Dissecting the requirements for Hepatitis C Virus RNA synthesis using a minus strand replication system

Dissertation

vorgelegt von

Lyudmila Shalamova

Dipl.-Biol.

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

Biochemisches Institut
der Medizinischen Fakultät
der Justus-Liebig-Universität Gießen

Gießen, Oktober 2018

This work was accomplished from August 2013 to October 2018 under the supervision of Prof. Dr. Michael Niepmann in the Institute of Biochemistry, Faculty of Medicine, Justus Liebig University Gießen.

Supervisors:

Prof. Dr. Albrecht Bindereif
Institute of Biochemistry
Faculty of Biology and Chemistry
Justus Liebig University Gießen

Prof. Dr. Michael Niepmann
Institute of Biochemistry
Faculty of Medicine
Justus Liebig University Gießen

Contents

I. Summary

The presented research focuses on viral and cellular determinants of the Hepatitis C Virus (HCV) replication. In particular, a novel minus strand replication system was developed to study distinct aspects of the HCV minus strand synthesis initiation. The system enabled a revision of state-of-the-art requirements for the HCV antigenome synthesis when uncoupled from possibly overlapping functions in viral translation and/or plus strand synthesis. So far mainly addressing the genome RNA sequence and structural elements prerequisites, the system displays a universal potential for investigation of replication of other plus strand RNA virus genomes.

HCV is a positive-sense single-stranded RNA virus that infects human hepatocytes and causes both acute and chronic hepatitis. The challenge of elimination of HCV as a public health threat is largely complicated by absence of an efficient vaccine and by rapid emergence of virus quasispecies resistant to the existing treatment. Therefore, novel approaches are required for further understanding of molecular mechanisms of the HCV life cycle. Development of the autonomous replicon and full-length HCV cell culture systems has enabled a substantial breakthrough in understanding of the HCV translation and replication. However, these systems allow analysis of *cis*-elements and *trans*-acting factors required for genome replication only in presence of both genomic ends. Thus, it remains unclear, whether an element identified in the annotated HCV genome exerts its function when physically present on the plus or the minus strand, or on both. Additionally, given an entangled nature of viral RNA and protein synthesis, a number of overlapping *cis*-elements cannot be assigned to a specific step of the viral life cycle. To overcome the above limitations, the minus strand replication system was designed to specifically focus on prerequisites for the HCV antigenome production.

The system, which represents a set of replication constructs assembled in agreement with the current knowledge on the HCV replication, uncouples to various extents the HCV minus strand synthesis from the plus strand synthesis and translation, thereby empowering diversified applications. The very 5'-end of the annotated HCV genome constituted by the stem-loop (SL) I and II domains was confirmed to be essential for the minus strand synthesis initiation at the genomic 3'-end. The positive regulation of the antigenome production was found to extend up to inclusion of the SL III domain; however, enabling of functional translation initiation from the HCV internal ribosome entry site (IRES) resulted in a profound negative effect on RNA replication. The latter observation has encouraged in-depth analysis of a balance between the HCV genome translation and replication that suggested an importance of the genomic RNA ends communication. Yet lacking an experimental confirmation, a possible circularization of the HCV RNA for efficient replication is supported by the importance of long-range RNA-RNA interactions between *cis*-elements, which were addressed by mutational analysis. Along with validation of the previously reported regulatory RNA elements, a comprehensive screening for *cis*-acting replication elements within the protein-coding sequence was undertaken. In addition to addressing the genome sequence and structure determinants, the minus strand replication system was utilized to examine a regulation of the HCV RNA synthesis by selected cellular factors. A positive role of the most of the liver-specific microRNA-122 (miR-122) binding sites on efficiency of the HCV minus strand synthesis initiation was demonstrated.

In conclusion, an assay system for the specific analysis of requirements for the HCV minus strand RNA synthesis was developed. Uncoupling of the HCV minus and plus strand replication from each other and from translation enabled a dissection of essential *cis*-elements and assignment of their functions specifically to antigenome synthesis. Ultimately, the versatility of the system enables further characterization of regulatory *trans*-factors and investigation of the interplay of molecular processes during the HCV life cycle as well as of other RNA viruses.

II. Zusammenfassung

Diese Arbeit befasst sich mit viralen und zellulären Voraussetzungen der Hepatitis-C-Virus (HCV) Replikation. Ein neuartiges Minusstrang-Replikationssystem wurde entwickelt, um spezifische Aspekte der Initiation der Minusstrangsynthese zu untersuchen. Dieses System ermöglicht eine genaue Betrachtung der Voraussetzungen für die HCV Antigenom-Synthese, wenn diese nicht mit potentiell überlappenden Funktionen der Positiv-Strang-Synthese und/oder der viralen Translation gekoppelt ist. Dadurch besteht ein hohes Potential, dieses System als Modell für Untersuchungen der Replikation anderer Positivstrang-RNA-Viren Genomen zu verwenden, da es hauptsächlich die Erfordernisse der RNA-Genomsequenz sowie strukturelle Elemente einbezieht.

HCV ist ein einzelsträngiges RNA-Virus positiver Polarität, welches menschliche Hepatozyten infiziert und sowohl akute als auch chronische Hepatitis verursacht. Die Herausforderung der Bekämpfung von HCV als Gefahr für die öffentliche Gesundheit wird durch das Fehlen eines wirksamen Impfstoffs und das schnelle Auftreten von Virus-Quasispezies, die gegen die bestehende Behandlung resistent sind, erheblich erschwert. Daher sind neue Ansätze erforderlich, um die molekularen Mechanismen des HCV-Lebenszyklus besser zu verstehen. Die Entwicklung autonomer Replikon- und HCV-Zellkultursysteme mit Genomen in voller Länge ermöglichte einen wesentlichen Durchbruch beim Verständnis der HCV Translation und Replikation. Allerdings ermöglichen diese Systeme die Analyse von *cis*-Elementen und *trans*-wirkenden Faktoren, welche auf die Genomreplikation nur in Gegenwart beider Genomenden wirken. Somit bleibt unklar, ob ein Element, welches im HCV-Genom identifiziert wurde, seine Funktion ausübt, wenn es physikalisch auf dem Plusstrang oder auf dem Minusstrang, oder auf beiden liegt. Angesichts der komplexen Natur der viralen RNA- und Proteinsynthese können überlappendende *cis*-Elemente nicht einem spezifischen Schritt des viralen Lebenszyklus zugeordnet werden. Um diese Einschränkungen zu überwinden, wurde ein Minusstrang-Replikationssystem entwickelt, um speziell die Voraussetzungen für die HCV-Antigenom-Produktion zu untersuchen.

Dieses System, welches eine Reihe verschiedener Replikationskonstrukte beinhaltet, die in Übereinstimmung mit dem aktuellen Verständnis der HCV-Replikation zusammengestellt wurden, entkoppelt die HCV-Minusstrangsynthese von der Plusstrangsynthese und der Translation, wodurch vielfältige Anwendungsmöglichkeiten bestehen. Das 5'-Ende des HCV-Genoms, welches aus den *Stem-Loop* (SL) I- und II-Domänen besteht, wurde als essentiell für die Initiation der Minusstrangsynthese am genomischen 3'-Ende bestätigt. Die Region, die für den positive Einfluss auf die Antigenom-Produktion verantwortlich ist, schließt die SL III-Domäne mit ein, jedoch führte die Einführung funktioneller Translationsinitiation durch die HCV interne ribosomale Eintrittsstelle (IRES) zu einem negativen Effekt auf die RNA-Replikation. Aufgrund dieser Beobachtung wurde eine eingehende Analyse des Gleichgewichts zwischen HCV-Genom-Translation und -Replikation durchgeführt, welche eine Interaktion zwischen den beiden genomischen RNA-Enden nahelegt. Obwohl bisher nicht experimentell bestätigt, spricht viel dafür, dass das HCV-Genom zirkularisiert, um effizient zu replizieren. Die Signifikanz der vorhergesagten RNA-RNA-Interaktion zwischen *cis*-Elementen über weite Distanzen wurde durch eine Mutationsanalyse überprüft. Neben zuvor beschriebenen regulatorischen RNA-Elementen wurde zusätzlich eine umfassende Suche nach *cis*-aktiven Replikationselementen innerhalb der Protein-kodierenden Sequenz durchgeführt. Zusätzlich zur Untersuchung der Genomsequenz und von Strukturelementen wurde das Minusstrang-Replikationssystem verwendet, um eine Regulation der HCV-RNA-Synthese durch ausgewählte zelluläre Faktoren zu untersuchen. Eine wichtige Rolle der meisten Bindesstellen der leberspezifischen microRNA-122 (miR-122) auf die Effizienz der HCV Minusstrangsynthese-Initiation wurde gezeigt.

Zusammenfassend wurde in dieser Arbeit ein experimentelles System für die Analyse von spezifischen Anforderungen für die HCV-Minusstrang-RNA-Synthese entwickelt. Die Entkoppelung der HCV-Minus- und Plusstrangreplikation voneinander und von der viralen Translation ermöglichte eine präzise Analyse wichtiger *cis*-Elemente und ihrer spezifischen Funktionen bei der Antigenomsynthese. Letztlich ermöglicht die Vielseitigkeit dieses Systems eine Charakterisierung von regulatorischen *trans*-Faktoren und die Untersuchung des Zusammenspiels molekularer Prozesse während des HCV-Lebenszyklus sowie bei anderen RNA-Viren.

1. Introduction

1.1 General information on Hepatitis C Virus

1.1.1 HCV history

HCV existence was first fully recognized in 1975 (Feinstone et al. 1975) as an unidentified infectious agent causing the so-called non-A/non-B (NANB) transfusion-associated hepatitis. In 1989 an experimental chimpanzee model was applied for the study of NANB hepatitis that allowed demonstrating the transmissible, viral nature of the agent. Via a blind immunoscreening approach using high-titer samples, a cDNA clone encoding an antigen specific for NANBH-infected patients was isolated (Choo et al. 1989). Sharing the properties with *Togaviridae* and *Flaviviridae*, such as a 10 kb RNA genome with a positive polarity with respect to the encoded antigen, the novel infection agent was designated as Hepatitis C Virus (HCV). Simultaneously, a cDNA clone from plasma of patients chronically infected with NANBH was derived with a characterization of a different viral epitope (Maéno et al. 1990).

This progress has rapidly led to a development of a serological assay improving the diagnostics of HCV and blood-donor screening (Kuo et al. 1989). In a short period of time the nucleotide sequence of HCV full-length genome was arranged from multiple overlapping cDNA clones. A single ORF encoding a polyprotein of 3011 amino acids and its main domains were described and homology to other members of the *Flaviviridae* family was postulated (Choo et al. 1991). HCV was later classified as a separate genus Hepacivirus within the *Flaviviridae* family and as a very distant relative of the Flavivirus genus and the Pestivirus genus (Robertson et al. 1998).

Conducted in 1993 a vast phylogenetic analysis of partial HCV sequences from a worldwide range of clinical isolates enabled a classification into 6 genotypes (nucleotide sequence differs by 30–35 %) including a various number of subtypes (20-25 % sequence difference; designated a, b, etc.) (Simmonds et al. 1993; Simmonds et al. 2005). In 1997 a consensus sequence of the genotype 1a (strain H77) was determined and lead to a generation of the first infectious molecular cDNA clones (Kolykhalov et al. 1997; Yanagi et al. 1997). Following the classification verification by the analysis of full-length sequences, a genotype 7 was newly discovered and affiliated (Murphy et al. 2015) resuming the existence of 7 major genotypes and 67 subtypes (Smith et al. 2014; Tarr et al. 2015). Genotype 1 and 3 are the most prevalent worldwide as well as in Europe (mostly subtypes 1a, 1b and 3a) (Bukh 2016), however a higher risk of progressive liver disorder is associated with genotype 3 (Smith et al. 2014).

In May 2016 the World Health Assembly on viral hepatitis proposed to eliminate viral hepatitis as a public health threat by 2030. The elimination is defined as a 90 % reduction in new chronic infections and a 65 % reduction in mortality comparing to the 2015 statistics. According to the WHO report from 2018, 71 million people globally have chronic hepatitis C infection (WHO 2018), comparing to 130-150 million people in 2016 (WHO 2016), with yearly 399 000 death cases, from hepatitis C-related liver diseases (cirrhosis, hepatocellular carcinoma and liver failure), comparing to former 0.7 million death rates (Lozano et al. 2012). Therefore, the Hepatitis C elimination still remains to be a great challenge (WHO 2018).

1.1.2 HCV genome, virology and life cycle overview

The genome organization of HCV RNA is similar to that of the other members of the *Flaviviridae* family and is represented by a positive-sense single-stranded RNA composed of a 9.6 kb long open reading frame (ORF) flanked by two untranslated regions (UTRs) at its 5'- and 3'- ends. The ORF encodes a polyprotein precursor of about 3000 amino acids that is cleaved co- and post- translationally into three structural and seven non-structural (NS) proteins by host or viral proteases (Fig. 1.1.1). The structural proteins – Core and envelope glycoproteins E1 and E2 – are involved in viral particle formation. The non-structural proteins include the p7 viroporin, the NS2 protease, the NS3-4A complex with protease and NTPase/RNA helicase activities, the NS4B and NS5A proteins, and the NS5B RNA-dependent RNA

polymerase (RdRp) that enable viral processing, replication and assembly (Moradpour et al. 2007; Moradpour and Penin 2013). Translation from an alternative "core +1" reading frame results in a production of the F protein of various sizes with yet unclear function (Branch et al. 2005).

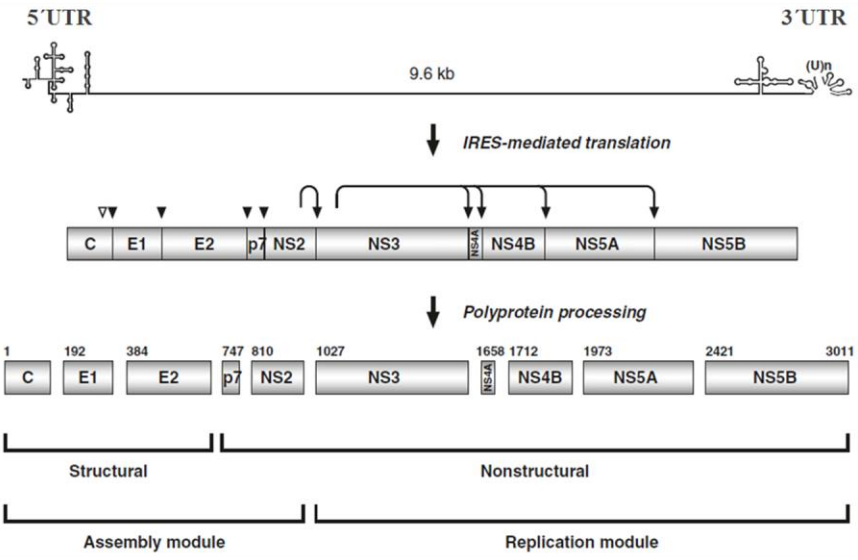


Figure 1.1.1: HCV genome and polyprotein processing.
The 9.6 kb RNA genome with flanking 5'- and 3'- untranslated regions (UTRs) is depicted above. Undergoing IRES-mediated translation and polyprotein processing by proteases results in three structural and seven non-structural proteins assigned to the assembly or to the replication modules. Solid arrowheads refer to posttranslational cleavage by the endoplasmic reticulum signal peptidase; the arrows stand for a cleavage by the viral NS2 protease co-translationally and by the NS3-NS4A protease posttranslationally. The open arrowhead indicates further C-terminal processing of the Core protein by the signal peptide peptidase.
(modified from Moradpour and Penin 2013)

The highly structured 5'- and 3'- UTRs play predominant roles in HCV translation and replication. In the 5'UTR stem-loops (SL) II to IV constitute the major part of the IRES, whereas SL I and II are essential for replication (Friebe et al. 2001) and harbor the two vital microRNA-122 (miR-122) binding sites contributing considerably to both processes (Jopling et al. 2005; Henke et al. 2008). The 3'UTR is composed of a variable region with two stem-loop structures, a 30-90 nt long polyU/UC tract and an almost invariant 98-nt X-tail (Fig. 1.1.2). All elements to a different extent facilitate initiation and regulation of the minus strand synthesis (Friebe and Bartenschlager 2002), as does an additional *cis*-acting replication element (CRE) 5BSL3.2 within the NS5B coding region that is engaged in a kissing-loop interaction with SL2 (Friebe et al. 2005).

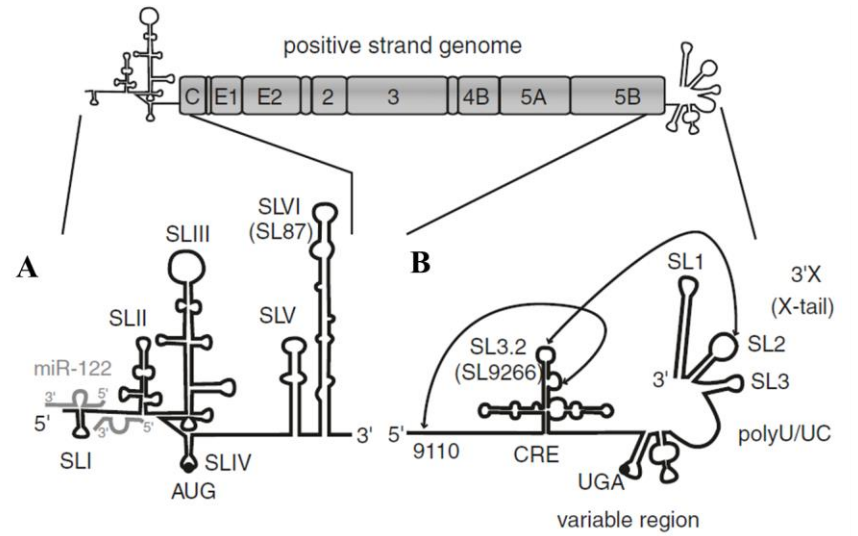


Figure 1.1.2: The 5'- and 3'- untranslated regions in HCV genomic RNA (A and B, respectively).
Arrows indicate long range interactions at the 3'UTR; two miR-122 molecules bound at the 5'UTR are depicted in grey. The details are provided in 1.1.2.
(from Lohmann 2013)

Although HCV is related to the broadly investigated Flaviviruses, the morphology of its virions remains poorly characterized due to their unusual low buoyant density. Their enveloped nature was validated by chloroform inactivation followed by infection of chimpanzees (Houghton 2009), whereas gel filtration and electron microscopy studies identified the mature particles to be 50-80 nm in diameter (Wakita et al. 2005). An infectious virion contains a spherical nucleocapsid of about 30 nm in diameter that

is formed by the Core protein and serves to contain supposedly a single copy of the HCV genomic RNA. The capsid is surrounded by a lipid-containing host cell derived double-layered lipid envelope supplemented with anchored E1 and E2 viral glycoproteins which are involved in entry of the HCV into hepatocytes via a number of host cell receptors (Moradpour et al. 2007; Dubuisson and Cosset 2014). Besides its primary target cells – hepatocytes - HCV was also reported to infect B cells, dendritic cells and some other cell types using CD81, the LDL receptor (LDLR), scavenger receptor class B type I (SR-BI), heparan sulfate proteoglycan (HSPG) and claudin-1 (Zeisel et al. 2013). HCV circulates in various forms in the infected host and can be associated with low-density lipoproteins (LDL) and very-low-density lipoproteins (VLDL), or with cholesterol esters accounting for almost half of the total HCV lipids (Merz et al. 2011).

In more detail, the HCV life cycle (Fig. 1.1.3) starts when the primary infection HCV particles are transported by the blood stream and access hepatocytes by entering the space of Disse through fenestrae between hepatic endothelial cells. The predominant transmission mode during ongoing HCV infection is a direct cell-to-cell spread. First interactions with attachment factors and receptors on the cell surface are conducted by both the lipoprotein components and the viral envelope glycoproteins. Docking of the virions due to ApoE interaction with HSPG, SR-BI or LDLR is then followed by the tetraspanin CD81 interaction with the a core of the E2 protein. CD81 acts in association with a co-receptor Claudin-1 (CLDN1) and primes the HCV envelope proteins for low pH-dependent fusion (Bartosch et al. 2003b; Zeisel et al. 2013; Dubuisson and Cosset 2014). The viral particles internalize via a clathrin-dependent endocytosis (Blanchard et al. 2006) and occur in early endosomes to undergo an endosomal membrane fusion. In turn, that leads to the HCV genome release into the cytosol where both translation and replication take place (Lindenbach and Rice 2013; Zeisel et al. 2013; Dubuisson and Cosset 2014).

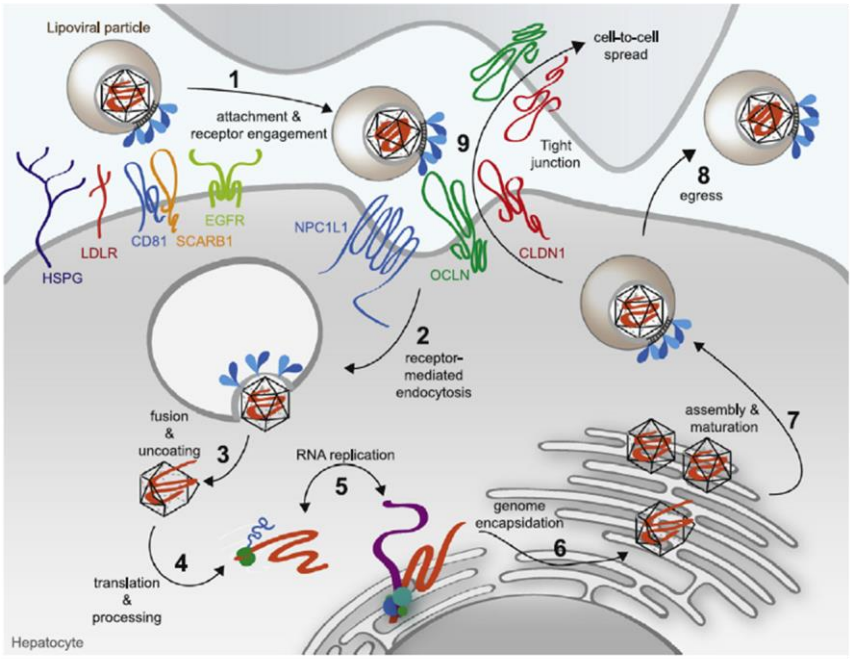


Figure 1.1.3: An overview on the HCV lifecycle.
(1) Primary infection of a hepatocyte by the HCV viral particle requires interaction with attachment factors and receptors. **(2)** Internalization of the virus via clathrin-dependent endocytosis and fusion leads to **(3)** the virus uncoating in the cytoplasm where **(4)** IRES-mediated translation and polyprotein processing take place. **(5)** HCV minus and plus strand RNA synthesis at the membranous web result in either genomic RNA accumulation or **(6)** genome encapsidation. **(7)** Along the maturation virions acquire the viral envelope and associate with endogenous lipoproteins. Lipoviroparticles are either **(8)** released from the cell or **(9)** infect neighboring cells via a direct cell-to-cell spread.
(from Catanese and Dorner 2015)

HCV RNA translation is entirely dependent on the cellular translation machinery as well as on a number of non-canonical RNA-binding proteins that are not involved in the regulation of cap-dependent eukaryotic translation initiation (Niepmann 2013). A distinct tertiary structure of the HCV IRES mediates direct binding and positioning of the 40S ribosomal subunit resulting in the 80S initiation complex assembly assisted by the eukaryotic initiation factor 3 (eIF3) (the detailed mechanism of HCV translation is elucidated in 1.2.2). After translation and polyprotein processing by cellular and viral proteases (Fig. 1.1.1) the HCV proteins remain associated with the endoplasmic reticulum (ER), where on site the replication machinery is arranged by non-structural proteins NS3/4A, NS4B, NS5A, and NS5B (the detailed mechanism of HCV replication is described in 1.3). In brief, NS4B protein mediates vesicular membrane alterations resulting in

formation of a so-called membranous web with double- or multi- membrane vesicles (DMVs/MMVs) that play a role of replication factories. The viral RNA-dependent RNA polymerase (RdRp) NS5B initiates synthesis of minus strand intermediates that in turn serve as a template for multiple copies of genomic plus strands. Nascent RNA genomes can either re-enter a new translation/replication round or can be assembled into infectious virions (Lohmann 2013). A number of host factors (e.g. RNA-binding proteins, miR-122 and lipoproteins) are involved in regulation of different steps of HCV replication. Lipid droplets (LDs) are located in proximity of the membrane vesicles in close association with the HCV Core and NS5A and are currently hypothesized to coordinate viral RNA synthesis and virion morphogenesis through the physical association of replication and assembly sites (Miyanari et al. 2007).

Typically for the *Flaviviridae* family, HCV particle assembly requires engagement of the viral non-structural proteins and is coupled with lipid metabolism. HCV assembly is conducted by a coordinated action of the envelope E1-E2 glycoprotein complex and LD-associated Core protein. Posttranslationally retained in the ER, E1-E2 heterodimers migrate to the virion assembly site via interaction with NS2 and p7. The C-terminal domain of NS5A is responsible for the switch from replication to assembly due to an interaction with the LD-bound Core protein. Other viral proteins such as NS3/4A enzyme complex, NS4B and NS5B along with host factors are also implicated in the assembly process. Virions presumably form by budding and exit the cell through the secretory pathway. In addition to producing extracellular virus particles, HCV has been reported to directly infect neighboring cells without releasing detectable virus particles (Lindenbach 2013).

1.1.3 HCV cell culture systems

Despite the rapid progress in characterization of HCV genome and polyprotein organization, a generation of an experimental cell culture system appeared to be a major challenge. In fact, clinical isolates of HCV failed to induce productive infection in cell culture: infections were originally reported in continuous human T and B cell lines, but remained inefficient and often irreproducible (Lohmann and Bartenschlager 2014). Although a fully permissive cell culture system was developed only in 2005 (Lindenbach et al. 2005; Wakita et al. 2005; Zhong et al. 2005), a few auxiliary *in vitro* model systems have been successfully applied by then for the investigation of individual steps of the HCV life cycle.

The primary step towards the generation of recombinant experimental systems was creation of infectious molecular cDNA clones for the genotype 1a in 1997 (Kolykhalov et al. 1997; Yanagi et al. 1997). An ability of an intrahepato-injected RNA transcript to infect chimpanzees defined the genetic elements' prerequisites for the further design of HCV replicons. In 1999 Lohmann and co-workers defined the minimal viral sequence required for the virus replication in human cells to be constituted by the 5'UTR, NS3-NS5B and the 3'UTR (strain Con1, genotype 1b) (Fig. 1.1.4). Representing a selectable bicistronic construct, where an antibiotic resistance gene was driven from the HCV IRES and the non-structural proteins - from the Encephalomyocarditis Virus (EMCV) IRES, these subgenomic replicons were capable of self-replication in HuH-7 hepatoma cell lines (Lohmann et al. 1999) (Fig. 1.1.5, B). The relatively low initial levels of replication were improved by adaptive mutations of the replicating HCV RNA and by increased host cell permissiveness (Lohmann et al. 2001). Therefore, selection for viral genomes with replication-enhancing

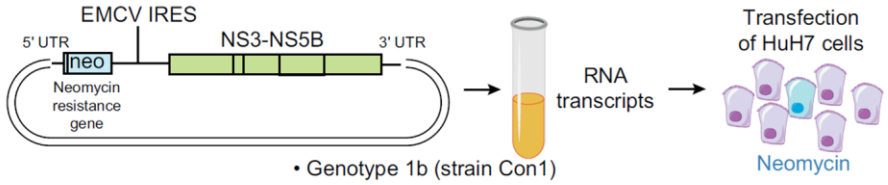


Figure 1.1.4: Structure and generation of HCV subgenomic replicons.
The selectable bicistronic constructs consist of the HCV 5'UTR, the neomycin (*neo*) gene, the EMCV-IRES, NS3-NS5B and the 3'UTR (genotype 1b) (Lohmann et al. 1999). After transfection of HuH-7 cells autonomously replicating clones are isolated using neomycin selection.
(from Bukh 2016)

mutations (REMs) and for permissive clones of a human hepatoma cell line with subsequent curing of the cells with interferon (IFN) lead to the generation of the highly permissive HuH-7.5 cells (Blight et al. 2002; Zhong et al. 2005). This clone appeared to be highly permissive for HCV RNA due to a single point mutation in Retinoic Acid Inducible Gene I (RIG-I) resulting in defective IFN signaling and therefore enhanced production and spread in cell culture (Sumpter et al. 2005).

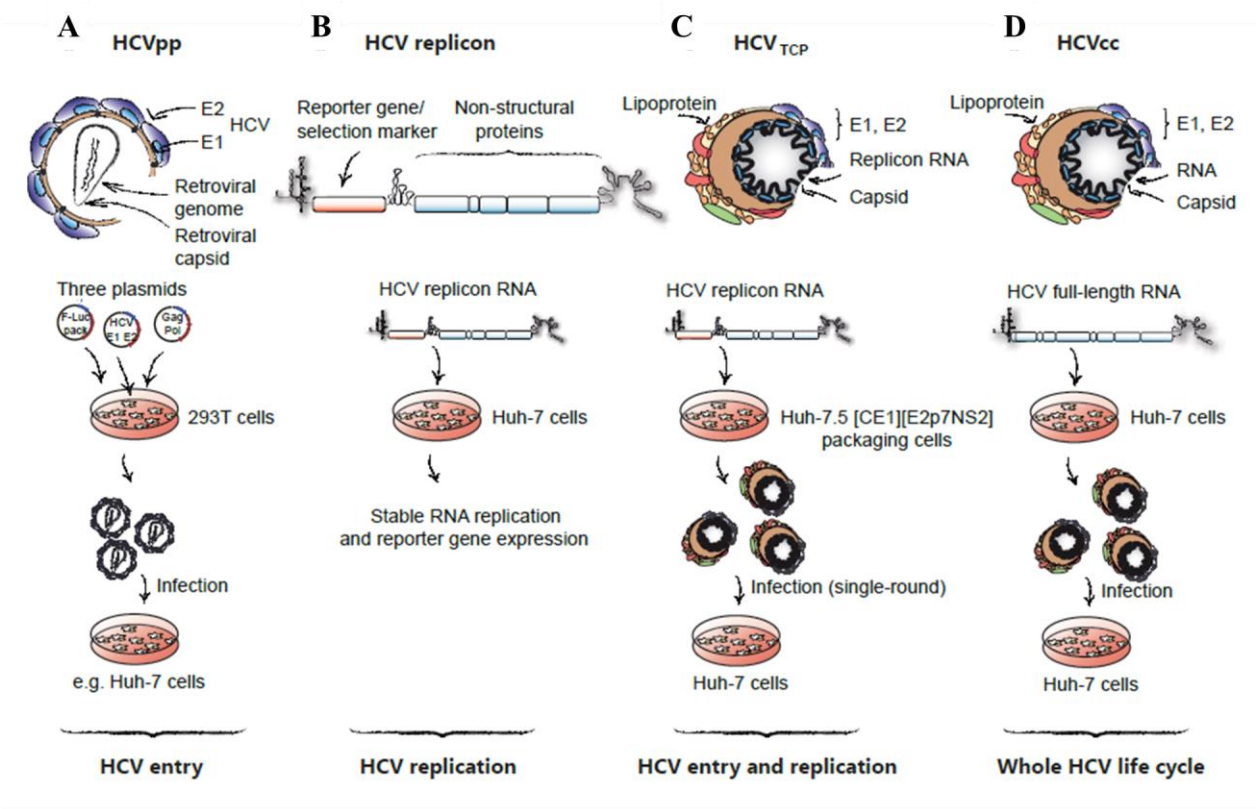


Figure 1.1.5: The main HCV cell culture model systems.
(A) The retroviral pseudoparticle system (HCVpp) (Bartosch et al. 2003a).
(B) The subgenomic HCV replicon system (Lohmann et al. 1999).
(C) The *trans*-complemented HCV particle system (HCV_{TCP}) (Steinmann et al. 2008).
(D) The cell culture infectious system (HCVcc) (Lindenbach et al. 2005; Wakita et al. 2005; Zhong et al. 2005).
The further details are provided in 1.1.3.
(from Gerold and Pietschmann 2014)

Another important *in vitro* system addressing the mechanism of the HCV attachment and entry is the pseudoparticle system (HCVpp; Fig. 1.1.5, A). Designed in 2003, it is represented by defective retroviral particles presenting the HCV E1-E2 glycoproteins on their surface (Bartosch et al. 2003a). Such pseudoparticles are generated in HEK 293T cells by co-transfection of three expression plasmids encoding E1/E2, retroviral Gag-Pol proteins and a retroviral provirus with a reporter gene. Upon entry into susceptible cells, followed by processing and integration of proviral RNA, the expression level of a reporter gene allows quantitative evaluation of the HCV entry. The HCVpp model system was an indispensable tool for investigation of E1-E2-mediated HCV entry and for identification of receptors and attachment factors (Ziesel et al. 2013).

Commonly, an infection with patient-derived HCV isolates is inefficient in cell culture due to the lack of a key host factor SEC14L2 that permits replication of different HCV genotypes without a need for adaptive mutations (Saeed et al. 2015). The only exception so far was a genotype 2a strain from a Japanese patient with fulminant hepatitis (Kato et al. 2001), termed JFH1, that within a subgenomic replicon could replicate in original HuH-7 cells without a requirement for cell culture adaptive mutations (Kato et al. 2003). This unique strain has eventually paved the way for the first cell culture infectious HCV system (HCVcc; Fig. 1.1.5, D). Described independently by three research groups (Lindenbach et al. 2005; Wakita et al. 2005;

Zhong et al. 2005) a recombinant genotype 2a HCV genome was capable of replication and infectious virus assembly in HuH-7.5 cells. Combination of the original JFH1 NS3-NS5B region with the Core-NS2 region from another genotype 2a isolate - J6 - resulted in intragenomic chimeras - JFH1/J6 (genotype breakpoint is between NS2 and NS3) (Lindenbach et al. 2006) or Jc1 (genotype breakpoint is within NS2) (Pietschmann et al. 2006) – characterized by enhanced replication, spread and particle production. Further on, chimeric HCV genomes (encoding structural proteins of a genotype of interest and non-structural proteins from the JFH1 with certain adaptive mutations) were developed for all 7 genotypes (Gottwein et al. 2009).

In order to overcome the difficulties with adaptation of different isolates in cell culture, an alternative *trans*-complemented particle system (HCV_{TCP}; Fig. 1.1.5, C) was engineered. This approach requires a JFH1 subgenomic replicon providing non-structural proteins to be transfected into so-called packaging cell lines. The *trans*-complemented assembly module can be derived from any HCV isolate to elucidate isolate-specific mode of viral entry and replication dissected from assembly, however being capable of a single-round infection only (Steinmann et al. 2008).

A development of *in vitro* cell culture systems - the HCVcc system on the first place - enabled studying of all aspects of the HCV life cycle *in vitro* and so far remains to be the most widely used experimental system in the field. However, the generation of cell systems permissive to HCV replication *in vitro* appeared to be a challenge for many years. Nowadays, the most permissive cell line for efficient replication *in vitro* is the HuH-7 human hepatoma cell line and its derivatives, in particular HuH-7.5 cells (Blight et al. 2002). One of the major limitations of such experimental systems is a non-polarity of HuH-7-based cultures in comparison to highly polarized hepatocytes in the liver. HepG2 cells, in contrast, polarize in culture, but require ectopic expression of human CD81 and miR-122 to become fully permissive (Steinmann and Pietschmann 2013). Recently primary hepatocyte cell culture models have become available to shed light on the role of host genetics in HCV infection (Ploss et al. 2010). Nevertheless, investigation of many aspects of the HCV infection still remains limited by the existing models.

1.2 HCV translation

1.2.1 Canonical versus IRES-mediated translation initiation

The eukaryotic translation initiation on cellular mRNA is a complex, strictly regulated process that involves a large number of participants, such as nine eukaryotic initiation factors (eIF), 40S and 60S ribosomal subunits, Met-$tRNA_i$ and other auxiliary factors.

The canonical translation (Fig. 1.2.1) typically begins with the 5'cap (m^7G) recognition by eIF4E followed by recruitment of the scaffold protein eIF4G, which in turn mediates unwinding of the 5'-terminal region of the mRNA by eIF4A, eIF4B and eIF4F in order to prepare it for ribosomal attachment. Interaction of the poly(A)-binding protein (PABP) bound to the poly(A) tail at the 3'-end with eIF4G directs mRNA circularization. The subsequent recruitment of the 43S preinitiation complex is mainly driven by interaction between eIF4G and eIF3. The 43S complex itself is comprised of the 40S subunit, the eIF2/GTP/Met-$tRNA_i$ ternary complex, eIF3, eIF1, eIF1A and eIF5 and serves for scanning of the mRNA downstream the cap to locate an initiation codon (which is usually the first AUG triplet in an optimal context GCC(**A/G**)CCAUG**G**) (Kozak 1987). Establishment of a stable codon-anticodon base-pairing is mediated by conformational changes of the ribosomal complex resulting in a "closed" conformation locked onto the mRNA (the 48S initiation complex). These events lead to a displacement of eIF1, permitting eIF5-mediated hydrolysis of eIF2-bound GTP and P_i release. Association of the 60S subunit to the 48S complex and eIF5B-mediated displacement of eIF2/GDP and other factors (eIF1, eIF3, eIF4B, eIF4F and eIF5) ultimately yield an elongation-competent 80S ribosome (Pestova and Kolupaeva 2002; Altmann and Linder 2010; Jackson et al. 2010; Voigts-Hoffmann et al. 2012).

Viruses, being naturally not capable of encoding their own translation machinery, have developed a range of peculiar mechanisms driven by specific genomic RNA sequences. In the absence of canonical

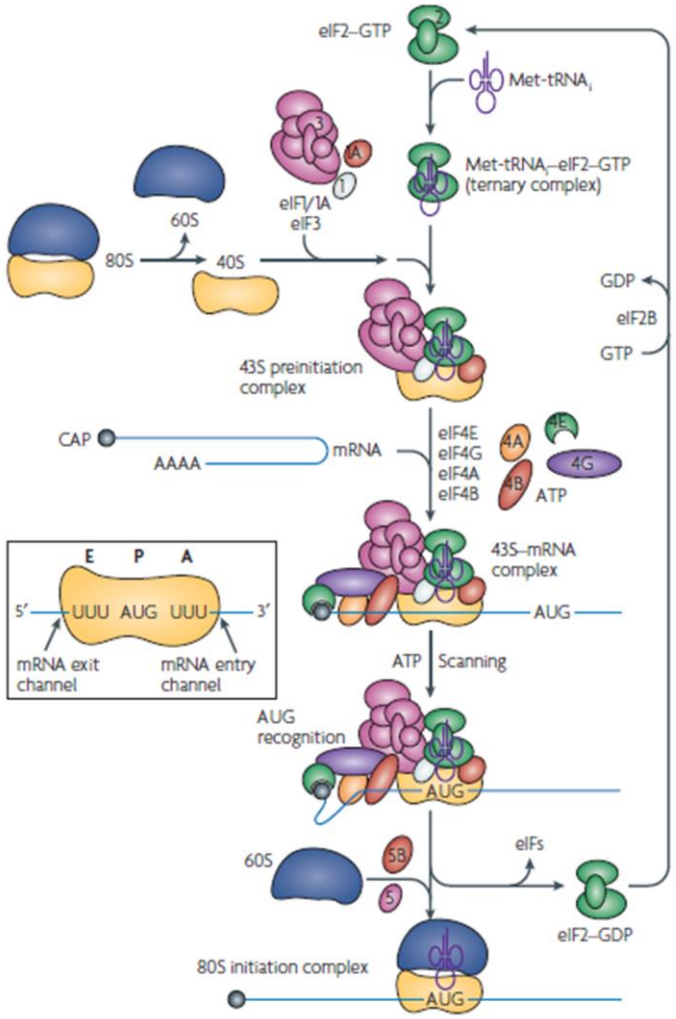


Figure 1.2.1: A schematic mechanism of the canonical eukaryotic translation initiation (the details are given in section 1.2.1).
(from Fraser and Doudna 2007)

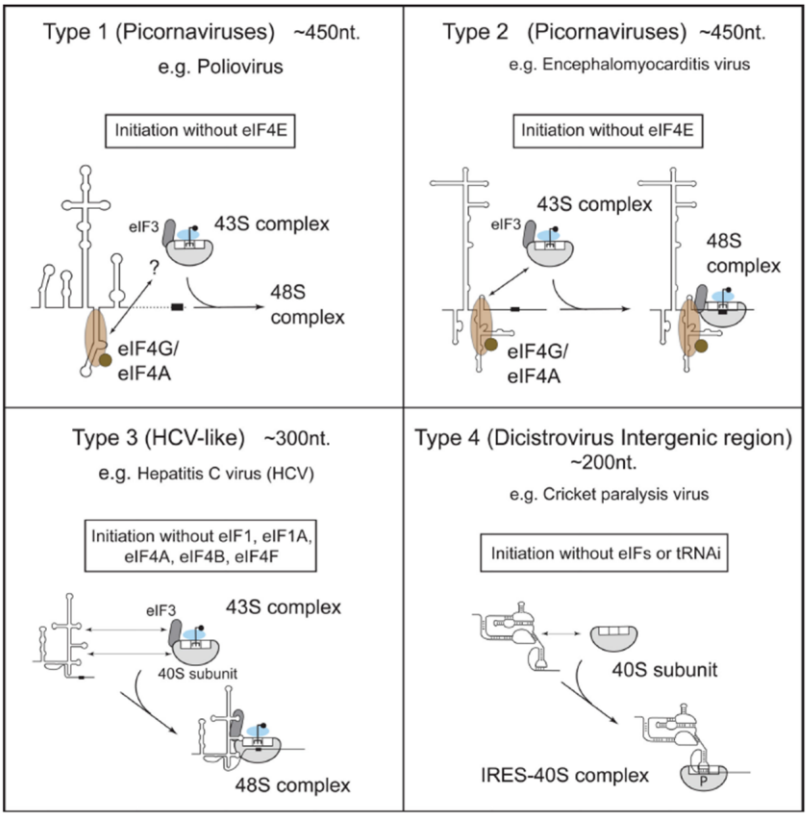


Figure 1.2.2: Comparison of ribosome recruitment mechanisms by four types of viral IRESs (a detailed description is provided in section 1.2.1).
(from Jackson et al. 2010)

signals for ribosome recruitment virus appear to be concurrently and entirely dependent on their hosts for the required components and have to achieve an advantageous customized mode of action as well as to circumvent immune surveillance mechanisms.

The first mechanism of a non-canonical cap-independent translation initiation was disclosed in 1988 for Poliovirus (PV) and Encephalomyocarditis Virus (EMCV) RNA referring to as an internal translation initiation mechanism promoted by uniquely structured 5'UTR RNA sequence termed internal ribosome entry site (IRES) (Jang et al. 1988; Pelletier and Sonenberg 1988). This discovery rapidly led to identification of a vast number of other examples of IRES-mediated translation initiation in both viral and cellular mRNAs (Kieft 2008). Moreover, to further compact their genomes and maximize coding capacity, viruses were found to manipulate host ribosomes to shift reading frames and translate overlapping open reading frames (Kerr and Jan 2016). Potential cellular IRESs share little structural similarity to each other and their underlying mechanism remains largely unknown, however, believed to bear resemblance to the one of Picornaviruses (Jackson et al. 2010).

An immediate in-depth detailing of the IRES structure and function was of high importance for a number of pathogenic viruses such as Hepatitis A Virus (HAV) (Glass and Summers 1992), HCV (Tsukiyama-Kohara et al. 1992), Foot-and-Mouth Disease Virus (FMDV) (Kuhn et al. 1990), HIV (Buck et al. 2001) and many others. Currently, viral IRESs are classified into four types according to their secondary structure and interaction with eIFs and the 40S subunit (Fig. 1.2.2). Initiation on Type 1 (Poliovirus, Rhinovirus) and 2 (EMCV, FMDV) IRESs involves most of the canonical eIFs, Met-tRNA$_i$ and additional IRES *trans*-acting factors (ITAFs) and starts downstream of the IRES. In contrast, Type 4 IRES (Cricket Paralysis Virus; CrPV) is capable of binding directly to the 40S subunits independently of any auxiliary proteins: the IRES domain occupies the P-site and there mimics codon-anticodon base-pairing (Wilson et al. 2000). The HCV IRES represents a typical example of the Type 3 IRES; it binds to the 40S ribosomal subunit and requires fewer canonical factors (eIF2; eIF3 and Met-tRNA$_i$) than Type 1 and 2 IRESs (Tsukiyama-Kohara et al. 1992).

1.2.2 IRES-mediated translation initiation in HCV

1.2.2.1 HCV IRES structure

Following a functional identification of the HCV IRES in 1992 (Tsukiyama-Kohara et al. 1992), multiple approaches enabled a rapid identification of its secondary structure. It was shown to consist of three main structural domains II-IV comprising ~340 nucleotides that adopt a tertiary fold under physiological salt conditions (Kieft et al. 1999) (Fig. 1.2.3). The 5'- and 3'- boundaries of the HCV IRES have been mapped using dicistronic reporter assays that located it to the residues 40 to 372 of the genomic 5'UTR overlapping with the start codon (342-344) and a part of the Core-coding region (Lukavsky 2009). SL I of the 5'UTR is not involved in translation, while the apical part of the SL II is engaged in a close contact with the 40S subunit inducing conformational changes (Spahn et al. 2001) important for ribosomal translocation (Filbin and Kieft 2011). The larger domain III consists of branching hairpin stem–loops (IIIabcdef) organized in 3- and 4-way junctions. The latter structure at the basal part of the domain III is referred to as the core of the HCV IRES with a double pseudoknot (IIIf). This region of the IRES is highly conserved both in structure and primary sequence and serves for binding of the platform of the 40S small ribosomal subunit (Spahn et al. 2001; Joseph et al. 2014) and AUG positioning at the 40S (Berry et al. 2010). The middle IIId subdomain interacts with the 18S rRNA, consequently playing an anchoring role for ribosome recruitment (Boehringer et al. 2005). The 4-way junction domain IIIabc, together with loops IIIa and IIIb, is essential for an interaction with eIF3 and recruitment of Met-tRNA$_i$–eIF2 (Fraser and Doudna 2007). The affinity of this binding is highly enhanced once the 40S is already bound to the IRES (Siridechadilok et al. 2005). The domain IV is represented by a small stem–loop structure and a highly conserved downstream coding sequence (Khawaja et al. 2015) and provides the AUG initiation codon for interaction with the ternary complex (Ji et al. 2004).

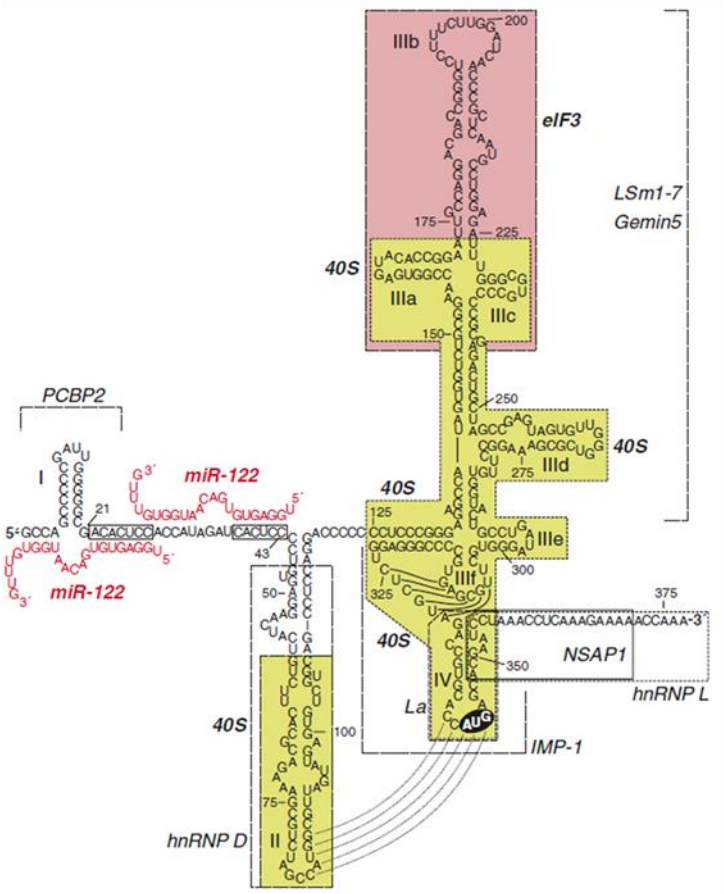


Figure 1.2.3: Secondary structure of IRES-containing HCV 5'UTR.
The HCV IRES consists of three domains II-IV, subdomains within the domain III are indicated IIIa-IIIf. The elements involved in the 40S binding are highlighted in yellow; the binding site for the eIF3 is highlighted in red. Other boxes indicate binding sites for IRES *trans*-acting factors (ITAFs), and binding of two miR-122 molecules (in red) at the 5'-end is depicted.
(from Niepmann 2013)

1.2.2.2 Model of HCV IRES translation initiation

The HCV IRES mediated translation initiation is remarkably different from canonical eukaryotic mRNA translation in two major aspects: 7-methyl guanosine 5'-cap recognition and scanning are not required for the 48S complex formation and only a subset of four canonical eIFs is sufficient (Kieft et al. 2001; Otto and Piglisi 2004). Instead, in a biphasic way of association, at first, the high-affinity interaction of the structured IRES element with the 40S subunit takes place followed by the second slow phase. The current binding model is based on the IRES flexibility and involves repositioning of the domain II (Fuchs et al. 2015) that further plays a role in tRNA recruitment and the transition to elongation (Kieft et al. 2001; Lukavsky 2009). Moreover, ribosomal toeprinting has demonstrated the direct positioning of the initiation codon in the mRNA binding cleft of the 40S ribosomal subunit upon its recruitment to the pseudoknot (Berry et al. 2010); therefore, no initiation factors or ribosome scanning are needed at this step. Next, the positioning is followed by the binding of eukaryotic translation initiation factor 3 (eIF3) and the eIF2/GTP/Met-tRNA$_i$ ternary complex that stabilizes the pre-initiation translation complex assembly (Ji et al. 2004). In the resulting 48S complex the tRNA is positioned in the P site of the 40S subunit and base-paired to the initiation codon. Subsequent eIF5-catalyzed GTP hydrolysis leads to a release of the initiator tRNA and eIF2 dissociation. The second GTP hydrolysis step involving the initiation factor eIF5B promotes association of the 60S ribosomal subunit to form the 80S initiation complex capable of directing viral protein synthesis (Ji et al. 2004; Fraser and Doudna 2007; Khawaja et al. 2015) (Fig. 1.2.4). Upon cellular response to viral infection, the activity of eIF2 can be reduced via eIF2α subunit phosphorylation in order to inhibit viral protein synthesis. Under these stress conditions the HCV IRES can use eIF2-independent pathways to generate the 80S ribosomes. One of the possibilities is to utilize an initiator tRNA-binding protein eIF2A that under stress conditions directly interacts with the domain IIId of the HCV IRES committing the recruitment of Met-tRNA$_i$ to the P site of the 40S (Kim et al. 2011). Similarly, eIF2D protein

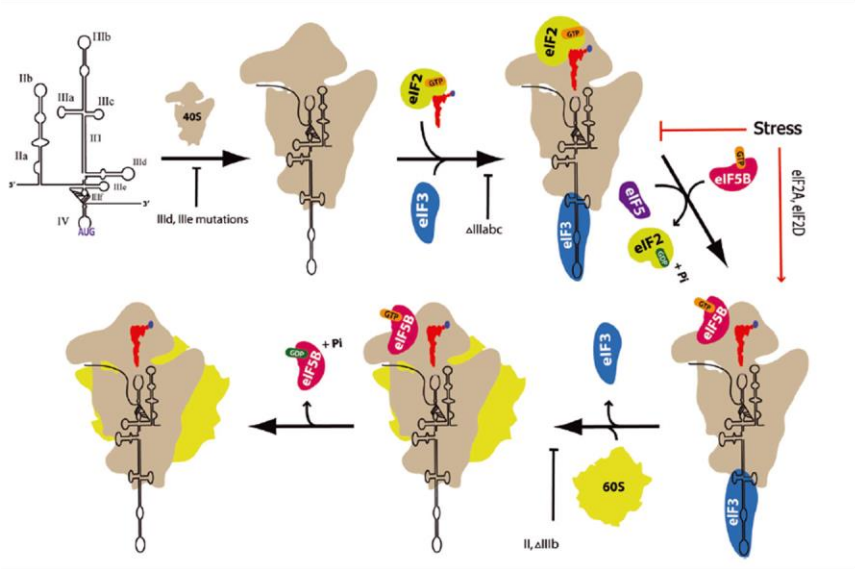


Figure 1.2.4: The mechanism of the HCV IRES translation initiation and its modulation.
The HCV IRES binds the 40S subunit (in brown), then eIF3 (in blue) and the ternary complex (eIF2/GTP/Met-tRNAi) and finally, after GTP hydrolysis and eIFs release, the 60S subunit (in yellow) joins to form the 80S ribosome. Mutations disabling particular steps of translation complex formation are shown under the arrows. An alternative eIF2-GTP-independent pathway is indicated by a red arrow.
(from Khawaja et al. 2015)

was reported to facilitate Met-tRNA$_i$ assembly with the 40S under stress conditions (Dmitriev et al. 2010). Alternatively, IF2-like protein eIF5B may assist in this initiator tRNA recruitment process and/or the subsequent 60S ribosome-joining step (Terenin et al. 2008).

1.2.3 Modulation of HCV translation initiation

The ability to initiate translation in the absence of many of the canonical initiation factors reflects the significance both of the IRES tertiary structure and primary nucleotide sequence that in part makes the IRES sensitive to mutations or inhibitors targeting the key binding sites. A number of studies have demonstrated an importance of the conservation and specificity of domains II-IV that are used at particular steps for interactions with the cellular translational machinery.

A high-affinity interaction with the 40S small ribosomal subunit requires binding surfaces formed by 4-way junctions between domains IIIabc and domains IIIef with the crucial importance of domain IIId (Kieft et al. 2001; Spahn et al. 2001; Joseph et al. 2014). However, neither deletion of the domains II and IV interacting with the 40S subunit nor deletion of the domain IIIb (part of the eIF3 binding site) affect the 40S binding affinity (Kieft et al. 2001; Otto and Piglisi 2004). Incorporated in a 3-way junction, domain IIId contains a hairpin loop (5'- 264UUGGGU269 -3'). The GGG triplet in its apical loop was found to be indispensable for the recruitment of the 40S (Lukavsky et al. 2000; Kieft et al. 2001). RNA aptamers with a consensus loop sequence of ACCCA, which is complementary to the apical loop of domain IIId, abrogated the binging of the 40S (Kikuchi et al. 2005). In particular, a GGG to CCC mutation in the IIId loop at positions 266–268 was demonstrated to reduce the 40S binding affinity by 25-fold (Kieft et al. 2001; Ji et al. 2004). Moreover, the deletion of the domain IIId or mutations in the IIId apical loop cause a low binding capability towards eIF2A, thus affecting translation activity under stress conditions (Kim et al. 2011). Crucial importance of a tetraloop (5'- GA[U/C]A -3') within the domain IIIe was also pointed out in several studies. Any deletions or mutations in the IIIe tetraloop strongly reduce both the 40S binding affinity and translational activity (Psaridi et al. 1999).

After the IRES association with the 40S ribosomal subunit, eIF3 mediates ternary complex binding and Met-tRNA$_i$ positioning in the P site. The binding affinity of eIF3 to the entire HCV IRES equals the one to an isolated domain IIIabc (Kieft et al. 2001), indicating that the 4-way helical junction IIIabc serves as a binding platform for eIF3. Deletions of both domains IIIabc and domain IIIb alone were demonstrated to strongly impair eIF3 binding and to obstruct translation initiation at the IRES-40S complex stage. In the absence of the IIIb region this effect is mainly achieved due to a lowered thermodynamic stability of complexes with eIF3 and of eIF2 with the 48S complex and the deposition of Met-tRNA$_i$ to the AUG binding site (Ji et al. 2004; Otto and Piglisi 2004). Domain IIIb consists of an apical loop and an internal

loop of variable sequence that is nevertheless characterized by a conserved three-dimensional secondary structure. Structural studies indicated that the secondary structure of the IIIb rather than conservation of individual bases is recognized by eIF3, and its recruitment depends on interactions within the properly folded IIIabc junction (Collier et al. 2002; Lukavsky 2009).

Another reported target for a disruption of the HCV IRES translation is subdomain IIb. It is constituted by an apical loop and a loop E motif characterized by a high degree of sequence conservation (Lukavsky et al. 2003). Chemical probing experiments underline an importance of intact domain IIb for the IRES-40S complex arrangement at the AUG start codon. Mutations in the apical loop of domain IIb, disturbing an interaction with rpS5, affect the 40S subunit conformation in a way that the most of mutant-bound ribosomes stall at the start site. Nonetheless, it does not lead to a complete abrogation of the translation-competent 80S complex formation. Since the ribosome samples conformations, an ability to translocate can be occasionally acquired allowing for the functional 80S ribosome formation (Filbin and Kieft 2011; Filbin et al. 2013).

1.3 HCV replication: state-of-the-art

1.3.1 Models for HCV replication study

The establishment of functional HCV replicons in cell culture (Lohmann et al. 1999) had an indispensible impact on the beginning of the era of *in vitro* studies on HCV RNA replication. Previously available infectious cDNA clones (Kolykhalov et al. 1997; Yanagi et al. 1997) suffered from a number of shortcomings such as the low efficiency in cell culture and the need for a controversial chimpanzee model. The novel approach was based on the transfection of cloned viral consensus genome sequences constructed of the HCV 5'UTR, NS3-NS5B and the 3'UTR derived from total liver RNA isolated from a chronically HCV-infected patient (genotype 1b). In order to select only for the replicon-containing cells, a neomycin phosphotransferase (*neo*) gene was included in replicon constructs in a bicistronic manner under control of the HCV-IRES (Fig. 1.1.4). Following transfection into the HuH-7 cells, neomycin selection and validation of replicon population, several autonomously replicating HCV RNA clones were isolated. Corresponding mutated constructs with an in-frame deletion of the NS5B polymerase active site served as a negative control (Lohmann et al. 1999). However, only a small number of efficiently replicating clones was consistently obtained that indicated a gain of adaptive mutations. Indeed, in the later studies these replication-enhancing mutations (REMs) were identified within nearly every HCV non-structural protein (but neither in 5'- nor in 3'- UTRs), although the level of adaptation was very variable. The most adaptive mutation in the study - 2884Gly in NS5B - was found to be conserved in all replicons derived from one cell line (Lohmann et al. 2001).

The further characterization of conserved mutations and their combinations revealed a dual adaptation nature. The mutations in NS4B, NS5A and NS5B were found to be highly adaptive, but incompatible with each other, whereas mutations within the N-terminus of NS3 displayed only a low or no effect on HCV replication efficiency, however cooperatively increased replication efficiency when combined with highly adaptive mutations and each other (Lohmann et al. 2003). A cluster of REMs, in particular an amino acid substitution Ser2204Ile, was found within NS5A, however an outstanding replication facilitation was observed only in about 10 % of transfected HuH-7 cells (Blight et al. 2000; Lohmann et al. 2003) that highlighted an importance of host cell environment. Significant variations in HuH-7 cells permissiveness were noted even among individual passages of the same stock, speculating on variations in the abundance and/or activity of cellular factors engaged in the HCV RNA replication (Lohmann et al. 2003). Suggesting a cellular environment maintaining a persistent replicon to be optimal, removal of the replicon from such adapted cells would create a perfect host cell for transient replication assays.

Consequently, several HuH-7 lines harboring subgenomic HCV replicons were cured of the HCV RNA by prolonged treatment with IFN-α: two of the clones (HuH -7.5 and -7.8) harbored replicons without any REMs and one contained a substitution Ser2204Ile (HuH-7.4). After a complete elimination of HCV

replicons and re-electroporation with HCV RNA, the first two clones demonstrated a significantly higher number of G-418 resistant colonies than the parental HuH-7 cells and the HuH-7.4 clone. The HuH-7.5 line appeared to be the most permissive of those tested, although curing of other replicon-containing cells did not always result in highly permissive lines (Blight et al. 2002). Further investigation of IFN-treated cells revealed that a mutational inactivation of RIG-I mediated cellular permissiveness to HCV replication. RIG-I is an interferon-inducible cellular RNA helicase that serves as a receptor to transduce signals to activate innate antiviral defenses. Upon HCV infection the helicase domain of RIG-I binds structured genomic RNA and dsRNA and activates interferon regulatory factor 3 (IRF3), thereby inducing interferon production and expression of IFN-stimulated genes (ISGs) in cells. This pathway was found to be defect in cells selected for permissiveness for the HCV RNA replication (Stumpler et al. 2005).

From the early 2000s, even before an establishment of the first cell culture infectious HCV system in 2005, replicon systems provided a valuable platform for identification of the role of HCV genome segments and proteins essential for RNA replication as well as of virus-host interactions and candidate therapeutic agents and direct-acting antivirals (DAAs) (Bartenschlager et al. 2013). Already the first replicons defined that the components of the replication module (NS3-NS5B), unlike the assembly module (C-NS2), were sufficient for their autonomous replication in the cell culture (Lohmann et al. 1999).

1.3.2 *Cis*-acting RNA elements

Enabled by the use of the HCV replicon system, a comprehensive mapping of the HCV 5'- and 3'-UTRs revealed basic requirements for the efficient RNA replication (Fig. 1.3.1), in particular for the initiation of minus strand synthesis. With regard to the 5'-end, a 40-nt deletion of the very 5'-terminus entirely abrogated replication, however only to a certain extent affected translation. The first 125 nt were found to be sufficient for the HCV RNA replication with a significant improvement of the efficiency when further 5'-end sequences (up to the complete 5'UTR) were introduced (Friebe et al. 2001). At the 3'-end the variable region and polyU/UC tract were found to be conditionally essential for replication: a complete deletion of the variable region was tolerated and resulted only in reduced replication efficiency, whereas at least a 26-nt homouridine stretch was required. Remarkably, neither point mutations nor stem-loop deletions within the X-tail were viable, proposing its crucial importance for efficient RNA replication (Friebe and Bartenschlager 2002). Intolerance to mutations in the latter region (Yi and Lemon 2003a) and a strong preference for the uridine residue at the very 3'-end (Yi and Lemon 2003b) indicated the significance of both structure and sequence of the X-tail in the HCV replication.

