes and co-localize with E2 in specific areas close to LDs .................................................................................. 109
Fluorescent reporter viruses facilitate dynamic tracking of NS5A for mutant viruses and during treatment with antiviral compounds .............................................. 112
METHODS ....................................................................................................................... 115
FIGURES ......................................................................................................................... 120
REFERENCES ................................................................................................................ 136
CHAPTER 5 ..................................................................................................................... 142
Summary and future outlook ....................................................................................... 142

Specific Aim 1: To systematically identify essential and non-essential areas within the entire HCV genome at high resolution ........................................................................ 143
Specific Aim 2: To acquire high-resolution snapshots of viral assembly sites in HCV-infected cells .................................................................................................................. 143
Future Outlook ............................................................................................................. 144
References ..................................................................................................................... 148
LIST OF TABLES

Table 3-1. Summary of high-throughput sequencing of recovered HCV libraries ..................55
Table 3-2. Gene-by-gene summary of high-resolution functional profile .........................56
Table 3-3. Summary of non-lethal insertions in globular domains of NS5B polymerase ........57
Table 3-4. Overview of inserted epitope tag sequences ...............................................58
Table 3-5. Filtering of functional profile produces a candidate list of non-structural HCV protein regions with roles in post-replication steps .........................................................59
LIST OF FIGURES

Figure 3-1. Strategy for functional profiling of HCV insertion libraries by massively parallel sequencing ..........................................................60

Figure 3-2. Functional map of the entire HCV genome ........................................61

Figure 3-3. Examination of P1/P0 and P2/P0 ratios across both the entire HCV genome and individual viral genes ..........................................................62

Figure 3-4. Graphing P1/P0 and P2/P0 values for AAA, CGR and RPH/Q insertion motifs across the HCV genome identifies regions of preference for certain motifs .........................63

Figure 3-5. Effect of location of 15-nt insertion relative to reading frame in select areas of viral genome .................................................................................64

Figure 3-6. Functional annotation of the C-terminal region of the HCV core protein ..........65

Figure 3-7. Validation of an essential role for C-terminal region of HCV core in infectious virus production .............................................................................66

Figure 3-8. Functional annotation of viral transmembrane proteins p7 and NS2 using P2/P1 ratios ..............................................................................................67

Figure 3-9. Surface views of annotated p7 viroporin structure ........................................68

Figure 3-10. Mapping tolerated insertions to the structure of NS5B polymerase ...............69

Figure 3-11. Screening dataset for hotspots tolerating 5-amino-acid-insertions ..................70

Figure 3-12. Engineering of infectious epitope-tagged HCV ........................................71

Figure 3-13. Multi-color fluorescence microscopy of cells infected with epitope-tagged NS5A virus .................................................................................72
Figure 3-14. Validation of NS4B insertion mutants with defects in post-replication steps .............................................................................................................................................73

Figure 4-1. Electron microscopy analysis of ultrastructural changes close to lipid droplets (LDs) in Huh-7.5.1 cells infected with HCV .............................................................................................................................................120

Figure 4-2. Electron microscopy analysis of ultrastructural changes within LD-rich region in Huh-7.5.1 cells infected with HCV .............................................................................................................................................121

Figure 4-3. Electron tomography of infected cell provides three-dimensional snapshots of membrane structures close to LDs .............................................................................................................................................122

Figure 4-4. Electron microscopy analysis captures snapshots of virus-like particles in HCV-infected cells .............................................................................................................................................123

Figure 4-5. Electron microscopy provides snapshots of VLPs close to LD boundary .............124

Figure 4-6. Electron tomography of infected cell delivers three-dimensional views of VLPs close to LDs .............................................................................................................................................125

Figure 4-7. Immunolocalization of HCV core and NS5A proteins in infected cells .............126

Figure 4-8. Superresolution light microscopy analysis of HCV core protein in infected cells .............................................................................................................................................127

Figure 4-9. Examination of core-positive clusters in the vicinity of LDs by superresolution light microscopy .............................................................................................................................................128

Figure 4-10. Fluorescent reporter viruses facilitate tracking of NS5A protein in live HCV-infected cells .............................................................................................................................................129

Figure 4-11. Fluorescent reporter proteins co-localize with viral NS5A in infected cells .............131

Figure 4-12. STED microscopy visualizes NS5A subcellular localization at high resolution ......132
Figure 4-13. Multi-color fluorescence microscopy reveals areas of NS5A, E2 and core protein co-localization close to LDs ................................................................. 133

Figure 4-14. Phenotypic characterization of NS5A I52N mutant using FNX\textsuperscript{YPET-NS5A} fluorescent imaging system ................................................................. 134

Figure 4-15. Live-cell images of daclatasvir-treated cells showing redistribution of NS5A in punctate pattern similar to I52N phenotype ................................................................. 135
As I am coming to the end of the road of the PhD journey, there are numerous people who have helped me along the way. Advisors, role-models, mentors, friends, family, collaborators, supporters, teammates, labmates, the list goes on and on but the message stays the same: I couldn’t have done it without them.

First and foremost, I thank my advisor Ren Sun for creating a wonderful environment for the pursuit of exciting scientific questions. I think back fondly to the moment we had our first discussion about research interests during a hike at Mt. Baldy my first year as a graduate student. It has been a privilege to train as his student. I am thankful for his constant support and guidance over the years. His commitment to doing excellent science and striving to ceaselessly break new grounds are qualities that I only hope to replicate in the career ahead of me.

Many faculty members have been instrumental in this achievement. I especially would like to thank my committee members, Drs. Sam Chow, Tom Graeber, Asim Dasgupta and Otto Yang. Their comments have made this body of work better. Their support has made me stronger. Extra thanks has to go to my pharmacology rotation advisors Drs. Sam Chow, Anna Wu, Sanggu Kim and Eric Lepin. Thanks also to Dr. Vaithilingaraja Arumugaswami, who took me under his wings initially and sparked my passion for the study of HCV. Without has patience and careful instruction none of this would have been possible. Dr. Ting-Ting Wu has also provided steady, caring guidance during my time as a graduate student and her love for science inspired me every day.

I would like to thank all the collaborators. All of the electron microscopy work would not have been possible without Drs. Hong Zhou, Li (Tony) Peng, Louise Hughes, Ivo Atanasov, Jiansen Li and Sergey Ryazantsev. Thanks has to also go out to the light microscopy core facility (ALMS) and its directors Laurent Bentolila and Matthew Schibler.
I am grateful for the past and present members of the Sun lab. Dr. Hangfei Qi has been instrumental in the progression of this thesis, always sharing reagents, always available for discussion about HCV and always there to provide scientific and emotional support. Part of the experiments in chapter 3 were designed and carried out by Dr. Qi entirely and previously presented in her dissertation (Qi, H., 2012. High-Resolution Functional Profiling Hepatitis C Virus Genome. ProQuest Dissertations and Theses. Ann Arbor, University of California, Los Angeles: 130). Dr. Qi graciously shared the next-generation sequencing data for the additional data analysis presented in my dissertation and the application of the functional profile described here. Moreover, she initially discovered the I52N mutant virus in the above thesis work and was open to a collaboration for the additional phenotypic characterization of the virus mutant in Chapter 4 of this dissertation. Similarly Nicholas Wu played a crucial role in the data analysis of this project. I truly admire his ability to bridge bioinformatics with virology. I also have to thank my undergraduate students, Tamar Stokelman and Jerry Lo, who contributed important reagents and who have made teaching and mentoring fun. And to all of the Sun lab members for their technical and research support, as well as the many laughs shared over the past years.

Many thank you wishes have to go out to the pharmacology student affairs officers, who have made the administrative part of being a graduate run smoothly. Sarah Starrett has contributed to my career path from the moment I interviewed at UCLA and Aija Gamburg and Emily Fitch have continued to be an invaluable resource for pharmacology graduate students.

To my pharmacology classmates, to dear friends at UCLA or outside, who have enriched my life in so many ways and kept me going when the going got hard. These friendships have kept me grounded and allowed me to never settle for less.

Finally, I thank my parents, Elizabeth and Eugen Remenyi. Their love, foresight, sacrifices and unending optimism have made this achievement possible. They have taught me that races can be won with the slow and steady approach of the turtle. I am forever thankful.
This research was supported by the UCLA Virology and Gene Therapy training program.

Chapter 3 is a version of the following manuscript:


Chapter 4 is a version of the following manuscript:


We thank Dr. Francis Chisari (the Scripps Research Institute, La Jolla) for kindly providing the Huh-7.5.1 cell line. We are grateful to Dr. S. Foung for provision of the human anti-E2 antibody CBH5 and to Dr. Charles Rice for the 9E10 antibody. Confocal laser scanning microscopy was performed at the Department of Anesthesiology (thanks to Yibing Wang and Hongmei Ruan) and the CNSI Advanced Light Microscopy/Spectroscopy Shared Resource Facility at UCLA, supported with funding from NIH-NCRR shared resources grant (CJX1-443835-WS-29646) and NSF Major Research Instrumentation grant (CHE-0722519). We thank Dr. Laurent Bentolila and Dr. Matthew Schibler for their assistance in microscopy experiments. Electron microscopy imaging was performed at the electron imaging center for nanomachines (EICN). We would like to thank Dr. Hong Zhou and Ivo Atanasov for their assistance in the experimental design and carrying-out of EM imaging. We are also grateful to Sergey Ryazantsev at the EMi for providing consultation for plastic embedding. Finally, we have to give thanks for the technical consultation for high-pressure freezing given by Zhuo Li at City of Hope and Paul Webster at the House Ear Institute.
VITA

2006        Sc.B. with Honors in Molecular Biology and Biochemistry
Brown University, Providence, RI

2006-2013  Graduate Student Researcher
Department of Molecular and Medical Pharmacology
UCLA

2007-2010  Graduate Student Trainee
Interdisciplinary Training Program in Gene Therapy/Virology
UCLA

2011        Teaching Assistant
Department of Molecular, Cell and Developmental Biology
UCLA

PUBLICATIONS AND PRESENTATIONS


Southern California Society of Microscopy and Microanalysis annual symposium (Oral Presentation), Los Angeles. Title: High-resolution imaging of HCV-infected cells. 3/30/2013

17th International Symposium on Hepatitis C Virus and Related Viruses (Oral Presentation), Yokohama. Title: Ultrastructural characterization of hepatitis C virus infection, 09/10-09/14/2010

Advanced Electron Microscopy in NanoMedicine Symposium (Poster Presentation), Los Angeles. Title: Correlative cryo-EM of hepatitis C virus infection. 10/02-10/03/2009

15th International Symposium on Hepatitis C Virus and Related Viruses (Poster Presentation), San Antonio. Title: HCV core - how a high-throughput functional profile lead to the identification of residues critical for infectious virus production, 10/05-10/09/2008
CHAPTER 1

Overview and specific aims
Upon infecting a host cell, the hepatitis C virus (HCV) induces a myriad of cellular changes and orchestrates distinct stages through which the virus needs to progress for its own survival. Following viral entry, uncoating and release of the viral RNA genome, this RNA genome serves as the template for the translation of viral proteins at the endoplasmic reticulum (ER). These viral proteins, through interactions among viral proteins themselves or through interactions with cellular factors, sequentially participate in the ensuing steps of the viral life cycle, namely genome replication and assembly of viral protein components into progeny virions. Before the first infectious cell culture system for the study of the HCV life cycle became available in 2005, the study of individual viral proteins was only possible in surrogate systems that lacked the capacity to produce infectious particles. With the arrival of an infectious cell culture system, many studies have since strived to address how viral protein components come together in infected cells, establish replication compartments, and assemble progeny virions. However, virus assembly remains one of the less-understood steps in the virus life cycle. My thesis work aims to fill this gap of knowledge by applying technical advances in high-resolution genetics and high-resolution imaging to the study of HCV assembly. Moreover, in order to dissect the various steps involved in the assembly process, we aim to develop protocols to track viral proteins in infected cells and to selectively block assembly steps through genetic mutations or perturbations with chemical compounds.

SPECIFIC AIMS FOR DISSERTATION

The main objective of this dissertation was to combine high-resolution genetics and high-resolution imaging in the study of HCV proteins in virus-producing cells. While the high-resolution genetics approach allowed us to systematically probe the entire HCV genome for viral protein regions essential at various steps of the HCV life cycle, including post-genome-
replication steps, the high-resolution imaging approach allowed us to acquire snapshots of the intracellular microenvironment in which the viral proteins carried out their functions.

A related objective is to further characterize the HCV assembly site and better understand the role of important mediators of HCV assembly, in particular the NS5A protein, whose dual role in both virus replication and assembly remains incompletely understood. To do so, we used high-resolution imaging techniques and advances in protein labeling to carry out imaging studies in both live and fixed cells. We show the application of these imaging techniques in the phenotypic characterization of a point mutant in domain I of NS5A. The characteristic phenotype, a change in subcellular localization of NS5A, also arises following treatment with a potent viral inhibitor targeting NS5A domain I. The mutation was originally discovered in our laboratory by saturation mutagenesis of domain I of NS5A. Therefore, we provide a proof-of-concept that high-resolution genetics can be successfully combined with high-resolution imaging to discover amino acid residues essential for later steps of the viral life cycle, such as HCV assembly.

The following paragraphs will describe the specific aims of my doctoral dissertation and present the two studies that address these aims.

Aim 1: To systematically identify essential and non-essential areas within the entire HCV genome at high resolution

Project 1: Combining 15-nucleotide insertion mutagenesis, genetic footprinting and next-generation sequencing to provide high-resolution functional information in a high-throughput manner.

Chapter 3: Profiling of 15-nucleotide insertion mutations by next-generation sequencing reveals a comprehensive functional map of the hepatitis C virus genome

(Remenyi et al., Ready for submission to Genome Research)
Pairing high-throughput sequencing technologies with high-throughput functional studies enables genome-wide studies of pathogenic organisms. A single experiment can uncover a comprehensive landscape of these pathogens’ essential and non-essential genes. In this chapter, we describe the application of next-generation sequencing (NGS) to a complete viral genome. We generated a transposon insertion library that contains mutants with 15-nucleotide (15-nt) insertions randomly distributed across the entire hepatitis C virus (HCV) genome. Using a genetic footprinting approach, we passaged the library in hepatic cells, recovered passaged virus pools, and used Illumina sequencing to simultaneously assay the abundance of mutant viruses in each pool. We provide evidence that the library’s genetic footprints reflected the biology of HCV. Furthermore, we show how these genetic footprints can serve as a resource for gaining new insight on functional HCV regions. With this approach we identified a region in the non-structural protein NS4B that plays a role in post-RNA-replication steps of the viral life cycle. This resource, which we are making freely available to the public, will allow researchers to easily evaluate the effect of 15-nt insertions at 99.3% of HCV codons. Once cloned as DNA, a viral genome can serve as the starting point for this transposon insertion sequencing approach, opening the door to systematic functional analyses of viruses for which infectious clones are available. Such analyses are now precise enough to determine both the exact amino acid sequence of insertion mutations as well as the resulting effect on viral fitness. Thus, this approach, which we refer to as high-resolution genetics, creates comprehensive lists detailing how insertion mutations affect various infection processes.

**Aim 2: To acquire high-resolution snapshots of viral assembly sites in HCV-infected cells**

**Project 2:** Characterizing the ultrastructure of infected cells and locating viral proteins at high resolution by electron microscopy, light microscopy and live-cell imaging.

**Chapter 4:** High-resolution imaging of the LD microenvironment in HCV-infected cells (Remenyi et al., In Preparation)
Current models of the assembly of hepatitis C virus (HCV) propose that cellular fat storage organelles, called lipid droplets (LDs), provide a platform for potential assembly sites. The colocalization of various viral proteins within this LD microenvironment has been shown to be essential for infectious virus production. However, the topological organization of this LD microenvironment and the exact location of HCV assembly sites remain unclear. In this chapter, we used a combination of electron microscopy, electron tomography, immunoelectron microscopy, superresolution light microscopy and multi-color confocal laser microscopy to analyze the spatial architecture and distribution of viral proteins in the LD microenvironment. Ultrastructural changes occurring near LDs included an accumulation of membrane vesicles (reminiscent of a “membranous web” structure), ER cisternae and particles resembling virions with diameters ranging from 30nm-80nm. Electron tomography clearly identified double-membrane vesicles in this environment, as well as particles budding from ER cisternae. To characterize the localization of viral proteins around LDs, we performed immuno-EM, superresolution light microscopy and multi-fluorescence microscopy of HCV-infected cells. We confirmed that HCV core protein formed clusters close to the edge of LDs, which often overlapped with areas containing E2 glycoprotein. Using superresolution light microscopy we were able to detect core-positive and NS5A-positive clusters with sizes between 100-150nm. Multi-color fluorescence microscopy also revealed co-localization of core, E2 and NS5A in larger 400-800nm clusters close to LDs. To complement the static light and electron microscopy images of fixed cells, we developed a panel of fluorescent reporter viruses for live-cell tracking of the NS5A protein, an essential protein for virus assembly. We show that a substitution mutation in the NS5A protein blocks the HCV life cycle at a specific step, leading to accumulation of core, E2 and NS5A close to LDs. Pharmacological treatment with Daclatasvir, a potent HCV antiviral targeting the NS5A protein, resulted in a similar change in NS5A localization. Our results provide high-resolution snapshots of HCV-induced structures in the LD microenvironment, improving our understanding of the sequence of HCV assembly events.
Moreover, we demonstrate an essential function of a NS5A domain I residue in post-replication steps of the HCV life cycle. Taken together, our results demonstrate the advantage of combining viral genetic approaches with imaging modalities such as superresolution light microscopy, electron tomography or live-cell fluorescence microscopy in dissecting complex cellular processes of viral infection.
CHAPTER 2

Background
Hepatitis C infection, a modern-day pandemic

Hepatitis C virus (HCV) has emerged as a major cause of human liver disease worldwide, with an estimated 160 million individuals (2.35% of the world’s population) persistently infected with the virus (Lavanchy 2011). Moreover, back-calculation models based on US epidemiological data estimate a rise in HCV mortality over the next 20 years (Deuffic-Burban, Poynard et al. 2007). The recent wave of increased HCV-related morbidity and mortality can be traced back to an unprecedented increase in the spread of HCV during the 20th century; this rise appears to be associated with the widespread availability of injectable therapies and the illicit use of injectable drugs (Alter 2007). Most infections become chronic: a condition that is incurable in many patients, leading to chronic hepatitis, cirrhosis, and hepatocellular carcinoma. Chronic HCV infection is a major risk factor for end stage liver disease and liver transplantation. Currently there is no vaccine, and the only available treatment option with pegylated interferon is expensive, has significant side effects and offers only limited response rates. Consequently, each year, there are an estimated 18,000 deaths attributable to HCV in the United States (El-Kamary, Jhaveri et al. 2011).

The recent federal approval of the first direct-acting HCV antivirals (targeting the viral protease NS3 and polymerase NS5B) surely marks a first step in the eradication of the HCV pandemic. Nonetheless, the lessons from treatment of human immunodeficiency virus (HIV) infections, such as the emergence of drug resistance and the need for combination therapy, warrant a continuation of the search for additional therapeutic targets for HCV. Also, it remains unclear how both patients with existing liver disease and patients co-infected with additional viruses such as HIV or Hepatitis B virus (HBV), who have been among the hardest to treat, will respond to next-generation antivirals. Effective treatment of these patients will almost certainly require a combination of antivirals that minimizes toxicities, drug-drug interactions and
debilitating side effects. Therefore it is imperative to stockpile the arsenal of HCV antivirals with as many weapons targeting as many steps in the viral life cycle as possible.

**The molecular virology and life cycle of HCV**

Hepatitis C virus (HCV) is a member of the *Flaviviridae* and contains a single-stranded positive-sense RNA genome 9.6kb in size. Extracellular HCV virions undergo receptor-mediated endocytosis into a low-pH vesicle once they interact with receptor molecules at the cell surface of host cells, which are mainly hepatocytes in the liver. Genomic RNA released into the cytoplasm is then translated using an internal ribosome entry site (IRES) to generate a single large polyprotein that is processed into ten mature HCV proteins. HCV researchers classify the viral proteins into two groups: the structural proteins (core protein, envelope proteins E1 and E2), which form the proteinaceous components of viral particles, and non-structural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, NS5B), which have accessory functions within the host cell but are not incorporated into virions. Non-structural proteins are important in formation of replication complexes (RC) found associated with a virus-derived ER-like membrane structure termed the membranous web (MW). Mature HCV proteins replicate the RNA genome via a minus-strand replicative intermediate to produce progeny RNA. A portion of this newly synthesized RNA is packaged into nucleocapsids formed by core protein and associated with the HCV glycoproteins, presumably leading to budding into the ER, as observed for related flaviviruses. Current models for HCV egress propose that mature virions then exit from the cell by hijacking the cellular secretory pathway, thereby completing the life cycle.

**Surrogate systems enable the study of HCV infection in cell culture**

To-date, the routine propagation of serum-derived HCV in cell culture has not been successful. Instead, HCV researchers have turned to surrogate systems allowing the study of HCV in cell culture. For example, the development of subgenomic replicon systems opened the door to
molecular studies of RNA replication. Groundbreaking findings showed that subgenomic replicons of subtypes 1b replicated autonomously in selected subclones of the Huh-7 hepatoma cell line (Lohmann, Korner et al. 1999). Further improvements led to a robust, cell-culture-based system for genetic and functional analyses of RNA replication (Blight, Kolykhalov et al. 2000). In subsequent years, several groups reported model systems to study HCV entry in cell culture, using replication-deficient retroviruses pseudotyped with HCV envelope glycoproteins (HCVpp) (Bartosch, Dubuisson et al. 2003; Drummer, Maerz et al. 2003; Hsu, Zhang et al. 2003). However, neither the replicon system nor the HCVpp system could generate infectious virus; thus, no robust model existed to study later steps of the viral life cycle, such as assembly. For this reason, our understanding of HCV assembly has always lagged behind our understanding of HCV entry and genome replication. This gap narrowed continuously following the establishment of the first cell culture system producing infectious virus (HCVcc) in 2005 (Lindenbach, Evans et al. 2005; Wakita, Pietschmann et al. 2005). Virus particles collected from cells transfected with a viral genome based on the Japanese fulminant hepatitis type 1 (JFH-1) clone are infectious in naive Huh-7 cells (Wakita, Pietschmann et al. 2005). Furthermore, these HCVcc particles productively infect chimpanzee livers in vivo (Wakita, Pietschmann et al. 2005). In addition to the JFH-1 clone, independent groups have reported the development of chimeric genomes with capacities for high viral titers (Lindenbach, Evans et al. 2005) (Pietschmann, Kaul et al. 2006). The most efficient of these proved to be a chimeric virus containing a region of the genotype 2a isolate J6 (from core protein to the first transmembrane segment of NS2) fused to a JFH-1 genome (Pietschmann, Kaul et al. 2006). Since then, numerous studies have used these cell culture systems to characterize viral assembly from genetic and biochemical angles, adding to a growing body of knowledge regarding HCV assembly events.
Cultured Huh-7 cells have an inefficient VLDL assembly pathway

Although the HCVcc system provides a powerful tool to study HCV infection within controlled \textit{in vitro} settings, it is important to be aware of the differences between hepatocytes grown \textit{in vitro} and those found \textit{in vivo}. Whereas hepatoma cell lines like Huh-7, Huh-7 Lunet or Huh-7.5.1 are rapidly dividing and poorly differentiated, hepatocytes are normally non-dividing and fully differentiated. In addition, Huh-7 and Huh-7-derived cells cannot respond to intracellular double-stranded RNA (Zhong, Gastaminza et al. 2005). Hepatocytes, however, are fully responsive to double-stranded RNA (Lanford, Guerra et al. 2003). Another important caveat in conducting experiments with Huh-7 derived cells is a possible defect in cellular pathways such as the assembly pathway of very-low-density lipoprotein (VLDL). Expression of only glycoproteins E1 and E2 in Caco-2 or HepG2 cells, two cell lines used to model VLDL synthesis and secretion, leads to production of particles containing E1-E2 / apolipoprotein B (apoB) complexes (Icard, Diaz et al. 2009). In contrast, experiments in Huh-7 cells generated only low-density, predominantly apoB-negative particles, suggesting a defect in authentic VLDL production (reviewed in) (Bartenschlager, Penin et al. 2011). This VLDL pathway has been shown to have close ties to the production of infectious particles, potentially leading to the formation of lipoviral particles (LVPs). Apolipoproteins could play important roles in the HCV budding process, in particle maturation in the ER or post-ER compartments (Popescu, Rouille et al. 2011). HCV may also associate with VLDL in the ER and exit cells by tagging onto VLDL in the secretory pathway. In conclusion, these later steps of the virus life cycle may be more accurately represented in model systems that are more physiologically relevant than the current HCVcc system. Nonetheless, the HCVcc system still remains suitable for elucidating earlier steps in the HCV assembly process.
Multiple viral proteins appear to play a role in the HCV assembly process occurring close to lipid droplets (LDs)

The complexity of intracellular events during the life cycle described above is staggering. Presumably these events occur at distinct locations within distinct subcellular compartments through distinct viral proteins. Along with the core protein, the envelope glycoproteins E1 and E2 make up the proteinaceous components of the virus particle. The non-structural proteins NS3-NS5B form the viral replication complex. The non-structural proteins p7 and NS2 play important roles in the HCV assembly process. In fact, numerous studies have uncovered the involvement of virtually every HCV protein, both structural and non-structural, in the HCV assembly process (for a review, see) (Popescu, Rouille et al. 2011). The importance of non-structural proteins appears to be not only peculiar to HCV, but also related flaviviruses. Another important feature of the HCV life cycle is the role played by intracellular storage organelles, called lipid droplets (LD), in infectious virus production and virus assembly (Boulant, Targett-Adams et al. 2007; Miyanari, Atsuzawa et al. 2007; Shavinskaya, Boulant et al. 2007; Boulant, Douglas et al. 2008). Non-structural HCV proteins and HCV RNA are recruited to LDs containing core and E2 proteins (Miyanari, Atsuzawa et al. 2007). Nonetheless, the precise purpose for the localization of viral proteins to this LD microenvironment is unclear and remains an active area of investigation.

The potential of advanced microscopy techniques for imaging of HCV assembly

To better understand the HCV assembly, researchers have tried to visualize the process. Microscopy has always been an important tool in the study of cellular structures and interactions between pathogens and host cells. However, only few studies have applied electron microscopy to the HCV infectious cell culture system and a comprehensive ultrastructural characterization of the HCV assembly process is still lacking. Instead, most of the current models of HCV assembly are based on confocal microscopy of fixed cells. The resolution limit of confocal microscopy
currently stands at about 250-300nm. Higher-resolution approaches such as stimulated emission depletion (STED) microscopy, stochastic optical reconstruction microscopy (STORM) or photoactivated localization microscopy (PALM) have pushed this resolution limit to the tens-of-nanometer range (Klar, Jakobs et al. 2000; Betzig, Patterson et al. 2006; Hess, Girirajan et al. 2006; Rust, Bates et al. 2006). These superresolution microscopy approaches can improve image resolution by up to ten times compared to traditional optical microscopy (Popescu, Rouille et al. 2011). An increase in the precision of image acquisition to the nanometer scale, using electron microscopy, or tens of nanometer scale, using superresolution light microscopy, has the potential to reduce some of the ambiguities in current models HCV assembly, since a more accurate picture of the localization of viral proteins can be obtained. Thus, these higher-resolution imaging technologies allow researchers to break new ground in the visualization of complex processes such as virus-host interactions.