Identified in a number of previous studies, six highly structured RNA elements in the NS5B region were challenged by site-directed mutagenesis of a subgenomic HCV replicon to determine their role in replication (You et al. 2004). When the mutagenesis strategy was disrupting the predicted RNA secondary

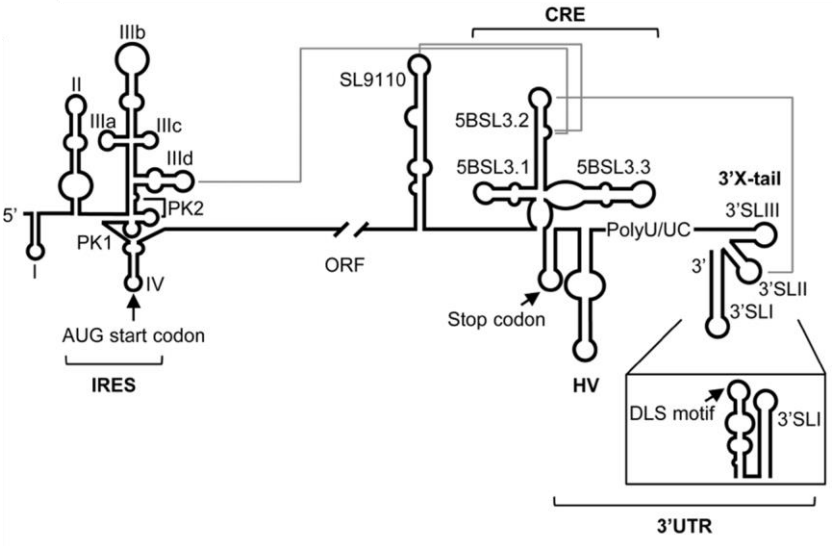


Figure 1.3.1: Secondary structures at the 5'- and 3'- ends of the plus strand HCV RNA engaged in long-range interactions.
Validated conformational structures of the IRES domains at the 5'-end and of the HCV NS5B-coding region and the 3'UTR at the 3'-end are depicted. RNA-RNA interactions between the IRES, the SL9110, the 5BSL3.2 (CRE) and the X-tail are represented by solid lines.
(from Romero-López et al. 2014)

structures (retaining the amino acid sequence of NS5B) solely mutations in 5BSL3.2 (Fig. 1.3.1) caused replication abrogation. This structural element embedded in a cruciform-like structure was defined to act as a *cis*-acting replication element (CRE) (designated variously as SL3.2, SL-V or lately SL9266). Further in-depth analysis verified an importance of both loops and the upper stem in replicon viability (You et al. 2004, Friebe et al. 2005). Separation of a genetic role of the 5BSL3.2 element from its role in replication, performed by translocation to the variable region in the 3'UTR, allowed rescuing the efficient replication and demonstrated an importance of the secondary structure rather than the primary sequence. Additionally, a complementarity between CACAGC motif within the apical loop of the 5BSL3.2 and the loop sequence in the SL2 within the X-tail provided genetic evidence for a mid-range kissing-loop interaction: abrogation of RNA replication by mismatches in the loops was rescued when complementarity was restored (Friebe et al. 2005). The subterminal 8-nt bulge at the 5BSL3.2 was discovered to establish two additional long-range interactions: one with the apical loop of the SL9110 located upstream the 5BSL3.2 (so-called Alt sequence), shown to be likewise crucial for replication (Diviney et al. 2008), and another interaction with the apical loop of the subdomain IIId of the IRES (Shetty et al. 2013) (see Fig. 1.3.2). Due to a mutually exclusive nature of these long-range interactions, the 5BSL3.2 was hypothesized to serve as a molecular switch between translation and replication (Diviney et al. 2008; Tuplin et al. 2012; Shetty et al. 2013). Intriguingly, the 5BSL3.2 element was shown to inhibit IRES-dependent, but not cap-dependent translation, in particular deletion of the domain 5BSL3.2 was associated with at least 5-fold increase in the IRES activity (Romero-López et al. 2012). These results are in agreement with a complex interplay between the IRES SL IIId, SL9110, bulge or loop of the 5BSL3.2 and SL2, since mutually exclusive RNA-RNA interactions are formed (see Fig. 1.3.2). Consistently with the enhancement of the HCV translation by the 3'UTR only in *cis* (Ito et al. 1998), the 5BSL3.2 regulatory activity of the IRES function was as well only achieved in the presence of the 3'UTR in *cis* (Romero-López et al. 2012). Conversely, more complex model

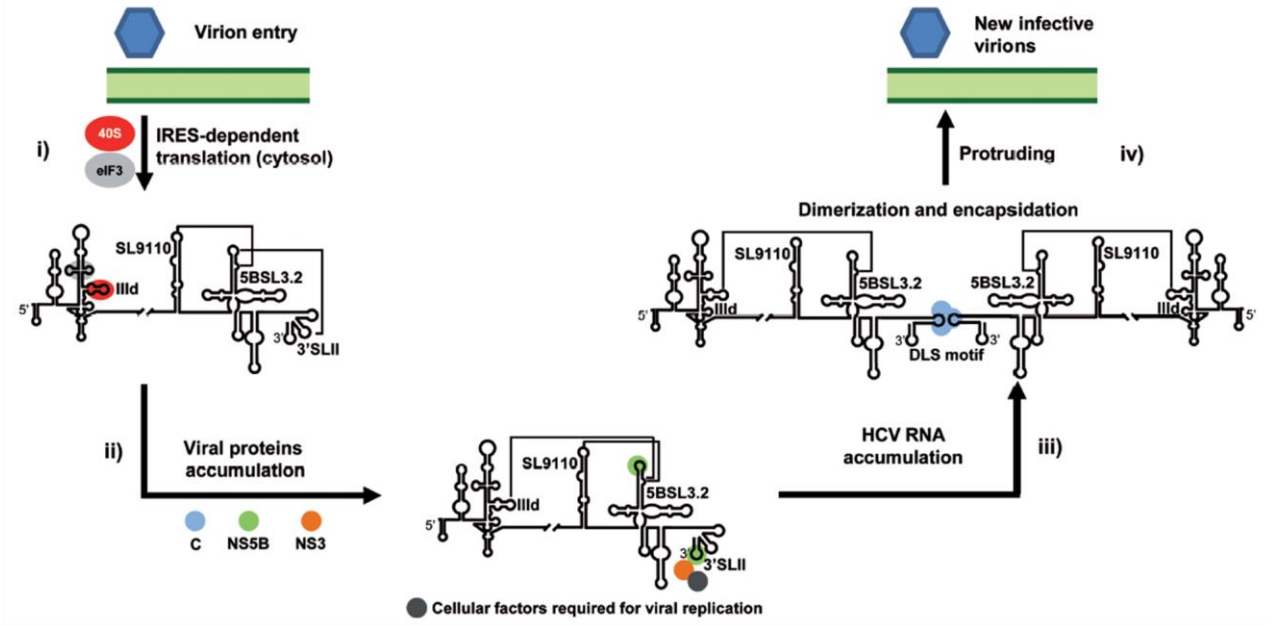


Figure 1.3.2: A model of the HCV IRES, 5BSL3.2 and 3'UTR intercommunication and its contribution to switching between the steps of the viral life cycle.
(i) After the virus internalization, translation of the polyprotein takes place in the cytoplasm and the IRES is occupied by the 40S and eIF3; therefore SL9110-5BSL3.2-SL2 interaction at the 3'-end is favored.
(ii) As a result of translation, viral proteins accumulate at the ER and eventually promote membrane alterations leading to formation of the replication sites. At this stage replication complex assembling at the 3'UTR impairs SL2-based long-range interactions, therefore allowing an establishment of the SLIIId-5BSL3.2 interaction (the SL9110-5BSL3.2 interaction is retained) that interferes with the HCV translation and conducts the switch to genome amplification.
(iii) Local accumulation of the HCV RNA promotes structural switch of the X-tail into the two stem-loops conformation possibly leading to dimerization of genomes via the dimerization linkage sequence (DLS).
(iv) Interaction of dimer forms with the HCV Core proteins further conducts the switch to genome encapsidation.
(from Romero-López et al. 2014)

suggests that structural rearrangements between open and closed conformation of the 5BSL3.2 element enhance and reduce the HCV translation, respectively (Tuplin et al. 2015), sustaining its function as a molecular switch (Fig. 1.3.2). The study demonstrated that both inactivation of the apical loop of the 5BSL3.2 and the apical loop of the 3'UTR SL2 by LNA antisense oligonucleotides interferes with translation. It emphasizes stimulatory function of the 5BSL3.2(loop)-SL2 "kissing-loop" interaction that comprises the closed conformation of the 5BSL3.2 (Tuplin et al. 2015). Nevertheless, neither mutagenesis nor biophysical methods have yet confirmed the formation of the 5BSL3.2 element cruciform and dynamic nature of the long-range RNA-RNA interactions *in vivo* (Romero-López and Berzal-Herranz 2017).

More detailed understanding of the *cis*-elements' intercommunication was obtained using SHAPE (selective 2'-hydroxyl acylation and primer extension) analysis that enables mapping of both local and long-range interactions in RNA molecule. An intricate interplay of long-range RNA-RNA interactions was proposed to operate between IRES, 5BSL3.2 and 3'UTR. The existence of the complexes SL9110-5BSL3.2-SL2 and IIId-5BSL3.2 was verified by several techniques in replication-competent RNA transcripts. Tertiary conformations of the IRES (domains IIIb, IIId and IV), the 5BSL3.2 and the X-tail of the 3'UTR were found capable of mutual regulation and reorganization regardless viral or cellular proteins (Romero-López et al. 2012; Tuplin et al. 2012; Romero-López et al. 2014).

With regard to the structure of the X-tail, a previously predicted co-existence of two stable conformations (Ivanyi-Nagy et al. 2006) seems to play a key role in switching between different functions during the HCV life cycle. In more detail, both predicted conformations retain the SL1 domain as a site of the primer-dependent minus strand synthesis initiation, whereas the upstream sequence forms either two canonical stem-loops 2 and 3 (Blight and Rice 1997) (Fig. 1.1.3) or a single stem-loop exposing a so-called dimerization linkage sequence (DLS) (Fig. 1.3.1). Located within an apical loop of this alternative stem-loop, the DLS region is a palindromic RNA motif that was proposed to enable formation of genomic RNA homodimers in the presence of the Core protein *in vitro* (Shetty et al. 2010).

The IRES and the 5BSL3.2 element were shown to perform a structural reorganization within the X-tail by switching to the two stem-loops conformation containing the DLS motif that in turn might be related to the switch between viral replication and genomic dimerization (Romero-López et al. 2014). A homodimer form may in turn be an intermediate conformation facilitating a switch between a process of genome replication and encapsidation. Indeed, viruses often establish a dynamic network of conserved interactions between functional genomic RNA domains to implement regulation of different steps of viral life cycle with the minimal requisite for proteins (Shi and Suzuki 2018).

For instance, direct RNA-RNA interactions between the 5'- and 3'- ends of some Flaviviruses (such as Dengue virus and Yellow fever virus) induce genome circularization mediating initiation of the RNA replication (Villordo and Gamarnik 2009). In contrast, circularization in Poliovirus requires RNA-protein-RNA interactions fulfilled by PCBP2 and PABP proteins bound to the 5'- and 3'- genomic ends, respectively (Herold and Andino 2001). Any definite proof whether a circularization of the HCV genome takes place and/or regulates RNA synthesis remains missing, however, indirect evidence and computational predictions allow to take it into account when speculating on a possible model of HCV replication (discussed below).

1.3.3 Viral and cellular determinants

Translated directly in the cytoplasm, HCV genome gives rise to a polyprotein that is processed by viral and cellular proteases into structural and non-structural proteins (detailed in 1.1.2). While the HCV translation fully relies on the host cell machinery, the viral replication complex is formed by encoded NS3-NS5B viral proteins. Embedded into the ER, multiple copies of the NS proteins induce typical for Flaviviruses intracellular membrane rearrangements that form vesicle-like structures referred to as "membranous web". Combined immunofluorescence analysis of the HCV NS proteins and metabolic labeling of newly synthesized viral RNA demonstrated their co-localization, therefore postulating the membranous web as a "replication factory" (Gosert et al. 2003) (Fig. 1.3.3).

1.3.3.1 Viral components of the HCV replication complex

In contrast to well-characterized individual NS3-NS5B proteins, the molecular architecture of the replication complex remains unclear. The multifunctional NS3 protein requires the co-factor NS4A anchoring the heterodimer in the membrane and mediating its activities: the N-terminal serine protease domain of the NS3 catalyzes processing of the downstream polyprotein and the C-terminal RNA helicase/NTPase domain is essential for RNA replication (Beran et al. 2009). NS4B is an integral membrane protein that induces formation of the membranous web via oligomerization and interaction with other NS proteins and viral RNA (Paul et al. 2011). Anchored by its N-terminus, serine phosphoprotein NS5A associates in homodimers to implement its multiple activities: the domains 1 and 2 are primarily involved in phosphorylation-dependent modulation of the HCV RNA replication, whereas the domain 3 serves viral assembly (Ross-Thriepland and Harris 2015). The viral RNA-dependent RNA polymerase (RdRp) NS5B is anchored by its C-terminal tail and harbors features that are canonical for all RdRp subdomains: fingers, palm and thumb of a right hand (Moradpour et al. 2004; Bartenschlager et al. 2010; Moradpour and Penin 2013). Recently revised mutation rate of the HCV replicase *in vivo* was computed to reach on average $2.5*10^{-5}$ that is 1.3 nucleotides per each round of replication (Ribeiro et al. 2012).

The formation of a functional replication complex has various *cis*- and *trans*- requirements for the NS HCV components. In transient complementation studies, a co-transfection of helper and various mutant replicon RNAs identified that *trans*-complementation was limited to certain lethal NS5A mutations, however only an expression of the minimal coding sequence of NS3 to NS5A, and not NS5A alone, from a helper RNA rescued a defective replicon RNA (Appel et al. 2005). Moreover, replacement of the HCV 5'- or 3'-UTRs by the sequences from heterologous genotypes dramatically impaired efficiency of the HCV plus or minus strand synthesis, respectively. In particular, a combination of the NS3 helicase, NS5A and NS5B with the homologous 5'UTR was required as well as a combination of the homologous X-tail and NS5B polymerase (Binder et al. 2007). Interestingly, requirements for the virion assembly appeared somewhat different from the ones for replication: *trans*-expression of NS5A alone was sufficient to complement defective mutations indicating an existence of a minor virion assembly pathway that bypasses the need for RNA replication (Herod et al. 2014). Further identified, strict requirements for the HCV NS3 helicase/NTPase activities in *cis* suggest its role in recruitment of the RNA template for replication. Complementation group analysis, conducted in the same study, identified that the NS3-NS4A and NS4B activities are coordinated, while the RdRp activity is dependent on the complete set of the upstream NS3-NS5A genes. Although NS5B can perform binding and polymerase activity when provided in *trans*, its expression is required in *cis* underlining a structural role in replication complex assembly (Kazakov et al. 2015). Common for other members of the *Flaviviridae* family, examples of *cis*-preferences represent a mechanism for genome quality control.

1.3.3.2 Organization of the HCV replication site

Biogenesis of the membranous web is induced by a coordinated action of the HCV non-structural proteins together with several host factors. The three dimensional organization of HCV-altered membranes was found to resemble those of Corona- or Picorna- viruses, rather than those of the related Flaviviruses (e.g. Dengue virus), adopting double- or multi- membrane configuration (Fig. 1.3.3). An electron tomography and 3D-reconstruction analysis identified that, in contrast to Dengue virus, the predominant membrane species - double-membrane vesicles (DMVs) – are on average 150 nm in diameter and about a half of them is connected by a neck to the ER membrane (Romero-Brey et al. 2012). The narrow connection to cytoplasm on the one hand allows an access of metabolites (e.g. NTPs) and an egress of newly synthesized RNA; on the other hand it provides an environment protective from cellular proteases, nucleases and receptors of innate immunity (Lohmann 2013; Chatel-Chaix and Bartenschlager 2014).

Morphologically the membranous web is a cytoplasmic accumulation of highly heterogeneous membranous vesicles that are embedded into an amorphous matrix (Fig. 1.3.3). Formerly, NS4B was

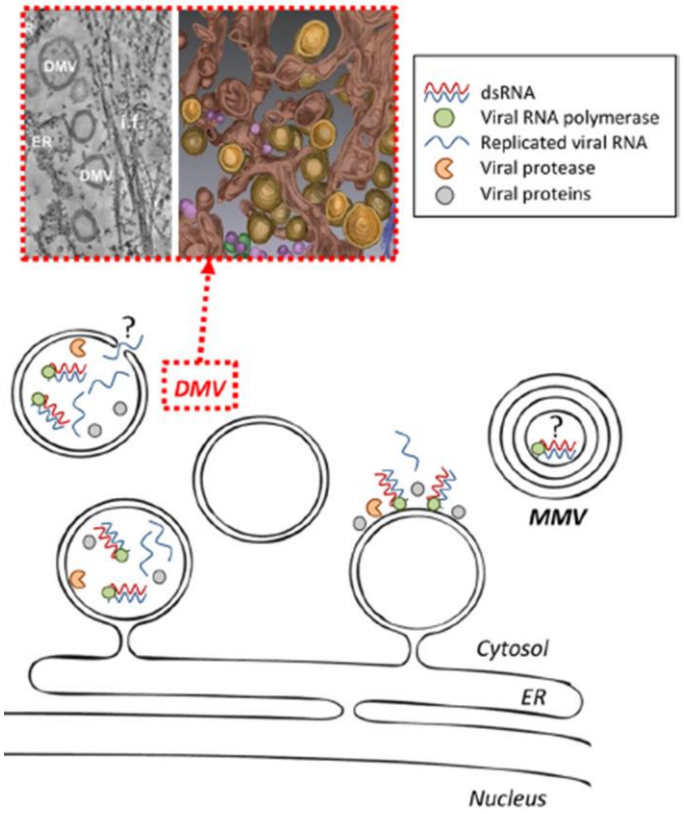


Figure 1.3.3: An electron microscopy image, a 3D-reconstitution from electron tomography and a corresponding schematic representation of the HCV replication sites (referred to as membranous web).
The HCV replication factories are organized within double or multiple membrane vesicles (DMV/MMV) by multiple copies of the HCV non-structural proteins with structural and enzymatic functions (more details are in 1.3.3.2). (from Chatel-Chaix and Bartenschlager 2014)

believed to play a primary role in a biogenesis of the web: its oligomerization induces analogous membrane alterations independently of the other NS proteins (Egger et al. 2002, Paul et al. 2013). Although also NS3-NS4A and NS5B may to a certain extent induce a formation of single membrane vesicles, NS5A is essential for the biogenesis of DMVs (Romero-Brey et al. 2012). Several host factors are the major contributors to the membrane reorganization. Interaction of NS5A and NS5B with phosphatidylinositol-4-kinase IIIα (PI4KIIIα) induces its lipid kinase activity and results in intra-vesicle accumulation of phosphatidylinositol-4-phosphate (PI4P) essential for normal morphology of the membranous web (Reiss et al. 2013). Cyclophilin A (CypA) induces conformational changes in NS5A through its peptidyl-prolyl *cis-trans* isomerase activity facilitating the recruitment of NS5B into the replication complex (Chatel-Chaix and Bartenschlager 2014). Formation of the HCV replication sites demands increased levels of lipids to perform membrane proliferations. In contrast to intact ER membranes, DMVs are enriched in cholesterol and sphingolipids and therefore dependent on cellular lipid-binding proteins, for example, vesicle-associated membrane proteins A and B (VAPA and VAPB) that are engaged in sphingolipid metabolism and vesicle transport (Paul et al. 2013; Li et al. 2015). In ultrastructural studies, lipid droplets (LDs) were found to be abandoned in proximity of the membranous web and play a central role in the HCV RNA replication and assembly coordination (Miyanari et al. 2007). Overall, the membranous web represents a balanced environment and its isolated nature favors both higher local concentration of components and physical separation of different steps of the viral life cycle.

1.3.4 Mechanism and regulation of the HCV RNA synthesis

1.3.4.1 Current model of minus strand synthesis initiation

The HCV RdRp NS5B is a key enzyme for the minus and plus strand viral RNA synthesis. The crystal structure of the NS5B replicase demonstrated that its catalytic domain has a closed conformation accommodating only a single-stranded RNA template and nucleotides for *de novo* initiation (Simister et al. 2009). A recombinant HCV RdRp was shown to be capable of copy-back and primer-dependent as well as of primer-independent *de novo* synthesis initiation *in vitro* (Luo et al. 2000; Zhong et al. 2000). Under special conditions the enzyme was functional on poly(C) and poly(U) homopolymeric templates, indicating that specific *cis*-elements are dispensable (Luo et al. 2000). Nonetheless, these *in vitro* experiments required very

high nucleotide concentrations, not supplied *in vivo*, suggesting that *in vivo* replication by the NS5B replicase is promoted by auxiliary factors.

In vivo the NS5B polymerase is anchored into membranes by the highly hydrophobic C-terminal transmembrane helix connected by a 40-nt linker to the catalytic core (Moradpour and Penin 2013). A catalytic cleft is situated between the fingertips and the linker interacting with a beta flap of the thumb domain and incurs a closed initiation conformation. This orientation of the beta flap and the linker restricts an access to the active site in order to ensure a selective and precise positioning of a 3'-end of genomic RNA together with incoming nucleotides along the finger RNA-binding cleft (Harrus et al. 2010). Several amino acid motifs within the NS5B polymerase were identified to be crucial for its activity. Most of single substitutions within those motifs were demonstrated to significantly reduce or inactivate the enzymatic activity; in particular, changes of the absolutely conserved aspartic acid residue within motif C at position 318 completely abolished the RdRp activity (reflects a well-known GND mutation) (Lohmann et al. 1997).

The *de novo* initiation starts upon binding of the first two priming nucleotides to the 3'-end and dinucleotide primer synthesis; this process is slow and demands high nucleotide concentrations. At the second - rate-limiting - step the dinucleotide serves as a primer for an accommodation of the third nucleotide that requires high GTP concentration and conformational reorganization of the polymerase structure. The latter is mediated by removal of the linker allowing the fingertips to directly contact the thumb; as a result, the catalytic core assumes an open conformation allowing an accommodation of the double-stranded RNA and egress of the elongating product (Harrus et al. 2010; Lohmann 2013). After a switch to an efficient elongation, relatively low nucleotide concentrations are sufficient for a processive synthesis of the entire HCV genome (Jin et al. 2012). The termination of the HCV RNA synthesis is achieved simply when the polymerase reaches the end of the template. After completion of synthesis the minus strand serves as a template for production of multiple copies of nascent plus strands.

While synthesis of the plus strand clearly occurs from an ideal (with a single-stranded overhang) template for *de novo* initiation, the initiation process for the minus strand still raises a number of questions. The 3'-end of the plus strand forming a strong stem structure (Blight and Rice 1997) represents, at least from the first sight, an unsuitable template for *de novo* initiation. Furthermore, the uridine is a strongly favorable residue at the very 3'-end *in vivo* (Yi and Lemon 2003b), whereas *in vitro* the polymerase demonstrates a preference for poly(C) templates (Lohmann et al. 1997). At last, proceeding in the opposite direction protein synthesis has to be prevented during ongoing replication. The above mentioned undeniably suggests a contribution of other viral and cellular determinants for the initiation and regulation of the minus strand synthesis.

1.3.4.2 Plus strand HCV RNA synthesis

The exact sequence and dynamics of HCV RNA synthesis following virus internalization and translation vary in the available cell culture replication models due to differences in cell permissiveness and viral isolates. The data obtained in a single-step growth *in vitro* system using a highly infectious variant of the JFH1 isolate, presented the first minus strands to be generated 4 and 6 h post transfection or infection, respectively, and to reach the maximum at 24-48 h. The HCV minus strand synthesis is limited by the number of replication sites where the genomic RNA is thought to be retained in a double-stranded intermediate form (Targett-Adams et al. 2008). Meanwhile, plus strand production is retarded; therefore the ratio between the strands is about 1:1. The production of the plus strands is boosted after 24-48 hours and leads to the ratio of 10:1 plus strands to minus strands (Keum et al. 2012).

As for the minus strand initiation discussed above, the plus strand RNA synthesis is performed by the NS5B polymerase and initiated *de novo* at the 3'-end of the minus strand (Luo et al. 2000). Numerous *cis*-signals on the HCV genome appeared to be genotype-specific in regulation of RNA synthesis: intergenotypic replicon chimeras harboring a 5'UTR of a heterologous HCV genotype manifested an impaired plus strand synthesis. Additionally this approach identified that the helicase NS3, NS5A, and NS5B are essential

components of an initiation complex for genomic RNA synthesis (Binder et al. 2007). A combination of computational and experimental analysis allowed visualization of the secondary structure of the 3'-end of the minus strand and revealed the lack of resemblance to the antisense sequence of the 5'UTR due to different G:U base pairs. The proposed domain I contains the first 220 nt and folds into five stable stem-loops, whereas the domain II (corresponding to the HCV IRES pseudoknot) adopts a varying secondary structures (Schuster et al. 2002; Smith et al. 2002) (Fig. 1.3.4). The minimal region sufficient for the plus strand synthesis initiation was identified by a genetic mapping of replicon constructs and constituts the first 125 nt of the 3'-end (comprising SL-I' SL-IIz'), however, an optimal replication required 341 nt (Friebe et al. 2001; Kim et al. 2002). As it was confirmed later, SL-IIy' is crucial for the efficient replication while subdomains SL-IIIa', SL-IIIb', SL-IIIcdef' and SL-IV' serve auxiliary functions. Interestingly, folding of the required subdomains appeared to be more important than the underlying primary sequence (Friebe and Bartenschlager 2009).

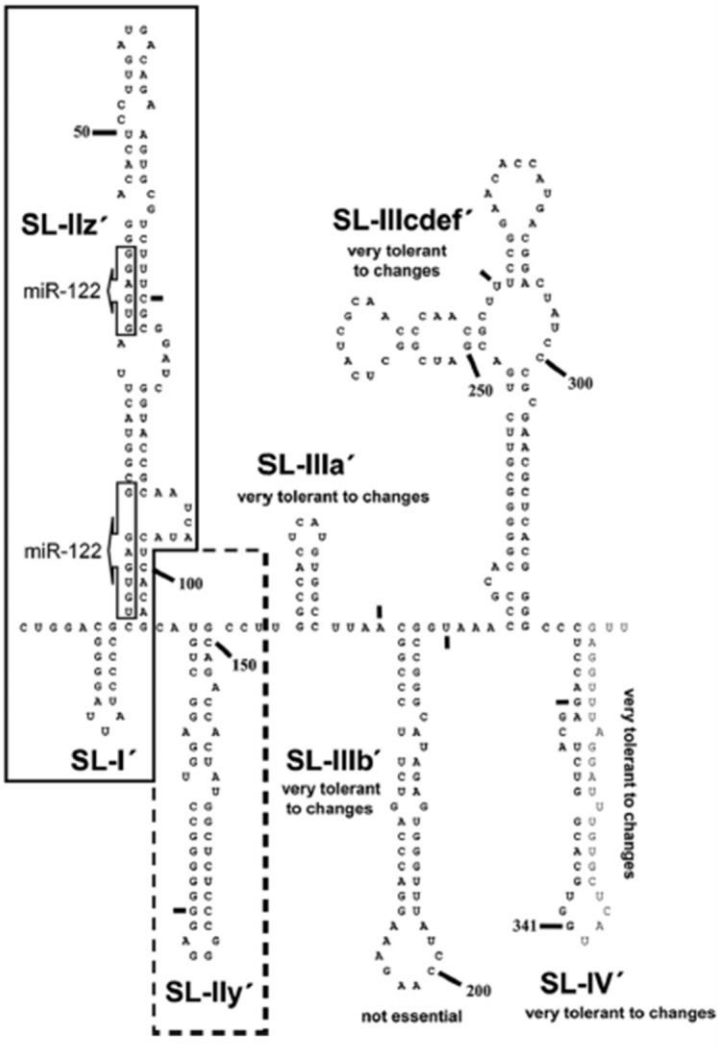


Figure 1.3.4: Predicted secondary structure of the 3'-end of the HCV minus strand RNA (JFH1 isolate).
The miR-122 binding sites within the subdomain SL-IIz' are indicated; the minimal sequence required for the plus strand synthesis (SL-I' and SL-IIz') is highlighted by a box; the SL-IIy' sequence required for efficient replication is indicated by a dashed box (proposed by Smith et al. 2002).
(from Friebe and Bartenschlager 2009)

Partially contrary requirements were derived by biochemical *in vitro* approaches. Various RNA fragments of the 3'-end of the HCV minus strand and their mutated variants were challenged *in vitro* in the presence of viral polymerase. The study demonstrated the high significance of the SL-I', whereas modifications into the SL-IIz' rather improved the RNA synthesis efficiency (Astier-Gin et al. 2005). Further *in vitro* analysis revealed that modifications into the SL IIIb' reduce RNA synthesis catalyzed by recombinant NS5B. In the minigenome system, its intact secondary structure appeared to be required for the efficient HCV minus strand RNA synthesis from the 3'-end by the viral replication complex (Mahias et al. 2010), in contrast to the genetic studies (Friebe and Bartenschlager 2009). Eventually, it

remains elusive whether miR-122 target sites located within SL-IIz' on the minus strand (complementary to the SL I-II) (Fig. 1.3.4) play a role in the plus strand genome replication.

1.3.4.3 Role of microRNA-122 and other host factors

Taking advantage of development of the fully permissive cell culture system (Lindenbach et al. 2005; Wakita et al. 2005; Zhong et al. 2005), transcriptomic and proteomic approaches enabled a comprehensive screening for host factors engaged at different stages of the HCV life cycle *in vitro*. Numerous cellular proteins take part in the formation of the membranous web and therefore indirectly regulate HCV replication, whereas others act directly on the genomic RNA. In a study using a sequence-specific biotinylated peptide nucleic acid (PNA)-neamine conjugate targeted to specifically capture the HCV genomic RNA *in situ*, 83 cellular factors associated with the viral genome were identified (Upadhyay et al. 2013). Multiple studies utilizing a systematic siRNA screening empowered assessment of candidate host factors' effects on viral fitness. For instance, inhibition of the HCV replication through silencing of the components of the RNAi pathway (Drosha, DGCR8, Dicer, TRBP, and Ago 1–4) was a leap forward in understanding of the previously suggested role of miR-122 in HCV RNA synthesis (Jopling et al. 2005; Randall et al. 2007; Wilson et al. 2011; Zhang et al. 2012).

In addition to the host factors described above - PI4KIIIα (Reiss et al. 2013), CypA (Chatel-Chaix and Bartenschlager 2014), VAPA/VAPB (Paul et al. 2013; Li et al. 2015) and LDs (Miyanari et al. 2007) - vital for the replication site organization (see 1.3.3.2), there is a number of other regulatory cellular determinants. Modulation proteins directly binding to the HCV RNA commonly contain multiple RNA-binding motifs and predominantly act at the 5'- and/or 3'- UTRs. Among those, Lupus antigen (La) (Ali et al. 2000; Mondal et al. 2008) and Poly(rC) binding protein 2 (PCBP2) (Wang et al. 2011; Masaki et al. 2015), were reported to oligomerize upon binding to the 5'- and 3'- UTRs, respectively, and act as chaperones for stabilization of the HCV RNA secondary and tertiary structures. Some members of the hnRNP group, such as the polypyrimidine tract-binding protein (PTB) and hnRNP C, were reported to take part in initiation and/or regulation of the HCV replication upon their binding to the 3'UTR (Gontarek et al. 1999). Stabilization of the secondary structure of stem-loops 2 and 3 by PTB is suggested to shift the equilibrium to the three stem-loops conformation of the X-tail (Fig. 1.3.1). Since both PTB and PCBP2 proteins bind with different affinities to the 5'- and 3'- UTRs they were hypothesized to mediate long-range interaction of the genomic ends that may promote the HCV RNA circularization (Shetty et al. 2013). An ability of La protein to interact with both of them may contribute to the same effect upon a protein bridge formation (Spangberg et al. 1999; Fontanes et al. 2009).

Known as a regulator of alternative splicing, EWSR1 protein upon HCV infection colocalizes with replication sites and interacts with the 5BSL3.2 structure, preferentially in the absence of the kissing interaction. This interaction appeared to be required for the HCV replication, but not for initial translation, suggesting the EWSR1 function in switching between open and closed conformations of the 5BSL3.2 element (Oakland et al. 2013). Next, generally localized to the nucleus, cellular proteins of the dsRBM family (NF90, NF110, NF45 and RNA helicase A) were found to accumulate in the cytoplasmic viral replication sites and promote HCV genome circularization (Isken et al. 2007). Moreover, exhibiting a substrate-selective RNA chaperone activity, a complex NF90-NF45 was shown to stimulate the initiation of HCV RNA synthesis (Schmidt et al. 2017). At last, it was recently discovered that a cDNA SEC14L2 (for tocopherol-associated protein 1, TAP1) is essential for the efficient HCV replication. Promoting HCV infection by enhancing vitamin E-mediated protection against lipid peroxidation, it enables RNA replication of all HCV genotypes in culture system without a need for adaptive mutations (Saeed et al. 2015).

The most remarkable host factor modulating almost all stages of the HCV life cycle is microRNA-122 (miR-122). Generally, microRNA (miRNA) duplexes are the post-translational down-regulators of eukaryotic gene expression that - in association with a miRNA-induced silencing complex (miRISC) - repress translation of target mRNA and promote its degradation upon binding to the 3'UTR (Fabian and

Sonenberg 2012). In contrast, miR-122 directly binds to both the 5'- and 3'- UTRs of the HCV genome and displays a positive effect on viral translation (Henke et al. 2008), replication (Jopling et al. 2005) and stability (Shimakami et al. 2012a). Unlike in other tissues, miR-122 is highly abundant in the liver, constituting over 70 % of the total miRNA pool that comprises about 66 000 copies per cell in adult hepatocytes (Lagos-Quintana et al. 2002; Jopling 2012). Encoded in a single genomic locus on human chromosome 18, it is transcribed by RNA polymerase II as pri-miRNA and undergoes a processing by Drosha into pre-miRNA followed by a cytoplasmic cleavage by Dicer, resulting in the 22 bp mature miRNA duplex with extremely conserved sequence (Kim et al. 2009). In hepatocytes miR-122 regulates the expression of many genes involved in fatty acid and cholesterol metabolism and displays a profound tumor suppressor activity (Tsai et al. 2012).

A number of conserved miR-122 binding sites were identified within the HCV genomic RNA (Fig. 1.3.5). Historically, the two tandem binding sites - S1 and S2 - were identified at the HCV 5'-end upstream of the IRES and reported to promote HCV replication (Jopling et al. 2005; Jopling et al. 2008). Subsequently, additional miR-122 binding consensus sequences were revealed: the conserved S3 site located in the otherwise variable region of the HCV 3'UTR and three sites – 5B.1-3 – within the NS5B coding region; their functions remain to be clarified (Nasheri et al. 2011; Gerresheim et al. 2017).

Binding of miR-122 to its target sequence occurs in a following manner. Upon association with argonaute proteins (mainly, Ago2) in the cytoplasm, mature miR-122 duplexes are unwound and one of the strands (the "passenger" strand) is discarded. The "guide" strand induces formation of the miRISC complex where it is positioned to expose its "seed" sequence (GGAGUGU at positions 2-8 of the miR-122) for interaction with a target RNA. An auxiliary binding region (UGGUGUU at positions 14-20 of the miR-122) contributes to base-pairing to various extents (Jopling et al. 2008; Shimakami et al. 2012b; Masaki et al. 2015) (Fig. 1.3.5).

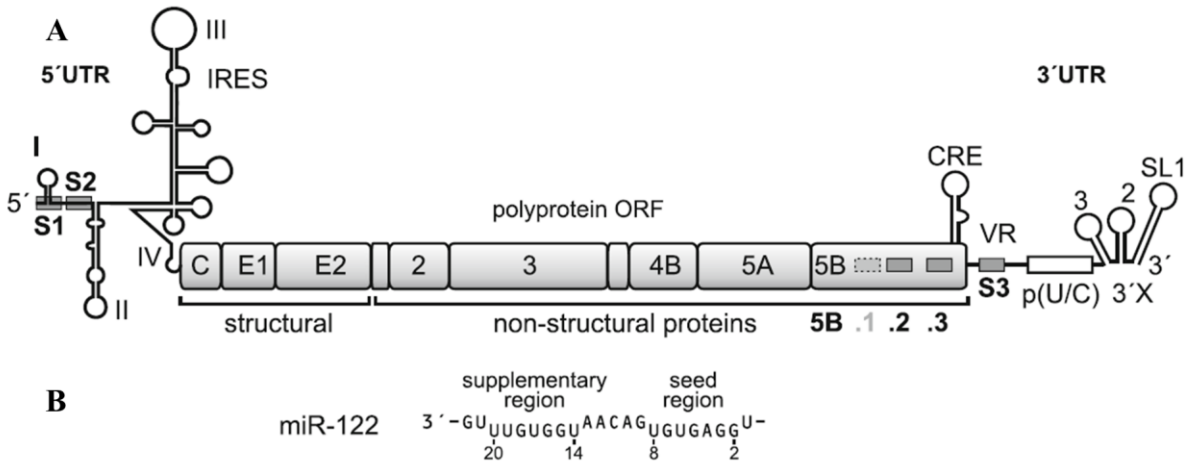


Figure 1.3.5: Distribution of miR-122 binding sites in the HCV genomic RNA.
(A) The HCV genomic RNA with depicted miR-122 target sites: S1 and S2 in the 5'UTR and 5B.1-3 and S3 at the 3'-end (grey boxes). The non-conserved 5B.1 site is indicated by a dashed box.
(B) The sequence of the miR-122 "guide" strand with the seed sequence (positions 2–8) and the potential auxiliary binding region (positions 14-20).
(from Gerresheim et al. 2017)

Effects of miR-122 binding at the HCV 5'UTR are well investigated. In association with Ago2, miR-122 was proven to protect the viral genome from 5'-exonuclease activity of the host mRNA decay machinery, substituting the need for the 5'-capping (Shimakami et al. 2012a). Similarly to the degradation of cellular uncapped mRNAs, the HCV RNA decay is mediated by the cellular 5'-3' exoribonuclease Xrn1 (Li et al. 2013). In addition, miR-122 was demonstrated to protect HCV RNA from degradation by the nuclear exoribonuclease Xrn2, playing rather a minor role in HCV genome decay (Sedano and Sarnow

2014). Notably, the S1 and S2 miR-122 binding sites seem to contribute more to RNA stability and translation stimulation, respectively, but always in a cooperative fashion (Thibault et al. 2015; Nieder-Röhrmann et al. 2017). Indirectly, increased stability of the HCV genome also impacts viral translation and replication.

A direct dose-dependent effect of miR-122 binding at the HCV 5'UTR on translation was demonstrated in several studies. Complementation of miR-122 in non-liver cells (lacking endogenous miR-122) was able to rescue translation of the HCV reporter RNA independently of the viral RNA replication (Henke et al. 2008). The enhancement of infectious virus production was observed when mutated S1 and/or S2 sites were addressed together with compensating miR-122 variants with an impact on both translation and replication (Jangra and Lemon 2010). Ago2-knockdown and co-localization data allowed proposing that miR-122 fulfills its functions at the HCV 5'UTR within an Ago2-containing miRNP complex (Roberts et al. 2011; Wilson et al. 2011; Conrad et al. 2013).

Similarly, stimulation of the HCV replication was revealed to be an effect of direct interaction of miR-122 with the genomic RNA. Severely impaired production of autonomous replicons upon sequestration of endogenous miR-122 or genetic modifications within target sequences was rescued by corresponding compensatory miR-122 variants (Jopling et al. 2005). Sequestration of miR-122 *in vivo* had a suppressive effect on the HCV viremia, paving the way to a promising treatment approach using anti-sense locked nucleic acids (LNA), e.g. Miravirsen (Lanford et al. 2010; Janssen et al. 2013). The phase 2 of the clinical trials on HCV-infected individuals reflected a significant decrease of HCV RNA up to non-detectable level after four-week therapy (Jansen et al. 2013).

The HCV RNA itself was discovered to exert a "sponge" effect, resulting in functional de-repression of host mRNAs that are normally targeted by miR-122 (Luna et al. 2015). Only few studies attempted so far to elucidate a role of the miR-122 binding sites at the 3'-end of the HCV genome. High-affinity binding of miR-122 to these sites was found to be cooperative and result in negative impact on overall HCV production, most likely through translation repression (Nasheri et al. 2011). Recently demonstrated, a dependence of miR-122/Ago2 binding affinity on local HCV RNA structure may suggest a link between these complexes and the current function of the viral RNA, such as translation or replication. Moreover, in opposition to the latter study, miR-122 binding to the 5B.2 site had a positive effect on overall genome replication (Gerresheim et al. 2017). Proposed association of miR-122/Ago2 complexes at the characterized binding sites within the HCV coding region and the 3'-end may support a previous assumption that highly structured nature of these regions serve to pause active translation by the ribosome (Xu et al. 2001). Thus, it is still of high priority to investigate relative contributions of each site to translation, replication and infectious virus production.

1.3.4.4 HCV replication versus translation: possible models

Many processes of the HCV life cycle are tightly associated with each other and regulated by the same set of viral and cellular factors. Ribosomes proceeding from 5' to 3' end during translation inevitably collide with the RNA replicase complex assembled at the 3'-end of the genomic RNA. In order to uncouple those processes, the virus may employ one or more molecular switches at both RNA and protein levels. Although the exact mechanism regulating a transition from translation to replication in the HCV life cycle is not defined yet, a number of clues as well as validated interactions allow proposing possible models.

Since liver-specific miR-122 is a key host factor modulating HCV translation, replication and stability, several studies suggest it to play a central role in altering translation versus replication. As described above, miR-122 was demonstrated to recruit Ago2 protein to the 5'UTR S1 and S2 sites and therefore stimulate IRES-mediated translation (Roberts et al. 2011; Wilson et al. 2011; Conrad et al. 2013; Nieder-Röhrmann et al. 2017) and protect the genomic RNA from Xrn1-mediated degradation (Shimakami et al. 2012a; Li et al. 2013). In Xrn1-depleted cells introduction of exogenous miR-122 enhances HCV replication. In the context of mutated binding sites, corresponding miR-122 mutant variants, however, fail to restore the replication rate

(Li et al. 2013). This indicates a complex function of miR-122 also in other processes tightly linked to RNA synthesis. Previously, it was shown for Flaviviruses that ongoing protein translation is not required for RNA synthesis when essential viral proteins were produced in abundance (Westaway et al. 1999). Interestingly, a replication stimulation effect of miR-122 was found to be dependent on active protein translation (Masaki et al. 2015). In the experiment cellular protein synthesis in Xrn1-depleted cells was blocked by treatment with either cycloheximide (that freezes ribosomes on viral RNA) or puromycin (that releases translating ribosomes from viral RNA). As a result, puromycin treatment led to a short-term stimulation of the RNA replication. Under these conditions miR-122 did not demonstrate its stimulating effect on replication, suggesting that it acts similarly and thus redundantly to puromycin, serving a switch from translation to replication by a non-additive increase of available templates for RNA synthesis. Supposedly, miR-122 may serve to reduce the circularized state of the viral genome by displacing PCBP2 (see below) from the genome RNA, and therefore freeing 3'UTR sequences for interactions with the replication complex (Masaki et al. 2015).

In addition to miR-122, the HCV 5'UTR serves as a binding platform for several cellular proteins, including PCBP2. This factor was reported to facilitate both translation and replication and to mediate HCV RNA circularization upon binding to the 5'- and 3'- UTRs (Wang et al. 2011). Another study demonstrated that PCBP2 competes with miR-122 for binding to the HCV 5'UTR since a major PCBP2 binding site overlaps one of the two functional miR-122 binding sites (Li et al. 2014). Taking this into account, replacement of PCPB2 at the HCV 5'UTR by miR-122 was hypothesized to promote an opening of a circularized form and to conduct a switch from translation to replication (Masaki et al. 2015). Indeed, depletion of PCPB2 from the cells resulted in reduction of translation efficiency, but not replication, and both processes under this condition were not responsive to supplementation of miR-122 (Masaki et al. 2015). Besides PCPB2 oligomerization, a circular form can be stabilized by the PCPB2-La-PTB protein bridge (Fontanes et al. 2009; Shetty et al. 2013). Alternatively, retention of the 40S ribosomal subunits at the HCV 3'UTR enhances translation re-initiation and maintains the circularized form of the genomic RNA that favors a multi-round translation and prevents an assembly of the replicase complex (Bai et al. 2013). Once

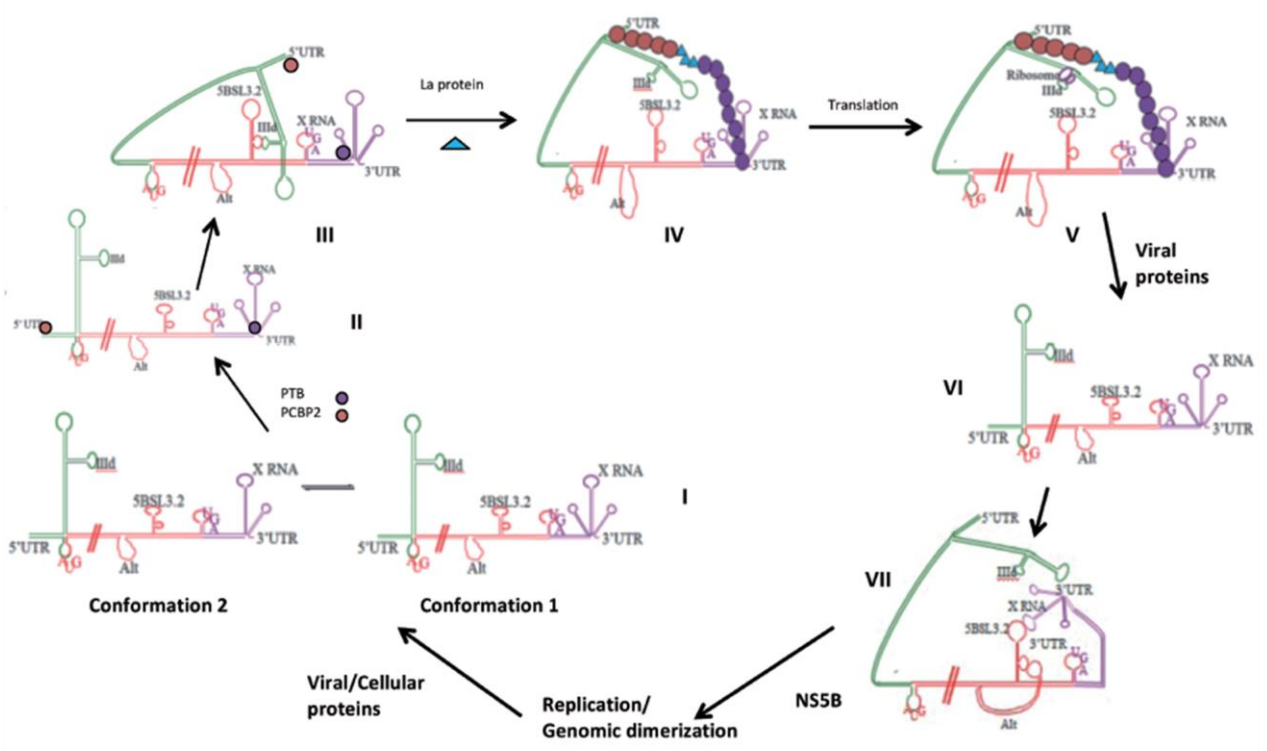


Figure 1.3.6: A speculative model of HCV switch from translation to replication based on biophysical evidences for the RNA-RNA and RNA-protein interactions.
A detailed description is provided in section 1.3.4.3.
(from Shetty et al. 2013)

established by any means, the circularized form of the HCV RNA is maintained and regulated via IRES-5BSL3.2-3'UTR long-range interactions (see 1.3.2). Confirmed by SHAPE analysis, rearrangement of these contacts may mutually regulate interactions of the HCV genomic RNA with cellular factors, therefore tuning protein synthesis versus RNA synthesis.

Taken all together, the following model of the switch between HCV translation and replication can be suggested (Shetty et al. 2013; Romero-López et al. 2014). After viral internalization and release of the HCV genome into the cytoplasm, cellular RNA-binding proteins PTB and PCPB2 occupy its 5'- and 3'- UTRs, respectively, (Fig. 1.3.6, II). At this stage the three stem-loops conformation of the HCV X-tail (Fig. 1.3.6, I, Conformation 1) is stabilized by PTB, thus favoring 5BSL3.2(loop)-SL2 and IIId-5BSL3.2(bulge) RNA-RNA interactions. The latter long-range interaction as well as the assembly of the PCPB2-La-PTB protein bridge brings the HCV 5'UTR in closer proximity of the 3'UTR (Fig. 1.3.6, III and IV). Engagement of the 40S ribosomal subunits disrupts the contact IIId-5BSL3.2(bulge) due to a higher affinity of the 40S subunit to the IRES. Thus, the translation machinery is assembling at the HCV IRES and the SL9110-5BSL3.2(bulge) RNA-RNA interaction takes place (Fig. 1.3.6, V). Protein synthesis initiation and the HCV polyprotein processing result in accumulation of viral non-structural proteins that induce cellar membranealterations and membranous web formation. Accumulating viral protease cleaves PCPB2 and PTB (Fontanes et al. 2009; Wang et al. 2011), leading to disintegration of the protein bridge and linearization of the HCV genome (Fig. 1.3.6, VI). Accumulating Core protein competes with the 40S for interaction with the IRES IIId subdomain, therefore slowing down the translation (Shimoike et al. 2006). In the absence of PTB, a stable kissing-loop interactions SL9110-5BSL3.2(bulge) and 5BSL3.2(loop)-SL2 are established and provide a suitable platform for replicase complex assembly and the HCV minus strand synthesis initiation (Fig. 1.3.6, VII). Accumulation of newly synthesized genomic plus strand RNAs shifts the equilibrium towards the two stem-loops conformation of the X-tail (Fig. 1.3.6, I, Conformation 2) that exposes the DLS motif and favors genome dimerization while disrupting the 5BSL3.2(loop)-SL2 interaction. This conformation slows down the HCV replication and promotes genome packaging in the presence of the Core chaperone protein. Nearly all stages are modulated by miR-122 action. Importantly, within the Ago2 complex miR-122 protects the 5'-end from Xrn1 degradation when the genome is present in linearized from and not shielded by the protein complexes. Decreasing the fraction of viral RNAs engaged in translation and proportionally increasing the fraction available for replication, miR-122 promotes a switch from protein to viral RNA synthesis acting at the 5'- and possibly at the 3'- genomic ends. Many aspects of the HCV replication remain elusive and require experimental validation. It is only clear that regulation of the HCV replication is an intricate interplay between cellular proteins and viral RNA that reciprocally and indivisibly tune the engagement of the RNA in protein translation versus viral RNA synthesis.

The most recent model (Tuplin et al. 2015; Romero-López and Berzal-Herranz 2017) implies the 5BSL3.2 element as a molecular switch and a number of mutually exclusive long-range RNA-RNA and protein-RNA interactions it is engaged in – to balance HCV translation and replication (Fig. 1.3.7). The5BSL3.2 represents a core of a complex pseudoknot structure formed by dynamic long-range RNA-RNA contacts involving structural elements of the HCV IRES, 5BSL3.2 and the 3'UTR (Tuplin et al. 2012). As described above, the three stem-loops configuration of the HCV 3'UTR enables 5BSL3.2(loop)-SL2 interaction and is referred to as a closed conformation (see Fig. 1.3.2, i), whereas the two stem-loops X-tail form – to as an open conformation (see Fig. 1.3.2, ii) that favors genome dimerization (Ivanyi-Nagy et al. 2006). Following endocytosis of HCV virions, the genomic RNA is released to the cytoplasm and associates at the ER. At this stage the HCV IRES, and in particular SL IIId, are engaged in interactions with assembling ribosome, that favors a closed conformation of the 5BSL3.2 pseudoknot (contacts SL9110-5BSL3.2(bulge) and 5BSL3.2(loop)- SL2); see Fig. 1.3.7, "Translation"). Polyprotein translation and its processing, as described above, results in accumulation of the viral replication proteins at the 3'-end of the genome as well as of cellular 5BSL3.2-binding proteins. The latter disrupts the former RNA-RNA contacts and facilitates the open conformation, in which 5BSL3.2(bulge) interacts with domain IIId that in turn brings the genomic ends into a close proximity (see Fig. 1.3.7, "Replication"). A dynamic shift to the IIId-5BSL3.2(bulge) contact

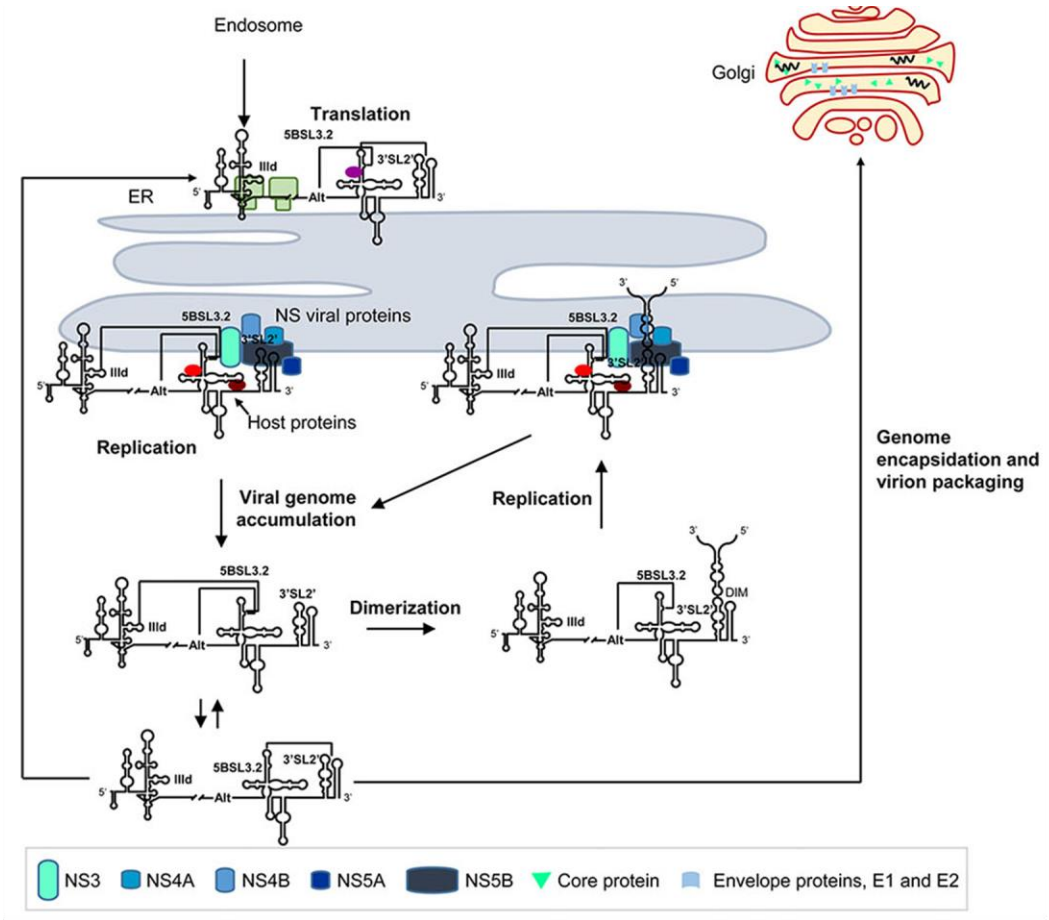


Figure 1.3.7: A recent model of regulation of HCV translation, replication and other steps of the viral life cycle based on conformational rearrangements in essential long-range RNA-RNA interactions and balancing between closed and open conformations of the 5BSL3.2 pseudoknot.
Further details are provided in section 1.3.4.3.
(modified from Romero-López and Berzal-Herranz 2017)

now prevents the ribosome recruitment to the IRES, therefore conducting a switch from viral translation to replication. During accumulation of the HCV genomic RNA thermodynamic equilibrium balances formation of open and closed conformations providing a possibility for new rounds of HCV translation (see Fig. 1.3.7, "Viral genome accumulation"). On the late stages of replication genome dimerization becomes more thermodynamically probable and, mediated by a contact of DLS elements between two genomic RNAs in open conformation of the 5BSL3.2 pseudoknot, it facilities both – further replication and shuffling for encapsidation (see Fig. 1.3.7, "Dimerization").

Both models virtually propose an analogous mechanistic view on the regulation of encountering intracellular stages of the HCV propagation. Evidently, a certain feedback mechanism enables a switch from translation, after sufficient amounts of non-structural proteins have been produced, to replication or when sufficient amounts of genomic RNA require relocation - for packaging. Proposed as such molecular sensors, miR-122 or 5BSL3.2 pseudoknot, inevitably function cooperatively and implicate a complex network of cellular factors.

1.4 Aims of the work

The life cycle of the HCV *in vitro* and especially *in vivo* is certainly more complex than we currently picture it. Virus RNA replication is inevitably linked to polyprotein synthesis and processing, to membranous web formation and to the *cis*/*trans* interplay of *cis*-elements with viral and cellular factors. Existing *in vitro* models are often incapable of addressing specific interactions of a *cis*-element or a host factor serving multiple and overlapping functions at different stages of the HCV life cycle.

Although, the autonomous replicon system and full-length HCV cell culture system have been indispensable for HCV studies, the presence of both 5'- and 3'- genomic ends limits their capacity to assign the exact determinants for each individual step of the HCV life cycle. Assessment of replication efficiency via an overall viral genome amplification or virus production does not provide details on whether a *cis*-element identified within the annotated HCV genome exerts its function through physical presence on the plus or the minus strand HCV RNA. If a sequence or structure of an element appears vital at one or numerous overlapping steps of the virus life cycle, mutations within this element may completely abolish RNA synthesis and/or generation of virus progeny. Thus, mutational analysis attempted in such systems is restricted by retention of efficient translation, minus strand RNA synthesis and plus strand RNA synthesis. Various regions within the HCV genome are known to play direct or indirect roles in multiple molecular processes during virus propagation. In particular, sequences within the HCV 5'- and 3'- UTRs were proven to engage at virtually every step, including replication of the plus and the minus strand as well as translation. Similar complications are encountered in attempt to dissect functions of a certain viral or cellular *trans*-acting factor contributing in several ways to the HCV viral cycle. It can be illustrated by entangled miR-122 interactions with the HCV RNA that contribute to the HCV genome translation, stability and replication upon binding to the target sites at both the 5'- and 3'- ends. Lastly, accumulating reports on *cis*-elements, other than the 5BSL3.2 element, located within the HCV protein-coding region and acting via long-range RNA-RNA interaction, require assignment of their functions to distinct steps of the viral life cycle.

The goal of the present doctoral work is to overcome the above mentioned limitations and to develop a novel minus strand replication system to dissect the requirements for the HCV minus and plus strand RNA synthesis. The system aims at uncoupling overlapping functions of *cis*-elements in viral replication from each other and from possible functions in translation initiation and other steps of the viral life cycle. The main focus of the current work is the analysis of minus strand RNA synthesis uncoupled from all other steps. In particular, revisiting the state-of-the-art requirements for the HCV antigenome synthesis initiation, aims at dissection of the essential *cis*-elements and clarification of their functions, also in scope of the interplay with plus strand RNA synthesis, translation and miR-122 function.

2. Materials and methods

2.1 Materials

2.1.1 Bacterial strains and cell lines

Strain	Company	Genotype
One Shot® TOP10 Chemically Competent E. coli	Thermo Scientific	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

The strain is highly transformable and ideal for stable replication of high-copy plasmids. Along with the usage of commercial stocks, self-made chemically competent E.coli TOP10 was also applied for transformations.

Cell line	Source	Origin
HeLa	Ralf Bartenschlager laboratory; Charles M. Rice laboratory	Human negroid cervix epitheloid carcinoma cells
HuH-7.5		Human hepatocellular carcinoma cells

HuH-7.5 is a derivative of HuH-7 cells generated after removal of the HCV replicon by IFN treatment (Blight et al. 2002). This clone appeared to be highly permissive for the HCV RNA due to a single point mutation in the dsRNA sensor retinoic acid inducible gene I (RIG-I) (Sumpter et al. 2005); the HuH-7.5 cell line was generated by Zhong and co-workers (Zhong et al. 2005).

2.1.2 Materials for bacterial growth and cell culture

2.1.2.1 Materials for bacterial growth

Material	Company
LB-Broth (Lennox)	Roth
Agar-Agar, Kobe I	Roth
SOC Outgrowth Medium	NEB
Ampicillin sodium salt	Roth
Glass spatula	Sigma

2.1.2.2 Materials for cell culture

Material	Company
Dulbecco's Modified Eagle's Medium (DMEM)	Thermo Scientific
100 x Penicillin-streptomycin solution (Pen/Strep; 10 000 U/ml penicillin and 10 000 µg/ml streptomycin)	Thermo Scientific
Fetal Bovine Serum (FBS)	Thermo Scientific
0.5 % Trypsin-EDTA (10 x; 5 g/l Trypsin, 2 g/l EDTA)	Thermo Scientific
Dimethyl Sulfoxide (DMSO)	Roth
Roti-Stock 10 x Phosphate buffered saline (PBS)	Roth
Lipofectamine 2000	Thermo Scientific

2.1.3 Enzymes

2.1.3.1 Restriction endonucleases

Enzyme	Units/µl	Buffer	Recognition sequence	Company
*Afl*II	20	CutSmart	C↓TTAAG	NEB
*Age*I-HF	20	CutSmart	A↓CCGGT	NEB
*Asc*I	10	CutSmart	GG↓CGCGCC	NEB

*Bam*HI-HF	20	CutSmart	G↓GATCC	NEB
*Bbv*CI	2	CutSmart	CC↓TCAGC	NEB
*Bsr*GI	20	CutSmart	T↓GTACA	NEB
*Eco*RI-HF	20	CutSmart	G↓AATTC	NEB
*Eco*RV-HF	20	CutSmart	GAT↓ATC	NEB
*Fse*I	2	CutSmart	GGCCGG↓CC	NEB
*Hin*dIII-HF	20	CutSmart	A↓AGCTT	NEB
*Mlu*I-HF	20	CutSmart	A↓CGCGT	NEB
*Msc*I	5	CutSmart	TGG↓CCA	NEB
*Not*I-HF	20	CutSmart	GC↓GGCCGC	NEB
*Nsi*I-HF	20	CutSmart	ATGCA↓T	NEB
*Sap*I	10	CutSmart	GCTCTTC(N)↓	NEB
*Sbf*I-HF	20	CutSmart	CCTGCA↓GG	NEB
*Xba*I	20	CutSmart	T↓CTAGA	NEB

2.1.3.2 Modifying enzymes

Enzyme	Units/µl	Company
Antarctic Phosphatase (supplied with 10 x Antarctic Phosphatase reaction buffer)	5	NEB
DNase I (RNase-free) (supplied with 10 x DNase I reaction buffer)	2	NEB
T4 DNA Ligase (supplied with 10 x T4 DNA Ligase buffer)	400	NEB
One Taq DNA Polymerase (supplied with 5 x OneTaq standard reaction buffer)	5	NEB
Proteinase K	0.8	NEB
RNase A (10 mg/ml)		Thermo Scientific
SP6 RNA polymerase (supplied with 10 x RNAPol reaction buffer)	20	NEB
T7 RNA polymerase (supplied with 10 x RNAPol reaction buffer)	50	NEB

2.1.4 Molecular biological consumables

Consumable	Concentration	Company
dNTPs (separate solutions of dATP, dGTP, dCTP and dTTP)	100 mM each	Roth
NTPs (separate solutions of ATP, GTP, CTP and UTP)	100 mM each	Roth
Radioactive [α-^{32}-P]-UTP (> 400 Ci/mmol, 10 mCi/ml)		Amersham Biosciences
$MgCl_2$	25 mM	NEB
GeneRuler DNA Ladder Mix		Thermo Scientific
HyperLadder 1 kb		Bioline
GeneRuler 50 bp and 100 bp	0.5 µg/µl	Thermo Scientific
5 x DNA Loading Buffer		Bioline
2 x RNA Loading Dye		NEB
PageRuler Prestained Protein Ladder (10 to 250 kDa)		Thermo Scientific
GlycoBlue	15 mg/ml	Thermo Scientific

2.1.5 Antibodies and beads

Antibody	Description	Company
Anti-Hepatitis C Virus NS3 antibody [8 G-2]	A mouse monoclonal antibody to HCV NS3 that efficiently reacts towards the JFH-1 strain (genotype 2a). Utilized in western blot at 1:500 dilution.	Abcam
Anti-Hepatitis C Virus Core Antigen antibody [C7-50]	A mouse monoclonal antibody to Hepatitis C Virus Core Antigen that recognizes an epitope between amino acid residues 21-40 of the HCV Core protein (conserved among different HCV strains). Used for	Abcam

	western blot at 1:1 000 dilution.	
Anti-GAPDH antibody	A mouse monoclonal antibody to a cytoplasmic housekeeping protein GAPDH (Glyceraldehyde-3-phosphate dehydrogenase). Applied in western blot as a loading control at 1:20 000 dilution.	Sigma
Anti-mouse-HRP antibody and anti-goat-HRP antibody	Horseradish peroxidase conjugated secondary antibodies that react with primary mouse / goat antibodies and allow visualization in presence of the HRP substrate. Both used at 1:20 000 - 1:40 000 dilution.	Antibodies Online
Anti-PTBP1 antibody	A goat monoclonal antibody to a Polypyrimidine tract-binding protein 1 that was utilized as a positive control of binding to the HCV 3'UTR at a working concentration 25 μg/ml.	Abcam
Anti-FLAG antibody	A mouse monoclonal antibody to a FLAG epitope that was utilized as a negative control of binding to the HCV 3'UTR at a working concentration 200 μg/ml.	Abcam
Anti-RARS antibody (ab31537)	A rabbit polyclonal antibody to Arginyl-tRNA synthetase. Applied in RNA immunoprecipitation assays at a working concentration 25 μg/ml.	Abcam
Anti-DARS antibody (ab151974)	A rabbit polyclonal antibody to Aspartyl-tRNA synthetase. Applied in RNA immunoprecipitation assays at a working concentration 25 μg/ml.	Abcam
Anti-QARS antibody (ab72957)	A mouse polyclonal antibody to Glutaminyl-tRNA synthetase. Was applied in RNA immunoprecipitation assays at a working concentration 25 μg/ml.	Abcam
Dynabeads Protein G magnetic beads	Recombinant Protein G (~17 kDa) covalently coupled to the beads' surface allows binding to Fc-region of a wide range of primary antibodies for the purposes of immunoprecipitation.	Thermo Scientific

2.1.6 Kits

Kit	**Purpose**	**Company**
GeneJET Gel Extraction Kit	Purification of DNA fragments from agarose gels	Thermo Scientific
GeneJET PCR Purification Kit	Purification and concentration of DNA fragments from PCR reactions	Thermo Scientific
GeneJET Plasmid Maxiprep Kit	Endotoxin-free plasmid preparation; maxi scale	Thermo Scientific
GeneJET Plasmid Miniprep Kit	Plasmid preparation; mini scale	Thermo Scientific
GeneJET RNA Cleanup Micro Kit	RNA cleanup and concentration after DNase I treatment	Thermo Scientific
qScript Flex cDNA Kit	Reverse transcription of a specific gene from total RNA samples	Quanta Biosciences
PerfeCTa SYBR Green FastMix	Real-time quantitative PCR	Quanta Biosciences
Qubit quantification assay Kits (dsDNA BR, RNA BR)	Quantification of DNA and RNA concentration	Thermo Scientific
SuperScript III Reverse Transcriptase	Reverse transcription of a specific gene from total RNA samples	Thermo Scientific
SuperSignal West Femto Chemiluminescent Substrate	Western blot substrate	Thermo Scientific

2.1.7 Plasmids

Map Number	Plasmid name	Source
6.1.1	pFK-JFH1-J6 C-846_dg_12961 (Jc1)	Kindly provided by Ralf Bartenschlager laboratory (Pietschmann et al. 2006)
6.1.2	pUC18_Plus_strand_backbone_4374	Inserts were chemically synthesized and introduced into pUC18 vector by eurofinsgenomics.eu Scrambling is done by Yutong Song and Steffen Mueller
6.1.3	pUC18_Minus_strand_backbone_4685	
6.1.4	pUC18_Fragment 1_NS5B_SCR_4157	
6.1.5	pUC18_Fragment 2_NS3-NS5B_SCR_5148	
6.1.6	pUC18_Fragment 3_NS3_SCR_4515	
6.1.7	pUC18_P.s_WT_hp_9899	Constructed by L. Shalamova
6.1.8	pUC18_P.s_SCR_hp_9899	
6.1.9	pUC18_Minus_strand_F1_6092	
6.1.10	pUC18_Minus_strand_F1_F2_8524	
6.1.11	pUC18_M.s_WT_hp_10334	
6.1.12	pUC18_M.s_SCR_hp_10334	
6.1.13	pUC18_P.s_WT_SL I-II_wt_9968	
6.1.14	pUC18_P.s_WT_SL I-II_S1mS2m_9968	
6.1.15	pUC18_P.s_WT_SL I-III_10182	
6.1.16	pUC18_P.s_WT_SL I-III_IIId mut_10182	
6.1.17	pUC18_P.s_WT_SL I-III_IIIb del_10136	
6.1.18	pUC18_P.s_WT_5'UTR_sinfr_stop_10242	
6.1.19	pUC18_P.s_WT_5'UTR_sinfr_stop_IIId mut_10242	
6.1.20	pUC18_P.s_WT_5'UTR_sinfr_stop_IIIb del_10196	
6.1.21	pUC18_P.s_SCR_SL I-II_wt_9968	
6.1.22	pUC18_P.s_SCR_SL I-II_S1mS2m_9968	
6.1.23	pUC18_P.s_WT_hp_GND_9899	
6.1.24	pUC18_P.s_WT_SL I-II_wt_GND_9968	
6.1.25	pUC18_P.s_WT_SL I-II_wt_5B.2m_9968	Constructed by N. Dünnes
6.1.26	pUC18_P.s_WT_SL I-II_wt_5B.3m_9968	
6.1.27	pUC18_P.s_WT_SL I-II_wt_S3m _9968	
6.1.28	pUC18_P.s_WT_SL I-II_S1mS2m_5B.2m_9968	
6.1.29	pUC18_P.s_WT_SL I-II_S1mS2m_5B.3m_9968	
6.1.30	pUC18_P.s_WT_SL I-II_S1mS2m_S3m_9968	
6.1.31	pUC18_P.s_WT_SL I-III_GND_10182	Constructed by L. Shalamova
6.1.32	pUC18_P.s_WT_5'UTR_sinfr_stop_GND_10242	
6.1.33	pUC18_P.s_WT_SL I-II_wt_8680mut_9968	
6.1.34	pUC18_P.s_WT_SL I-II_wt_9170mut_9968	
6.1.35	pUC18_P.s_WT_5'UTR_Core_10779	Constructed by G. Gerresheim
6.1.36	pUC18_P.s_WT_5'UTR_Core_GND_10779	
6.1.37	pUC18_P.s_WT_5'UTR_Core_IIIb del_10733	
6.1.38	pHCV-SIN_3235	AG Niepmann plasmid collection
6.1.39	pHCV-3'UTR only_3571	

2.1.8 Oligonucleotides and primers

All DNA and RNA oligonucleotides were supplied by Biomers. Locked nucleic acids (LNA) mixmer oligonucleotides were supplied by Exiqon; a "+" indicates a respective LNA residue.

2.1.8.1 RNA oligonucleotides

Name	Sequence (5' - 3')	Source
miR-122 mat	(phos)UGGAGUGUGACAAUGGUGUUUG	Conrad et al. 2016

miR-122*	(phos)AACGCCAUUAUCACACUAAAUA	
miR-122 mat S1m	(phos)UGUAGUCUGACAAAGUCGUUUG	Nieder-Röhrmann et al. 2017
miR-122* S1m	(phos)AACGACUUUAUCAGACUCAAUA	
miR-122 mat S2m	(phos)UGGAUAGUGACAACUGUGUUUG	
miR-122* S2m	(phos)AACGCAGUUAUCACUAUAAAUA	
miR-122 mat 5B.2m	(phos)UGCUGUGUGACAAUGGUGUUUG	Gerresheim et al. 2017
miR-122* 5B.2m	(phos)AACGCCAUUAUCACACGAAAUA	
miR-122 mat 5B.3m	(phos)UGGUGUCUGACAAUGGUGUUUG	
miR-122* 5B.3m	(phos)AACGCCAUUAUCAGACAAAAUA	
miR-122 mat S3m	(phos)UGCAGAGUGACAAUGGUGGGUG	
miR-122* S3m	(phos)CCCGCCAUUAUCACUCUAAAUA	

2.1.8.2 DNA oligonucleotides

Name	**Purpose**	**Sequence (5' - 3')**
NS5B_minus_RT	RT primer targeting the NS5B sequence in the minus strand	TGAGGTGTTCTGCGTGGAC
NS5A_minus_RT/for	RT / forward qPCR primer targeting the NS5A sequence in the minus strand	GCCTCCCCTTCATCTCTTGT
Spinach_minus_RT	RT primer targeting the Spinach sequence in the minus strand	CCGTCCTTCACCATTTCATT
Spinach_minus_RT_snap	RT primer targeting the Spinach sequence in the minus strand that forms a stem (underlined sequence) and a loop, unless hybridizes to the target	CTGAATGAAATGCCGTCCTTCACCATTTCATTCAG
EMCV_plus_norm_RT	RT primer targeting the EMCV sequence in the plus strand (used for normalization)	CCCCTTGTTGAATACGCTTG
NS5B_minus_qPCR_for	A pair of qPCR primers for quantification of the minus strand targeting the NS5B region	ACATTTTTCACAGCGTGTCG
NS5B_minus_qPCR_rev		GTACCTAGTGTGTGCCGCTCT
NS5A_minus_qPCR_rev	A reverse qPCR primers for quantification of the minus strand targeting the NS5A region	GATGGCGGTCTTGTAGTTCG
Spinach_minus_qPCR_for	A pair of qPCR primers for quantification of the minus strand targeting the Spinach region	GGAACTGCTTCCTTCACGAC
Spinach_minus_qPCR_rev		ACCATATTGCCGTCTTTTGG
Spinach_minus_qPCR_rev 2	An extended reverse qPCR primer for quantification of the minus strand targeting the Spinach region	CCACCATATTGCCGTCTTTTGG
EMCV_plus_norm_qPCR_for	A pair of qPCR primers for quantification of the plus strand targeting the EMCV region (used for normalization)	AGACCCCTAGGAATGCTCGT
EMCV_plus_norm_qPCR_rev		CCGTCCTTCACCATTTCATT
NS5B_minus_RT_rev	Reverse end-point PCR primer targeting the NS5B sequence in the minus strand	CGCCCCAAGTTTTCTGAGGG
GAPDH_for	RT / forward qPCR primer targeting cellular GAPDH	GAGTCAACGGATTTGGTCGT
GAPDH_rev	Reverse qPCR primer targeting cellular GAPDH	GATCTCGCTCCTGGAAGATG

RC_NS5B-3'UTR_SP6_for	A pair of primers for a PCR template generation serving in synthesis of control RNA template (HCV minus strand)	ATTTAGGTGACACTATAGGCGATATCTCTTCAATTGG
RC_ NS5B-3'UTR_rev		AGTTAGCTATGGAGTGTACC
as-3'UTR_RPA (probe 1)_SP6_for	A pair of primers for a PCR template generation used for a short sense 3'UTR probe synthesis (probe 1)	ATTTAGGTGACACTATAGATCTGCAGAGAGACCAGTTACGG
as-3'UTR_RPA (probe 1)_rev		GAGGTTACACGGGCTTGACG
as-3'UTR_RPA (probe 2)_SP6_for	A pair of primers for a PCR template generation used for a short sense 3'UTR probe synthesis (probe 2)	ATTTAGGTGACACTATAGGTGGCTCCATCTTAGCCC
as-3'UTR_RPA (probe 2)_rev		CTCGGAATGTTGCCCAGC
sense-5'UTR_RPA_SP6_for	A pair of primers for a PCR template generation used for antisense 5'UTR probe synthesis	ATTTAGGTGACACTATAGTGCACGGTCTACGAGACCTCC
sense-5'UTR_RPA_rev		GCGCCATTCGCCATTCAGGC

2.1.8.3 LNA oligonucleotides

Name	Sequence (5' - 3')	Source
anti-miR-122 LNA/DNA mixmer	+CCA+TTG+TCA+CAC+TC+C	Miravirsen (Janssen et al. 2013)

2.1.9 Buffers and solutions

2.1.9.1 Bacterial growth and cell culture solutions

LB growth medium: in order to obtain 1 l of the medium, 20 g of LB-Broth needs to be dissolved in deionized H_2O and sterilized by autoclaving. The growth medium is usually supplemented with a selective antibiotic (ampicillin) at working concentration 100 µg/ml directly before use.

2 % Agar-LB plates: for 200 ml stock solutions, 4 g of LB-Broth and 4 g of agar are supplemented with deionized H_2O; dissolution and sterilization are conducted by autoclaving. If necessary, antibiotic is added to a warm solution at working concentration 100 µg/ml prior to pouring dishes.

Tfb1 solution: 100 mM rubidium chloride, 50 mM manganese chloride ($MnCl_2$-$4H_2O$), 30 mM potassium acetate, 10 mM calcium chloride ($CaCl_2$-$2H_2O$), 15 % glycerol; pH is adjusted to 5.8 by 1 M acetic acid. The solution is sterilized by filtering and stored at 4°C.

Tfb2 solution: 10 mM rubidium chloride, 10 mM MOPS, 75 mM calcium chloride ($CaCl_2$-$2H_2O$), 15 % glycerol; pH is adjusted to 6.5 by 1 M potassium hydroxide. The solution is sterilized by filtering and stored at 4°C.

Complete 10 % DMEM: contains 100 U/ml penicillin, 100 µg/ml streptomycin and 10 % FBS (per 500 ml: 5 ml Pen/Strep and 50 ml FBS).

Cryomedium: DMEM that contains 20 % FBS and 10 % DMSO (per 50 ml: 10 ml FBS and 5 ml DMSO).

Cytomix: 120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4 (pH 7.6), 25 mM HEPES, 2 mM EGTA, 5 mM $MgCl_2$, pH 7.6; supplemented by 2 mM ATP and 5 mM reduced glutathione (GSH) before use.

0.5 % Trypsin-EDTA: 10 x stock is diluted 1:10 with sterile 1 x PBS.

1 x PBS: 10 x PBS stock is diluted 1:10 with deionized H_2O and sterilized by autoclaving.

2.1.9.2 Gel electrophoresis buffers

50 x TAE buffer: 2 M Tris (pH 7.6), 1 M acetic acid, 50 mM EDTA.

10 x TBE buffer: 1 M Tris (pH 7.6), 1 M boric acid, 20 mM EDTA.

Stacking gel buffer: 1.5 M Tris-HCl, pH 6.8, adjusted with HCl.

Resolving gel buffer: 1.5 M Tris-HCl, pH 8.8, adjusted with HCl.

10 x SDS Running buffer: 0.25 M Tris-HCl, 1.92 M glycine, 1 % SDS.

2 x SDS Sample buffer: 125 mM Tris-HCl (pH 6.8), 20 % glycerol, 20 % ß-Mercaptoethanol, 4 % SDS, few grains of bromphenol blue.

FA loading buffer: 80 % formamide, 10 % glycerol, 50 mM EDTA (pH 7.5), few grains of bromphenol blue and xylene cyanol.

2.1.9.3 Western blot buffers

NP-40 lysis buffer: 25 mM Tris-HCl (pH 7.5), 150 mM KCl, 2 mM EDTA (pH 7.5), 0.5 mM DTT, 0.5 % NP-40.

Anode I buffer: 0.3 M Tris-HCl, 10 % Methanol, pH 10.4, adjusted with HCl.

Anode II buffer: 25 mM Tris-HCl, 10 % Methanol, pH 10.4, adjusted with HCl.

Cathode buffer: 25 mM Tris-HCl, 10 % Methanol, pH 9.4, adjusted with HCl.

10 x TBS buffer: 0.2 M Tris-HCl, 0.15 M NaCl, pH 7.5, adjusted with HCl.

1 x TBST buffer: 10 x TBS is used to prepare 1 x solution and supplemented with 0.25 % Tween 20.

Blocking solution: 10 x TBS is used to prepare 1 x TBS(T) solution and supplemented with 10 % of dry milk.

2.1.9.4 IP and RPA buffers

RNA hybridization buffer: 80 % Formamide, 40 mM Tris-Cl (pH 7.5), 400 mM NaCl, 1 mM EDTA.

RNase digestion buffer: 10 mM Tris-Cl (pH 7.5), 300 mM NaCl, 5 mM EDTA; supplemented by 40 µg/ml RNase A directly before use.

Proteinase K Buffer: 200 mM Tris-Cl (pH 7.5), 300 mM NaCl, 25 mM EDTA, 2 % SDS.

IP Wash Buffer: 50 mM Tris-Cl (pH 7.5), 300 mM NaCl, 5mM $MgCl_2$, 0.05 % NP-40.

Fixer solution: 7.5 % acetic acid, 5 % ethanol.

2.1.10 Chemicals and reagents

Chemical	Company
Acetic acid	Sigma
Agarose	Roth
Ammonium persulfate (APS)	Roth
Boric acid	Roth
Bromophenol blue	Roth
Calcium chloride ($CaCl_2$)	Roth
Chloroform	Sigma
Dimethyl sulfoxide (DMSO)	Roth
Dithiothreitol (DTT)	Sigma
Ethanol (≥99.8 %, p.a.)	Roth
Ethidiumbromide	Roth
Ethylenediaminetetraacetic acid (EDTA)	Roth

Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid (EGTA)	Roth
Formamide	Roth
Glycerol	Roth
Glycine	Roth
Hydrochloric acid min. 32 % (HCl)	Roth
Isoamylalcohol	Roth
Isopropanol	Roth
Magnesium chloride ($MgCl_2$)	Roth
Manganese chloride ($MnCl_2$-$4H_2O$)	Sigma
Methanol	Roth
MOPS (3-(N-morpholino) propanesulfonic acid)	Roth
Nonidet P-40 (NP-40)	AppliChem
Phenol "Roti Phenol"	Merck
Potassium acetate (KOAc)	Roth
Potassium chloride (KCl)	Fluka BioChemika
Potassium hydroxide (KOH)	Roth
Rotiphorese Gel 40 (acrylamide/bisacrylamide 29:1)	Roth
Rubidium chloride (RbCl)	Sigma
Skimmed milk powder	EDEKA Krenschker
Sodium acetate (NaOAc)	Roth
Sodium chloride (NaCl)	Roth
Sodium Dodecyl Sulfate (SDS)	Roth
TEMED (N,N,N',N'-Tetramethylethylendiamine)	Roth
Tris-hydroxymethylaminomethane (Tris)	Roth
TRIzol reagent	Thermo Scientific
Tween 20	Sigma
Xylene cyanol	Roth
β-Mercaptoethanol	Sigma

2.1.11 Laboratory and cell culture plastic

Consumable	**Company**
6- and 12- well microplates	Sarstedt
10 cm dishes	Sarstedt
Electroporation cuvettes (4 mm)	Sigma
Eppendorf tubes (1.5 ml, 2 ml; Safe Lock)	Sarstedt
Falcon tubes (15 ml, 50 ml)	Sarstedt
Filter pipette tips (10 µl, 20 µl, 200 µl, 1250 µl)	Sarstedt
qPCR Seal optical clear film	VWR
qPCR Semi-Skirted Plate	VWR
Pipette Tips (10 µl)	Sarstedt
Sterile serological pipettes (5 ml, 10 ml, 25 ml, 50 ml)	Greiner
Tissue culture flasks (25 cm^2, 75 cm^2, 175 cm^2)	Sarstedt

2.1.12 Laboratory equipment

Equipment	**Company**
Accublock Digital Dry Bath	Labnet
Autoclave V150	Systec
B15 petri dish incubator	Thermo Scientific
CB series CO_2 Incubator	Binder
Centrifuge 5415 C	Eppendorf
Destamat, bi-destiller	Heraeus
Digital pH-Meter 644	Knick Elektronische Messgeräte
Duomax 1030 shaker	Heidolph

ED240 hot-air cabinet	Binder
FastBlot B44 semidry blotting chamber	Biometra
Geiger counter LB 124	Berthold Technologies
GelDoc XR gel documentation system	BioRad
GenePulser Xcell	Bio-Rad
HA 2448 BS LaminAir lamina flow	Heraeus
Heat-stir US152 magnetic stirrer	Stuart
Heraeus Biofuge Fresco	Heraeus
Herolab UV Transilluminator	Herolab
Julabo 7A water bath	Julabo
Labquake Rotator	Thermo Scientific
Leica DM IL inverted microscope	Leica Microsystems
LKB/BROMMA 2002 Power Supply	LabX
MagnaRack magnetic separation rack	Thermo Scientific
Mastercycler ep Realplex² S	Eppendorf
Micropipettors (0.5 µl to 1000 µl)	Gilson
Pipetboy comfort pipettor	Integra Biosciences
Phosphor Imager Cyclone Plus	Perkin Elmer
Qubit 2.0 Fluorimeter	Thermo Scientific
Sharp R 202 Microwave	Sharp
TProfessional PCR cycler	Biometra
Vortex Genie 2	Scientific Industries

2.1.13 Tools, software and internet resources

BioEdit Sequence Alignment Editor version 7.2.0 (Hall 1999)

Clone Manager Professional Suite version 8 (Small Work Group License)

Clustal MUSCLE: https://www.ebi.ac.uk/Tools/msa/muscle/

Clustal Omega: https://www.ebi.ac.uk/Tools/msa/clustalo/

Double Digest Finder by NEB: https://www.neb.com/tools-and-resources/interactive-tools/double-digest-finder

OriginPro8 SR4 software, v8.0951, OriginLab Corporation

Primer3web (v. 4.0.0) (Rozen and Skaletsky 2000; Koressaar and Remm 2007; Untergasser et al. 2012): http://bioinfo.ut.ee/primer3/

Tm Calculator by NEB: http://tmcalculator.neb.com

2.2 Methods

2.2.1 Microbiological methods

2.2.1.1 Cultivation of bacteria

All steps involving bacteria were carried out under sterile conditions using a Bunsen burner. All solutions and growth media were sterilized either by filtering or autoclaving (20 min at 120°C, 2 bars). Glassware and metal spatulas were disinfected by dry heat sterilization (4 h at 280°C), and only sterile plastic consumables were applied. A working surface was disinfected before and after bacteria cultivation with 70 % ethanol. Bacteria were predominantly grown under selective antibiotic conditions – 100 µg/ml of ampicillin - according to the resistance gene encoded in plasmids used for transformation. Cultivation is conducted either on 2 % agar-LB plates or in liquid LB-medium in an incubator or a shaker, respectively; at 30 or 37°C.

2.2.1.2 Preparation of competent bacterial cells: rubidium chloride method

Based on commercially available One Shot TOP10 E.coli cells, additional chemically competent bacterial stocks were prepared. TOP10 cells from one vial were plated on an antibiotic-free LB-agar plate and incubated overnight at 37°C. Then a starter mini culture was produced by an inoculation of a single colony into 10 ml of antibiotic-free growth medium, followed by overnight incubation in 37°C shaker. On the next day a mini culture was transferred into 1 l of antibiotic-free LB-medium and grown under the same conditions for about 5 h until the cultures OD_{600} reaches 0.3-0.5. From this point every step is to be performed on ice and only ice-cold Tbf1/Tbf2 solutions to be used. The culture is kept on ice for 5 min and then centrifuged at 5000 rpm for 10 min at 4°C. Next, cells are gently resuspended in 200 ml ($1/5^{th}$ from initial culture volume) of Tfb1 solution and kept on ice for another 5 min. Eventually, cells are spun down at 5000 rpm for 10 min at 4°C and resuspended in 5 ml ($1/20^{th}$ from initial culture volume) of Tfb2 solution and additionally incubated on ice for 15 min. 100, 200 and 500 µl aliquots are dispensed into pre-chilled tubes and snap-frozen in liquid nitrogen; storage is at -80°C. In order to characterize a current stock, one vial is transformed by pUC19 plasmid and transformation efficiency is estimated according to a formula:

Number of transformants / DNA [µg] = number of colonies on the plate / (total amount of DNA transformed [µg] x a dilution of a culture before plating)

2.2.1.3 Transformation of competent cells

For transformation, a plasmid or a ligation reaction is added to 100 µl of competent cells in required amount and incubated on ice for 30 min. Transformation is conducted by a heat-shock for 1 min at 42°C followed by a cool-down for 2-3 min on ice. Next, cells are carefully mixed into 250 µl of warm SOC outgrowth medium and being recovered for 1 h in a 37°C shaker. Transformed cells are plated on ampicillin-containing agar plates and cultivated as described above.

2.2.2 Cell culture methods

2.2.2.1 Passaging of eukaryotic cells

All steps involving eukaryotic cells were carried out under sterile conditions using a biosafety cabinet and certified sterile solutions, growth media and plastic ware. Everything that was taken into the cabinet as well as a working surface were disinfected before and after with 70 % ethanol. Cells were maintained in a 37°C, 5 % CO_2 humidified incubator and subcultured every 2-3 days depending on a cell type and on current confluence. All the solutions and media for subculturing need to be pre-warmed to room temperature. First cells are washed once with 1 x PBS, then a required amount of 0.5 % trypsin-EDTA is applied (3 and 5 ml for 75 and 175 cm^2 flasks, respectively). Culture flasks are kept at 37°C until cells are detached, and then trypsin is inactivated by addition of complete DMEM containing 10 % FBS. After resuspension of cells in the growth media, they are counted, if necessary, and seeded into new flasks or plates, as desired.

2.2.2.2 Thawing and freezing of eukaryotic cells

In order to create a long-term storage stock of freshly purchased or kindly provided cell lines, cryostocks need to be prepared. Every new cell line first is passaged several times in increasing amounts for cells adaptation and multiplication until they reach about 90 % confluence. Then all cells are trypsinized, resuspended in a small volume of serum-containing DMEM, transferred into a sterile tube and pelleted at 1000 rpm for 5 min at room temperature. After supernatant removal, cells are gently resuspended in cryomedium at a concentration of about 2-4*10^6 cells per ml; and 1 ml aliquots are transferred into cryotubes and snap-frozen in liquid nitrogen; storage is in a liquid nitrogen tank.

Viability of cells is then tested by thawing one of the vials. For this reason, a vial is briefly thawed in 37°C water-bath; cells are to be pelleted at 1000 rpm for 5 min, then resuspended in pre-warmed 10 % DMEM and placed into a 25 cm^2 flask containing complete growth medium. This medium needs to be exchanged within 4-6 h (after attachment of viable cells) and if necessary from time to time until the cells reach confluence of ~90 %.

2.2.2.3 Counting and seeding of eukaryotic cells

For all the experiments in which cells had to reach a certain percent of confluence they were counted prior seeding. A concentration of cells in a suspension obtained after trypsinization was determined using a Neubauer improved hemocytometer. Characteristics of this counting chamber are the following: a depth is 0.1 mm, a counting grid is 3 mm x 3 mm in size with 9 square subdivisions of 1 x 1 mm^2 and the central square is devided into 25 squares of width 0.2 mm. To calculate a total amount of cells, the number of cells in 2-3 bigger squares is counted, and the mean value is multiplied by 10^4 and by a suspension's volume (in ml).

For transfection cells are to be seeded one day before, in either 6- or 12- well microplates, at amount of 2-2.5*10^5 and 0.7-0.75*10^5 cells per well, respectively, that provides 70-90 % cells confluence.

2.2.2.4 Transfection of eukaryotic cells: Lipofectamine 2000 method

To perform (co-)transfection of RNA substrates with microRNA or other RNA oligonucleotides, Lipofectamine 2000 method was applied. Transfection was only conducted on the cells reached 70-90 % confluence; in 6-well microplates for RT-qPCR experiments and in 12-well microplates for protein analysis by western blot. The amounts given below correspond to transfection in 6-well format, whereas for 12-well format all the amounts should be reduced by 3-fold.

The day before transfection, cells were trypsinized and seeded into 6-well plates (2.5*10^5 cells/well). In all experiments 0.3 pmol of RNA template per well was transfected; when required, microRNA(s) or LNA were co-transfected in various amounts. RNA components and Lipofectamine reagent were first pre-mixed in separate tubes (in 100 µl of serum/antibiotic-free DMEM per sample) and then combined and mixed well by vortexing for 10 s. Mixtures were incubated at room temperature for 15 min and then transferred dropwise on cells; no change of media was conducted. Cells were further incubated at standard conditions, usually for 48 h or according to a chosen time-course schedule, until harvesting and total RNA isolation.

2.2.2.5 Transfection of eukaryotic cells: Electroporation method

Evaluation of stability of RNA substrates in experimental cells transfected by Lipofectamine (see 2.2.2.4) is unflavored since liposomes provide an uneven nucleic acid uptake resulting in differential transfection efficiency or/and retention of RNA at the outer cellular surface and gradual release leading to artificially extended lifetime values. A more optimal approach is electroporation, which allows precise and instant delivery of RNA. Calculation of half-life time of nucleic acid species requires quantification of its remaining fraction during an expedient time-course. In the present work stability of RNA templates was evaluated within 36 h period. At first, HuH-7.5 cells were trypsinized, pelleted by centrifugation at 1500 rpm for 5 min at room temperature, then washed 1-2 times with PBS and resuspended in Cytomix to 10^7 cells/ml.

For the electroporation, 400 µl ($4*10^6$ cells) of HuH-7.5 suspension were combined with 1 pmol of the RNA of interest, gently mixed by pipetting up and down, transferred into 4 mm cuvettes (Sigma) and then pulsed using a GenePulser Xcell (Bio-Rad) with the following conditions: square wave, 270 V, 20 ms. Next, cells were carefully resuspended in 12 ml of complete DMEM and seeded in 6- or 12- well plates (2 or 1 ml per well, respectively). The input samples, representing a time-point 0 h (in fact, about 30 min post electroporation), were directly harvested by Trizol lysis of pelleted cells. In order to remove unviable cells, 6 h post electroporation growth media exchange was conducted. According to the selected time-course, cells were further harvested by Trizol at 6, 16, 24 and 36 h post electroporation; the lysis was followed by total RNA isolation (see 2.2.6.1). Remaining fractions of transfected RNA constructs were quantified by RT-qPCR (see 2.2.6.3), targeting the EMCV IRES region on plus strands. For individual constructs, mean Ct values for each time-point were normalized to a value for the input sample (100 %) and then mean remaining fractions (derived from at least three independent experiments) were fitted to exponential decay function using OriginPro8 software (ExpDec1 function: y = A1*exp(-x/t)+y0). Half-life values $T_{1/2}$ were calculated from decay curves using a t1 (lifetime) characteristic parameter and its standard error values, as $T_{1/2} = t1*\ln(2)$.

2.2.3 Basic molecular biological methods

2.2.3.1 Preparation of plasmid DNA

Preparation of plasmid DNA was performed according to the following standardized procedure. For the primary colony screening, small-scale DNA preparation was applied. Colonies were inoculated and transferred into 15 ml Falcon tubes containing 2-2.5 ml of pre-warmed LB medium with an appropriate antibiotic (100 µg/ml ampicillin). Afterwards samples were incubated in 30 or 37°C shaker overnight; caps of the tubes were loosened. It is important to note that an incubation at 30°C instead of 37°C served to reduce activity of bacterial endonucleases and recombination negatively affecting long plasmids (over 10 kb) during cultivation. Next day 1 ml of cultures was transferred into Eppendorf tubes and spun down at maximum speed for 5 min. Further procedures were conducted using a GeneJET Plasmid Miniprep Kit according to the manufacturer's instructions. The kit enables a convenient SDS/alkaline lysis of pelleted bacteria followed by a neutralization step providing optimal high-salt environment for plasmid DNA binding and protein/chromosomal DNA precipitation. While cell debris are precipitated and removed by centrifugation, the plasmid DNA is bound to the silica-based membrane of the column. Consequent washing with ethanol serves to remove contaminants, and plasmid DNA is ultimately eluted by a small volume of nuclease-free water. Eventually, integrity of isolated plasmid DNA was verified by restriction analysis (see 2.2.3.2) and by agarose gel-electrophoresis (see 2.2.4.1).

Large-scale DNA preparation served to produce highly concentrated, pure and endonuclease-free stocks of a desired plasmid. Usually, one selected small-scale bacterial clone per plasmid was inoculated from an overnight mini-culture into a glass flask filled (to maximum $1/3^{rd}$) with pre-warmed LB medium containing an appropriate antibiotic (100 µg/ml ampicillin). Similarly, incubation was conducted overnight in a 30 or 37°C bacterial shaker (the incubation temperature mode conforms to the one for the primary cultures). Next day the whole amount of a culture was distributed and carefully equilibrated in centrifuge flasks, and centrifuged at 5000 rpm for 10 min at 4°C. Further procedures were performed using a GeneJET Plasmid Maxiprep Kit according to the manufacturer's instructions. The principle of the procedure is identical to the one of the small-scale plasmid DNA isolation, with addition of the endotoxin-binding step, since the large-scale preparation is to be applied to cells. Prior the stock creation, integrity of isolated DNA was confirmed by restriction analysis (see 2.2.3.2) and agarose gel-electrophoresis (see 2.2.4.1); plasmid DNA concentration was precisely measured using Qubit (see 2.2.3.7).

2.2.3.2 Restriction endonuclease digest of DNA

Restriction digest of DNA with endonucleases is an essential method to generate sticky ends for ligation procedures as well as an analytical approach for verification of integrity of DNA plasmids: for

screening and after a large-scale preparation. Enzymatic reaction is set up according to the manufacturer's instructions that include specific guidelines on a compatible buffer and an enzymatic activity (see 2.1.3.1). The volume of a restriction digest reaction is customized depending on the downstream application of a product. For an *in vitro* transcription DNA template preparation, usually 50 µg of plasmid DNA was used for the digest with 100 U of *Eco*RI endonuclease in 200 µl reaction volume for 2 hours at 37°C. For analytical purposes smaller amounts of DNA are sufficient: about 1 µg of DNA was digested within 25 µl reaction mixture. Larger amounts of DNA fragments are normally used for ligation procedures after subsequent gel-extraction (see 2.2.4.2). When a double digest is required, the optimal conditions can be defined at the manufacturer's website (Double Digest Finder). A standard digest procedure is usually carried out at 37°C for 1 h, but may significantly vary depending on the selected restriction endonuclease.

2.2.3.3 The polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) is a commonly known indispensable procedure for *in vitro* amplification of a DNA segment flanked by two defined sequences. Here, PCR was predominantly used for cloning purposes, particularly, for site-directed mutagenesis (described in detail in 2.2.5.3). The standard PCR procedure is outlined below. The reaction was performed using One Taq DNA Polymerase together with a provided OneTaq Standard Reaction Buffer (20 mM Tris-HCl, 22 mM NH_4Cl, 22 mM KCl, 1.8 mM $MgCl_2$, 0.06 % IGEPAL CA-630, 0.05 % Tween 20; pH 8.9) abiding the manufacturer's instructions. The general composition (for a standard 25 µl reaction is given in brackets) is following: 1 x OneTaq Standard Reaction Buffer (5 µl of 5 x stock), 0.2 mM dNTPs (0.5 µl of 10 mM stock), 0.2 µM of each forward and reverse primers (0.5 µl of 10 µM stocks), 0.625 U One Taq DNA Polymerase (0.125 µl of 5 U/µl stock), <1 µg template DNA (volume is variable); the total volume is adjusted to 25 µl with nuclease-free water. When necessary, PCRs can be scaled up to 2 x (50 µl) or 4 x (100 µl) for preparative purposes. The reactions were performed using one of the listed thermal cyclers (see 2.1.12) applying the standardized temperature profile: initial denaturation – for 2 min at 94°C; 30 cycles of subsequent denaturation (for 15-30 sec at 94°C), annealing (for 30 sec at 45-68°C) and elongation (1 min per 1 kb, at 68°C); final extension - for 5 min at 68°C; hold is at 4°C. The denaturation time should be increased when dealing with templates with a high GC-content. The annealing temperatures are to be defined for every primer pair (NEB Tm Calculator was used); in case of difficulties, the optimal conditions were determined via a gradient PCR program. The number of cycles may vary depending on expected copy number of a target. All PCR products are further analyzed by agarose gel-electrophoresis (see 2.2.4.1) before proceeding with a downstream application.

2.2.3.4 *In vitro* RNA transcription

Replication template RNAs for both HCV minus and plus strand synthesis initiation were generated by *in vitro* transcription. Corresponding plasmid DNAs were linearized with *Eco*RI digest (see 2.2.3.2), followed by extraction with phenol and chloroform and precipitation with ethanol (see 2.2.3.5 and 2.2.3.6); finally DNAs were dissolved in 50 µl nuclease-free water. DNA concentration was measured by Qubit using the dsDNA BR Assay Kit, and completeness of linearization was tested by 1 % agarose gel electrophoresis (see 2.2.4.1).

In vitro transcription for replication constructs generation was performed using T7 RNA polymerase with the following modifications to the manufacturer's protocol: 3.5 mM of each ATP, GTP, CTP and UTP, additional 5 mM $MgCl_2$ and 10 mM DTT, 25 ng/µl of linearized plasmid DNA and 1 U/µl of T7 RNA polymerase. After 3 h at 37°C, an additional 0.5 U/µl of T7 RNA polymerase was added, and the reaction was incubated for another 2 h. Termination of transcription and template DNA removal was conducted by addition of 2 U of RNase-free DNase I per 1 µg of template DNA, followed by incubation for 15 min at 37°C. RNA was further extracted with phenol and chloroform and precipitated with ethanol (see 2.2.3.5 and 2.2.3.6); finally RNA transcripts were dissolved in RNase-free water. RNA concentration was measured by Qubit using RNA BR Assay Kit (see 2.2.3.7), and integrity of transcripts was tested by 1 % agarose gel electrophoresis (see 2.2.4.1).

For the purposes of radioactively labeled RNA production, *in vitro* transcription was performed according to the standard manufacturer's protocol with a modification to the concentration of UTP. To achieve an efficient incorporation of radio-labeled [α-^{32}-P]-UTP into a transcribed RNA, the amount of cold UTP was reduced to 0.1 mM comparing to a standard 0.5 mM final concentration. It is important to note that *in vitro* transcription of certain RNAs – such as containing the U-rich HCV 3'UTR sequence – requires optimization of reaction conditions or a change of radioactively labeled nucleotide in order to ensure a generation of a full-length product. For an average sequence the *in vitro* transcription conditions were as following: 1 x T7 RNA polymerase reaction buffer, 0.5 mM of each ATP, GTP and CTP, 0.1 mM of cold UTP, 0.5 μM [α-^{32}-P]-UTP (400 Ci/mmol, 10 mCi/ml), additional 10 mM DTT, 10 ng/μl of DNA template and 0.1 U/μl of T7 RNA polymerase. The reaction was carried out at 37°C for 1.5-2 h for T7 RNA polymerase as well as for SP6 RNA polymerase that was required in several applications. Subsequent removal of DNA template and purification of radioactive RNA transcripts is performed similarly to cold transcription products, as mentioned above and in 2.2.3.5 and 2.2.3.6. Verification of RNA integrity and concentration estimation was conducted via denaturing polyacrylamide gel electrophoresis (see 2.2.4.4). The amount of radioactively labeled RNA was calculated through estimation of percentage of incorporated radioactive nucleotides.

2.2.3.5 Phenol-chloroform extraction

A procedure of nucleic acids extraction with phenol and chloroform serves to denature and remove proteins from an enzymatic reaction (such as restriction digest, DNase I digest or *in vitro* transcription). In order to complete purification and concentrate the sample, the extraction was always followed by ethanol (or isopropanol) precipitation. The procedure of nucleic acids separation from a protein fraction is based on a formation of two phases – watery and organic. Nucleic acids remain in the watery phase, whereas proteins are denatured and trapped in an organic phase. At first, 1.5 volume of phenol is added to 1 volume of an aqueous sample of interest. After mixing thoroughly, the sample is centrifuged at maximum speed (13 000 rpm) for 5 min at room temperature. The watery phase is then transferred into a clean tube (avoiding taking out any of the organic phases) and mixed with 1.5 volume of phenol/chloroform mixture (1:1), and the procedure is repeated. Similarly, at the last step the watery phase is mixed with 1.5 volumes of chloroform only to remove residual phenol, final centrifugation is performed and eventually the watery phase is collected into a clean tube. This multiplicity of steps serves to increase the purity of nucleic acid, however, the procedure may be reduced to two steps (omitting the second step).

2.2.3.6 Precipitation of nucleic acids with ethanol/isopropanol

The precipitation with ethanol or isopropanol usually comes after phenol-chloroform extraction (see 2.2.3.5) to both concentrate nucleic acids and remove water-soluble contaminants that remain from the previous step (salts, sugars, etc.). The aqueous phase derived at the last step of phenol-chloroform extraction is thoroughly mixed with 2.5 volumes of an absolute ethanol (or 1 volume of isopropanol; the highest degree of purity). Usually 1/10th volume of 3 M sodium acetate is added to neutralize charges on the nucleic acid, however, when a low yield of nucleic acids is expected, 15 μg of GlycoBlue agent can be additionally applied as a co-precipitant. For precipitation, samples are incubated at -20°C for at least 1 h; and nucleic acid is then recovered by centrifugation at maximum speed (13 000 rpm) for 15 min at 4°C. After removing supernatant, pellets are washed 1-2 times with 70 % ethanol and air-dried at 37°C. Pellets are then dissolved in a reasonable amount of nuclease-free water and, when necessary, nucleic acid concentration is measured according to 2.2.3.7.

2.2.3.7 Nucleic acid concentration measurement

All steps involving measurements of nucleic acids concentration were performed using a Qubit fluorometric quantification approach. The Qubit 2.0 Fluorimeter together with corresponding assay kits allows quantification of both RNA and DNA, in either high-sensitive or broad range. Measurements were conducted according to the manufacturer's instructions: 1 μl of a nucleic acid sample is well-mixed into

199 µl of a working solution and incubated for 2 min. Calculation of a final concentration is automatized and based on a programmed calibration curve.

2.2.3.8 Preparation of microRNA duplexes

Co-transfection of microRNAs – original miR-122 and its mutated versions (see 2.1.8.1) – required application of mature imperfect duplexes. All miRNA duplexes were generated from purchased single stranded RNA oligonucleotides by annealing equimolar amounts of a guide (mat) and a corresponding passenger (*) strands. In mutated duplexes, the sequence was designed to maintain the original mis-pairing to assure that the duplex is unwound from the correct end. Annealing was performed in a PCR cycler by a steady temperature decrease from 90°C to 4°C (1°C per minute). Generated miR-122, miR-122_S1m/S2m/5B.2m/5B.3m/S3m duplexes were then aliquoted and stored at -20°C.

2.2.3.9 DNA sequencing

For analysis and verification of DNA segments after cloning or mutagenesis, sequencing service by GATC-biotech (and later SeqLab) was applied. Both companies provided primer synthesis service prior to sequencing and required only shipment of 50 to 500 ng of DNA sample (sample requirements vary depending on the DNA source).

2.2.4 Gel electrophoresis methods

2.2.4.1 Agarose gel electrophoresis

A standard agarose gel electrophoresis procedure was applied for analysis and separation of both RNA and DNA. Multiple molecular biological techniques require a size and integrity verification by this method during their experimental timeline: visualization of generated PCR products, digested DNA fragments or transcribed RNA (see 2.2.3.2-4) – are just some of the applications. Meanwhile, preparative agarose gels served to separate DNA fragments for further cloning purposes.

A concentration of agarose – from 1 to 2 % - was chosen depending on the expected size of the fragments to be separated: 1 % gels for the fragments from 0.5 to 10 kb and 2 % - for the 0.1-2 kb range. Gels are prepared in 1 x TAE buffer by agarose melting in a microwave oven. Before loading a gel, nucleic acid sample is to be mixed with an appropriate – RNA or DNA - loading dye. Standard commercially available loading dyes usually contain bromphenol blue and xylene cyanol and allow tracking of the nucleic acid migration process. The electrophoresis is performed in 1 x TAE running buffer at 20-25 mA per gel; a choice of the current depends on the size of separated fragments and a preferred resolution. A suitable length marker is to be loaded on each gel. To stain separated nucleic acids, the gels were soaked in ethidium bromide solution (0.5 µg/ml in 1 x TAE buffer) for 20-40 min. Visualization of fragments was carried out under UV-light illumination using a GelDoc XR gel documentation system.

2.2.4.2 Recovery of DNA fragments from agarose gels

In order to separate certain DNA species from a mixture of fragments, agarose gel electrophoresis followed by a gel extraction was conducted. A preparatory procedure was consistent with 2.2.4.1. However, to avoid an unwanted DNA damage by UV-light, visualization was conducted using a table transilluminator and a protective layer of aluminum foil underneath a gel; an exposure time was reduced to minimum. A fragment of interest was defined in accordance with a marker ladder, excised using a blade and transferred into a clean tube. Subsequent purification of DNA from a gel slice was carried out with a GeneJET Gel Extraction Kit, following the manufacturer's instructions.

2.2.4.3 Polyacrylamide gel electrophoresis under non-denaturing conditions

Analytical non-denaturing polyacrylamide gel electrophoresis (PAGE) was applied as an alternative method for shorter DNA fragments separation (below 500 bp; for instance, after RT-PCR, see 2.2.6.2) instead of 2 % agarose gel electrophoresis. Electrophoresis was performed under non-denaturing conditions;

the concentration of acrylamide was chosen in accordance to the expected size of the fragments to be separated: 6 or 8 %. Gels were prepared in 1 x TBE buffer supplemented with an appropriate amount of acrylamide solution (from a 40 % stock solution), 0.1 % APS and 0.1 % TEMED just before use. Prior loading on a gel, samples were premixed with a loading dye containing bromphenol blue and xylene cyanol as well as a suitable length marker to control a migration process. The electrophoresis was performed in 1 x TBE running buffer at 20-25 mA per gel; gels were subsequently stained in the ethidium bromide bath (0.5 μg/ml in 1 x TBE) for 5-10 min. Visualization of fragments was carried out under UV-light illumination using a GelDoc XR gel documentation system.

2.2.4.4 Polyacrylamide gel electrophoresis under denaturing conditions

Analytical polyacrylamide gel electrophoresis conducted under denaturing conditions was applied for evaluation of *in vitro* transcribed radio-labeled RNAs integrity (see 2.2.3.4) as well as at final steps of RPA and RIP experiments (see 2.2.8 and 2.2.9). Depending on size of expected RNA products, 6-12 % polyacrylamide gels supplemented with 50 % Urea, 0.1 % APS, 0.1 % TEMED in 1 x TBE were poured. A careful washing of the wells of a gel and a pre-run of electrophoresis at standard conditions (for at least 10 min) are highly recommended to improve a resolution and a final image quality. Prior loading the samples are to be supplemented with 1 (or higher) volume of Formamide-containing loading buffer and denatured for 2 min at 90°C at the heating block. Electrophoresis is performed at 30 mA per gel for a suitable period of time according to expected bromphenol blue and xylene cyanol fronts migration for a chosen acrylamide concentration. Eventually radioactively labeled RNA is visualized by autoradiography. For the analysis of *in vitro* transcribed RNA gels were usually wrapped in a plastic foil and exposed directly to Kodak X-ray film for about 5 min. In order to conduct a prolonged exposure, gels were usually fixed in fixer solution for 30-40 min and dried on a gel-drier for 2 h at 80°C. Dehydrated gels were exposed to X-ray film or to a PhosphorImager screen for a required period of time to obtain an optimal image.

2.2.4.5 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

For the separation and analysis of proteins under SDS-denaturing conditions, the discontinuous Laemmli polyacrylamide gel system was utilized (Laemmli 1970). Such SDS gels consist of an upper stacking gel (usually 5 %) and lower resolving gel (8-12 %); the percentage of acrylamide in a resolving gel depends on the size of separated proteins. For visualization of the HCV NS3 protein together with cellular GAPDH (as a loading control), 10 % resolving gel was used in the Laemmli system.

First, a resolving gel solution is to be poured; its composition is the following: 8-12 % Acrylamide/bisacrylamide (29:1), 375 mM Tris-HCl (pH 8.8), 0.1 % SDS, 0.1 % APS and 0.1 % TEMED. After a complete polymerization of a resolving gel, a stacking gel solution is to be added on top; it has a lower pH and consists of: 5 % Acrylamide/bisacrylamide (29:1), 125 mM Tris-HCl (pH 6.8), 0.1 % SDS, 0.1 % APS and 0.1 % TEMED. Protein-containing samples were diluted 1:4 in 4 x SDS sample buffer and denatured by cooking on a dry bath at 95°C for 10 min. After a brief spin-down samples were directly loaded on a gel together with a size marker. Electrophoresis was performed using a vertical gel chamber in 1 x SDS running buffer, first, at 25 mA per gel for a stacking gel and at 35 mA per gel for a resolving gel; regularly, the gels were run until the bromphenol blue dye front exited the gel. After electrophoresis gels were immediately applied for the western blot analysis (see 2.2.7).

2.2.5 Molecular cloning and mutagenesis methods

A multi-step preparatory cloning procedure was conducted. The DNA templates for transcription were meticulously designed in this work and then commercially synthesized in parts by EuroFins and finally assembled as follows.

The minimal template for minus strand synthesis initiation - pUC18_P.s_WT_hp_9899 (designated as “hp”; see 6.1.7) - was assembled in one step from a chemically synthesized backbone plasmid (pUC18_Plus_strand_backbone_4374; 6.1.2) and a plasmid encoding the full-length HCV JFH1/J6 genome

(pFK-JFH1-J6 C-846_dg (Jc1)_12961; 6.1.1; Pietschmann et al. 2006). Alternatively, most of the ORF was scrambled, i.e. the primary sequence was permutated to disable RNA *cis*-signals while retaining amino acid coding specificity, codon usage and codon pair bias (Song et al. 2012). Scrambled sequence design was done by Yutong Song and Steffen Mueller at Stony Brook University (NY, USA). The scrambled version of the DNA template encoding for plus strand initiation construct was assembled in three steps using four chemically synthesized constructs: a backbone plasmid pUC18_Minus_strand_backbone_4685 (6.1.3) and three plasmids containing fragments of scrambled NS3-NS5B sequence (pUC18_Fragment 1_NS5B_SCR_4157, pUC18_Fragment 2_NS3-NS5B_SCR_5148 and pUC18_Fragment 3_NS3_SCR_4515; 6.1.4-6). Construction of the template for minus strand synthesis in scrambled context (pUC18_P.s_SCR_hp_9899; 6.1.8) and the one for plus strand initiation in wild-type context (pUC18_Minus_strand_WT_hp_10334; 6.1.11) was performed by an exchange of backbones/cassettes (also see Supplementary materials 6.2.1 and Suppl. Fig. 6.1-6.5). Since the *Asc*I site is located 152 nt upstream of the NS5B-3'UTR junction, this exchange site is located downstream of the NS5B miR-122 target sites, but directly upstream of the 5BSL3.2 element. Therefore, both resulting minimal constructs (wild-type or scrambled) contain the wild-type 5BSL3.2 sequence, but the miR-122 target sites in the NS5B coding region are retained only in original constructs.

The required modifications were introduced into the hp minimal constructs by site-directed mutagenesis. The template for minus strand replication initiation was supplemented with partial HCV 5'UTR sequences – SL I-II (position 1-117; harboring wild-type or mutated miR-122 sites; mutations are as described in Nieder-Röhrmann et al. 2017), SL I-III (position 1-330) or the complete 5'UTR (position 1-376, followed by 12 codons of the HCV Core-coding sequence and UGA) – resulting in plasmids pUC18_P.s_WT_SL I-II_wt_9968, pUC18_P.s_WT_SL I-II_S1mS2m_9968, pUC18_P.s_WT_SL I-III_10182 and pUC18_P.s_WT_5'UTR_sinfr_stop_10242, respectively (6.1.13-15 and 6.1.18). Accordingly, the SL I-II sequence was also inserted into the same hp construct with a scrambled context (resulting in pUC18_P.s_SCR_SL I-II_wt_9968 and pUC18_P.s_SCR_SL I-II_S1mS2m_9968; 6.1.21-22). The constructs with individually mutated *cis*-elements in the HCV NS5B region - pUC18_P.s_WT_SL I-II_8680mut_9968 and pUC18_P.s_WT_SL I-II_9170mut_9968 – were generated by site-directed mutagenesis (6.1.33-34).

The miR-122 binding sites in the HCV NS5B coding region and 3'UTR were mutated within SL I-II constructs (6.1.13-14), as in Gerresheim et al. 2017, resulting in plasmids with individually mutated 5B.2m, 5B.3m or S3m miR-122 binding sites: pUC18_P.s_WT_SL I-II_wt_5B.2m_9968, pUC18_P.s_WT_SL I-II_wt_5B.3m_9968, pUC18_P.s_WT_SL I-II_wt_S3m_9968 (6.1.25-27) and pUC18_P.s_WT_SL I-II_S1mS2m_5B.2m_9968, pUC18_P.s_WT_SL I-II_S1mS2m_5B.3m_9968, pUC18_P.s_WT_SL I-II_S1mS2m_S3m_9968 (6.1.28-30).

The modifications within the SL III domain - a mutation of GGG to CCC in the apical loop of SL IIId and a deletion of the domain SL IIIb – were in turn introduced into the SL I-III or the complete 5'UTR constructs by site-directed mutagenesis resulting in plasmids pUC18_P.s_WT_SL I-III_IIId mut_10182, pUC18_P.s_WT_SL I-III_IIIb del_10136, pUC18_P.s_WT_5'UTR_sinfr_stop_IIId mut_10242 and pUC18_P.s_WT_5'UTR_sinfr_stop_IIIb del_10196 (6.1.16-17 and 6.1.19-20). The constructs including the full-length Core-coding sequence downstream the HCV 5'UTR – unmodified or lacking the apical SL IIIb – are designated as pUC18_P.s_WT_5'UTR_Core_10779 and pUC18_P.s_WT_5'UTR_Core_IIIb del_10733 (6.1.35 and 6.1.37).

The replicase-deficient plasmid variants were generated for the minimal hp and SL I-II, SL I-III, 5'UTR, 5'UTR-Core constructs (6.1.7 and 6.1.13, 6.1.15, 6.1.18, 6.1.35) by introduction of a GND mutation within the NS5B RdRp gene (318D→N; Lohmann et al. 1997) using site-directed mutagenesis resulting in plasmids pUC18_P.s_WT_hp_GND_9899, pUC18_P.s_WT_SL I-II_wt_GND_9968, pUC18_P.s_WT_SL I-III_GND_10182, pUC18_P.s_WT_5'UTR_sinfr_stop_GND_10242 and pUC18_P.s_WT_5'UTR_Core_

GND_10779 (6.1.23-24, 6.1.33-34 and 6.1.36). These constructs served as negative controls of the minus strand synthesis initiation.

Some of the essential procedures for molecular cloning and mutagenesis, like preparative PCR and restriction digest, were already described above (see 2.2.3.2 and 2.2.3.3). The following steps are discussed below. Maps and annotations for all plasmids are provided in the Appendix (see 6.1). All cloning procedures are detailed in the Supplementary materials (see 6.2.1).

2.2.5.1 Dephosphorylation of DNA fragments

A standard cloning procedure engages preparation of the fragments – a vector and an insert - by sticky ends generation via a digest with specific restriction endonucleases. This approach, however, engenders terminal 5'-phosphate on both fragments that may lead to a self-ligation and re-circularization of a linearized vector. To reduce the vector background, a linearized vector is usually dephosphorylated prior ligation. Therefore, the gel-purified vector fragment (see 2.2.4.2) is treated with Antarctic Phosphotase according to the manufacturer's instructions. For a standard 20 µl dephosporylation reaction, 5 U of the enzyme per 1 pmol of DNA ends is applied in 1 x Antarctic Phosphatase reaction buffer; the incubation is carried out at 37°C for 30 min and followed by the Phosphotase inactivation at 80°C for 2 min. After DNA concentration measurement (see 2.2.3.7) the vector fragment can be directly applied for a ligation reaction.

2.2.5.2 Ligation of DNA fragments

To complete the formation of a desired plasmid, ligation of a linearized dephosphorylated vector and an insert was carried out using T4 DNA Ligase in accordance with the manufacturer's instructions. A typical 20 µl reaction mixture contains the following components: 1 x T4 DNA Ligase buffer, 50 ng of vector, a 3-5-fold molar excess of insert DNA, 1 µl (400 U) T4 Ligase; the reaction volume is adjusted with nuclease-free water. Depending on complexity of the ligation strategy, incubation can be performed either for 10 min at room temperature or overnight at 16°C. Afterwards, the ligation mixture is directly utilized for transformation of competent bacteria cells (see 2.2.1.3). Selection of the clones harboring the construct of interest is conducted using various molecular biological approaches: from restriction or PCR analysis (see 2.2.3.2 and 2.2.3.3) to sequencing of the insert and its flanking regions (see 2.2.3.9).

2.2.5.3 Site-directed mutagenesis

In order to introduce a defined mutation – substitution, insertion or deletion – into a plasmid of interest, several site-directed mutagenesis approaches were applied. Some mutations can be performed by one-step PCR if a restriction site is close to the site to be mutated, whereas most modifications require two-step mutagenesis PCR. In the first situation, one of the PCR primers is designed to harbor a mutation and a restriction site for further replacement. The opposite primer is chosen in a way that it flanks another appropriate restriction site. After the segment of interest is amplified, it is cleaved with the corresponding restriction endonucleases, as in the initial plasmid (see 2.2.3.2). The vector and the insert are then purified and combined as described above (see 2.2.5.1 and 2.2.5.2).

The primer extension approach, however, enables introduction of longer modifications or two remote mutations at once or if no appropriate restriction site is close to the mutated region. Therefore, mutation-containing primers are incorporated via independent nested PCRs to eventually combine them in the final product. The first PCR requires two types of primers: flanking primers that are complementary to the ends of a segment of interest and internal primers that contain the mismatched sequence (either the sequence to be mutated or inserted or the sequence from both sides of the deletion). In any case, the mismatched sequence mediates the further annealing of the products during the second PCR using flanking primers only. To be able to replace the mutated segment in the plasmid of interest, appropriate restriction sites are either retained within the segment or introduced with the flanking primers.

2.2.6 Reverse transcription and quantitative PCR (RT-qPCR) methods

In all experiments efficiency of the HCV minus or plus strand synthesis initiation from a certain RNA construct was estimated using a reverse transcription and quantitative PCR approach. First, as a proof of principle, synthesis initiation was demonstrated by RT-PCR and then – for quantitative comparison – all measurements were carried out via RT-qPCR.

2.2.6.1 Trizol-based cell lysis and total RNA isolation

The first step in the analysis of newly synthesized HCV RNA strands is an isolation of total RNA from transfected cells. HuH-7.5 cells were seeded and transfected as described in 2.2.2.4 or 2.2.2.5. After a required time interval (usually, 48 h post transfection), the cells were rinsed once with PBS and lysed by Trizol (1 ml per well). Phenol and guanidine isothiocyanate – the major components of the Trizol reagent – enable disruption of cells and isolation of RNA of wide-scale molecular size. The lysates were transferred into tubes and mixed thoroughly with 200 µl of chloroform for RNA extraction. After an incubation for 3 min at room temperature, the fractions were separated by centrifugation at maximum speed (13 000 rpm) for 15 min at 4°C. The RNA-containing upper aqueous fraction was collected into a clean tube; while DNA and proteins remain in the interphase and lower organic phases and discarded. Total RNA was precipitated by addition of 1 volume (500 µl) of isopropanol in the presence of 1 µl GlycoBlue reagent for at least 1 h at -20°C. Next, the recovery of total RNA was conducted by centrifugation at maximum speed (13 000 rpm) for 15 min at 4°C, followed by two subsequent washing steps (with 700 µl of 80 % ethanol) and reconstitution in 50 µl of RNase-free water. Residual DNA removal was conducted with of DNase I with subsequent enzyme removal by the RNA Clean-up Kit according to the manufacturer's instructions. Eventually, total RNA samples were eluted in equivalent amounts of RNase-free water. Final measurement of total RNA concentration usually was not conducted due to the downstream normalization strategy.

2.2.6.2 Reverse transcription and PCR (RT-PCR)

Reverse transcription and PCR were performed for only a limited amount of samples using a SuperScript III Reverse Transcriptase and the supplemented components following the manufacturer's instructions. Later on, all reverse transcription reactions were performed using the qScript Flex cDNA Kit (Quanta Biosciences) according to the manufacturer's instructions. For each RT reaction 1 µg of DNase I-treated total RNA from the step 2.2.6.1 was used in 8 µl mixture containing 2 µM of each gene-specific RT primer (see 2.1.8.2): NS5A/NS5B_minus_RT, Spinach_minus_RT or EMCV_plus_RT. Following incubation for 5 min at 65°C, 2 µl of reverse transcriptase reaction solution was supplemented and cDNA synthesis was performed for 30-60 min at 42°C, followed by the enzyme inactivation for 5 min at 85°C. Distinct conditions were applied when a snap RT primer was utilized for cDNA synthesis of Spinach region to ensure the most specific primer annealing. In detail, after heating the mixture of total RNA and primers up to 70°C, temperature was decreased at a rate 1°C per minute until reached 60°C. After that samples were brought to room temperature and then a reverse transcriptase reaction solution was added; cDNA synthesis was shortened up to 20 min and enzyme inactivation was carried out at standard conditions.

The downstream end-point PCR was performed by OneTaq DNA Polymerase using 1 µl of obtained cDNA together with a pair of gene-specific nested primers (for the detection of NS5A/NS5B/Spinach regions on minus strands or EMCV IRES region on plus strands). The general composition of a standard 25 µl reaction was as described above (see 2.2.3.3) using 1 µl of obtained cDNA together with a pair of gene-specific primers (see 2.1.8.2). An optimal temperature profile was the following: initial denaturation for 2 min at 95°C; 25 cycles of subsequent denaturation (for 15 s at 95°C), annealing (for 15 s at 60°C) and elongation (for 30 s at 68°C); final extension for 3 min at 68°C; hold is at 4°C. The PCR products were visualized by 1.5 % agarose gel-electrophoresis (see 2.2.4.1) or by 6 % PAGE (see 2.2.4.3).

2.2.6.3 Quantitative RT-PCR (RT-qPCR)

Since RT-PCR is only a semi quantitative method, in order to discover and evaluate small differences in HCV RNA synthesis, a quantitative RT-PCR protocol was established. Reverse transcription was performed using the qScript Flex cDNA Kit according to 2.2.6.2. The qPCR was performed using the PerfeCTa SYBR Green FastMix according to the manufacturer's instructions. In more detail, separate SYBR Green master mixes were prepared for each nested qPCR primer pair (2 µM each; for the detection of NS5A/NS5B/Spinach regions on minus strands or EMCV IRES region on plus strands) and distributed onto a colorless 96-well PCR plate. 2 µl of each cDNA sample produced with specific RT primers were applied into corresponding wells in 2 or 3 technical replicates (each final reaction volume is 20 µl). All measurements were conducted using Eppendorf Mastercycler ep Realplex^2 S with the following temperature profile: initial denaturation for 2 min at 95°C; 40 cycles of subsequent denaturation (for 15 s at 95°C) and elongation (for 30 s at 60°C); melting curve for 20 min; hold is at 4°C. Specifically for quantification of cDNA obtained with a snap RT primer (for Spinach region only) annealing/elongation step was carried out at 61°C in order to improve assay target-specificity.

The values for the quantitative analysis were obtained from at least three independent experiments. Data was processed using the provided with the Realplex^2 software to determine the threshold cycle (Ct) and analysis was performed according to Pfaffl (Pfaffl 2001) as described below.

2.2.6.4 RT-qPCR data analysis

To begin the analysis, amplification efficiencies were determined for each of the primer pairs used: for the NS5A/NS5B or Spinach regions on minus strands and EMCV IRES region on plus strands (also see Supplementary Materials 6.2.4). To accomplish this, a 10-fold dilution series of cDNA generated from the SL I-III construct transfected sample was amplified in qPCR. Measured Ct values were plotted against the common logarithm of the dilution factor, and the slope was derived from the plot. Amplification efficiency (E) was calculated using the following formula: $E = 10^{(-1/slope)}$. Mock-transfected cells were used as a control of primers' specificity in melting curves analysis (see 6.2.4).

The value used to compare expression of certain RNA species ("target", t; corresponds to either NS5A, NS5B or Spinach region on minus strands) to an expression of a reference gene of choice ("reference", ref; corresponds to EMCV IRES region on input plus strands) is designated as a "Relative Expression Ratio" (RER) and can be calculated in accordance to the formula:

$$RER = \frac{\left(E_t\right)^{\Delta Ct_t\,(control-sample)}}{\left(E_{ref}\right)^{\Delta Ct_{ref}\,(control-sample)}}$$

E_t and E_{ref} are the respective amplification efficiencies; ΔCt_t and ΔCt_{ref} are the Ct deviations of [control - sample] of the target or reference RNA, respectively.

Calculations were made using Microsoft Excel 2010. Means and standard deviations (SD) were calculated from the relative expression ratios (RER). Data is represented as Mean +/- SD. Statistical significance was calculated by a two-tailed Student's T-Test (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$), with a p-value > 0.05 considered not significant. For all RT-qPCR reactions, abundance of detected minus strands was normalized to abundance of plus strand RNAs in the total RNA recovered from the cells. In the resulting figures, error bars show standard deviation between at least three independent experiments.

2.2.7 Protein analysis methods: western blot

For the detection of specific proteins, western blot analysis was applied. All experiments aimed to compare the HCV NS3 protein abundance in cell lysates at a certain time-point after transfection with RNA replication constructs. Detection of GAPDH was conducted in parallel and served as a loading control for

evaluation of HCV protein expression. Prior the analysis, cells transfected in a 12-well plate were lysed by NP-40 lysis buffer (150 µl per well) for 30 min at 4°C. Afterwards cell debris were pelleted by centrifugation at maximum speed (13 000 rpm) for 30 min at 4°C. Next, the supernatant was transferred into a clean tube for the analysis. Separation of the proteins was performed by SDS-PAGE (see 2.2.4.5) usually using 10 % of the obtained lysate. Directly after the gel run, proteins were transferred to a polyvinylidene difluoride (PVDF) membrane by semi-dry electroblotting. The order of layering is the following: first, three pieces of Whatman paper soaked in Anode buffer I, next, three pieces soaked in Anode buffer II, then a PVDF membrane (needs to be activated my immersion in methanol shortly before use), after that the gels are carefully placed on the membrane (a stacking gel is usually removed) and covered by three pieces of paper soaked in Cathode buffer. The size of a membrane and Whatman paper is chosen in accordance to the number of gels. After assembly protein transfer to the membrane is performed electrophoretically by application of 43 mA per gel for 1.5 h. When the transfer is completed, the membrane is blocked in a blocking 1 x TBS solution for 1 h at room temperature (or overnight at 4°C). At the next step, the membrane is to be incubated subsequently with primary antibodies (a mixture of anti-NS3 and anti-GAPDH antibodies) and then with the secondary (anti-mouse) HRP antibodies. All antibodies are diluted in a blocking TBST solution to their working concentrations (see 2.1.5). All incubations are carried out at room temperature under constant agitation and followed by three-step washing with 1 x TBST (each step takes 5 min). At last, the membrane is washed twice in 1 x TBS to remove detergent and incubated with a SuperSignal West Femto Chemiluminescent Substrate solution for 5 min at room temperature (1 ml of working solution per analyzed gel was applied). Eventually the membrane was thoroughly cleared from the solution and exposed to X-ray film for various time intervals (usually, 15 s to 15 min).

2.2.8 Ribonuclease protection assay (RPA)

Ribonuclease protection assay (RPA) is a sensitive technique that allows detection, quantification and mapping of specific RNAs within total cellular RNA pool. Principally, the assay is based on hybridization of a radioactively labeled target-specific single-stranded RNA probe to RNA species of interest and therefore on probe's protection from degradation upon treatment with one or a mixture of endoribonucleases (RNases). In a classical procedure RNase A or a mixture of RNases A/T1 are applied: RNase A mediates specific cleavage within a single-stranded RNA at C and U residues, RNase T1 specifically hydrolyzes at G residues. Following hybridization, all remaining RNA species and unbound probe RNA are hydrolyzed and removed from a solution together with RNases by subsequent purification procedure including Proteinase K. In order to quantify protected fragments reflecting an amount of target RNA, the probe/target hybrids are precipitated and separated on a denaturing polyacrylamide gel. Denaturing conditions combined with detection by autoradiography visualize exclusively a radioactive probe RNA that remains intact proportionally to initial amount of RNA of interest in experimental total RNA sample.