In addition, application of electron tomography to study different stages of the virus life cycle is a recent development that has delivered novel insights into the architecture of several virus-host interactions (Kopek, Perkins et al. 2007; Welsch, Miller et al. 2009; Felts, Narayan et al. 2010; Peng, Ryazantsev et al. 2010; Romero-Brey, Merz et al. 2012; Miorin, Romero-Brey et al. 2013). The draw of electron tomography is that it is comparable to medical tomographic techniques like magnetic resonance imaging (MRI) or computed axial tomography (CAT) in the sense that it provides three-dimensional (3-D) spatial information. Moreover, it does so at a cellular scale and with nanometer resolution. Thus, tomography of infected cells has the potential to provide structural insight at resolutions in-between those obtained with X-ray crystallography and conventional light microscopy, thereby surmounting a vital gap in the biomedical imaging spectrum. Previous efforts in our laboratory have visualized the life cycle of the murine gammaherpesvirus-68 (MHV-68) in cultured cells by dual-axis electron tomography (Peng, Ryazantsev et al. 2010). Three-dimensional reconstructions of viral attachment, entry,
capsid maturation, formation of nuclear inclusion bodies, tegumentation, and exit provided snapshots of MHV-68 morphogenesis inside host cells at high resolution. Transient processes such as incoming herpesvirus capsids injecting viral DNA through the nuclear pore complex could be imaged with unprecedented detail. This study demonstrated the feasibility of applying electron tomography to the study of viral infection in cultured cells.

With the development of new fluorescent probes and microscopes, the analysis of dynamic processes within cells has been greatly facilitated. Fluorescence light microscopy allows the researcher to follow his chosen target in living cells over time, revealing functional information through the use of fluorescent markers. The discovery of the green fluorescent protein (GFP) and the various fluorescent derivatives (for a review on fluorescent proteins, see) (Shaner, Steinbach et al. 2005) greatly facilitated the study of viruses in infected cells. Viral proteins tagged with fluorescent reporters become directly visible in living cells without the need for fixation. In one approach, the coding sequence for GFP was inserted into the coding region of both a genotype 1a and 1b HCV replicon (Jones, Gretton et al. 2007; Wolk, Buchele et al. 2008). These replicons harboring GFP inserted into the NS5A protein allowed for the study of HCV replication complexes in live cells and showed that NS5A localized in either large, static clusters likely representing membranous webs or small, motile structures, likely representing small replication complexes. Moreover, analogous studies of GFP-tagged NS5A in the HCVcc system visualized NS5A fluorescence around LDs and in perinuclear structures (Nevo-Yassaf, Yaffe et al. 2012). These structures were incorporated into a highly immobile platform superimposed over the ER membrane (Nevo-Yassaf, Yaffe et al. 2012). Nonetheless, how this subcellular localization of NS5A relates to the location of the HCV assembly site still remains an open question.
References


CHAPTER 3

Profiling of 15-nucleotide insertion mutations by next-generation sequencing reveals a comprehensive functional map of the hepatitis C virus genome
ABSTRACT

Pairing high-throughput sequencing technologies with high-throughput functional studies enables genome-wide studies of pathogenic organisms. A single experiment can uncover a comprehensive landscape of these pathogens’ essential and non-essential genes. In this chapter, we describe the application of next-generation sequencing (NGS) to a complete viral genome. We generated a transposon insertion library that contains mutants with 15-nucleotide (15-nt) insertions randomly distributed across the entire hepatitis C virus (HCV) genome. Using a genetic footprinting approach, we passaged the library in hepatic cells, recovered passaged virus pools, and used Illumina sequencing to simultaneously assay the abundance of mutant viruses in each pool. We provide evidence that the library’s genetic footprints reflected the biology of HCV. Furthermore, we show how these genetic footprints can serve as a resource for gaining new insight on functional HCV regions. With this approach we identified a region in the non-structural protein NS4B that plays a role in post-RNA-replication steps of the viral life cycle. This resource, which we are making freely available to the public, will allow researchers to easily evaluate the effect of 15-nt insertions at 99.3% of HCV codons. Once cloned as DNA, a viral genome can serve as the starting point for this transposon insertion sequencing approach, opening the door to systematic functional analyses of viruses for which infectious clones are available. Such analyses are now precise enough to determine both the exact amino acid sequence of insertion mutations as well as the resulting effect on viral fitness. Thus, this approach, which we refer to as high-resolution genetics, creates comprehensive lists detailing how insertion mutations affect various infection processes.
INTRODUCTION

Next-generation sequencing (NGS) techniques have promoted strong growth in available DNA sequences, boosting deposition of new sequences to public databases. In contrast, increases in our knowledge of viral gene function has been trailing this type of growth. Advances primarily rely on generating and testing individual mutant viruses. For genome-scale studies, this approach would be very time-consuming and labor-intensive. Thus, virologists could benefit from rapid, high-throughput methods capable of uncovering, in a single experiment, the relationship between sequences and encoded functions.

The microbiology field has already witnessed how a multitude of studies broke new ground by employing a genetic footprinting approach that combined next-generation sequencing with traditional transposon mutagenesis (van Opijnen, Bodi et al. 2009). In brief, these strategies recover representative transposon insertion libraries grown in defined selection conditions. Next-generation sequencing of insertion site junctions then determines relative frequencies of each insertion mutant at various points during the selection. In this way, one experiment can, in a quantitative manner, report the contribution of thousands of genome positions on the outcome of any selection (Gawronski, Wong et al. 2009; Goodman, McNulty et al. 2009; Langridge, Phan et al. 2009; van Opijnen, Bodi et al. 2009).

The addition of transposons to the molecular biology tool kit enabled large-scale mutagenesis studies in a variety of systems, with pioneering work conducted in bacteria and yeast (Smith, Botstein et al. 1995; Smith, Chou et al. 1996; Akerley, Rubin et al. 1998; Haapa, Taira et al. 1999). Early work on viruses include transposon-mediated 15-nucleotide (15-nt) insertion mutagenesis of Potato virus A (Kekarainen, Savilahti et al. 2002), the 5′ end of human immunodeficiency virus (Laurent, Olsen et al. 2000), and a murine gamma-herpesvirus (Moorman, Lin et al. 2004; Song, Hwang et al. 2005). A major difference of these types of studies compared to insertional mutagenesis studies in bacteria lies in their scale: relatively
small insertions (only 15-bp compared to entire transposons for bacterial studies) are introduced at the majority of possible sites across the entire region-of-interest. Therefore, this approach is particularly well-suited for near-saturating insertional mutagenesis and subsequent analysis of viruses with smaller genomes. However, the aforementioned studies carried out genetic footprinting of insertion sites either through urea-polyacrylamide-gels, analyses of colony blots, or analyses of individual mutants. We and others have increased the throughput of the footprinting analysis by using a massively parallel profiling platform based on capillary electrophoresis; this resulted in functional profiles for the entire viral genomes of hepatitis C virus (HCV), Venezuelan equine encephalitis virus (VEEV) and norovirus (Arumugaswami, Remenyi et al. 2008; Beitzel, Bakken et al. 2010; Thorne, Bailey et al. 2012). However, despite these improvements, detection accuracy and sensitivity remained limiting factors. For example, capillary electrophoresis comes with an unavoidable detection inaccuracy of one to two nucleotides, which adds uncertainty when trying to determine the exact amino-acid-identity of a profiled mutant. Moreover, the low capacity of fragment analysis by capillary electrophoresis prevents the detection of lower-abundance insertion mutants. Thus, striving to improve the quality of genetic footprinting, Beitzel et al. were the first to use massively parallel sequencing to obtain a high-resolution functional profile of a viral genome, namely VEEV (Beitzel, Bakken et al. 2010). In these experiments, the authors used Roche 454 sequencing. In contrast, transposon insertion sequencing of bacterial genomes took advantage of the increased sequencing capacity of Illumina technology. Thus, the application of Illumina sequencing to the functional profiling of viral genomes would represent a further technological advance for viral profiling platforms. Recently, this approach has been reported for the influenza virus (Heaton, Sachs et al. 2013), but no such mutagenesis and in-depth NGS-based analysis has been performed for a positive-stranded RNA virus such as HCV.
To improve the resolution of functional viral profiling, we performed Illumina sequencing of large viral mutant pools. We also re-generated these pools using transposon insertion mutagenesis of an HCV construct, resulting in a higher density of insertion than our previously published study. We selected the HCV mutant library in cultured cells, followed by genetic footprinting using Illumina sequencing. We chose HCV for our genome-wide studies because of the relevance to public health, the availability of molecular biology tools, and the rich body of existing literature, which includes experimentally determined protein structures. An estimated 130-170 million people worldwide suffer from hepatitis C virus infection (Alter 2007). Over time, liver infection wreaks havoc through chronic inflammation and liver cancer. HCV research lacked infectious model systems to study the virus life cycle until 2005, when several groups reported successful propagation of HCV in cell culture (Lindenbach, Evans et al. 2005; Wakita, Pietschmann et al. 2005). In host hepatocytes, translation of the relatively compact RNA genome (9600 base pairs) produces ten viral proteins. Ascribing functions to these ten viral proteins remains an active area of investigation.

In this report, we show that our transposon insertion library can serve as a resource that provides information on the effect of 5-amino acid (5-aa) insertions on virus viability. These insertions are spread randomly across the entire HCV genome, with a single insertion for every mutant virus, at 99.3% of HCV codons. We address whether the genetic footprints of various HCV proteins capture HCV biology. These applications include identifying amino acids located in a prototypical replication-complex protein that appear to play a role in later steps of virus replication, as well as using the resource to guide us in the construction of infectious epitope-tagged viruses. By taking full advantage of advances in sequencing technology, this systems-level analysis of the HCV genome provides a comprehensive resource for the entire research community. Thus, we envision this resource as a tool to aid future functional studies of the HCV life cycle.
RESULTS

To create a comprehensive resource that would allow any researcher to retrieve information regarding the effect of small exogenous insertions on HCV fitness, we employed a strategy that combined transposon insertion sequencing with genetic footprinting (Figure 3-1). First, using bacteriophage Mu transposase, we generated a transposon library for HCV (Figure 3-1A). The final library consisted of a pool of over one million mutants. Each of these mutants harbored one 15-nucleotide (15-nt) insertion at one random location of the HCV genome. The 15-nt insertions themselves contained a unique NotI restriction site as well as a 5-nt duplication of the sequence adjacent to the insertion site. Note that this method does not delete any nucleotides, nor does it introduce stop codons. In addition, analysis of NotI digestion products confirmed the absence of individual plasmids containing two or more insertion sites (data not shown).

Second, to carry out genetic footprinting analysis, we selected the mutagenized HCV library in cell culture (Figure 3-1B). Following electroporation of the in vitro-transcribed RNA library into Huh-7.5.1 cells, we recovered total RNA pools at two distinct points in time. At 96 hours post-transfection (96 hpt), we collected the total RNA from these electroporated cells, which represented pool 1 (P1). We also collected cell culture media, which we used to infect naïve cells. After allowing this infection to progress for 72 hours, we collected the total RNA from these infected cells, which represented pool 2 (P2). Along with the original mutagenized RNA library, which represented the unselected input pool (P0), we processed the three pools for genetic footprinting analysis.

To perform the genetic footprinting analysis of the pool 0, pool 1, and pool 2 libraries, we used a NGS-based digital counting approach. Figure 3-1C and D outline the key steps in the preparation of the sequencing samples. After we generated thirteen overlapping amplicons spanning the entire HCV genome by RT-PCR (F1 through F13), we randomly fragmented these amplicons by sonication (Figure 3-1C). Next, we ligated sequencing adaptors (Adaptor A) to the
resulting fragments (Figure 3-1D). To enrich for fragments harboring 15-nt insertions, we exposed 15-nt ends using NotI digestion, followed by ligation to a biotin-modified adaptor (Adaptor B) through NotI overhangs. Purification using streptavidin-coated magnetic beads then removed any fragments lacking the biotin tag. Thus, only fragments with insertion sites could continue on through the subsequent PCR amplification and NGS preparation steps.

We ran our sequencing samples on an Illumina platform, which generated 6.4 million reads for pool 0, 3.0 million reads for pool 1 and 4.5 million reads for pool 2 (Table 3-1). The initial library from pool 0 contained 8398 unique insertion positions. Thus, our experiment covered 87% of the entire 9600 base pairs of genomic space. This resulted in a change of 99.3% of HCV codons. At the amino-acid level, the pool 0 library contained at least one unique insertion mutant at 98% of HCV amino-acid positions. The strong negative pressure during passaging of the HCV library already became apparent in pool 1, where the sequencing reads of 5290 mutants equaled zero (Table 3-1). This negative pressure was also apparent in pool 2, where 8260 insertion mutants displayed zero values (Table 3-1). In conclusion, the combination transposon insertion sequencing and genetic footprinting yielded a dataset large enough to assay the majority of nucleotide positions in the HCV genome and nearly every encoded amino acid position across multiple passaging pools.

**Transposon insertion sequencing precisely maps complex HCV mutant libraries across different selection pools**

To visualize the genetic footprints of this comprehensive insertion library, we generated an insertion map using the Illumina sequencing data (Figure 3-2A). For each unique insertion position, we determined the number of sequence reads in pool 0, pool 1, and pool 2. In addition to the global footprint of the entire HCV genome, we could also obtain close-up views of the effect of 15-nt insertions on individual genes, such as the p7 and NS2 genes, shown as
representative examples (Figure 3-2B). The precision of determining the exact position of each insertion mutant allowed us to line up the resulting graphs with the known domain organization of the p7 and NS2 proteins. Note that for HCV, the core protein and envelope glycoproteins E1 and E2 make up the structural components of the virus. The non-structural proteins NS3 to NS5B, often referred to as “replicase” proteins, form components of the viral replication complex. The p7 and NS2 proteins most likely play non-structural roles. However, these proteins are dispensable for genome replication. At the same time, more and more studies have been implicating both p7 and NS2 in virion formation. Overall, the functional map of the entire HCV genome showed a clear reduction of detected insertion mutants in pool 1 (compared to pool 0) and a further reduction in pool 2. The reduction of insertion mutants in pool 1 was most pronounced for insertions in the NS3, NS4A, NS4B, and NS5B genes (Figure 3-2A, purple, cyan, pink and black segments). Moreover, we detected a steep decline for insertions in the 5’ untranslated regions (5’UTR, red), which contains an internal ribosomal entry site (IRES). In contrast, insertions in the core, E1, E2 and p7 genes (gray, blue, green, light blue) persisted in pool 1 and underwent a strong reduction in pool 2 instead. Insertions in the NS2 and NS5A genes (brown, olive) displayed a mixed selection pattern, with mutants in certain domains of the encoded protein persisting in pool 1 and even pool 2, while others dropped out of the population as early as pool 1. Finally, a large proportion of insertions in the 3’ untranslated region (3’UTR, orange) persisted both in pool 1 and pool 2. In summary, the functional map of the HCV genome displayed very distinct patterns of negative and positive selection. Moreover, the precision of insertion site sequencing allowed us to associate these patterns with individual HCV genes and underlying protein domains.
Selection profiles within individual HCV genes vary according to gene’s encoded function

The results above indicated that insertions in particular genes displayed a selection profile specific for that gene. To gain further insight into the behavior of insertion mutants relative to the mutated gene, we examined selection profiles on a per-gene basis. To examine the frequency of 15-nt insertion mutants across pools 0, 1, and 2, we based our analysis on use of one million sequencing reads. This was necessary to normalize the data from different library preparations. In particular, we examined the ratios between the normalized sequencing reads from pool 1 and the reads from input pool 0 (P1/P0). This allowed us to assign a value to the amount of selection taking place during library passaging. Similarly, the ratio of the pool 2 value and pool 0 value (P2/P0) gave us an estimate of selection during the second passage of the library. Whereas insertions in NS5A (not including the RNA-binding Domain I) maintained the highest average P1/P0 and P2/P0 ratios out of all HCV genes (7.9 and 9.2), insertions in the NS2 autoprotease domain, NS3, NS4A, NS4B and NS5A Domain I and NS5B (not including the C-terminal membrane anchor) were at near-zero levels (Table 3-2). The majority of insertions in Core, E1, E2, p7 and the NS2 transmembrane domain displayed nonzero P1/P0 values (range of 0.95-1.67). While the average P2/P0 ratios for insertions in core, E1 and E2 genes was close to zero, they were higher for insertions in p7 and the NS2 transmembrane domain (1.10 and 1.19), consistent with the persistence of numerous p7- and NS2-insertion mutants in pool 2 (also seen in Figure 3-2B). Moreover, we further confirmed these trends by counting the number of insertion mutants dropping to zero in pool 1 or pool 2 (Table 3-2, Columns 7 and 8). This gave us a further estimate of the selection taking place during passaging of the library. In summary, we saw clear differences between the ten viral genes with regards to the associated P1/P0 and P2/P0 ratios.
These gene-by-gene differences in P1/P0 and P2/P0 ratios prompted us to get a clearer picture of the variation among insertion mutant’s ratios across HCV genes. Histograms of these ratios for every profiled insertion mutant showed that about 22% of all insertion mutants displayed P1/P0 ratios between 0.11 and 1, while 47% of mutants had values of 0 (Figure 3-3A). With regards to P2/P0 ratios, 83% of mutants had values of 0, underscoring the additional negative selection in pool 2. However, when we sorted the insertion mutants gene by gene, we identified two unique distribution patterns. The first pattern was characteristic for insertions in the structural proteins (core, E1, E2). P1/P0 ratios for insertions in these genes fell into a range between 0.5 and 2 (Figure 3-3B, upper histogram). Corresponding P2/P0 ratios, however, concentrated at the lower end of the range, closer to zero (Figure 3-3C, lower histogram). This pattern was consistent with the expected pattern for structural proteins, which are dispensable for genome replication in pool 1 but essential for virus propagation in pool 2.

We observed a second pattern for insertions in the NS3, NS4A, NS4B, NS5B genes. Both P1/P0 and P2/P0 for these genes displayed a strong tilt towards 0 and 0.01 values, making up 76-92% and 94-99% of the population, respectively (Figure 3-3C, upper and lower histograms). Note that these genes encode replicase proteins, which are required for the early steps of the virus life cycle such as genome replication.

The p7 protein displayed histogram patterns comparable to the structural proteins (Figure 3-3D), which was consistent with a protein that is essential for later steps in the viral life cycle but not essential for the early steps of genome replication. On the other hand, the NS2 protein had a higher proportion of zero values compared to core, E1, E2 and p7 (34% versus 4-5%, Figure 3-3E). This pattern was consistent with NS2’s additional role in protein processing, which is essential for early replication in P1. Finally, insertions in the NS5A gene displayed a unique pattern as well. Here, 41% of insertion mutants had P1/P0 with values of 2.1 or more (Figure 3-3E, upper histogram). Moreover, for P2/P0, this subpopulation still made up 37% of the total
population (Figure 3-3C, lower histogram), suggesting toleration of a large number of insertions. The NS5A gene encodes a multifunctional protein that plays a role in both replication and post-replication steps of the virus life cycle. Lastly, insertions in the non-translated regions of the genome (5’UTR and 3’UTR) displayed a distribution pattern that matched the structural protein pattern and NS5A pattern respectively (data not shown). In summary, we could group the histograms of the distribution of average P1/P0 and P2/P0 values according to similarities in the profile. Most importantly, the separation into these groups largely overlapped with separation based on the genes’ encoded function, namely as structural components (Core, E1, E2) and replicase components (NS3, NS4A, NS4B, NS5B).

**HCV profile reveals bias for particular amino-acid-insertions within special functional areas of the genome**

After we had examined the gene-by-gene differences in the pool 1 and pool 2 selection profiles, we wanted to assess if amino-acid differences of encoded insertions could also have an effect on these selection profiles. The actual encoded amino-acid insertion represented one of three possible frames depending on an insertion’s location relative to the overall HCV reading frame (Arumugaswami, Remenyi et al. 2008). This resulted in 5-aa insertions with the following constant motifs: alanine-alanine-alanine (AAA), cysteine-glycine-arginine (CGR) and arginine-proline-histidine/glutamine (RPH/Q). When we took the entire dataset’s P1/P0 and P2/P0 ratios, separated these ratios by each frame, and examined histograms of the ratio distribution, we could not detect any differences between the three frames (Figure 3-5A). Conversely, when plotting the P1/P0 and P2/P0 ratios for every encoded reading frame in relation to the genome position, we noted two regions with biases for particular reading frames (Figure 3-4, black arrows). Upon closer inspection, these areas of difference overlapped with two known functional stretches encoded by these HCV proteins, namely a signal peptide stretch at the C-terminus of the core protein as well as a hypervariable region in the E2 protein. Histograms of P1/P0 ratios
for the signal peptide area of HCV core showed that about 94% of insertions encoding the AAA frame resulted in high P1/P0 ratios of 2.1 or more and non-zero P2/P0 ratios (Figure 3-5B). In contrast, 58% of P1/P0 ratios for the RPH/Q frame displayed values between 0.11 and 1, while 100% of P2/P0 ratios were between 0 and 0.1. This suggested a higher tolerance of the AAA frame compared to the RPH/Q frame. The second region matched a hypervariable region in the E2 protein. For this stretch, 80% of insertions that contained the cysteine-glycine-arginine sequence had P2/P0 ratios of 0. This lower preference of the CGR sequence compared to the AAA and RPH/Q frames was already seen in P1, with none of the CGR insertions displaying P1/P0 ratios higher than 5. In contrast, 77% and 62% of AAA and RPH/Q insertions had P1/P0 ratios higher 5. In conclusion, this analysis provided evidence that our genetic footprinting approach was sensitive enough to pick up biases for the types of amino acids inserted. In addition, only particular regions of the HCV genome exhibited these biases, with the most prominent examples being the C-terminal signal peptide of the core protein and the hypervariable region of E2.

**Functional annotation of the C-terminal region of the core protein supports biological relevance of protein structure**

To further confirm that we could match our screen data with already-known functional regions, we closely examined our genetic footprints in areas where experimental structural information was available. The signal peptide region of the core protein represented our first region of interest due to the previously described bias for alanine residues. We used the log of the P2/P1 ratios as a measurement of fitness and displayed the log values for each frame as a heat map (Figure 3-6). This also allowed us to determine the effect of the three types of amino-acid insertions at each amino acid residue by color-coding the previously-published NMR structure (Oehler, Filipe et al. 2012) according to our heat map. The annotated tentative model of the peptide’s orientation within the ER membrane reaffirmed that P2/P0 ratios for insertions in the
membrane hydrophobic core, namely the helix extending from amino acids 175-186, were highest when the insertion coded for alanine residues (Figure 3-6, cyan/blue colors for AAA frame).

We noticed that the boundary between tolerated and non-tolerated amino acids at amino acid position 176 (isoleucine) and 177 (phenylalanine) bracketed a signal peptide peptidase cleavage site. Processing of the core protein by signal peptide peptidase (SPP) represents an important step in this protein’s maturation. Our data suggested that 5-aa insertions positioned upstream of this cleavage site interfere with the function of the core protein, most likely by affecting core processing. To test this hypothesis we constructed individual mutant viruses containing amino acid substitutions in the C-terminal region of the core protein. By using substitution mutations (Figure 3-7A) instead of insertion mutations, we wanted to further distinguish between essential and non-essential residues with regard to infectious virus production. As controls, we included the ASC180/3/4VLV (ASC) and the IF176/7AA (IF) mutations. The ASC mutation was previously shown to abolish cleavage of the mutant core protein by SPP in the HCV strains Glasgow and J1, but not in the JFH-1 strain (Okamoto, Mori et al. 2008; Targett-Adams, Hope et al. 2008). The second control mutation, IF176/7AA, inhibited core protein processing in genotype 1a and 1b strains and abolished infectious virus production when introduced into a JFH-1-based virus (Okamoto, Moriishi et al. 2004; Okamoto, Mori et al. 2008). We transfected RNA transcripts of these controls, along with the three alanine mutants NFT, GNLP and PVS, into Huh-7.5.1 cells. As expected, none of the mutations had an effect on intracellular viral RNA levels, as measured in a quantitative RT-PCR assay at multiple time points following transfection (data not shown). In contrast, measurements of extracellular infectivity at 48 and 72 hours post-transfection indicated that only the ASC and PVS mutants were capable of infectious virus production (Figure 3-7B). Immunofluorescence assays showed that transfected cells produced readily detectable levels of the NS5A protein, suggesting that
viral protein translation was functional in all mutants (data not shown). In contrast, detection of core protein levels varied in cells transfected with these mutants, with the IF, GNLP and NYT mutants exhibiting lower staining for the core protein (data not shown). To better compare core protein levels across mutants and to distinguish between processed and unprocessed forms of core protein, we examined intracellular levels of core protein by Western blotting. To ensure expression of core protein regardless of the ability to produce infections virus particles, we subcloned DNA fragments coding for the IRES, Core and part of E1 into a pFLAG-CMV expression plasmid. We then transfected mutant core plasmids into 293T cells and harvested total protein lysates at 20h post-transfection. Cell lysates were separated by SDS-PAGE (15%) for detection of core and beta-actin. Doublets for NFT and GNLP indicated the presence of the processed p21 and unprocessed p23 form of the core protein, whereas the PVS mutant and WT constructs mostly contained the lower-weight p21 form (Figure 3-7C). Moreover, the IF mutation appeared to drastically affect the stability of the protein, which prevented us to evaluate core protein processing status.

In conclusion, we validated the importance of particular amino acids in the correct processing of the C-terminal end of the HCV core protein. The correct processing was essential for the production of infectious virus particles. Moreover, the importance of these amino acids was already suggested by the genetic footprints of our initial screen. Therefore, this result provided another example in which the genetic footprints indeed reflected the underlying biological functions of HCV residues.

**Functional annotation of transmembrane protein structures supports their biological relevance**

Another characteristic of the C-terminal helix, which forms part of the core protein’s signal peptide, was its membrane association. This prompted us to obtain more detailed views of the
genetic footprints in regions encoding additional membrane-associated segments.

Consequently, we matched a previously-published NMR structure of the p7 protein (OuYang, Xie et al. 2013) to our genetic footprints (Figure 3-8A). We observed a concentration of tolerated insertion mutants within an unstructured region connecting the N-terminal helix and first transmembrane helix TM1. In general, the p7 protein did not tolerate insertions located within the transmembrane helices themselves, except for a portion of transmembrane helix 2 (TM2). Similar to the toleration pattern in the C-terminal helix of the core protein, TM2 of p7 tolerated alanine insertions in particular (stretch of four blue boxes, LTGL). Overall, these results suggested that we could indeed zoom into our genetic footprint with single-residue resolution. Moreover, the profile itself was consistent with the results from independent structural studies.

To further extend our functional analysis to areas of the HCV genome coding for membrane-spanning portions, we wanted to examine the footprint of an additional HCV transmembrane protein, NS2. When overlaying known structural information on the membrane-bound portion of the NS2 protein (Jirasko, Montserret et al. 2008; Jirasko, Montserret et al. 2010) with our functional profile, we found that the most tolerated region was located proximal to a short helix in TMS1 (Figure 3-8B). A second region of toleration lined up with a short loop connecting TMS2 and TMS3. We also detected a limited number of mutants with insertions located within the helices of TMS1, TMS2 and TMS3. Their rather low P2/P1 ratios (purple colors in heat map), however, suggested a reduced viral fitness. In conclusion, our genetic footprints of the C-terminal domain of core, the entire p7 protein, and the N-terminal half of the NS2 protein showed remarkable consistency with solved transmembrane protein structures. In turn, we could now provide functional evidence in the context of the entire infectious life cycle. This supported the biological relevance of these solution structures, which, with the exception of p7, were based on fragments of the studied protein.
Mapping of tolerated insertion mutants to protein structures of p7 viroporin and the viral polymerase NS5B reveals structural context for insertion toleration

After we observed that areas of toleration in the p7 protein were mostly within turns, loops and outside of transmembrane helices, we wanted to examine where these areas of toleration located on the structure of the entire p7 viroporin. While the p7 protein is rather small (63 amino acids), six p7 monomers can self-assemble to form larger ion channel complexes (OuYang, Xie et al. 2013). After annotating surface views of the viroporin structure with the heat map of P2/P1 ratios, we could visualize an area tolerating all three insertion motifs in the middle part of the ion channel (Figure 3-9, side view, black arrows). While most of the tolerated insertions were in a region forming an unstructured loop between the N-terminal helix and TM1 (amino acids 15-18), we also saw toleration at the C-terminal end of the N-terminal helix, especially for mutants encoding the RPH/Q insertion motif. Transparent views from the top and bottom of the channel revealed that the unstructured connecting loop localized to the channel periphery, whereas the tolerated helix residues faced the inside of the channel. Altogether, our functional information from the genetic footprints allowed us to define the boundaries of an area of toleration for 5-aa insertions within the structural landscape of the oligomerized p7 ion channel.

Next, we wanted to further examine the characteristics of tolerated insertion positions within the context of a globular protein, the NS5B polymerase. Out of a total of over a thousand insertion mutants located in the globular part of NS5B, only six mutants persisted at appreciable levels in P1 and P2 (Table 3-3). Mapping tolerated sites to the published crystal structure of HCV polymerase NS5B (Schmitt, Scrima et al. 2011) placed these insertions on solvent-exposed surfaces of the protein (Figure 3-10, upper panel). Also, mutants with insertions within the loop regions of the crystal structure showed a tolerated phenotype (Leucine 60, Alanine 150, and Arginine 379). Similarly, Leucine 443 tolerated a cysteine-glycine-arginine-asparagine-leucine insertion, even though this residue formed the beginning of a beta-sheet (Figure 3-10,
lower panel). Lastly, Alanine 305 and Arginine 523 both formed part of an alpha helix but tolerated a cysteine-glycine-arginine-lysine-alanine and threonine-valine-cysteine-glycine-arginine insertion (Figure 3-10, lower panel). A closer look at the insertion at Arginine 523 revealed that the insertion duplicated the existing valine-cysteine-glycine-arginine residues while adding only one exogenous residue (threonine). Overall, functional annotation of the NS5B structure showed that insertions in solvent-exposed loops, as well as a few peripheral portions of helices and beta-sheets produced tolerated mutants. However, the exact sequence of the insertion continued to matter, as seen for insertions after Arginine 523.

**Identification of nonessential regions aids in the design of viable epitope-tagged viruses**

Results from our functional annotation of existing protein structures already hinted at the usefulness of the high-resolution genetic footprints for identifying regions of insertion toleration, such as surface-exposed loops or unstructured linker regions. Thus, we wanted to use these genetic footprints to guide us in the insertion of other exogenous motifs into the HCV genome. Genetic insertion of small motifs (6-20 amino acid residues) can provide specific “tags” for viral proteins that are useful for protein detection. Using specific antibodies that bind to these epitope tags, researchers can purify viral proteins and associated complexes from infected cells; also, antibodies can aid in the visualization of viral protein subcellular localization. In order to identify candidate regions for epitope-tag insertion and produce a map of “hot spots” tolerating insertions, we filtered the functional profile with stringent criteria. Hot spot regions had to tolerate 5-aa insertions regardless of the insertion reading frame. Therefore, the resulting plot only contained data points that fell within a cluster of at least three consecutive insertion mutants showing P2/P1 ratios >0.01 (Figure 3-11A). Moreover, we disregarded any insertion mutant that dropped to P1/P0 ratios of 0.1 or lower. This analysis revealed hot spots in Core, E2, p7, NS2 and NS5A (Figure 3-11B-E).
After taking into account the known structural organization of these proteins, we picked specific locations for the insertion of epitope tags (Figure 3-11B-E, red bars). Because the length and amino-acid sequence of our inserted epitopes differed from the screen’s 5-aa insertions, we inserted a panel of tags with varying sizes, sequences, and utilities. This panel included the FLAG, HA, myc, tetracysteine and OLLAS tags (Table 3-4). This variety served to increase the likelihood of producing viable epitope-tagged mutants. After constructing the individual epitope-tagged mutants, we measured their fitness by infectivity assays. The majority of epitope-tagged mutants remained viable (Figure 3-12A). Peak infectivity for the wild-type virus can reach $10^5$ infectious units per ml, so that even viruses with a 100-fold reduction in infectivity produce relatively high levels of 1000 infectious particles per ml. Moreover, we confirmed that the tags remained functional by staining infected cells with antibodies against the respective tag, followed by immunofluorescence microscopy (Figure 3-12B). The subcellular localization of tagged proteins (Core for 562-HA, E2 for 1489 FLAG, and NS2 for 2792 OLLAS) was consistent with previous studies. Moreover, anti-tag antibodies only detected protein in cells infected with tagged viral genomes, such as myc-tagged NS2 (Figure 3-12C, 2792-myc). For viral genomes with tagged p7, three constructs with different epitope tags had the highest infectivity (2590-HA2x, 2590-OLLAS and 2644-HA). However, staining of 2644-HA by immunofluorescence microscopy produced a much weaker signal compared to staining of 2590-HA2x (Figure 3-12D). On the other hand, 2590-OLLAS had an even higher fitness compared to 2590-HA2x while allowing robust detection of the OLLAS tag, making this construct a useful tool for the study of p7 protein in infected cells. Lastly, epitope-tagged constructs provided added flexibility in multi-color fluorescence microscopy experiments due to the wide range of antibodies available against epitope-tagged proteins. Multi-labeling experiments rely on the binding of primary antibodies raised in different species. Subsequently, species-specific secondary antibodies conjugated to different fluorochromes allow for specific detection of multiple proteins in the same sample. Using the 7177-FLAG virus, which produced FLAG-tagged NS5A protein, we
could detect three viral proteins simultaneously using primary antibodies against FLAG, HCV core and HCV E2 (Figure 3-13). In addition, staining of infected cells with a fluorescent lipophilic dye, which incorporates into LDs, provided an additional cellular marker. This 4-color fluorescent labeling protocol revealed the presence of triple-positive areas containing NS5A-FLAG, core, and E2 in close proximity to LDs. In conclusion, this work confirmed that we could tap our HCV genetic footprint resource to guide the successful placement of epitope-tags for the production of replication-competent reporter viruses, which have a wide range of applications, such as the visualization of viral protein subcellular localization by fluorescence microscopy.

**Further filtering of transposon mutant database uncovers new functional regions encoded by HCV genes**

The above analyses confirmed that the dataset’s genetic footprints agreed with known functions of HCV genes, as well as known protein structures. This provided evidence that our profiling approach produced a valid readout of HCV biology. Next, we wanted to test whether this resource could produce biological information beyond known functions of HCV genes. Therefore, we added distinct screening criteria to filter out specific insertion mutants. Our screening criteria limited our search to insertions in non-structural proteins with P1/P0 ratios between 0.65 and 1.85. In addition, P2/P0 ratios had to be 0.01 or less. We picked the 0.65-1.85 P1/P0 range based on the observation that insertions in core, E1, and E2 showed average P1/P0 ratios of 1.6, 1.4, and 1.0. Therefore, we expected that insertion mutants with comparable P1/P0 ratios would be competent for genome replication. However, P2/P0 ratios of 0.01 or less would be consistent with a defect at a later step of the viral life cycle. Applying these screening criteria produced a candidate list of 18 insertion mutants (Table 3-5). Because of the overrepresentation of insertions in NS4B (8 out of 18 mutants), we focused subsequent validation efforts on the NS4B region. Individually cloned NS4B mutants showed a defect in spreading in cell culture, as evidenced by immunofluorescence studies examining the
percentage of HCV-positive cells following transfection of recombinant viral genomes (Figure 3-14A). Moreover, NS4B mutants had markedly reduced infectivity, as evidenced by measuring infectious particles released from transfected cells (Figure 3-14B). Western blot analysis of HCV-transfected cells revealed a slight drop in NS5A levels compared to the E1 E2 knockout virus, which served as a replication-competent, assembly-deficient control (data not shown). Note that no NS5A protein can be detected by either immunofluorescence or Western blotting following transfection of mutant viral RNA containing a mutation in NS5B polymerase that prevents RNA replication. The production of NS5A for all NS4B insertion mutant viruses suggested that these viruses were indeed RNA replication-competent and that the up to 100 fold drops in infectivity must have been due to additional defects in post-RNA-genome replication steps such as assembly, egress or entry. In summary, we identified a new region in the NS4B protein that is essential for a post-replication step. Importantly, we showed that our dataset could serve as a resource to screen for insertion mutants that reveal regions of the genome with previously unrecognized biological functions.

DISCUSSION

Matching high-throughput sequencing technologies with high-throughput functional studies has become both an opportunity and a challenge with the rapid technological advances in next-generation sequencers and explosion in available genome sequences. Approaches that successfully complement high-throughput sequencing with high-throughput functional data have not yet been broadly utilized in the study of microorganisms, in particular viruses.

Here, we report an HCV transposon insertion library that covers 98% of amino acids encoded by the HCV genome. Second, the genetic footprints of this library, captured by next-generation sequencing, reflect HCV biology remarkably. Third, the genetic footprints serve as a resource from which we can extract new information about HCV biology.
Input library contains at least one insertion mutant between nearly every amino acid encoded by the HCV genome

The goal of this study was to obtain a comprehensive functional map of every amino acid position in the HCV genome. Indeed, our results show that out of 9686 possible insertion mutants, we detected 8398 in the input pool P0. This represented a coverage of 87% of nucleotide positions, 98% of codon triplets and 98% of amino acid positions. The input pool did not contain mutants with insertions in long nucleotide repeats, such as the 3’ untranslated region. Previously, we observed similar cold spots in these poly-nucleotide stretches, suggesting an inability of the mini-Mu transposon to insert within these target sequences (Arumugaswami, Remenyi et al. 2008). This is consistent with previous studies showing that MuB, an accessory protein of phage Mu, forms stable filaments on A/T rich DNA, thereby inhibiting Mu transposition within bound A/T sequences (Ge and Harshey 2008). When we surveyed missing insertion mutants in coding regions of the genome, it became clear that most gaps consisted of only one missing insertion mutant. Because encoded amino acids are defined by codons consisting of three nucleotides, a gap of one nucleotide would still leave functional information on two other insertions at the same codon. Thus, our initial library truly contained at least one insertion mutant at 98% of amino acid residues encoded by the HCV genome. Finally, this dense coverage was made possible by the increased sensitivity of the NGS-based counting method.

Profiling method improves on previous approaches to determine functional maps of entire viral genomes

Recent studies, which mutagenized entire virus genomes using *in vitro* DNA transposition, demonstrated the promise of high-throughput functional approaches in virology (Arumugaswami, Remenyi et al. 2008; Beitzel, Bakken et al. 2010; Thorne, Bailey et al. 2012).
Their analysis of genetic footprints, however, relied on capillary or 454 sequencing. To our knowledge, the study described here represents the first to employ Illumina sequencing in the genetic footprinting of an entire positive-stranded viral genome. This platform offers deeper sampling than the aforementioned methods, which we view as a major advantage for capturing the complexity of transposon libraries. The current profile improves on our previously published profile of the HCV genome (Arumugaswami, Remenyi et al. 2008) in several ways. First, transposon insertion sequencing removed the uncertainty we experienced when determining insertion locations with capillary electrophoresis. We and others reported inherent detection inaccuracies of one to two nucleotides (Arumugaswami et al. 2008, Beitzel et al. 2010). This possible shift of the insertion relative to the HCV reading frame could call into question the exact amino-acid-identity of a profiled mutant. Second, our current input library of 8398 mutant viruses covered most of the HCV genome, compared to the 2399 mutant viruses in our previous profile. Third, the current study used an infectious cell culture system based on a chimeric J6/JFH-1 virus instead of the original JFH-1 virus. Other groups showed that J6/JFH-1 produced 100-1000 fold higher levels of virus compared to JFH-1 virus (Pietschmann, Kaul et al. 2006). This improved infectious cell culture system had important implications for our library passaging conditions. Previously, we did not perform selection by infection with a mutant virus library, but instead passaged the viral RNA in transfected cells for 21 days. The limited replication capacity of JFH-1 necessitated this prolonged passaging. In contrast, the J6/JFH-1 system allowed us to complete the entire selection in one week, limiting the chance of secondary-site-mutations to evolve during extensive passaging. Although such mutations do occur naturally due to the error-prone nature of the viral RNA polymerase, we would not expect them to confound our footprinting results due to the relatively short duration of the passaging experiment. Another aspect of the passaging experiment was the inclusion of a supernatant passaging step. This allowed us to distinguish mutants that caused an overall defect in genome replication from mutants defective in virus particle formation.
Genetic footprints reflect HCV biology

After we confirmed that the input library was indeed comprehensive, we asked whether resulting genetic footprints accurately reflected HCV biology. Indeed, our genetic footprints were consistent with data gathered with fundamentally different methods. First, we could clearly distinguish selection profiles for insertions in genes encoding the structural proteins, which make up the protein components of viral particles, from those in genes encoding non-structural proteins, which make up replication complexes. The Core, E1, E2, p7 and NS2 proteins are dispensable for genomic replication. Our observation that insertions in these proteins persisted in P1 was consistent with this non-essential role in genomic replication. Similarly, negative selection in P2 implicated these proteins in post-replication functions, such as assembly of virus particles. In contrast, the non-structural proteins NS3 to NS5B form the viral components of the replication complex. Thus, the observed reduction of pool 1 mutants for genes encoding these proteins reflected the essential role of these proteins for genome replication. Interestingly, a closer look at the NS2 profile revealed pool-1-persistence of insertions in the N-terminal half. In contrast, the majority of insertions in the C-terminal half of the protein dropped out of the population. The NS2 protein carries out multiple functions during viral infection. The globular C-terminal half contains an autoprotease activity, while the membrane-spanning N-terminal half was recently described as an essential factor for the assembly process (reviewed in Popescu, Rouille et al. 2011). Thus the different genetic footprints we observed in the N- and C-terminus of NS2 match the different functions for the N- and C-terminus during virus replication. Interestingly, the C-terminal half of NS2 contained two short stretches at the C-terminus that persisted in pool 1 but disappeared in pool 2 (Figure 3-2B, NS2 P1 profile at genome positions 3300 and 3400), indicating that these regions were non-essential for early steps of the viral life cycle, such as autoprotease cleavage, but played a role in the later steps of virus production.
Consistent with a tolerated phenotype, comparison with the previously published crystal structure of the C-terminal domain of NS2 (Lorenz, Marcotrigiano et al. 2006) revealed that these two regions either fell within loop regions or contain only short structured segments. However, at present we cannot rule out that disappearance in P2 was a consequence of an initial moderate impairment in viral fitness in P1 that was sufficient to prevent persistence of the insertion mutants in the competitive environment of the en masse passaging experiment. Future studies will be needed to confirm the dispensability of these residues for autoprotease cleavage and essentiality for later steps of the virus life cycle, such as particle assembly.

NS5A is another multifunctional protein with a dual role in both replication and assembly (Lohmann, Hoffmann et al. 2003; Appel, Zayas et al. 2008; Tellinghuisen, Foss et al. 2008). The protein contains three domains with unique roles. The well-structured Domain I participates in viral replication (Tellinguisen et al. 2005). On the other hand, Domains II and III are natively flexible and play roles in HCV assembly (Appel, Zayas et al. 2008; Masaki, Suzuki et al. 2008; Tellinghuisen, Foss et al. 2008; Hanoulle, Badillo et al. 2010; Verdegem, Badillo et al. 2011). Indeed, our study confirms this separation of function with a strong P1 selection against insertions in Domain I. In contrast, insertions in Domains II and III represented the most tolerated stretches in the entire genome in our profile, consistent with the idea that unstructured regions are more accommodating of short peptide insertions. A notable exception was the C-terminal end of Domain II (Figure 3-2A, genome positions 7189-7283), which did not tolerate insertions as evidenced by the dramatic reduction of insertion mutants in pool 1. The suggested essential role for this region in genome replication was consistent with a recent independent mutagenesis study of the C-terminal 30 amino acids of NS5A domain II using a comparable infectious cell culture system (Ross-Thriepland, Amako et al. 2013). In summary, an agreement between the genetic footprints obtained in this study and previous knowledge on the functions of HCV proteins support the validity of this resource.
Genetic footprints complement structural studies of HCV proteins

To provide a second line of evidence supporting the validity of this resource, we surveyed the profile from a structural angle. Previous studies have already reported the structures of various HCV proteins, among them NMR structures of individual inter-membrane segments of HCV core, p7 and NS2 (Jirasko, Montserret et al. 2008; Jirasko, Montserret et al. 2010; Oehler, Filipe et al. 2012; OuYang, Xie et al. 2013). In addition, the HCV JFH-1 strain used to determine a recent crystal structure of the NS5B polymerase (Schmitt, Scrima et al. 2011) matches the genotype of our FNX-24 strain. Overall, our genetic footprints showed remarkable consistency with these structures. When mapping insertion sites to the protein structures, we saw insertion tolerance in external surface loops, unstructured linker regions and even a short unstructured region separating two closely spaced transmembrane helices. These areas may be better positioned to accommodate small additional peptides. We already saw evidence of this pattern in our previously published functional profiling of NS5B (Arumugaswami, Remenyi et al. 2008). Moreover, tolerance in unstructured loops seems to be a general theme for viral proteins, since insertional profiling of other viruses such as poliovirus and norovirus arrived at similar conclusions (Teterina, Lauber et al. 2011; Teterina, Pinto et al. 2011; Thorne, Bailey et al. 2012). Finally, our functional annotation of transmembrane proteins such as p7 and NS2 demonstrates the promise this method has in complementing the structural biology of membrane proteins. Current estimates place the fraction of membrane proteins encoded by sequenced genomes at 20-30% (Boyd, Schierle et al. 1998; Wallin and von Heijne 1998). At the same time, they reportedly represent in excess of 50% of all drug targets (Korepanova, Gao et al. 2005). Yet, difficulties in the crystallization of membrane proteins and difficulties in NMR structure determinations of larger proteins (Huang and Mohanty 2010) often prevent the determination of high-resolution structures, which is the first step to study structure-function relationships. Therefore, we think that the resource described in this paper, which examines the phenotypes of insertions in membrane proteins within the context of their native biological
environment, contains a wealth of information that can enhance the current understanding of membrane proteins’ structure-function relationships and support the biological relevance of solved protein structures.

Mapping the functional profile onto protein structures gives us a global look of how structure relates to phenotypic data. In addition, the resolution of our profile is high enough to examine the amino-acid identity of an insertion within the context of the local structure. The most obvious explanation for insertion toleration is that the additional amino acid sequences do not disturb the local protein structure. In turn, protein function remains unaltered. Finally, in this scenario, our ultimate read-out, viral fitness, would remain unchanged as well. Certainly, our data suggests that there are different degrees of toleration, which suggests that the positioning and amino-acid identity of the insertion can affect the local protein structure in varying ways. For example, we noticed that membrane-associated helices in Core, NS5B and p7 primarily tolerated insertions that contained an alanine motif. In contrast, insertions that resulted in insertion of proline or glycine were less tolerated. Whereas alanine is known to have high helix-forming propensities, the propensities of proline and glycine are poor (Pace and Scholtz 1998). Thus, the observed genetic footprints are consistent with alanine insertions being readily incorporated into the local alpha-helical structure. On the other hand, insertion of disruptive proline and glycine residues would compromise helical structures, thereby impairing viral fitness. Note that the N-terminal helix of core forms a signal peptide that is important for the ER localization and further processing of the protein through a series of cleavage events. Thus, any insertion disrupting this structure can affect optimal protein processing and in turn, viral fitness.

To test for these proposed core protein processing defects, we complemented the functional profile of the C-terminal core-E1 signal peptide by analyzing the phenotypes of individual mutants with amino acid substitutions at locations in this region. We have identified several residues that are essential for infectious virus production in the J6/JFH-1 background and
confirmed residues 188-190 as dispensable for infectious virus production. We showed that core mutants that did not produce infectious viral particles had reduced overall levels of core protein and an increased amount of unprocessed p23 core. The IF mutant showed a dramatic decrease in core protein levels, consistent with independent findings that reported that this mutation abolished viral replication and core localization in the detergent-resistant membrane (Okamoto, Mori et al. 2008). We conclude that the GNLP and NFT mutations impaired core protein processing. This was sufficient to abrogate infectious virus production. Our results are consistent with individual Alanine scanning studies of the entire core protein as well as the C-terminal region of the core protein (Murray, Jones et al. 2007; Kopp, Murray et al. 2010).

Collectively, our validation studies underscore the usefulness of the mutational analysis system for obtaining a functional profile of specific regions-of-interest and locating essential and non-essential residues within that region.

The hypervariable region of the E2 protein was another area in which a particular insertion frame appeared to more disruptive. All insertion mutants containing the Cysteine-Glycine-Arginine motif were dramatically reduced, indicating that these insertions compromised viral fitness. We hypothesize that the extra cysteine residue mediates this negative effect. The E2 protein forms nine disulfide bonds through strictly conserved cysteine residues (Krey, d’Alayer et al. 2010). During the course of viral protein translation, the E2 hypervariable region becomes exposed to reducing conditions in the ER lumen. Under these conditions the extra cysteine residue may affect the overall folding of the E2 protein, which relies on forming specific disulfide bonds between pre-defined cysteine residues. As a result, altered folding would affect protein function and ultimately compromise incorporation of E2 into viral particles.

Taken together, our genetic footprints show remarkable consistency with existing knowledge about the overall function of HCV proteins, the functions of specific protein domains, and the
corresponding protein structures. We conclude that this functional profile provides a valid resource that captures HCV biology.

**Genetic footprint resource enables extraction of new information regarding HCV biology**

To show two applications of this resource, we first created maps of non-essential areas of the HCV genome. Second, we used the genetic footprints to aid in the identification of new functional regions.

The first application of this resource focused on quickly identifying regions of the HCV genome with increased tolerance to insertions. This information can guide the exact placement of functional, exogenous sequences into a viral genome with less impact on the viability of modified viruses. By choosing more stringent criteria for regions to qualify as hot spots for toleration, we produced a map of candidate regions for small peptide insertion. Traditionally, the N- and C-termini of proteins represent the most intuitive locations to place small peptide tags, since these areas are the least likely to affect protein folding or interfere with local inter-protein interactions. However, for viruses such as HCV, placement at these sites runs the risk of altering the presentation of cleavage sites during protein processing. Nonetheless, our hot spot analysis identified the N-termini of E2, p7 and NS2 as prime candidates to insert small peptides with limited fitness cost. This is consistent with the findings from independent studies, which engineered individual epitope-tagged viruses at these locations (Jirasko, Montserret et al. 2010; Ma, Anantpadma et al. 2011; Merz, Long et al. 2011; Popescu, Callens et al. 2011; Prentoe and Bukh 2011; Stapleford and Lindenbach 2011; Vieyres, Brohm et al. 2013). While we view this as further validation of our genetic footprints, we also expanded on these previous results by showing that a wide variety of epitope tags are tolerated at these termini, in addition to the originally-published tags. Moreover, our hot spot analysis for the core and p7 proteins yielded two unpublished internal locations for epitope-tag placement producing viable virus.
We also found that insertion into internal sites was more sensitive to the size and amino-acid identity of the actual inserted tag. This represents a limitation of our approach, which only informed us of candidate positions for epitope-tag insertion, since the library mutants contain 5-aa insertions that differ from the available epitope tags in both size and amino-acid composition. However, for studies with a focus on the identification of sites tolerating epitope-tag insertion, the protocol for library construction can easily be modified following a method outlined by Moradpour et al. (Moradpour, Evans et al. 2004). Briefly, these modifications would use the original library’s inserted NotI restriction sites to clone designed DNA sequences into the HCV transposon library. DNA sequences would contain an epitope-tag sequence, such as the one encoding a FLAG tag, and a complementary NotI “sticky end”. Selection of the modified, epitope-tagged library and NGS-based counting of insertion sites would then provide a direct phenotypic read-out of epitope-tagged mutants displaying minimal effects on virus viability. This direct screening approach may be necessary to aid in the design of reporter viruses containing epitope tags in the E1, NS3, NS4A and NS5B proteins. Our genetic footprints indicated that insertions in these proteins resulted in strong negative selection in P2. Also, toleration was very dependent on the reading frame at which the transposon inserted. In addition, our efforts to use tolerated P1 insertion locations for positioning epitope tags did not yield infectious viruses (R Remenyi, unpublished). In many cases, these tagged viruses were not competent for genome replication; this suggests that these proteins are inherently less tolerant of insertions. Therefore, re-acquiring genetic footprints using libraries already containing epitope-tags could provide a better candidate list of tolerated insertion sites. In conclusion, the genetic footprinting approach outlined in this study can not only identify new functional regions of the HCV genome, but also assist the rational design of epitope-tagged viruses. The inserted tags provide a molecular handle that can be probed by various techniques, including Western blotting, immunocapture, immunoprecipitation and immunofluorescence assays. Finally, these reporter viruses will be
useful tools for the analysis of protein complexes forming during viral infection and further dissection of interactions between the tagged viral protein and other viral or cellular proteins.