The initial step of the assay comprises a generation of α-^{32}P-labeled RNA probe that is complementary to the target. Probe RNA is usually synthesized by *in vitro* transcription from a suitable template using a T7 or SP6 promoter. When promoter is encoded within a plasmid DNA template, a linearized form of the latter is utilized directly as a transcription template (see 2.2.3.4). Alternatively, a promoter can be introduced within a PCR primer, and then a PCR product serves as a template for *in vitro* RNA synthesis. *In vitro* transcription from any of template types is conducted according to a standard protocol for radio-labeled RNA synthesis using a relevant - T7 or SP6 – RNA polymerase and followed by the DNA template removal with DNase I enzyme (see 2.2.3.4). The resulting probe is purified by phenol-chloroform extraction and ethanol precipitation (see 2.2.3.5 and 2.2.3.6), resuspended in an appropriate amount of RNase-free water and its integrity is verified by denaturing polyacrylamide gel electrophoresis (see 2.2.4.4).

Preparation of experimental total RNA is conducted on the next step. Following transfection and incubation of cells as required for an experiment, cells are lysed using the NP-40 lysis buffer for 30 min at 4°C with a subsequent centrifugation similarly to as described in 2.2.7. The supernatant is then transferred into a new tube and supplemented with 1 % SDS, 10 µg tRNA, $CaCl_2$ (to a final concentration of 1 mM) and

30 µg of Proteinase K. To facilitate a Proteinase K digest, the samples are then incubated at 50°C for 1 h. Proteinase K is a subtilisin-related serine protease that hydrolyzes a variety of peptide bonds in a wide range of temperatures and buffers and serves to remove enzymes and proteins from enzymatic reactions or cell lysates, respectively. Isolated total RNA is eventually extracted with phenol-chloroform, precipitated with ethanol (see 2.2.3.5 and 2.2.3.6) and resuspended in RNA hybridization buffer.

Hybridization of combined isolated experimental total RNA and radio-labeled probe RNA (in amounts equivalent between all samples) is initiated by rapid denaturation for 5 min at 85°C and subsequently performed on a heating block at 42°C for at least 16 h (usually overnight). Shortly before use, RNase A is diluted in RNase digestion buffer to working concentration of 40 µg/ml. At the end of hybridization, 10-fold amount of completed RNase digestion buffer is added to the samples; the final solutions is then carefully mixed and incubated for 90 min at 30°C. In order to terminate the RNase digest and to purify the remaining RNA, samples are supplemented with 0.5 % SDS, 10 µg tRNA, $CaCl_2$ (to a final concentration of 1 mM) and 100 µg of Proteinase K. The Proteinase K digest and a following purification procedure are conducted as described above (with an exception of RNA precipitation in the presence of NaCl instead of NaOAc).

At the last step of the procedure remaining probe RNA protected by the target RNA is visualized after separation on a denaturing polyacrylamide gel (polyacrylamide concentration depends on the size of a probe) by autoradiography (see 2.2.4.4).

2.2.9 RNA immunoprecipitation (RIP)

RNA immunoprecipitation (RIP) is a standard approach applied in attempt to demonstrate RNA-protein interactions. In a simplified experimental setup the procedure can be conducted *in vitro*, however most commonly a cellular lysate is used as a source of protein: either containing the RNA-protein complex of interest or the complex is formed upon introduction of a target RNA by transfection. For the latter, the principle of the assay is based on visualization of radioactively labeled target RNA that is pulled down in a complex with its interaction partner - protein - due to binding by a specific antibody. In the current work HCV 3'UTR interaction with a number of aminoacyl-tRNA synthetases (ARS) was challenged. A role of a positive control was given to the PTB protein that is well-known to bind to polyU/C tract at the HCV 3'-end. When the selected ARSs demonstrate an ability to stably bind to predicted sites within the HCV 3'UTR, it can be detected and evaluated by autoradiography.

At the initial step radio-labeled HCV 3'UTR RNA is generated by *in vitro* transcription (see 2.2.3.4), quantified and transfected into eukaryotic cells (HeLa cells) using Lipofectamine 2000 (see 2.2.2.4) for 4-6 h. A fraction of the radioactive RNA transcript is saved to further serve as a size marker. Next, transfected cells are lysed by NP-40 lysis buffer (similar to as in 2.2.7). A portion of cell lysate at this step is also saved to serve as an input loading control. In the meanwhile, antibodies specific to candidate interaction proteins as well as to the PTB (positive control) and FLAG-epitope (negative control) are bound to protein G magnetic beads. In detail, 100 µl of protein G beads suspension is washed with 0.5 ml PBS using magnetic rack and eventually resuspended in 1 ml of PBS. Washed beads are combined with each antibody in a separate tube and incubated on a rotating wheel for at least 3 h (or overnight) at 4°C. This step allows binding of conservative Fc-regions of antibodies to the protein G and therefore mediates their immobilization at the beads surface. Importantly, an amount of each antibody applied for pull-down has to be optimized in advance based on the manufacture's recommendations. Unbound antibodies are removed by washing with 1 ml of IP washing buffer, the beads are again resuspended in 0.5 ml PBS.

At the next step, beads coupled to various antibodies are combined with identical portions of cell lysate and incubated similarly on a rotating wheel for at least 3 h at 4°C. After that the beads are washed 4 times using 1 ml of IP washing buffer to eliminate unspecific binding and resuspended in 1 ml of PBS. Subsequent Proteinase K treatment and purification procedures are carried out for all experimental samples and for RNA transcript and input cell lysate: Proteinase K is added to a final concentration of 20 mg/ml to release bound RNA during incubation for 15 min at 65°C. Further enzyme removal and target RNA isolation

is performed by phenol-chloroform extraction and ethanol precipitation (see 2.2.3.5 and 2.2.3.6), and resulting pellets are resuspended in appropriate amounts of formamide-containing loading buffer. Lastly, the results are analyzed by denaturing polyacrylamide gel electrophoresis (see 2.2.4.4), followed by autoradiography. Efficiency of target RNA recovery mediated by binding to a candidate protein interaction partner is evaluated when radioactive signals are compared to the ones in intact cell lysate (equivalent relative amounts loaded).

3. Results

3.1 Design and establishment of the minus strand replication system

The concept of a novel *in vitro* system that is able to bypass limitations of the current model systems was carefully considered in accordance to the existent information on the requirements for the HCV minus and plus RNA synthesis. Therefore, minimal constructs for separate initiation of both strands were designed in a way to be relatively easily modified (extended or diminished). As emphasized above, the roots of major limitations of the available *in vitro* systems – the replicon system and its derivatives as well as the full-length infectious system (Fig. 3.1.1, A) – are in simultaneous presence of both 5'- and 3'- genomic ends that enable HCV translation and full replication cycle. Since quantification of an output from a certain modification depends on the translation or/and replication performance, these systems yield no mechanistic insight when a mutation abrogates any of the processes. Even if a mutation is not "lethal", it still remains challenging to define at which exact stage the affected *cis*-element performs its function. In order to clarify the uncertainties provided by the current replication models, the novel reduced replication system was developed in the present work.

The separate minus and plus strand replication systems were designed to specifically uncouple minus strand synthesis from the plus strand replication and both from translation and other stages of the HCV life cycle. Each HCV RNA strand is to be initiated from an individual RNA template (Fig. 3.1.1, B and C) that contains the essential (according to the current knowledge) elements for a desired strand synthesis: the 3'-end of the corresponding complementary strand and the HCV non-structural (NS) proteins encoded in *cis*. Since the present work predominantly focuses on the dissection of requirements for the HCV minus strand synthesis, this system (Fig. 3.1.1, B) is further featured in detail.

The DNA elements for both systems were obtained by a commercial chemical synthesis and assembled into the minimal replication constructs, as thoroughly described and illustrated in the Supplementary material 6.2.1. Most of the functional elements are common for both systems. In the absence of a 5'-cap – as it is for the HCV genome – prevention of a rapid exonuclease degradation of experimental RNA construct is accomplished via a stable artificial stem-loop encoded directly downstream the T7 promoter on the DNA template. The hairpin (hp) is followed by a Streptavidin S1 aptamer (not depicted) (Srisawat and Engelke 2001) included to enable isolation and analysis of replication complexes that assemble at the 3'-end of the replicating strand. Subsequent Spinach aptamer (Paige et al. 2011; Ouellet 2016) encoded as a reverse complementary sequence surrounded by an appropriate scaffold to eliminate interference of neighboring sequences on the aptamer folding. Obtaining a functional secondary structure upon a complementary strand synthesis, this aptamer could be utilized for an intracellular visualization, localization and/or quantification of the newly produced HCV RNA strands. Additional information on structure and function of the Streptavidin and Spinach aptamers is given in the Supplementary material 6.2.2. The next vital component of the replication constructs is an Encephalomyocarditis Virus (EMCV) IRES that drives translation of the HCV non-structural proteins NS3 to NS5B independently of the HCV IRES. The replication module proteins are indispensable for the HCV replication *in vitro* and most of them are required in *cis*, however translation initiated by the HCV IRES situated on the same RNA substrate would largely disadvantage dissection of overlapping functions. The EMCV IRES in turn is not regulated by the same mechanisms as the HCV IRES, therefore performs a consistent and unbiased by unrelated mutations translation of the essential HCV proteins. The encoded NS3-NS5B gene cassette fully corresponds to the JFH1 isolate since it was derived from the Jc1 plasmid (for the plasmid map see 6.1.1) encoding full-length HCV J6/JFH1 chimeric genome (Pietschmann et al. 2006). Besides the wild-type NS3-NS5B sequence, a respective scrambled gene cassette was considered in the design. Scrambling of a protein coding region here stands for permutation of a nucleotide sequence in a way to disable genomic RNA *cis*-signals while retaining amino acid sequence, coding specificity, codon usage and codon pair bias. The resulting experimental constructs containing the scrambled NS3-NS5B sequence give raise to an unaffected replication complex,

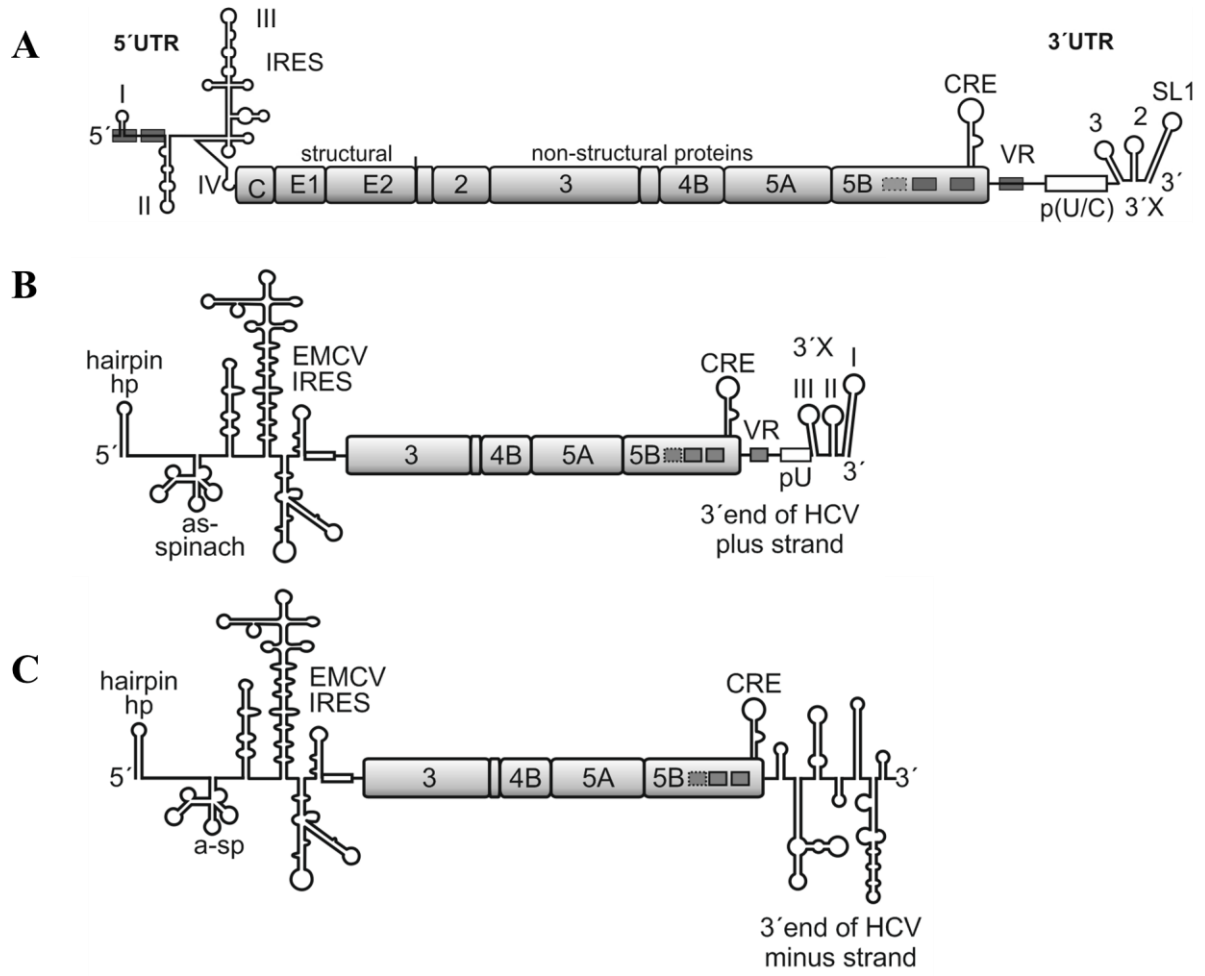


Figure 3.1.1: Functional elements of the replication system.
(A) A schematic of the HCV plus strand RNA genome with depicted 5'- and 3'- UTRs, the IRES, an open reading frame encoding structural and non-structural proteins. The IRES domains are labeled by roman numbers, the 3'UTR structural elements – by arabic numbers. The *cis*-acting replication element (CRE) 5BSL3.2, variable region (VR), poly U/C tract (p(U/C)) and the X-tail (3'X) are depicted as important determinants of the viral replication.
(B-C) The basic components of the developed replication system: template RNAs for the minus (B) and plus (C) strand synthesis sharing common structural elements such as a 5'-terminal hairpin (hp), a reverse complementary sequence for the Spinach aptamer (a-sp), the EMCV IRES followed by the HCV NS3-NS5B coding region containing the 5BSL3.2 element (CRE). The system for the minus strand synthesis contains the 3'-end of the HCV plus strand, but not the HCV 5'UTR; the system for the plus strand synthesis - the 3'-end (including a partial Core-coding sequence) of the HCV minus strand.

but lack the *cis*-acting replication elements. Such version of the NS3-NS5B sequence was implemented to both constructs at the assembly step and required a synthetic production due to a complexity of mutations to be introduced. The computational procedure of scrambling was fully performed by Yutong Song and Steffen Mueller. The full-length sequences of the wild-type and scrambled NS3-NS5B cassettes are detailed in the Supplementary material 6.2.3. In order to maintain a wild-type sequence of the 5BSL3.2 element, which was reported to be essential for the HCV replication (Friebe et al. 2005; You and Rice 2008), and to optimize the complex assembly procedure, the scrambling was applied to the coding sequence between the *Bbv*CI and *Asc*I sites (also see Supplementary material 6.2.1). This retains 238 nt of the wild-type sequence upstream the *Bbv*CI site and 152 nt downstream of the *Asc*I site, which can be further mutated by conventional methods if required. Notably, in all scrambled constructs the NS5B miR-122 binding sites (5B.1, 5B.2 and 5B.3) become inactivated. In both basic DNA templates (Fig. 3.1.1, B and C) the 3'-end of the ORF is directly merged to the HCV 3'UTR of a complementary strand followed by the Hepatitis Delta Virus (HDV) ribozyme (Shih and Been 2002). It is of a great importance to have the ribozyme fused to the very last 3'-nucleotide of the functional sequence to obtain an RNA substrate with the exact HCV 3'-end after *in vitro* transcription. The HDV ribozyme acts in a self-cleavage manner and does not require any protein factors. As an additional measure, the ribozyme sequence was accompanied with the T7 Terminator (T7T) to prevent

continuous transcription and to ease a cleavage by the ribozyme. Additional information on structure and function of the HDV ribozyme is provided in the Supplementary material 6.2.2.

In the basic construct for the plus strand synthesis (Fig. 3.1.1, C; for the plasmid map see 6.1.11) the 3'-end of the template RNA comprises the complete reverse complementary sequence of the plus strand HCV 5'UTR and the first 325 nt of the HCV Core-coding sequence intending to maintain the folding of the 3'-end of the minus strand as well as to include the *cis*-signals opposite to the coding region downstream the IRES. The minimal – designated as "hp" - construct for the minus strand synthesis (Fig. 3.1.1, B; for the plasmid map see 6.1.7) harbors the 3'-end of the plus strand which is indispensable for the initiation with the HCV replication complex. Generated in wild-type or scrambled version (6.1.7 and 6.1.8) of the NS3-NS5B coding sequence, these constructs represent the basis for further modifications and mutations. The replicase-deficient version of this construct (shortly, hp_GND; 6.1.23) was generated to serve as a negative control of minus strand synthesis initiation: the point mutation 318D→N within the NS5B RdRp gene completely inactivates the NS5B replicase (Lohmann et al. 1997).

Prior to any modifications of the basic minus strand synthesis construct, preliminary tests were conducted to ensure functionality of the T7 promoter and of the EMCV IRES. Following the plasmid template linearization by *Eco*RI (the unique site located downstream the T7T), the *in vitro* transcription with a T7 RNA polymerase was established. Efficient and high-yield *in vitro* transcription of long RNAs required optimization of a standard protocol provided for the T7 polymerase; the conditions are detailed in 2.2.3.4. The *in vitro* transcribed RNA construct was visualized by 1 % agarose gel electrophoresis comparing the sample RNA before and after treatment with DNase I (Fig. 3.1.2, A). The polyprotein production (driven by the EMCV IRES) was challenged by a western blot detection of the HCV NS3 protein (Fig. 3.1.2, B). The HCV NS3 content was monitored in HuH-7.5 cell lysates derived at 6 h, 12 h and 24 h post transfection (h.p.t.) and compared on a gel to the lysate of mock-transfected cells (24 h.p.t. only). The rapid accumulationof the viral protein was observed already 6 hours post transfection followed by a steady decay of both thetranslation template and the viral protein, consistently to the research on half-lives of viral non-structural proteins shown to be 11–16 h (Pietschmann et al. 2001; Pause et al. 2003). Also in the

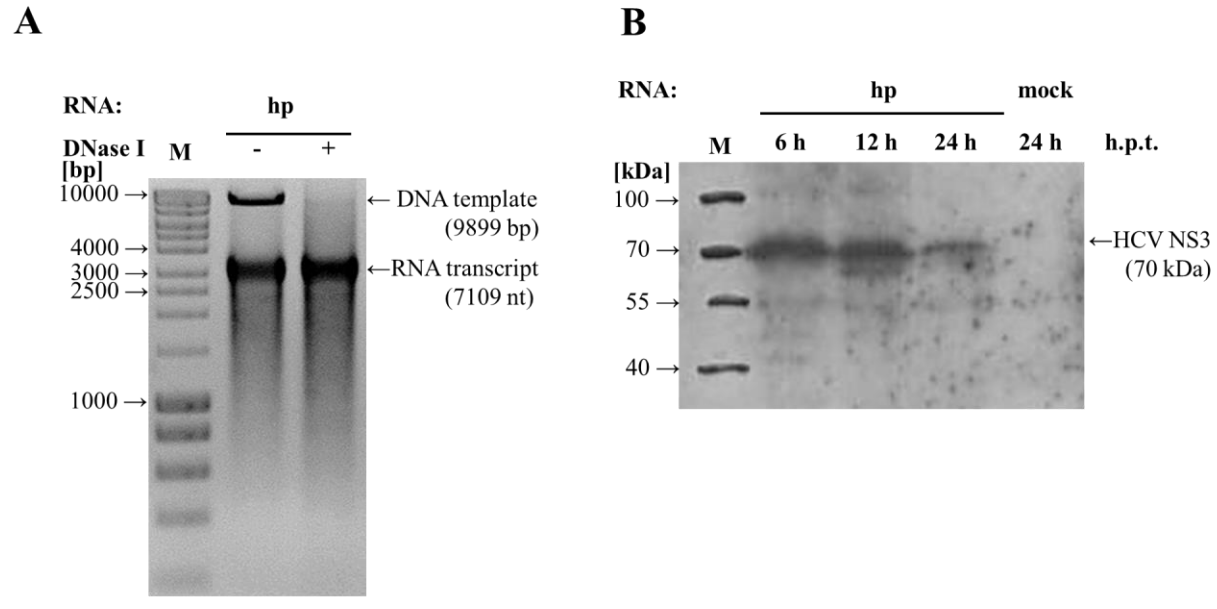


Figure 3.1.2: Validation of the designed minimal hairpin (hp) construct functionality.
(A) *In vitro* transcribed RNA from the linearized plasmid DNA template: prior (-) and after (+) a removal of the DNA template upon treatment with DNase I. Visualization is conducted on 1 % agarose gel using ethidium bromide staining. The RNA transcript for the hp construct is 7109 nt long, however it migrates at the 3 kb level, since compared to the dsDNA marker (M, GeneRuler DNA Ladder Mix).
(B) Western blot detection of the HCV NS3 protein (70 kDa) in HuH-7.5 cells lysate 6, 12 and 24 h post transfection (h.p.t.) with the hp RNA construct. No HCV NS3 is detected in mock-transfected cell lysate (24 h post transfection). M, protein marker ladder.

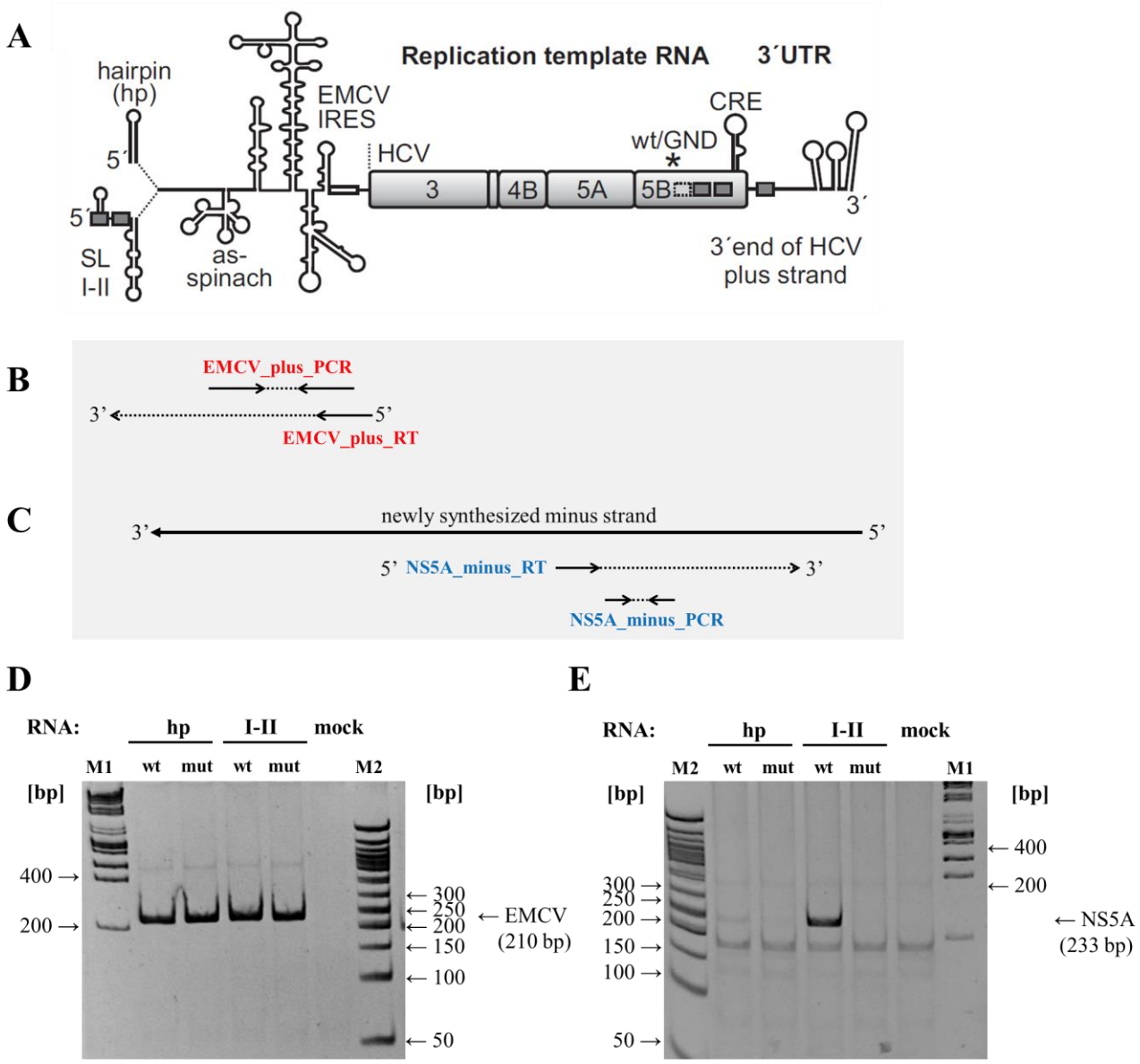


Figure 3.1.3: The basic experimental system for the analysis of the HCV minus strand synthesis initiation.
(A) The replication RNA templates used in the experimental setup: the minimal hairpin (hp) construct and its derivative containing SL I-II sequence from the HCV 5'UTR. Within those constructs the NS5B replicase is either wild-type (wt) or inactivated by the GND mutation (mut) of the active site.
(B) A scheme of RT-PCR detection of transfected plus strand templates (input) using a specific RT primer and a pair of nested PCR primers targeting the EMCV IRES sequence.
(C) A corresponding RT-PCR detection of the newly produced minus strands with a specific RT primer and a pair of nested PCR primers targeting the reverse complementary sequence of the NS5A coding region. All primers are universal for the examined constructs.
(D-E) Visualization of the HCV plus (D) and minus strand (E) content in total RNA derived from HuH-7.5 cells 2 days post transfection with the experimental constructs (from A). The product of the nested PCR (210 bp) targeting the EMCV IRES region on plus strands obtained for all, except mock-transfected, samples indicating the presence of input template. In contrast, the PCR product targeting the complementary NS5A coding sequence (233 bp) was identified only in the sample transfected with the wild-type SL I-II template RNA. M1, a DNA HyperLadder 1 kb; M2, GeneRuler 50 bp DNA ladder.

infectious HCV system, large fraction of the NS proteins is eliminated after early translation rounds when the replications sites are established. However, the hp construct here lacks an ability to perform continuous replication, therefore new plus strands are not generated to serve as translation templates.

After successful transfection of the HuH-7.5 cells with the *in vitro* generated HCV RNA replication template, the basic hp construct was challenged for ability for the HCV minus strand synthesis initiation. The minimal hp construct lacks any HCV 5'UTR sequences, therefore in the replication assay it wascompared to its extended version containing HCV SL I-II at the construct's 5'-end (Fig. 3.1.3, A). The

latter RNA construct (designated as "SL I-II") was generated from the linearized plasmid (for the plasmid map see 6.1.13) following the analogous procedure. The HCV SL I-II was reported to be the minimal essential 5'UTR sequence in a replicon system required for HCV replication (Friebe et al. 2001), in part due to the S1 and S2 miR-122 binding sites located at the single-stranded stretches of the SL I. For the primary replication assay, the replicase-deficient variants of both – the hp and the SL I-II constructs – were utilized as negative control of the minus strand synthesis initiation; generated from plasmids 6.1.23-24. The four RNA constructs were transfected to the HuH-7.5 cells via the Lipofectamine method (see 2.2.2.4); 48 hours post transfection the cell lysis with Trizol reagent was performed (2.2.6.1), followed by total RNA isolation and DNA removal. The total RNA after transfection is expected to contain the input RNA plus strand templates (Fig. 3.1.3, A) and, if the construct is capable of the minus strand synthesis initiation, also the newly synthesized minus strands. Detection of the plus and minus strands was accomplished by region specific RT-PCR targeting common for all experimental constructs: EMCV IRES and NS5A (reverse complementary of the coding sequence) regions, respectively (Fig. 3.1.3, B and C). The resulting PCR products (after 25 and 30 cycles, respectively) were visualized by 8 % PAGE (Fig. 3.1.3, D and E). As a result, for all tested plus strand constructs – hp_wt/GND and SL I-II_wt/GND (wt stands for the intact replicase in contrast to the GND-mutant) – the input template RNA was still present and detectable in cellular total RNA fraction 2 days post transfection (Fig. 3.1.3, D). In the sample corresponding to the mock-transfected cells no PCR product was obtained underlining that the nested primer pair detecting EMCV IRES is specific to the HCV, but not cellular, RNA. Notably, PCR targeting the complementary (on the minus strand) sequence of the NS5A coding region resulted in a specific signal for the SL I-II_wt construct only (Fig. 3.1.3, E). Neither the replication-deficient constructs (hp_GND and SL I-II_GND) nor mock-transfected sample led to amplification of specific fragment of expected length. Since minus strands upon HCV infection are generally significantly less abundant than plus strands, the visualization of the minus strand synthesis required high amount of total RNA material and high amplification cycle number. This inevitably results in background bands which however appear for all experimental samples, including the mock control, and migrate unrelatedly to the expected product length. However, it is not clear from the electrophoretic image whether the minimal hp construct yields any minus strand synthesis or in absence of the HCV 5'UTR sequences becomes incapable of the synthesis initiation. Answering this question would require a more sensitive approach. Nevertheless, the present replication assay clearly demonstrates that SL I and II of the HCV 5'UTR are sufficient for the minus strand synthesis initiation in accordance to the data obtained in replicon system. The RT-PCR was found to be a useful preliminary method to estimate minus strand replication ability of experimental RNA constructs, however some signals may appear beyond its detection level and/or overlaid by the background. In course of the project a detection of the HCV plus and minus strands was attempted using the ribonuclease protection assay (2.2.8), however a sensitivity of the method appeared insufficient for detection in the current experimental setup (for more details refer to Supplementary results 6.3.1). Eventually a quantitative RT-PCR method (2.2.6.3) was utilized in the present study replacing the semi-quantitative RT-PCR.

To conclude, the system attempting to separate HCV minus and plus strand synthesis from each other and from HCV translation was designed and proved to be functional. Mindfully generated RNA substrates are able to drive the HCV translation from heterologous IRES upon transfection into hepatoma cells and after certain sequence improvement – to serve as a template for the minus strand synthesis. Further sequence extensions and modifications and their effects on replication are to be presented in the following section.

3.2 Dissection of the HCV 5'UTR sequence requirements for minus strand synthesis initiation

In agreement with the known requirements for the HCV replication, the minus strand synthesis initiation required a presence of the SL I-II sequence of the HCV 5'UTR within the replication template. However, due to multiple functions that those and other domains of the HCV 5'UTR serve in the viral life cycle, it is a challenge to define if these sequences actually are involved in minus strand initiation. While some sequences and structures only have an indirect effect on the antigenome synthesis, the others are essential for the replication complex activity at the genome's 3'-end. This section provides an insight into the role of the HCV 5'UTR functional blocks in the initiation of the minus strand synthesis.

The minimal construct for the minus strand synthesis initiation (Fig. 3.1.1, B) was extended by partial HCV 5'UTR sequences replacing the terminal hairpin at the template's 5'-end (Fig. 3.2.1, A). The modification of the basic construct harboring SL I-II of the HCV 5'UTR was already featured above; it allows engaging of the two miR-122 binding sites at the 5'-end of the experimental construct that were reported to promote HCV translation (Henke et al. 2008) and replication (Jopling et al. 2005) and to protect the genomic RNA from Xrn1-mediated degradation (Shimakami et al. 2012a). The second modification

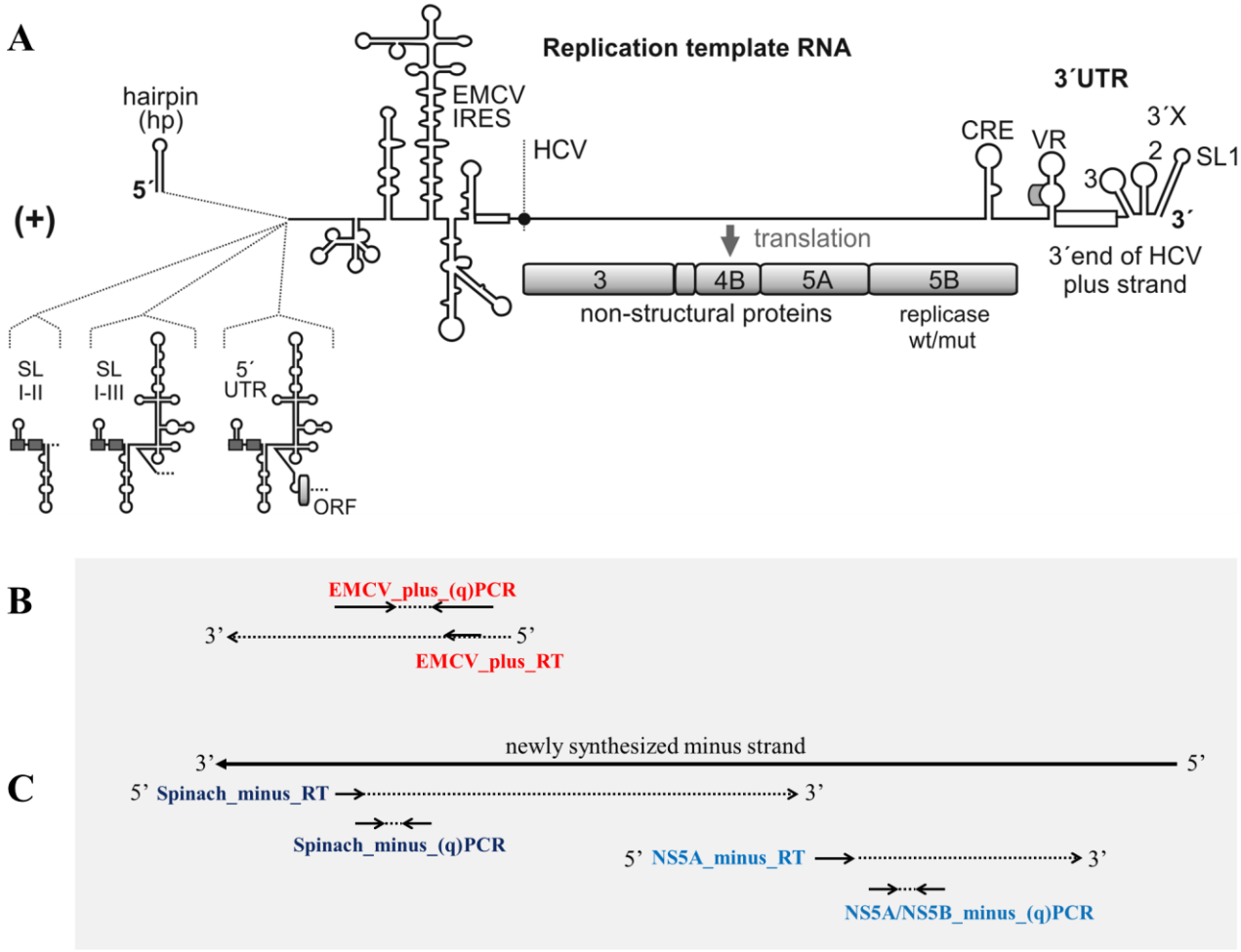


Figure 3.2.1: The 5'-end modifications of the basic replication construct for the analysis of the HCV minus strand synthesis initiation.
(A) Replication RNA templates with distinct sequences at the 5'-end containing either a hairpin (the minimal hp construct), the HCV SL I-II, SL I-III or the complete HCV 5'UTR containing an initiation codon and a short open reading frame.
(B) A scheme of RT-(q)PCR detection of input plus strand templates using a specific RT primer and a pair of nested (q)PCR primers targeting the EMCV IRES sequence.
(C) A corresponding RT-(q)PCR detection of the newly produced minus strands with a specific RT primer and a pair of nested (q)PCR primers targeting the reverse complementary sequence of either the HCV NS5A/NS5B coding region or the Spinach region. All primers are universal for the examined constructs.

comprises the entire SL I-III region (the construct is designated as "SL I-III") of the HCV 5'UTR and therefore permits, in addition to the miR-122 binding, assembly of the 40S ribosomal subunit and of the translation initiation factor eIF3 at the 5'-end of the replication template (Spahn et al. 2001). Lacking the AUG start codon located within the SL IV, this construct is incapable of initiation of functional translation from the HCV IRES. The third modification represents an extension with the complete HCV 5'UTR followed by a short open reading frame (36 nt) confined by an UGA stop codon. Thus, in virtue of the in-frame truncated Core-coding region, the "5'UTR" construct includes the complete and stable HCV IRES structure and an authentic AUG to fulfill a functional translation initiation event. Without interference with the downstream elements and their functions, the translation initiated at the template's 5'-end enables studying an intricate translation versus replication balance.

These replication constructs were generated by *in vitro* transcription from linearized DNA templates as previously (for plasmid maps see 6.1.7, 6.1.13, 6.1.15 and 6.1.18, respectively). The integrity of the full-length RNA was ensured by agarose gel electrophoresis (Fig. 3.2.2, A) and the functionality of the EMCV IRES, which mediates the HCV replication proteins production, was demonstrated by western blot (Fig. 3.2.2, B). Peculiarly, the effect of miR-122 binding on the stability of a replication template (SL I-II construct) was likely illustrated (Fig. 3.2.2, B, lane 2 vs. 3). Cellular lysates applied for the analysis were derived from HuH-7.5 cells transfected for 8 hours with each of the four described RNA templates either alone or together with a Miravirsen-like anti-miR-122 LNA/DNA mixmer, designated as "LNA" from here on, (applied with the SL I-II construct only). This LNA functionally sequesters endogenous miR-122, therefore preventing its function on the HCV RNA or cellular mRNAs. As clearly reflected by western blot, the basic function of the EMCV IRES is not affected by introduction of various 5'-end sequence modifications (lanes 1, 3-5), however the removal of miR-122 does profoundly reduce the HCV NS3 content (lanes 2 and 3). Since the EMCV IRES function is miR-122 independent, the observed drop in NS3 production is likely related to availability of a template after transfection, i.e. a template's stability. Considering this effect is observed independently of the nature of the 5'-terminal sequence, other indirect functions of the miR-122 or its function at the 3'-end of the HCV genome could have had an impact.

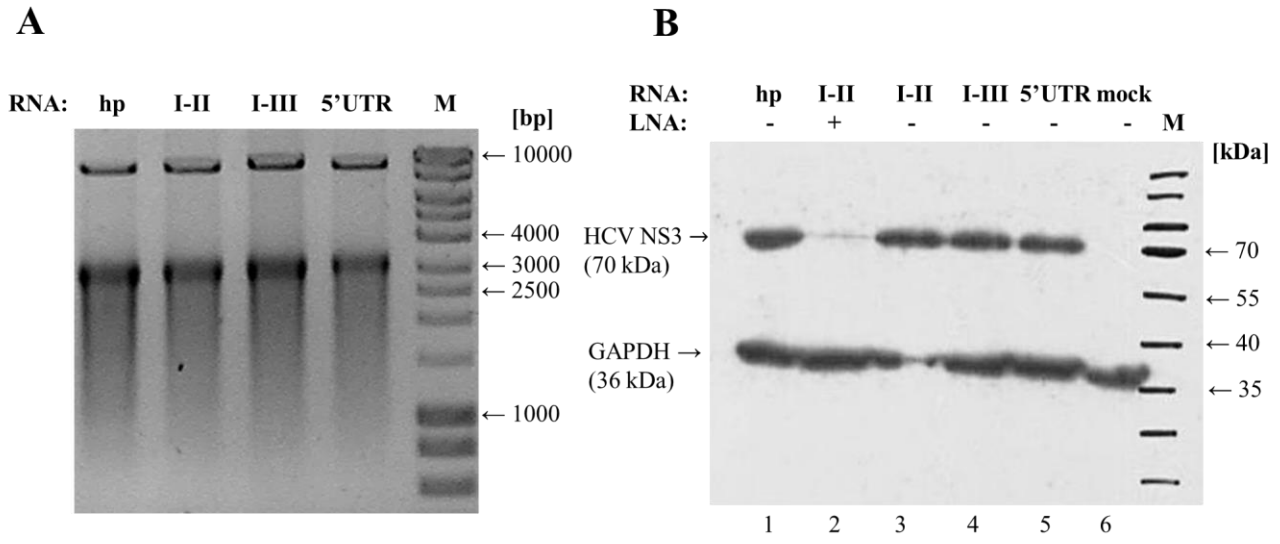


Figure 3.2.2: Verification of functionality of the extended 5'-end constructs.
(A) *In vitro* transcribed RNA for the minimal hp construct and its 5'-end extended versions containing either SL I-II, SL I-III or the complete HCV 5'UTR. Visualized prior the DNase I treatment, samples represent a mixture of integral RNA transcript and initial linearized DNA template. The analysis is conducted on 1 % agarose gel using ethidium bromide staining; the molecular length of both DNA templates and RNA transcripts is slightly variable in accordance to sequence extensions. The RNA transcripts are compared to the dsDNA marker (M, GeneRuler DNA Ladder Mix).
(B) Western blot detection of the HCV NS3 protein (70 kDa) and cellular GAPDH (36 kDa) as a reference protein in HuH-7.5 cells lysate 8 h post transfection with each of the experimental construct (hp or its 5'-end modified versions, as in A) or mock-transfected cells. Co-transfection of anti-miR-122 LNA together with a replication template is indicated as "+". M, protein marker ladder.

Further on, the extended replication constructs were compared in a replication assay. As previously, preliminary analysis was conducted by RT-PCR, later on fully replaced by a quantitative PCR method. Universal primers targeting the input plus strands and newly produced minus strands were applied for both kinds of PCR analysis. The specific RT primers are indispensable for the synthesis of cDNAs corresponding to one or the other viral strand: due to the relatively low abundance of the HCV plus and especially minus strands (the ratio between HCV plus and minus strands is from 1:10 to 1:100) random primers failed to generate consistent amounts of cDNA in the reverse transcription reaction from cellular total RNA pool (data not shown). Notably, in order to reveal slight changes in minus strand synthesis efficiency from different RNA constructs, the reverse transcriptase requires RNase H activity to remove HCV strands that have fulfilled a template function. The input plus strands in the current and all the downstream replication assays are detected by RT-(q)PCR targeting the EMCV IRES region (Fig. 3.2.1, B), as previously; the same nested PCR primers are used for both kinds of PCR. The newly synthesized minus strands, if can be generated from an experimental RNA construct, are detected by RT-(q)PCR targeting a reverse complementary sequence of either the HCV NS5A/NS5B or the Spinach coding region (Fig. 3.2.1, C) using the same nested PCR primers for all constructs analyzed by both types of PCR. Quantification of minus strands by targeting either Spinach or NS5A/NS5B regions provides different information. While nascent minus strands detected via the Spinach region stand for nearly full-length construct's complementary strands, the strands detected via the NS5A/NS5B region might not have resulted in the complete antigenome, if synthesis was abrogated for a certain reason. Therefore, the latter detection reflects to a higher extent the number of initiation events, whereas quantification of nearly full-length products represents a relative count of completed synthesis events.

Initially, ability of the extended 5'-end constructs to initiate minus strand synthesis was assessed semi-quantitatively by RT-PCR. Plus and minus RNA strands were detected in total RNA fraction isolated from HuH-7.5 cells transfected with each of the four RNA templates (Fig. 3.2.1, A) 2 days post transfection. The input plus strands were revealed for each experimental template (Fig. 3.2.3, A), while the newly synthesized minus strands were found only for the extended versions of the minimal hp construct (Fig. 3.2.3, B). Total RNA from mock-transfected cells sample does not provide any signals underlining the primers' specificity to HCV sequences.

Next, similarly derived cDNA samples were utilized for a quantitative analysis. Since the efficiency of minus strand synthesis initiation from a particular RNA construct is directly linked to its own abundance, the detected minus strand species were routinely normalized to the plus strand, i.e. input, content in the same total RNA sample. Thence, in spite of varying stability or potential ability to continuous replication (e.g. when both HCV genome ends are simultaneously present), minus strands count is relative to available plus strand templates. In the current section both – shorter and longer - elongation products are quantified via targeting a reverse complementary sequence of the NS5B and Spinach region (Fig. 3.2.1, B and C, respectively). The hp and SL I-II constructs harboring an inactivating GND mutation in the RNA-polymerase gene are regarded to as negative control, since any GND-mutant construct is incapable of antigenome synthesis. The values derived for this construct are indicated by dotted lines. In the present and most of the assays below, the mean relative expression ratio (RER) for the SL I-II construct was set to 1 to ease the comparison of the experimental constructs replication abilities. More aspects of the quantification approach are discussed in the Supplementary material 6.2.4.

As evident from the graphs (Fig. 3.2.3, C and D), the results derived from targeting either NS5B or Spinach regions, respectively, appear to be largely in agreement. The RER value obtained for the basic hp RNA template appeared not to significantly differ from both analyzed replicase-deficient mutants (hp_GND and SL I-II_GND; indicated by dotted line), implying that the minus strand synthesis initiation is not possible in the absence of essential HCV 5'UTR sequences at the 5'-end of the experimental construct. The previously demonstrated ability of the 5'-end extended derivatives of the basic hp construct to synthesize minus strand (Fig. 3.2.3, B) was hereby quantified. Verified by targeting both, NS5B and Spinach regions,

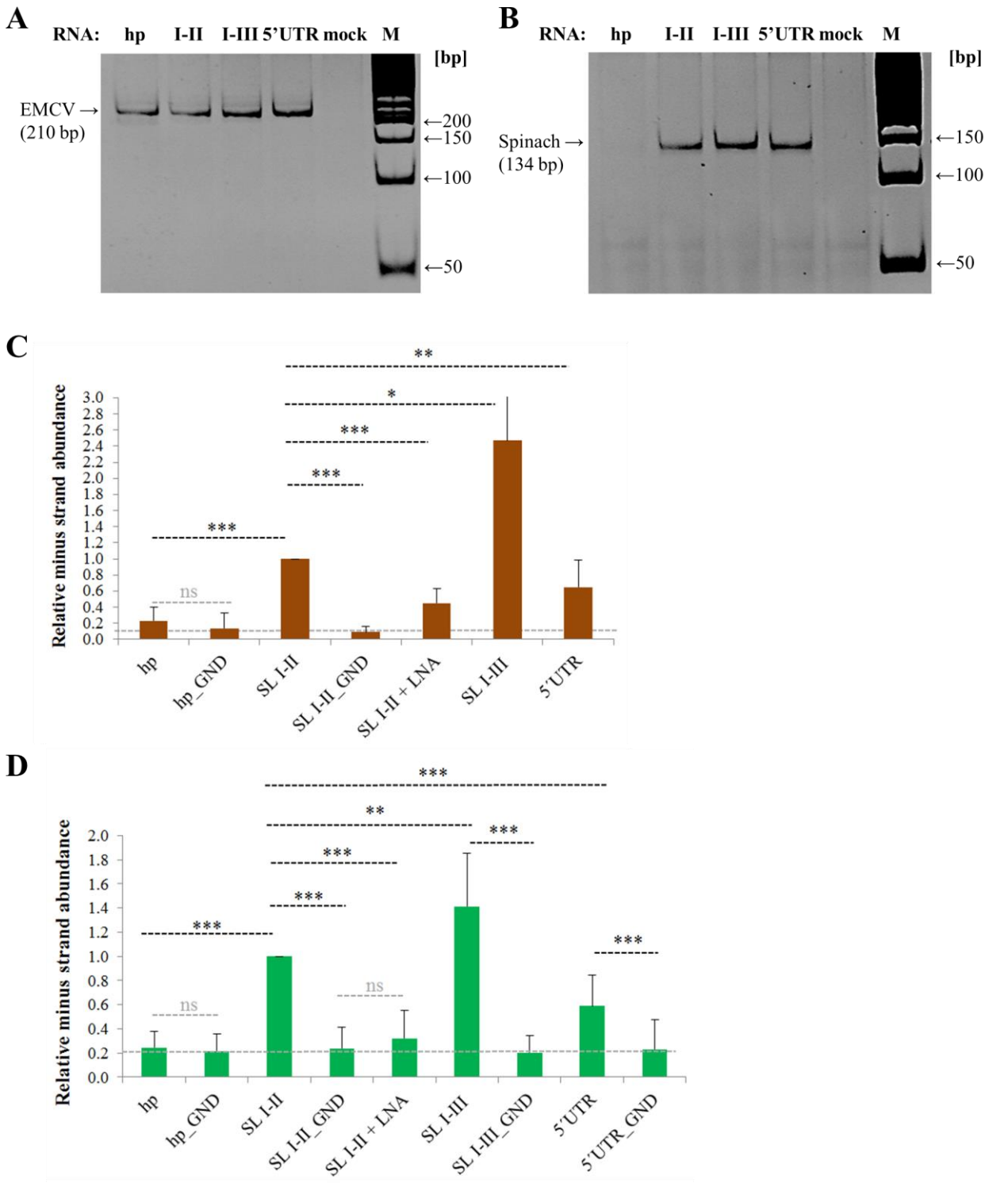


Figure 3.2.3: Illustration of the requirement for the HCV 5'UTR sequences in the minus strand synthesis initiation.
(A-B) Visualization of the plus (A) and minus strands (B) by RT-PCR in total RNA from HuH-7.5 cells 2 days post transfection with the experimental 5'-end constructs: hp, SL I-II, SL I-III and 5'UTR. The nested PCR product targeting the EMCV IRES region on plus strands (210 bp) indicates the presence of input templates. A product of nested PCR targeting the Spinach sequence (134 bp) is detected on the nascent minus strands. Total RNA from mock-transfected cells does not result in any signal supporting the primers specificity. M, GeneRuler 50 bp DNA ladder.
(C-D) Quantification of newly synthesized minus strands by RT-qPCR targeting the reverse complementary sequence of either the NS5B (C) or the Spinach coding region (D). Abundance of minus strands is normalized to plus strand abundance in the same RNA sample. The background signal derived for the replication-deficient hp_GND and SL I-II_GND constructs is labeled as a grey dashed line. The RER values obtained for the SL I-II construct are set to 1 and the fold difference serves as a measure for comparison. LNA stands for a Miravirsen-like anti-miR-122 LNA. Asterisks mark p-values from two-tailed student's T-tests (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ns: not significant). On the diagrams, error bars show standard deviation between at least three independent experiments. In both assays all primers are universal for the examined constructs.

the relative minus strand abundance increases with the addition of SL III sequence, but surprisingly drops, relatively to SL I-II and SL I-III constructs, with the inclusion of the complete HCV 5'UTR followed by a short ORF. As tested only in the Spinach region (Fig. 3.2.3, D), the GND mutants corresponding to SL I-III and 5'UTR constructs are unable of progeny genome synthesis. Expectedly, co-transfection of the SL I-II construct with anti-miR-122 LNA leads to dramatic reduction of minus strand synthesis, which is in correspondence with the observed above HCV NS3 protein drop (Fig. 3.2.2, B), and related to vital functions of the miR-122 on the HCV genome such as genome stabilization and replication stimulation. Interestingly, the ratio between minus strands generated from SL I-II or SL I-III constructs differs upon targeting either NS5B or Spinach regions. It can only be speculated if this inclines on elevated initiation capacity from the SL I-III template and in more similar full-length antigenome production ability, when compared to the SL I-II template. As a nearly full-length antigenome synthesis ability is a more important measure, in most of the experiments below only the Spinach region is targeted in RT-qPCR assays.

Taken together, the visual and quantitative representation provided evidence that the HCV minus strand synthesis is initiated in the developed minus strand replication system when the essential 5'UTR SL I-II sequence is present. Further extensions of the 5'-end with the HCV 5'UTR sequences were expected to demonstrate either a simulative or no effect on the antigenome synthesis. An inclusion of the SL I-III sequence permits, in addition to functional miR-122 binding, the proposed long-range interactions with the *cis*-elements at the genome 3'-end as well as the assembly of the translation complex. Indeed, the replication assay results indicate significant increase of the relative minus strand abundance, disclosing a positive role of those interactions in minus strand synthesis initiation. The direct and indirect premises are to be further dissected by additional mutational approaches. Remarkably, a further extension of the replication template with the complete HCV 5'UTR sequence followed by a short ORF results in a significant drop of minus strand synthesis comparing to both SL I-II and SL I-III constructs. In fact, the 5'UTR construct resembles a full-length or replicon situation when both genomic ends are present simultaneously and the functional translation from the HCV IRES takes place: the shortage can be compensated by a thoughtful mutational approach disabling selectively one or more functions of the construct. The controversy of the observed results may serve as an illustration of a balance between mutually exclusive HCV translation and replication: once permitted, the functional translation slows down the overall replication rate. A closer insight into this observation is provided in the following section.

3.3 Analysis of the HCV translation impact on the minus strand synthesis

Although the HCV translation initiation is entirely dependent on the cellular machinery, its mechanism is remarkably different from the one for eukaryotes and engages fewer canonical factors. The intact HCV IRES element directs and positions the 40S ribosomal subunit without a need for an initial recognition and scanning. The HCV IRES SL III domain represents a binding platform for the small ribosomal subunit, while the domains II and IV contribute to this interaction and its regulation (Fig. 3.3.1, A, in yellow). These domains are sensitive to mutations affecting both sequence and structure, however the apical loop of the domain IIId was found to be indispensable for the recruitment of the ribosome: a 266–268GGG to CCC mutation reduces 40S binding affinity by 25-fold (Kieft et al. 2001; Ji et al. 2004). Following the 40S association, the eukaryotic initiation factor 3 (eIF3) mediates attachment of the tertiary complex and positioning of tRNA in the P site. Thereby, the correct recruitment and function of the eIF3 is crucial for the functional translation initiation. The eIF3 binding platform on the HCV IRES is comprised by the 4-way helical junction IIIabc (Fig. 3.3.1, B, in red). The translation initiation can be disrupted at the IRES-40S complex stage by the removal of the stem-loop domain IIIb (Otto and Piglisi 2004; Ji et al. 2004). This knowledge was applied in the present study to dissect a stage of the HCV translation initiation impeding the minus strand RNA synthesis.

The HCV translation and replication are indefeasibly linked together and essentially vital for each other. However, since regulated by the similar set of viral and cellular determinants, these processes need to be uncoupled in time and space. Thus, the above observed relative interference of the HCV translation

initiation with the minus strand RNA synthesis was not entirely unexpected. The minus strand replication system appeared to be helpful in shedding the light on a possible mechanism. The original SL I-III and 5'UTR replication templates (Fig. 3.3.2, A and B, E) are capable of the HCV minus strand synthesis initiation at its 3'-end (Fig. 3.2.3) as well as of an assembly of the translation initiation complex at the 5'-end. Moreover, the 5'UTR construct is additionally competent to drive a functional translation of a short open reading frame from the HCV IRES. In order to dissect the functions overlapping in this configuration, disruptive mutations were introduced in both constructs.

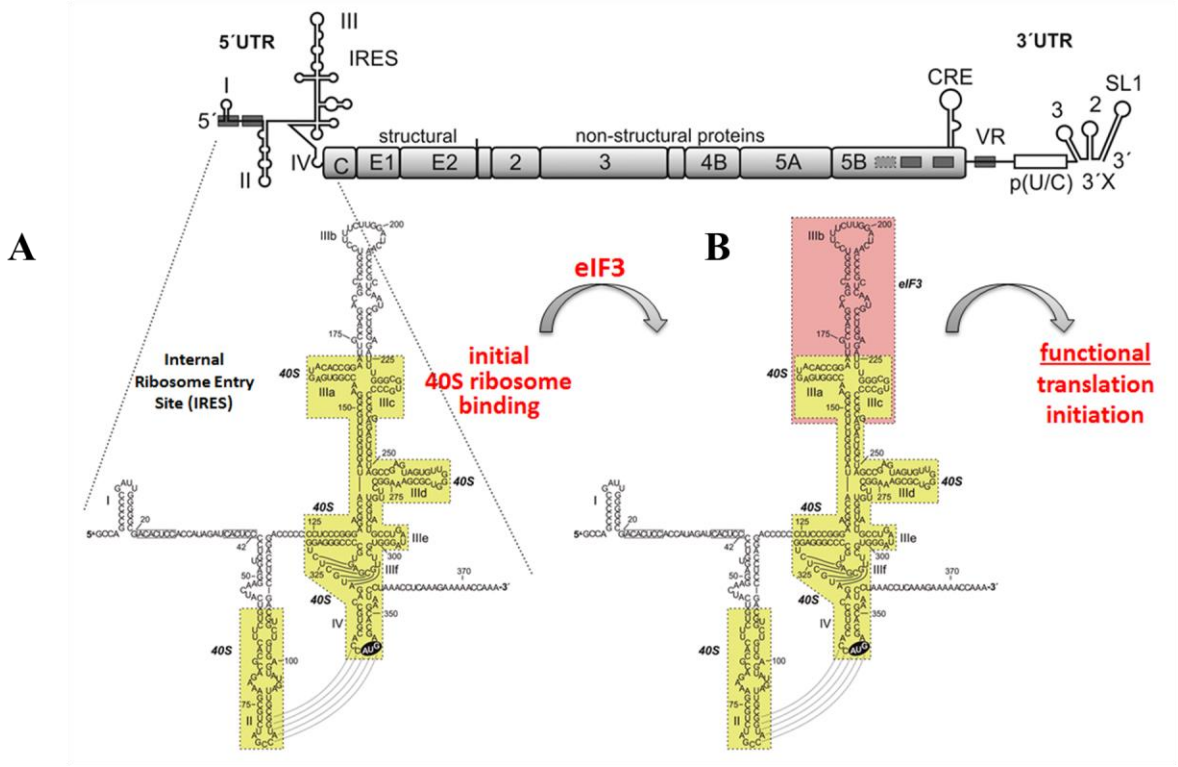


Figure 3.3.1: Main stages of translation initiation from the HCV IRES.
(A) Detailed illustration of the HCV IRES primary and secondary structure. Stem-loops II – IV (highlighted in yellow) are involved in binding of the 40S small ribosomal subunit. Attachment of the 40S takes place without any auxiliary factors and accomplished due to the IRES tertiary structure.
(B) Recruitment of the eukaryotic initiation factor eIF3 occurs to the apical part of the domain III (highlighted in red) and mediates the functional translation initiation.
(modified from Niepmann 2013).

Disabling of translation initiation from the HCV IRES was achieved by a deletion of the domain IIIb that abolishes attachment of the eIF3, without seriously affecting the 40S subunit binding affinity. This modification was applied to both wild-type constructs (ΔIIIb variants; Fig. 3.3.2, C and F), although the translation initiation is not possible in SL I-III construct by default. The binding of the 40S ribosomal subunit alone may affect the replication initiation (through altering the HCV 5'- and 3'- end long-range interactions) either negatively, by interfering with the proposed genomic ends' hybridization (Fricke et al. 2015), or positively by bringing the HCV genome ends together (Bai et al. 2013). Its affinity to both ends leads to their accommodation in a close proximity that may impede the replication complex assembly at the 3'-terminus. Additionally, the 40S competes with the 5BSL3.2 for binding to the IRES domain IIId permitting an interaction between 5BSL3.2 and SL9110. Thus, in both experimental constructs the 40S attachment was restricted by the GGG to CCC mutation in the apical loop of the IIId (IIId* variants; Fig. 3.3.2, D and G).

The replication templates harboring the described modifications in SL I-III or 5'UTR context were generated by *in vitro* transcription, as previously. Verification of experimental RNA integrity was accomplished by agarose gel electrophoresis (Fig. 3.3.3, A). The HCV replication proteins expression in transfected cells was demonstrated by western blot detection of the NS3 protein (Fig. 3.3.3, B). For the latter

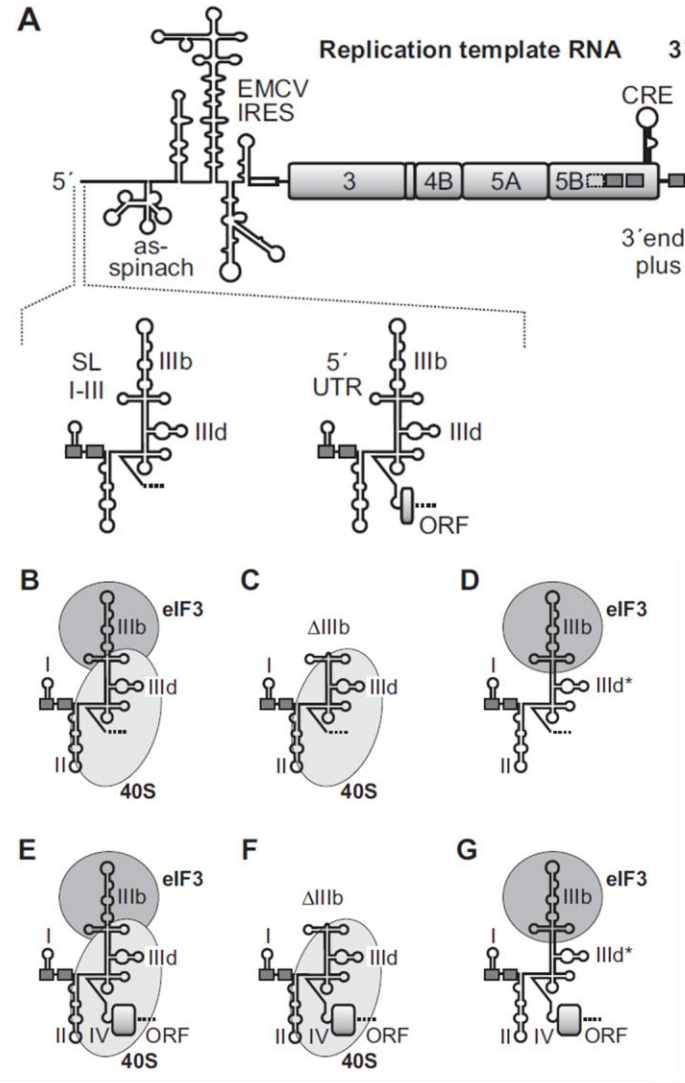


Figure 3.3.2: Schematic representation of the replication constructs for analysis of the balance between HCV translation and replication.
(A-B, E) The replication RNA templates containing either SL I-III (B) or the complete HCV 5'UTR with the initiation codon and a short ORF (E) were mutated to disrupt one of the stages of translation initiation from the HCV IRES.
(C, F) The deletion of the HCV IRES domain IIIb (ΔIIIb) in the context of either the SL I-III (C) or the 5'UTR (F) construct aims to abolish binding of the eIF3 and therefore prevent the functional translation initiation from the 5'UTR construct. However, both ΔIIIb RNA templates remain capable of the 40S ribosomal subunit recruitment. The latter is abrogated via conversion of 266–268GGG nucleotides within the apical loop of the domain SL IIId to CCC sequence (IIId*).
(D, G) The IIId* mutation in the SL I-III (D) or the 5'UTR (G) context disables both the 40S binding and the translation initiation. Although these templates may still bind the eIF3, the affinity in the absence of the 40S is lowered.
The potential of the 40S and/or eIF3 binding is depicted at the 5'-terminal fragments of corresponding constructs.

test, the HCV NS3 protein was targeted in HuH-7.5 cells lysate 2 days post transfection with each of the constructs. The wild-type SL I-III and 5'UTR constructs (Fig. 3.3.3, B; lanes 1 and 5) were additionally co-transfected with anti-miR-122 LNA (lanes 2 and 6). The similar effect, as for the SL I-II transfected sample (Fig. 3.2.2, B), was observed: sequestration of endogenous miR-122 from the transfected hepatoma cells seems to reduce stability or/and replication ability of the input plus strand templates, therefore less target protein can be produced from the EMCV IRES (at later time-points) due to a lack of templates. Providing that deterioration of the translation function from the HCV IRES should only improve the replication rate, the direct positive impact of miR-122 on RNA synthesis was presumably affected. In both the SL I-III and the 5'UTR experimental constructs all annotated miR-122 target sites are present: S1 and S2 sites are located within the SL I sequence at the 5'-end and the 5B.1-3 and S3 sites – at the 3'-end of plus strands (the precise impact of miR-122 binding to either of these sites is the main issue of the next section). Impairment of miR-122 binding to these sites by sequestering LNA reduces the efficiency of minus and plus strand synthesis from the SL I-III and 5'UTR replication templates, therefore limiting genomic RNA amounts remaining by 2 days post transfection that could serve as a template for the HCV protein synthesis. In turn, the modifications - ΔIIIb and IIId* - both did not impair the HCV NS protein expression (lanes 3, 4 and 7, 8), even resulting in more intense signals comparing to their wild-type replication templates. The possible explanation here may again refer to an alteration of the input RNA stability: the formation of the translation initiation complex at the HCV IRES, and in particular a performance of active translation, may be linked to elevated activity of cellular exonucleases at the template's 5'-end.