The second application of this resource took advantage of the high resolution of the genetic footprints, allowing us to retrieve functional information with high precision. Not only does the resource provide the exact position of the insertion, but it also delivers semi-quantitative measures of viral fitness. Ratios of sequencing counts between the input pool 0 and selected pools 1 and 2 can then be used to search for interesting insertion mutants. By specifying our searches according to particular P1/P0 or P2/P0 ratios, as well as the regions of the HCV genome encoding non-structural proteins, we could easily populate a candidate list for areas in non-structural proteins mediating late steps in the virus life cycle. This approach can also be applied to other steps of the viral life cycle, such as entry. For our search of non-structural HCV protein regions critical in post-genome-replication steps, insertions in the NS4B gene dominated the resulting candidate list. This non-structural protein is a key player in the formation of the HCV replication complex. In addition, independent studies reported NS4B’s role in post-replication of the HCV life cycle (Elazar, Liu et al. 2004; Jones, Patel et al. 2009; Han, Aligo et al. 2011; Paul, Romero-Brey et al. 2011; Han, Manna et al. 2013). Insertions from the candidate list locate in between a predicted amphipathic alpha helix (Elazar, Liu et al. 2004; Gouttenoire, Montserret et al. 2009) and a structurally resolved alpha helix at the N-terminus of the protein. Although a clear role of this region in mediating later steps in the HCV life cycle, such as particle assembly, has not yet been reported, a recent study suggested that particular residues in the N-terminus could play minor roles in HCV assembly (Blight 2011). However, this study analyzed substitution mutations that altered native charged residue to alanine residues. Thus, varying strengths of observed mutant phenotypes could be due to differences in the type of mutations analyzed (substitution vs. insertion mutations). Future studies will need to further investigate the exact mechanism of action underlying the role of this NS4B region in virus assembly.
Altogether, we envision our footprinting resource to serve as a tool aiding the quick identification of new functional candidate regions, which prompts the initiation of more detailed individual studies addressing possible mechanisms of action.

**Perspectives for transposon insertion sequencing in HCV biology**

The genetic footprints from this study make up a database, from which the sequenced phenotypes of individual insertion mutants can be easily retrieved. After looking up insertion mutants with interesting genetic footprints, researchers can then reconstruct the corresponding insertion mutant or introduce variations of mutations at the profiled amino acid location for follow-up studies, as we have done for mutation studies of the C-terminal region of the core protein and a new functional region in NS4B. These types of studies will be critical for in-depth examinations of the impact of interesting amino-acid-level disruptions on various HCV protein functions. Thus, a forward genetic tool (a mutagenized virus population whose phenotypes can be screened en masse) can function as a launching pad for reverse genetics studies (Goodman, McNulty et al. 2009).

Although this report used genome-wide transposon insertion sequencing to assess the basic requirements for HCV viability in cell culture, the approach lends itself to analyze requirements for growth or survival under any selective condition. For example, studies comparing genetic footprints in the presence and absence of an antiviral agent such as interferon are ongoing in our laboratory. Also, hepatic cell lines could be manipulated to overexpress viral or cellular proteins. Moreover, cellular host factors can be knocked down in analogous experiments. Fitness analyses of HCV mutants in these cell lines could shed light on detailed interactions of HCV with the overexpressed/knocked-down factors. Also, even though hepatic cell lines provide a robust infectious system for proof-of-concept studies, future transposon insertion sequencing experiments should strive to perform selections of the HCV insertion library in additional model systems. Examples of these systems would include primary hepatocytes as well as hepatocyte-
like cells, which can be derived from human embryonic stem cells or induced pluripotent stem cells (Wilson and Stamataki 2012). Ultimately, applying the transposon insertion sequencing approach to animal models of HCV infection would be truly exciting. The emergence of new small animal models and improvement on existing models (Billerbeck, de Jong et al. 2013) make high-throughput functional studies of viruses using in vivo models a tangible goal in the near future.

In summary, high-throughput sequencing of transposon insertion sites in the complete genome of HCV provided us with a resource that captured HCV biology and led to identification of new functional regions. The resource’s resolution was high enough to provide information at the amino-acid level. Future characterization of compelling insertion mutants and tagged mutants will further uncover specific functions of HCV proteins. Transposon insertion sequencing can be a tool for the systems-wide study of any virus. The compact genome size of most viruses allows researcher to mutagenize and analyze entire genomes at high density. Therefore, the virology field should benefit from future resources created by high-throughput screening methods.

METHODS

Cells
The Huh-7.5.1 cell line was kindly provided by Dr. Francis Chisari from the Scripps Research Institute, La Jolla. The cells were cultured in Dulbecco’s Modified Eagle Medium (DMEM, Life Technologies, Carlsbad, CA) supplemented with 10% of fetal bovine serum (FBS), 10mM non-essential amino acids (Life Technologies), 10mM HEPES, penicillin (100 units/ml), streptomycin (100 mg/ml), and 2mM L-glutamine at 37°C with 5% CO₂.

Plasmids
pFNX-HCV is the plasmid we synthesized based on the chimeric sequence of J6/JFH1 virus. We introduced seven nucleotide substitutions, resulting in synonymous mutations to the genome. To generate the 15-nt insertion library, the NotI site was also mutated. The construct and sequence are available upon request. Individual mutant viruses with 15-nt insertion or epitope tags were constructed by site-directed PCR mutagenesis. Primer sequences for each 15-nt insertion or tagged virus are available upon request.

**Antibodies**

Mouse anti-HA and anti-FLAG, as well as rabbit anti-FLAG antibodies were purchased from Sigma (St. Louis, MO, F1804, H9658, F7425). The anti-OLLAS antibody was obtained from Novus Biologicals (Littleton, CO), while the rabbit anti-myc antibody 71D10 was acquired from Cell Signaling (Danvers, MA). The mouse antibody C7-50, used to detect HCV core protein, was purchased from Abcam (Cambridge, MA). Furthermore, the anti-NS5A mouse antibody 9E10 (Lindenbach, Evans et al. 2005) and the human anti-E2 antibody, CBH5 (Hadlock, Lanford et al. 2000) were generous gifts from C. M. Rice (Rockefeller University) and S. Foung (Stanford University), respectively. Anti-mouse secondary antibodies conjugated to fluorophores Alexa 405, 488 or 594 were purchased from Life Technologies (Carlsbad, CA), as was the anti-human secondary antibody conjugated to Alexa 555. Anti-rabbit secondary antibody conjugated to Atto 647N, used in four-color fluorescence microscopy, was obtained from Active Motif (Carlsbad, CA).

**Transposon mutagenesis**

The plasmid carrying FNX-HCV was subjected to *in vitro* bacteriophage Mu transposon mutagenesis (Mutation generation system kit, Thermo Scientific, Waltham, MA), followed by incubation on plates containing kanamycin and ampicillin to select for bacteria transformed with transposon-inserted plasmids. Pooled plasmids were isolated and digested with NotI enzyme (New England Biosciences, Ipswich, MA) to remove the transposon fragment, followed by a
self-ligation step, which left fifteen nucleotides (15-nt, 5′-NNNNNTGCGGCCGCA-3′ with N= duplicated 5 nucleotides from target DNA) randomly inserted across the virus genome. We obtained a mutant plasmid library isolated from more than one million individual bacterial colonies, which is 100-fold greater than the number of all possible insertions.

In vitro transcription

A total of 16 µg of the pFNX library DNA was used for in vitro transcription using the T7 Ribomax Express kit (Promega, Madison, WI). RNA was further treated with DNase to remove template DNA.

Transfection

120µg of RNA was delivered into 4.8×10^7 Huh-7.5.1 cells by electroporation as described before (Arumugaswami et al. 2008). Electroporation was conducted with the settings of 270 V, 100 ohms, and 960 µF. Cells were resuspended in 40 ml of complete DMEM and plated in T-75 culture flasks. At 8 hours post-transfection (8hpt), media containing dead cell debris in the culture flasks and plates were replaced with fresh complete DMEM. Following an incubation period of 96h, cell culture media was removed and tested for infectivity. The resulting titer of this reconstituted mutant library was 8.4x10^4 focus forming units (ffu) per ml.

Insertion library passage

We passaged the library in Huh7.5.1 cells for another round at a multiplicity of infection (MOI) of 0.2. Total RNA was isolated from the cells using Tri-reagent (Molecular Research Center Inc. Cincinnati, OH). The DNase-treated and -purified RNA was used for functional profiling analysis. Total RNA from these infected cells at the end of the experiment constituted pool 2 (P2). The in vitro transcribed RNA mutant library served as the input library pool 0 (P0). Finally, the total RNA of transfected cells collected at 96 hours post transfection (96hpt) represented pool 1 (P1).

Library preparation
Total extracted RNA from P0, P1 and P2 was reverse transcribed into cDNA with Superscript III (Life Technologies). The cDNA then served as template for PCR to amplify thirteen overlapping fragments covering the entire virus genome. These amplicons were mixed in an equimolar ratio to a final amount of 5ug. The mixtures were randomly fragmented by sonication and the ends were repaired with T4 DNA polymerase, Klenow DNA polymerase, and T4 polynucleotide kinase to produce blunt, double-stranded DNA fragments. These fragments were ligated to one of the sequencing adaptors (adaptor A), followed by restriction with NotI enzyme (NEB). The digested fragments were ligated to a biotin-modified adaptor (adaptor B) through NotI overhangs, GGCC, with *E.Coli* ligase at 16°C overnight. The ligated products were purified by streptavidin labeled magnetic beads (Dynabeads®M-270 Streptavidin, Life Technologies), and enriched by PCR for the adapter-modified DNA fragments. The pooled PCR products were pooled together and processed for Illumina library preparation according to the manufacturer’s instructions. Only the fragments with insertion sites could be amplified and sequenced, which allowed us to quantitatively and precisely determine the mutants in the libraries.

**Illumina sequencing**

Sequencing of multiplexed insertion mutant libraries was carried out using the Illumina Genome Analyzer IIX (Illumina, San Diego, CA).

**Sequencing analysis**

We mapped the HCV fragment from the reads to identify the base (N5) preceding the 15-nt sequence (N₁N₂N₃N₄N₅TGCGGCCGCAN₁N₂N₃N₄N₅, insertion being underlined) as the insertion site. Briefly, raw data was first processed to eliminate bad quality reads by SeqTrim (Falgueras et al., 2010). In this study, different populations are linked with distinct barcodes; a successfully enriched HCV sequence read would have started with an 11-mer segment: 5’-DDDDggccgca N₁N₂N₃N₄N₅ . . . -3’. DDDD stands for the 4-mer barcode sequence and N represents the duplicated HCV sequence from the transposon insertion. For each pool (P0, P1 and P2), we
recorded the number of reads at each transposon insertion location. To assign a phenotype for each insertion mutant, the ratio of sequencing reads between selected (P1 or P2) and the non-selected P0 pool was calculated. For additional detail, refer to the supplemental methods.

*Mapping functional data to protein structures*

The functional profile data was incorporated into the NMR and crystal structures of HCV proteins using the PyMOL molecular visualization program (DeLano Scientific, Palo Alto, CA). All structures were obtained by downloading deposited PDB files from the Research Collaboratory for Structural Bioinformatic (RCSB) PDB protein data bank at [http://www.rcsb.org/pdb/](http://www.rcsb.org/pdb/). For structures based on strains that differed from our FNX virus, we performed an amino acid sequence alignment using the ClustalW program, available online (Bernstein, Koetzle et al. 1977; Larkin, Blackshields et al. 2007; Goujon, McWilliam et al. 2010). This allowed us to better evaluate whether amino acid residues from our screen could be matched to the amino acid residues from deposited protein structures.

*Cloning of individual insertion mutants*

Individual mutant viruses with 15-nt insertion or epitope tags were constructed using the PCR overlap extension method. Insertion and substitution mutations were confirmed by individual Sanger sequencing.

*Viral titration*

The virus titer was assayed by measuring the foci forming units of infectious virus particles per ml of supernatant. Culture supernatant taken from different time points, e.g. 48hpt, 72hpt and 96hpt were diluted in a 10-fold serial manner before applying onto naive Huh7.5.1 cells seeded in 96-well plates. 72hpt, the cells were fixed with 100% methanol and immunostained for core antigen with antibody C7-50 (Abcam), followed by goat anti-mouse
Alexa 488-conjugated secondary antibody. Virus titer was determined by counting the number of core antigen-positive foci at the highest dilutions.

**Immunofluorescence assay**

At 48h post transfection, Huh7.5.1 cells in 48-well plates or 8-well chamber slides (Millicell EZ slide, EMD Millipore, Billerica, MA) were fixed with 4% paraformaldehyde for 5 min at room temperature followed by overnight incubation at 4°C. After three washes with PBS, the cells were permeabilized with 0.05% Triton-x 100 for 10min and with blocked with 10% FBS, 3% BSA. All incubations with primary antibody were carried out overnight. Secondary antibodies conjugated to respective fluorescent dyes were added and incubated for 1 hour at room temperature. Cells were washed three times with PBS and nuclei was stained with Hoechst 33342 (Life Technologies) before imaging by epifluorescence microscopy. For four-color confocal microscopy, secondary antibodies were conjugated to Alexa 405, Alexa 555 and Atto 647N respectively (anti-mouse, anti-human, anti-rabbit). In addition, a separate incubation with 50 mg/ml of Bodipy 493/503 (Life Technologies) for 1h was used to stain lipid droplets. Lastly, slides were mounted in Prolong Gold Antifade reagent. Slides were analyzed by confocal microscopy using either a Nikon Eclipse Ti imaging system or Leica TCS SP2 AOBS system.
### Table 3-1 | Summary of high-throughput sequencing of recovered HCV libraries

<table>
<thead>
<tr>
<th>Sample</th>
<th>Total no. of reads</th>
<th>No. of unique insertion positions detected</th>
<th>No. of insertion positions with reads = 0</th>
<th>Average read count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input (P0)</td>
<td>6,415,885</td>
<td>8,398</td>
<td>1,283</td>
<td>662</td>
</tr>
<tr>
<td>Passage 1 (P1)</td>
<td>3,048,022</td>
<td>4,396</td>
<td>5,290</td>
<td>315</td>
</tr>
<tr>
<td>Passage 2 (P2)</td>
<td>4,547,714</td>
<td>1,426</td>
<td>8,260</td>
<td>468</td>
</tr>
<tr>
<td>Gene region</td>
<td>Genome Location (nucleotides)</td>
<td>Total insertions in respective region</td>
<td>Average P1/P0 ratio</td>
<td>Average P2/P0 ratio</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------</td>
<td>---------------------------------------</td>
<td>---------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Core</td>
<td>341-914</td>
<td>574</td>
<td>497</td>
<td>1.58</td>
</tr>
<tr>
<td>E1</td>
<td>915-1490</td>
<td>576</td>
<td>498</td>
<td>1.38</td>
</tr>
<tr>
<td>E2</td>
<td>1491-2591</td>
<td>1100</td>
<td>983</td>
<td>0.95</td>
</tr>
<tr>
<td>p7</td>
<td>2589-2780</td>
<td>191</td>
<td>176</td>
<td>1.67</td>
</tr>
<tr>
<td>NS2 TM Domain</td>
<td>2781-3082</td>
<td>302</td>
<td>273</td>
<td>1.65</td>
</tr>
<tr>
<td>NS2 Protease Domain</td>
<td>3083-3430</td>
<td>348</td>
<td>309</td>
<td>0.06</td>
</tr>
<tr>
<td>NS3</td>
<td>3429-5324</td>
<td>1895</td>
<td>1657</td>
<td>0.01</td>
</tr>
<tr>
<td>NS4A</td>
<td>5325-5486</td>
<td>161</td>
<td>142</td>
<td>0.1</td>
</tr>
<tr>
<td>NS4B</td>
<td>5484-6269</td>
<td>784</td>
<td>695</td>
<td>0.08</td>
</tr>
<tr>
<td>NSSA Domain I</td>
<td>6270-6907</td>
<td>638</td>
<td>542</td>
<td>0.02</td>
</tr>
<tr>
<td>NSSA Rest</td>
<td>6908-7667</td>
<td>760</td>
<td>664</td>
<td>7.9</td>
</tr>
<tr>
<td>NSSB</td>
<td>7668-9377</td>
<td>1710</td>
<td>1481</td>
<td>0.07</td>
</tr>
<tr>
<td>NSSB C-terminus</td>
<td>9378-9442</td>
<td>64</td>
<td>61</td>
<td>5.6</td>
</tr>
<tr>
<td>Validated Wild-type-like Mutants</td>
<td>7443-7627</td>
<td>-</td>
<td>7</td>
<td>5.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>341-9442</td>
<td>9102</td>
<td>7980</td>
<td>1.1</td>
</tr>
</tbody>
</table>
Table 3-3 | Summary of non-lethal insertions in globular domains of NS5B polymerase

<table>
<thead>
<tr>
<th>Genome position</th>
<th>Preceding amino acid</th>
<th>Amino Acid sequence</th>
<th>Normalized P0 Reads</th>
<th>Normalized P1 Reads</th>
<th>Normalized P2 Reads</th>
<th>P1/P0</th>
<th>P2/P0</th>
</tr>
</thead>
<tbody>
<tr>
<td>7844</td>
<td>L60</td>
<td>FDRTQVRPQVLDAYHD</td>
<td>0.6</td>
<td>3.3</td>
<td>0.7</td>
<td>5.3</td>
<td>1.1</td>
</tr>
<tr>
<td>8112</td>
<td>A150</td>
<td>EVFCVDPAADPAKGGK</td>
<td>15</td>
<td>60</td>
<td>2</td>
<td>4</td>
<td>0.2</td>
</tr>
<tr>
<td>8581</td>
<td>A305</td>
<td>LAACKACGRKAGIVAP</td>
<td>3</td>
<td>9</td>
<td>2</td>
<td>2.8</td>
<td>0.7</td>
</tr>
<tr>
<td>8803</td>
<td>R379</td>
<td>LGPRGRCGSRREYLT</td>
<td>347</td>
<td>990</td>
<td>69</td>
<td>2.9</td>
<td>0.2</td>
</tr>
<tr>
<td>8995</td>
<td>L443</td>
<td>TLDQNLCGRLNFEMYG</td>
<td>38</td>
<td>101</td>
<td>16</td>
<td>2.7</td>
<td>0.4</td>
</tr>
<tr>
<td>9226</td>
<td>R523</td>
<td>GGKAACGRTVCGRYLF</td>
<td>456</td>
<td>2548</td>
<td>2052</td>
<td>5.6</td>
<td>4.5</td>
</tr>
<tr>
<td>Epitope Tag</td>
<td>Amino Acid Sequence</td>
<td>Length (No. of amino acid residues)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
<td>------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>HHHHHHH</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLAG</td>
<td>DYKDDDDDK</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA</td>
<td>YPYDVPDYA</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myc</td>
<td>EQKLISEEDL</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>FLNCCPGCCMEP</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLLAS</td>
<td>SGFANELGPRLMGK</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA2x</td>
<td>YPYDVPDYAYPYDVPDYA</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLAG3x</td>
<td>DYKDHGDGYKDHDIDYKDDDK</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3-5 | Filtering of functional profile produces a candidate list of non-structural HCV protein regions with roles in post-replication steps

<table>
<thead>
<tr>
<th>Genome Position</th>
<th>HCV Protein ID</th>
<th>Amino Acid Sequence</th>
<th>Input reads</th>
<th>P1 Reads</th>
<th>P2 Reads</th>
<th>P1/Input</th>
<th>P2/Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>5328</td>
<td>NS4A</td>
<td>LEVMTSTAAASTWVLAG</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>5421</td>
<td>NS4A</td>
<td>GRLHVNAANQRVVVA</td>
<td>12</td>
<td>8</td>
<td>0</td>
<td>0.7</td>
<td>0</td>
</tr>
<tr>
<td>5428</td>
<td>NS4A</td>
<td>HVNQRVCGRVVAVPKD</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>5571</td>
<td>NS4B</td>
<td>GLLQQASAAAASKQAQD</td>
<td>211</td>
<td>264</td>
<td>2</td>
<td>1.2</td>
<td>0.007</td>
</tr>
<tr>
<td>5574</td>
<td>NS4B</td>
<td>LLQQASNAAAASKQAQDI</td>
<td>24</td>
<td>26</td>
<td>0</td>
<td>1.1</td>
<td>0</td>
</tr>
<tr>
<td>5583</td>
<td>NS4B</td>
<td>QASKQAHAAAAQDIQA</td>
<td>73</td>
<td>87</td>
<td>0</td>
<td>1.2</td>
<td>0</td>
</tr>
<tr>
<td>5584</td>
<td>NS4B</td>
<td>ASKQAQCGRTQDIQPAM</td>
<td>699</td>
<td>471</td>
<td>3</td>
<td>0.7</td>
<td>0.004</td>
</tr>
<tr>
<td>5594</td>
<td>NS4B</td>
<td>QAQDIQLRPOQPAMQS</td>
<td>20</td>
<td>24</td>
<td>0</td>
<td>1.2</td>
<td>0</td>
</tr>
<tr>
<td>5596</td>
<td>NS4B</td>
<td>AQDIQPGRPKAMQS</td>
<td>39</td>
<td>42</td>
<td>0.2</td>
<td>1.1</td>
<td>0.006</td>
</tr>
<tr>
<td>5597</td>
<td>NS4B</td>
<td>AQDIQPVRPOQPAMQS</td>
<td>190</td>
<td>134</td>
<td>0</td>
<td>0.7</td>
<td>0</td>
</tr>
<tr>
<td>5607</td>
<td>NS4B</td>
<td>IQPAMQAAAQASWPKV</td>
<td>535</td>
<td>521</td>
<td>2</td>
<td>1.0</td>
<td>0.003</td>
</tr>
<tr>
<td>6847</td>
<td>NS5A</td>
<td>LPCEPECGRTEPDADV</td>
<td>174</td>
<td>122</td>
<td>0.9</td>
<td>0.7</td>
<td>0.005</td>
</tr>
<tr>
<td>7290</td>
<td>NS5A</td>
<td>VAGCALPAAALPPKKA</td>
<td>35</td>
<td>38</td>
<td>0</td>
<td>1.1</td>
<td>0</td>
</tr>
<tr>
<td>7671</td>
<td>NS5B</td>
<td>TTVCCSIAASMSYSWT</td>
<td>82</td>
<td>112</td>
<td>0</td>
<td>1.4</td>
<td>0</td>
</tr>
<tr>
<td>8793</td>
<td>NS5B</td>
<td>VSVALGPAAGPRGRRR</td>
<td>8</td>
<td>13</td>
<td>0</td>
<td>1.6</td>
<td>0</td>
</tr>
<tr>
<td>8979</td>
<td>NS5B</td>
<td>ILMVQDTAAADTLQNL</td>
<td>36</td>
<td>41</td>
<td>0.2</td>
<td>1.1</td>
<td>0.006</td>
</tr>
<tr>
<td>8991</td>
<td>NS5B</td>
<td>QDTLDQNAAQNLNFS</td>
<td>31</td>
<td>31</td>
<td>0</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>9373</td>
<td>NS5B</td>
<td>VSRARPGRPRSLF</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>0.7</td>
<td>0</td>
</tr>
</tbody>
</table>

* Filtering Criteria: P0>5; 0.65<P1/P0<1.85; P2/P0 ≤ 0.01; HCV proteins NS3-5B only. P0, P1 and P2 total reads normalized to 1,000,000.
Figure 3-1. (A) Generation of an HCV insertion library by random insertion mutagenesis of a plasmid carrying the HCV genome (pFNX-HCV). (B) Passaging the HCV insertion library. Input, P0=untransfected RNA library. First passage, P1=96h post-transfection. Second passage, P2=72h post-infection. (C, D) Library preparation and Illumina sequencing.
Figure 3-2. (A) Digital counting of each insertion mutant indicated by different insertion sites, in the input and selected libraries. Bar graph of raw sequencing reads from all 15-bp-insertion mutants recovered at each passage (Input [P0], passage 1 [P1], passage 2 [P2]). Each bar represents a unique mutant with bar heights indicating the sequencing reads for each mutant. Matching colors between the map and a cartoon of HCV genome organization further illustrate the location of insertions relative to each HCV gene. (B) Closer look at the genetic footprint of insertion sites located in the p7 and NS2 genes, presented as a bar graph. X-axis shows genome location, while the y-axis indicates sequencing counts. Furthermore, a cartoon of p7 and NS2 protein domain organization serves as a reference. TM1, TM2 helix: Transmembrane helix 1 and 2. TMS1, TMS2, TMS3: Transmembrane segment 1, 2 and 3.
Figure 3-3 | Examination of P1/P0 and P2/P0 ratios across both the entire HCV genome and individual viral genes.

Figure 3-3. (A) Histogram containing all insertion mutants (=all) and their corresponding P1/P0 and P2/P0 values. Data is binned according to a selected range (0, 0.01, 0.1, 0.5, 1, 2, 5, More). Black bar height indicates frequency of insertion mutants at corresponding P1/P0 bin values. Lower graphs report P2/P0 values. (B-E) Histograms separated by HCV genes. Further grouping is based on known function of encoded protein (structural proteins, replication-complex, viroporin or multifunctional proteins). N=Number of insertion mutants within the corresponding gene.
Figure 3-4. Graphing P1/P0 and P2/P0 values for AAA, CGR and RPH/Q insertion motifs across the HCV genome identifies regions of preference for certain motifs. 

Shown are P1/P0 (upper two panels) and P2/P0 (lower two panels) ratios across the entire protein-coding regions of the HCV genome. Numbers on the x-axis refer to nucleotide positions of the HCV genome. Each point represents one insertion mutant. Colors further group insertion mutants into three categories (blue, red, grey) based on the identity of the first three inserted amino acids: Alanine-Alanine-Alanine (AAA), Cysteine-Glycine-Arginine (CGR) or Arginine-Proline-Glutamine/Histidine (RPQ/H). Black arrows point towards areas that contain an uneven distribution of the AAA, CGR and RPH/Q categories. To better visualize the uneven distribution around nucleotide number 841, we set the y-axis boundary to a maximum of 15 and 5 for P1/P0 and P2/P0 ratios respectively.
Figure 3-5 | Effect of location of 15-nt insertion relative to reading frame in select areas of viral genome.

(A) Histogram containing all insertion mutants (=all HCV proteins) and their corresponding P1/P0 and P2/P0 values. Data is binned according to a selected range (0, 0.01, 0.1, 0.5, 1, 2, 5, More). Black bar height indicates frequency of insertion mutants at corresponding P1/P0 bin values. Lower graphs report P2/P0 values. Graphs further group insertion mutants into three categories based on the identity of the first three inserted amino acids: Alanine-Alanine-Alanine (AAA), Cysteine-Glycine-Arginine (CGR) or Arginine-Proline-Glutamine/Histidine (RPQ/H).

(B) Histogram containing insertion mutants within Core domain III helix and E2 Hypervariable region and their corresponding P1/P0 and P2/P0 values. N=Number of insertion mutants within the corresponding gene.

Figure 3-5. (A) Histogram containing all insertion mutants (=all HCV proteins) and their corresponding P1/P0 and P2/P0 values. Data is binned according to a selected range (0, 0.01, 0.1, 0.5, 1, 2, 5, More). Black bar height indicates frequency of insertion mutants at corresponding P1/P0 bin values. Lower graphs report P2/P0 values. Graphs further group insertion mutants into three categories based on the identity of the first three inserted amino acids: Alanine-Alanine-Alanine (AAA), Cysteine-Glycine-Arginine (CGR) or Arginine-Proline-Glutamine/Histidine (RPQ/H). (B) Histogram containing insertion mutants within Core domain III helix and E2 Hypervariable region and their corresponding P1/P0 and P2/P0 values. N=Number of insertion mutants within the corresponding gene.
**Figure 3-6 | Functional annotation of the C-terminal region of the HCV core protein**

**Figure 3-6.** Heat map of P2/P1 values in the C-terminal region of the HCV core protein. Each box provides the phenotype of mutants containing a 5-aa insertion following the indicated amino acid. Gray colors indicate an absence of the respective insertion mutant in the original input pool. Mutants are separated by the motif contained in the actual insertion (AAA, CGR or RPH/Q). Color bar ranging from cyan to magenta corresponds to a range of P2/P1 values (Log P2/P1=1 to Log P2/P1=−4). Previous studies solved the structure of this region (PDB ID: 2LIF) by NMR spectroscopy (Oehler, Filipe et al. 2012). The assigned secondary structure from the Define Secondary Structure of Proteins (DSSP) algorithm provides an overview of helices (zig zag lines) and unstructured loops (straight lines). The top panel aligns HCV core amino acid sequences of HCV strains from this study (FNX24) and the NMR spectroscopy experiment. ClustalW alignment shows asterisks (*) for conserved residues, a colon (:) for highly similar residues and a dot (.) for similar residues. Ribbon diagrams below the heat map provide an additional view of the annotated structure, along with a tentative model of the peptide’s orientation within the ER membrane. Black arrows: cleavage sites of signal peptide peptidase (SPP) and signal peptidase (SP).
Figure 3-7 | Validation of an essential role for C-terminal region of HCV core in infectious virus production

(A) Shown are the locations of the mutations and amino acid substitutions made in the context of a HCV JC1 clone resulting in five recombinant viruses (nomenclature: NFT, GNLP, IF, ASC and PVS). Numbers represent the HCV polyprotein residue number. Alanine and Val-Leu-Val substitutions were introduced by extension PCR mutagenesis with primers containing the desired changes. The resulting PCR fragments representing the entire core region and part of E1 were subcloned into a pFLAG-CMV vector, sequence-verified and finally swapped with the core segment of a plasmid containing the complete HCV genome. (B) To confirm lethal phenotypes, we transfected RNA transcripts of each viral mutant into Huh-7.5.1, collected cell culture media 48h and 72h after transfection, and measured infectivity by a limiting dilution assay. (C) To confirm core processing defects, we lysed cells 48h post-transfection in protein gel sample buffer and separated lysates by SDS-15% polyacrylamide gel electrophoresis (PAGE). Western blot analysis was carried out using antibodies against HCV core and cellular actin (loading control).
Figure 3-8 | Functional annotation of viral transmembrane proteins p7 and NS2 using P2/P1 ratios

A

Figure 3-8. (A) Heat map of P2/P1 values in the p7 protein. Heat map and ClustalW alignment were generated as in Figure 3-7. DSSP secondary structure also indicates a turn (arc) towards the C-terminal end of the p7 protein. A recent study used p7 from a genotype 5a strain in their structural determinations (OuYang, Xie et al. 2013). PDB ID: 2M6X. (C) Heat map of P2/P1 values in the N-terminal region of NS2. Helical segments (zig zag lines), turns (arcs) and unstructured loops (straight lines) were deduced from NMR analyses (Jirasko, Montserret et al. 2008; Jirasko, Montserret et al. 2010) of NS2 peptides ([aa 1-27], [27-59], [60-99]). PDB ID’s: 2JY0, 2KW1, 2KZ. A tentative model of NS2 topology within the ER membrane serves as a reference.
Figure 3-9. Surface views of annotated p7 viroporin structure. Gray colors indicate an absence of the respective insertion mutant in the original input pool. Amino acid residues in p7 structure were colored according to the heat maps generated in Figure 3-9A. Each row separates the functional data by the motif contained in the actual insertion (AAA, CGR or RPH/Q). Color bar ranging from cyan to magenta corresponds to a range of P2/P1 values. We used the same color range as shown in the color bar of Figure 3-8A. Surface views were generated from the recently published NMR structure of the p7 viroporin (OuYang, Xie et al. 2013). Shown are a side view of the viroporin as well as views from the top and bottom. Surface transparency was set to 65% for top and bottom views to reveal ribbon representations inside the viroporin. PDB ID: 2M6X.
Figure 3-10. Mapping tolerated insertions to the structure of NS5B polymerase. Upper panel: The surface view at the figure top maps the locations of amino acid positions tolerating 5-aa insertions to a NS5B crystal structure lacking the C-terminal membrane anchor (PDB ID: 3I5K) (Schmitt, Scrima et al. 2011). Red color marks amino acids preceding tolerated insertions. Left and right views differ by 180°. Lower panel: Zoom-in of amino acid positions tolerating 5-aa insertions. Here, we used a ribbon view to present the NS5B structure.
Figure 3-11. (A) P2 values of insertion mutants across the entire HCV genome. This plot only contains data points that fall within a cluster of at least three consecutive insertion mutants with nonzero P2/P0 ratios. To increase stringency, we discarded all insertion mutants with P1/P0 ratios less than 0.1. A cartoon of HCV genome organization provides a reference. (B) P2/P0 ratios of insertion mutants in the HCV core protein. (C) P2/P0 ratios of insertion mutants in the HCV p7 protein. (D) P2/P0 ratios of insertion mutants in the HCV NS5A protein. Insertions highlighted in red represent positions that we subsequently tested for epitope-tag-insertion.
Figure 3-12 | Engineering of infectious epitope-tagged HCV.

(A) Viral fitness of epitope-tagged HCV. Bar graph shows the fitness of individually cloned epitope-tagged viruses. Fitness is defined as the ratio of mutant virus infectivity to the wild-type infectivity, as determined by a limited dilution assay of supernatants taken from transfected Huh-7.5.1 cells. Numbers (349-7177) refer to the genome position after which the sequence encoding the respective epitope tag was inserted. HA= Human influenza hemagglutinin. His= Polyhistidine. OLLAS= *Escherichia coli* OmpF Linker and mouse Langerin fusion sequence. TC=Tetracysteine.

(B) Immunofluorescence assay of cells infected with epitope-tagged HCV. Huh-7.5.1 cells were infected with indicated epitope-tagged viruses. Three days after infection, we fixed cells and processed samples for an immunofluorescence assay using antibodies against the inserted epitope tag. We examined immunostained cells using epifluorescence microscopy. Note that Hoechst dye stains the nuclei of both infected and uninfected cells. (C) Immunofluorescence assay of cells infected with either a representative epitope-tagged virus (2792-myc) or untagged FNX24 virus. To detect tagged proteins, we used a specific antibody that binds specifically to the myc-tag. (D) Immunofluorescence assay of cells infected with HA-tagged p7 mutants. To detect tagged proteins, we used an antibody that binds specifically to the HA-tag epitope. Scale bar = 50µm.
Figure 3-13 | Multi-color fluorescence microscopy of cells infected with epitope-tagged NS5A virus

Composite: Anti-FLAG, Anti-E2, Anti-Core, LD

Figure 3-14. We fixed cells infected with the 7177-FLAG epitope-tagged virus at 48h post-infection. This virus contains a FLAG tag within the NS5A reading frame. We then performed an immunostaining protocol that would allow us to simultaneously look at core, E2, LD and FLAG subcellular localization. We stained fixed cells sequentially with anti-FLAG, anti-E2 antibody CBH-5 and anti-core antibody C7-50. We then carried out secondary staining using Bodipy 493/503 (white, LDs) as well as antibodies anti-human-IgG-Alexa555 (red, E2), anti-rabbit-IgG-Atto647N (blue, FLAG), and anti-mouse-IgG-Alexa405 (green, core). Images show a four-channel sequential laser scan of one optical confocal section through a representative cell. Images were captured at 60nm pixel size with a Nikon Ti microscope and a 100x oil objective. For publication purposes, images were median-filtered and contrast-enhanced. Scale bar = 5µm.
Figure 3-14 | Validation of NS4B insertion mutants with defects in post-replication steps

A

<table>
<thead>
<tr>
<th></th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anti-NS5A</td>
<td>Hoechst</td>
</tr>
<tr>
<td>WT</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
</tr>
<tr>
<td>5571</td>
<td><img src="image7" alt="Image" /></td>
<td><img src="image8" alt="Image" /></td>
</tr>
<tr>
<td>5584</td>
<td><img src="image13" alt="Image" /></td>
<td><img src="image14" alt="Image" /></td>
</tr>
<tr>
<td>5597</td>
<td><img src="image19" alt="Image" /></td>
<td><img src="image20" alt="Image" /></td>
</tr>
<tr>
<td>5607</td>
<td><img src="image25" alt="Image" /></td>
<td><img src="image26" alt="Image" /></td>
</tr>
<tr>
<td>5615</td>
<td><img src="image31" alt="Image" /></td>
<td><img src="image32" alt="Image" /></td>
</tr>
<tr>
<td>E1E2 Del</td>
<td><img src="image37" alt="Image" /></td>
<td><img src="image38" alt="Image" /></td>
</tr>
</tbody>
</table>

B

![Image](image43)

Figure 3-14. (A) Immunofluorescence assay of cells transfected with individual NS4B insertion mutant RNA.

Numbers indicate genome positions at which 5-aa insertions were introduced and correspond to mutants on candidate list shown in Table 3-5. Two and three days after infection, we fixed cells and processed samples for an immunofluorescence assay using antibodies against NS5A. We examined immunostained cells using epifluorescence microscopy. Note that Hoechst dye stains the nuclei of both infected and uninfected cells. E1E2 Del mutant contains
a large deletion in the E1 and E2-coding segments of the HCV genome and serves as an assembly-deficient control. Scale bar: 200µm. (B) Using a limited dilution assay, we determined the infectivity of supernatants harvested at various days after transfection of NS4B mutant RNA into Huh-7.5.1 cells. Ffu, focus-forming units.

REFERENCES


CHAPTER 4

High-resolution imaging of the LD microenvironment in HCV-infected cells
ABSTRACT

Current models of the assembly of hepatitis C virus (HCV) propose that cellular fat storage organelles, called lipid droplets (LDs), provide a platform for potential assembly sites. The co-localization of various viral proteins within this LD microenvironment has been shown to be essential for infectious virus production. However, the topological organization of this LD microenvironment and the exact location of HCV assembly remain unclear. In this chapter, we used a combination of electron microscopy, electron tomography, immunoelectron microscopy, superresolution light microscopy and multi-color confocal laser microscopy to analyze the spatial architecture and distribution of viral proteins in the LD microenvironment. Ultrastructural changes occurring near LDs included an accumulation of membrane vesicles, ER cisternae and particles resembling virions. Electron tomography clearly identified double-membrane vesicles in this environment, as well as particles budding from ER cisternae. To characterize the localization of viral proteins around LDs, we performed immuno-EM, superresolution light microscopy and multi-fluorescence microscopy of HCV-infected cells. We confirmed that HCV core protein formed clusters close to LDs, which often overlapped with areas containing E2 glycoprotein. Using superresolution light microscopy we were able to detect core-positive and NS5A-positive clusters with sizes between 100-150nm. Multi-color fluorescence microscopy also revealed co-localization of core, E2 and NS5A in larger 400-800nm clusters close to LDs. To dynamically track the subcellular localization the NS5A protein, an essential protein for virus assembly, we developed a panel of fluorescent reporter viruses for live-cell imaging studies. We show that a substitution mutation in the NS5A protein blocks the HCV life cycle at a specific step, leading to accumulation of core, E2 and NS5A close to LDs. Pharmacological treatment with daclatasvir, a potent HCV antiviral targeting the NS5A protein, resulted in a similar change in NS5A localization. Our results provide high-resolution snapshots of HCV-induced structures in the LD microenvironment, improving our understanding of the sequence of HCV assembly
events. Moreover, we demonstrate an essential function of a NS5A domain I residue in post-replication steps of the HCV life cycle.

INTRODUCTION

Hepatitis C virus (HCV), a positive-strand RNA virus and member of the Flaviviridae family, poses a global health problem, with an estimated 160 million individuals (2.35% of the world’s population) chronically infected (Lavanchy 2011). In these individuals, HCV can lead to chronic hepatitis, liver disease and hepatocellular carcinoma (Shepard, Finelli et al. 2005). Moreover, back-calculation models based on US epidemiological data estimate a rise in HCV mortality over the next 20 years (Deuffic-Burban, Poynard et al. 2007). Current HCV therapies, which mostly involve prolonged treatments with ribavirin and pegylated interferon injections, are expensive and effective only in a subset of patients. No vaccine is available yet. Even though the entry of first-generation HCV-specific protease inhibitors into the clinic has greatly complemented the current standard of care, therapeutic options remain hampered by significant side effects and the occurrence of drug resistance (Vermehren, Susser et al. 2012). Therefore, the development of additional selective antiviral strategies represents an urgent need. Targeting of specific steps of the viral life cycle, on the other hand, necessitates a fundamental understanding of the basic principles governing the various steps of the HCV life cycle.

A general feature of all positive-strand RNA viruses is the remodeling of intracellular membranes creating an optimized environment where RNA replication and virion assembly can occur (Miller and Krijnse-Locker 2008). While the establishment of such ‘microenvironments’ facilitates the coordination of different steps in the viral life cycle, it might also serve to shield viral RNA, especially double-stranded RNA intermediates, from recognition by host immune responses. Many studies have analyzed these specific membrane changes, referred to as the “membranous web” (MW), and their function as RNA replication sites (Egger, Wolk et al. 2002;
Visualization of the virion assembly site, however, has remained much more elusive, despite an understanding that HCV assembly likely uses similar membrane platforms for the successful incorporation of virion components.

Mechanistic studies on HCV assembly have lagged studies of other steps of the viral life cycle, namely virus entry and RNA genome replication; a main reason for this has been the lack of an experimental cell culture system to study assembly of infectious viral particles. This changed in 2005, when two independent groups reported successful propagation of HCV in cell culture (Lindenbach, Evans et al. 2005; Wakita, Pietschmann et al. 2005). Since then, using this infectious cell culture system, numerous studies have uncovered the involvement of virtually every HCV protein, both structural and non-structural, in the HCV assembly process (for a review, see Popescu, Rouille et al. 2011). Current models of this process suggest that HCV virions assemble on or near lipid droplets (LDs). Assembled viral particles then acquire their envelope at the endoplasmic reticulum (ER), before hijacking the cellular secretory pathway to exit the host cell. Even though the critical role of LDs in this process has been well-documented, it remains unclear whether HCV assembly occurs within juxtaposed ER membranes or at the LD membrane itself.

To achieve the goal of improving the understanding of the organization of HCV assembly sites and viral proteins, various groups have been using electron microscopy and light microscopy to visualize viral particles in the infectious cell culture system (Popescu, Rouille et al. 2011). Due to the nanometer-resolution afforded by electron microscopy, this imaging method is well-suited for the structural exploration of viral particle assembly. However, previous ultrastructural work using the infectious cell culture system has either focused on the structure of extracellular viral particles (Gastaminza, Dryden et al. 2010; Catanese, Uryu et al. 2013) or on the formation of the virus-induced membranous replication compartments (Ferraris, Gosert, Egger et al. 2003; Ferraris, Blanchard et al. 2010; Ferraris, Beaumont et al. 2012; Romero-Brey, Merz et al. 2012).
Beaumont et al. 2012; Romero-Brey, Merz et al. 2012). While these studies have established that purified HCV particles display a size variability with a range of 30nm-86nm, with infectious particles having a specific size of 55-65nm (Gastaminza, Dryden et al. 2010; Catanese, Uryu et al. 2013), the intracellular assembly of these particles inside an infected host cell has not yet been visualized. Moreover, the spatial relationship of the assembly site, the virus-induced membranous web and cellular organelles such as the ER and LDs remains poorly understood.

Among the various imaging techniques available to HCV researchers, confocal microscopy has been the method-of-choice for visualization of viral proteins in infected cells. Much of what is known about HCV assembly stems from confocal microscopy studies on fixed cells. While limited in its resolution to ~300nm, the ability to determine the subcellular localization of viral proteins and their co-localization with cellular markers has shed light on the elusive HCV assembly process. Core protein, which encapsidates the viral RNA, localizes to both the ER and the surface of LDs (Moradpour, Englert et al. 1996; Barba, Harper et al. 1997). LDs are dynamic organelles that store intracellular lipid reserves and play roles in cellular lipid homeostasis (Targett-Adams, Chambers et al. 2003; Fujimoto and Ohsaki 2006; Martin and Parton 2006). Prevention of core’s localization to LDs severely impairs virus infectivity, as well as viral RNA and core protein release (Boulant, Targett-Adams et al. 2007; Miyanari, Atsuzawa et al. 2007; Boulant, Douglas et al. 2008). Moreover, over the course of an infection, NS3, NS4A, NS4B, NS5A and NS5B accumulate in the proximity of the LDs, at the putative site of viral assembly. HCV envelope glycoproteins E1 and E2 are retained in the ER (Rouille, Helle et al. 2006), but co-localize with NS2-dotted structures during the course of an infection (Popescu, Callens et al. 2011). This suggests a complex interplay between viral proteins during the establishment of the viral assembly site. To fully understand this complex set of viral protein interactions occurring in the microenvironment surrounding LDs, imaging approaches that allow tracking of multiple viral proteins at high resolution are required.
While the exact composition of the viral assembly site still remains to be determined, NS5A appears to be the most likely candidate component along with the virion building blocks core, E1, and E2. Numerous studies have demonstrated NS5A’s essential role in HCV RNA replication (Blight, Kolykhalov et al. 2000; Lohmann, Korner et al. 2001; Tellinghuisen, Foss et al. 2008). Importantly, virion morphogenesis also requires NS5A functions (Miyanari, Atsuzawa et al. 2007; Appel, Zayas et al. 2008; Hughes, Griffin et al. 2009). Nonetheless, the precise roles of this multifunctional protein in the HCV life cycle have not yet been fully established. Among the various biological properties ascribed to NS5A are the ability to exist in basal and hyperphosphorylated forms (Evans, Rice et al. 2004; Huang, Staschke et al. 2007; Lindenbach, Pragai et al. 2007; Tellinghuisen, Foss et al. 2008; Reiss, Harak et al. 2013), to bind to HCV RNA (Huang, Hwang et al. 2005; Foster, Belyaeva et al. 2010; Hwang, Huang et al. 2010) and to interact both with other HCV proteins (Dimitrova, Imbert et al. 2003; Shimakami, Hijiikata et al. 2004; Masaki, Suzuki et al. 2008) as well as distinct host factors (Hamamoto, Nishimura et al. 2005; Okamoto, Nishimura et al. 2006; Lim and Hwang 2011; Ploen, Hafirassou et al. 2013; Salloum, Wang et al. 2013; Vogt, Camus et al. 2013). Despite the current lack of a complete understanding of HCV assembly, NS5A seems to function as a key non-enzymatic determinant of this process, likely linking RNA replication and virion assembly steps (Evans, Rice et al. 2004; Appel, Zayas et al. 2008; Tellinghuisen, Foss et al. 2008).

The focus of the experiments presented in this chapter is to dissect the composition and 3-D architecture of the microenvironment surrounding LDs in infected cells. This microenvironment not only houses RNA replication sites, but also HCV assembly sites, whose visualization has eluded previous studies. To image virus-induced intracellular changes, our experiments combined emerging imaging technologies, such as electron tomography and superresolution light microscopy, with traditional approaches of electron microscopy, immunoelectron microscopy, multi-fluorescence confocal light microscopy and time-lapse fluorescence...
microscopy. Since we wanted to apply these imaging technologies to a model system that reproduced the entire infectious HCV life cycle, we chose the Huh-7.5.1-J6/JFH-1 infectious cell culture model. Our results provided snapshots of the spatial relationships within a remodeled microenvironment containing LDs, virus-induced membrane vesicles, ER, and virus-like particles. Moreover, in both fixed and live cells, we visualized the distribution of viral proteins that are critical for the HCV assembly process. Lastly, we demonstrate how a mutation in the NS5A protein as well as treatment with a NS5A-targeting molecule promotes specific phenotypic alterations. Our imaging platform allowed us to detect these changes in NS5A behavior, namely an accumulation of the NS5A protein along with structural proteins in cytoplasmic punctae close to LDs. Collectively, our findings illustrate the need for multi-modal imaging along with advanced viral genetics approaches in the study of complex intracellular processes such as HCV assembly. We suggest that the imaging platform described here should be generally applicable for phenotypic characterization of viral mutants and determining working mechanisms of newly developed drugs.

RESULTS

**Electron microscopy analysis of infected cells reveals the juxtaposition of lipid droplets, membrane vesicles and endoplasmic reticulum**

To analyze the interactions of HCV with its host cell, we performed transmission electron microscopy (TEM) on plastic-embedded preparations of HCV-infected cultured cells. Due to the high resolution afforded by electron microscopy, this approach would allow us to characterize the intracellular environment of infected cells at an ultrastructural level. We infected Huh-7.5.1 cells with the FNX-24 virus, which is based on a highly assembly-competent Jc1 genome. To ensure that nearly all cells in our sample would contain actively replicating virus, we allowed the infection to progress for three days. At that point, we chemically fixed the cells and processed
them for EM using a traditional protocol involving dehydration, plastic embedding and ultrathin sectioning.

In the resulting electron micrographs of ultrathin plastic sections (thickness of 50-80nm), LDs were clearly identifiable as round structures with an amorphous electron-dense interior (Figure 4-1A-E). Because of the known importance of LDs in HCV morphogenesis (Bartenschlager, Penin et al. 2011), we focused our efforts on characterizing the ultrastructural changes occurring near LDs in infected cells. In the cytoplasmic areas near LDs of infected cells, we observed striking morphological changes, which were absent in uninfected control cells. The most notable abnormalities in infected cells consisted of an accumulation of membrane vesicles, reminiscent of a membranous web (MW) structure. We observed this accumulation of membrane vesicles in perinuclear areas, which also contained LDs (Figure 4-1A, area surrounded by dash line). Membrane vesicles displayed single-, double- and multi-layer membranes (representative example, Figure 4-1B, black arrows). Close to the LDs of infected cells, we also noticed several areas with an increase in rough ER (rER) cisternae (Figure 4-1C, black arrow). The rER was non-linear and did not show any organized pattern. Instead, the ER appeared rather convoluted and often overlapped within the plastic section (Figure 4-1D). In some instances, single ER stacks lined the periphery of LDs (Figure 4-1E, black arrows). In summary, these images confirmed the proximity of LDs with membrane vesicles, which are viewed as likely sites of RNA replication, and ER, which is known to be the site of viral protein translation.

**Electron tomography provides three-dimensional views of membrane alterations surrounding LDs**

To increase the amount of cellular space surveyed in one plastic section, we prepared thicker plastic sections (200-300nm). Moreover, we searched the sample for sectioned cellular areas exhibiting an abundance of LDs. Lower-magnification images of a LD-rich infected cell revealed
a multitude of LDs clustered in the cytoplasmic space (Figure 4-2A). Furthermore, we observed numerous changes in the membrane ultrastructure around these LDs (Figure 4-2B). However, the ultrastructural organization of these membranes structures could not be fully appreciated in the close-up views obtained using 2-D transmission electron microscopy, since membranes often consisted of multi-layered, overlapping structures within the 200-nm-thick section (Figure 4-2C-E).

To better characterize the spatial organization of these membrane alterations occurring within the LD-rich region shown in Figure 4-2, we obtained three-dimensional (3-D) reconstructions of several regions-of-interest (ROIs) using electron tomography (ET). We obtained tilt series of the three areas corresponding to Figure 4-2’s close-up views. Following alignment and reconstruction of the entire tomographic volume, we could then create a gallery of digital z-series single-slice images with an approximate thickness of 1-2 nm (Figure 4-3). We also provide videos displaying these tomograms section-by-section (supplemental videos S1-S3). The superior contrast in these single-slice images allowed us to clearly visualize the double-membranes of putative replication compartments (Figure 4-3A-C, blue arrows) as well as invaginations in membrane compartments (Figure 4-3A and C, red arrows). We also observed a double-membrane vesicle containing additional spherical membrane vesicles in its interior (Figure 4-3B, hash tag). These internal vesicles had an electron-lucent interior and appeared to be in contact with adjacent vesicles; moreover, they made contact with the larger double-membrane compartment itself. Our high-resolution snapshots also showed an ER cisternae and filamentous structures making contact with double-membrane vesicles (DMVs) (Figure 4-3B and C, white arrows). In fact, slice z=120 (Figure 4-3C) captured a snapshot of a small, electron-dense, neck-like connection forming a bridge between DMVs and a smaller membrane compartment (marked with red * symbol). Altogether, these images confirmed the presence of DMVs in the immediate vicinity of LD edges. DMVs made contacts with other
membrane structures, as well as rER suggesting that they are part of a larger network and originally derived from ER compartments. Filament bundles further suggested that the area around LDs underwent dynamic changes during infection.

**Particles with sizes of 50-75nm are found at various subcellular localizations, including areas enriched in membrane vesicles and ER**

Having characterized the intracellular environment of infected cells around LDs, we wanted to take advantage of the high resolution of EM to identify intracellular viral particles. Previous studies of purified virus particles and ultrastructural characterizations of HCV-infected cultured cells or tissues reported diameters in the range of 50-75nm for HCV particles (Gastaminza, Dryden et al. 2010; Catanese, Uryu et al. 2013). The high resolution of electron microscopy resolution allows for clear identification of structural features in this size range. After careful screening of cellular sections from infected cells, we obtained snapshots of putative virus particles at various areas of the cell. Representative examples included an extracellular particle close to the plasma membrane (Figure 4-4A, black arrow), intracellular particles contained in membrane compartments (Figure 4-4B and C, black arrows), and free particles in the cytoplasm (Figure 4-4D, black arrow). These particles displayed a round morphology with a thick electron-dense coat. The size of the particles varied from 50 to 75nm. In addition, the electron density of the interior showed some variability from particle to particle (see particles in Figure 4-4C and D). We were not able to clearly assign structural features such as capsid structures, protein spikes or lipid bilayer envelopes to these particles. Due to our inability to unmistakably identify HCV particles based on structural features, we will refer to these putative virus particles as virus-like particle (VLP) instead.

Next, we were interested in determining whether we could image VLPs close to areas showing an accumulation of membrane vesicles, the putative sites of RNA replication. The successful assembly of progeny virions would not only require structural protein components but
also the coordinated incorporation of viral RNA produced at genome replication sites. Indeed, areas rich in DMVs (Figure 4-4E, center image and lower left zoom-in insert) also contained free particles both throughout the same cytoplasmic space and in the immediate surroundings (Figure 4-4E, zoom-in inserts). Moreover, our EM snapshots captured numerous budding events at ER membrane stacks (Figure 4-4E, upper left, upper right and lower right inserts). ER stacks themselves were non-linear and rather convoluted. Taken together, our 2-D electron microscopy analysis spotted VLPs close to potential sites of viral RNA replication (DMVs), as well as sites of viral protein translation (ER membrane).