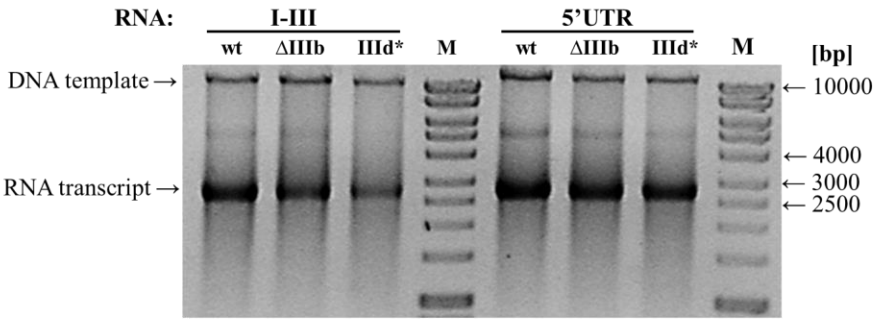
Since the expression of the HCV replication proteins appeared not to be negatively affected for the mutated templates, the further quantitative analysis of the minus strand synthesis efficiency was conducted. The varying stability of input replication RNA templates does not bias the data obtained by RT-qPCR, since in each sample the newly generated minus strand count is always normalized to abundance of the plus strand templates. Hence, although less stable plus strand species result in fewer minus strands in absolute count, the relative efficiency of the particular RNA construct to serve as a temple remains solid. A similar adjustment principle is useful when a construct is capable of continuous replication: although the template plus strands are amplified over the time post transfection, the count of progeny minus strands per one template RNA is unaffected.

In the replication assay the original SL I-III and 5'UTR constructs were compared to their corresponding mutant variants and the effect on minus strand synthesis efficiency was evaluated by RT-qPCR targeting the NS5B and Spinach regions on newly generated minus strands (Fig. 3.3.3, C and D, respectively). The ΔIIIb and IIId* modifications aimed to dissect the overlapping functions exerted by the SL I-III or the complete HCV 5'UTR sequence in scope of translation from the HCV IRES versus antigenome synthesis. The outcome of the introduced mutations was evaluated for each context (SL I-III or 5'UTR) contrasted to a wild-type SL I-III construct (the RER values were set to 1). In the described assay the basic hp construct with a GND mutation (hp_GND) serves as a negative control of the minus strand synthesis: the signals obtained for this construct are indicated by a dotted line on the graphic representation.

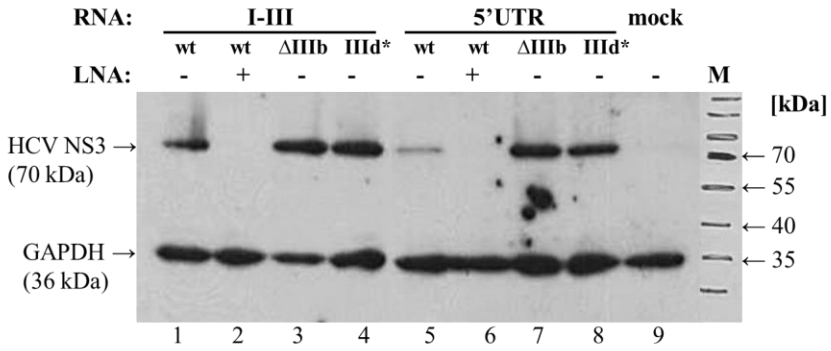
Contrasting results were acquired for the series of the SL I-III and the 5'UTR experimental constructs (Fig. 3.3.3, C and D, left). In the SL I-III context deletion of the SL IIIb that prevents an association of the eIF3 did not have an effect on the minus strand synthesis efficiency. On the other hand, restriction of the 40S binding to the HCV IRES significantly lowered the relative minus strand abundance. Indeed, incapable of a functional translation initiation, the SL I-III construct is not affected by arrest of the eIF3 binding. However the mutation in IIId loop sequence leads to a lack of the 40S subunit association and/or a number of suggested long-range interactions. On the contrary, in the context of the 5'UTR construct, an elevation of a relative minus strand production was observed when translation initiation from the HCV IRES was prevented: both by the IIId* mutation and more profoundly – by ΔIIIb deletion (Fig. 3.3.3, C and D, right). The latter modification disrupting an interaction with the eIF3 demonstrated the more pronounced effect than the inhibition of the 40S binding, which underlines the impact of an actual translation initiation event in negative regulation of the minus strand synthesis initiation. Indeed, according to published literature, the 5'UTR_ΔIIIb completely disables translation initiation, whereas the IIId* mutation within the same RNA template only reduces the efficiency of the 40S binding by 25-fold (Ji et al. 2004). It can be hypothesized that some initiation events in the context of 5'UTR_IIId* construct may still take place that slightly disadvantages the replication. Nevertheless, the difference between the 5'UTR construct mutants was missing a statistical significance.

The complex interplay between distant *cis*-elements at the 3'-end of the HCV genome and the 5'UTR, together with assembly of the translation complex, drives a switch between translation and replication. The long-range contact between the domain IIId and the bulge of 5BSL3.2 *cis*-element is thought to bring the genome ends into a close proximity and therefore facilitate replication (Shetty et al. 2013). The latter structural element thereby normally has to compete with the ribosome, in particular with the 40S subunit, that in turn favors a translation event. Taking these into account, the GGG to CCC mutation in the apical loop of the IIId is to impair both – translation and replication. For this reason, in the SL I-III series, where translation cannot play a role as an inhibitory factor, disruption of the IIId-5BSL3.2(bulge) interaction plainly interferes with a formation of an open 5BSL3.2 element conformation and genome circularization (Romero-López et al. 2012; Romero-López and Berzal-Herranz 2017). Additionally, binding of the 40S ribosome itself was described to promote the ends communication due to its interaction of non-competing nature with both 5'UTR and poly(U/C) track in the 3'UTR via different sites on the 40S (Bai et al. 2013). Either of the two mechanisms or both may result in the observed decrease in the efficiency of the

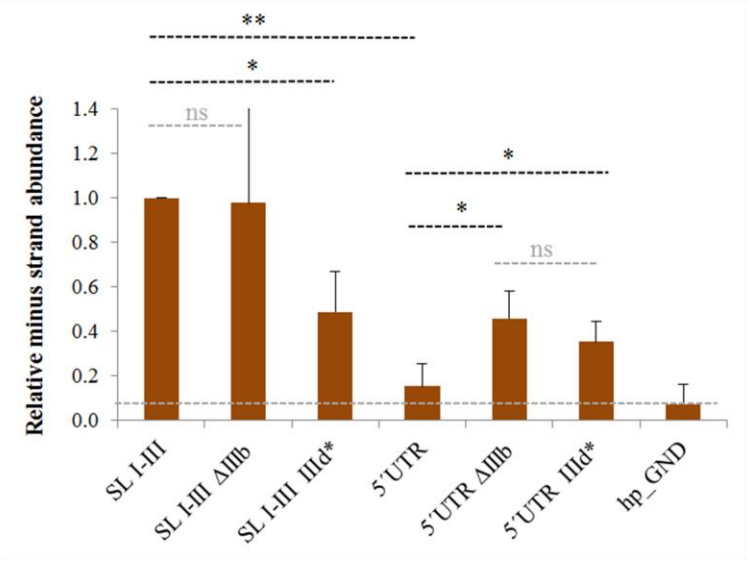
A


RNA:
I-III
5'UTR
wt
ΔIIIb
IIId*
M
wt
ΔIIIb
IIId*
M
[bp]
DNA template →
RNA transcript →
← 10000
← 4000
← 3000
← 2500

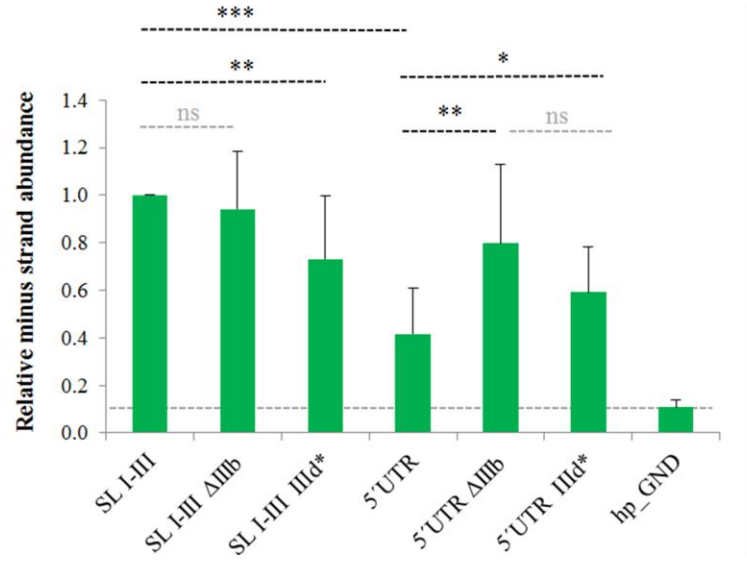
B


RNA:
I-III
5'UTR
mock
wt
wt
ΔIIIb
IIId*
wt
wt
ΔIIIb
IIId*
LNA:
- + - - - + - - -
M
[kDa]
HCV NS3 →
(70 kDa)
GAPDH →
(36 kDa)
← 70
← 55
← 40
← 35
1 2 3 4 5 6 7 8 9

C


**
*
ns
*
*
ns
Relative minus strand abundance
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
SL I-III
SL I-III ΔIIIb
SL I-III IIId*
5'UTR
5'UTR ΔIIIb
5'UTR IIId*
hp_GND

D


***
**
*
ns
**
ns
Relative minus strand abundance
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
SL I-III
SL I-III ΔIIIb
SL I-III IIId*
5'UTR
5'UTR ΔIIIb
5'UTR IIId*
hp_GND

◄Figure 3.3.3: Replication RNA templates for dissection of the effects of HCV translation on minus strand RNA synthesis initiation.
(A) *In vitro* transcribed RNA constructs harboring SL I-III or the complete HCV 5'UTR at the 5'-end (wild-type or ΔIIIb/IIId*) are visualized on 1 % agarose gel using ethidium bromide staining. The series of the SL I-III or the 5'UTR RNAs are represented by an original wild-type construct and its mutants with either a deletion of the domain IIIb (ΔIIIb) or with a mutation of the nucleotides 266-268 (IIId*). Samples are analyzed before the DNase I treatment, thus contain a mixture of RNA transcript and initial linear DNA template. The molecular weight varies depending on the size of the 5'-end insert; the RNA transcripts are compared to the dsDNA marker (M, GeneRuler DNA Ladder Mix).
(B) Western blot detection of the HCV NS3 protein (70 kDa) and cellular GAPDH (36 kDa) as a reference protein in HuH-7.5 cells lysate 2 days post transfection with each of the SL I-III/5'UTR construct series (wild-type or mutated versions, as in A). No HCV NS3 protein was detected in RNA samples co-transfected with an anti-miR-122 LNA (lane 2 and 6) as well as in a mock-transfected cell lysate (lane 9). M, protein marker ladder.
(C-D) Quantification of the newly synthesized minus strands by RT-qPCR targeting the reverse complementary sequence of the NS5B (in brown) or the Spinach (in green) regions. The hairpin construct harboring a replicase-inactivating GND mutation serves as a negative control of minus strand synthesis and indicates background signal (a grey dashed line). The RER values (in C and in D) for the wild-type SL I-III construct are set to 1. In relation to these, the fold differences of the relative synthesis efficiency are presented on diagrams. The mutated variants are analyzed separately in opposition to their wild-type precursors. Asterisks mark p-values from two-tailed student's T-tests (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ns: not significant). On the diagrams, error bars show standard deviation between at least three independent experiments.

SL I-III_IIId* antigenome production. Contrariwise, the SL IIId* and the ΔIIIb mutations led to an elevation of the 5'UTR replication potency. The principal explanation here is a relief of the inhibition mediated by enabled functional translation initiation. Since the 40S subunit has a higher affinity to the HCV IRES than the 5BSL3.2(bulge), the 5BSL3.2 pseudoknot will preferably be in a closed conformation and the RNA template undergoes translation. Therefore, a more prominent elevation of minus strand synthesis is observed for the 5'UTR_ΔIIIb due to a total arrest of translation, and a lesser – for the 5'UTR_IIId* construct that occasionally is capable of translation initiation. Disruption of the long-range SLIIId-5BSL3.2(bulge) interaction and genome circularization also might contribute to the latter effect.

Overall, the conducted experiments provide only a slight glance at the entangled interplay between the HCV translation and genome synthesis. Impairing action of the functional HCV translation on the 5'UTR construct indicates that the HCV translation impedes the replication at the stage of ongoing protein synthesis. Evidently, when the interaction of the eIF3 with the HCV 5'-end is disabled, a higher potency of the minus strand production from the 3'-end is displayed. On the contrary, disruption of the SL IIId may have a dual effect: preventing a translation initiation, it also detunes essential communication between the genome ends. Further in-depth investigation should be attempted in order to resolve this duality. One possibility to distinguish between the effects of either a physical binding of the 40S or of long-range interactions is to introduce modifications in the IIIe tetraloop, retaining the SL III intact. Disruptions within the IIIe prohibit the 40S subunit binding as well as reduce overall translational activity (Psaridi et al. 1999). A combination of double SL IIId* and the ΔIIIb mutations may appear useful to reveal any unexpected replication inhibitory factors or events when comparing such mutant in the 5'UTR context with the wild-type SL I-III construct. Eventually, the other less investigated *cis*-elements, i.e. the SL 588 within the Core-coding region (Pirakitikulr et al. 2016), should be incorporated into experimental constructs to complete the state-of-the-art 5'-end *cis*-signals engaged in long-range interactions with the 3'-end elements with a proposed impact on the HCV replication.

3.4 Effects of miR-122 acting at the HCV 5'- and 3'- UTRs on the HCV minus strand replication

The vital role of liver-specific miR-122 on HCV replication (Jopling et al. 2005), translation (Henke et al. 2008) and RNA stability (Shimakami et al. 2012a) has been proposed for a long period of time. The mechanism of miR-122 action on the HCV RNA is well-described for the 5'UTR target sites S1 and S2, however its function at the 3'-end candidate binding sites (5B.1-3 and S3) remains unclear. Moreover, a transition from serving one function to another upon miR-122 binding is largely elusive. The main limitation originates from the multiple functions of those 5'UTR sites in the HCV life cycle that are challenging to dissect. Therefore, in the current research it was attempted to unravel the miR-122 impact on the HCV minus

strand synthesis efficiency upon binding to the target sites at either of the HCV genomic ends and to determine the requirements for the antigenome production using the minus strand replication system.

In order to uncouple the miR-122 functions in the HCV translation, the SL I-II construct containing all – the 5'- and 3'- ends' binding sites – was utilized for the analysis (Fig. 3.4.1, A). Role of the miR-122 binding sites at either HCV genomic end in minus strand synthesis were investigated separately bymutational approach disabling binding of the miR-122, but retaining an original secondary structure of the HCV RNA. The mutations within the 5'UTR S1 and S2 miR-122 binding sites were originally designed and generated by Anika Nieder-Röhrmann, as described in her doctoral dissertation (also in Nieder-Röhrmann et al. 2017) (Fig. 3.4.1, B and C). The details on the mutated sequence are specified in the Appendix together with plasmid maps corresponding to the wild-type and its S1/S2-mutated version (designated as SL I-II_S1mS2m; see 6.1.13-14). In the scope of the present work only a double – S1mS2m – mutant (Fig. 3.4.1, C) in the context of the SL I-II construct was generated; the in-depth analysis of each of the two HCV 5'UTR miR-122 binding sites is not included. Apart from a complexity reason, these sites function cooperatively and play a more profound role when addressed simultaneously (Nieder-Röhrmann et al. 2017), which justifies application of a double mutant in the proof-of-principle experiment.

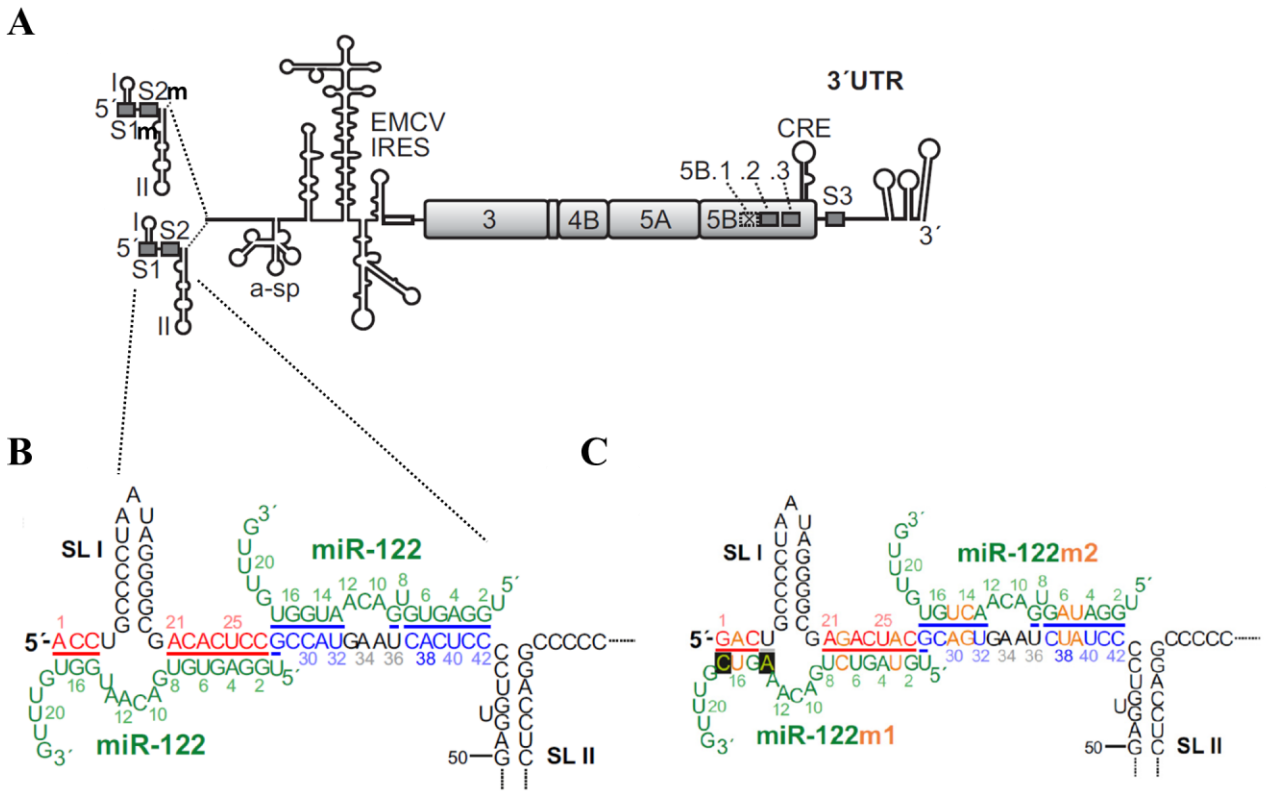


Figure 3.4.1: The system for analysis of miR-122 role in HCV minus strand synthesis upon binding to the 5'- and 3'- end target sites at the HCV genome.
(A) The replication RNA template harboring the HCV SL I-II at the 5'-end is a basic construct for the analysis. The S1 and S2 sites at the 5'-end and the 5B.2, 5B.3 and S3 sites at the 3'-end of the HCV genome are indicated by grey boxes. The 5B.1 site was inactivated in all 3'-end mutant variants and depicted as a dashed box. For the 5'UTR miR-122 binding sites the double S1/S2 mutant is generated for the analysis (S1mS2m); whereas the variants disabled of miR-122 binding at the 3'-end are utilized as single mutants (5B.2m, 5B.3m or S3m), also in the SL I-II_S1mS2m context.
(B) Representation of the current knowledge on the miR-122 binding to the S1 (red) and S2 (blue) target sites in the HCV 5'UTR. The nucleotides within the miR-122 molecule are numbered in green from 5'- to 3'- end. Some nucleotides of the HCV SL I are numbered in red and blue; the sequence belongs to the J6 isolate (genotype 2a).
(C) Representation of the mutated S1m (red) and S2m (blue) miR-122 binding sites in the HCV 5'UTR. The binding of correspondingly modified miR-122 mutants (miR-122_S1m and miR-122_S2m, in green) serves to compensate and restore functional assembly of the miR-122/Ago2 complex, as in wild-type situation. The mutations within the S1/S2 target sites and the compensatory miRNAs are labeled in orange (designed and verified in the doctoral dissertation of Anika Nieder-Röhrmann).

With regard to the HCV 3'-end target sites – 5B.1-3 and S3 – mutations (Fig. 3.4.1, A) sufficient to disable the miR-122 binding and function were introduced into the same wild-type replication construct, and were presented in the doctoral work of Nadia Dünnes (also in Gerresheim et al. 2017). The detailed representation of mutated 3'-end sites' sequences in the context of the SL I-II construct (designated as SL I-II_5B.2m, SL I-II_5B.3m and SL I-II_S3m) is provided in the Appendix (see 6.1.25-27). In the present work the potential miR-122 site 5B.1 located within the HCV NS5B coding region was inactivated in all generated constructs since this site is not conserved among HCV isolates (only in 26 out of 106 selected HCV isolates; Fricke et al. 2015). In contrast, the 5B.2, 5B.3 and the S3 miR-122 target sites located downstream the 5B.1 in the HCV NS5B coding region and the 3'UTR were found to be highly conserved (99/106, 106/106 and 106/106 among selected HCV isolates respectively; Fricke et al. 2015), therefore included in the mutational analysis. The mutations were introduced within seed site regions separately for each of the binding sites resulting in single mutants. Similarly, the analogous mutations were generated within the SL I-II_S1mS2m construct (designated as SL I-II_S1mS2m_5B.2m, SL I-II_S1mS2m_5B.3m and SL I-II_S1mS2m_S3m; see 6.1.28-30).

Generated by *in vitro* transcription, replication template RNAs corresponding to the above described SL I-II construct and its 5'- and/or 3'- UTR miR-122 binding mutants (Fig. 3.4.2, A) were tested upon transfection into HuH-7.5 cells. Here, the cells were lysed 2 days post transfection and the HCV NS3 protein detection was conducted by western blot (Fig. 3.4.2, B). Visualization of the viral protein at this time-point only for few of the transfected RNA templates indicates that the introduced mutations crucially affect

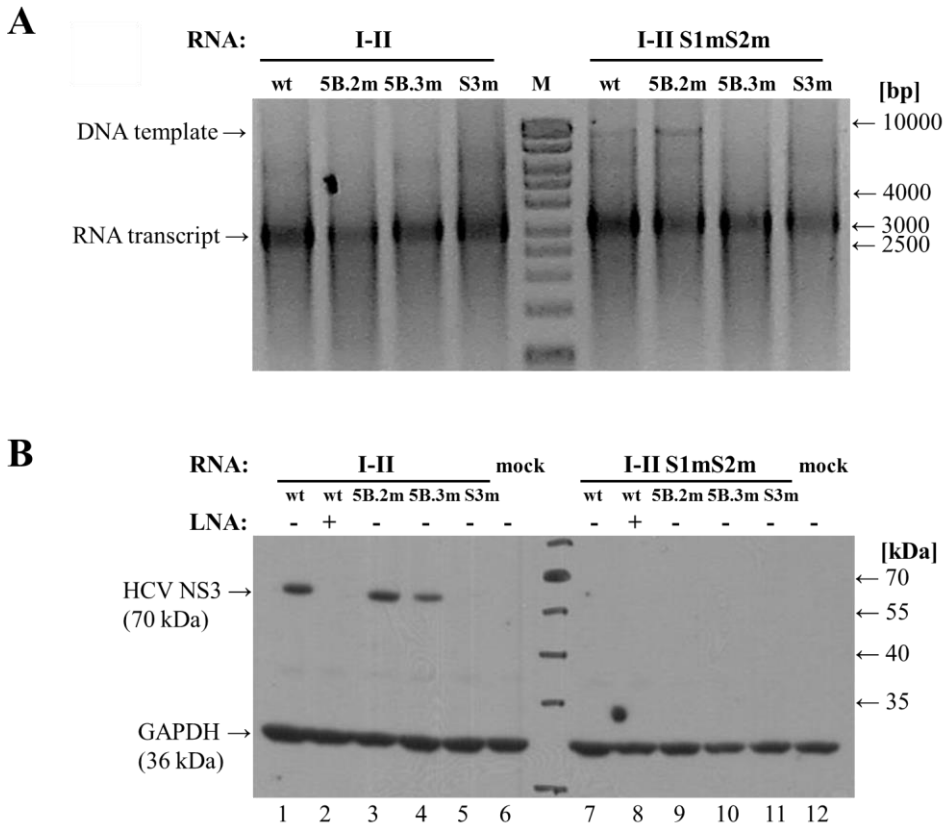


Figure 3.4.2: Replication RNA templates for investigation of miR-122 role in the HCV minus strand synthesis initiation.
(A) Visualization of *in vitro* transcribed RNAs corresponding to the SL I-II construct (wt) and its mutated variants on 1 % agarose gel using ethidium bromide staining. The double S1/S2 miR-122 binding mutant is designated as S1mS2m; single mutants of the HCV 3'-end 5B.2, 5B.3 and S3 sites are annotated as 5B.2m, 5B.3m and S3m, respectively. Loaded prior the DNase I treatment, the samples contain both input DNA template and RNA transcript. Based on the same construct all RNA transcripts are 7180 nt; the transcripts are compared to the dsDNA ladder (M, GeneRuler DNA Ladder Mix).
(B) Western blot detection of the HCV NS3 protein (70 kDa) and cellular GAPDH (36 kDa) as a reference protein in HuH-7.5 cells lysate 2 days post transfection with the SL I-II (lanes 3-5) and SL I-II_ S1mS2m (lanes 9-11) miR-122 binding mutants (labeling as in A). The HCV NS3 protein is not detected when RNA templates are co-transfected with anti-miR-122 LNA (lanes 2 and 8) as well as for most of the mutants and in mock-transfected cells (lanes 6 and 12). M, protein marker ladder.

replication driven from these constructs. At the later time-points post transfection, i.e. 2 days, an availability of templates for translation rather than templates' stability appears critical for the HCV protein accumulation. Indeed, only the constructs that are capable of continuous replication may still retain in a cell 2 days post transfection and serve as a template for protein synthesis. Clearly, apart from the wild-type SL I-II template RNA, only its 5B.2 and to a lesser extent 5B.3 mutants remain viable in terms of ongoing rounds of replication (lanes 1, 3 and 4).

As illustrated in the above sections, co-transfection of sequestering anti-miR-122 LNA reduces the HCV NS proteins expression already at the 8 h post transfection, most likely due to a negative impact on RNA stability. On the long run, such LNA rather affects the efficiency of replication from transfected RNA construct since miR-122 binding is abolished to all 5'- and 3'- UTR binding sites (Fig. 3.2.3, C and D). Therefore, any replication template co-transfected with the anti-miR-122 LNA resembles its mutant with all five sites disabled from miR-122 binding perspective, but remains wild-type regarding the primary HCV sequence. None of the SL I-II_S1mS2m derivatives (lanes 7; 9-11) results in continuous replication, suggesting that the S1 and S2 sites are important for the HCV replication at the first place. The 5B.2 and 5B.3 single mutants of the SL I-II construct maintain a capacity for ongoing replication, since a template for translation of the HCV NS proteins is still available 2 days post transfection, thus the 5B.2 and 5B.3 miR-122 target sites are not essential for the HCV replication, whereas the S3 site appears to be evidently required. Whether this effect is mediated by the miR-122 binding or/and is a consequence of the primary sequence alteration requires clarification by a miR-122 complementation approach.

To elucidate the role of miR-122 binding to the above-mentioned sites along the HCV genome, the replication assay was carried out. The wild-type SL I-II construct and its SL I-II_S1mS2m derivative were transfected in series in parallel with the corresponding miR-122 binding mutants at the 3'-end (5B.2m, 5B.3m and S3m). The efficiency of the HCV minus strand synthesis was evaluated by RT-qPCR (targeting only Spinach region) using the total RNA isolated from the cells 2 days post transfection. Transfection with the SL I-II construct harboring the GND mutation in the replicase coding-region served as a negative control of minus strand synthesis initiation; the signals derived for SL I-II_GND are indicated with a dotted line at the presented schematics. The mean RER value for the wild-type SL I-II construct was set to 1 to enable visual comparison to all generated – single and multiple – miR-122 binding mutants (Fig. 3.4.3, A). Results of this replication assay to a high extent resemble the HCV NS3 protein detection by western blot (Fig. 3.4.2, B). Relative minus strand synthesis efficiency of the intact SL I-II construct is significantly higher than of all the mutants, except for the SL I-II_5B.2m. Already in wild-type S1/S2 context, disabling of miR-122 binding to the 5B.3 or S3 sites severely impairs replication. A similar negative effect is displayed when the 5'UTR miR-122 target sites are disabled (SL I-II_S1mS2m). Further mutations of the 3'-end target sites – 5B.2, 5B.3 and S3 – do not significantly impair antigenome synthesis further on. This observation emphasizes the primary importance of the miR-122 binding to the S1 and S2 sites for efficient HCV replication. However, a fundamental reason for this importance may be either direct physical consequences of the binding, i.e. in Ago2 complex formation and stabilization of RNA templates (Shimakami et al. 2012a; Li et al. 2013), or/and in indirect effects on *cis*-elements' communication due to primary structure alterations.

This uncertainty can be addressed by a compensation approach. In the current experimental situation, either application of miR-122 variants harboring compensatory substitutions to restore its binding to mutated sites or introduction of corresponding mutations within *cis*-elements that are proposed to engage in RNA-RNA interactions. The latter – genetic approach – is to a large degree more challenging, not only due to substantial mutagenesis required, but most importantly due to involvement of the other known (and unknown) functions of *cis*-elements along the HCV genome. Particularly, the SL I primary sequence, containing both 5'UTR miR-122 binding sites, is highly conserved among HCV isolates and was predicted to engage in long-range interactions with the 3'-end region (also remarkably conserved) potentially resulting

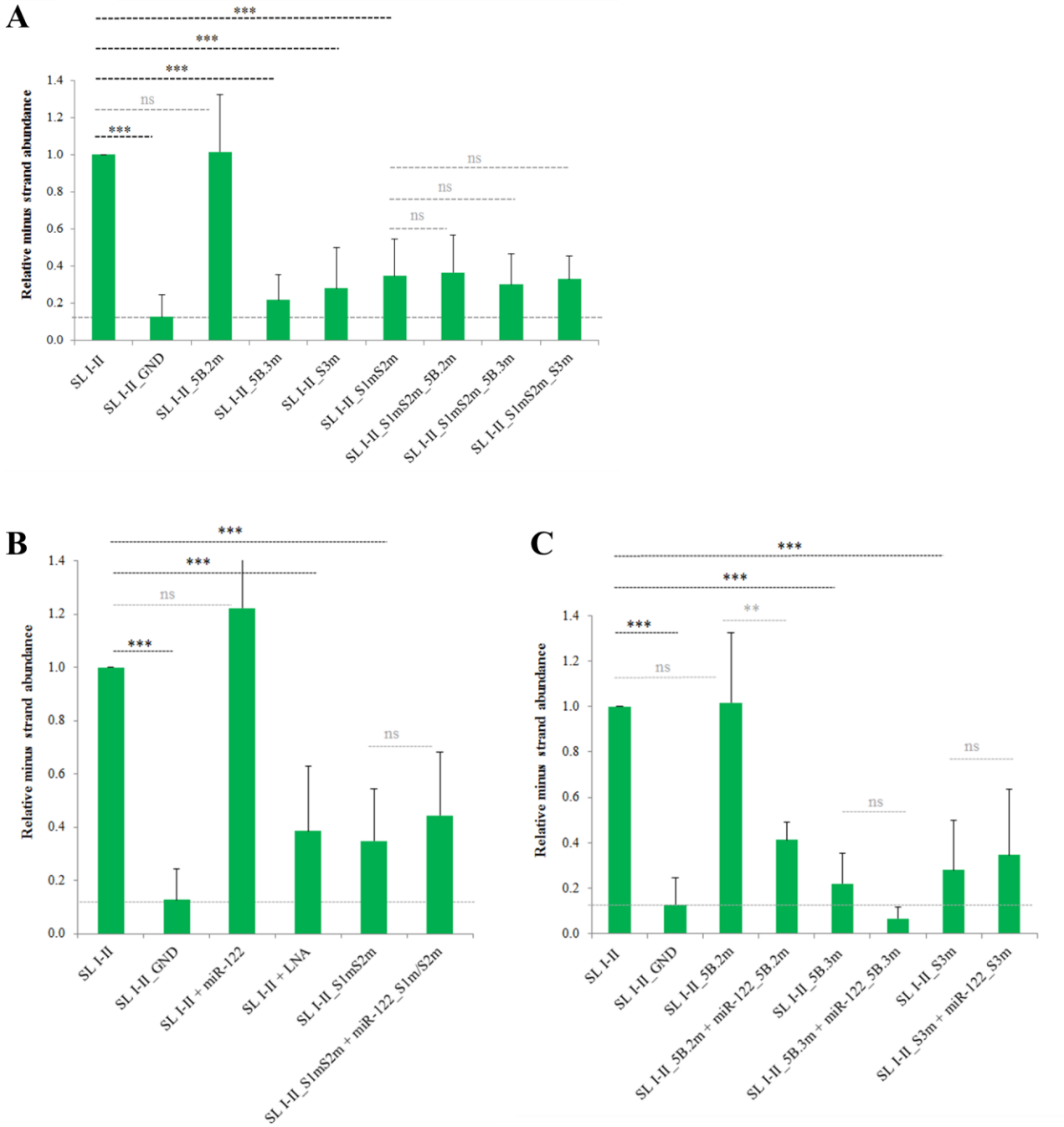


Figure 3.4.3: Effects of miR-122 binding at the HCV 5'- and 3'- end on the HCV minus strand synthesis initiation.
(A) Role of the miR-122 binding sites in the HCV 5'- and 3'- UTRs. Quantification of newly synthesized HCV minus strands by RT-qPCR 2 days post transfection with SL I-II replication template containing either wild-type or mutated S1 and S2 target sites as well as with derivatives of both constructs containing mutations of the 3'-end miR-122 binding sites (annotated as in Fig. 3.4.2).
(B) Replication assay challenging a function of the 5'UTR miR-122 binding sites in more detail. Similar transfection of the wild-type SL I-II replication RNA template and its S1mS2m derivative alone or together with corresponding miR-122 variants (wt or miR-122_S1m/S2m) in attempt to compensate miR-122 action at the 5'UTR target sites. Co-transfection of RNA templates with anti-miR-122 LNA mimics a substantial miR-122 binding abrogation.
(C) Replication assay challenging a function of the HCV 3'-terminal miR-122 binding sites when the 5'UTR sites S1 and S2 are retained intact. The replication template SL I-II and its single-site mutant derivatives are compared alone or in presence of correspondingly mutated miR-122 variants (miR-122_5B.2m/5B.3m/S3m).
In all assays all primers are universal for the examined constructs; the nascent minus strands were quantified targeting the Spinach region (identification of long elongation products) and normalized to the input plus strand abundance. The background signal derived for the replication-deficient SL I-II_GND construct is labeled as a grey dashed line. The RER values for the SL I-II construct are set to 1 and the fold difference serves as a measure for comparison. Asterisks mark p-values from two-tailed student's T-tests (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ns: not significant). On the diagrams, error bars show standard deviation between at least three independent experiments.

in the HCV genome circularization (Fricke et al. 2015). Any attempts of genetic alterations at the conserved and indispensable 3'-end are expected to result in, by default, unviable genomes.

Thence, compensation by correspondingly mutated synthetic miR-122 variants was carried out. At first, to challenge the importance of miR-122 binding to the 5'UTR, wild-type and S1mS2m SL I-II constructs were transfected either alone or with a respective miR-122 (wild-type or a composition of mutated) or withanti-miR-122 LNA (Fig. 3.4.3, B). The compensatory mutations in the miR-122 "guide" strand were designed in accordance to introduced S1m and S2m mutations within the S1 and S2 sites (see 6.1.13-14) to restore the binding. These miR-122 variants were demonstrated to bind to their corresponding mutated target sites and to rescue the miR-122 binding at the HCV 5'UTR (dissertation of Anika Nieder-Röhrmann). An equimolar mixture of the modified miR-122_S1m and miR-122_S2m duplexes was used in the current work for the co-transfection with the SL I-II_S1mS2m replication construct. In the replication assay additional wild-type miR-122 did not significantly increase the relative minus strand abundance displayed by the SL I-II construct, most likely due to a saturated effect of miR-122 on the HCV genome, however sequestration of endogenous miR-122 from HuH-7.5 cells by LNA resulted in a dramatic drop of synthesis efficiency (Fig. 3.4.3, B). As shown above, the minus strand production was drastically affected when the S1mS2m version of the SL I-II construct was used. Remarkably, compensatory mutated miR-122 S1m and S2m variants applied together with the SL I-II_S1mS2m replication construct could not improve antigenome production. Resulting in a similarly low minus strand synthesis efficiency upon application of compensatory miR-122 variants, this experiment proves the importance of the primary sequence constituting the S1 and S2 target sites in the SL I region. Actual binding of miR-122 to those sites, if contributes at all, seems to play a secondary role in regulation of the HCV minus strand synthesis initiation. Indeed, the engagement of the SL I into the long-range interaction with the genomic 3'-end (initiated between the apical regions of the SL II and the DLS and extended up to 62 nt) and a subsequent genome ends hybridization was proposed to result in the HCV genome circularization (Fricke et al. 2015), which is known to facilitate replication in other members of the *Flaviviridae* family, i.e. in Dengue virus (Alvarez et al. 2005).

The role of miR-122 binding to the HCV 3'-end sites 5B.2, 5B.3 and S3 was analyzed challenging the single site mutants in replication assay upon application of compensatory miR-122 variants. These miR-122 variants were designed in accordance to introduced mutations within the 5B.2, 5B.3 and S3 sites (see 6.1.28-30) during the doctoral work of Nadia Dünnes and their binding to the corresponding mutated target sites was validated (dissertation of Nadia Dünnes). Since all constructs from the SL I-II_S1mS2m series displayed very low functionality in terms of minus strand synthesis (Fig. 3.4.3, A), most likely resulting already from mutations in the 5'UTR, these constructs were not subjected to a compensation assay. For the series of the SL I-II construct with original S1 and S2 sites, minus strand abundance was quantified in total RNA fraction from HuH-7.5 cells 2 days post transfection via RT-qPCR targeting the Spinach region (Fig. 3.4.3, C). As was shown above, 5B.3m and S3m mutants of the original SL I-II replication template demonstrate significantly lower relative abundance of nascent minus strands. Co-transfection of individually mutated constructs with their corresponding compensatory miR-122 variants failed to rescue the antigenome synthesis rate in all mutants. Resulting effects of application of a compensatory miR-122 were comprised by either insignificant change in the replication efficiency (for 5B.3m and S3m mutants) or by a surprising (and significant) impairment – as for the 5B.2m mutant. The latter anomaly could hypothetically originate from less effective transfection of the input RNA templates or more likely by dilution of the cellular pool of the wild-type miR-122 molecules leading to its decreased activity at the functional S1 and S2 sites. To sum up, in a manner similar to compensation of the miR-122 binding at the 5'-end, physical binding of miR-122 variants and Ago2 complex assembly at the 3'-end target sites were found insufficient for effective rescue of minus strand synthesis initiation. Further research is required to unravel genuine roles of these – yet under investigated – miR-122 target sites. Their importance in HCV replication was confirmed using the full-length HCV culture system and contrasting degree of contribution of each individual binding site was demonstrated (Gerresheim et al. 2017). Supposedly, the miR-122 binding and Ago2 recruitment to the 3'-end sites of the

HCV genome (especially the more downstream) may contribute to conformational changes at the genome terminus in order to facilitate replication initiation. However, the obtained results in turn imply the significance of the primary RNA sequence, which can be utilized as a scaffold for surrounding structural elements or to play a direct role.

In conclusion, miR-122 is one of the most influential host factors in HCV life cycle. Its contribution at various stages directly and indirectly complicates the dissection of its functions and mechanisms of action. Its activity at the HCV 3'-end remains poorly understood, although the high degree of conserved target sites in this region is striking. The developed minus strand replication system is applied in the present work to gain knowledge on the miR-122 role in the HCV minus strand synthesis initiation upon binding to the 5'- and 3'- end target sequences. Importantly, it provided evidence on the positive role of miR-122 binding at either of the HCV ends on antigenome production. A more comprehensive complementation analysis is required to clarify the mechanism of the promoting effect at both ends, since compensatory miR-122 binding was found to be insufficient for efficient replication in contrast to the primary sequence of this vital region. In analogy to the observed importance for the S1 and S2 miR-122 binding sites and their possible role in genome circularization, high sequence conservation of certain regions along the HCV genome (and in particular of the miR-122 binding sites) point out their functional role in the viral life cycle. Further insight on potential of the minus strand replication system in unraveling of functions of other HCV *cis*-elements is presented in the following section.

3.5 A glance over the role of *cis*-elements located in the HCV coding region in minus strand synthesis initiation

An importance of the *cis*-elements contained at the HCV 5'- and 3'- UTRs in plus and minus strand synthesis had been recognized for long time (You et al. 2004, Friebe et al. 2005). In addition, the *cis*-acting replication element 5BSL3.2 was identified in the NS5B coding region. The complex tertiary structure of the 5BSL3.2 element was found to engage in several alternating long-range RNA-RNA interactions and proposed to serve as a molecular switch between translation and replication (Diviney et al. 2008; Tuplin et al. 2012; Shetty et al. 2013). Later on, more signals that are essential for the HCV replication were identified within the NS5B coding region (Chu et al. 2013; Mauger et al. 2015), however they still require further investigation. Moreover, even fewer information is available concerning other potential *cis*-elements located along the rest of the HCV polyprotein coding sequence. Identification and validation of *cis*-elements is a meticulous process, which is often complicated by overlapping functions of an element in the viral life cycle.

In the present work the developed minus strand replication system was adapted for an assessment of possible roles of other *cis*-acting replication elements within the HCV coding region in minus strand synthesis initiation. In the designed replication templates the most downstream 152 nt of the NS5B coding region were not mutated to retain the 5BSL3.2 element unaffected, as long as it is known to be essential for the HCV replication. When required, further mutagenesis can be conducted within this unmodified sequence. The rest of the HCV NS3-NS5B coding sequence was scrambled *in silico* (conducted by Yutong Song and Steffen Mueller) and produced by commercial chemical synthesis. The scrambling procedure aims to disable all possible RNA *cis*-signals that may act at the input RNA template, but retaining the amino acid sequence to obtain operative replication machinery, as well as codon usage and codon pair bias to prevent local differences in translation speed, which might have impact on protein folding and by that on activity and functionality of the HCV proteins. As a basis for analysis, the hp (as a part of cloning and assembly procedure) and SL I-II constructs were modified accordingly resulting in pairs of wild-type and scrambled (SCR) replication templates (Fig. 3.5.1, A), designated as SL I-II (when wild-type) and SL I-II_SCR (for plasmid maps see 6.1.13 and 6.1.21). Additionally, the nearly fully scrambled NS3-NS5B cassette was incorporated into the SL I-II construct containing mutated miR-122 binding sites at the 5'UTR (SL I-II_S1mS2m_SCR; see 6.1.22).

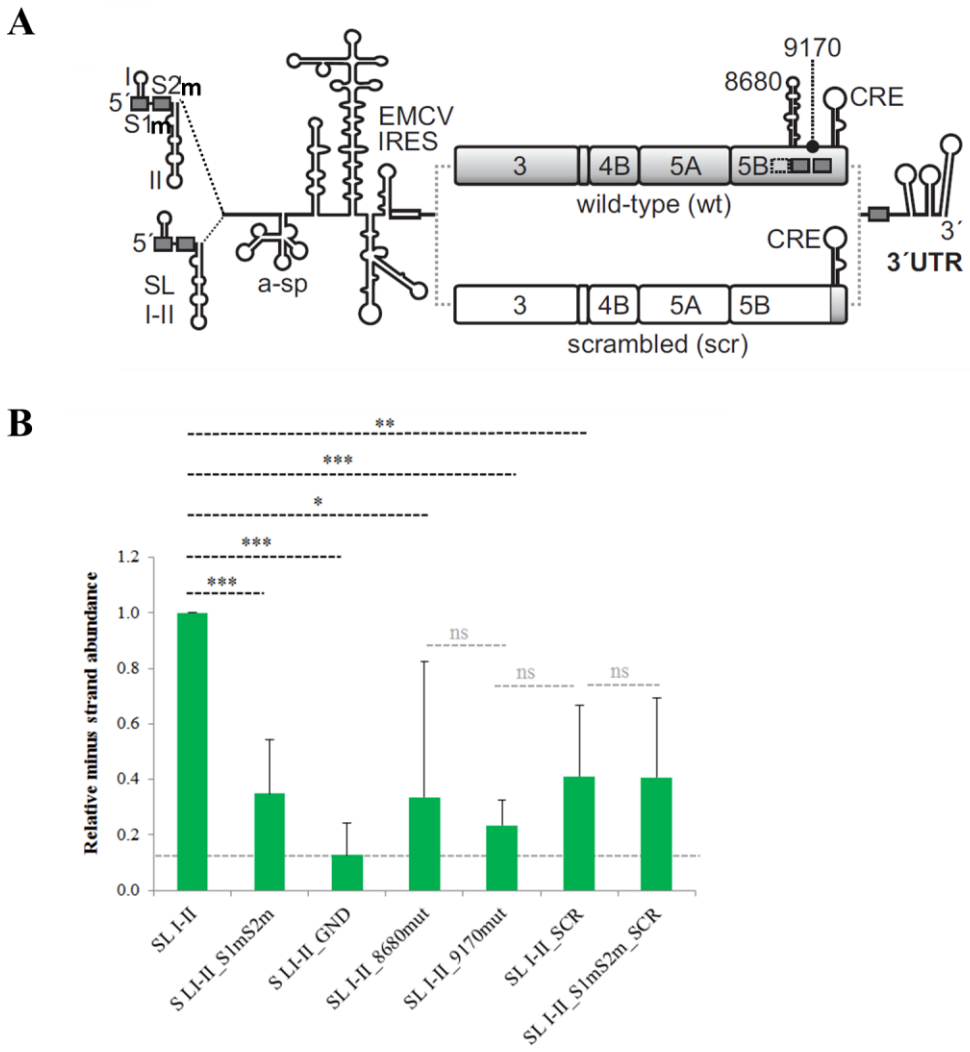


Figure 3.5.1: Role of the HCV NS3-NS5B coding sequence in the minus strand synthesis initiation.
(A) The experimental replication SL I-II construct contains either wild-type or scrambled (the ORF is depicted in white) HCV NS3-NS5B coding sequence; the 5BSL3.2 element (CRE) within the most downstream 152 nt is always retained wild-type (in grey) to enable replication. An analogous scrambled ORF is also incorporated into the miR-122 binding mutant SL I-II_S1mS2m. In construct SL I-II_8680mut only the apical loop at position 8680 within the NS5B region was mutated. In construct SL I-II_9170mut only a sequence at position 9170 is mutated, disabling its hybridization with the bulge of the 5BSL3.2 element.
(B) Quantification of newly synthesized HCV minus strands was performed by RT-qPCR targeting the Spinach region in HuH-7.5 cells lysate 2 days post transfection with derivatives of the SL I-II replication construct, listed in (A). The SL I-II_GND construct stands for a negative control of antigenome production and indicates background signal (a grey dashed line). The RER values for the SL I-II construct are set to 1 and the fold difference serves as a measure for comparison. All primers are universal for the examined constructs; the nascent minus strand abundance is normalized to the input plus strand abundance. Asterisks mark p-values from two-tailed student's T-tests (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ns: not significant). On the diagrams, error bars show standard deviation between at least three independent experiments.

This series of SL I-II constructs - in wild-type or scrambled context - was challenged in the replication assay, where, as before, the SL I-II RNA template harboring inactivating replication GND mutation was utilized as a negative control of the HCV minus strand synthesis initiation (indicated as a dotted line on the graphic representation). The HuH-7.5 cells, as in the experiments above, were transfected with each construct and total cellular RNA was isolated 2 days post transfection. Results provided by the RT-qPCR quantification of relative minus strand abundance targeting the Spinach region (Fig. 3.5.1, B) demonstrated a significant abrogation of the nascent HCV minus strand RNA production after non-directed scrambling. Apparently, potentially occurring *cis*-elements within the SL I-II_SCR construct (excluding the vital 5BSL3.2 element) were inactivated. Thus, an importance of *cis*-elements located upstream the 5BSL3.2 element in the HCV minus strand synthesis initiation can be recognized from this experiment; however, in-depth identification of the essential sequences and structures is necessary.

In the fashion of the 5BSL3.2, it can be speculated that the remaining *cis*-signals are also engaged in intramolecular RNA communication with the sequences at the HCV 5'- and 3'- ends to regulate minus strand synthesis. Since scrambling of the entire NS3-NS5B sequence (excluding the very last 152 nt in the NS5B coding region) results in drastic reduction of minus strand synthesis efficiency, next step requires addressing of *cis*-elements mutations separately. A comprechensive approach would implement division of the entire scrambled sequience into several parts and analysis of each part mutated independently of the others. This is inevitable in a course of very preliminary screeening of essential *cis*-elements. However, a reasonable number of publications on the HCV *cis*-acting replication elements suggested candidate target sequences. Two *cis*-elements within the HCV NS5B coding regions, which had been shown to be involved in overall HCV replication in the published literature, were selected for mutational analysis. The first element, known as J8640 (Mauger et al. 2015) or SL8647 (Chu et al. 2013), is located in the NS5B-coding region at position 8680 of the JFH-1 HCV isolate. This *cis*-element represents an apical loop within a stem-loop that is involved in genome replication and virus particle production (Chu et al. 2013). Another characterized *cis*-element, known as SL9110 or Alt-sequence (Diviney et al. 2008), is located within the NS5B sequence at position 9170 in the JFH-1 genome. The apical loop of the SL9110 resides upstream of the 5BSL3.2 and engages in the long-range interaction with the bulge of the latter. This SL9110-5BSL3.2(bulge) contact was reported to be critically involved in replication and virus production (Diviney et al. 2008).

Therefore, these mutations were introduced only into the SL I-II construct (designated as SL I-II_8680mut and SL I-II_9170mut; see 6.1.33-34) in order to be compared with the wild-type SL I-II construct and its entirely scrambled version. Challenged in a replication assay, the RNA templates containing individual mutations of the 8680 and 9170 *cis*-elements displayed a drastic drop in relative minus strand abundance, similar to the fully-scrambled SL I-II_SCR construct (Fig. 3.5.1, B). No significant difference in replication ability was observed between all three *cis*-element mutants - SL I-II_8680mut/9170mut or SL I-II_SCR. Additionally tested 5'UTR miR-122 binding mutant (SL I-II_S1mS2m; in wild-type or scrambled contex) also presented similarly low replication capabilities: likely the major inactivating effect in this case originated from mutation of the essential 5'UTR sequence.

To sum up, the 5BSL3.2 element appears to be not the only crucially important *cis*-element within the HCV polyprotein coding region for the efficient minus strand synthesis initiation. Comprehensive non-specific scrambling of the coding region upstream the 5BSL3.2 element as well as individual directed distraction of known vital elements 8680 and 9170 (in the JFH-1 HCV isolate) resulted in severely reduced replication activity of the characterized SL I-II constructs. Indeed, disabling hybridization of the 9170 element's loop to a bulge of the 5BSL3.2 unbalances the dynamic conformational rearragement in the 5BSL3.2 element and formation of essential long-range interractions; however, the exact mechanistic details still remain to be disclosed. In turn, modifications introducted within an apical loop of the 8680 element are thought to disrupt its communication with the SL2/DLS in the X-tail (Fricke and Marz 2016). While the latter is not known to have a direct effect on replication efficiency, retaning the complex interplay between RNA *cis*-elements may be the key. Several previously characterized structural RNA elements remain candidates for further analysis. Among them, for example, are so-called SL6038 element located in the NS4B region (Pirakitikulr et al. 2016) as well as the J7880 and J8880 elements distributed along the HCV NS5B coding region (Mauger et al. 2015). All these elemnents are described to a certain extent to play a role in replication. A prominent *cis*-element designated as SL 588 is located downstream the HCV IRES and was also suggested to act as a *cis*-acting replication element (Pirakitikulr et al. 2016). Further research on these *cis*-elements may definitely take advantage of using the minus strand replication system that uncouples the major HCV life cycle events from the HCV RNA synthesis.

3.6 Dynamics of experimental HCV RNA constructs: synthesis and decay

The above chapters have illustrated broad applications of the developed minus strand replication system in dissecting determinants and essential components for the HCV replication. Plasticity of the system, in terms of enabling effortless modifications of functional elements, suggests a broad range of future applications, i.e. further unraveling of the miR-122 function on the HCV genome, a trial of non-structural proteins' mutants as well as partial provision of replication proteins in *trans*, screening for *cis*-elements within the structural proteins (especially, in Core) coding regions, and many others. However, so far the system was utilized within a time period post transfection of RNA templates (around 48 h) that implies the active HCV plus and minus strand replication. It was shown in several studies on highly infectious virus variants, such as the JFH1, that minus strand production peaks at 24-48 h post transfection/infection and followed by plus strands ascent (Targett-Adams et al. 2008, Keum et al. 2012). Indeed, in the context of the developed system, only highly variable and low singnals were derived in attempt to detect minus strand production 24 h post transfection, likely due to various degrees of sequence and/or functional elements deficiency within the experimental HCV RNAs that to a certain extent postpones a replication uprise. Moreover, given that only a few of the developed replication RNAs contain sufficient sequence elements to undergo continuous replication (meaning that they use newly synthesized minus strands as templates for nascent plus strand production), the dynamics of HCV RNA supplied by the minus strand replication system should differ from the one of a wild-type viral RNA. The current chapter attempts to elucidate the dynanic changes after transfection into cells of the main RNA replication templates presented in this research.

The importance of unraveling of the HCV RNA dynamics can not be underrated. A rise and a decay of minus and/or plus HCV RNA provided by transfection of a certain RNA template are directly linked to their possible applications and limitations. Every of the main replication RNAs used in the above sections – hp, SL I-II, SL I-III and 5'UTR (Fig. 3.2.1, A) – was subjected to transfection in a time-course manner together with the corresponding replication-deficient GND mutants. An additonal construct was introduced in this analysis: the so-called 5'UTR-Core construct (for the plasmid map see 6.1.35) that resembles the characterized 5'UTR replication template, with an exception that its open reading frame is represented by the complete HCV Core-coding sequence in contrast to the short upstream fragment provided in the 5'UTR construct. This construct was developed to allow the possible effector functions of the structural elements contained within the Core-coding sequence as well as of the Core protein itself.

To minimize bias, transfection reactions related to different time-points originated from a master-mix and the normalization procedure was conducted as following. Plus strand abundance was measured by RT-qPCR targeting the EMCV region at the plus strand (universal for all contructs) (Fig. 3.2.1, B) and, instead of being normalized to a certain cellular mRNA abundance, was normalized to input plus strand (i.e. transfected RNA template) abundance individually within each construct's series; in other words, per well of cells. This approach allowes to avoid serious bias rising from amplification of cells during a prolonged time-course as well as from the quality and relative amounts of input RNAs. Instead, it provides a fraction of remaining input RNA originated from the same master-mix. Therefore, the plus strand abundance for each construct at the time-point 0 is set to 1. Normalization of minus strand content derived by targeting the Spinach region by RT-qPCR (Fig. 3.2.1, C) was performed, as previously, to plus strand abundance, resulting in a comparison of each construct's ability to give rise to minus strands. All transfections in the current section were achieved by electroporation in order to ensure that all detected RNA had entered the cell and did not reside at the cellular surface or within endosomes, as for Lipofectamine-based transfection. The cells transfected with each of the tested constructs were lysed directly after transfection and further at 6-8 h, 16 h, 24 h, 36-48 h and 96 h post transfection (full time-course data available for most of the below-mentioned constructs).

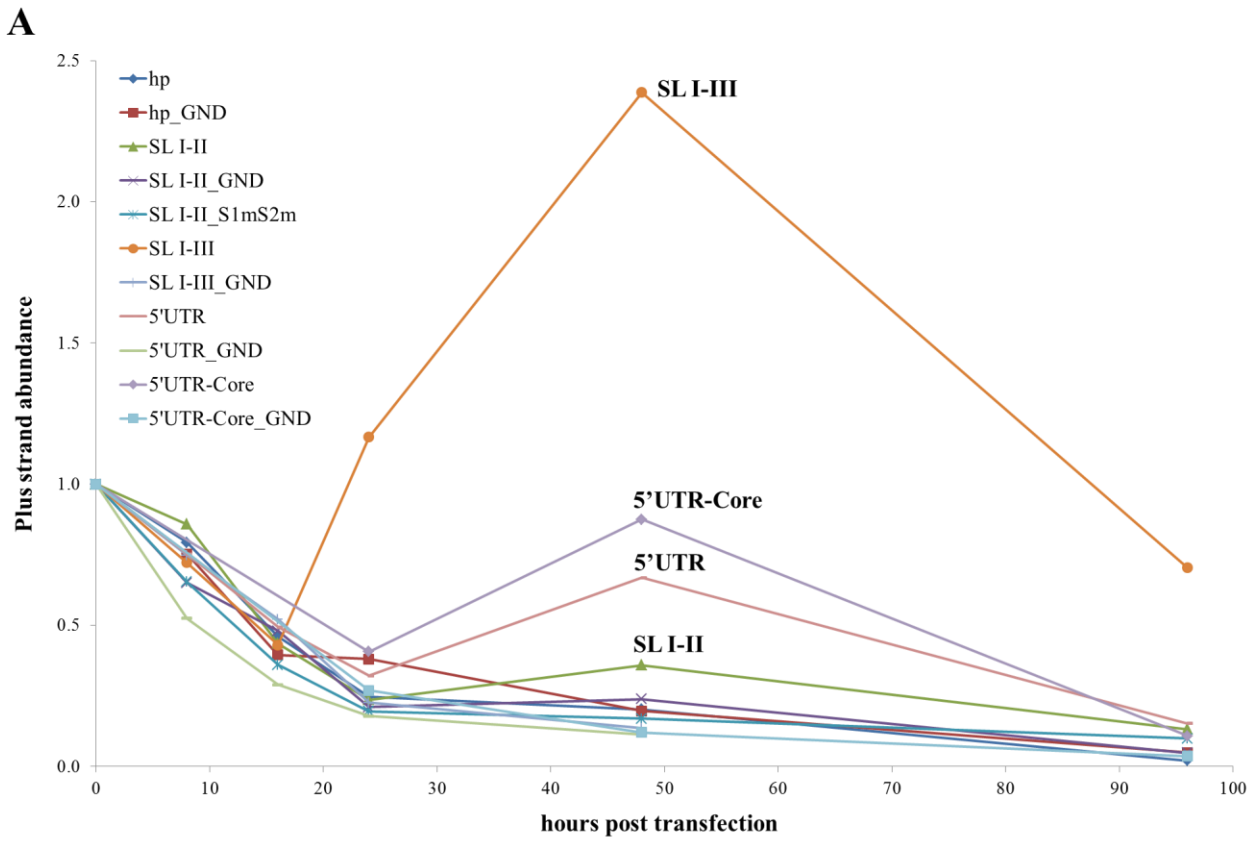
A summary of several time-course experiments (n = 3-6) is illustrated on Fig. 3.6.1. Regardless exact starting amounts of input plus strand templates (that may differ from experiment to experiment depending on integrity of *in vitro* transcribed RNA, efficiency of transfection and cell passage), a stable gradual decay of

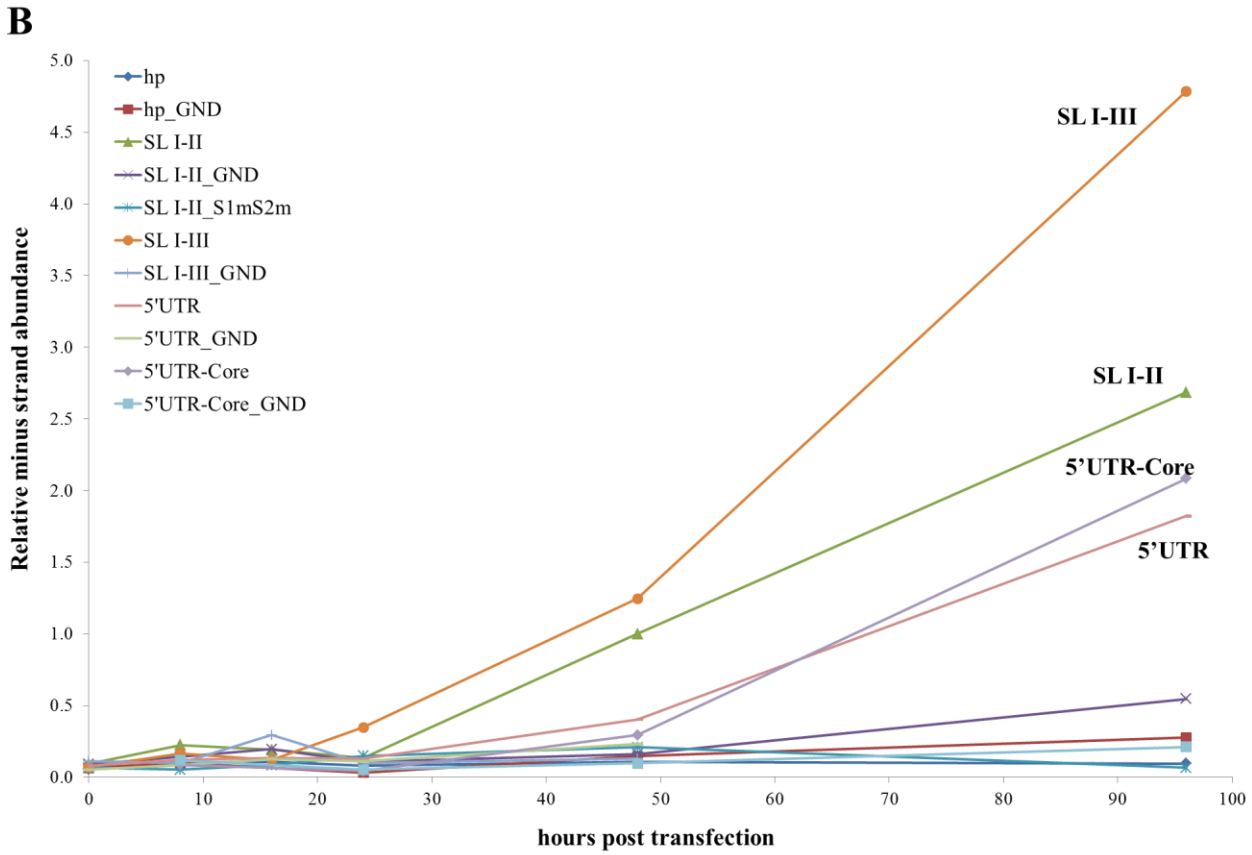
plus strand RNAs is observed within the first 16 h post transfection for all studied constructs: in both wild-type or GND context (Fig. 3.6.1, A). The minimum of the plus strand content is found to reside at 24 h time-point for most constructs; however no detailed time-course was conducted between 24 h and 48 h time-points to investigate it in more detail. Then, the plus strand abundance for some of the constructs elevates by the 48 h post transfection: significantly – for the SL I-III construct (starting at 24 h), slightly and variably – for the SL I-II, 5'UTR and 5'UTR-Core constructs. This observation strongly supports the capability of these constructs (especially of the SL I-III) to utilize newly synthesized minus strands (detected in the sections above) as templates for nascent plus strands synthesis. Importantly, no significant difference in plus strand accumulation was observed upon addition of the HCV Core-coding sequence to the 5'UTR constructs, emphasizing an inhibitory effect of functional translation from the HCV IRES on the construct's replication. Insignificant or no increase in plus strand abundance followed by the further decline was demonstrated by the rest of the constructs (hp and all GND mutants), verifying their inability to produce neither of the HCV genomic strands. Nevertheless, eventual decline of plus strand abundance for all experimental constructs by 96 h time-point takes place and is in agreement with descent of overall replication displayed in full-length viral systems. No further information on the fate of transfected RNA was obtained in scope of the current research.

Relative abundance of minus strands in corresponding total RNA samples is illustrated by Fig. 3.6.1, B. Unlike transfected plus strands, which are at their maximum at time-point 0, minus strands remain below detection level until 24 h post transfection, inclusively. As the graph clearly illustrates, relative minus strand abundance rises for the SL I-II, SL I-III, 5'UTR and 5'UTR-Core constructs at time-points 48 h and 96 h only. The replication-incapable GND mutants and the hp construct result only in background signals. The derivative of the SL I-II construct, harboring the mutated miR-122 binding sites at the 5'-end, demonstrates no accumulation of either of the strands, consistently with the results from section 3.4. Interestingly, the peak does not reside at 48 h time-point, as for corresponding plus strands' quantification, neither for relative minus strand abundance nor for raw minus strand signals prior normalization (Fig. 3.6.1, C). Likely, it originates from the combination of general decrease of replication rates after 48 h and of postponed synthesis dynamics of the studied constructs. When the replication slows down, plus strands are no longer synthesized in high excess over minus strands and all available minus strands are retained in double-stranded form. It would be of high interest to investigate the fate of both strands in a longer time-course; however it is largely complicated by viability of transfected cells at late time-points.