**Free VLPs as well as membrane-vesicle-associated VLPs locate close to LD interfaces**

Currently, it is not known whether HCV particles assemble at the surface of LDs or at LD-associated ER membranes. In order to characterize particles located near LD interfaces, we took a series of snapshots of LDs exhibiting VLPs in the immediate surroundings. We observed a variety of free VLPs that did not appear to be contained in compartments (Figure 4-5A-E, black arrows). Some of the VLPs clearly displayed an electron-dense coat and an interior that was more electron-lucent. We also found VLPs that appeared associated with membrane compartments (Figure 4-5D, red arrows). We also saw 55 to 60nm-VLPs with well-defined envelopes and rather hollow interiors at the edge of a LD (Figure 4-5E and F). This heterogeneity in size and appearance suggested that these VLPs represented a mix of virions at different stages of the assembly pathway. In fact, some of the imaged VLPs may represent assembly intermediates. In conclusion, we detected VLPs of variable density and size that were associated with membrane structures in close proximity to LDs as well as VLPs that appeared to be unassociated and free. Thus, we were not able to disprove either of the two current models of where HCV particles assemble (ER or LD surface). However, we should note that the complexity of the cellular architecture around LDs, especially with regards to membrane...
structures (see Figures 4-5A-C), made it difficult to unequivocally categorize VLPs as membrane-associated or free-standing using traditional 2-D ultrathin-section EM.

**Electron tomography provides three-dimensional views of spherical particles close to LDs**

To increase the likelihood of capturing entire VLPs in the proximity of LDs in a plastic section, we prepared thicker plastic sections with a thickness of 200-300nm and selected two ROIs as representative examples (Figure 4-6A and B). When analyzed with traditional 2-D EM, the first ROI appeared to contain particles with an electron-dense coat in the immediate vicinity of an LD and ER stack (Figure 4-6A, white arrow). However, the projection image produced by traditional TEM did not clearly resolve the envelopes of overlapping particles. Moreover, in a second ROI, the traditional EM images clearly showed two VLPs (Figure 4-6B, white arrows), which overlapped with an electron-dense area reminiscent of a LD. However, the exact spatial organization between particle, LD and adjacent membrane compartments could not be fully appreciated with the 2D projection image.

To fully resolve structural features and gain a clearer understanding of the spatial relationships of these particles around LDs, we obtained 3-D reconstructions of the above ROIs using electron tomography. We also provide videos displaying these tomograms section-by-section (supplemental video S4 and S5). The surveyed area contained ten spherical VLPs adjacent to the LD boundary (Figure 4-6C, purple colors in rendered volume). Single-slice images allowed us to measure the diameters of these particles. Approximate diameters ranged from 30 to 35nm (Figure 4-6C, white arrows), consistent with the size of assembled capsids without a viral envelope. VLPs were rather pleomorphic and displayed only one discernible outer layer. Within the same cytoplasmic space, we observed two ER membrane budding events, suggesting that VLPs originated from the nearby ER (Figure 4-6C, red arrows). Single-slice images of the second ROI clearly resolved two additional VLPs with diameters of 30nm
and 35nm respectively (Figure 4-6D, white arrows). One of these particles was within ~13 nm of a membrane compartment showing an invagination reminiscent of a viral genome replication site. Adding resolution in the z-dimension allowed us to clarify that the particles were located at the bottom of the spherical LD. The characteristic additional electron density of the LD only became apparent in the later z-slices of the tomographic volume (Figure 4-6C, black arrows). In summary, our ET analysis captured 30-35nm unenveloped particles directly adjacent to LDs as well as potential membrane compartments involved in genome replication. Moreover, we captured budding events at ER cisternae, suggesting that particles may form at the ER.

**Immunoelectron microscopy confirms presence of HCV core and NS5A proteins within LD microenvironment**

Core and NS5A proteins play critical roles during HCV assembly. Core not only functions as a structural component of the virion, but also regulates the formation of assembly sites and recruitment of replication complexes to the LD. Similarly, the NS5A protein plays a critical role in bridging virion assembly and genome replication. To allocate these viral proteins to specific virus-induced structures within the LD microenvironment, we performed immunolabeling of low-temperature-embedded plastic sections. We infected cells with a reporter virus containing NS5A tagged with a modified yellow fluorescent protein, YPet. To guarantee that this reporter virus spread to nearly all cells in the sample, we allowed the infection to proceed for four additional days. At four days post-infection, we fixed cells briefly with formaldehyde to inactivate infectious HCV. This enabled safe handling during the subsequent high-pressure freezing, freeze-substitution, low-temperature embedding, and ultrathin sectioning steps. Indirect immunolabeling with a primary mouse anti-core antibody and secondary gold-conjugated anti-mouse antibodies confirmed the preservation of core epitope antigenicity. In addition, for the detection of YPet-tagged NS5A, we used an anti-GFP antibody. This antibody binds to an epitope that is conserved between GFP and YPet.
We found that both the core and NS5A antibody showed specific labeling on sections from infected cells. Overall, membranes were less preserved in our immuno-EM preparation protocol compared to the traditional plastic embedding protocol. Consequently, most of the structural features of membrane rearrangements were lost. Moreover, we were no longer able to detect VLPs in these samples, which suggested that these structures were not stable enough to allow preservation during the immuno-EM sample preparation. Despite these shortcomings in ultrastructural preservation, we detected the gold-label in clusters adjacent to structures resembling LDs when labeling with the core antibody (Figure 4-7A). Note that we detected nonspecific background labeling that was rather uniformly distributed in the rest of the cell. In contrast, the label was concentrated in clusters at distinct sites close to LDs. For the NS5A antibody, we found labels on structures resembling LDs (Figure 4-7B, black arrows). In addition, we observed label on both the exterior and interior of round structures made up of a vacant interior and an electron-dense halo (Figure 4-7B, red arrows). At lower magnifications these structures appeared to be part of larger networks reminiscent of the MW (Figure 4-7B, left-hand images). Thus, the NS5A-positive structures we observed could be DMVs, albeit with poorly-preserved membrane layers. If these structures indeed represented genome replication sites, they should contain viral RNA as well. However, despite several attempts, our efforts to label viral RNA metabolically for detection by immuno-EM were unsuccessful. In conclusion, both core and NS5A localized to LDs, with NS5A localizing to presumed DMVs as well. Unfortunately, the structural information gained from immuno-EM was very limited due to poor structural preservation. Therefore we were not able to gain any further insight into the ultrastructural details of core- and NS5A-positive structures.
Clusters of HCV core protein are close to LDs but show variability in size, shape, distribution patterns and E2 co-localization

Although the ultrastructural preservation of the immuno-EM experiments were suboptimal, the resolution of our images was sufficient to suggest that core protein formed clusters close to LDs. To gain better insight into the frequency and exact subcellular localization of these clusters, we used superresolution light microscopy to reveal the size of core-positive clusters. Images taken with a stimulated emission depletion (STED) microscope had a resolution of ~80nm. For these imaging studies of chemically-fixed, HCV-infected cells we used the primary antibody against the HCV core protein and a secondary antibody coupled to a STED-compatible dye, Atto 647N. In addition, we labeled LDs with the fluorescent lipophilic dye Bodipy 493/503. To examine co-localization of core clusters with other structural components of the virion, we also labeled samples with a primary monoclonal antibody against the E2 glycoprotein and a secondary antibody coupled to the fluorescent dye Alexa Fluor 555. Although the Bodipy 493/503 and Alexa Fluor 555 dyes were not compatible with STED microscopy, we would still be able to examine the subcellular localization of LDs and E2 protein using the regular confocal mode of the STED microscope. Three days after infection of Huh-7.5.1 cells with HCVcc, we fixed cells and performed immunofluorescence labeling against core and E2, followed by labeling of LDs with Bodipy 493/503.

HCV core punctae were well-resolved in images taken by STED microscopy (Figure 4-8A, middle images). To reduce the unavoidable blurring effects during image acquisition, we also processed these images with a digital deconvolution software (Figure 4-8A, right-most images). Additional multi-channel fluorescence microscopy then allowed us to overlay the fluorescent signals from the super-resolved anti-core channel with the signals from the confocal anti-E2 and LD channels (Figure 4-8B). Core punctae were close to the surface of LDs and overlapped with regions that were positive for the E2 glycoprotein. When surveying the entire cell, we noticed
two characteristic patterns of core subcellular localization. In the first pattern, core protein formed reticular networks that wrapped around LDs (Figure 4-9A, two left-most images), compatible with an ER association of core protein. In addition, the majority of these core-positive areas overlapped with E2-positive patches, indicated by the magenta areas in the red-and-blue overlay (Figure 4-9B, two left-most images). In the second pattern, core formed rather localized clusters close to the LD surface (Figure 4-9A, two right-most images). Interestingly, areas with high levels of core and clustering of multiple punctae close to LDs tended to overlap with areas showing high E2 levels (Figure 4-9B, two right-most images).

Next, we manually selected well-resolved core punctae and measured their diameters (Figure 4-9A, boxes #1-4; zoom-ins shown in Figure 4-9D). We selected a total of 300 intracellular punctae localizing close to LDs. Punctae displayed variability in both size and shape. Counted punctae had an average diameter of 130nm (±32 nm). We also observed larger clusters with diameters in a range of 150 to 220nm. These larger clusters also tended to be more asymmetric (Figure 4-9D, #1 and #2 zoom-in). In contrast, smaller clusters in the 80-100nm range tended to have more symmetric, round shapes (Figure 4-D, #3). Note that 80nm was also the resolution limit for this imaging technique. In some cases, core punctae overlapped with E2-positive regions (Figure 4-D, #1). In other cases, core punctae were adjacent to E2-positive regions (Figure 4-D, #2 and #4). We also found instances with no apparent co-localization of the core cluster with E2 (Figure 4-D, #3). In summary, our STED microscopy analysis of core protein localization confirmed the presence of core in both individual clusters as well as larger reticular networks around LDs. Some of these clusters also overlapped with areas of high E2 levels, consistent with these being areas of high viral protein production and potentially sites of viral assembly.
Development of a panel of infectious reporter viruses containing NS5A tagged with fluorescent proteins

In our immunolocalization experiments, we observed NS5A label on the surface of LDs as well as putative genome replication sites. To complement these findings and acquire additional images of NS5A subcellular localization, we developed a panel of reporter viruses containing fluorescently-tagged NS5A. This allowed us to analyze infected cells by fluorescence microscopy. Although NS5A is not a structural component of progeny virions, it plays crucial roles in mediating virus assembly. Moreover, it is a component of the viral replicase and therefore provides a marker for RNA genome replication sites as well.

The functional profile of NS5A domain III, presented in Chapter 3 of this thesis, indicated that the entire domain tolerated small, 5-aa insertions regardless of reading frame. This was consistent with previous work, which showed that insertion of green fluorescent protein (GFP) in domain III produced a replication-competent reporter virus in the replicon system (Moradpour, Evans et al. 2004). However, insertion in NS5A domain III of a Jc1 virus resulted in a ~50-fold drop in supernatant infectivity and reduction of viral spread (Schaller, Appel et al. 2007). Similarly, insertion of EGFP in domain III at a downstream site resulted in a significant defect in the assembly of intracellular infectious virions (Gottwein, Jensen et al. 2011). In the same study, adaptation to cell culture by sequential passaging rescued this defect by introducing a 40-amino-acid deletion in domain II (Gottwein, Jensen et al. 2011).

To develop comparable infectious reporter viruses with high peak infectivity titers and efficient viral spread in the context of our Jc1-based FNX-24 virus, we chose the position reported in Gottwein et al. for insertion of fluorescent proteins (Figure 4-10A). We also introduced the 40-amino-acid deletion into FNX-24 to rescue potential assembly defects. As a reporter protein we selected the yellow-fluorescent YPet protein due to its 2.4-fold higher brightness compared to EGFP (Nguyen and Daugherty 2005; Shaner, Steinbach et al. 2005). Second, we engineered a
red-fluorescent mCherry reporter virus to allow for imaging of NS5A produced from multiple viruses in the same sample. Using confocal microscopy filter sets specific for YPet and mCherry, it is possible to clearly distinguish fluorescent signals from YPet-labeled and mCherry-labeled reporter viruses. Third, we inserted the miniSOG fluorescent protein, which allows for correlative light and electron microscopy (CLEM) approaches (Shu, Lev-Ram et al. 2011). In CLEM experiments, the subcellular localization of a protein can be determined using the inherent fluorescence of the miniSOG genes, which can then be converted to an electron dense signal that becomes traceable at the ultrastructural level as well. To further increase the fluorescent signals of miniSOG, which has a modest green fluorescence compared to GFP (quantum yield of 0.37 versus 0.6, Shu, Lev-Ram et al. 2011) we also inserted a tandem version of the miniSOG gene, containing two copies of the fluorescent gene. Lastly, we included the OLLAS epitope tag at the N-terminus of each fluorescent protein, which allows for additional detection of the tagged protein using anti-OLLAS antibodies (Park, Cheong et al. 2008). In Western Blot experiments, the OLLAS tag offered 100-fold higher sensitivity compared to comparable anti-FLAG M2 and anti-V5 epitope tag detection (Park, Cheong et al. 2008).

To analyze infectious virus production in cells transfected with the different reporter genomes, we determined infectivity titers of cell culture supernatants harvested at various time points after transfection of tagged genomes. Growth curves of the FNX^{NS5A-YPet}, FNX^{NS5A-mCherry}, FNX^{NS5A-miniSOG} and FNX^{NS5A-tdminiSOG} viruses indicated moderate decreases in peak infectivity compared to the wild-type FNX-24 virus (Figure 4-10B). Nonetheless, the viruses still produced robust levels of extracellular virus with peak infectivities greater than 10^4 particles per ml. Next, to demonstrate the usefulness of these reporter viruses for tracking NS5A protein in infected cells, we visualized subcellular localization of tagged NS5A using fluorescence microscopy. We detected fluorescent signals in live cells as early as 24 hours following transfection of FNX^{NS5A-YPet} and FNX^{NS5A-mCherry} RNA genomes (data not shown). Fluorescent signals were strongest
from the 48 hour time point onwards, allowing reliable detection of NS5A in live cells (Figure 4-10C). For FNX\textsuperscript{NS5A-miniSOG} and FNX\textsuperscript{NS5A-tdminiSOG} reporter viruses, however, fluorescence was much weaker compared to the FNX\textsuperscript{NS5A-YPet} and FNX\textsuperscript{NS5A-mCherry} viruses and could only be detected starting at 72 hours (data not shown). In fact, fluorescence levels were barely above the background fluorescence (measured in control cells transfected with viral RNA of untagged wild-type HCV), especially for FNX\textsuperscript{NS5A-miniSOG}. This complicated imaging of live cells infected with the miniSOG-based reporter viruses.

To confirm that fluorescent proteins are indeed tracking NS5A localization, we performed immunofluorescence assays (IFA) using the 9E10 monoclonal antibody, which recognizes NS5A from both genotype 1b and 2a subtypes (Lindenbach, Evans et al. 2005). To allow us to carry out co-localization studies, we used secondary anti-mouse antibodies labeled with either Alexa Fluor 488 or Alexa Fluor 594. We fixed cells after three days of incubation with cell-culture produced FNX\textsuperscript{NS5A-YPet} and FNX\textsuperscript{NS5A-mCherry}, followed by IFA. To quantitate the amount of co-localization between the signals from the anti-NS5A antibody and reporter protein fluorescence, we determined Pearson’s correlation coefficients. The Pearson’s coefficient provides the degree of overlap between two images acquired with different fluorescence channels and ranges from 0 to 1 (Manders, Stap et al., 1992; Adler, Pagakis et al., 2008). In cells infected with FNX\textsuperscript{NS5A-YPet}, signals from the YPet fluorescent protein co-localized with signals from the Alexa Fluor 594 anti-NS5A label, producing a high Pearson’s coefficient of 0.98 (Figure 4-11A, first panel). Similarly, in cells infected with FNX\textsuperscript{NS5A-mCherry}, signals from the mCherry fluorescent protein co-localized with signals from the Alexa Fluor 488 anti-NS5A label at a high Pearson’s coefficient of 0.92 (Figure 4-11A, second panel). NS5A exhibited a perinuclear punctate pattern, which was similar to the localization patterns seen in cells infected with wild-type FNX24 (data not shown). This suggests that the fluorescent-protein insertions did not affect the overall localization of the NS5A protein. Note that we were not able to detect green fluorescence in parallel IFA experiments of
cells infected with $\text{FNX}^{\text{NS5A-miniSOG}}$ and $\text{FNX}^{\text{NS5A-tdminiSOG}}$, presumably due to the lower initial green fluorescence and a possible loss of fluorescence during the IFA protocol. However, immunostaining using the monoclonal anti-OLLAS antibody confirmed production of the OLLAS-tagged tdminiSOG reporter protein (Figure 4-11B). Furthermore, the reporter protein showed a characteristic perinuclear punctate localization, indicating that the OLLAS-tdminiSOG insertions did not alter the subcellular localization of NS5A during the viral life cycle.

Taking together the results from viral infectivity assays, live-cell fluorescence microscopy, and immunofluorescence assays, we conclude that HCV reporter viruses containing NS5A tagged with fluorescent proteins YPet, mCherry, miniSOG and tdminiSOG produce functional viral proteins and replicate efficiently. These reporter viruses exhibited robust levels of infectious virus production and allowed for live tracking of fluorescent NS5A protein in infected cells. However, insertion of reporter genes reduced infectivity titers by about 10-fold. Moreover, fluorescent signals from the miniSOG-NS5A and tdminiSOG-NS5A reporter proteins were rather faint, limiting their usefulness in live imaging studies.

**STED microscopy of $\text{FNX}^{\text{NS5A-YPet}}$-infected cells reveals presence of 110-150nm-sized NS5A clusters**

To obtain the subcellular localization of NS5A at high resolution, we combined STED microscopy with our fluorescent reporter tagging approach. We infected Huh-7.5.1 cells with the $\text{FNX}^{\text{NS5A-YPet}}$ virus and allowed the infection to progress for three days. Next, we chemically fixed the cells and carried out an immunostaining protocol using the monoclonal 9E10 antibody along with a secondary antibody conjugated to the STED-compatible Atto647N dye. This approach allowed us to acquire super-resolved STED images of the anti-NS5A signal along with regular confocal images of YPet fluorescence. In addition, we further processed images by applying digital deconvolution filters to remove image noise.
Fluorescent signals from the YPet reporter protein served as a “counterstain” reference for NS5A subcellular localization. Visualization of the anti-NS5A signal in the microscope’s STED mode resulted in a higher image resolution and revealed distinct, small, punctate signals that were not apparent in the images taken with the confocal microscopy set-up (Figure 4-13A, green arrows; middle image vs. left- and right-hand images). Next, we measured the sizes of these smaller (<200nm), NS5A-positive punctae (Figure 4-13A, white arrows pointing towards representative punctae). The majority of punctae were between 111-150nm in diameter (Figure 4-13B). NS5A was also present in larger clusters, which were equally well-resolved in confocal and STED images (Figure 4-13C). The majority of these larger clusters displayed diameters of 351-450nm. In addition to these smaller punctae and larger clusters, we also observed NS5A in reticular networks and occasionally forming ring-like structures (Figure 4-13A, ring-like structure marked with red arrows). In summary, STED microscopy proved useful for the detection of smaller, 115-150nm-sized NS5A punctae. However, NS5A was also prevalent in larger 351-450nm clusters. Moreover the resolution of confocal microscopy was sufficient to visualize these larger structures, especially when imaging YPet fluorescence.

**Infectious reporter viruses containing NS5A tagged with YPet facilitates analysis of NS5A, core and E2 localization around LDs**

Taking advantage of the ability to easily detect NS5A protein in cells infected with the FNX<sup>NS5A-YPet</sup> virus, we next examined the subcellular localization of the structural proteins core and E2 in relation to NS5A using multi-color fluorescence microscopy. For these experiments, we used a standard confocal microscopy set-up for all three fluorescent channels. In particular, we were interested whether NS5A present in larger (>300nm) clusters (which were well-resolved in confocal images) co-localized with core and E2. Because of the spectral overlap of the YPet fluorescent protein and Bodipy 493/503 dye, we used an alternative approach for the visualization of LDs in FNX<sup>NS5A-YPet</sup> cells. When imaging cells using the microscope’s
transmission filter setting, we noticed the presence of darker, round structures dispersed throughout the cytoplasm (Figure 4-12A, green arrows). To confirm that these structures indeed represented LDs, we performed an additional staining step using Bodipy 493/503. Next, we visualized green fluorescence in the corresponding field-of-view. Binary conversion of both transmission and green-fluorescence images indicated that the transmission mode produced additional signals apart from the cytoplasmic round structures. Nonetheless, we could still clearly identify these round structures in the midst of the image background (Figure 4-12A, red arrows). Importantly, the round structures overlapped with the signals from the Bodipy 493/503 staining (Figure 4-12B, yellow color), indicating that they indeed represented LDs. Thus, we used the microscope’s transmission mode to provide a reference of the LD organization in subsequent multi-color fluorescence microscopy experiments.

To detect core and E2 proteins alongside YPet fluorescence and LDs, we immunostained fixed cells with monoclonal anti-core and anti-E2 antibodies, followed by incubation with secondary Atto647N- and Alexa Fluor 555-conjugated antibodies. Sequential four-channel laser scanning then allowed us acquire images of YPet, E2, core and LD signals from the same sample. Note that LDs were visualized using the microscope’s transmission mode. Using this method, we could detect LDs as darker areas with a circular shape. Core protein (green color) localized to the surface of these LDs (Figure 4-13C and D) and could be found throughout the cytoplasm in a more reticular pattern (Figure 4-13C, left-hand image of “core” channel). Signals from the red E2 channel and blue YPet-NS5A channel co-localized, as indicated by the prevalent magenta color in the image overlay. Moreover, these proteins showed a reticular distribution pattern reminiscent of ER-localized proteins. Triple-positive areas appeared as white patches in the image overlay. Comparison with the transmission channel revealed that triple-positive protein clusters were located close to LDs (for an example, see Figure 4-13D left-most image). Some triple-positive clusters were located within a larger network of magenta E2- and
NS5A-positive patches (Figure 4-13D, middle image), suggesting that core protein formed distinct clusters within an E2- and NS5A-containing ER network. Next, we counted a range of diameters for triple-positive clusters, resulting in an average diameter of 594nm (±130nm, n=232, Figure 4-13D, graph). Taken together, our multi-color fluorescence microscopy experiments using the FNX^{NS5A-YPet} reporter virus demonstrated the accumulation of core, E2 and NS5A in ~600nm sized clusters close to LDs. These clusters showed high levels of these HCV protein and could therefore potentially represent sites of high viral protein translation or sites of viral assembly.

**Phenotypic characterization of a point mutant in NS5A domain IA using the FNX^{NS5A-YPet}-based fluorescent imaging system**

Light-microscopy techniques offer great flexibility in analyzing specimens under ambient conditions. This allows for the collection of real-time information in live cells. Due to the large field-of-view, light microscopy also allows for quick examination of the subcellular localization of large areas of the cell population. Finally, acquisition of 3-D (X, Y, Z) information becomes possible using confocal laser scanning microscopy. In a proof-of-concept experiment, we therefore used the FNX^{NS5A-YPet}-based fluorescent imaging system for an initial phenotypic characterization of a viral mutant originally discovered in our laboratory using a high-resolution genetics approach.

In addition to the 15-nt-insertion mutagenesis platform outlined in Chapter 3 of this thesis, our laboratory has developed a high-throughput profiling system that integrates saturation mutagenesis and next-generation sequencing (Qi 2012). Briefly, saturation mutations (leading to mutations to all possible amino acids at each position) were introduced into domain IA of NS5A (amino acids 18-103). Following selection of the saturated viral library in Huh-7.5.1 cells, we examined the effect of point mutations on viral fitness by determining changes in frequency of each mutant virus using next-generation sequencing. In resulting fitness profiles, we noted that
the isoleucine residue at position 52 (I52) of NS5A domain IA did not tolerate changes to non-hydrophobic amino acids. Moreover, substitution to an equally-sized, yet polar Asparagine residue (mutant "I52N") completely abrogated infectious virus production, while only moderately affecting genome replication.

To further investigate the mechanism of the infectivity defect mediated by I52N, we introduced this mutation into the FNX^{NS5A-YPet} genome and analyzed the subcellular localization of NS5A in cells transfected with I52N FNX^{NS5A-YPet} RNA genomes. Confocal microscopy analysis showed that NS5A was concentrated in cytoplasmic clusters (Figure 4-14A) and lacked the typical reticular staining pattern seen in FNX^{NS5A-YPet}-infected cells (for comparison, see Figure 4-13C). Interestingly, YPet-positive clusters co-localized with signals from anti-core and anti-E2 antibodies (Figure 4-14A, white color in overlay). Moreover, NS5A and E2-positive clusters were located close to LDs (stained with Bodipy 493/503) and in some cases even wrapped around LDs (Figure 4-14B). We confirmed this association with LDs in digitally deconvolved 3-D reconstructions of z-stacks acquired using scanning laser microscopy (supplemental video S6). This approach allowed for a more in-depth analysis of the spatial relationships between NS5A, E2 and LDs than that offered by 2-D confocal microscopy. We detected clusters containing both NS5A and E2 protein either adjacent to or coating individual LDs.

Overall, the total fluorescent intensity of the YPet signal per I52N-transfected cell showed a ~3-fold decrease compared to cells transfected with FNX^{NS5A-YPet} RNA (representative images shown in Figure 4-14C). In all of our fluorescence microscopy experiments, NS5A-positive clusters represented a characteristic phenotypic feature of cells transfected with I52N-FNX^{NS5A-YPet} RNA. To examine the dynamic nature of these NS5A-positive clusters, we collected real-time information in live, infected cells. In short-term (t=10min) epifluorescence imaging studies,
we did not detect large movements of NS5A-clusters in the examined time frame (Figure 4-14D). This suggested that NS5A clusters represented static structures.

The antiviral drug daclatasvir (BMS-790052) has been implied to target NS5A functions in domain 1A (Belda and Targett-Adams 2012), which is the same domain that the I52N mutation was located in. Thus, we wanted to ask the question whether the I52N mutation and daclatasvir treatment targeted a similar function mediated by NS5A domain IA. To do so, we examined the subcellular localization of the NS5A protein following treatment with daclatasvir. We infected Huh-7.5.1 cells with the fluorescent FNXNS5A-YPet reporter virus and incubated the cells for an additional three days to allow the infection to spread. Next, we added cell culture media containing either 200pM or 50pM daclatasvir resuspended in dimethyl sulfoxide (DMSO). Previous experiments indicated an EC$_{50}$ value of 50pM of daclatasvir in our HCVcc system (data not shown). Cell media only containing 1% DMSO served as a negative control. As early as 15h post-treatment with daclatasvir we could detect groups of NS5A-expressing cells displaying NS5A in cytoplasmic clusters (Figure 4-15A). This effect was more pronounced for cells treated with 200pM of daclatasvir and become even more apparent after 24h of drug treatment. Figure 4-15B illustrates the similarities between the subcellular localization of the I52N mutant and daclatasvir-treated cells. These results suggest that the I52N mutation and daclatasvir treatment lead to a similar phenotype with regard to redistribution of the NS5A protein in cytoplasmic clusters.

In summary, our imaging studies revealed that NS5A containing an I52N mutation retained the ability to localize in core-E2-NS5A-positive clusters. Moreover, these clusters were found associated with LDs as well, suggesting that the ability of NS5A to target to LDs was not compromised by the I52N mutation. Live-cell imaging studies showed that NS5A clusters remained static within a 10-min time frame. Finally, treatment with daclatasvir resulted in a
redistribution of NS5A protein similar to the one seen in cells transfected with I52N-FNXNS5A-Ypet RNA.