In order to monitor dynamic changes in HCV protein levels along a similar time-course, the HCV NS3 protein was monitored up to 96 h post transfection of replication RNA templates. Since all non-structural HCV proteins encoded in replication templates (except for the HCV Core in the 5'UTR-Core construct) are translated from the EMCV IRES independently of the constructs' nature and the HCV IRES sequence, ratios between detected NS3 protein should resemble availability of plus strand templates – input or nascent – that serve as templates for translation. Therefore, given transfection of equimolar amounts of RNA and identical transfection efficiency, protein levels detected at early time-points (as the first rounds of HCV protein translation start 4 h post transfection when driven from HCV IRES) should be similar for all constructs, including replication-deficient GND mutants. Considering a half-life of the HCV NS3 protein *in vivo* to be 11–16 h (Pietschmann et al. 2001; Pause et al. 2003) and decay of input plus strands within a similar time period (Fig. 3.6.1, A), relative increase in the HCV protein content should originate from translation from decreasing amount of input plus strands and emerging, when possible, nascent genomic strands.

As illustrated by the Fig. 3.6.2, A, at the time-point 16 h post transfection (upper panel), levels of detected HCV NS3 protein appear rather variable and systematic difference between wild-type and GND constructs is not observed. Additionally, the presented western blot illustrates emergence of background bands after longer exposure time, also in mock-transfected lane, which reside at the same level as the HCV NS3 protein and complicate the interpretation. However, corresponding transfection reactions result in fewer background and indicates profoundly different levels of the NS3 protein 36 h post transfection (Fig. 3.6.2, A,


A
hp
hp_GND
SL I-II
SL I-II_GND
SL I-II_S1mS2m
SL I-III
SL I-III_GND
5'UTR
5'UTR_GND
5'UTR-Core
5'UTR-Core_GND
SL I-III
5'UTR-Core
5'UTR
SL I-II
Plus strand abundance
hours post transfection


B
hp
hp_GND
SL I-II
SL I-II_GND
SL I-II_S1mS2m
SL I-III
SL I-III_GND
5'UTR
5'UTR_GND
5'UTR-Core
5'UTR-Core_GND
SL I-III
SL I-II
5'UTR-Core
5'UTR
Relative minus strand abundance
hours post transfection

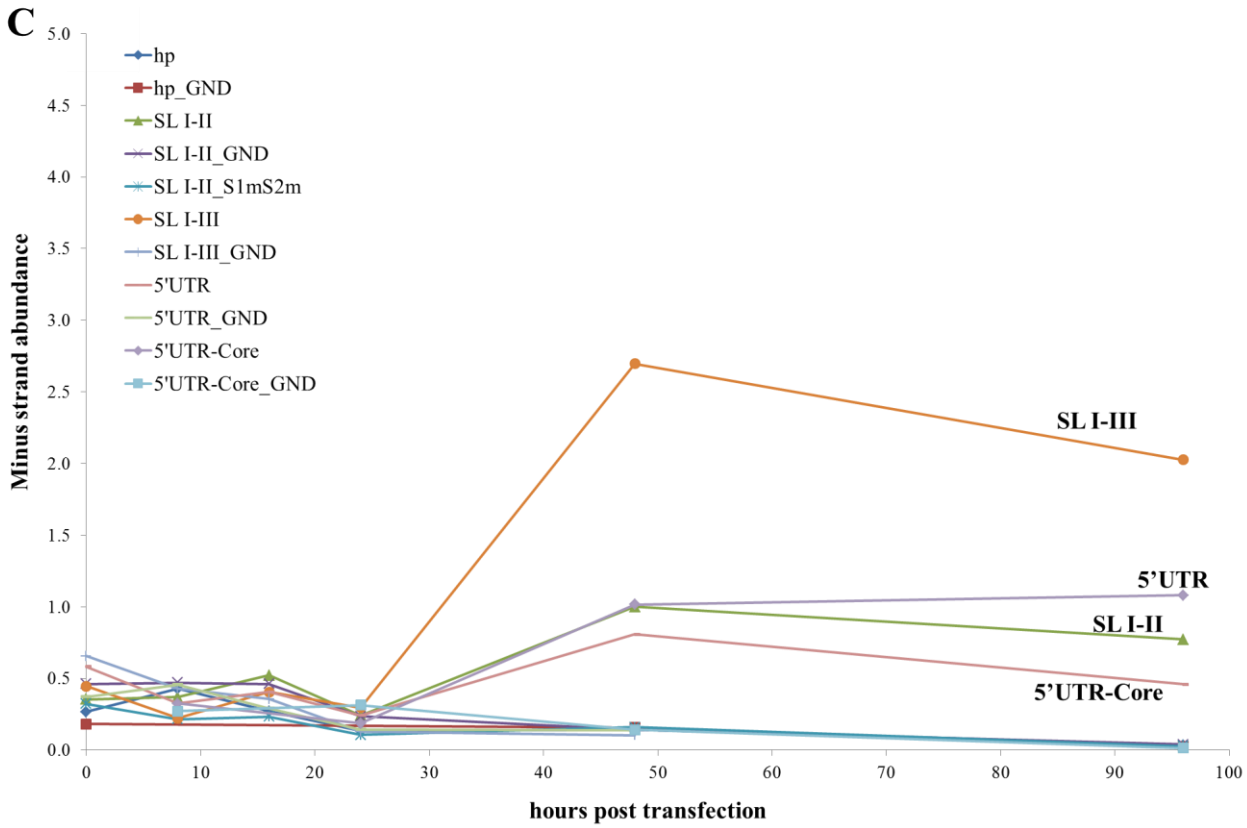


Figure 3.6.1: Dynamics of plus and minus strand synthesis by the RNA replication templates.
(A-C) *In vitro* transcribed RNA constructs harboring a hairpin (hp), SL I-II (wild-type or S1mS2m), SL I-III, a complete HCV 5'UTR with a short ORF or the HCV 5'UTR followed by a full-length Core-coding sequence – in wild-type or GND context - transfected by electroporation to assess the dynamics of input plus strands' decay (A) and of minus strand accumulation (B) during a time-course up to 96 h.
(A) Plus strand abundance is quantified by RT-qPCR targeting the EMCV IRES region on plus strands and normalized for each construct to its input RNA template amounts (time-point 0, set to 1). Fractions of remaining RNA for individual constructs were calculated accordingly and plotted together on a graph (see the legend). The means of remaining RNA fractions along the time-course are derived from three to six independent experiments. Error bars are not shown to simplify the representation, due to high variability between replicates.
(B) The graph represents quantification of the newly synthesized minus strands in the corresponding total RNA samples, obtained in time-course (A). Relative minus strand abundance results from the normalization of RT-qPCR targeting the Spinach region on the minus strand to plus strand abundance. The RER value for the wild-type SL I-II construct at time-point 48 h is set to 1; in relation to it, the fold differences of the relative synthesis efficiency are presented on diagrams. Error bars are not shown to simplify the representation; the means of relative minus strand RNA amounts in the time-course are derived from three to six independent experiments. In both assays all primers are universal for the examined constructs.
(C) Minus strand abundance quantified by RT-qPCR targeting the Spinach region prior normalization to plus strand abundance as shown in (B). The means of RER values for each construct are calculated in relation to the wild-type SL I-II construct at time-point 48 h (set to 1); accordingly, the fold differences of the relative synthesis efficiency are presented on diagrams. Error bars are not shown to simplify the representation. The data set is derived from three to six independent experiments.

lower panel). Evidently, a strong increase can be observed for the SL I-III construct, which is in agreement with the plus strand abundance time-course. Thus, pronounced elevation of plus strands in the SL I-III construct transfected sample supplies new templates for translation, whereas lesser represented plus strands in the SL I-II and 5'UTR samples seem to only maintain, but not increase viral protein content in cells at later time-points. Agreed to steadily lowering levels of input plus strands in GND samples (Fig. 3.6.1, A), these cannot keep the previous levels of the NS3 protein, and both – input template and newly synthesized protein – are subjected to consecutive degradation.

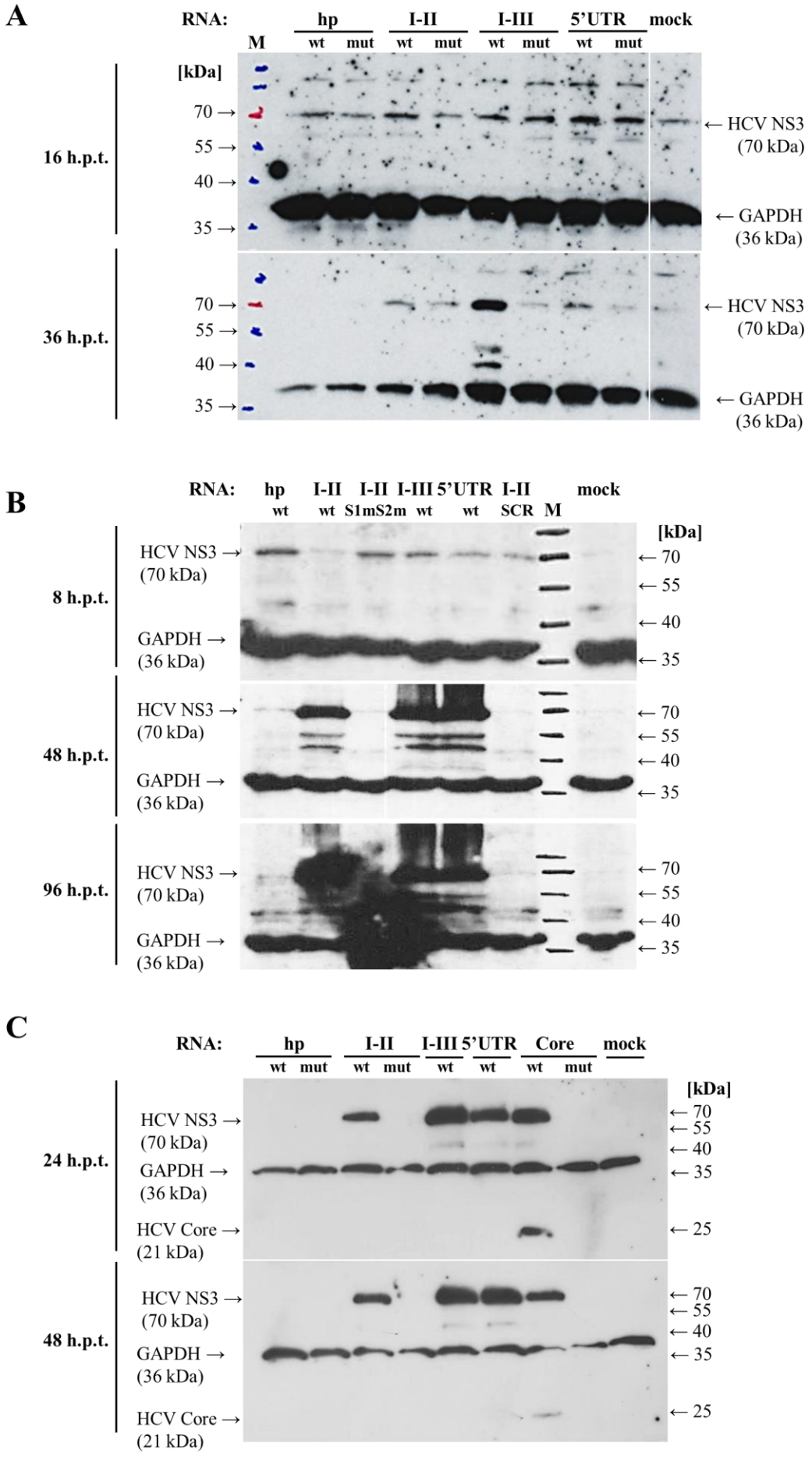
In an extended up to 96 h time-course similar fate of the HCV NS3 protein was observed (Fig. 3.6.2, B). At the early time-point – 8 h post transfection – low and variable levels of the HCV protein were detected (upper panel). At 48 h time-point the NS3 levels elevated dramatically for the similar set of constructs (SL I-II, SL I-III and 5'UTR) and remained unchanged by 96 h (exposure to one X-ray film). A

decrease of the HCV NS3 protein to background level by 48 h time-point supports the previously shown (at plus and minus strand RNA level) incapability of selected constructs (hp, SL I-II_S1mS2m and a scrambled version of the SL I-II) to undergo efficient minus and plus strand replication.

Another time-course aimed at direct comparison of the HCV protein abundance at time-points most relevant for the minus and plus strand RNA quantification upon transfection with the main set of functional constructs, including the 5'UTR-Core (Fig. 3.6.2, C). As evident from the RNA constructs dynamics (Fig. 3.6.1), already at 24 h time-point input template RNA becomes limiting, unless the RNA construct is capable of plus strand production. Limitation in templates for the HCV non-structural proteins translation from the EMCV IRES reflects dissimilar amounts of detected HCV NS3 protein. Neither of the replication constrcuts that were shown incapable of plus strand synthesis, such as all tested GND variants and also the hp construct, display detectable amounts of the NS3 protein (Fig. 3.6.2, C). The direct correlation between the number of genomic RNAs and protein synthesis is reflected by the highest NS3 protein content displayed for the SL I-III construct and the lowest – for the SL I-II. Importantly, the production of the HCV Core protein from the HCV IRES is demonstrated in amounts comparable to the HCV NS3 that is synthesised from the EMCV IRES, as the whole gene cassette. A decrease of both the NS3 and the Core protein amounts along the time-course imply that an addition of a vital Core-coding sequence at the constructs 5'-end nevertheless does not permit a persistence of the construct, which undergoes a subsequent decay (Fig. 3.6.1, A).

These illustrations may underline how tightly the processes of protein synthesis and replication of genomic and antigenome RNA strands are linked. Affecting either of the steps often completely disables the others process resulting in no readout. This emphasizes an importance of development of experimental systems, in which as few functions as possible are linked to each other. Uncoupling HCV replication from translation and minus strand synthesis from plus strand production (in many, but not all, of the constructs), the minus strand replication system allows determining the roles of selected elements through an introduction of one at a time mutations and tracking their effects on minus strand synthesis and overall HCV RNA and protein dynamics. One indispensable characteristic that should be taken into account when comparing various constructs and their derivatives is an overall stability of replication template RNAs. Since the abundance of minus strands in the current work is routinely normalized to plus strand abundance, stability appears a vital characteristic of experimental construct to account for. Indeed, results can be misinterpreted when the stability of examined templates, especially of the ones that are incapable of continuous replication, are drastically different. Based on the published data, certain modifications are already expected to affect stability of the HCV genome strand, such as mutations within the miR-122 binding sites at the 5'UTR that are well-known to serve in protection of the HCV genome from 5'-exonucleolytic degradation (Shimakami et al. 2012a). In this context, normalization of minus strand abundance to plus strands count still provides a more unbiased picture than normalization to cellular house-keeping gene, in particular for the constructs with high plus strand replication rates. Nevertheless, the need to account for

►Figure 3.6.2: Dynamics of the HCV NS3 protein synthesis from RNA replication templates.
(A) Western blot detection of the HCV NS3 protein (70 kDa) and cellular GAPDH (36 kDa) as a reference protein in HuH-7.5 cells lysates collected at 16 and 36 h post transfection (h.p.t.) with RNA constructs harboring a hairpin (hp), SL I-II, SL I-III, a complete HCV 5'UTR or with their replication-deficient GND mutants. A certain background signal is observed also in a mock-transfected lane at expected size for the HCV NS3. M, protein marker ladder.
(B) Western blot detection of the HCV NS3 protein (70 kDa) and cellular GAPDH (36 kDa) as a reference protein in HuH-7.5 cells lysates collected at 8, 48 and 96 h post transfection (h.p.t.) with the following RNA replication templates: hp, SL I-II (wt or S1mS2m), SL I-III, 5'UTR and SL I-II with a scrambled ORF (labeled as SCR). No or nearly no HCV NS3 protein was detected in lanes corresponding to hp construct or any of the mutants as well as in a mock-transfected cell lysate at 48 and 96 h post transfection. M, protein marker ladder.
(C) Western blot detection of the HCV NS3 protein (70 kDa) and cellular GAPDH (36 kDa) as a reference protein in HuH-7.5 cells lysates collected at 24 and 48 h post transfection (h.p.t.) with RNA constructs harboring a hairpin, SL I-II, SL I-III, HCV 5'UTR, HCV 5'UTR followed by a full-length HCV Core-coding sequence or with their corresponding replication-deficient GND mutants. HCV Core antibody allows visualization of the Core protein (21 kDa) expressed from the HCV IRES by 5'UTR-Core construct. M, protein marker ladder.


A
RNA:
hp
I-II
I-III
5'UTR
mock
M
wt
mut
[kDa]
70 →
55 →
40 →
35 →
16 h.p.t.
← HCV NS3
(70 kDa)
← GAPDH
(36 kDa)
36 h.p.t.
B
RNA:
hp
I-II
I-II S1m
I-II S2m
I-III
5'UTR
I-II SCR
M
mock
[kDa]
8 h.p.t.
48 h.p.t.
96 h.p.t.
HCV NS3 →
(70 kDa)
GAPDH →
(36 kDa)
← 70
← 55
← 40
← 35
C
RNA:
hp
I-II
I-III
5'UTR
Core
mock
[kDa]
24 h.p.t.
48 h.p.t.
HCV NS3 →
(70 kDa)
GAPDH →
(36 kDa)
HCV Core →
(21 kDa)
← 25

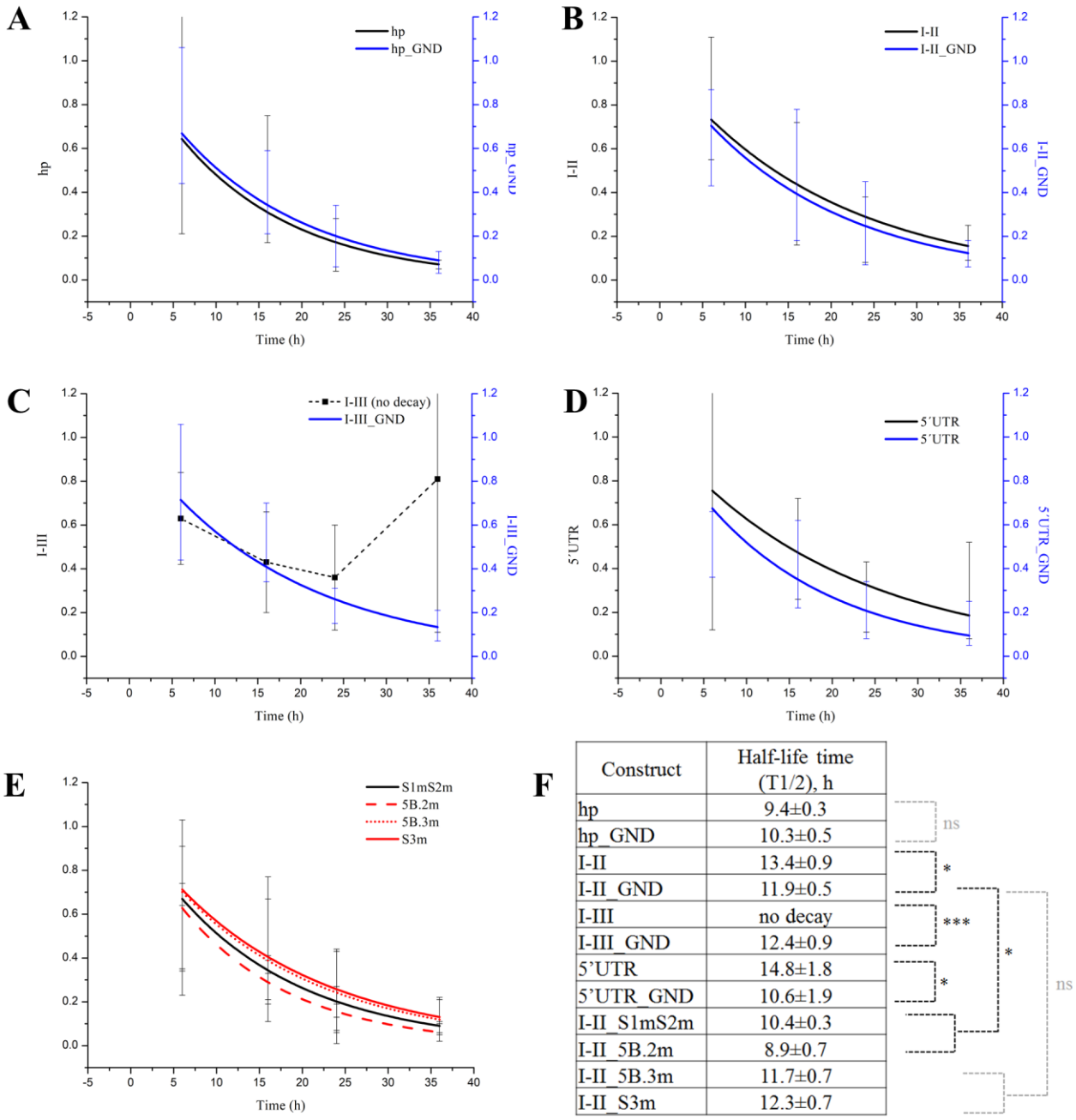
various stabilities of experimental constructs, in general, remains debatable in the field.

Characterization of the stability of the main RNA templates applied in the work was performed via a time-course experiment. As it was acquired in the above experiments (Fig. 3.6.1, A), input plus strand RNAs largely decay within the first 24 h post transfection. It can be approximated from the graphs that for the most of the constructs a half-life value is in a range between 8 and 16 h. Therefore, a detailed time-course was conducted to track input plus strands abundance at 0, 6, 16, 24 and 36 h post electroporation. As before, plus strand abundance was quantified by RT-qPCR targeting the EMCV region at the input RNA template.

Normalization was conducted individually for each construct to the plus strand signal derived for the time-point 0 (set to 1). Such normalization, i.e. to a well of transfected cells, results in fractions of remaining plus strand RNA that enable an approximation to an exponential decay function (conducted using OriginPro8 software and a ExpDec1 function) and subsequent calculation of half-life values. Cellular GAPDH mRNA was in parallel quantified in the same RNA samples to verify a comparable efficiency of total RNA isolation from cells (data not shown). Comparison of decay rates of the replication constructs hp, SL I-II, SL I-III and 5'UTR with their GND mutants and a number of miR-122 binding mutants (derivatives of the SL I-II construct) provided interesting results. The half-life time values for the hairpin (hp) constructs – both wild-type and GND version – demonstrate one of the lowest stability values among the investigated RNAs: 9.4±0.3 and 10.3±0.5 h, respectively (Fig. 3.6.3, A and F). For the pairs of the SL I-II and 5'UTR (wild-type and GND) constructs, the tendency of each replication-deficient construct to display a slightly lower half-life time was observed: 13.4±0.9 and 11.9±0.5 h for the SL I-II series and 14.8±1.8 h and 10.6±1.9 h – for the 5'UTR (Fig. 3.6.3, B, D and F). A fundamentally analogous situation is observed for the pair of the SL I-III (wild-type and GND) constructs (Fig. 3.6.3, C): as already shown in Fig. 3.6.1, A, the replication-competent SL I-III construct efficiently generates new plus strands and therefore cannot be fitted to a decay curve, however its GND version decays in a similar manner like other GND mutants and displays a comparable half-life time - 12.4±0.9 h (Fig. 3.6.3, F). Virtually, taking standard deviations into account, half-lives of all GND mutants vary around 10-12 h, with the lowest value for the hp_GND construct. As for the wild-type SL I-II, SL I-III and 5'UTR constructs, which are able to produce new progeny plus strands (Fig. 3.6.1, A), the calculated half-lives do not represent a true decay rate, unlike the ones for GND mutants, and can only be used to ensure a plus strand dynamics' difference when compaired to their replication-deficient versions. To sum up, concerning the above-mentioned constructs, exclusively their GND mutants should be utilized to elucidate possible alterations of RNA stability originating from sequence extensions.

Another group of RNA constructs subjected to stability assay was composed of the derivatives of the SL I-II construct with mutations affecting miR-122 binding to the HCV 5'UTR S1 and S2 sites (S1mS2m) as well as to the 3'-end binding sites (5B.2m, 5B.3m and S3m) (as in section 3.4). Providing that miR-122 functions on the HCV RNA in association with Ago2, it is expected to protect the original SL I-II construct from degradation at the 5'-end. Indeed, stability of the S1mS2m variant is slightly lower than of the SL I-II_GND RNA (10.4±0.3 and 11.9±0.5 h, respectively) (Fig. 3.6.3, E and F). The mutants of miR-122 binding to the two most downstream target sites - 5B.3m and S3m – do not demonstrate any significant difference in half-life time values when compared to the SL I-II_GND RNA (11.7±0.7 and 12.3±0.7 h), whereas the 5B.2m construct displays a substantial decrease in stability (Fig. 3.6.3, E and F): 8.9±0.7 h. Surprisingly, the 5B.2m mutant was the only miR-122 binding mutant that did not demonstrate any significant reduction of minus strand synthesis efficiency when compared to the wild-type SL I-II construct (Fig. 3.4.3, A). Nevertheless, the difference in stabilities of the studied constructs are not substantial enough (not exceeding 25 %) to affect the conclusions made from comparison of relative minus strand abundance in the above experiments. On the other hand, RNA half-life values need to be always minded – in experiments with reduced and full-length systems.

In summary, the replication-competent constructs featured in the current study unambiguously illustrate the tight association of the HCV protein and RNA synthesis: translation of substantial amounts of protein requires a sufficient amount of translation templates, whereas the HCV non-structural proteins are



F

Construct	Half-life time (T1/2), h
hp	9.4±0.3
hp_GND	10.3±0.5
I-II	13.4±0.9
I-II_GND	11.9±0.5
I-III	no decay
I-III_GND	12.4±0.9
5'UTR	14.8±1.8
5'UTR_GND	10.6±1.9
I-II_S1mS2m	10.4±0.3
I-II_5B.2m	8.9±0.7
I-II_5B.3m	11.7±0.7
I-II_S3m	12.3±0.7

Figure 3.6.3: Assessment of stability of experimental replication template RNAs.
(A-F) *In vitro* transcribed RNA constructs harboring a hairpin (hp), SL I-II (wt or with each of the miR-122 binding site mutants), SL I-III or a complete HCV 5'UTR – in wild-type or in GND context – transfected into HuH-7.5 cells by electroporation. Plus strand abundance is quantified by RT-qPCR targeting the EMCV IRES region in total RNAs isolated at 0, 6, 16, 24 and 36 h post transfection. For each construct plus strand abundance is normalized to its input RNA template amounts (time-point 0, set to 1), resulting in fractions of remaining RNA (y-axis) at each of 6, 16, 24 or 36 h time-points. For each construct means of these fractions were plotted (not shown) and fitted to exponential decay function (ExpDec1) using OriginPro8 software. Decay curves are presented on graphs (A) to (E) and were utilized to calculate half-life time values (T1/2) (F). Error bars display variability between means of remaining RNA fractions within three to six independent experiments.
(A) Decay curves for hp RNA template and its GND mutant.
(B) Decay curves for the SL I-II RNA construct and its GND mutant.
(C) A decay curve for the SL I-III_GND mutant, the wild-type SL I-III construct does not undergo decay and its plus strand abundance is plotted on the graph as a dashed black line.
(D) Decay curves for the 5'UTR RNA construct and its GND mutant. In (A-D) black curves correspond to wild-type constructs and blue – to their GND mutants.
(E) Decay curves for miR-122 binding mutants of the SL I-II RNA construct. A decay curve for the S1mS2m mutant is depicted in black, the curves for the 3'-end miR-122 binding mutants (5B.2m, 5B.3m and S3m) are in red, according to the legend.
(F) Half-lives (T1/2) and their standard deviations for the RNA templates calculated from (A-E).

essential for an arrangement of the HCV replication sites in transfected cells. Stability of RNA replication constructs is an important characteristic, which upon its increase may enable a more efficient overall replication rate. Dynamic at different stages of the HCV life cycle, the HCV genome stability depends on the current conformation of the plus strand RNA and on occupation of binding platforms on the RNA by various cellular factors. Thus, evaluation of half-life time of experimental RNA templates comes with a high degree of variation. The following can be summarized and hypothesized from the performed time-courses. The hairpin construct, protected from degradation at the 5'-end only by an artificial stem-loop, lacks the HCV SLs I and II and miR-122/Ago2 recruitment; hence has the lowest half-life time. Relatively low stability is demonstrated by the replication-deficient 5'UTR_GND construct; since this construct is capable of functional translation initiation from the HCV IRES, it may undergo transformational rearrangements coupled with translation, in turn exposing the RNA to degradation. This aspect should be addressed using another translation-competent construct – the 5'UTR-Core – as well as its translation mutants. A study on plus and minus strand RNA and HCV protein dynamics of constructs harboring translation abrogating mutations (ΔIIIb and IIId* in the SL III) remains to be performed. Comparison of stability of the wild-type translation-deficient 5'UTR RNA constructs can assist in unraveling the mechanistic aspects of switching between translation and replication. At last, the 5'- and 3'-end miR-122 binding mutants displayed the most variable half-lives among the studied RNA templates. While the decreased stability of the 5'-terminal miR-122 binding mutant is expected from published literature, the role of miR-122 binding at 3'-end requires further in-depth investigation.

4. Discussion

In this study an innovative reduced replication system for specific analysis of the HCV minus strand synthesis initiation was developed and applied for revisiting some of the vital requirements for the HCV replication and translation. A vast number of former studies focusing on mechanisms and regulation of the HCV life cycle events have been performed using a replicon or full-length HCV systems. These have provided indispensable insight into essential viral and cellular determinants for the HCV entry, translation, genome replication and assembly; however since harboring both genomic ends, the obtained results were often missing a precise assignment of each individual element to a specific function.

The current work was predominantly focused on uncoupling of the antigenome synthesis prerequisites from the ones for the genome strand synthesis and/or for translation, known to be all together engaged in an intricate interplay. The developed minus strand replication system applied in the study reveals that sequences present at the 5'-end of the plus strand HCV RNA are indeed positively involved in the regulation of the minus strand synthesis initiation at the 3'-end of the genome. The system enables a revision of functions of well-studied liver-specific miR-122 in HCV replication and provides a novel insight into its function upon binding to the 3'-end target sites. A vital contribution to the substantially controversial aspect of a switch between HCV translation and replication has been made. An importance of selected *cis*-acting RNA elements located within the protein coding region is reviewed and underlined for efficient antigenome synthesis. Nevertheless, it remains debatable whether the HCV RNA undergoes major genome rearrangements, i.e. circularization and/or dimerization, in analogy to other members of the *Flaviviridae* family.

4.1 RNA *cis*-signals requirements for the HCV genome replication

In the course of molecular evolution RNA virus genomes acquire optimal primary and secondary structures essential to fulfill their lifecycle in a particular cell type of a particular host organism. The primary RNA sequence is tightly linked to its structure, they co-evolve and advantageous combinations are retained and conserved within virus generations. For decades, members of the *Flaviviridae* family are known to contain the vital conserved *cis*-elements involved in virus translation, RNA replication and encapsidation within the 5'- and 3'- untranslated regions. A primary role assigned for the 5'UTR is a cap-dependent (genus Flavivirus) or cap-independent (genera Hepacivirus and Pestivirus) translation initiation. Notably, sequences that are not directly involved in translation are frequently found to be essential for overall virus production and characterized by high degree of conservation among virus isolates. 5'UTR elements on a genomic strand may contribute to a switch from translation to initiation of antigenome synthesis or to packaging. Moreover, complementary sequences located at the 3'-end of the minus strand are directly involved in the initiation of plus strand synthesis.

Taking advantage of a bicistronic replicon system (Lohmann et al. 1999; Fig. 1.1.4), comprehensive mapping of the HCV 5'UTR previously revealed that SL I and II constitute a region that is not essential for the HCV translation, but rather for RNA replication: up to 125 nt were found essential and the complete 5'UTR - required for efficient colony formation by replicons (Friebe et al. 2001; Kim et al. 2002). Such read-out enabled an uncoupling of the investigated sequences' role in translation from possibly overlapping functions in replication: in replicon systems the HCV non-structural gene cassette is driven from a heterologous IRES element, whereas a reporter gene is dependent on the HCV 5'UTR sequences. A related concept was applied in design of the replication system in the current study; however, in contrast to replicons, our system also attempts uncoupling of minus and plus strand synthesis, which is crucial since a number of elements were found to execute their functions on both genomic and antigenomic strands. The minimal hairpin (hp) construct appeared incapable of minus strand synthesis initiation, whereas the RNA construct containing the HCV SL I-II displayed ability for antigenome synthesis (Fig. 3.1.3, E), in agreement with the above mentioned studies. Nevertheless, requirement for the SL I-II sequence remains to be further challenged in order to determine the exact sequence and/or structure that are indispensable for fulfillment of

replication. A glance at an aspect of primary sequence requirements is discussed below in the context of miR-122 binding sites.

The SL I-II virtually represents the minimal 5'UTR sequence requirements for RNA replication initiation from the landmark publications, albeit contrasting results on the role of the downstream sequence have been obtained. Unlike in a replicon system (that cannot discriminate between minus and plus strand replication), our results demonstrate a strong enhancement of the minus strand synthesis in presence of the SL I-III sequences and a subsequent impairment with an addition of further downstream 5'UTR and protein-coding elements, i.e. SL IV harboring a translation initiation codon followed by only a short open reading frame or by the complete HCV Core-coding sequence (Fig. 3.2.3, Fig. 3.6.1, B). Neither of the two sequence extension variants following the SL I-III (5'UTR and 5'UTR-Core constructs) managed to restore the replication efficiency of the SL I-III construct, unless the translation function was impaired (Fig. 3.3.3, C and D). Notably, a role of a number of other reported candidate *cis*-acting RNA elements within the 5'-end on minus strand synthesis was not displayed in the context of our system. Mutational disruption of the structural integrity of the SL V and IV was shown to affect viral replication to various extents (McMullan et al. 2007; Vassilaki et al. 2008); however, the nature of the experimental systems used does not allow drawing a definite conclusion, at which step of the HCV replication cycle these elements actually are involved. Advantages of SHAPE analysis, enabling identification of novel conserved and biologically functional RNA elements, suggested a clustering of RNA regulatory elements within the Core and NS5B coding regions positively involved in HCV replication (Mauger et al. 2015; Pirakitikulr et al. 2016) (Fig. 4.1). Although a directed mutational analysis of these elements in the context of the minus and plus strand synthesis assay is still to be performed, it can already be concluded that the structural elements downstream the annotated SL III sequence are not essential for the HCV minus or plus strand synthesis and may rather play a role in translation. Whilst efficient translation elevates overall virus production (when analyzed coupled to replication as in the presented publications), when uncoupled from replication it appears to have an inhibitory impact on RNA synthesis.

A slight advantage of the 5'UTR-Core construct over the 5'UTR construct, encoding only for 36 nt of the Core-coding region, was observed in our replication assays (Fig. 3.6.1). Giving the similar ability for functional translation initiation of both RNA templates, they however differ in a number of above-mentioned *cis*-elements. Enhancement of RNA replication may originate from either a direct positive effect of structural regulatory elements or alternatively from an inhibitory effect of a region within the SL IV on translation (Wang et al. 2000; Kim et al. 2003). The latter is mediated by a long-range RNA-RNA interaction between a sequence flanked by the two miR-122 binding sites (position 24-38) and a stretch within the SL VI stem in the Core-coding region. Notably, the inhibitory effect on translation remains after frame-shift mutations, but not mutations of the primary sequence (Honda et al. 1999). An answer is to be obtained by follow-up experiments upon a relief of a proposed inhibitory effect of active translation from the HCV IRES: both constructs should be challenged in a replication assay when contain the mutations disabling the translation function (i.e. ΔSL IIIb; Ji et al. 2004; Otto and Piglisi 2004).

With regard to contribution of the genomic 3'UTR to the HCV minus strand synthesis, its crucial importance has been understood and thoroughly studied. Various requirements were proposed for different regions with an emphasis on the very 3'-terminal X-tail sequence conservation and its essential role in replication (Friebe and Bartenschlager 2002; Yi and Lemon 2003a,b). In our replication system the intact HCV 3'UTR sequence was retained (Fig. 3.2.1, A). On the other hand, an increasing number of reports on functional roles of *cis*-elements within the protein coding region in replication, prevalently NS5B, were found intriguing. Yet many of the proposed regulatory signals (Fig. 4.2) require revisiting, since their exact function and mechanism of action remain elusive. Relatively early reports on a group of highly conserved and stable stem-loop structures at the C-terminal NS5B coding region – 5BSL1-3 (alternatively named SL-VII to SL-V) – suggested their role in replication (Lee et al. 2004; You et al. 2004). Both a bulge and an apical loop within the *cis*-replication element 5BSL3.2 were proclaimed essential for the HCV replication

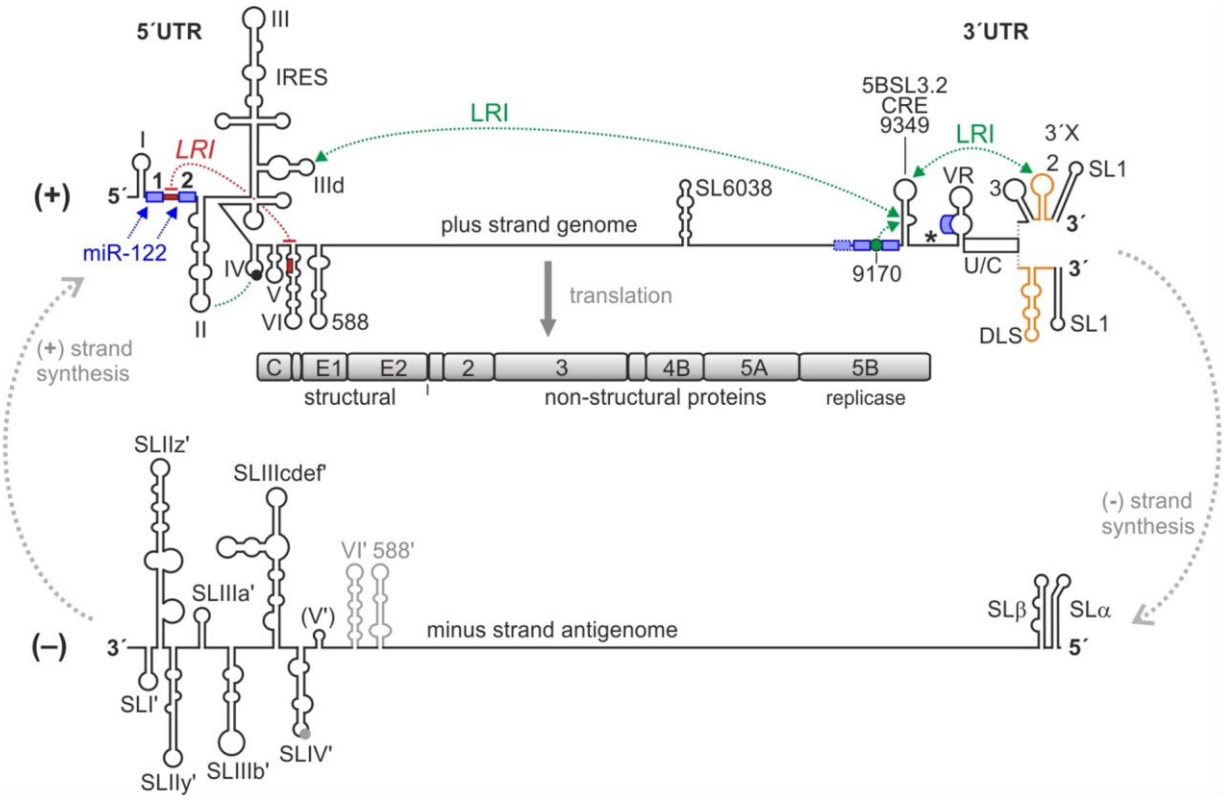


Figure 4.1: ***Cis*-acting replication elements in the Hepatitis C Virus genome replication.**
Plus strand RNA (top) serves as a template for a polyprotein translation (middle) and the minus strands synthesis (bottom), 3'-end of the antigenome strand, in turn, directs synthesis of nascent plus strands. Plus strand HCV RNA contains highly structured elements at 5'- and 3'- genome ends. At the 5'-end IRES-containing 5'-untranslated region is followed by structural RNA elements in the Core-coding region. At the 3'-end selected *cis*-elements within the NS5B coding region are depicted, followed by the structural determinants in the 3'-untranslated region. Structural elements at the 3'-end of the minus strand differ from the ones on the complementary 5'-end of the plus strand. Details on each element are provided in 4.1. Two alternating conformations of the 3'-X-tail are illustrated. The main long-range interactions (LRI) between the key *cis*-elements engaged in regulation of the HCV replication are indicated by arrows (in green – stimulatory, in red – inhibitory); elements thought to play a role in packaging are in orange. The binding sites for miR-122 are labeled by blue boxes at the 5'- and 3'- ends of the genomic RNA.
(from Niepmann et al. 2018)

due to formation of a "kissing-loop" interaction with the SL2 at the 3'UTR (Friebe et al. 2005). Moreover, both sequence and structure conservation is crucial in this element, suggesting multiple overlapping functions of the 5BSL3.2 (You and Rice 2008) (Fig. 4.2). Later, an importance of the other long-range RNA interactions engaging the 5BSL3.2 was understood: the subterminal bulge-loop interaction with a loop of an upstream *cis*-element SL9110 (named "9170" in Fig. 4.2) was also found crucial for the HCV replication (Diviney et al. 2009). A significance of the latter and other *cis*-replication elements was challenged using the developed minus strand replication system (Fig. 3.5.1). The 5BSL3.2 element was so far always retained unmodified due to a genetic proof of its utmost importance (Friebe et al. 2005). Notably, mutations affecting the interaction between the 5BSL3.2(bulge) and the SL9110 resulted in a reduction of minus strand synthesis efficiency, but not a complete abrogation of replication (Fig. 3.5.1, B; the SL9110 is designated as SL9170 according to its position in the JFH-1 isolate sequence). Another proposed long-range interaction established between the bulge of the 5BSL3.2 and an apical loop of subdomain IIId of the IRES (Shetty et al. 2013), when affected by mutations within the IIId loop sequence (266–268GGG to CCC), resulted in an attenuated minus strand production (Fig. 3.3.3, C and D). The latter mutation is known to impair another interaction at the HCV RNA: the integrity of the subdomain IIId apical loop sequence was found critical for the recruitment of the 40S small ribosomal subunit (Kieft et al. 2001; Ji et al. 2004) (Fig. 4.3). In agreement with a proposed function of the 5BSL3.2 as a molecular switch between viral translation and replication (Tuplin et al. 2015; Romero-López and Berzal-Herranz 2017), the mutation of the SL IIId prevents a long-range interaction that aims bringing the genomic ends together and therefore impair a switch to replication.

The above-mentioned interaction of the 5BSL3.2 with the SL2 (Fig. 4.2) that was reported to be required for overall viral replication in infectious HCV system (You and Rice 2008) was not challenged in scope of this work and can be addressed in further investigations. Controversial conclusions had been made on the exact role of this interaction in the HCV life cycle. Utilization of the full-length viral system complicates a dissection of a role in translation and/or replication by mutational analysis. An increase in observed viral replication efficiency when the 5BSL3.2(loop)-SL2 interaction is either intact or restored by compensatory mutations (You and Rice 2008) can indeed be due to the direct function of this LRI in replication or can be a result of impaired translation (Romero-López et al. 2012) that, in turn, indirectly shifts the balance to RNA synthesis. The latter observation is supported by the fact that an inhibitory effect on translation was relieved when the 5BSL3.2 or the 3'UTR regions alone were provided (Lourenço et al. 2008; Romero-López et al. 2012). Other studies, however, demonstrated either no 5BSL3.2-dependent effect on translation (Lee et al. 2004) or a contrasting positive effect (Tuplin et al. 2015); thus, further clarifications of the 5BSL3.2 function are required.

Findings on the role of RNA regulatory elements located within the protein coding region on the HCV genome postulate tuning of the overall virus production at different stages of the life cycle, however exact mechanisms are largely speculative. A majority of the *cis*-element candidates discovered recently by a combination of chemical probing, RNA structure modeling and functional assays (Mauger et al. 2015; Pirakitikulr et al. 2016) (Fig. 4.2) was not possible to challenge in scope of the present study. Some of them are claimed to engage only at late stages of virus production; another, such as SL8001, although predicted to limit the initial steps of replication, displays an indifferent effect in the cell culture model (Pirakitikulr et al. 2016). One of the elements, annotated as SL8647 (Chu et al. 2013) and claimed to be required for efficient genome replication, for *de novo* infectivity (Chu et al. 2013) and for production of infectious virus (Mauger et al. 2015), has especially drawn our attention. A sequence located in an apical loop of the stem-loop structure was predicted to interact with the 3'UTR dimerization linkage sequence (DLS) (Fricke and Marz 2016). Therefore, this sequence may play a role in conformational rearrangements during switching between replication and packaging and/or in genome dimerization and encapsidation (Ivanyi-Nagy et al. 2006; Shetty et al. 2010; Romero-López et al. 2014) (Fig. 4.2). In our experiments a construct harboring a mutated apical loop sequence within the element (designated as SL8680 according to its position in the JFH-1 isolate sequence) displayed a significantly lowered minus strand synthesis efficiency (Fig. 3.5.1, B), supporting the importance of this *cis*-element at the earlier stages of replication independently from packaging. Our replication system in further experiments can be applied for characterization of the other structural elements – J7880 and J8880 – predicted to positively affect HCV replication (Mauger et al. 2015) (Fig. 4.2); however possible interactions of these elements with the other RNA elements or cellular factors have not yet been determined.

A comprehensive screening for novel potential *cis*-replication elements has been initiated in the current study. RNA sequences with a regulatory function on various stages of the viral life cycle are predominantly highly conserved among virus isolates and frequently exhibited as single-stranded stretches on highly-ordered structures. So far a relatively highly-structured NS5B coding region was a target for a screening for *cis*-elements (Fig. 4.2). Additionally, a replication-suppressive function was assigned to the SL6038 element within the NS4B coding sequence (Pirakitikulr et al. 2016), and a few functional RNA elements were discovered within the HCV Core-coding sequence (Fig. 4.1), as mentioned above. Nonetheless, other genomic regions largely stayed out of focus. In the present work, in the context of the minimal replication-competent SL I-II construct (Fig. 3.5.1, A), most of the HCV NS3 to NS5B protein coding sequence was permutated to essentially disable all potential RNA *cis*-elements, while retaining amino acid coding specificity, codon usage and codon pair bias (in analogy to Song et al. 2012), though retaining the known to be essential for replication 5BSL3.2 and downstream signals unaffected. Recoding of the ORF upstream the 5BSL3.2 element led to a pronounced reduction in replication efficiency (Fig. 3.5.1, B) emphasizing the need for more detailed research. In follow-up experiments the entire scrambled sequence should be subdivided and separately analyzed as chimeras with wild-type sequences in order to narrow down

prominent regions. Although of an immense effort, these experiments will assist a discovery of novel regulatory elements and likely a validation of the recently published ones. Additionally, the scrambled RNA template can also appear useful in follow-up *trans*-complementation experiments: such construct when providing intact replication proteins, but only selected *cis*-elements, may restore a function of a reduced reporter RNA construct.

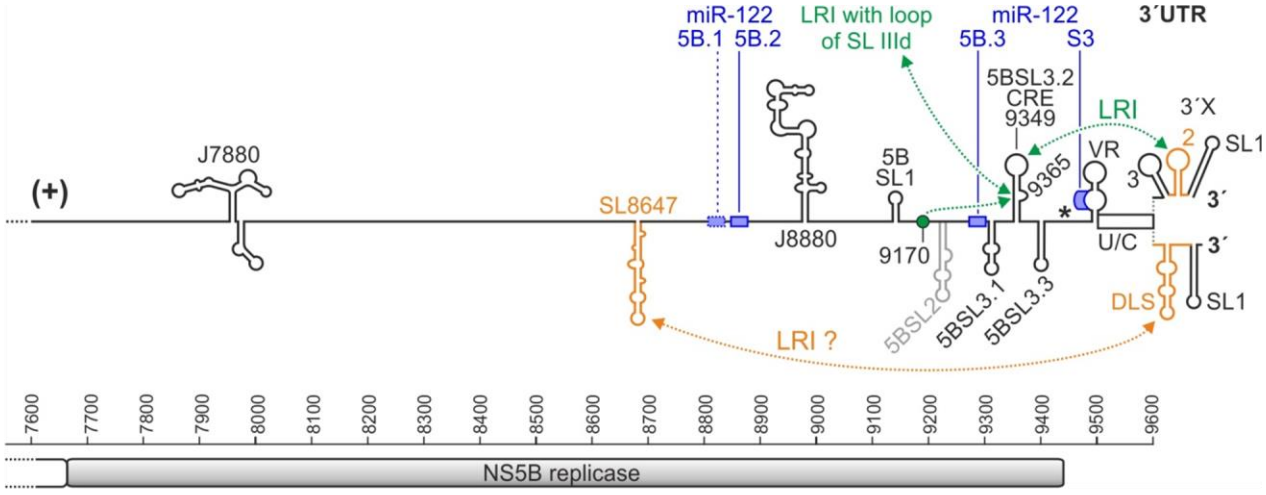


Figure 4.2: *Cis*-acting replication elements at the 3'-end of the Hepatitis C Virus genome.
A detailed representation of the annotated *cis*-replication elements at the 3'-end of the HCV genome. The 3'-terminus of the HCV ORF is indicated by the NS5B coding sequence. Further details on each element are provided in section 4.1. Two alternating conformations of the 3'-X-tail are illustrated. The main long-range interactions (LRI) between the key *cis*-elements engaged in regulation of the HCV replication are indicated by arrows (in green – stimulatory, in red – inhibitory); elements thought to play a role in packaging are in orange. The binding sites for miR-122 are labeled by blue boxes at the 5'- and 3'- ends of the genomic RNA.
(from Niepmann et al. 2018)

A discovery of novel *cis*-replication elements requires further investigation into their mechanism of action. Evolution of viral genomes maintains such elements conserved in sequence and structure, often as loops or internal bulges. The RNA regulatory elements on viral genomes serve for fine-tuning of molecular processes during a course of infection as well as to minimize requirements for proteins. Among the regulatory RNA elements some depend in their function on both sequence and structure, for others – structure plays a primary role and functionality of such elements can be rescued in mutational analysis by compensatory changes within a sequence of their interaction partner. RNA secondary structures may interact when located on a genome in close proximity or give rise to long-range interactions (LRI) that may bridge distances between a few hundred to several thousands of nucleotides. LRIs and their ultimate form – genome circularization – are well-understood to assist viral replication and other stages of life cycle, e.g. packaging. Genome circularization is proposed for many members of the *Flaviviridae* family, including HCV (Romero-López et al. 2009), some picornaviruses and retroviruses. Advantages of bioinformatic analysis nowadays enable computational prediction of novel RNA-RNA interactions, which however require meticulous validation. Potential LRIs occurring in the HCV genome were recently revisited by bioinformatic analysis using the "LRIscan" program, which has largely confirmed the existing knowledge as well as proposed new possible interactions (Fricke and Marz 2016). The replication system presented in the current study can represent a useful tool for verification of long-range RNA interactions in the HCV context.

Up to date only few RNA-RNA interactions were experimentally addressed by the system, such as interactions of the SL IIId and of RNA elements within the Core-coding region (see above). An attempt to experimentally demonstrate a circularization of the HCV genome has been undertaken. Generally, circularization of a viral genome is known to be mediated by RNA *cis*-elements and recruitment of *trans*-acting proteins; however protein-independent bridging of 5'- and 3'- genomic ends was reported for Dengue virus (Alvarez et al. 2005). For HCV, cellular and viral proteins are thought to assist a process of

genome circularization; although it remains unclear what is established first - the ends' communication via RNA-RNA hybridization - or their approximation by proteins interacting with both ends. Fulfillment of sequence prerequisites for circularization of the HCV genome has been bioinformatically demonstrated (Fricke et al. 2015). Formation of an alternative stem-loop within the 3'UTR X-tail and exposure of the DLS sequence (Ivanyi-Nagy et al. 2006; Shetty et al. 2010) allow hybridization to the complementary sequence within the apical loop of the SL II in the 5'UTR. This primary contact was predicted to extend up to 62 nt leading to hybridization of the HCV 5'- and 3'- ends (Fricke et al. 2015). Mutational approaches to validate this interaction remain challenging due to the crucial importance of primary sequence conservation at the HCV 3'-end; one such attempt to prove circularization failed, most likely due to the fact that any changes in the extremely conserved 3'UTR renders the genome inactive and cannot be rescued by any compensatory mutations at the other genome end (Fricke et al. 2015).

Indirect conclusions supporting the functional importance of the described interaction have been made from our experiments. Mutations introduced within a single-stranded stretch between SL I and II in the 5'UTR, aiming to disable miR-122 binding at this region, abrogated the minus strand synthesis (Fig. 3.4.3, A), as expected due to the known function of miR-122 binding for the HCV replication (Jangra and Lemon 2010). However, application of correspondingly mutated miR-122 variants that restore miR-122/Ago complexes assembly at the RNA template could not rescue HCV replication (Fig. 3.4.3, B) that evidently implies a critical importance of primary sequence conservation for the minus strand synthesis in our experimental system. Taking together that, firstly, a utilization of an RNA construct lacking the HCV IRES uncouples HCV replication and translation, secondly, that an impact of mutations on functions of the 3'-end of the minus strand in replication was minimized - the wild-type SL I-II construct does not support efficient plus strand synthesis (Fig. 3.6.1, A), and, lastly, that the miR-122 function in RNA stability was recovered due to Ago2 recruitment by mutated miR-122, the primary sequence itself must be concluded to be essential for replication. Several studies on miR-122 independent propagation support this assumption. An adaptation of the full-length HCV to mutations within the 5'UTR affecting miR-122 binding does not result in compensatory mutations to restore the miR-122 binding, but rather in alternative mutations that, nevertheless, reverted high replication rates (Hopcraft et al. 2016). Specific nucleotide changes introduced into the annotated very 5'-terminal HCV 5'UTR sequence (without altering the IRES activity) resulted in reduction of plus strand RNA synthesis on the minus strand 3'-end RNA template to various extents (Masante et al. 2008). Thus, apart from physical binding of miR-122 at the 5'UTR, efficient HCV replication appears to depend on RNA-RNA and/or RNA/protein interactions mediated by the 5'-end primary sequence of the annotated genome. A hypothesis that the primary sequence directly contributes to genome ends' contacts initiated either by long-range RNA interactions or by RNA-binding proteins requires experimental evidence. Attempts to facilitate replication by artificial bridging of the HCV RNA constructs' ends using complementary oligonucleotides did not provide any conclusive results likely due to rather subtle nature of genome rearrangements assisted by many different factors.

Notably, a computational prediction of long-range interactions for HCV and other members of the *Flaviviridae* family underlined their dissimilar distribution along viral genomes. The HCV genome was characterized by abundance of LRIs at the 5'- and 3'- UTRs, whereas for Flavivirus genomes LRIs are presented at the 3'-end as well as in the protein coding sequence. Nevertheless, interactions between the very genomic ends are predicted for both Hepacivirus and Flavivirus RNA (Fricke and Marz 2016). Such differences in LRI distribution may suggest different strategies applied by these viruses for genome circularization. Flaviviruses, such as Dengue virus (DENV), West Nile Virus (WNV) and Yellow Fever Virus (YFV), are known to circularize via the 8 nt motifs (that can be extended) located within the ORF at the 5'-end and at the very terminus of the 3'-end, and the degree of base pairing rather than exact sequence affects replication efficiency (Khromykh et al. 2001). The Human Immunodeficiency Virus type 1 (HIV-1) was found to be dependent on the viral nucleic acid chaperone protein NC and the cellular poly(A)-binding protein (PABP) for genome circularization and rearrangements. For HIV, spontaneous protein assembly facilitates the bridging between the 5'- and 3'- ends (Beerens and Kjems 2010). In Luteovirus and

Tombusvirus direct base-pairing of genomic ends and formation of specific kissing complexes regulate multiple processes including the switch between translation and RNA replication (Edgil and Harris 2006). At last, undertaking of a circular conformation by cellular mRNAs is essential for efficient translation, for which it is promoted by a PABP interacting with 5'-cap binding proteins (Mazumder et al. 2003). While indispensable for mRNA, protein bridges often play a secondary role in circularization of viral genomic RNA. Given the contrast between organization of the Flavivirus and HCV genomic ends as well as distinct mechanisms of translation initiation, the HCV genome circularization is thought to be primed by RNA-binding proteins.

Functions of selected *cis*-elements have been also extrapolated for the *trans* genomic RNA interactions. Existence of two alternative conformations of the 3'UTR X-tail (Fig. 4.2) had been shown both *in vitro* (Ivanyi-Nagy et al. 2006, Shetty et al. 2010) and bioinformatically (Fricke et al. 2015). The stem-loop originating from combination of SLs 2 and 3 exposes a short palindromic DLS motif (CUAG) in its apical loop that enables formation of RNA homodimers, leading to the speculation that HCV RNA genomes may dimerize. A validated kissing-loop interaction of the preceding SL2 with the apical loop of the 5BSL3.2 (Fig. 4.2) implies dynamic rearrangements of the X-tail depending on a state of the viral genome. Interplay between either of the RNA elements shifts equilibrium to one of the processes with a dimerized genome being an intermediate form. Essential functional regions 5BSL3.2 and the HCV IRES were found to exert contrary control over RNA dimer formation (Romero-López and Berzal-Herranz 2017). Our reduced replication system can represent a tool for clarification of HCV genome dimerization requirements for the minus strand RNA synthesis. A possibility to uncouple the HCV antigenome synthesis from translation, while retaining essential RNA regulatory elements, may assist dissection of primary switch determinants. It has been clearly demonstrated in our experimental system that enabling of translation from the HCV IRES has a negative impact on the efficiency of viral RNA synthesis (Fig. 3.3.3, C and D). Solely disabling the translation event without affecting RNA-RNA interactions largely, but not entirely, restored the replication rates, underlining that translation and replication are mutually exclusive processes. Mechanistic aspects of a switch between HCV translation and replication are discussed below.

Although specific aspects of the HCV plus strand synthesis initiation were out of focus of this work, extension of functional elements in RNA constructs enabled nascent plus strands production. All unmodified successors of the SL I-II construct (Fig. 3.2.1, A) displayed a various degree of genome RNA synthesis efficiency at the late time-points (Fig. 3.6.1, A). Largely controversial information had been published on RNA-element requirements for the HCV plus strand synthesis (Fig. 1.3.4). The studies utilizing *in vitro* system in the presence of the HCV RdRp (Astier-Gin et al. 2005; Masante et al. 2008) or replicon system (Friebe and Bartenschlager 2009) agree that SL I' is indispensable for RNA synthesis initiation at the 3'-end of the HCV minus strand. However, reports on a role of the following SL IIz' element are conflicting: the biochemical studies reveal an inhibitory impact (Astier-Gin et al. 2005), whereas in replicon system this sequence together with the third structural element SL IIy' were found to contribute to efficient replication (Friebe and Bartenschlager 2009). Lastly, unlike genetic studies, the *in vitro* experiments emphasize the role of the SL IIIb' structure in plus strand synthesis. Our data on a follow-up plus strand production indirectly points out the importance of the elements complementary to SL I and II. Indeed, the SL I-II construct demonstrated, although very low, ability for continuous replication (Fig. 3.6.1, A). It would be of a high interest to challenge a further truncated version of this RNA containing only the SL I at the 5'-end in order to rule out whether the basic levels of plus strand synthesis from it are due to inhibitory effect of the SL IIz' sequence or a requirement for more downstream elements. Nevertheless, the SL I-III construct in our experiments displayed an utmost ability for ongoing replication even comparing to the 5'UTR and 5'UTR-Core constructs capable of functional translation from the HCV IRES (Fig. 3.6.1, A). It is likely that determinants of the efficient minus strand RNA synthesis indirectly impact on abundance of plus strands, thus plus strand replication still remains to be dissected from the minus strand production. By the date translation is found to generally interfere with RNA synthesis: both versions of the 5'UTR containing constructs result in lowered plus strand synthesis rates. Although it remains to be confirmed by additional

experiments, the 5'UTR-Core construct displays a slightly higher abundance of both minus and plus strands, likely due to the presence of additional *cis*-elements (i.e. SL VI and SL 588) (Fig. 4.1) and therefore their complementary analogs on the minus strand, as well as due to an expression of the HCV Core protein. These aspects should be challenged by subsequent mutational analysis and assessed by tracking of the strands' dynamics.

Overall, the constructed replication system displays a high potential for investigation and clarification of debated aspects of HCV biology. In the setup presented in the current study the experimental system was not limited in viral (non-structural) and cellular protein determinants to focus on regulation of viral RNA synthesis. In particularly, with an emphasis on *cis*-replication RNA element prerequisites for antigenome synthesis, the following conclusions have been made. Uncoupled from other stages of the HCV life cycle, synthesis of the HCV minus strand RNA depends on certain 5'UTR sequences. Previously utilized replicon or full-length genome systems, although demonstrating strict requirements for these sequences, contained both HCV genomic ends and therefore could not clearly assign their role to minus or plus strand synthesis, or both. Our experiments evidently confirm that the initiation of the HCV antigenome synthesis from the 3'-end requires the SL I and II sequence at the 5'UTR. Moreover, minimal requirements for the HCV minus strand synthesis include a conservation of primary sequence within these elements, possibly to enable a genome circularization. The RNA replication template, lacking the HCV 5'UTR sequences, was confirmed to be incapable of minus strand synthesis, which was not possible to address in former systems due to a readout via overall replication and/or virus production, since HCV plus strand synthesis has to be initiated from the complementary sequences on the opposite strand. The more detailed investigation into the requirements within the SL I and II can be attempted in the future. Next, our replication system distinctly shows how further addition of the SL III significantly stimulates minus strand synthesis and enables a robust nascent plus strand production. The HCV SL I-III sequence not only allows an experimental construct for miR-122 binding at the 5'UTR and for possible genome circularization, but also for interaction with a number of HCV RNA-binding proteins including the 40S small ribosomal subunit. Partial addressing these factors and mutations of known regulatory sequences within these SLs is key for the dissection of exact determinants for the HCV minus and plus strand synthesis. A 3-nt mutation in the subdomain IIId seriously impaired the minus strand replication that, regarding the absence of the HCV IRES directed translation, underlines an importance of the HCV genomic ends communication and approximation driven by the long-range interaction 5BSL3.2-SLIIId and/or the 40S subunit binding at both ends. An even more severe effect on replication is observed when functional translation from the HCV IRES is permitted. Notably, impairment of the genomic ends communication by the same mutation in the apical loop of the SL IIId, which primarily restricts translation, then improves RNA synthesis, pointing out a strong inhibitory effect of translation on replication. Overall, plus strand synthesis within the translation-competent construct is also comparably lower than of the translation-deficient construct. So far, mostly hypothetical models are available to illustrate a balance between the HCV translation and RNA synthesis. Lastly, in the context of minimal required 5'-end sequences the role of selected *cis*-elements located within a protein coding sequence on minus strand synthesis efficiency was demonstrated. RNA elements SL8680 and SL9170, which were previously annotated to be vital for overall replication efficiency, are suggested to exert their positive effects via interactions with the DLS and the 5BSL3.2(bulge), respectively. The latter, when enabled, may play a role in stabilization of genome conformation that favors RNA synthesis rather than translation. More yet undiscovered regulatory RNA elements may reside within the HCV coding sequence, as was demonstrated after non-directed scrambling of the HCV NS3-NS5B ORF. Nonetheless, an impact of a number of *cis*-elements on HCV replication is still to be clarified in future studies using the developed replication system.

The concept of the system specifically focusing on minus strand synthesis – yet with a possibility to enable plus strand replication – can be potent for applications in context of the other plus strand RNA virus genomes. The minus strand replication system could serve to refine and analyze in-depth the information

obtained by full-length studies on such specific molecular step in viral life cycle as antigenome RNA synthesis.

4.2 Role of miR-122 in HCV replication

With discovery of the vast – and still expanding – variety of cellular microRNAs, enslavement of host miRNAs by viruses had been reported as one of the strategies for efficient virus propagation. Apart from a direct binding to the viral genome, miRNAs were found to exert their function via complex regulatory pathways. Viruses in turn can suppress an expression of antiviral and induce an expression of proviral cellular miRNAs. Moreover, many DNA viruses (e.g. HSV, HCMV and EBV) and some RNA viruses (e.g. WNV, HIV-1) even encode their own miRNAs to shape the virus-host interaction (Kincaid and Sullivan et al. 2012).

There are several limiting factors, such as insufficient abundance of miRNAs in a particular cell type, that prevent its utilization by viruses. Among RNA viruses, HCV and its homologs (e.g. the GB virus B, GBV-B, the equine Non-Primate Hepacivirus, NPHV) take advantage of the highly abundant and liver-specific miR-122 that itself is virtually the major regulator of lipid metabolism and lipoprotein assembly in hepatocytes (Tsai et al. 2012), which is exploited by the HCV on site for lipoviral particle formation (Andre et al. 2005). MiR-122 appears to be the most influential host factor in the HCV life cycle that (together with receptor specificity and a dependence on lipid metabolism) determines a tropism of the HCV to hepatocytes (Yu et al. 2018). It can only be speculated, which of these factors have predetermined the tropism, however, the exploitation of miR-122 was most likely a secondary adaptation to enhance exploitation of the liver environment: the related Hepaciviruses that are only partly dependent on miR-122 display a broader tissue tropism (Yu et al. 2018).