DISCUSSION

In the last seven years the HCV field has made significant strides towards gaining insight into the determinants of the HCV assembly process. The application of traditional electron microscopy and light microscopy approaches to the HCV infectious cell culture system has contributed greatly towards this endeavor. Nonetheless, where the assembly site is exactly located has so far remained an unsolved question.

Taking advantage of electron microscopy and electron tomography approaches, we show here that the LD microenvironment displays a juxtaposition of LDs, membrane vesicles (including double membrane vesicles) and ER cisternae, providing membranous platforms for both genome replication and viral protein translation. We also observe viral-like particles close to the LD surface, associated with membrane compartments and budding from the ER. At the light microscope level, superresolution microscopy reveals a subcellular distribution of core protein both in smaller clusters close to the LD surface as well as larger reticular networks overlapping with ER-localized E2 glycoprotein. Similarly, the NS5A protein forms smaller and larger clusters; we also identified areas in the LD microenvironment of an HCV-infected showing a co-localization of core, E2 and NS5A. We still detected triple-positive areas in proximity of LDs for a NS5A mutant deficient in infectious virus production. Finally this NS5A mutant displayed a characteristic localization of NS5A in large, static, cytoplasmic dots. Treatment with a promising inhibitor targeting NS5A appeared to phenocopy this observed NS5A localization.
Infected cells contain an interconnected microenvironment of LDs, membrane vesicles and ER

Our ultrastructural characterization of cultured cells infected with a highly assembly-competent HCV clone showed the area around LDs was interlaced with membrane vesicles as well as extensive and often fragmented ER stacks (both rough ER and smooth ER). First, our traditional thin-section EM analysis detected these ultrastructural membrane changes around readily identifiable LDs only in infected cells. Second, individual digital slices of 3-D electron tomography reconstructions allowed us to confirm the double-membrane nature of membrane vesicles, the presence of invagination events in single-membrane vesicles and the connection between DMVs and separate membrane compartments, likely to represent ER, through an electron-dense neck-like structure. Invagination events could represent a critical step in the biogenesis of DMVs and a mechanism for single-membrane vesicles to acquire a double membrane by inward budding, as proposed previously (Ferraris, Beaumont et al. 2012; Romero-Brey, Merz et al. 2012). Furthermore, our finding that DMVs are connected to adjacent membranes via a neck-like structure and thus appear as protrusions of the ER are consistent with independent studies using electron tomography to characterize the membranous web (Romero-Brey, Merz et al. 2012). The authors point out that these types of structures are surprisingly similar to the replication sites of distantly related picornaviruses, coronaviruses and arteriviruses (Romero-Brey, Merz et al. 2012). In contrast, related flaviviruses like dengue virus (DENV), West Nile virus (WNV) and tick-borne encephalitis virus (TENV) have been shown to form invaginations, not exvaginations, of ER membrane (Welsch, Miller et al. 2009; Gillespie, Hoenen et al. 2010; Miorin, Romero-Brey et al. 2013). Nonetheless, future studies detecting newly-synthesized RNA are needed to show whether these DMV-ER-exvaginations represent active replication sites.
Our results are also in agreement with previous studies that have shown the association of the MW with ER membranes and LDs (Gosert, Egger et al. 2003; Targett-Adams, Boulant et al. 2008; Wolk, Buchele et al. 2008). The first recognizable ultrastructural change following infection of a cell with HCV consists of membrane remodeling leading to an accumulation of membranous vesicles. These vesicles have been previously identified as the principal site of viral RNA replication in cultured cells harboring a subgenomic replicon from a 1b genotype virus (Gosert, Egger et al. 2003). Analogous studies in the infectious cell culture system (Ferraris, Beaumont et al. 2012) observed three types of membrane alterations: vesicles in clusters (ViCs), contiguous vesicles (CVs), and double-membrane vesicles (DMVs). Early in infection, Ferraris et al. noted the formation of a network of CVs surrounding LDs, whereas later stages of infection showed a large increase in the number of heterogenous DMVs (Ferraris, Beaumont et al. 2012). Similarly, Romero-Brey et al. observed that DMVs, residing primarily in an area surrounding ER cisternae early in infection and increasing both in number and diameter during the progression of infection (Romero-Brey, Merz et al. 2012). The MW itself contained cellular markers of rough ER along with markers of early and late endosomes, COP vesicles, mitochondria and LDs (Romero-Brey, Merz et al. 2012). Furthermore, an observed correlation between the kinetics of viral RNA amplification and DMV formation pointed towards the involvement of DMVs in RNA replication (Romero-Brey, Merz et al. 2012). Finally, Ferraris et al. used 3-D reconstructions of serial thin sections to show that DMVs formed a massive network closely associated with LD clusters. ViCs were contained in several small clusters, while numerous small CVs tended to surround LDs (Ferraris, Beaumont et al. 2012). Taken together, our results along with the findings from independent groups strongly suggest that the LD microenvironment contains the membranous platform facilitating RNA replication (through DMVs, CVs or ViCs), which is interconnected with a platform allowing for viral protein translation (through the rER membrane) and viral assembly (LD surfaces or ER derived membrane).
LD microenvironment contains virus-like particles of varying sizes and densities close to LD surface

Another outcome of our study is a visualization of spherical virus-like particles of varying sizes close to LD surfaces. Although we were able to detect virus-like particles at various locations in the cell, we focused our EM efforts on the LD microenvironment due to its known role in HCV assembly. We were not able to detect any budding events from the LD surface. Instead, detected particles were always spherical, but displayed a variable diameter and thickness of the particle coat. We were not able to detect specific viral structural features such as glycoprotein spikes, hexameric capsid assemblies and instead had to characterize particles as “virus-like” by their size and spherical nature. While some particles appeared to be free-standing in the cytoplasm, others appeared to make contacts with neighboring membrane compartments. Furthermore, we observed budding events from rER membranes located close to LDs. In contrast, we could not obtain images of virus-like particles budding into the ER or being contained in the ER lumen, which has been a prevalent model for how virions acquire the viral lipid envelope. Our inability to detect these specific events may be a consequence of an overall low assembly efficiency in the HCVcc system. Since electron microscopy can only survey a limited amount of cellular space (only 50-200nm slices at a time), a large number of events are required to be detectable. Judging by the amount of infectious virus released into the supernatant, HCV assembly efficiency on a per cell basis is very low (Bartenschlager, Penin et al. 2011).

Despite the proposed low assembly efficiency of the HCVcc system, we were able to observe budding events from rER membranes located close to LDs. This suggests a mechanism of particle formation similar to the one revealed by ET analysis of cells infected with a recombinant baculovirus expressing HCV structural system (Badia-Martinez, Peralta et al. 2012). Due to the self-assembling properties of core-E1-E2 proteins, HCV-like particles form in this experimental
system and allow the study of events ranging from formation of small buds to fission of particles at the ER site. Our snapshots of budding events in the infectious cell culture system agree with Badia-Martinez et al.’s proposed model that an increasing concentration of core-E1-E2 molecules initiates formation of relatively small sized-particles (30-60nm) from the ER membrane. Larger particles could originate from coalescing buds. Finally, the heterogeneity in size and density of our imaged intracellular virus-like particles mirrors what is seen in ultrastructural studies of purified extracellular particles. These studies report that particles displayed varying sizes within a range of 40-100nm (Gastaminza, Dryden et al. 2010; Catanese, Uryu et al. 2013). However, only particles with a specific size (~55-60nm) were infectious (Gastaminza et al. 2010). This suggests that infectious virus formation requires a defined particle architecture. This also means that many of the viral-like particles we imaged in the intracellular environment could represent non-infectious assembly intermediates.

Accumulation of HCV core protein on the LD surface is associated with the production of infectious virus in Huh-7.5.1 cells (Boulant, Montserret et al. 2006; Boulant, Targett-Adams et al. 2007; Miyanari, Atsuzawa et al. 2007). Furthermore, microscopy studies of cells containing HCV genomes have visualized HCV RNA accumulating around the surfaces of LDs (Miyanari, Atsuzawa et al. 2007). Less is known about the nature of viral particles close to the surfaces of LDs. Electron microscopy not only allows for an ultrastructural characterization of the membrane changes around LDs, but can also clearly identify viral particles due to its nanometer-scale image resolution. Previous snapshots of spherical virus-like particles with an average diameter of 50nm showed a close association with membranes in close proximity to LDs (Miyanari, Atsuzawa et al. 2007). In immuno-EM experiments, these particles reacted with anti-E2 and anti-core antibodies, suggesting that these particles represent true virions (Miyanari, Atsuzawa et al. 2007). Unfortunately, despite repeated attempts, experiments designed to acquire immunoelectron micrographs of viral particles labeled with either E2 or core antibodies were
unsuccessful. While we showed specific immunolabeling of cellular sections prepared by cryosubstitution using core antibodies, we could no longer detect viral particles in these samples, suggesting that these structures were not stable enough to allow preservation during the immuno-EM sample preparation.

The presence of cellular coated vesicles, which can display size ranges similar to HCV particles, is known to complicate the clear identification of virus particles in infected cells. Not only can hepatocytes be naturally engaged in numerous transport functions, but active virus-induced remodeling of intracellular membranes to create RNA replication sites could also lead to an up-regulation of vesicular transport. For example, we noted that small particles in the 50-nm range could indeed be detected close to Golgi compartments even in uninfected cells. Therefore, in order to avoid misinterpretation of coated vesicles as virus particles, we avoided areas with nearby Golgi complexes when screening for the presence of viral particles in the LD microenvironment. Also, we limited our search to particles with dense coats and interiors, as well as rather uniform size. By doing so, we hoped to avoid imaging cellular vesicles, which are known to differ from viral particles by a thinner and more readily distinguishable coat, a rather clear interior and variable diameters. Nonetheless, future immuno-EM studies will be needed to unequivocally locate core-E2-positive particles around LDs and determine the 3-D architecture of the LD microenvironment containing these particles.

**Core and NS5A proteins localize in clusters of varying sizes and co-localize with E2 in specific areas close to LDs**

While the subcellular localization of individual HCV proteins has been well-studied by confocal microscopy, the limit of optical resolution (~250nm) restricts the analysis of protein distribution to structures greater than this resolution limit, since objects that are smaller than ~250nm appear to be the same size as the resolution limit of the microscope. Furthermore, the microscope user would not be able to separate objects that are closer together than 250 nm.
Using stimulated emission depletion (STED) microscopy, a superresolution microscopy approach, we identify protein clusters of both core and NS5A protein with average diameter in a size range less than the resolution limit of conventional light microscopy. Core clusters with sizes <250nm have an average diameter of 130nm (±32 nm, n=300). For these immunofluorescence experiments, we note that the fluorescent STED-compatible dye is coupled to a secondary antibody with a size of ~15nm, which in turn is separated from the target protein by the bridging primary antibody. The length of this antibody tree thus adds an extra ~30nm to labeled structures. Currently we do not know whether these smaller core clusters represent assembled particles or clusters of accumulated core protein prior to assembly. The symmetric shape of some of these clusters, however, suggests that at least a subset of the detected clusters could indeed represent core protein assembled into core capsids, especially for punctae in the 80nm range, which also represents the detection limit of our superresolution approach. Co-localization analysis with the ER-resident E2 glycoprotein and LDs indicated that core clusters were part of a reticular network of E2 positive areas, consistent with ER retention of E2, along with interlaced LDs. Indeed, core clusters were mostly found on LD surfaces either as isolated clusters or larger reticular networks. We propose that core protein primarily exists in two different forms within the LD microenvironment: associated with E2-positive ER that wraps around LDs or associated with the LD surfaces in areas with little E2. Of note, LD-associated polarized caps of core protein were previously seen in both live imaging studies as well as fixed immunostained cells (Boulant, Targett-Adams et al. 2007; Boulant, Douglas et al. 2008; Counihan, Rawlinson et al. 2011). Future STED microscopy studies will be needed to visualize changes in the distribution pattern of the core protein in mutant viruses, which are blocked at particular stages of virus particle assembly.

Prior genetic and biochemical studies suggest that NS5A represents an important connecting element between the replication and assembly steps (Evans, Rice et al. 2004; Appel, Zayas et
al. 2008; Tellinghuisen, Foss et al. 2008). It plays a pivotal role in the recruitment of replication complexes around LDs, to proposed assembly sites (Miyanari, Atsuzawa et al. 2007). Our superresolution images of NS5A protein localization reveal the presence of smaller NS5A-positive clusters with average diameters of 136nm (±28 nm, n=150). Larger NS5A-positive clusters had an average diameter of 429nm (±74 nm, n=92). Based on our ultrastructural studies of HCV-infected cells, smaller clusters could represent membrane vesicles and therefore single virus-induced replication sites. Consistent with this hypothesis, our immuno-EM studies showed labeling of membrane-vesicle-like structures in the corresponding size range of 100-200nm. In previous studies, DMVs were shown to have an average diameter of ~125nm early in infection and 147nm later on in infection (Romero-Brey, Merz et al. 2012). While it is very promising that these smaller clusters, which could not be separated using conventional confocal microscopy, can now be visualized using superresolution microscopy, more imaging studies will be needed to examine the biogenesis and functional significance of smaller clusters and to confirm that they indeed represent individual membrane vesicles.

Since larger NS5A clusters were well-resolved in standard confocal images, we used multicolor fluorescence microscopy to examine the co-localization of NS5A with structural components of the viral particle, core and E2. In particular, we developed a protocol that allows for the simultaneous detection of core, NS5A, E2 and LDs. To our knowledge, this is the first report of a protocol providing fluorescently-stained samples for the acquisition of 4-channel confocal images. We were able to clearly identify core-E2-NS5A positive areas with an average diameter of 594nm (±130nm, n=232). Triple-positive areas varied in size from 305nm to 1140nm. Clusters were also contained in areas that were rich in LDs and in many cases, juxtaposed to the LD surface. When infectious virus production was blocked by a mutation in domain I of the NS5A protein, triple-positive areas could still be detected close to LDs. However, compared to the wild-type virus, all three proteins displayed an excessive accumulation in a
limited number of these triple-positive clusters for the NS5A I52N mutant virus. Average
diameters of 670nm (±158nm, n=50) were comparable with the average diameters seen for the
wild-type virus. Overall, these results indicate that the co-localization of core, E2, and NS5A
close to LDs is necessary but not sufficient for infectious virus production. Bukh and coworkers
have previously analyzed core/NS5A protein co-localization for various JFH-1-based
recombinants and demonstrated condensed signals of NS5A and core at 12h post-infection
(Galli, Scheel et al. 2013). Furthermore core protein almost perfectly overlapped with the LD
signal (Galli, Scheel et al. 2013). In contrast, signals for both core and NS5A showed a more
diffuse distribution at later time points, compatible with ER association (Galli, Scheel et al.
2013). The authors speculate that in the early stage of the viral life cycle, new particle assembly
of recombinant viruses such as J6/JFH-1 is restricted to the LD surface; thus, LDs might serve
as initial storage areas for the accumulation of core before the actual assembly of particles is
initiated (Galli, Scheel et al. 2013). At later time points, virion assembly may be transferred to
the surface of the ER (Galli, Scheel et al. 2013). In light of these findings, we propose that the
I52N mutation blocks viral assembly at a step following the initial accumulation of core, NS5A
and E2 at the LD surface, presumably preventing transfer of the viral assembly site to ER
membranes.

**Fluorescent reporter viruses facilitate dynamic tracking of NS5A for mutant viruses and
during treatment with antiviral compounds**

Even though our images taken by EM, ET, superresolution light microscopy, and multi-color
light microscopy provide high-resolution snapshots of the LD microenvironment in fixed cells,
they do not provide any information on the dynamics of viral protein localization within this
specialized compartment. Therefore, we developed a panel of reporter viruses containing NS5A
tagged with a variety of fluorescent proteins. Of these, the FNXNS5A-YPet and FNXNS5A-mCherry
reporter constructs produced the brightest signals in live cells and released high levels of
infectious virus (Figure 4-11). FNX^{NS5A-miniSOG} and FNX^{NS5A-tdminiSOG} produced high levels of infectious viral particles. Fluorescent signals in live cells, however, tended to be much weaker and difficult detect. Nonetheless these constructs still hold great promise for correlated light and electron microscopy. Future studies should address whether the levels of miniSOG-tagged NS5A reach sufficient levels to mediate the formation of electron-dense deposit that can be imaged by electron microscopy, as has been demonstrated recently with miniSOG-tagged SynCAM1 in cultured cortical neurons (Shu, Lev-Ram et al. 2011).

By analyzing the I52N NS5A mutant virus using the aforementioned FNX^{NS5A-YPet} imaging platform, we confirmed that cytoplasmic NS5A-positive clusters were static over a period of ten minutes. This proof-of-concept experiment illustrates the usefulness of the fluorescent imaging platform for a phenotypic characterization of viral mutants that doesn’t require fixation of cells and additional immunolabeling approaches. The static nature of the NS5A-containing clusters is consistent with our hypothesis that the I52N mutation causes a block in viral assembly and accumulation at initial viral protein storage sites. In the future, more detailed mobility analyses using extended periods of examination as well as quantitative assessments by methods such as fluorescence recovery after bleaching (FRAP) can shed further light on the biogenesis of these static clusters. While we can only speculate on how the I52N mutation causes the observed block in assembly, our finding that treatment with the antiviral compound daclatasvir causes a similar accumulation of NS5A in cytoplasmic dots (Figure 4-15) suggests that both the I52N mutation and daclatasvir treatment might be impairing a similar function of NS5A. A unified picture of the working mechanism of NS5A-targeting molecules, such as daclatasvir, is still missing, although there is an agreement that they perturb domain I-related functions of NS5A (Lemm, O’Boyle et al. 2010; Targett-Adams, Graham et al. 2011; Belda and Targett-Adams 2012). Previously, our laboratory performed selection of a saturation mutagenesis library within domain I of NS5A and, using next-generation sequencing, detected a panel of residues that
influenced the drug sensitivity of domain I mutants (Qi 2012). Along with independent studies mapping daclatasvir resistance mutations to the N-terminal domain I (Fridell, Qiu et al. 2010; Gao, Nettles et al. 2010; Lemm, O’Boyle et al. 2010; Wang, Huang et al. 2012), these findings provide genetic evidence linking the antiviral activity of daclatasvir to the same domain in which the I52N mutation is located. Moreover, a redistribution of intracellular NS5A in response to NS5A inhibitors has been demonstrated previously (reviewed in Belda and Targett-Adams 2012). In particular, NS5A is redistributed from the ER to the surface of LDs in replicon-containing cells treated with NS5A inhibitors (Targett-Adams, Graham et al. 2011). In a system expressing viral proteins in the absence of RNA replication, NS5A also exhibited an altered subcellular localization following daclatasvir treatment (Lee, Ma et al. 2011). In this system, the NS5A protein showed a more diffuse localization pattern after daclatasvir treatment (Lee, Ma et al. 2011). Current models propose that the binding of daclatasvir could render NS5A incapable of incorporation in newly forming replication complexes by locking or altering the conformation of the protein (Belda and Targett-Adams 2012). No longer sequestered to the ER, NS5A may then be free to redistribute to the ER. Of note, recent mathematical models built using in vivo viral kinetics data predicted that daclatasvir blocks both intracellular viral HCV synthesis and virion assembly/secretion in vivo (Guedj, Dahari et al. 2013). Therefore, we propose that both the daclatasvir and I52N mutation may be inhibiting this assembly-related function of NS5A. However, future assays that will allow researchers to probe defined NS5A functions will be necessary to further refine our hypotheses of the working mechanism of both the I52N mutant and NS5A inhibitors.

In conclusion, the results presented in this chapter increase our understanding of the ultrastructure of the LD microenvironment and the distribution of viral proteins associated with it. Electron microscopy, electron tomography, superresolution light microscopy and multi-color fluorescence microscopy provided snapshots that illustrated a juxtaposition of LDs, VLPs,
membrane vesicles, core-positive clusters and triple-positive core-E2-NS5A. Therefore, these snapshots contribute to a growing body of evidence pointing to the importance of the LD microenvironment in the HCV life cycle. Superresolution microscopy and multi-color fluorescence microscopy procedures (allowing simultaneous detection of E2, core and NS5A) expand the number of approaches available to probe the viral life cycle. Finally, we show the promise of combining high-resolution viral genetics (which lead to the identification of the I52N mutation by a high-throughput functional screen) and a fluorescent HCV reporter imaging system for the identification of essential functions related to assembly.

METHODS

Cell Culture and Viral Infection

Huh-7.5.1 were cultured in Dulbecco’s Modified Eagle Medium (DMEM, Life Technologies, Carlsbad, CA) supplemented with 10% of fetal bovine serum (FBS), 10mM non-essential amino acids (Life Technologies), 10mM HEPES, penicillin (100 units/ml), streptomycin (100 mg/ml), and 2mM L-glutamine at 37°C with 5% CO2. After plating cells at about 25% confluence and following an overnight incubation, cells were infected with cell-culture-produced HCV (HCVcc) at 37°C at an MOI of 0.1. Media was changed the following day and cells were allowed to grow for another two days to allow the virus to spread through the culture. We used a cell-culture adapted version of the FNX24 virus, which we synthesized based on the chimeric sequence of the J6/JFH1 virus, also referred to a Jc1. We introduced 7 nucleotide mutations, resulting in synonymous mutations to the genome. The construct and sequence are available upon request.

Electron Microscopy and Electron Tomography

HCV-infected cells were processed for electron microscopy following a traditional plastic embedding approach outlined in Peng et al. (Peng, Ryantzantsev et al., 2010). Briefly, cells
cultured on 10 cm plates were washed with PBS and fixed for 1 hour in 2% glutaraldehyde in 1x PBS (pH 7.4) on ice. Cells were scraped, pelleted and subjected to osmium post-fixation (1% OsO4 in PBS) for 1 hour (on ice) and 2% uranyl acetate en bloc staining for 1h (on ice). Next, samples were dehydrated in an ascending ethanol series, infiltrated and embedded in Spurr’s resin, and finally sectioned for ultrathin-section-microscopy using an UCT ultratome (Leica Microsystems, Mannheim, Germany) with glass knives. Glass knives were prepared using a KMR2 knifemaker (Leica). Resulting sections had a thickness in the range of 50-80nm. Sections used for electron tomography had a thickness of about 200 nm. Ultrathin sections were collected on grids coated with a Formvar membrane, whereas thicker 200nm sections were collected on naked grids (100 mesh, copper) and stained with saturated aqueous uranyl acetate and lead citrate from both sides.

For thin-section-microscopy, cell sections were visualized using either a JEOL JEM1200EX or FEI CM120. TEM imaging and ET on thick sections were performed using an FEI Tecnai F20 electron microscope operated at 200 kV. Tilt series were recorded with a defocuses around 2 μm on a TVIPS F415MP 16 megapixel CCD camera at a magnification of 40,600x by tilting the specimen from −70° to 70° with 1° interval using the FEI tomography software Batchtomography.

Data Processing

Data processing was performed on Linux servers with dual Intel Xeon 5355 (Quad-Core) 2.66 GHz processors and 16 GB memory Images in each tomography tilt series were aligned and combined to generate 3D tomograms using eTomo (IMOD, Boulder, CO). Tilt series were reconstructed using eTomo in IMOD package (Kremer, Mastronardo et al., 1996). We used both the fiducialless alignment and patch tracking functions of the software, with patches set to 400x400 pixel size. Using the median function of the clip IMOD program, we filtered tomograms, thereby reducing image noise. To generate videos of tilt series and tomograms, we used
3Dmod (IMOD, Boulder, CO). To combine images into a video, we used VideoMach (Version 5.4.8, Gromada).

**Immunolabeling of resin-embedded sections**

At 3 days post-infection, FNX-YPET-infected cells were briefly fixed by adding 16% Formaldehyde straight to cell culture medium to give a final concentration of 2% Formaldehyde. This also ensured inactivation of infectious virus and facilitated further handling of the samples. We collected cells by scraping and centrifugation. We then resuspended cells in a low volume of resuspension medium (20% FBS, 20% BSA, 20mM HEPES in DMEM, 0.15M Sucrose). Membrane carriers filled with cell paste were placed into the loader of a Leica EMPACT2 high-pressure freezer (Leica). Samples were frozen according to manufacturer’s instructions. Frozen membrane carriers were stored in liquid nitrogen until further processing. Freeze substitution was carried out in Acetone containing 0.4% Glutaraldehyde for 58h at -90°C, followed by a warm-up phase to -25°C at 5°C/h for 13h in a Leica EM AFS2 freeze-substitution system (Leica Microsystems, Mannheim, Germany). At -25°C cells were rinsed with 100% Acetone, followed by sequential infiltration with 30%, 50%, 70% and 100% Lowicryl HM20 in Acetone for 4-5h each. Polymerization of HM20 resin was carried out at -25°C under UV lamp for 50h. Samples were further polymerized at room temp 25°C for 3 days. Ultrathin sections were cut and placed on Formvar carbon-coated nickel EM grids. We blocked these section by incubating a blocking solution consisting of 0.8% Bovine serum albumin (BSA), 0.1% fish gelatin and 5% goat serum. The monoclonal antibody against the HCV core protein (C7-50) and the polyclonal antibody to the GFP protein (Abcam, Cambridge, MA) were then added, diluted 1:100 and 1:2000 respectively in blocking buffer. We performed primary antibody incubation overnight at 4°C. To provide negative controls, we also incubated sections with blocking buffer only. As an additional negative control, we labeled with an antibody against an influenza viral protein, which is not expressed in our HCV-infected samples. The following day, we washed sections with blocking
buffer and incubated with the appropriate secondary antibody conjugated to either 10-nm or 15-nm gold particles (Ted Pella, Redding, CA) diluted according to manufacturer’s instructions. To increase contrast, we post-stained sections with 5% uranyl acetate and 5% lead citrate and observed EM grids with a JEOL JEM1200EX microscope.

**Fluorescence microscopy**

Huh-7.5.1 cells were grown either in 4-well chamber slides (Millicell EZ slide, EMD Millipore, Billerica, MA) or on 9x9 coverslips placed in 24-well plates. Following infection with either FNX24 or FNX^{NS5A-YPet} HCVcc for three days, cells were fixed with 4% paraformaldehyde for 5 min at room temperature followed by incubation at 4°C for three hours. After three washes with PBS, the cells were permeabilized with 0.05% Triton-x 100 for 10min and with blocked with 10% FBS, 3% BSA. All incubations with primary antibodies anti-core C7-50 (Abcam, Cambridge, MA), anti-E2 CBH5 or anti-NS5A 9E10 were carried out sequentially overnight or for a minimum of 5 hours. Secondary antibodies were conjugated to Alexa 555 and Atto 647N respectively (anti-mouse, anti-human). For studies examining lipid droplets, we also stained with Bodipy 493/503. Lastly, slides were mounted in Prolong Gold Antifade reagent. Slides were analyzed by fluorescence microscopy using a Leica TCS SP5 STED (Leica Microsystems, Mannheim, Germany) inverted confocal laser scanning microscope equipped with a 100x/1.4NA oil objective. For acquisition, we used the LAS AF software (Leica Microsystems). We acquired images for each channel sequentially. For STED imaging, we acquired diffraction-limited YPet, Bodipy 493/503, or Alexa 555 signals prior to Atto 647N signals. These acquisitions were done in the conventional confocal microscope mode (with STED laser off). In contrast, the Atto 647N signals (anti-core or anti-NS5A) were acquired in STED or conventional confocal microscopy modes. Pixel size during image acquisition was set to 30nm. Images were noise-filtered (deconvolved) with the LAS AF software using a Lorentzian PSF of 70nm FWHM in XY dimensions and the signal energy filter. For publication purposes image brightness and contrast
were adjusted. For acquisition of Z-stacks, we used a Nikon Eclipse Ti imaging system. 3-D renderings of acquired Z-stack series were generated using the microscope's Nikon Elements software.