Along with providing a favorable environment, miR-122 promotes the HCV life cycle via a direct function on the HCV RNA during translation and replication. Only few other RNA viruses are characterized by direct functions of cellular miRNAs on their genomes. An efficient replication of the Hepacivirus called NPHV (related to HCV) requires miR-122 binding to a single site at the genome 5'UTR (Kapoor et al. 2011). For Bovine Viral Diarrhea Virus (BVDV), interaction of miR-17 and let-7 with the genomic 3'UTR was found critical for translation and viral RNA synthesis (Scheel et al. 2016). The miR-10a-3p was found to exert post-transcriptional regulation of Coxsackievirus B3 (CSB3) RNA via direct targeting of the 3D-coding sequence (Tong et al. 2013). Lastly, Poliovirus was recently reported to be dependent on miR-134 (Orr-Burks et al. 2017).

Its most well-known effects miR-122 exerts at the two adjacent conserved target sites S1 and S2 (Fig. 3.4.1; Fig. 4.1) in the HCV 5'UTR (Jopling et al. 2005), which is in agreement with the strongest Ago binding mapped to these sites by CLIP (Luna et al. 2015). Since miR-122 was found to enhance overall HCV RNA accumulation via a modulation of translation (Henke et al. 2008), replication (Jopling et al. 2008) and RNA stability (Shimakami et al. 2012a,b), it remains a challenge to designate precise mechanisms of each of the miR-122 actions. The current study focuses on re-evaluation of the requirements for miR-122 binding to the HCV RNA in fulfillment of virus minus strand RNA synthesis. The developed system allowed uncoupling of the HCV minus strand synthesis from translation and largely from the plus strand synthesis. Mutations affecting miR-122 binding were generated within the experimental construct (Fig. 3.4.1, A), which contains all annotated target sites both at the 5'- and 3'- ends of the HCV genome, and provided an information on requirements for each of the sites in the HCV minus strand production.

In agreement with published results, binding of miR-122 to the S1 and S2 binding sites at the 5'UTR was found essential for efficient antigenome synthesis (Fig. 3.4.3, A). Moreover, a weak ability for nascent plus strand synthesis, expressed by the wild-type SL I-II construct, was completely abrogated (Fig. 3.6.1, A). A confidently lower stability of the RNA replication template harboring the S1m and S2m sites in comparison to a replication-deficient GND construct in turn confirms the importance of these sites in protection of the viral RNA from exonucleolytic degradation (Fig. 3.6.3, F). However, the observed

difference in stability is not profound enough to be the only cause of impaired replication in our system. Indeed, a knockdown of the Xrn1 – the major 5'-exonuclease responsible for the HCV RNA degradation – fails to rescue a replication of HCV mutant defective in miR-122 binding to the 5'UTR (Li et al. 2013; Masaki et al. 2015). When an increased susceptibility to decay was compensated by application of correspondingly mutated miR-122 variants, replication was not restored (Fig. 3.4.3, B). This fact points out that, apart from physical binding of miR-122, the primary nucleotide sequence of the binding sites appears vital for efficient minus strand synthesis. This observation provides an indirect, but strong, argument for a role of SL I-II sequence in replication enhancement though genome circularization, as thoroughly discussed above. The latter can be primed by RNA-RNA long-range interactions (Fricke et al. 2015) and/or protein bridges, e.g. via PCBP2 protein that binds both to the SL I in the HCV 5'UTR and to the poly(U/C) tract in the 3'UTR (Wang et al. 2011). Notably, alternating occupation of the HCV 5'UTR by PCPB2 protein and miR-122 is proposed to mediate a switch between translation and replication (Masaki et al. 2015) that also underlines a critical role of primary sequence in regulation of the HCV replication by the SL I-II region.

Although the S1 and S2 miR-122 binding sites have been hypothesized to unevenly contribute to each of the proposed functions in the HCV life cycle upon binding of miR-122 (Thibault et al. 2015), in the present study they were so far addressed only simultaneously. Taking into account that the two sites are located in highly conserved proximity from each other (Fricke et al. 2015) and that miR-122 binding occurs cooperatively (Thibault et al. 2015; Nieder-Röhrmann et al. 2017), separate addressing of the S1 and S2 sites may lead to misinterpretation of their individual functions due to generally weaker binding when only one of the two sites is mutated. Both of the 5'UTR sites were shown to equally and cooperatively contribute to HCV replication in the replicon system (Thibault et al. 2015); however it would be interesting to revisit this statement in the context of minus strand synthesis using our experimental system. Controversial results obtained by the replicon system (Thibault et al. 2015) suggested that miR-122 binding to the S1 site (and therefore the effect on replication) is not saturated in HuH-7.5 cells. In our experiments, using the wild-type SL I-II construct transfected into the same cell system, an ectopic supplementation with additional miR-122 resulted in a highly variable, however insignificant, increase of replication efficiency (Fig. 3.4.3, B). This is an intriguing and under-investigated in the context of HCV replication matter that requires a further insight, when S1 and S2 miR-122 target sites are addressed separately.

An importance of another cluster of miR-122 binding sites located at the 3'-end of the HCV genome (Fig. 4.1) was recognized relatively recently. These sites demonstrate an exceptional degree of conservation among the HCV isolates (Fricke et al. 2015) and mediate recruitment of miR-122/Ago2 complexes in a similar manner to the 5'UTR sites (Luna et al. 2015), implying their functional roles. Up to date contradictory results have been published regarding the roles of the 5B.2, 5B.3 and the 3'UTR S3 miR-122 binding sites. There were indirect evidences of their functions in HCV translation and replication exerted independently of the miR-122 interaction with the 5'UTR target sites (Henke et al. 2008; Jangra and Lemon 2010). Another study in replicon system (Nasheri et al. 2011) assigns an inhibitory role to the 5B.2 and S3 target sites and mechanistically links it to slowing down actively translated ribosomes leading to negative modulation of the overall HCV production aiming at suppression of viral titers for purposes of escaping immune response evasion (Mahajan et al. 2009). A more recent study using the full-length HCV system demonstrated an overall stimulatory impact of the 5B.2 and - to a smaller extent - 5B.3 sites, but not S3, on total HCV RNA and protein levels (Gerresheim et al. 2017). In the present work roles of each individual miR-122 binding site at the HCV 3'-end in minus strand synthesis were challenged by mutational analysis. Mutations within each of the conserved sites (the non-conserved 5B.1 site was inactivated) were generated in the context of the replication construct harboring only SL I-II 5'UTR sequence (Fig. 3.4.1, A). In contrast to any of the previous studies, the 5B.3 and S3 site were found to be essential for antigenome synthesis, whereas the 5B.2 site was not found to contribute (Fig. 3.4.3, A and C). Strictly, this result is a unique finding and cannot be reviewed in the context of previous studies that were accounting for overall HCV production and incapable of dissection of virus antigenome synthesis from other steps of the life cycle.

Interestingly, an opposing effect on replication RNA constructs' stability was observed upon these mutations: a mutation in the 5B.2 site significantly decreased the half-live time of the SL I-II construct, whereas mutations within the 5B.3 and S3 sites had no effect on RNA stability when compared to a corresponding replication-deficient GND construct (Fig. 3.6.3, F). Until the more in-depth research on the function of these sites is conducted, we can only hypothesize that the interactions at the 5B.2 site somehow mediate preservation of viral RNA, and at the 5B.3 and S3 sites – a direct modulation of minus strand replication. When evaluating a role of a selected miR-122 target site, it is also crucial to take into consideration that effector functions of miR-122 binding sites were reported to be dependent on the sites' accessibility (Sagan et al. 2010; Gerresheim et al. 2017). Under the circumstances of highly ordered RNA structure of the NS5B coding region and the 3'UTR, a function of the 5B.1-5B.3 and S3 binding sites at a certain step of the viral life cycle may be determined by an actual conformation of the genome permitting or not an interaction of miR-122 with required sites. The RNA-RNA long-range interactions and RNA-protein interactions that reshape a genome conformation in favor of a particular process may play a key role in engaging miR-122 to its target sites, when needed. This fact can help an explanation why application of compensatory miR-122 variants (mutated in correspondence to each of the miR-122 binding sites) appeared insufficient to rescue an efficient minus strand synthesis (Fig. 3.4.3, C). Although the binding *in vitro* of mutated miR-122/Ago2 complexes to the targets was restored, the mutations may unpredictably affect the neighboring *cis*-elements' communication.

A dilemma whether miR-122 exerts its function via a physical interaction with a selected target site or indirectly by affecting the formation of RNA secondary structures and long-range interactions, or both, should be addressed in the future studies. It is not always possible to dissect functions of certain miR-122 site(s) when the other(s) are mutated, as it was demonstrated for a series of the SL I-II_S1mS2m construct, since the mutations within the S1 and S2 sites already completely abrogate the RNA synthesis and moreover could not be rescued by compensatory miR-122 variants (Fig. 3.4.3, A). However, a combination of mutational analysis and application of antisense LNA oligonucleotides can offer a possible solution. In such an approach individual miR-122 sites can be obstructed without affecting the primary sequence by genome-specific LNAs; or the anti-miR-122 LNA in combination with target sites' mutations and compensatory miR-122 can be applied. The latter approach may recreate conditions of limiting miR-122 and point at prioritized targets of miR-122 binding and their functions. So far, the anti-miR-122 LNA/DNA mixmer was only utilized together with unmodified SL I-II construct and resulted in abolishment of the HCV replication and protein synthesis from the construct. Such effect of sequestering of cellular miR-122 by this LNA in the context of wild-type RNA primary sequence underlines an importance of physical binding of miR-122 for efficient minus strand synthesis (Fig. 3.4.3, B) and robust HCV production (Gerresheim et al. 2017).

A more fundamental role of the miR-122 binding at the HCV genome 3'-end can be proposed. Often miRNAs target a 3'UTR of mRNAs, in part to prevent collusion of miR/Ago2 complexes with scanning or translating ribosomes. In contrast, an array of 3'-terminal miR-122 binding sites at the HCV genome is located just at the very end of a protein open reading frame. This suggests an inviting mechanism of functional ribosomes deceleration mediated by three to four miR-122/Ago2 complexes. A high degree of conservation of the 5B.2, 5B.3 and S3 binding sites among the HCV isolates as well as of distances between them (Fricke et al. 2015) strongly supports this assumption. Ribosome pausing is reported to occur during elongation or termination of translation and serve a quality control purpose in both pro- and eukaryotes (Buskirk and Green 2017). In HCV such elongation deceleration may serve to facilitate the nearly full-length, but still held at its C-terminus, NS5B RNA polymerase to find the 3'-end of an apparently intact genomic RNA target. Whence, the C-terminus of the HCV RdRp does not exert the enzymatic activity (Moradpour et al. 2004; Bartenschlager et al. 2010), the initiation of antigenome synthesis can be enabled even before completion of the NS5B protein assembly. This intriguing and yet to be confirmed function of the adjacent miR-122 binding sites in HCV replication is indirectly supported by our results. Indeed, dissimilar requirements for each of the 3'-end target sites suggest that the most upstream of the investigated sites 5B.2 is not required for minus strand synthesis initiation, whereas the 5B.3 and S3 site were found

indispensable for antigenome synthesis as well as for maintenance of continuous replication (Fig. 3.4.2, B). Controversially, in the full-length HCV system requirements for each of the 3'-end miR-122 binding sites on overall virus production were found to be right the opposite (Gerresheim et al. 2017). It has to be noted that our replication system provides an insight on the function of these sites selectively on minus strand synthesis uncoupled from other stages of the HCV life cycle. It can only be speculated whether the full-length virus possesses certain compensatory mechanisms to support efficient antigenome synthesis and which other yet undiscovered function do these binding sites implement during the HCV life cycle.

At last, only little is known about possible roles of other miRNAs associated with the HCV infection. Several cellular miRNAs, such as miR-199a, let-7b, miR-181c, miR-448 and miR-196, were reported to bind at the genome 5'UTR and at the 3'-end (within the coding sequence and 3'UTR) and be implicated in suppression of the HCV RNA replication (Ojha et al. 2016). A negative impact on HCV propagation of these miRNAs suggests a therapeutic potential alternative to miR-122 sequestration (Lanford et al. 2010; Janssen et al. 2013), which requires unraveling of mechanistic details of action of these miRNAs. The replication system presented in the current study may assist in understanding of the roles of these as well as the other host factors acting on the HCV RNA in virus replication and representing a possible target for anti-HCV therapy.

In conclusion, the developed replication system has succeeded to provide novel insights into such controversial aspects as a role of miR-122 in HCV replication, in particular in the minus strand synthesis. With regard to the influence of miR-122 on such aspects of the HCV life cycle as translation and stability, although thoroughly revised by the scientific community, open questions still remain. The current work confirms a positive role of miR-122 binding to the sites located at the 5'UTR, in addition underlining a requirement for intact 5'UTR SL I and II primary sequence for the HCV antigenome synthesis. Similar effects on efficiency of the HCV minus strand replication were observed when the downstream 3'-end miR-122 binding sites 5B.3 and S3 were challenged by mutational analysis. Principally, an action of miR-122 at both genome ends was demonstrated to modulate the HCV minus strand synthesis uncoupled from other molecular processes. On the one hand, this way of action is dependent on actual recruitment of miR-122/Ago2 complexes, as confirmed by sequestration of endogenous miR-122 by LNA. On the other hand, it depends on integrity of the primary sequence, as evident from the failure of rescue experiments using miR-122 variants with compensating mutations. Interestingly, supplementation of the experimental cell system with additional miR-122 did not result in a confident increase of replication rates. This points out that the amounts of miR-122 in HuH-7.5 cells are sufficient for its activities on the RNA replication constructs. Overall, a number of hypotheses require more detailed investigation empowered by further mutational analysis in the context of the developed system. One of them is an apparent significance of overlapping and/or adjacent RNA regulatory sequences and structures in a complex interplay between major molecular processes engaging viral RNA. Additional non-canonical roles of miR-122 binding to the HCV genome, such as ribosome pausing or switching the balance between translation and replication, require further studies.

4.3 A balance between translation and replication

A tight association and mutual dependence of all intracellular stages of the HCV life cycle together with a need for separating them in time and/or space suggest an existence of a strictly regulated balance between the processes. Early rounds of virus translation take place in the cytoplasm in association with the ER-membrane; however membrane compartments' refinement is subsequently induced by viral proteins to form an optimal and enclosed environment for the HCV RNA synthesis. In these compartments – double-membrane vesicles – antigenome synthesis followed by progeny plus strands replication take place. The latter induce new rounds of viral translation and replication and the switch between the two mutually exclusive processes is ought to be mindfully governed. Globally driven by shortage or excess of products of the HCV life cycle, the balance between translation and replication (and at the late stages – packaging) is currently thought to be regulated by an intricate interplay between structural RNA elements and host and

viral proteins simultaneously acting at one or two copies of the HCV genome. These *cis*-acting replication elements (as discussed above) and *trans*-acting factors (as described below) are generally entangled in multiple alternating interactions that profoundly complicate analysis of their individual function as well as of exact molecular prerequisites for each step of virus propagation. In the present study analysis of requirements for HCV minus strand synthesis enables a contribution to the state-of-the-art mechanism concerning a balance between translation and replication upon HCV infection.

According to the current model, the HCV genome assumes a circular conformation at various stages of the life cycle to facilitate viral protein and RNA synthesis, as it is common for other members of the *Flaviviridae* family, which is supported by multiple experimental and computational studies. The HCV IRES, 5BSL3.2 and X-tail are thought to be the major elements establishing the RNA-RNA and RNA-protein communication and subsequent structural rearrangements of the genome. Following entry of the HCV genome into the cytoplasm, the viral RNA instantly engages in interaction with cellular proteins. The HCV 5'UTR serves as a binding platform for recruitment of the host translation machinery, in which the GGG sequence within subdomain SL IIId plays a key role in efficient and stable 40S ribosomal subunit binding, whereas the domain SL IIIabc is critical for the functional translation complex assembly (Kieft et al. 2001). When the SL IIId is occupied by bound ribosomes and therefore is not available for long-range RNA-RNA interactions, the 5BSL3.2 is present in a closed conformation (i.e. interactions 5BSL3.2(loop)-SL2 and 5BSL3.2(bulge)-SL9110) (Diviney et al. 2008; Tuplin et al. 2012; Shetty et al. 2013; Romero-López et al. 2014); and a potential interaction site with the RNA polymerase is likely occupied by hnRNP A1 protein (Ríos-Marco et al. 2016). The poly(U/C) tract at the 3'-end is occupied by PTB protein (Ito et al. 1998) that can in turn bridge the genome ends via an interaction with hnRNP L protein residing at the 5'UTR (Kim et al. 2000). The ends' communication at the stage of translation is strongly supported by enhancement of translation in presence of the HCV 3'UTR (Song et al. 2006) and by a non-competing binding of the 40S subunit to both genomic ends (Bai et al. 2013). PCBP2 is another cellular protein that binds to both genomic ends at sites partially overlapping with miR-122 binding. All these RNA-protein interactions facilitate viral translation through protection of genome ends and recycling of ribosomes on a circular template (Fig. 4.3).

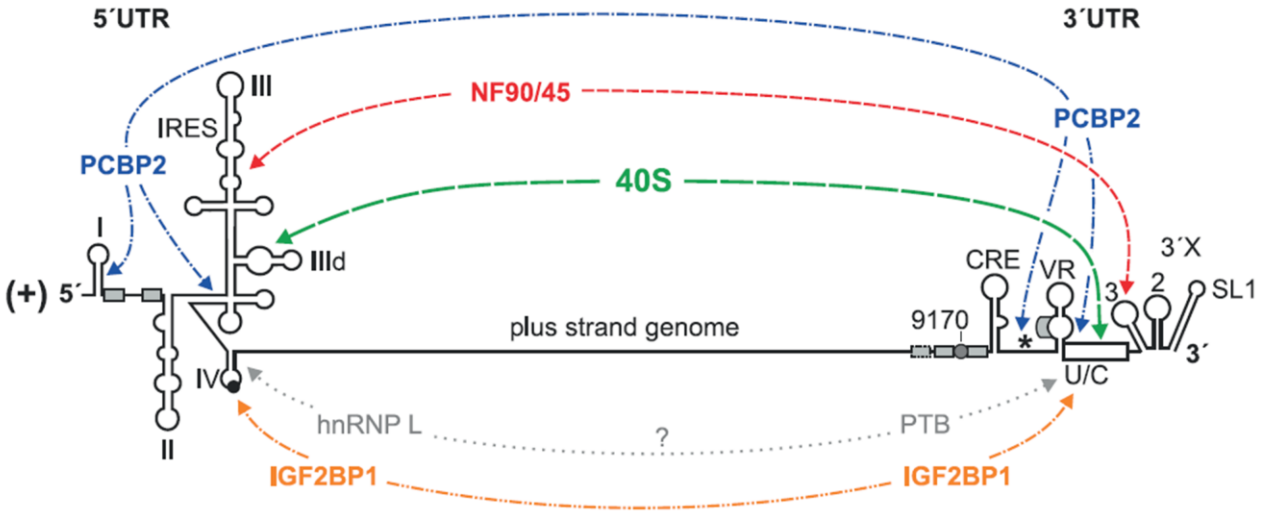


Figure 4.3: Possible long-range interactions of the Hepatitis C Virus genome ends mediated by *trans*-acting factors.
A representation of cellular proteins and the 40S small ribosomal subunit reported to interact with both viral genome ends and hypothesized to facilitate the RNA ends communication and/or HCV genome circularization. The details are given in 4.3.
(from Niepmann et al. 2018)

Accumulation of viral proteins after several rounds of translation is thought to initiate genome rearrangements. The viral NS3 protease acts to cleave the polyprotein and to disintegrate the translation-stimulating protein bridges (Fontanes et al. 2009; Wang et al. 2011). The latter is assisted by miR-122

displacing PCBP2 from the viral genome and freeing the 3'-end for purposes of replication complex formation (Masaki et al. 2015). Accumulating HCV Core protein displaces the 40S ribosomal subunit (Shimoike et al. 2006) freeing the subdomain IIId for a long-range interaction with the 5BSL3.2(bulge), thereby the 5BSL3.2 is present in an open conformation and restores the HCV genome circularization. This interaction in turn shifts equilibrium between 3'-X-tail conformations to the two stem-loops form with an exposure of the DLS region as well as of the three overhang nucleotides at the 3'-terminus of the HCV genome - the form that strongly favors RNA synthesis initiation (Romero-López et al. 2014). Originating from an interaction between inverted repeats within the DLS motifs, the homodimeric form of the 3'UTR was shown to be a preferential template for the NS5B replicase activity (Masante et al. 2015). Interestingly, biophysical studies indicate that in the absence of other interaction partners, the 5BSL3.2(loop)-SL2 is a preferred contact, in contrast to genomes' dimerization (Cantero-Camacho and Gallego 2015). Therefore, maintenance of the actively replicating genome form requires a structural disruption of the three stem-loops conformation and stabilization of the two-stem loops form through RNA-RNA and RNA-protein interactions. Ultimately, an efficient switch to the latter has an impact on the genotype-dependent virulence (Tuplin et al. 2012). Nonetheless, at this step interaction of the 5BSL3.2(bulge) with either SL IIId or SL9110 are thermodynamically equal and alternate during antigenome synthesis, which occurs from dimerized or isolated genomic RNA templates. Under conditions of gradually accumulating plus strands, the two-stem loop conformation of the 3'-end is even further promoted and stabilized through formation of genome dimers interacting via the DLS. The following events additionally shift a balance during active replication. The 5BSL3.2(bulge)-SLIIId interaction itself modifies the three-dimensional organization of the HCV IRES altering its functionality; therefore an impairment of the 40S and eIF3 recruitment leads to a translationally-repressed state (Collier et al. 2002; Romero-López et al. 2012). Furthermore, the apical loop of the 5BSL3.2 recruits the viral RdRp (Zhang et al. 2005) that again interferes with the interaction 5BSL3.2(loop)-SL2 and most importantly places the RNA polymerase in a close proximity to the 3'-overhang.

Undoubtedly, the recapitulated mechanism of a switch between translation and replication is an exceptionally entangled process, which is triggered by few key elements, but regulated and stabilized by a wide range of factors. Following early rounds of translation and replication sites' formation, it can only be speculated how many rounds of translation/replication each individual viral RNA undergoes before the equilibrium is shifted towards one of the molecular processes. Apparently, the whole range of different functional conformational states can be found in the cell at a selected time-point and even subtle changes in local concentration of a certain product can promote a change of the genome's function. As a substantial part of this work, an initiation of the HCV minus strand synthesis was dissected from translation and largely – from plus strand synthesis. Additionally, replication of the constructs containing various fragments of the HCV 5'UTR - and ultimately the complete 5'UTR followed by the Core-coding sequence – was investigated. The constructs including the HCV 5'UTR sequence SL I-III and further, although permitting both genomic and antigenomic strands production, shed light on the complex interplay between the functional elements balancing HCV RNA and protein synthesis.

An assessment of various *cis*-acting replication elements in the context of the SL I-II construct enabled narrowing down the HCV RNA synthesis requirements independently of translation driven from the HCV IRES, as discussed in detail above. Notably, an ability of this construct for minus strand and even weak plus strand production (Fig. 3.6.1, A and B) is in agreement with the knowledge that replication of Flaviviruses does not require ongoing translation when non-structural viral proteins are provided (Westaway et al. 1999). However, a number of reports underline apparently non-overlapping, but overall positive, impact of translation on replication and vice versa. For instance, replication stimulation by miR-122 was shown to be dependent on active protein translation (Masaki et al. 2015). On the other hand, IRES-mediated translation initiation at the HCV 5'UTR is promoted in presence of the 3'UTR (Song et al. 2006). As a matter of fact, the mutual dependence of the HCV translation and replication most likely occurs as a result of engagement of the same set of *cis*-elements and/or host factors and from using the overall virus production as an

experimental readout. The latter represents the major limitation of the HCV full-length or replicon systems that was attempted to overcome in the present study.

Our results demonstrate a substantial improvement of minus strand synthesis efficiency as well as establishment of robust plus strand synthesis solely upon addition of the SL III downstream the SL I-II sequence at the constructs' 5'-end (Fig. 3.2.3, C and D). The HCV SL I-III sequence at the 5'UTR represents an incomplete IRES and alone is not sufficient for functional translation initiation. In turn, SL III plays an essential role in the 40S subunit binding at the HCV 5'UTR (Spahn et al. 2001) as well as in binding of the translation initiation factor eIF3 (Kieft et al. 2001). Importantly, the affinity of eIF3 is largely enhanced when preceded by the 40S recruitment (Siridechadilok et al. 2005). Therefore, the high replication rates displayed by the SL I-III construct in our study (Fig. 3.6.1, A and B) most likely originate from increased ability for the genome ends to communicate via protein bridges. Indeed, the experimental constructs SL I-II and SL I-III harbor identical *cis*-replication elements at the genome's 3'-end, but the effector function of those may differ drastically depending on a dimensional conformation of the RNA template. Although the SL I-II construct provides sufficient sequences and structures for RNA synthesis initiation at the 5'- and 3'-ends, the protein binding platform at its 5'-end is largely reduced in comparison to the SL I-III construct. Such a strong requirement specifically for the SL III presence suggests the role of several cellular proteins in protein-mediated bridging of the RNA ends. As already mentioned above, the 40S small ribosomal subunit was found to bind in a non-overlapping manner to both 5'- and 3'- genome ends via SL III and poly(U/C) tract, respectively. In a situation when translation from the HCV IRES is permitted, such interaction would serve circularization of the RNA template for purposes of efficient translation. Conversely, in context of the translation-incompetent SL I-III construct the closed RNA form can facilitate the long-range interactions that are beneficial for replication, however omitting a competition with translation for the same template. The PCBP2 is another protein that can bind to both HCV genomic ends and potentially promote replication via a protein bridge formation. Notably, PCBP2 is characterized by two binding platforms at each of the genome ends; within the HCV 5'UTR these are the SL I and the SL III pseudoknot (Spangberg and Schwartz 1999; Wang et al. 2011). The first of these two platforms at the very 5'-end may mediate a circularization of the SL I-II replication template, although less efficiently as for the SL I-III construct. Thus, the PCBP2-mediated circularization of the SL I-II construct may contribute to replication rates observed for this RNA. Importantly, as for the 40S ribosomal subunit, the PCBP2 protein is known to stimulate translation in non-reduced systems, however in our experimental setup they appear beneficial for RNA synthesis due to an absence of competition for a template between translation and replication. Lastly, in presence of SL I-III sequence at the construct's 5'-end, the cellular RNA chaperones NF90/NF45 (Isken et al. 2007) can facilitate genomic RNA circularization via binding to the SL III pseudoknot and to SL3 at the 5'- and 3'- genome ends, respectively. The presence of the 5BSL3.2 was demonstrated to be required for replication stimulation by these chaperons (Schmidt et al. 2017) that additionally testifies for a role of protein bridges in the genome ends' approximation for purposes of establishment of beneficial long-range RNA-RNA interactions. The key LRI occurring between the 5'- and 3'- end in circularization is the 5BSL3.2(bulge)-SLIIId interaction. This interaction, established in the full-length HCV genome after displacement of the translating ribosome, is thought to be involved in genome's reshaping that represses translation and favors minus strand synthesis initiation.

Despite the fact that all described RNA-protein and RNA-RNA interactions are possible in context of the 5'UTR construct, it displays a seriously impaired ability for minus and plus strand synthesis (Fig. 3.3.3, C and D; Fig. 3.6.1, A and B). Besides, overall lower translation efficiency was observed at later time-points for this RNA template, which reflects the reduced amount of available templates for protein synthesis due to inefficient replication (Fig. 3.3.3, B; Fig. 3.6.2, C). The nature of these differences between the SL I-III and 5'UTR constructs lays in enabling of functional translation initiation from the HCV IRES. This translation virtually does not result in any useful protein product and the ORF sequence (36 nt) does not contribute with any vital RNA regulatory elements. Furthermore, when the vital elements contained within the Core-coding sequence and functions fulfilled by the Core protein itself are provided by the 5'UTR-Core replication

construct, the efficiency of RNA synthesis remains virtually the same (3.6.1, A and B). The comparison of the replication rates from the SL I-III and 5'UTR/5'UTR-Core constructs evidently demonstrates a repressive effect of translation from HCV IRES on antigenome and genome RNA synthesis. Apparently, the above-mentioned RNA-protein interactions, e.g. with the 40S subunit or the PCBP2, in presence of a complete HCV IRES followed by an ORF prioritize HCV translation over replication. The cellular and viral determinants essential for RNA synthesis initiation are now insufficient to override active translation, and the overall rates of replication decrease due to the need to compete for an RNA template.

In an attempt to shift a balance from the HCV translation to replication, two principal sites within the HCV IRES were targeted by mutational analysis. The three point mutations within the apical loop of the subdomain IIId (IIId*) had been found crucial for both long-range interaction with the 5BSL3.2(bulge) (Shetty et al. 2013) and for efficient binding of the 40S ribosomal subunit (Kieft et al. 2001; Ji et al. 2004). A deletion of the subdomain IIIb (ΔIIIb) was demonstrated to be sufficient for abrogation of the eIF3 binding, importantly without affecting the 40S subunit recruitment (Kieft et al. 2001; Ji et al. 2004). Despite an incomplete reduction of the affinity of the 40S binding by the IIId* mutation; it subsequently impairs the eIF3 attachment (Siridechadilok et al. 2005) that makes translation from the mutated IRES highly improbable. Strikingly, a deterioration of translation initiation by prevention of the eIF3 binding to the HCV IRES led to the most pronounced increase in replication rates of the 5'UTR_ΔIIIb construct (Fig. 3.3.3, C and D). A smaller positive effect acquired by the IIId* mutation is in agreement with an importance of the IIId loop for RNA-RNA communication. Indeed, for the SL I-III construct, whose replication is not slowed down by functional translation, this mutation resulted in a significant reduction of minus strand synthesis (Fig. 3.3.3, C and D), suggesting both the 40S subunit and LRIs to play an important, but not essential, role in RNA synthesis. Nevertheless, despite a negative impact of abrogation of these interactions in the context of the 5'UTR construct, disallowance of active translation overrides the converse effect and results in elevated replication.

An alternative explanation for the reduction of replication efficiency after inclusion of additional annotated 5'UTR sequences may include events taking place at the 3'-end of the antigenome minus strand. Inclusion of more sequences upstream of the 3'-end of the minus strand, compared to those sequences in the minus strand 3'-end provided by the SL I-III construct, may result in inhibition of initiation of plus strand synthesis at the minus strand 3'-end. However, in our opinion this explanation appears less likely, in particular since very distinct mutations (deletion of the SL IIIb sequences or mutation of only three nucleotides in the SL IIId) have profound effects on enhancing minus strand synthesis in the 5'UTR construct but not in the SL I-III construct.

The next degree of complexity to this interplay was added with the full-length Core-coding sequence included downstream the HCV IRES. The Core protein itself is known to shift equilibrium from translation to replication via a negative feedback mechanism: accumulating Core protein signals for sufficient protein synthesis by competing with the 40S subunit and eventually displacing it and promoting the *cis*-elements contacts and genome dimensional rearrangements in favor of RNA replication. Moreover, the Core-coding sequence contains *cis*-elements (Fig. 4.1) with various impact on the HCV replication (McMullan et al. 2007; Vassilaki et al. 2008), as discussed above. The similarly decreased minus strand synthesis efficiency displayed by the 5'UTR-Core, as for the 5'UTR constructs, indicates that the beneficial function of the Core-coding sequence and the protein itself are not sufficient to override the inhibiting effect of translation initiated from the HCV IRES. The mutational analysis in context of the 5'UTR-Core construct yet remains to be performed in follow-up experiments. Functions of additional *cis*-elements within the Core-coding sequence on antigenome synthesis and overall replication is to be evaluated in the absence of suppressive effect from functional translation initiation. Ultimately, effects of double-mutants IIId*/ΔIIIb should be assessed in comparison to single mutated variants for all developed SL III-containing constructs in order to challenge a contribution of the subdomain IIId when translation from the HCV IRES is restricted.

In conclusion, uncoupling of RNA and protein synthesis in scope of the designed experimental system allowed isolating the requirements for the HCV replication. It was evidently demonstrated upon extension of the 5'-end sequence – up to the full-length HCV IRES followed by an open reading frame – that a functional translation initiation event inhibits minus strand synthesis initiation at the 3'-end, likely due to corresponding genome rearrangements and template limitations. Indeed, the molecular processes of the HCV translation and antigenome RNA synthesis use the same genome RNA template for initiation. In order to avoid collusions of the HCV RNA polymerase with translating ribosomes, the regulatory RNA elements with assistance of *trans*-acting proteins reshape the genome to a conformation favoring one of the processes and repressing another. Both efficient HCV translation and replication appear to rely on genome template circularization mediated however by distinct RNA-RNA and RNA-protein interactions. Since the subsequent addition of the HCV 5'UTR sequences results in robust plus strand replication, future attempts to dissect this process from minus strand production in the presence and absence of active translation should be made. This would imply perspective dissection of requirements for the HCV genome synthesis independently of the antigenome synthesis, which is a general challenge for plus strand RNA viruses.

4.4 Conclusions and open questions

In scope of the present work vital aspects of the HCV minus strand synthesis uncoupled from plus strand synthesis and HCV translation were revisited using the minus strand replication system. This system attempts to overcome limitations of the full-length genome or replicon systems containing both genome ends and therefore being unable to assign the overlapping effects of *cis*-elements to distinct steps of RNA replication. Up to date the special focus on the *cis*-elements engaged in minus strand RNA synthesis provided evidence that sequences at the 5'-end of the HCV genome are required for minus strand synthesis initiation at the 3'-end. This process, when uncoupled from other steps of the HCV life cycle, appears to depend on liver-specific miR-122 and on several previously annotated RNA regulatory elements within the NS5B-coding region as well as on certain yet unknown *cis*-elements in the HCV polyprotein coding sequence. Enabling of functional translation initiation from the HCV IRES was demonstrated to have a strong suppressive effect on antigenome synthesis. Although this research has provided an insight into major debatable aspects of the HCV minus strand synthesis regulation, it raises a number of succeeding questions.

Among the questions that are to be answered by subsequent experiments is a more detailed investigation into the 5'UTR sequence requirements for efficient HCV minus strand synthesis, for which a comprehensive mutational analysis can be conducted. Moreover, this direction of the study would indirectly assist a determination of the requirements for a follow-up plus strand synthesis initiation, displayed by some of the experimental constructs. In this context the constructs characterized by robust continuous replication are ultimate targets for mutagenesis. Thereby, the separate analysis of minus and plus strand synthesis may be improved by a further increase in specificity and sensitivity, in particular of minus strand detection. Further insight into protein-mediated regulation of observed competition between translation initiation and RNA synthesis needs to be acquired, perhaps by knockdown approach to confirm the roles of potential *trans*-acting factors in this balance. The role of the proteins functioning on the HCV RNA in genome circularization is also a vital point for further investigation. Despite the fact that mutational analysis within essential *cis*-replication elements does not always yield results due to distraction of vital RNA-RNA interactions, targeting of the host factors acting on these elements may provide a solution. Roles of under-investigated cellular determinants of the HCV minus and plus strand synthesis, such as miRNAs, long non-coding RNAs and RNA-binding proteins, in particular tissue-specific host factors, can be addressed in long-term studies by the system developed in this work. Furthermore, the system offers a potential for *in situ* localization of replication complexes. Any defects upon application of reduced genome RNA template and/or provision of replication determinants in *trans* can be identified using a fluorescent Spinach aptamer. Controversial aspects of the HCV-induced autophagy and subsequent suppression of host immune response and their effects on virus replication can be addressed in cell culture using the reduced, but robustly replicating RNA constructs. Ultimately a panel of replication constructs addressing distinct points of RNA

synthesis is to be established to make its concept easily adjustable for the analysis in other plus strand RNA virus genomes.

5. References

Ali N, Pruijn GJM, Kenan DJ, Keene JD, Siddiqui A. Human La antigen is required for the hepatitis C virus internal ribosome entry site (IRES)-mediated translation. *J. Biol. Chem.* 2000; 275, 27531–27540. doi: 10.1074/jbc.M001487200.

Altmann M, Linder P. Power of Yeast for Analysis of Eukaryotic Translation Initiation. *J. Biol. Chem.* 2010; 285:31907–31912. doi: 10.1074/jbc.R110.144196.

Alvarez DE, Lodeiro MF, Ludueña SJ, Pietrasanta LI, Gamarnik AV. Long-Range RNA-RNA Interactions Circularize the Dengue Virus Genome. *J Virol*. 2005;79(11):6631-6643. doi: 10.1128/JVI.79.11.6631-6643.2005.

Andre P, Perlemuter G, Budkowska A, Brechot C, Lotteau V. Hepatitis C virus particles and lipoprotein metabolism. *Semin Liver Dis.* 2005; 25(1): 93-104. doi: 10.1055/s-2005-864785.

Andreev DE, Hirnet J, Terenin IM, Dmitriev SE, Niepmann M, Shatsky IN. Glycyl-tRNA synthetase specifically binds to the poliovirus IRES to activate translation initiation. *Nucleic Acids Res*. 2012;40(12):5602-5614. doi: 10.1093/nar/gks182.

Appel N, Herian U, Bartenschlager R. Efficient Rescue of Hepatitis C Virus RNA Replication by *trans*-Complementation with Non-structural Protein 5A. *J Virol*. 2005;79(2):896-909. doi: 10.1128/JVI.79.2.896-909.2005.

Astier-Gin T, Bellecave P, Litvak S, Ventura M. Template requirements and binding of hepatitis C virus NS5B polymerase during *in vitro* RNA synthesis from the 3'-end of virus minus strand RNA. *FEBS J*. 2005 Aug;272(15):3872-86. doi: 10.1111/j.1742-4658.2005.04804.x.

Bai Y, Zhou K, Doudna JA. Hepatitis C virus 3'UTR regulates viral translation through direct interactions with the host translation machinery. *Nucleic Acids Res*. 2013;41(16):7861-7874. doi: 10.1093/nar/gkt543.

Bartenschlager R, Cosset FL, Lohmann V. Hepatitis C virus replication cycle. *J Hepatol*. 2010 Sep;53(3):583-5. doi: 10.1016/j.jhep.2010.04.015.

Bartenschlager R, Lohmann V, Penin F. The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. *Nat Rev Microbiol.* 2013 Jul;11(7):482-96. doi: 10.1038/nrmicro3046.

Bartosch B, Dubuisson J, Cosset FL. Infectious Hepatitis C Virus Pseudo-particles Containing Functional E1–E2 Envelope Protein Complexes. *J Exp Med*. 2003a;197(5):633-642. doi: 10.1084/jem.20021756.

Bartosch B, Vitelli A, Granier C, et al. Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *J Biol Chem.* 2003b Oct 24;278(43):41624-30. doi: 10.1074/jbc.M305289200.

Beerens N, Kjems J. Circularization of the HIV-1 genome facilitates strand transfer during reverse transcription. *RNA*. 2010;16(6):1226-1235. doi: 10.1261/rna.2039610.

Belsham GJ, Jackson RJ. Translation initiation on picornavirus RNA. In N. Sonenberg, J. W. B. Hershey, and M. B. Mathews (ed.), Translational control of gene expression. 2000; p. 869–900.Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Beran RKF, Lindenbach BD, Pyle AM. The NS4A Protein of Hepatitis C Virus Promotes RNA-Coupled ATP Hydrolysis by the NS3 Helicase. *J Virol*. 2009;83(7):3268-3275. doi: 10.1128/JVI.01849-08.

Berry KE, Waghray S, Doudna JA. The HCV IRES pseudoknot positions the initiation codon on the 40S ribosomal subunit. *RNA*. 2010;16(8):1559-1569. doi: 10.1261/rna.2197210.

Binder M, Quinkert D, Bochkarova O, et al. Identification of Determinants Involved in Initiation of Hepatitis C Virus RNA Synthesis by Using Intergenotypic Replicase Chimeras. *J Virol*. 2007;81(10):5270-5283. doi: 10.1128/JVI.00032-07.

Blanchard E, Belouzard S, Goueslain L, et al. Hepatitis C Virus Entry Depends on Clathrin-Mediated Endocytosis. *J Virol*. 2006;80(14):6964-6972. doi: 10.1128/JVI.00024-06.

Blight KJ, Rice CM. Secondary structure determination of the conserved 98-base sequence at the 3' terminus of hepatitis C virus genome RNA. *J Virol*. 1997;71(10):7345-7352.

Blight KJ, Kolykhalov AA, Rice CM. Efficient initiation of HCV RNA replication in cell culture. *Science.* 2000 Dec 8;290(5498):1972-4.

Blight KJ, McKeating JA, Rice CM. Highly Permissive Cell Lines for Subgenomic and Genomic Hepatitis C Virus RNA Replication. *J Virol*. 2002;76(24):13001-13014. doi: 10.1128/JVI.76.24.13001-13014.2002.

Boehringer D, Thermann R, Ostareck-Lederer A, Lewis JD, Stark H. Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. *Structure.* 2005 Nov;13(11):1695-706. doi: 10.1016/j.str.2005.08.008.

Branch AD, Stump DD, Gutierrez JA, Eng F, Walewski JL. The hepatitis C virus alternate reading frame (ARF) and its family of novel products: the alternate reading frame protein/F-protein, the double-frameshift protein, and others. *Semin Liver Dis.* 2005 Feb;25(1):105-17. doi: 10.1055/s-2005-864786.

Buck CB, Shen X, Egan MA, Pierson TC, Walker CM, Siliciano RF. The Human Immunodeficiency Virus Type 1 *gag* Gene Encodes an Internal Ribosome Entry Site. *J Virol*. 2001;75(1):181-191. doi: 10.1128/JVI.75.1.181-191.2001.

Bukh J. The history of hepatitis C virus (HCV): Basic research reveals unique features in phylogeny, evolution and the viral life cycle with new perspectives for epidemic control. *J Hepatol.* 2016 Oct;65(1 Suppl):S2-S21. doi: 10.1016/j.jhep.2016.07.035.

Buskirk AR, Green R. Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Philos Trans R Soc Lond B Biol Sci*. 2017;372(1716):20160183. doi: 10.1098/rstb.2016.0183.

Cantero-Camacho Á, Gallego J. The conserved 3'X terminal domain of hepatitis C virus genomic RNA forms a two-stem structure that promotes viral RNA dimerization. *Nucleic Acids Res*. 2015;43(17):8529-8539. doi: 10.1093/nar/gkv786.

Catanese MT, Dorner M. Advances in experimental systems to study hepatitis C virus *in vitro* and *in vivo*. *Virology.* 2015 May;479-480:221-33. doi: 10.1016/j.virol.2015.03.014.

Chatel-Chaix L, Bartenschlager R. Dengue Virus- and Hepatitis C Virus-Induced Replication and Assembly Compartments: the Enemy Inside—Caught in the Web. *J Virol*. 2014;88(11):5907-5911. doi: 10.1128/JVI.03404-13.

Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science.* 1989;244:359–362.

Choo QL, Richman KH, Han JH, et al. Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci U S A*. 1991;88(6):2451-2455.

Chu D, Ren S, Hu S, et al. Systematic Analysis of Enhancer and Critical *cis*-Acting RNA Elements in the Protein-Encoding Region of the Hepatitis C Virus Genome. *J Virol*. 2013;87(10):5678-5696. doi: 10.1128/JVI.00840-12.

Collier AJ, Gallego J, Klinck R, et al. A conserved RNA structure within the HCV IRES eIF3-binding site. *Nat Struct Biol.* 2002 May;9(5):375-80. doi: 10.1038/nsb785.

Conrad KD, Giering F, Erfurth C, et al. Correction: microRNA-122 Dependent Binding of Ago2 Protein to Hepatitis C Virus RNA Is Associated with Enhanced RNA Stability and Translation Stimulation. *PLoS One*. 2016;11(7):e0160132. doi: 10.1371/journal.pone.0160132.

Diviney S, Tuplin A, Struthers M, et al. A Hepatitis C Virus *cis*-Acting Replication Element Forms a Long-Range RNA-RNA Interaction with Upstream RNA Sequences in NS5B. *J Virol*. 2008;82(18):9008-9022. doi: 10.1128/JVI.02326-07.

Dmitriev SE, Terenin IM, Andreev DE, et al. GTP-independent tRNA Delivery to the Ribosomal P-site by a Novel Eukaryotic Translation Factor. *J Biol Chem*. 2010;285(35):26779-26787. doi: 10.1074/jbc.M110.119693.

Dubuisson J, Cosset FL. Virology and cell biology of the hepatitis C virus life cycle: an update. *J Hepatol.* 2014 Nov;61(1 Suppl):S3-S13. doi: 10.1016/j.jhep.2014.06.031.

Edgil D, Harris E. End-to-end communication in the modulation of translation by mammalian RNA viruses. *Virus Res.* 2006 Jul;119(1):43-51. doi: 10.1016/j.virusres.2005.10.012.

Egger D, Wölk B, Gosert R, et al. Expression of Hepatitis C Virus Proteins Induces Distinct Membrane Alterations Including a Candidate Viral Replication Complex. *J Virol*. 2002;76(12):5974-5984. doi: 10.1128/JVI.76.12.5974-5984.2002.

Fabian MR, Sonenberg N. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nat Struct Mol Biol.* 2012 Jun 5;19(6):586-93. doi: 10.1038/nsmb.2296.

Feinstone SM, Kapikian AZ, Purcell RH, Alter HJ, Holland PV. Transfusion-associated hepatitis not due to viral hepatitis type A or B. *N Engl J Med.* 1975 Apr 10;292(15):767-70.

Filbin ME, Kieft JS. HCV IRES domain IIb affects the configuration of coding RNA in the 40S subunit's decoding groove. *RNA*. 2011;17(7):1258-1273. doi: 10.1261/rna.2594011.

Filbin ME, Vollmar BS, Shi D, Gonen T, Kieft JS. HCV IRES manipulates the ribosome to promote the switch from translation initiation to elongation. *Nat Struct Mol Biol*. 2013;20(2):150-158. doi: 10.1038/nsmb.2465.

Fontanes V, Raychaudhuri S, Dasgupta A. A cell permeable peptide inhibits Hepatitis C Virus Replication by Sequestering IRES Transacting Factors. *Virology*. 2009;394(1):82-90. doi: 10.1016/j.virol.2009.08.012.

Fraser CS, Doudna JA. Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat Rev Microbiol.* 2007 Jan;5(1):29-38. doi: 10.1038/nrmicro1558.

Fricke M, Dünnes N, Zayas M, Bartenschlager R, Niepmann M, Marz M. Conserved RNA secondary structures and long-range interactions in hepatitis C viruses. *RNA*. 2015;21(7):1219-1232. doi: 10.1261/rna.049338.114.

Fricke M, Marz M. Prediction of conserved long-range RNA-RNA interactions in full viral genomes. *Bioinformatics.* 2016 Oct 1;32(19):2928-35. doi: 10.1093/bioinformatics/btw323.

Friebe P, Lohmann V, Krieger N, Bartenschlager R. Sequences in the 5' Nontranslated Region of Hepatitis C Virus Required for RNA Replication. *J Virol*. 2001;75(24):12047-12057. doi: 10.1128/JVI.75.24.12047-12057.2001.

Friebe P, Bartenschlager R. Genetic Analysis of Sequences in the 3' Nontranslated Region of Hepatitis C Virus That Are Important for RNA Replication. *J Virol*. 2002;76(11):5326-5338. doi: 10.1128/JVI.76.11.5326-5338.2002.

Friebe P, Boudet J, Simorre J-P, Bartenschlager R. Kissing-Loop Interaction in the 3' End of the Hepatitis C Virus Genome Essential for RNA Replication. *J Virol*. 2005;79(1):380-392. doi: 10.1128/JVI.79.1.380-392.2005.

Friebe P, Bartenschlager R. Role of RNA Structures in Genome Terminal Sequences of the Hepatitis C Virus for Replication and Assembly. *J Virol*. 2009;83(22):11989-11995. doi: 10.1128/JVI.01508-09.

Fuchs G, Petrov AN, Marceau CD, et al. Kinetic pathway of 40S ribosomal subunit recruitment to hepatitis C virus internal ribosome entry site. *Proc Natl Acad Sci U S A*. 2015;112(2):319-325. doi: 10.1073/pnas.1421328111.

Gerold G, Pietschmann T. The HCV life cycle: *in vitro* tissue culture systems and therapeutic targets. *Dig Dis.* 2014;32(5):525-37. doi: 10.1159/000360830.

Gerresheim GK, Dünnes N, Nieder-Röhrmann A, et al. microRNA-122 target sites in the hepatitis C virus RNA NS5B coding region and 3' untranslated region: function in replication and influence of RNA secondary structure. *Cell Mol Life Sci*. 2017 Feb;74(4):747-760. doi: 10.1007/s00018-016-2377-9.

Glass MJ, Summers DF. A *cis*-acting element within the hepatitis A virus 5'-non-coding region required for *in vitro* translation. *Virus Res.* 1992 Oct;26(1):15-31.

Glažar P, Papavasileiou P, Rajewsky N. circBase: a database for circular RNAs. *RNA*. 2014;20(11):1666-70. doi: 10.1261/rna.043687.

Gontarek RR, Gutshall LL, Herold KM, et al. hnRNP C and polypyrimidine tract-binding protein specifically interact with the pyrimidine-rich region within the 3'NTR of the HCV RNA genome. *Nucleic Acids Res*. 1999;27(6):1457-1463.

Gosert R, Egger D, Lohmann V, et al. Identification of the Hepatitis C Virus RNA Replication Complex in Huh-7 Cells Harboring Subgenomic Replicons. *J Virol*. 2003;77(9):5487-5492. doi: 10.1128/JVI.77.9.5487-5492.2003.

Gottwein JM, Scheel TK, Jensen TB, et al. Development and characterization of hepatitis C virus genotype 1-7 cell culture systems: role of CD81 and scavenger receptor class B type I and effect of antiviral drugs. *Hepatology*. 2009 Feb;49(2):364-77. doi: 10.1002/hep.22673.

Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 1999; 41:95-98.

Hansen TB, Jensen TI, Clausen BH, et al. Natural RNA circles function as efficient microRNA sponges. *Nature*. 2013 Mar 21;495(7441):384-8. doi: 10.1038/nature11993.

Harrus D, Ahmed-El-Sayed N, Simister PC, et al. Further Insights into the Roles of GTP and the C Terminus of the Hepatitis C Virus Polymerase in the Initiation of RNA Synthesis. *J Biol Chem*. 2010;285(43):32906-32918. doi: 10.1074/jbc.M110.151316.

Henke JI, Goergen D, Zheng J, et al. microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J*. 2008;27(24):3300-3310. doi: 10.1038/emboj.2008.244.

Herold J, Andino R. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol Cell*. 2001 Mar;7(3):581-91.

Herod MR, Schregel V, Hinds C, Liu M, McLauchlan J, McCormick CJ. Genetic Complementation of Hepatitis C Virus Non-structural Protein Functions Associated with Replication Exhibits Requirements That Differ from Those for Virion Assembly. *J Virol*. 2014;88(5):2748-2762. doi: 10.1128/JVI.03588-13.

Honda M, Rijnbrand R, Abell G, Kim D, Lemon SM. Natural Variation in Translational Activities of the 5' Nontranslated RNAs of Hepatitis C Virus Genotypes 1a and 1b: Evidence for a Long-Range RNA-RNA Interaction outside of the Internal Ribosomal Entry Site. *J Virol*. 1999;73(6):4941-4951.

Hopcraft SE, Azarm KD, Israelow B, et al. Viral Determinants of miR-122-Independent Hepatitis C Virus Replication. *mSphere*. 2016;1(1):e00009-15. doi: 10.1128/mSphere.00009-15.

Houghton M. The long and winding road leading to the identification of the hepatitis C virus. *J Hepatol.* 2009 Nov;51(5):939-48. doi: 10.1016/j.jhep.2009.08.004.

Isken O, Baroth M, Grassmann CW, et al. Nuclear factors are involved in hepatitis C virus RNA replication. *RNA.* 2007;13(10):1675-1692. doi: 10.1261/rna.594207.

Ito T, Tahara SM, Lai MMC. The 3'-Untranslated Region of Hepatitis C Virus RNA Enhances Translation from an Internal Ribosomal Entry Site. *J Virol*. 1998;72(11):8789-8796.

Ivanyi-Nagy R, Kanevsky I, Gabus C, et al. Analysis of hepatitis C virus RNA dimerization and core–RNA interactions. *Nucleic Acids Res*. 2006;34(9):2618-2633. doi: 10.1093/nar/gkl240.

Jackson RJ, Hellen CUT, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol*. 2010;11(2):113-127. doi: 10.1038/nrm2838.

Jang SK, Kräusslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *J Virol*. 1988;62(8):2636-2643.

Jangra RK, Yi M, Lemon SM. Regulation of Hepatitis C Virus Translation and Infectious Virus Production by the MicroRNA miR-122. *J Virol*. 2010;84(13):6615-6625. doi: 10.1128/JVI.00417-10.

Janssen HL, Reesink HW, Lawitz EJ, et al. Treatment of HCV infection by targeting microRNA. *N Engl J Med.* 2013 May 2;368(18):1685-94. doi: 10.1056/NEJMoa1209026.

Ji H, Fraser CS, Yu Y, Leary J, Doudna JA. Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc Natl Acad Sci U S A*. 2004;101(49):16990-16995. doi: 10.1073/pnas.0407402101.

Jin Z, Leveque V, Ma H, Johnson KA, Klumpp K. Assembly, Purification, and Pre-steady-state Kinetic Analysis of Active RNA-dependent RNA Polymerase Elongation Complex. *J Biol Chem.* 2012;287(13):10674-10683. doi: 10.1074/jbc.M111.325530.

Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science.* 2005 Sep 2;309(5740):1577-81. doi: 10.1126/science.1113329.

Jopling CL, Schütz S, Sarnow P. Position-dependent Function for a Tandem MicroRNA miR-122 Binding Site Located in the Hepatitis C Virus RNA Genome. *Cell Host Microbe*. 2008;4(1):77-85. doi: 10.1016/j.chom.2008.05.013.

Jopling C. Liver-specific microRNA-122: Biogenesis and function. *RNA Biol*. 2012;9(2):137-142. doi: 10.4161/rna.18827.

Joseph AP, Bhat P, Das S, Srinivasan N. Re-analysis of cryoEM data on HCV IRES bound to 40S subunit of human ribosome integrated with recent structural information suggests new contact regions between ribosomal proteins and HCV RNA. *RNA Biol*. 2014;11(7):891-905. doi: 10.4161/rna.29545.

Jost I, Shalamova LA, Gerresheim GK, Niepmann M, Bindereif A, Rossbach O. Functional sequestration of microRNA-122 from Hepatitis C Virus by circular RNA sponges. *RNA Biol.* 2018;15(8):1032-1039. doi: 10.1080/15476286.2018.1435248.

Kapoor A, Simmonds P, Gerold G, et al. Characterization of a canine homolog of hepatitis C virus. *Proc Natl Acad Sci U S A.* 2011;108(28):11608-11613. doi: 10.1073/pnas.1101794108.

Kato T, Furusaka A, Miyamoto M, et al. Sequence analysis of hepatitis C virus isolated from a fulminant hepatitis patient. *J Med Virol.* 2001 Jul;64(3):334-9.

Kato T, Date T, Miyamoto M, et al. Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon. *Gastroenterology.* 2003 Dec;125(6):1808-17.

Kazakov T, Yang F, Ramanathan HN, Kohlway A, Diamond MS, Lindenbach BD. Hepatitis C Virus RNA Replication Depends on Specific *Cis*- and *Trans*-Acting Activities of Viral Non-structural Proteins. *PLoS Pathog*. 2015;11(4):e1004817. doi: 10.1371/journal.ppat.1004817.

Kerr CH, Jan E. Commandeering the Ribosome: Lessons Learned from Dicistroviruses about Translation. *J Virol*. 2016;90(12):5538-5540. doi: 10.1128/JVI.00737-15.

Keum SJ, Park SM, Park JH, Jung JH, Shin EJ, Jang SK. The specific infectivity of hepatitis C virus changes through its life cycle. *Virology.* 2012 Nov 25;433(2):462-70. doi: 10.1016/j.virol.2012.08.046.

Khawaja A, Vopalensky V, Pospisek M. Understanding the potential of hepatitis C virus internal ribosome entry site domains to modulate translation initiation via their structure and function. *Wiley Interdiscip Rev RNA*. 2015;6(2):211-224. doi: 10.1002/wrna.1268.

Khromykh AA, Meka H, Guyatt KJ, Westaway EG. Essential Role of Cyclization Sequences in Flavivirus RNA Replication. *J Virol*. 2001;75(14):6719-6728. doi: 10.1128/JVI.75.14.6719-6728.2001.

Kieft JS, Zhou K, Jubin R, Murray MG, Lau JY, Doudna JA. The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J Mol Biol.* 1999 Sep 24;292(3):513-29. doi: 10.1006/jmbi.1999.3095.

Kieft JS, Zhou K, Jubin R, Doudna JA. Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA*. 2001;7(2):194-206.

Kieft JS. Viral IRES RNA structures and ribosome interactions. *Trends Biochem Sci*. 2008;33(6):274-283. doi: 10.1016/j.tibs.2008.04.007.

Kikuchi K, Umehara T, Fukuda K, Kuno A, Hasegawa T, Nishikawa S. A hepatitis C virus (HCV) internal ribosome entry site (IRES) domain III–IV-targeted aptamer inhibits translation by binding to an apical loop of domain IIId. *Nucleic Acids Res*. 2005;33(2):683-692. doi: 10.1093/nar/gki215.

Kim JH, Hahm B, Kim YK, Choi M, Jang SK. Protein-protein interaction among hnRNPs shuttling between nucleus and cytoplasm. *J Mol Biol.* 2000 May 5;298(3):395-405. doi: 10.1006/jmbi.2000.3687.

Kim JH, Park SM, Park JH, Keum SJ, Jang SK. eIF2A mediates translation of hepatitis C viral mRNA under stress conditions. *EMBO J*. 2011;30(12):2454-2464. doi: 10.1038/emboj.2011.146.

Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol.* 2009 Feb;10(2):126-39. doi: 10.1038/nrm2632.

Kim YK, Kim CS, Lee SH, Jang SK. Domains I and II in the 5' nontranslated region of the HCV genome are required for RNA replication. *Biochem Biophys Res Commun.* 2002 Jan 11;290(1):105-12. doi: 10.1006/bbrc.2001.6167.

Kim YK, Lee SH, Kim CS, Seol SK, Jang SK. Long-range RNA–RNA interaction between the 5’ nontranslated region and the core-coding sequences of hepatitis C virus modulates the IRES-dependent translation. *RNA*. 2003;9(5):599-606. doi: 10.1261/rna.2185603.

Kincaid RP, Sullivan CS. Virus-Encoded microRNAs: An Overview and a Look to the Future. *PLoS Pathog.* 2012;8(12):e1003018. doi: 10.1371/journal.ppat.1003018.

Kolykhalov AA, Agapov EV, Blight KJ, Mihalik K, Feinstone SM, Rice CM. Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. *Science.* 1997 Jul 25;277(5325):570-4.

Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics.* 2007 May 15;23(10):1289-91. doi: 10.1093/bioinformatics/btm091.

Kozak M. An analysis of 5’-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res*. 1987;15(20):8125-8148.

Kühn R, Luz N, Beck E. Functional analysis of the internal translation initiation site of foot-and-mouth disease virus. *J Virol*. 1990;64(10):4625-4631.

Kuo G, Choo QL, Alter HJ, et al. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science.* 1989 Apr 21;244(4902):362-4.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970 Aug 15;227(5259):680-5.

Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T. Identification of tissue-specific microRNAs from mouse. *Curr Biol.* 2002 Apr 30;12(9):735-9.

Lanford RE, Hildebrandt-Eriksen ES, Petri A, et al. Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science*. 2010;327(5962):198-201. doi: 10.1126/science.1178178.

Lee H, Shin H, Wimmer E, Paul AV. *cis*-Acting RNA Signals in the NS5B C-Terminal Coding Sequence of the Hepatitis C Virus Genome. *J Virol*. 2004;78(20):10865-10877. doi: 10.1128/JVI.78.20.10865-10877.2004.

Li Y, Masaki T, Yamane D, McGivern DR, Lemon SM. Competing and noncompeting activities of miR-122 and the 5' exonuclease Xrn1 in regulation of hepatitis C virus replication. *Proc Natl Acad Sci U S A*. 2013;110(5):1881-1886. doi: 10.1073/pnas.1213515110.

Li Y, Masaki T, Shimakami T, Lemon SM. hnRNP L and NF90 Interact with Hepatitis C Virus 5'-Terminal Untranslated RNA and Promote Efficient Replication. *J Virol*. 2014;88(13):7199-7209. doi: 10.1128/JVI.00225-14.

Li Y, Yamane D, Masaki T, Lemon SM. The Yin and Yang of Hepatitis C: Synthesis and Decay of HCV RNA. *Nat Rev Microbiol*. 2015;13(9):544-558. doi: 10.1038/nrmicro3506.

Lindenbach BD, Evans MJ, Syder AJ, et al. Complete replication of hepatitis C virus in cell culture. *Science.* 2005 Jul 22;309(5734):623-6. doi: 10.1126/science.1114016.

Lindenbach BD, Meuleman P, Ploss A, et al. Cell culture-grown hepatitis C virus is infectious *in vivo* and can be recultured *in vitro*. *Proc Natl Acad Sci U S A*. 2006;103(10):3805-3809. doi: 10.1073/pnas.0511218103.

Lindenbach BD. Virion assembly and release. Hepatitis C Virus: From Molecular Virology to Antiviral Therapy. *Curr Top Microbiol Immunol.* 2013;369:199-218. doi: 10.1007/978-3-642-27340-7_8.

Lindenbach BD, Rice CM. The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol*. 2013;11(10):688-700. doi: 10.1038/nrmicro3098.

Lohmann V, Körner F, Herian U, Bartenschlager R. Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J Virol*. 1997;71(11):8416-8428.

Lohmann V, Körner F, Koch J, Herian U, Theilmann L, Bartenschlager R. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science.* 1999 Jul 2;285(5424):110-3.

Lohmann V, Körner F, Dobierzewska A, Bartenschlager R. Mutations in Hepatitis C Virus RNAs Conferring Cell Culture Adaptation. *J Virol*. 2001;75(3):1437-1449. doi: 10.1128/JVI.75.3.1437-1449.2001.

Lohmann V, Hoffmann S, Herian U, Penin F, Bartenschlager R. Viral and Cellular Determinants of Hepatitis C Virus RNA Replication in Cell Culture. *J Virol*. 2003;77(5):3007-3019. doi: 10.1128/JVI.77.5.3007-3019.2003.

Lohmann V. Hepatitis C Virus RNA Replication. Hepatitis C Virus: From Molecular Virology to Antiviral Therapy. *Curr Top Microbiol Immunol.* 2013;369:167-98. doi: 10.1007/978-3-642-27340-7_7.

Lohmann V, Bartenschlager R. On the history of hepatitis C virus cell culture systems. *J Med Chem.* 2014 Mar 13;57(5):1627-42. doi: 10.1021/jm401401n.

Lourenço S, Costa F, Débarges B, Andrieu T, Cahour A. Hepatitis C virus internal ribosome entry site-mediated translation is stimulated by *cis*-acting RNA elements and *trans*-acting viral factors. *FEBS J.* 2008 Aug;275(16):4179-97. doi: 10.1111/j.1742-4658.2008.06566.x.

Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2012 Dec 15;380(9859):2095-128. doi: 10.1016/S0140-6736(12)61728-0.

Lukavsky PJ, Otto GA, Lancaster AM, Sarnow P, Puglisi JD. Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. *Nat Struct Biol.* 2000 Dec;7(12):1105-10. doi: 10.1038/81951.

Lukavsky PJ, Kim I, Otto GA, Puglisi JD. Structure of HCV IRES domain II determined by NMR. *Nat Struct Biol.* 2003 Dec;10(12):1033-8. doi: 10.1038/nsb1004.

Lukavsky PJ. Structure and function of HCV IRES domains. *Virus Res*. 2009;139(2-2):166-171. doi: 10.1016/j.virusres.2008.06.004.

Luna JM, Scheel TKH, Danino T, et al. Hepatitis C virus RNA functionally sequesters miR-122. *Cell*. 2015;160(6):1099-1110. doi: 10.1016/j.cell.2015.02.025.

Luo G, Hamatake RK, Mathis DM, et al. *De Novo* Initiation of RNA Synthesis by the RNA-Dependent RNA Polymerase (NS5B) of Hepatitis C Virus. *J Virol*. 2000;74(2):851-863.

Maéno M, Kaminaka K, Sugimoto H, et al. A cDNA clone closely associated with non-A, non-B hepatitis. *Nucleic Acids Res.* 1990;18(9):2685-2689.

Mahajan VS, Drake A, Chen J. Virus-specific host miRNAs: antiviral defenses or promoters of persistent infection? *Trends Immunol*. 2009;30(1):1-7. doi: 10.1016/j.it.2008.08.009.

Mahias K, Ahmed-El-Sayed N, Masante C, et al. Identification of a structural element of the hepatitis C virus minus strand RNA involved in the initiation of RNA synthesis. *Nucleic Acids Res*. 2010;38(12):4079-4091. doi: 10.1093/nar/gkq109.