**Live imaging acquisition**

Huh-7.5.1 cells were infected with the $\text{FNX}^{\text{NS5A-YPet}}, \text{FNX}^{\text{NS5A-mCherry}}, \text{FNX}^{\text{NS5A-miniSOG}}$, $\text{FNX}^{\text{NS5A-tdminiSOG}}$ reporter viruses. For live-cell imaging under drug treatment, we added the inhibitor Daclatasvir (BMS790052) to the cell culture media at a working concentration of 50-200pM. Daclatasvir (Selleck Chemicals, Houston, Tx) was dissolved in anhydrous dimethyl sulfoxide (DMSO) before further dilution in culture media to the appropriate working concentration. Cells treated with DMSO only (1%) served as a negative treatment control. For analysis of the I52N mutant, we introduced in-vitro-transcribed viral RNA, containing the I52N mutation in a $\text{FNX}^{\text{NS5A-YPet}}$ background, into Huh-7.5.1 cells by electroporation. At the appropriate time post-infection/post-transfection, cells were transferred in phosphate-buffered saline containing 20mM HEPES for live cell imaging using a Leica DMIRB epifluorescence microscope. YPet and miniSOG fluorescence images were taken using a standard FITC filter set, while mCherry was captured using a Texas red filter set. To avoid cellular toxicity effects in the uncontrolled environment of our microscope set-up, we limited imaging to short sessions, after which cells were transferred back into complete DMEM-based media and returned to 37°C with 5% CO$_2$. 

119
Figure 4-1 | Electron microscopy analysis of ultrastructural changes close to lipid droplets (LDs) in Huh-7.5.1 cells infected with HCV

Figure 4-1. (A-E) 2-D electron micrographs of ultrathin (50-75nm) plastic sections of embedded Huh-7.5.1 cells fixed after three days of infection with a cell-culture adapted version of a Jc1 virus. Oval dashed area (A) indicates membrane alterations reminiscent of the membranous web. Arrows in each panel emphasize the double- and multimembrane nature of vesicle compartments (B). Arrows also point towards non-linear ER cisternae around lipid droplets (C). ER, endoplasmic reticulum; LD, lipid droplet; Ex, extracellular space.
Figure 4-2 | Electron microscopy analysis of ultrastructural changes within LD-rich region in Huh-7.5.1 cells infected with HCV

Figure 4-2. (A-E) 2-D electron micrographs of a ~200nm thick plastic section of embedded Huh-7.5.1 cells fixed after three days of infection with a cell-culture adapted version of a Jc1 virus. (A) Low magnification (12,000x) view of infected Huh-7.5.1 cell. Black box marks region shown in higher-magnification micrographs. (B) View of a LD-rich region at 40,000x. Black boxes indicate regions close to LDs that are shown in the higher-magnification zoom-ins in the right-hand panels. (C-E) Zoom-in views of LD boundaries at 80,000x magnification. Note the presence of overlapped structural features, especially in C.
Electron tomography of infected cell provides three-dimensional snapshots of membrane structures close to LDs

**Figure 4-3**. (A-C) Gallery displaying digital 1-nm-thick slices from different positions (numbered) in the z-dimension of reconstructed tomographic volumes. Images capture the structural features of membrane structures close to LDs. (A) Area imaged by electron tomography corresponds to area shown in the 2-D electron micrograph of Figure 4-3C. Blue arrows point towards double-membranes, while red arrows indicate an invagination event at a membrane compartment. (B) Area visualized by electron tomography corresponds to area shown in Figure 4-3D. Blue arrows mark double-membranes. White arrows denote a rER cisterna. Black arrow draws attention to single-membrane vesicles inside a DMV. (C) Z-series view of area previously shown in Figure 4-3E. Blue arrows emphasize double-membrane nature of vesicles, while red arrows mark invagination events. White arrows point towards filament structures. Scale bar: 50nm.
Figure 4-4 | Electron microscopy analysis captures snapshots of virus-like particles in HCV-infected cells

(A-E) 2-D electron micrographs of ultrathin (50-75nm) plastic sections of embedded Huh-7.5.1 cells fixed after three days of infection with a cell-culture adapted version of a Jc1 virus. Black boxes indicate regions shown in adjacent zoomed-in images. Arrows: extracellular particle (A), particle within intracellular compartment (B), multiple particles within an intracellular compartment (C), cytoplasmic particles (D). Diameters of these particles are indicated as well. Ex, extracellular space.
Figure 4-5 | Electron microscopy provides snapshots of VLPs close to LD boundary

Figure 4-5. (A-E) Electron micrographs depicting VLPs close to LD interface, as well as ultrastructural changes within LD microenvironment. Diameters of imaged particles are provided as well. Black arrows highlight VLPs, some of which appear to be associated with adjacent membrane compartments (red arrows). LD, lipid droplet. Note that even in this ultrathin sections, overlapping structural features can complicated the interpretation of the spatial relationships of VLPs relative to the LD microenvironment.
Figure 4-6 | Electron tomography of infected cell delivers three-dimensional views of VLPs close to LDs

Figure 4-6. (A-B) 2-D electron microscopy depicting complex, overlapping structures close to LDs. (C-D) Gallery displaying digital 1-nm-thick slices from different positions (numbered) in the z-dimension of reconstructed tomographic volumes. A colored surface rendering of a slab taken from the 3-D volume is provided for additional orientation. Green, lipid droplet. Purple, particles. Images capture both particle morphology and size (white arrows) as well as budding events (red arrows). (D) Analysis of multiple sections reveals the presence of an LD in later sections (black arrows), situated on top of two particles (white arrows). Particles are also close to a membrane compartment displaying an invagination (red arrow).
Figure 4-7. EM analysis of immunolabeled Lowicryl sections prepared from high-pressure frozen, freeze-substituted, low-temperature-embedded cells. Cells were infected with a reporter virus containing NS5A tagged with yellow fluorescent protein (YFP). (A) Immunogold labeling with an anti-core antibody revealed clusters of label close to the LD surface. In uninfected cells, background labeling was observed as single gold particles dispersed throughout the cytoplasm, but never in clusters. (B) Immunogold labeling with an anti-GFP antibody decorated both LDs (black arrows, zoom-in inset) and round structures reminiscent of membrane vesicles (red arrows). We did not detect any immunolabeling in uninfected control cells.
Figure 4-8 | Superresolution light microscopy analysis of HCV core protein in infected cells

Figure 4-8. (A) Comparison of images obtained by confocal microscopy and STED microscope equipped with a 100x objective. We infected Huh-7.5.1 cells with a cell-culture-adapted Jc1 virus, fixed after three days of incubation and performed IFA using primary antibodies against core protein and secondary antibodies conjugated to the STED compatible dye Atto 647N. Digital deconvolution with a signal energy filter further removed image noise (right-most images). The outer boundary of the cell cytoplasm as well as the nucleus were marked with solid and dashed white lines, respectively. Regions boxed in red in (A) are shown in the magnified view of (B). (C) Multichannel fluorescent microscopy of an HCV-infected cell. Cells were prepared as in (A). In addition, we labeled LDs with the lipophilic dye Bodipy 493/503 and E2 glycoprotein using the anti-E2 antibody CBH-5, followed by a secondary antibody conjugated to Alexa Fluor 555. Panel shows black-and-white images from each fluorescent channel. The anti-core image represents a deconvolved STED microscopy image. In contrast, images from LD and E2 channel were acquired with the microscope’s confocal model. (D) Co-localization of core, LD and E2 in an HCV-infected cell. Left hand overlay image provides a low-magnification view of the cell with each fluorescent channel pseudo-colored as indicated (core, red; LD, green; E2, blue). The boxed region is shown in left-hand images. To better illustrate co-localization, only two fluorescent channels are shown at a time. Scales and imaged regions are identical for A and C, B and D.
Figure 4-9 | Examination of core-positive clusters in the vicinity of LDs by superresolution light microscopy

Figure 4-9. (A) High magnification views of super-resolved core protein clusters around LDs. Immunofluorescence and STED microscopy were carried out as in Figure 4-8. Images were acquired with a 100x objective. After applying a digital zoom, images were median filtered to remove background noise. Brightness and contrast were adjusted as well. Numbered white squares show core clusters at low (C) and high (D) magnifications. (D) Graph showing size distribution of core clusters appearing close to LDs. We analyzed 300 clusters from the cell shown in Figure 4-8.
Figure 4-10 | Fluorescent reporter viruses facilitate tracking of NS5A protein in live HCV-infected cells

(A) Schematic of reporter virus genomes encoding fluorescent proteins. We inserted indicated fluorescent proteins (YPET, mCherry, tandem miniSOG and miniSOG) in domain III of the NS5A protein of a genotype 2A virus, FNX24. Moreover, FNX24 contained a 40-aa acid deletion in NS5A domain II for the YPET, mCherry and tandem miniSOG insertions. The smaller miniSOG gene was inserted within a FNX24 background containing the full NS5A protein. Moreover, reporter genes contained short, flexible linkers at their N- and C-termini as well as the OLLAS epitope tag at the N-terminus, which allows for highly sensitive detection of tagged NS5A using the anti-OLLAS monoclonal antibody (pink boxes). Colors of fluorescent proteins reflect the respective emission wavelengths of the chromophore. (B) Fluorescent reporter genomes produce robust levels of infectious virus following transfection of RNA genomes. We transfected Huh-7.5.1 cells with RNA transcripts of the indicated reporter viruses and a wild-type-virus control. We measured infectivity titers in culture supernatants taken from transfected cells at indicated time points following transfection. FFU: Focus-forming units. (C) Epifluorescence microscopy
visualizes expression of YPET and mCherry fluorescent reporter proteins in live cells. Following transfection of indicated reporter genomes, we imaged live cells using fluorescence microscopy. As early as two days after transfection, we detected cells exhibiting strong green fluorescence as well as strong red fluorescence for the NS5A-YPET and NS5A-mCherry genomes, respectively.
**Figure 4-11** | Fluorescent reporter proteins co-localize with viral NS5A in infected cells

(A) Immunofluorescence staining confirms co-localization of YPet and mCherry reporter proteins and NS5A. We infected Huh-7.5.1 cells with the indicated reporter viruses. At three days post-infection, cells were fixed and processed for immunofluorescence assays. After probing for NS5A using the anti-NS5A antibody 9E10 and either Alexa Fluor 488- or Alexa Fluor 594-conjugated secondary antibody (for mCherry and YPET viruses respectively), we visualized fluorescent signals using confocal microscopy. For each sample, we collected fluorescent signals with filter sets optimized for the respective fluorochrome (Alexa Fluor 488, Alexa Fluor 594, YFP, mCherry). Co-localization analyses were performed on the representative images shown in this figure. Pearson’s coefficients are indicated as a measure of the rate of association of the two fluorochromes, with 1 standing for complete positive correlation and 0 for no correlation. We also plotted the pixel gray values of the Anti-NS5A/YPET and mCherry/Anti-NS5A images against each other in a scatter plot. These cytofluorograms use the intensity of a given pixel in the green image as the x-coordinate of the scatter plot and the intensity of the corresponding pixel in the red image as the y-coordinate. (B) Immunofluorescence staining of OLLAS-tagged protein in cell infected with FNX<sup>NS5A-td miniSOG</sup>. At three days post-infection, cells were fixed and processed for immunofluorescence assays. After probing for the OLLAS epitope using an anti-OLLAS antibody and an Alexa Fluor 594-conjugated secondary antibody, we visualized fluorescent signals using confocal microscopy. We also collected transmission images to visualize cells in a bright-field mode.
Figure 4-12. NS5A subcellular localization visualized by STED microscopy. Cells infected with FNX^{NS5A-YPet} were chemically fixed, then immunostained using monoclonal antibodies against NS5A. (A) We acquired images using the microscope’s STED setup (middle images). As a comparison, we also acquired anti-NS5A signals in the standard confocal mode (left-hand images). The confocal microscope mode was also used to record the NS5A-Ypet signal (right-hand images). Images of represent the same field of view, with the lower panel representing zoomed-in views of the boxed regions in the upper panel. To further remove noise, all images were digitally deconvolved. (B) Size distribution of smaller NS5A-positive punctae. Left-hand images show representative punctae. We manually measured the diameters of ~150 punctae. (C) Size distribution of larger NS5A-positive clusters. Left-hand images show representative clusters. We manually measured the diameters of 92 clusters.
Figure 4-13. (A) Confocal microscopy of fixed Huh-7.5.1 cells, grown on chamber slides. We stained for LDs using Bodipy 493/503. Transmission images were acquired using the microscope’s transmission filter setting. We created binary (black and white) images using ImageJ software. (B) To examine co-localization, binary images were pseudo-colored (red: LDs; green: transmission signals) and overlayed. Yellow color indicates areas of co-localization. (C) Multi-color fluorescence microscopy of infected cells. We fixed cells infected with FNXNS5A-YPet at 72h post-infection. We then performed immunostaining using anti-core and anti-E2, followed by secondary staining using anti-human-IgG-Alexa555 (red, E2) and anti-mouse-IgG-Atto647N (green, core). YPet (blue) and transmission signals were acquired using the microscope’s YFP and transmission filter settings. Images show a four-channel sequential laser scan of one optical confocal section through a representative cell. Images were captured at 30nm pixel size with a Leica TCS SP5 microscope and a 100x oil objective. For publication purposes, images were median-filtered and contrast-enhanced. Binary image of transmission signal provides a reference of LD organization. (D) Size distribution of YPet, core and E2 triple-positive clusters. Areas of co-localization appear as white patches in the image overlays of fluorescent signals (left-hand images). We manually measured diameters of ~200 triple-positive clusters.
Figure 4-14. (A) Multi-color fluorescence microscopy of cells transfected with I52N mutant RNA. We fixed cells transfected with I52N-FNX^{NS5A-YPet} RNA genomes at 72h post-transfection. We then performed immunostaining as described in Figure 4-13C. Note the localization of triple-positive core-E2-NS5A clusters close to LDs (dark spots).

(B) Multi-color fluorescence microscopy of cells transfected with I52N mutant RNA. We fixed cells transfected with I52N-FNX RNA genomes at 72h post-transfection. For these experiments, we labeled LDs with Bodipy 493/503 (green) and immunostained for E2 (red) and NS5A (blue). Note that E2 and NS5A co-localize (magenta color in overlay) in cytoplasmic dots that remain in the proximity of LDs. (C) Confocal microscopy showing NS5A-YPet fluorescence. Cells were transfected either with I52N-FNX^{NS5A-YPet} or FNX^{NS5A-YPet} RNA genomes. Cells were fixed and processed for confocal imaging. Images were taken using the microscope’s YFP filter set. The NS5A-YPet signal is pseudo-colored with a spectrum bar to reveal low-intensity signals (colored like the lower end of the spectrum bar) along with high-intensity signals (colored like upper end of the spectrum bar). (D) Live-cell imaging of a cell transfected with I52N-FNX^{NS5A-YPet} RNA genomes. At four days post-transfection, cells were transferred in phosphate-buffered saline containing 20mM HEPES for live cell imaging using an epifluorescence microscope equipped with a standard FITC filter set. YPet fluorescence was captured at 1-min intervals. Nuc: Nucleus, indicated by dashed red line for reference purposes.
Figure 4-15 | Live-cell images of Daclatasvir-treated cells showing redistribution of NS5A in punctate pattern similar to I52N phenotype

Figure 4-14. (A) Live-cell imaging of FNX\textsuperscript{NS5A-YPet}-infected cells following drug treatment with daclatasvir. Cells were infected with FNX\textsuperscript{NS5A-YPet} and infection was allowed to spread to >90% of cells in the following three days. At that point, cell culture media was replaced with fresh media containing daclatasvir at the indicated concentrations (200pM, 50pM). At indicated time points (15h, 24h or 48h post-treatment) sealed plates containing live cells were transferred in PBS supplemented with 20mM HEPES to an epifluorescence microscope equipped with a 20x objective. Insets represent zoom-ins of the boxed regions. Note that the 15h and 24h time points show images taken on the same sample, but not the same viewing area. DMSO: Dimethyl sulfoxide, served as the negative control at 1%, since daclatasvir was resuspended in DMSO. (B) Side-by-side comparison of i) I52N-FNX\textsuperscript{NS5A-YPet}-transfected cells ii) cells infected with FNX\textsuperscript{NS5A-YPet} and treated with 200pM daclatasvir for 48h and iii) control cells with 1% DMSO in media. Images were acquired as described in A. Insets represent zoom-ins of the boxed regions.
REFERENCES


Foster, T. L., T. Belyaeva, et al. (2010). "All three domains of the hepatitis C virus non-structural


localization using novel cell culture systems expressing core-NS2 and NS5A of

Gao, M., R. E. Nettles, et al. (2010). "Chemical genetics strategy identifies an HCV NS5A

Gastaminza, P., K. A. Dryden, et al. (2010). "Ultrastructural and biophysical characterization of

Gillespie, L. K., A. Hoenen, et al. (2010). "The endoplasmic reticulum provides the membrane


viruses with genotype 1 to 7 core-non-structural protein 2 (NS2) expressing fluorescent

Guedj, J., H. Dahari, et al. (2013). "Modeling shows that the NS5A inhibitor daclatasvir has two
modes of action and yields a shorter estimate of the hepatitis C virus half-life." Proc Natl

Hamamoto, I., Y. Nishimura, et al. (2005). "Human VAP-B is involved in hepatitis C virus


CHAPTER 5

Summary and future outlook
Specific Aim 1: To systematically identify essential and non-essential areas within the entire HCV genome at high resolution

Study 1 applied next-generation sequencing (NGS) to the functional profiling of a complete viral genome. Findings show that the newly-generated transposon insertion library that contains mutants with 15-nucleotide (15-nt) insertions randomly distributed across the entire hepatitis C virus (HCV) genome represented a nearly saturated library, with one insertion between almost every amino acid residue. We provided evidence that the library’s genetic footprints following passaging in cell culture reflected the biology of HCV. Furthermore, we showed how these genetic footprints can serve as a resource for gaining new insight on functional HCV regions. With this approach we identified a region in the non-structural protein NS4B that plays a role in post-RNA-replication steps of the viral life cycle. This resource, which we are making freely available to the public, will allow researchers to easily evaluate the effect of 15-nt insertions at 99.3% of HCV codons. Once cloned as DNA, a viral genome can serve as the starting point for this transposon insertion sequencing approach, opening the door to systematic functional analyses of viruses for which infectious clones are available. Such analyses are now precise enough to determine both the exact amino acid sequence of insertion mutations as well as the resulting effect on viral fitness. Thus, this approach, which we refer to as high-resolution genetics, creates comprehensive lists detailing how insertion mutations affect various infection processes, including the less-understood process of HCV assembly, across all ten viral proteins.

Specific Aim 2: To acquire high-resolution snapshots of viral assembly sites in HCV-infected cells

Study 2 developed protocols to study HCV infection in cultured cells using a variety of imaging modalities, including electron microscopy, electron tomography, immunoelectron microscopy, superresolution light microscopy, multi-color fluorescence microscopy and live-cell fluorescence
microscopy. Our ultrastructural analysis reinforces models in which the LD microenvironment provides a spatially restricted platform for the coordination of RNA replication (occurring in membrane vesicles), viral protein production (occurring at ER membranes) and particle formation. Despite the detection of virus-like particles within the proximity of LDs we were not able to conclusively identify the virion assembly site at the ultrastructural level. At the light microscope level, we optimized sensitive assays to visualize viral proteins core and NS5A at superresolution. We also generated a panel of replication-competent fluorescent reporter viruses that facilitate the phenotypic characterization of viral mutants, as well as the detection of alterations in a phenotypic markers upon treatment with a small molecule inhibitor. Fluorescent NS5A reporter viruses are also compatible with multi-color confocal imaging; in particular, we presented a protocol for the simultaneous detection of core, E2 and NS5A proteins close to LDs.

**Future Outlook**

The overall question for the next stage of work is: “what is the chronology of events leading up to the virus-induced structures, viral protein cluster formations and viral protein co-localizations we can now image using the approaches described in this dissertation?”

To address this question, I propose the following experiments.

**Time-lapse imaging of assembly-deficient virus mutants with mutations in non-structural proteins**

In chapter 4 of this dissertation, we have shown how fluorescent HCV reporter viruses allow for dynamic tracking of the NS5A protein, a critical player in both viral RNA replication and virion assembly. We also show how the imaging system can be exploited to detect an abnormal subcellular localization of NS5A produced by the I52N NS5A mutant virus, which is incapable of producing infectious viral particles. Chapter 3 presented a systematic approach for identifying
HCV regions essential for the viral life cycle. We showed how the resulting information can be used to identify functional regions involved in later steps of the life cycle, such as a short stretch in the N-terminal region of the non-structural protein NS4B. For future experiments, we propose to further analyze the phenotypes of insertion mutants showing genetic footprints that are consistent with a defect in post-RNA replication steps. Our candidate list of insertions in non-structural proteins with potential defects in post-RNA-replication steps (Table 3-5) also included insertion mutants in the NS5A and NS5B proteins. Introducing these insertions, as well as the previously-identified NS4B insertions, to the FNX^{NS5A-YPet} reporter virus could address whether these insertions also affect the subcellular localization and movement of NS5A inside live host cells. Our rationale for these experiments is that in order to learn about the chronology of events in HCV assembly, we need to selectively block assembly at different steps of the pathway and visualize abnormally accumulating assembly intermediates.

**Time-lapse imaging of NS5A subcellular localization under the treatment of viral inhibitors and other small-molecules shown to have an effect on HCV assembly**

Experiments described in chapter 4 used a potent inhibitor targeting NS5A functions and showed that drug treatment caused a subcellular localization similar to the one seen when introducing the I52N mutation to domain I of NS5A. This indicates that functions related to viral assembly can be targeted using small molecules, leading to a redistribution of the NS5A protein. Future studies are needed to characterize how patterns of NS5A subcellular localization differ following treatment with various classes of HCV inhibitors, such as NS3 protease inhibitors, NS4A inhibitors, NS4B inhibitors, NS5B polymerase inhibitors, p7 ion channel inhibitors, or core dimerization inhibitors. Once the resulting patterns of NS5A subcellular localization and kinetics of any observed changes are established in this live-imaging system, the next step could involve matching these profiles with profiles of additional small molecules targeting host cell factors shown to be essential for the HCV life-cycle. NS5A is known to interact with many host factors,
yet the functional role for these interactions in a complex process such as HCV assembly has not been established. Thus, molecules that appear to target HCV assembly but for which the exact mechanism of action (MOA) is missing would be of particular interest. In the proposed experimental set-up, the pattern of changes in NS5A subcellular localization could serve as a marker for the MOA of molecules targeting NS5A or viral assembly in general, as previously suggested for the replicon 1b system (Targett-Adams, Graham et al. 2011)

**Quantitative imaging of NS5A protein interacting with other viral or cellular proteins**

A series of regulated protein-protein interactions form the basis of the HCV assembly process. Methods that can measure these interactions within the environment of the host cell, such as fluorescence resonance energy transfer (FRET) or fluorescence cross-correlation spectroscopy (FCCS) techniques, will be pivotal to work out this set of interactions. FRET analysis of Gag, a major structural component of HIV-1, has already allowed the quantitative analysis of the assembly of individual virions in living cells (Jouvenet, Bieniasz et al. 2008). Our panel of fluorescently-labeled reporter viruses will facilitate such kinds of assays, which have the potential to uncover the sequence of events of key protein interactions during HCV assembly in a dynamic fashion. In particular, the YPet gene forms a strong FRET pair with the cyan variant CyPet fluorescent protein (Nguyen and Daugherty 2005; Kukulski, Schorb et al. 2011). Thus the FNX\textsuperscript{NS5A-YPet} virus reporter presented in Chapter 4 will provide an opportunity to examine of protein interactions in FRET experiments of HCV-infected cells.

**Correlated light and electron microscopy in HCV-infected cells**

Bridging the gap between light microscopy and electron microscopy should be the ultimate goal of future experiments. Light microscopy offers the advantage of rapid screening of a large area of specimen, determining the specificity of labeling and easy detection of multiple antigens at once. Ultimately, we wish to employ a correlated fluorescence light microscopy-electron
microscopy (CLEM) approach, which takes advantage of these benefits of light microscopy to guide the search for structures/events of interest and then uses electron microscopy to zoom in on them. Fluorescent proteins have revolutionized live-cell imaging of dynamic events. In chapter 4 of this dissertation, I have presented genetically engineered viruses that allow for easy, robust and live imaging of viral infection using fluorescence light microscopy. Technical advances in EM sample preparation now allow for the preservation of fluorescence using cryopreparations, as demonstrated by Briggs and coworkers (Kukulski, Schorb et al. 2011). Moreover, Watanabe et al. demonstrated that optimized cryosubstitution protocols preserve fluorescence at high enough levels to perform PALM and STED imaging on the resin-embedded sections, while maintaining a well-preserved ultrastructure (Watanabe, Punge et al. 2011). Our YPet-tagged reporter virus would be well-suited for these CLEM approaches due to its superior brightness and robust infection in cell culture.

In addition, our miniSOG-based constructs provide an opportunity for correlated confocal and EM imaging. This approach would take advantage of the modest green fluorescence and photo-generated singlet oxygen from miniSOG for fluorescence photo-oxidation of DAB. However, future studies will be needed to improve the amount of fluorescent signal detected for the miniSOG-based reporter viruses, as these constructs showed much lower fluorescent signals compared to the analogous YPet- and mCherry-tagged viral genomes.

CLEM approaches would allow us to address important questions raised by our findings in chapter 4. Currently we can only speculate whether I52N NS5A clusters correspond to membranous web structures, remodeled ER stacks or other modified cellular organelles. Using CLEM, we could determine which structures the NS5A-positive clusters corresponded to at the nanometer resolution of EM. Using ET, we could then carry out a 3D-analysis of these sites. CLEM of cells treated with HCV inhibitors could also provide valuable information to deduce mechanisms of action. In the end, we envision an integrated approach where high-throughput
mutagenesis platforms like the one described in chapter 3 produce candidate lists of viral mutants with defects at specific steps of the viral life cycle. Time-resolved light microscopy imaging using the fluorescent reporter systems described in chapter 4 then aids in the further phenotypic characterization of the mutants on the candidate list. This narrows the list further down to mutants, for which a phenotypic alteration (such as abnormal subcellular localization of viral protein) can be detected. Finally, advanced imaging studies like superresolution light microscopy, electron microscopy and electron tomography would be initiated once an interesting phenotype is observed, and would allow the researcher to further "zoom into" the cellular environment to provide a cellular context.

References