Masaki T, Arend KC, Li Y, et al. miR-122 Stimulates Hepatitis C Virus RNA Synthesis by Altering the Balance of Viral RNAs Engaged in Replication Versus Translation. *Cell Host Microbe*. 2015;17(2):217-228. doi: 10.1016/j.chom.2014.12.014.

Masante C, Mahias K, Lourenço S, et al. Seven nucleotide changes characteristic of the hepatitis C virus genotype 3 5' untranslated region: correlation with reduced *in vitro* replication. *J Gen Virol.* 2008 Jan;89(1):212-21. doi: 10.1099/vir.0.83067-0.

Mauger DM, Golden M, Yamane D, et al. Functionally conserved architecture of hepatitis C virus RNA genomes. *Proc Natl Acad Sci U S A.* 2015;112(12):3692-3697. doi: 10.1073/pnas.1416266112.

Mazumder B, Seshadri V, Fox PL. Translational control by the 3'-UTR: the ends specify the means. *Trends Biochem Sci.* 2003 Feb;28(2):91-8. doi: 10.1016/S0968-0004(03)00002-1.

Memczak S, Jens M, Elefsinioti A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature.* 2013 Mar 21;495(7441):333-8. doi: 10.1038/nature11928.

Merz A, Long G, Hiet M-S, et al. Biochemical and Morphological Properties of Hepatitis C Virus Particles and Determination of Their Lipidome. *J Biol Chem.* 2011;286(4):3018-3032. doi: 10.1074/jbc.M110.175018.

Minor PD. Attenuation and reversion of the Sabin vaccine strains of poliovirus. *Dev Biol Stand.* 1993;78:17-26.

Miyanari Y, Atsuzawa K, Usuda N, et al. The lipid droplet is an important organelle for hepatitis C virus production. *Nat Cell Biol.* 2007 Sep;9(9):1089-97. doi: 10.1038/ncb1631.

Mondal T, Ray U, Manna AK, Gupta R, Roy S, Das S. Structural Determinant of Human La Protein Critical for Internal Initiation of Translation of Hepatitis C Virus RNA. *J Virol*. 2008;82(23):11927-11938. doi: 10.1128/JVI.00924-08.

Moradpour D, Brass V, Bieck E, et al. Membrane Association of the RNA-Dependent RNA Polymerase Is Essential for Hepatitis C Virus RNA Replication. *J Virol*. 2004;78(23):13278-13284. doi: 10.1128/JVI.78.23.13278-13284.2004.

Moradpour D, Penin F, Rice CM. Replication of hepatitis C virus. *Nat Rev Microbiol.* 2007 Jun;5(6):453-63. doi: 10.1038/nrmicro1645.

Moradpour D, Penin F. Hepatitis C Virus Proteins. From Structure to Function. Hepatitis C Virus: From Molecular Virology to Antiviral Therapy. *Curr Top Microbiol Immunol.* 2013;369:113-42. doi: 10.1007/978-3-642-27340-7_5.

Mu X, Greenwald E, Ahmad S, Hur S. An origin of the immunogenicity of *in vitro* transcribed RNA. *Nucleic Acids Res*. 2018;46(10):5239-5249. doi: 10.1093/nar/gky177.

Murphy DG, Sablon E, Chamberland J, Fournier E, Dandavino R, Tremblay CL. Hepatitis C Virus Genotype 7, a New Genotype Originating from Central Africa. *J Clin Microbiol*. 2015;53(3):967-972. doi: 10.1128/JCM.02831-14.

Nasheri N, Singaravelu R, Goodmurphy M, Lyn RK, Pezacki JP. Competing roles of microRNA-122 recognition elements in hepatitis C virus RNA. *Virology.* 2011 Feb 20;410(2):336-44. doi: 10.1016/j.virol.2010.11.015.

Nieder-Röhrmann A, Dünnes N, Gerresheim GK, Shalamova LA, Herchenröther A, Niepmann M. Cooperative enhancement of translation by two adjacent microRNA-122/Argonaute 2 complexes binding to the 5' untranslated region of hepatitis C virus RNA. *J Gen Virol.* 2017 Feb;98(2):212-224. doi: 10.1099/jgv.0.000697.

Niepmann, M. Internal initiation of translation of picornaviruses, hepatitis C virus and pestiviruses. *Recent Res. Dev. Virol.* 1999; 1:229–250.

Niepmann M. Hepatitis C virus RNA translation. From Structure to Function. Hepatitis C Virus: From Molecular Virology to Antiviral Therapy. *Curr Top Microbiol Immunol.* 2013;369:143-66. doi: 10.1007/978-3-642-27340-7_6.

Niepmann M, Shalamova LA, Gerresheim GK, Rossbach O. Signals Involved in Regulation of Hepatitis C Virus RNA Genome Translation and Replication. *Front Microbiol*. 2018;9:395. doi: 10.3389/fmicb.2018.00395.

Norcum MT, Warrington JA. Structural analysis of the multienzyme aminoacyl-tRNA synthetase complex: a three-domain model based on reversible chemical crosslinking. *Protein Sci*. 1998;7(1):79-87.

Oakland TE, Haselton KJ, Randall G. EWSR1 Binds the Hepatitis C Virus *cis*-Acting Replication Element and Is Required for Efficient Viral Replication. *J Virol*. 2013;87(12):6625-6634. doi: 10.1128/JVI.01006-12.

Ochs K, Zeller A, Saleh L, et al. Impaired Binding of Standard Initiation Factors Mediates Poliovirus Translation Attenuation. *J Virol*. 2003;77(1):115-122. doi: 10.1128/JVI.77.1.115-122.2003.

Ojha CR, Rodriguez M, Dever SM, Mukhopadhyay R, El-Hage N. Mammalian microRNA: an important modulator of host-pathogen interactions in human viral infections. *J Biomed Sci*. 2016;23:74. doi: 10.1186/s12929-016-0292-x.

Orr-Burks NL, Shim B-S, Wu W, Bakre AA, Karpilow J, Tripp RA. MicroRNA screening identifies miR-134 as a regulator of poliovirus and enterovirus 71 infection. *Sci Data*. 2017;4:170023. doi: 10.1038/sdata.2017.23.

Otto GA, Puglisi JD. The pathway of HCV IRES-mediated translation initiation. *Cell*. 2004 Oct 29;119(3):369-80. doi: 10.1016/j.cell.2004.09.038.

Ouellet J. RNA Fluorescence with Light-Up Aptamers. *Front Chem*. 2016;4:29. doi: 10.3389/fchem.2016.00029.

Paige JS, Wu K, Jaffrey SR. RNA mimics of green fluorescent protein. *Science*. 2011;333(6042):642-646. doi: 10.1126/science.1207339.

Park SG, Choi EC, Kim S. Aminoacyl-tRNA synthetase-interacting multifunctional proteins (AIMPs): a triad for cellular homeostasis. *IUBMB Life*. 2010 Apr;62(4):296-302. doi: 10.1002/iub.324.

Paul D, Romero-Brey I, Gouttenoire J, et al. NS4B Self-Interaction through Conserved C-Terminal Elements Is Required for the Establishment of Functional Hepatitis C Virus Replication Complexes. *J Virol*. 2011;85(14):6963-6976. doi: 10.1128/JVI.00502-11.

Paul D, Hoppe S, Saher G, Krijnse-Locker J, Bartenschlager R. Morphological and Biochemical Characterization of the Membranous Hepatitis C Virus Replication Compartment. *J Virol*. 2013;87(19):10612-10627. doi: 10.1128/JVI.01370-13.

Pause A, Kukolj G, Bailey M, et al. An NS3 serine protease inhibitor abrogates replication of subgenomic hepatitis C virus RNA. *J Biol Chem*. 2003 May 30;278(22):20374-80. doi: 10.1074/jbc.M210785200.

Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*. 1988 Jul 28;334(6180):320-5. doi: 10.1038/334320a0.

Pestova TV, Kolupaeva VG. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev*. 2002;16(22):2906-2922. doi: 10.1101/gad.1020902.

Pfaffl MW. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res*. 2001;29(9):e45.

Pietschmann T, Lohmann V, Rutter G, Kurpanek K, Bartenschlager R. Characterization of Cell Lines Carrying Self-Replicating Hepatitis C Virus RNAs. *J Virol*. 2001;75(3):1252-1264. doi: 10.1128/JVI.75.3.1252-1264.2001.

Pietschmann T, Kaul A, Koutsoudakis G, et al. Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. *Proc Natl Acad Sci U S A*. 2006;103(19):7408-7413. doi: 10.1073/pnas.0504877103.

Pirakitikulr N, Kohlway A, Lindenbach BD, Pyle AM. The coding region of the HCV genome contains a network of regulatory RNA structures. *Mol Cell*. 2016;62(1):111-120. doi: 10.1016/j.molcel.2016.01.024.

Piwecka M, Glažar P, Hernandez-Miranda LR, et al. Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. *Science*. 2017 Sep 22;357(6357). doi: 10.1126/science.aam8526.

Ploss A, Khetani SR, Jones CT, et al. Persistent hepatitis C virus infection in microscale primary human hepatocyte cultures. *Proc Natl Acad Sci U S A*. 2010;107(7):3141-3145. doi: 10.1073/pnas.0915130107.

Psaridi L, Georgopoulou U, Varaklioti A, Mavromara P. Mutational analysis of a conserved tetraloop in the 5' untranslated region of hepatitis C virus identifies a novel RNA element essential for the internal ribosome entry site function. *FEBS Lett*. 1999 Jun 18;453(1-2):49-53.

Randall G, Panis M, Cooper JD, et al. Cellular cofactors affecting hepatitis C virus infection and replication. *Proc Natl Acad Sci U S A*. 2007;104(31):12884-12889. doi: 10.1073/pnas.0704894104.

Reiss S, Harak C, Romero-Brey I, et al. The Lipid Kinase Phosphatidylinositol-4 Kinase III Alpha Regulates the Phosphorylation Status of Hepatitis C Virus NS5A. *PLoS Pathog*. 2013;9(5):e1003359. doi: 10.1371/journal.ppat.1003359.

Ribeiro RM, Li H, Wang S, et al. Quantifying the Diversification of Hepatitis C Virus (HCV) during Primary Infection: Estimates of the *In Vivo* Mutation Rate. *PLoS Pathog*. 2012;8(8):e1002881. doi: 10.1371/journal.ppat.1002881.

Ríos-Marco P, Romero-López C, Berzal-Herranz A. The *cis*-acting replication element of the Hepatitis C virus genome recruits host factors that influence viral replication and translation. *Sci Rep*. 2016;6:25729. doi: 10.1038/srep25729.

Roberts APE, Lewis AP, Jopling CL. miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components. *Nucleic Acids Res*. 2011;39(17):7716-7729. doi: 10.1093/nar/gkr426.

Robertson B, Myers G, Howard C, et al. Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. International Committee on Virus Taxonomy. *Arch Virol*. 1998;143(12):2493-503.

Romero-Brey I, Merz A, Chiramel A, et al. Three-Dimensional Architecture and Biogenesis of Membrane Structures Associated with Hepatitis C Virus Replication. *PLoS Pathog*. 2012;8(12):e1003056. doi: 10.1371/journal.ppat.1003056.

Romero-López C, Barroso-delJesus A, García-Sacristán A, Briones C, Berzal-Herranz A. The folding of the hepatitis C virus internal ribosome entry site depends on the 3'-end of the viral genome. *Nucleic Acids Res*. 2012;40(22):11697-11713. doi: 10.1093/nar/gks927.

Romero-López C, Barroso-delJesus A, García-Sacristán A, Briones C, Berzal-Herranz A. End-to-end crosstalk within the hepatitis C virus genome mediates the conformational switch of the 3'X-tail region. *Nucleic Acids Res*. 2014;42(1):567-582. doi: 10.1093/nar/gkt841.

Romero-López C, Berzal-Herranz A. The 5BSL3.2 Functional RNA Domain Connects Distant Regions in the Hepatitis C Virus Genome. *Front Microbiol*. 2017;8:2093. doi: 10.3389/fmicb.2017.02093.

Ross-Thriepland D, Harris M. Hepatitis C virus NS5A: enigmatic but still promiscuous 10 years on! *J Gen Virol*. 2015 Apr;96(Pt 4):727-38. doi: 10.1099/jgv.0.000009.

Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol*. 2000;132:365-86.

Saeed M, Andreo U, Chung H-Y, et al. SEC14L2 enables pan-genotype HCV replication in cell culture. *Nature*. 2015;524(7566):471-475. doi: 10.1038/nature14899.

Sagan SM, Nasheri N, Luebbert C, Pezacki JP. The efficacy of siRNAs against hepatitis C virus is strongly influenced by structure and target site accessibility. *Chem Biol*. 2010 May 28;17(5):515-27. doi: 10.1016/j.chembiol.2010.04.011.

Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One*. 2012;7(2):e30733. doi: 10.1371/journal.pone.0030733.

Scheel TKH, Luna JM, Liniger M, et al. A broad RNA virus survey reveals both miRNA dependence and functional sequestration. *Cell Host Microbe*. 2016;19(3):409-423. doi: 10.1016/j.chom.2016.02.007.

Schmidt T, Friedrich S, Golbik RP, Behrens S-E. NF90–NF45 is a selective RNA chaperone that rearranges viral and cellular riboswitches: biochemical analysis of a virus host factor activity. *Nucleic Acids Res*. 2017;45(21):12441-12454. doi: 10.1093/nar/gkx931.

Schuster C, Isel C, Imbert I, Ehresmann C, Marquet R, Kieny MP. Secondary Structure of the 3' Terminus of Hepatitis C Virus Minus strand RNA. *J Virol*. 2002;76(16):8058-8068. doi: 10.1128/JVI.76.16.8058-8068.2002.

Sedano CD, Sarnow P. Hepatitis C virus subverts liver-specific miR-122 to protect the viral genome from exoribonuclease Xrn2. *Cell Host Microbe*. 2014;16(2):257-264. doi: 10.1016/j.chom.2014.07.006.

Shetty S, Kim S, Shimakami T, Lemon SM, Mihailescu M-R. Hepatitis C virus genomic RNA dimerization is mediated via a kissing complex intermediate. *RNA*. 2010;16(5):913-925. doi: 10.1261/rna.1960410.

Shetty S, Stefanovic S, Mihailescu MR. Hepatitis C virus RNA: molecular switches mediated by long-range RNA–RNA interactions? *Nucleic Acids Res*. 2013;41(4):2526-2540. doi: 10.1093/nar/gks1318.

Shi G, Suzuki T. Molecular Basis of Encapsidation of Hepatitis C Virus Genome. *Front Microbiol*. 2018;9:396. doi: 10.3389/fmicb.2018.00396.

Shih IH, Been MD. Catalytic strategies of the hepatitis delta virus ribozymes. *Annu Rev Biochem.* 2002;71:887-917. doi: 10.1146/annurev.biochem.71.110601.135349.

Shimakami T, Yamane D, Jangra RK, et al. Stabilization of hepatitis C virus RNA by an Ago2–miR-122 complex. *Proc Natl Acad Sci U S A*. 2012a;109(3):941-946. doi: 10.1073/pnas.1112263109.

Shimakami T, Yamane D, Welsch C, Hensley L, Jangra RK, Lemon SM. Base Pairing between Hepatitis C Virus RNA and MicroRNA 122 3' of Its Seed Sequence Is Essential for Genome Stabilization and Production of Infectious Virus. *J Virol*. 2012b;86(13):7372-7383. doi: 10.1128/JVI.00513-12.

Shimoike T, Koyama C, Murakami K, et al. Down-regulation of the internal ribosome entry site (IRES)-mediated translation of the hepatitis C virus: critical role of binding of the stem-loop IIId domain of IRES and the viral core protein. *Virology.* 2006 Feb 20;345(2):434-45. doi: 10.1016/j.virol.2005.10.013.

Simister P, Schmitt M, Geitmann M, et al. Structural and Functional Analysis of Hepatitis C Virus Strain JFH1 Polymerase. *J Virol*. 2009;83(22):11926-11939. doi: 10.1128/JVI.01008-09.

Simmonds P, Holmes EC, Cha TA, et al. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J Gen Virol.* 1993 Nov;74 (11):2391-9. doi: 10.1099/0022-1317-74-11-2391.

Simmonds P, Bukh J, Combet C, et al. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology*. 2005 Oct;42(4):962-73. doi: 10.1002/hep.20819.

Siridechadilok B, Fraser CS, Hall RJ, Doudna JA, Nogales E. Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science.* 2005 Dec 2;310(5753):1513-5. doi: 10.1126/science.1118977.

Smith DB, Bukh J, Kuiken C, et al. Expanded Classification of Hepatitis C Virus Into 7 Genotypes and 67 Subtypes: Updated Criteria and Genotype Assignment Web Resource. *Hepatology*. 2014;59(1):318-327. doi: 10.1002/hep.26744.

Smith RM, Walton CM, Wu CH, Wu GY. Secondary Structure and Hybridization Accessibility of Hepatitis C Virus 3'-Terminal Sequences. *J Virol*. 2002;76(19):9563-9574. doi: 10.1128/JVI.76.19.9563-9574.2002.

Song Y, Liu Y, Ward CB, et al. Identification of two functionally redundant RNA elements in the coding sequence of poliovirus using computer-generated design. *Proc Natl Acad Sci U S A.* 2012;109(36):14301-14307. doi: 10.1073/pnas.1211484109.

Spahn CM, Kieft JS, Grassucci RA, et al. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science.* 2001 Mar 9;291(5510):1959-62. doi: 10.1126/science.1058409.

Spangberg K, Goobar-Larsson L, Wahren-Herlenius M, Schwartz S. The La protein from human liver cells interacts specifically with the U-rich region in the hepatitis C virus 3' untranslated region. *J Hum Virol.* 1999 Sep-Oct;2(5):296-307.

Spangberg K, Schwartz S. Poly(C)-binding protein interacts with the hepatitis C virus 5' untranslated region. *J Gen Virol*. 1999 Jun;80 (Pt 6):1371-6. doi: 10.1099/0022-1317-80-6-1371.

Srisawat C, Engelke DR. Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *RNA*. 2001;7(4):632-641.

Steinmann E, Brohm C, Kallis S, Bartenschlager R, Pietschmann T. Efficient *trans*-Encapsidation of Hepatitis C Virus RNAs into Infectious Virus-Like Particles. *J Virol*. 2008;82(14):7034-7046. doi: 10.1128/JVI.00118-08.

Steinmann E, Pietschmann T. Cell culture systems for hepatitis C virus. Hepatitis C Virus: From Molecular Virology to Antiviral Therapy. *Curr Top Microbiol Immunol*. 2013;369:17-48. doi: 10.1007/978-3-642-27340-7_2.

Sumpter R, Loo Y-M, Foy E, et al. Regulating Intracellular Antiviral Defense and Permissiveness to Hepatitis C Virus RNA Replication through a Cellular RNA Helicase, RIG-I. *J Virol*. 2005;79(5):2689-2699. doi: 10.1128/JVI.79.5.2689-2699.2005.

Swayze EE, Siwkowski AM, Wancewicz EV, et al. Antisense oligonucleotides containing locked nucleic acid improve potency but cause significant hepatotoxicity in animals. *Nucleic Acids Res*. 2006;35(2):687-700. doi: 10.1093/nar/gkl1071.

Tabak HF, Van der Horst G, Smit J, Winter AJ, Mul Y, Groot Koerkamp MJ. Discrimination between RNA circles, interlocked RNA circles and lariats using two-dimensional polyacrylamide gel electrophoresis. *Nucleic Acids Res*. 1988;16(14A):6597-605.

Targett-Adams P, Boulant S, McLauchlan J. Visualization of Double-Stranded RNA in Cells Supporting Hepatitis C Virus RNA Replication. *J Virol*. 2008;82(5):2182-2195. doi: 10.1128/JVI.01565-07.

Tarr AW, Khera T, Hueging K, et al. Genetic Diversity Underlying the Envelope Glycoproteins of Hepatitis C Virus: Structural and Functional Consequences and the Implications for Vaccine Design. *Viruses*. 2015;7(7):3995-4046. doi: 10.3390/v7072809.

Terenin IM, Dmitriev SE, Andreev DE, Shatsky IN. Eukaryotic translation initiation machinery can operate in a bacterial-like mode without eIF2. *Nat Struct Mol Biol.* 2008 Aug;15(8):836-41. doi: 10.1038/nsmb.1445.

Thibault PA, Huys A, Amador-Cañizares Y, Gailius JE, Pinel DE, Wilson JA. Regulation of Hepatitis C Virus Genome Replication by Xrn1 and MicroRNA-122 Binding to Individual Sites in the 5’ Untranslated Region. *J Virol*. 2015;89(12):6294-6311. doi: 10.1128/JVI.03631-14.

Tingting P, Caiyun F, Zhigang Y, Pengyuan Y, Zhenghong Y. Subproteomic analysis of the cellular proteins associated with the 3' untranslated region of the hepatitis C virus genome in human liver cells. *Biochem Biophys Res Commun.* 2006 Sep 1;347(3):683-91. doi: 10.1016/j.bbrc.2006.06.144.

Tong L, Lin L, Wu S, et al. MiR-10a* up-regulates coxsackievirus B3 biosynthesis by targeting the 3D-coding sequence. *Nucleic Acids Res*. 2013;41(6):3760-3771. doi: 10.1093/nar/gkt058.

Tsai W-C, Hsu S-D, Hsu C-S, et al. MicroRNA-122 plays a critical role in liver homeostasis and hepatocarcinogenesis. *J Clin Invest*. 2012;122(8):2884-2897. doi: 10.1172/JCI63455.

Tsukiyama-Kohara K, Iizuka N, Kohara M, Nomoto A. Internal ribosome entry site within hepatitis C virus RNA. *J Virol*. 1992;66(3):1476-1483.

Tuplin A, Struthers M, Simmonds P, Evans DJ. A twist in the tail: SHAPE mapping of long-range interactions and structural rearrangements of RNA elements involved in HCV replication. *Nucleic Acids Res*. 2012;40(14):6908-6921. doi: 10.1093/nar/gks370.

Tuplin A, Struthers M, Cook J, Bentley K, Evans DJ. Inhibition of HCV translation by disrupting the structure and interactions of the viral CRE and 3' X-tail. *Nucleic Acids Res*. 2015;43(5):2914-2926. doi: 10.1093/nar/gkv142.

Untergasser A, Cutcutache I, Koressaar T, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012;40(15):e115. doi: 10.1093/nar/gks596.

Upadhyay A, Dixit U, Manvar D, Chaturvedi N, Pandey VN. Affinity Capture and Identification of Host Cell Factors Associated with Hepatitis C Virus (+) Strand Subgenomic RNA. *Mol Cell Proteomics*. 2013;12(6):1539-1552. doi: 10.1074/mcp.M112.017020.

Villordo SM, Gamarnik AV. Genome Cyclization as Strategy for Flavivirus RNA Replication. *Virus Res*. 2009;139(2):230-239. doi: 10.1016/j.virusres.2008.07.016.

Voigts-Hoffmann F, Klinge S, Ban N. Structural insights into eukaryotic ribosomes and the initiation of translation. *Curr Opin Struct Biol.* 2012 Dec;22(6):768-77. doi: 10.1016/j.sbi.2012.07.010.

Wakita T, Pietschmann T, Kato T, et al. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med*. 2005;11(7):791-796. doi: 10.1038/nm1268.

Wang T-H, Rijnbrand RCA, Lemon SM. Core Protein-Coding Sequence, but Not Core Protein, Modulates the Efficiency of Cap-Independent Translation Directed by the Internal Ribosome Entry Site of Hepatitis C Virus. *J Virol*. 2000;74(23):11347-11358.

Wang L, Jeng K-S, Lai MMC. Poly(C)-Binding Protein 2 Interacts with Sequences Required for Viral Replication in the Hepatitis C Virus (HCV) 5' Untranslated Region and Directs HCV RNA Replication through Circularizing the Viral Genome. *J Virol*. 2011;85(16):7954-7964. doi: 10.1128/JVI.00339-11.

Westaway EG, Khromykh AA, Mackenzie JM. Nascent flavivirus RNA colocalized *in situ* with double-stranded RNA in stable replication complexes. *Virology*. 1999 May 25;258(1):108-17. doi: 10.1006/viro.1999.9683.

World Health Organization. Guidelines for the screening, care and treatment of persons with hepatitis C infection. Geneva; Updated version, April 2016. (http://apps.who.int/iris/bitstream/10665/205035/1/9789241549615_eng.pdf?ua=1).

World Health Organization. Guidelines for the care and treatment of persons diagnosed with chronic hepatitis C virus infection. Geneva; April 2018. Licence: CC BY-NC-SA 3.0 IGO. (http://apps.who.int/iris/bitstream/handle/10665/273174/9789241550345-eng.pdf?ua=1).

Wilson JE, Powell MJ, Hoover SE, Sarnow P. Naturally Occurring Dicistronic Cricket Paralysis Virus RNA Is Regulated by Two Internal Ribosome Entry Sites. *Mol Cell Biol*. 2000;20(14):4990-4999.

Wilson JA, Zhang C, Huys A, Richardson CD. Human Ago2 Is Required for Efficient MicroRNA 122 Regulation of Hepatitis C Virus RNA Accumulation and Translation. *J Virol*. 2011;85(5):2342-2350. doi: 10.1128/JVI.02046-10.

Xu Z, Choi J, Yen TSB, et al. Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *EMBO J*. 2001;20(14):3840-3848. doi: 10.1093/emboj/20.14.3840.

Yanagi M, Purcell RH, Emerson SU, Bukh J. Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. *Proc Natl Acad Sci U S A*. 1997;94(16):8738-8743.

Yi M, Lemon SM. 3' Nontranslated RNA Signals Required for Replication of Hepatitis C Virus RNA. *J Virol*. 2003a;77(6):3557-3568. doi: 10.1128/JVI.77.6.3557-3568.2003.

Yi M, Lemon SM. Structure–function analysis of the 3' stem-loop of hepatitis C virus genomic RNA and its role in viral RNA replication. *RNA*. 2003b;9(3):331-345. doi: 10.1261/rna.2144203.

You S, Stump DD, Branch AD, Rice CM. A *cis*-Acting Replication Element in the Sequence Encoding the NS5B RNA-Dependent RNA Polymerase Is Required for Hepatitis C Virus RNA Replication. *J Virol*. 2004;78(3):1352-1366. doi: 10.1128/JVI.78.3.1352-1366.2004.

You S, Rice CM. 3' RNA Elements in Hepatitis C Virus Replication: Kissing Partners and Long Poly(U). *J Virol*. 2008;82(1):184-195. doi: 10.1128/JVI.01796-07.

Yu Y, Scheel TKH, Luna JM, et al. Correction: miRNA independent hepacivirus variants suggest a strong evolutionary pressure to maintain miR-122 dependence. *PLoS Pathog*. 2018;14(9):e1007303. doi: 10.1371/journal.ppat.1007303.

Zhang C, Huys A, Thibault PA, Wilson JA. Requirements for human Dicer and TRBP in microRNA-122 regulation of HCV translation and RNA abundance. *Virology*. 2012 Nov 25;433(2):479-88. doi: 10.1016/j.virol.2012.08.039.

Zhong W, Uss AS, Ferrari E, Lau JYN, Hong Z. *De Novo* Initiation of RNA Synthesis by Hepatitis C Virus Non-structural Protein 5B Polymerase. *J Virol*. 2000;74(4):2017-2022.

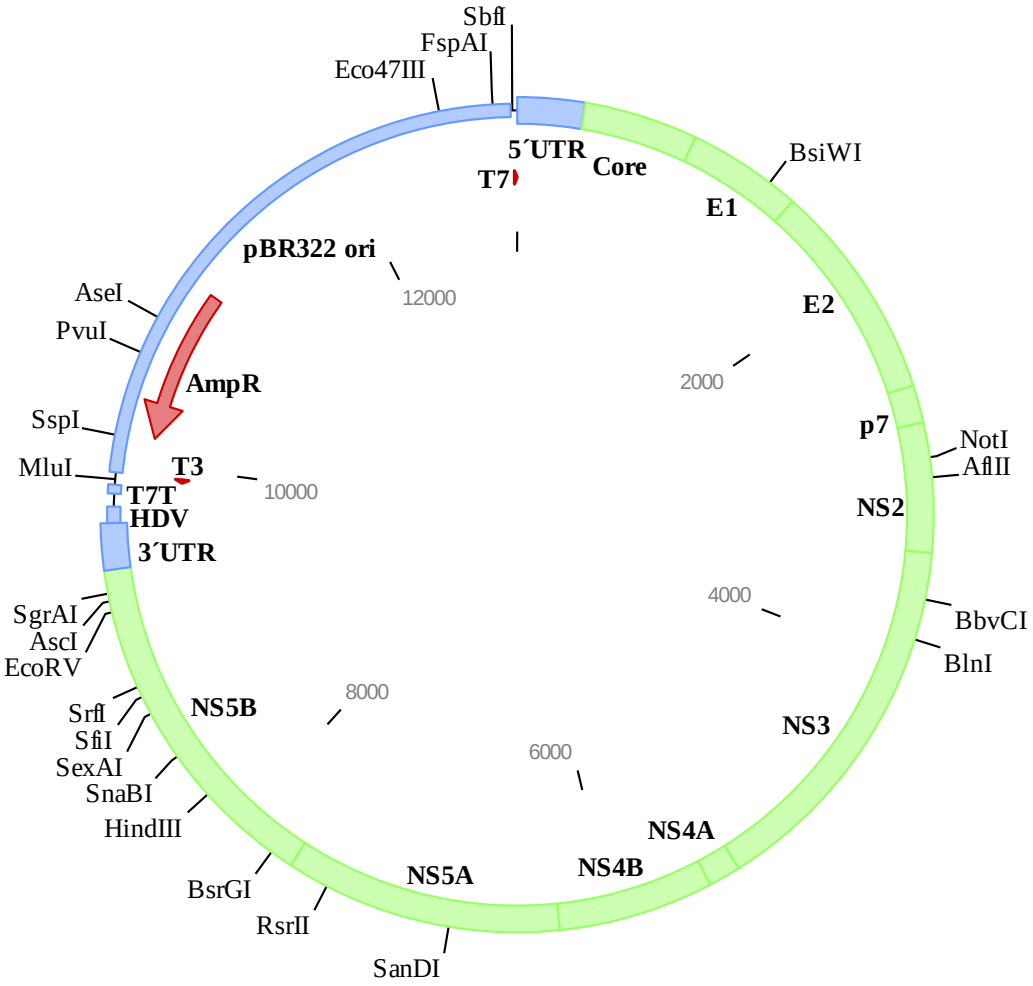
Zhong J, Gastaminza P, Cheng G, et al. Robust hepatitis C virus infection *in vitro*. *Proc Natl Acad Sci U S A*. 2005;102(26):9294-9299. doi: 10.1073/pnas.0503596102.

Zeisel MB, Felmlee DJ, Baumert TF. Hepatitis C virus entry. Hepatitis C Virus: From Molecular Virology to Antiviral Therapy. *Curr Top Microbiol Immunol*. 2013;369:87-112. doi: 10.1007/978-3-642-27340-7_4.

6. Appendix

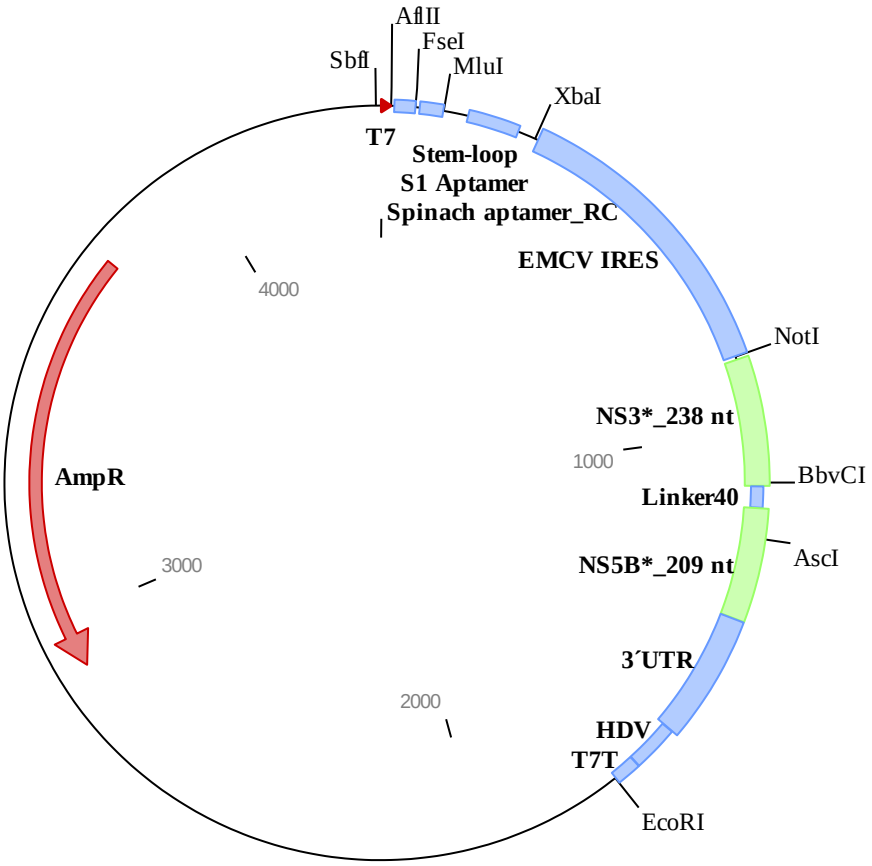
6.1 Plasmid maps

6.1.1 pFK-JFH1-J6 C-846_dg (JC1)_12961



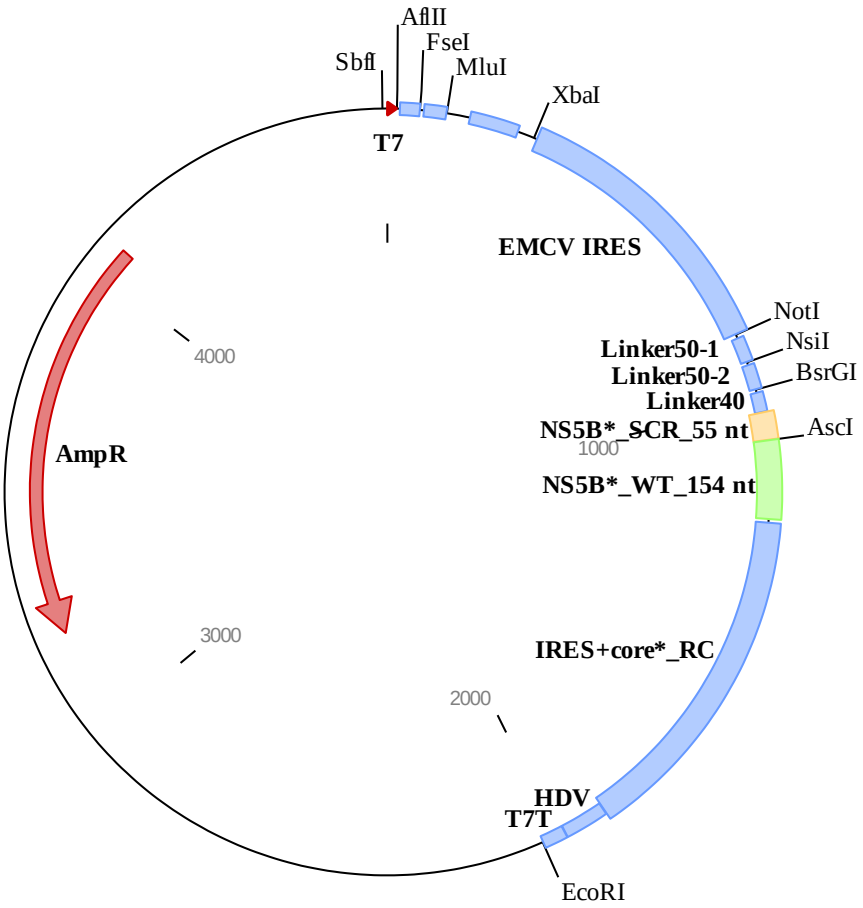
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
5’UTR	HCV 5’UTR untranslated region	18-357
Core	HCV Core-coding sequence	358-930
E1	HCV E1 glycoprotein coding sequence	931-1506
E2	HCV E2 glycoprotein coding sequence	1507-2607
p7	HCV p7 polypeptide coding sequence	2608-2796
NS2	HCV NS2 protein coding sequence	2797-3447
NS3	HCV NS3 protein coding sequence	3448-5340
NS4A	HCV NS4A protein coding sequence	5341-5502
NS4B	HCV NS4B protein coding sequence	5503-6285
NS5A	HCV NS5A protein coding sequence	6286-7683
NS5B	HCV NS5B protein coding sequence	7684-9459
3’UTR	HCV 3’-untranslated region	9460-9695
HDV	Hepatitis D Virus genomic ribozyme	9696-9779
T7T	T7 RNA Polymerase terminator	9845-9891
T3	T3 RNA Polymerase promoter (RC)	9931-9950
pBR322 ori	pBR322 origin of replication	9954-12945
Amp R	Ampicillin resistance gene (RC)	10162-11019

6.1.2 pUC18_Plus_strand_backbone_4374



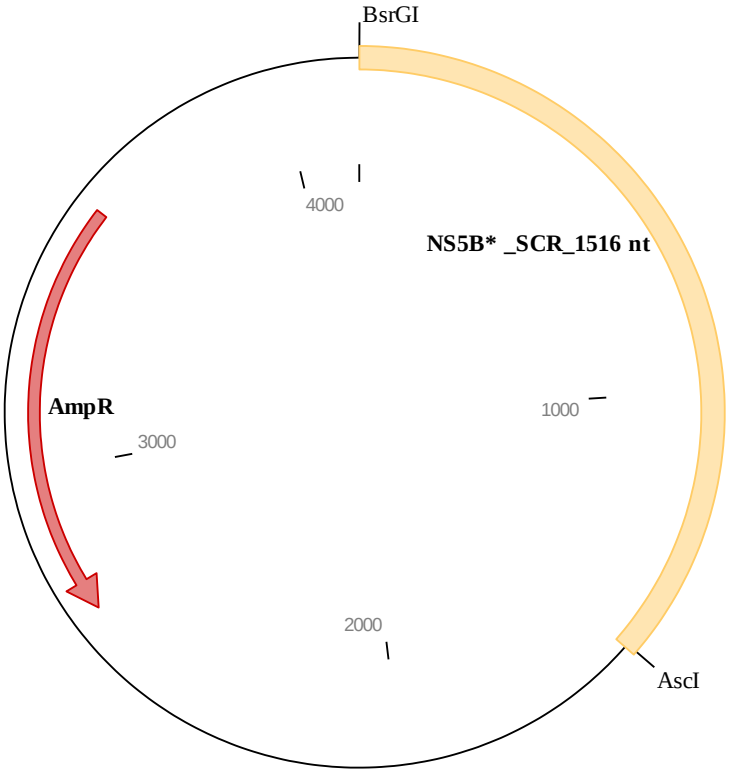
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
NS3*_238 nt	Partial wild-type NS3 sequence (238 nt)	862-1099
Linker40	A random 40 nt linker sequence	1100-1139
NS5B*_209 nt	Partial wild-type NS5B sequence (209 nt)	1140-1348
3'UTR	HCV 3'-untranslated region	1349-1584
HDV	Hepatitis D Virus genomic ribozyme	1585-1671
T7T	T7 RNA Polymerase terminator	1672-1718
Amp R	Ampicillin resistance gene (RC)	2895-3755

6.1.3 pUC18_Minus_strand_backbone_4685



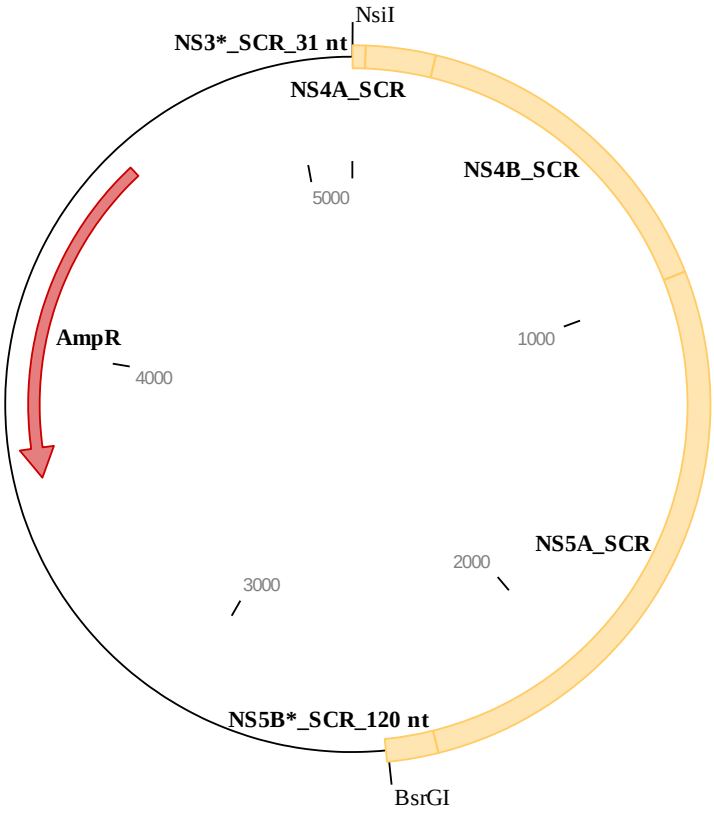
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
Linker50-1	A random 50 nt linker sequence	864-913
Linker50-2	A random 50 nt linker sequence	920-969
Linker40	A random 40 nt linker sequence	976-1015
NS5B*_SCR_55 nt	Partial scrambled NS5B sequence (55 nt)	1016-1070
NS5B*_WT_154 nt	Partial wild-type NS5B sequence (154 nt)	1071-1224
IRES+Core*_RC	Reverse complement of HCV Internal Ribosome Entry Site followed by a partial HCV Core-coding sequence (325 nt)	1231-1895
HDV	Hepatitis D Virus genomic ribozyme	1896-1982
T7T	T7 RNA Polymerase terminator	1983-2029
Amp R	Ampicillin resistance gene (RC)	3206-4066

6.1.4 pUC18_Fragment 1_NS5B_SCR_4157



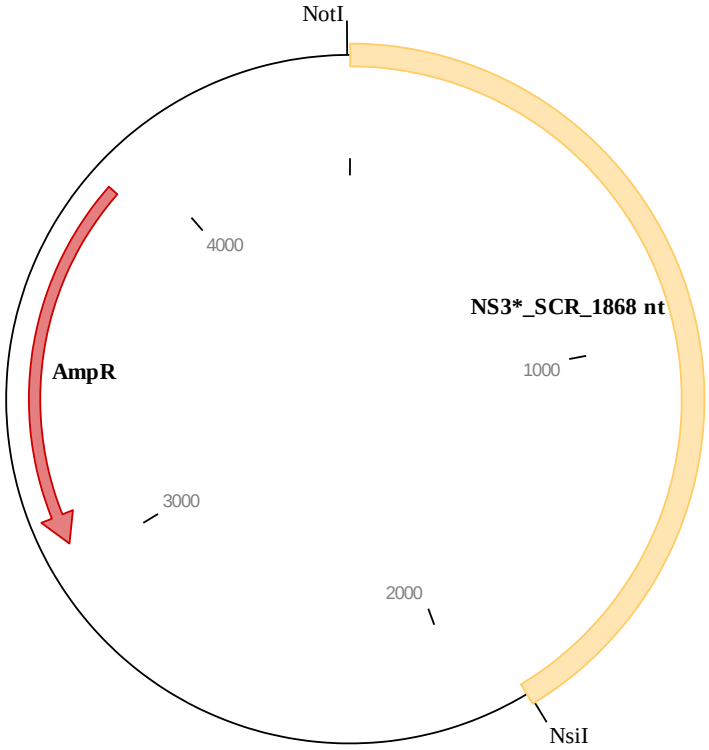
Feature	Description	Start – Stop (nt)
NS5B*_SCR_1516 nt	Partial scrambled NS5B sequence (1516 nt)	1-1516
Amp R	Ampicillin resistance gene (RC)	2693-3553

6.1.5 pUC18_Fragment 2_NS3-NS5B_SCR_5148



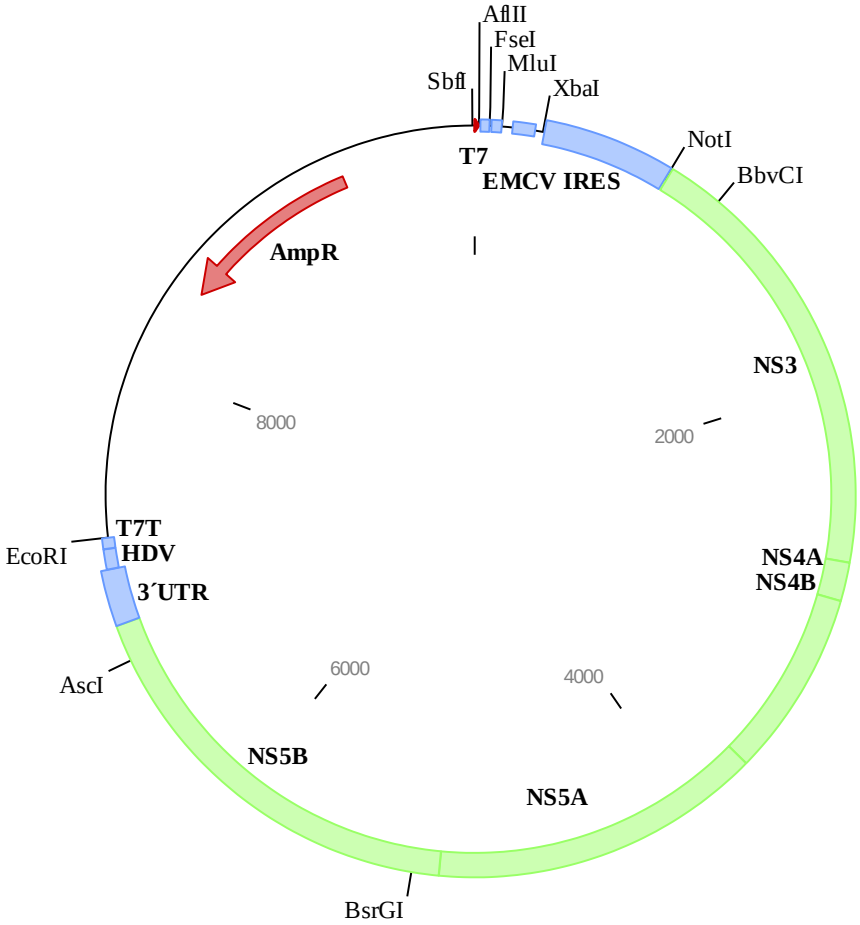
Feature	Description	Start – Stop (nt)
NS3*_SCR_31 nt	Partial scrambled NS3 sequence (31 nt)	1-31
NS4A_SCR	Scrambled NS4A coding sequence	32-193
NS4B_SCR	Scrambled NS54B coding sequence	194-976
NS5A_SCR	Scrambled NS5A coding sequence	977-2374
NS5B*_SCR_120 nt	Partial scrambled NS5B sequence (120 nt)	2375-2494
Amp R	Ampicillin resistance gene (RC)	3671-4531

6.1.6 pUC18_Fragment 3_NS3_SCR_4515



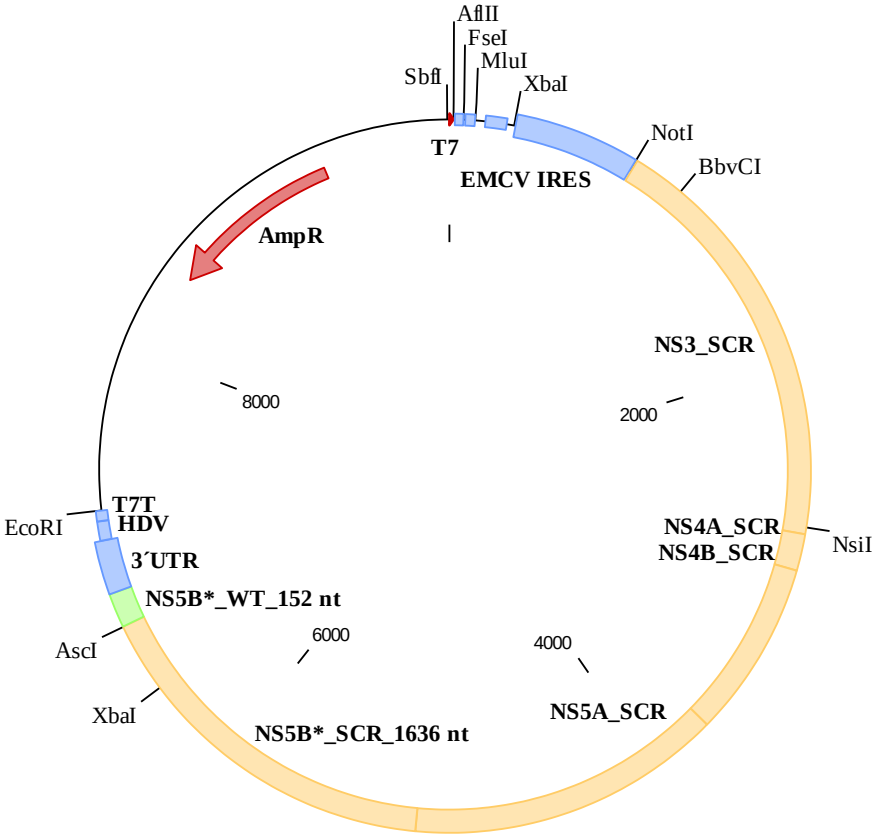
Feature	Description	Start – Stop (nt)
NS3*_SCR_1868 nt	Partial scrambled NS3 sequence (1868 nt)	1-1868
Amp R	Ampicillin resistance gene (RC)	3045-3905

6.1.7 pUC18_P.s_WT_hp_9899



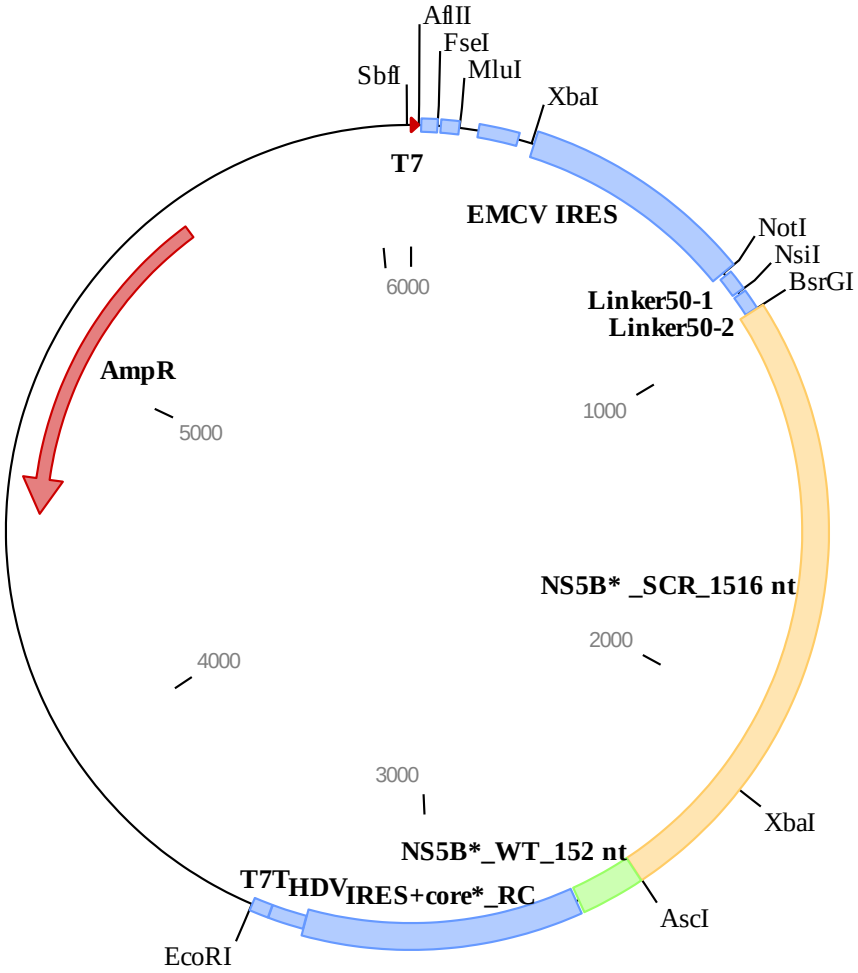
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
NS3	Wild-type NS3 coding sequence	862-2754
NS4A	Wild-type NS4A coding sequence	2755-2916
NS4B	Wild-type NS4B coding sequence	2917-3699
NS5A	Wild-type NS5A coding sequence	3700-5097
NS5B	Wild-type NS5B coding sequence	5098-6873
3’UTR	HCV 3’-untranslated region	6874-7109
HDV	Hepatitis D Virus genomic ribozyme	7110-7196
T7T	T7 RNA Polymerase terminator	7197-7243
Amp R	Ampicillin resistance gene (RC)	8420-9280

6.1.8 pUC18_P.s _SCR_hp_9899



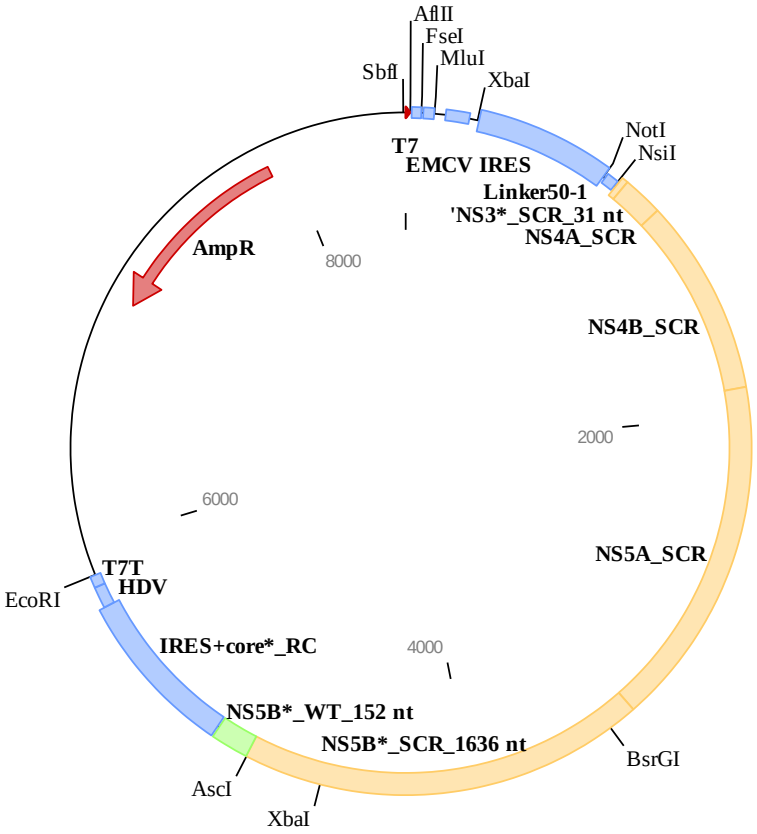
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
NS3_SCR	Scrambled NS3 coding sequence	862-2754
NS4A_SCR	Scrambled NS4A coding sequence	2755-2916
NS4B_SCR	Scrambled NS4B coding sequence	2917-3699
NS5A_SCR	Scrambled NS5A coding sequence	3700-5097
NS5B*_SCR_1636 nt	Partial scrambled NS5B sequence (1636 nt)	5098-6721
NS5B*_WT_152 nt	Partial wild-type NS5B sequence (152 nt)	6722-6873
3’UTR	HCV 3’-untranslated region	6874-7109
HDV	Hepatitis D Virus genomic ribozyme	7110-7196
T7T	T7 RNA Polymerase terminator	7197-7243
Amp R	Ampicillin resistance gene (RC)	8420-9280

6.1.9 pUC18_Minus_strand_F1_6092



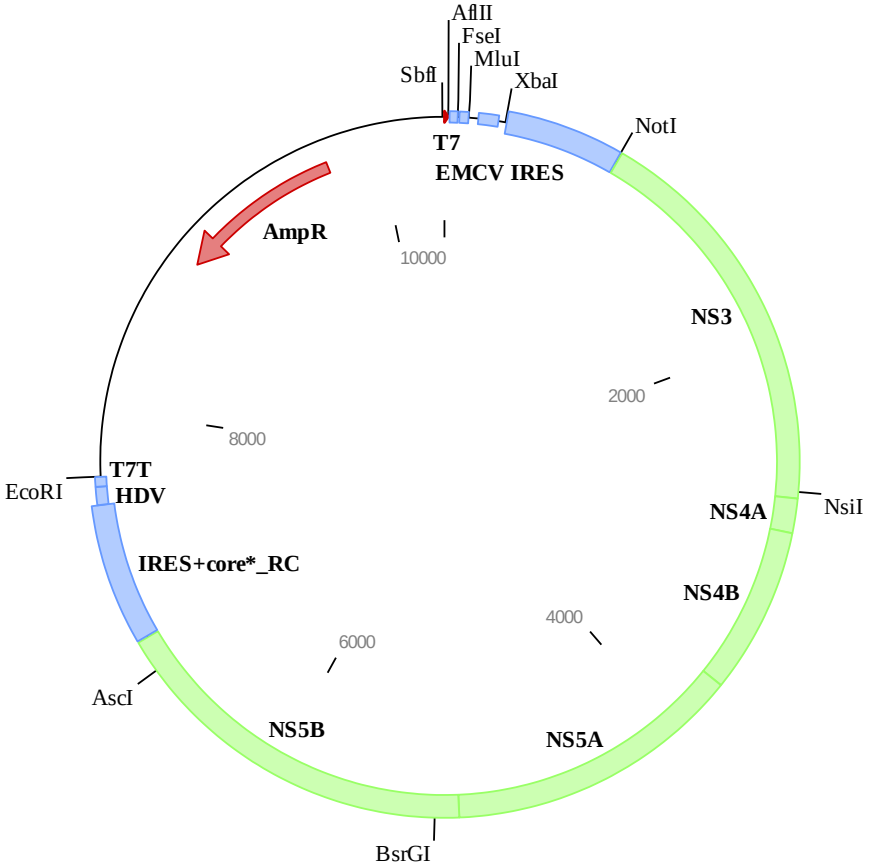
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
Linker50-1	A random 50 nt linker sequence	864-913
Linker50-2	A random 50 nt linker sequence	920-969
NS5B*_SCR_1516 nt	Partial scrambled NS5B sequence (1516 nt)	971-2479
NS5B*_WT_152 nt	Partial wild-type NS5B sequence (152 nt)	2480-2631
IRES+Core*_RC	Reverse complement of HCV Internal Ribosome Entry Site followed by a partial HCV Core-coding sequence (325 nt)	2638-3302
HDV	Hepatitis D Virus genomic ribozyme	3303-3389
T7T	T7 RNA Polymerase terminator	3390-3436
Amp R	Ampicillin resistance gene (RC)	5515-5473

6.1.10 pUC18_Minus_strand_F1_F2_8524



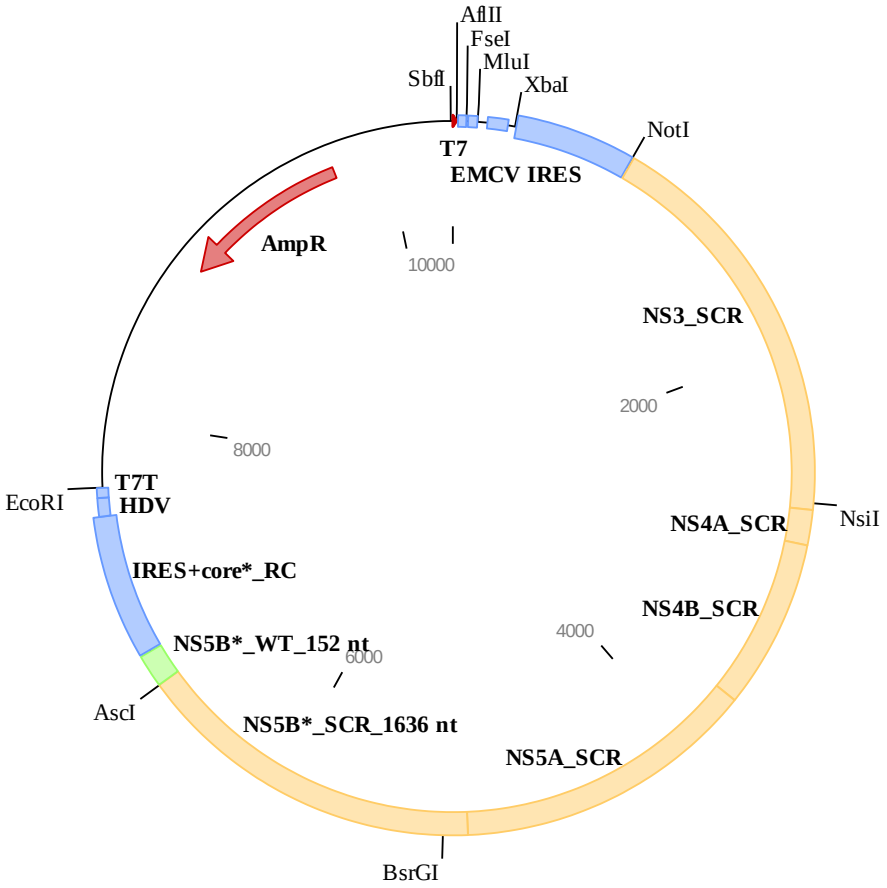
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
Linker50-1	A random 50 nt linker sequence	864-913
NS3*_SCR_31 nt	Partial scrambled NS3 sequence (31 nt)	919-944
NS4A_SCR	Scrambled NS4A coding sequence	945-1106
NS4B_SCR	Scrambled NS54B coding sequence	1107-1889
NS5A_SCR	Scrambled NS5A coding sequence	1890-3287
NS5B*_SCR_1636 nt	Partial scrambled NS5B sequence (1516 nt)	3288-4911
NS5B*_WT_152 nt	Partial wild-type NS5B sequence (152 nt)	4912-5063
IRES+Core*_RC	Reverse complement of HCV Internal Ribosome Entry Site followed by a partial HCV Core-coding sequence (325 nt)	5070-5734
HDV	Hepatitis D Virus genomic ribozyme	5735-5821
T7T	T7 RNA Polymerase terminator	5822-5868
Amp R	Ampicillin resistance gene (RC)	7045-7905

6.1.11 pUC18_M.s_WT_hp_10334



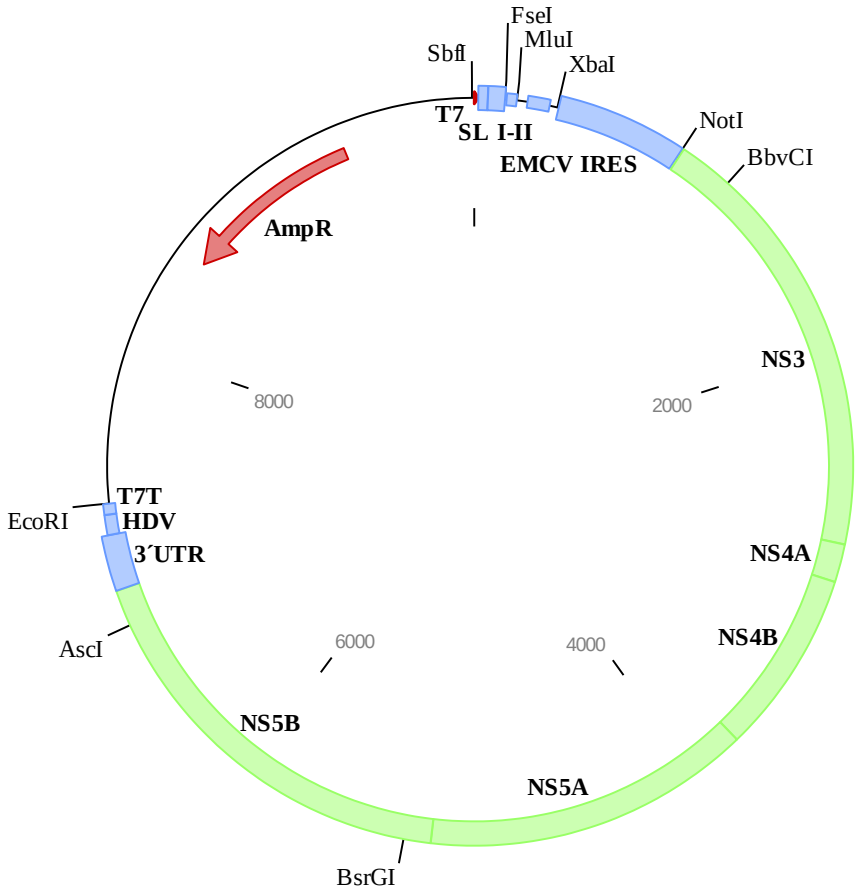
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
NS3	Wild-type NS3 coding sequence	862-2754
NS4A	Wild-type NS4A coding sequence	2755-2916
NS4B	Wild-type NS4B coding sequence	2917-3699
NS5A	Wild-type NS5A coding sequence	3700-5097
NS5B	Wild-type NS5B coding sequence	5098-6873
IRES+Core*_RC	Reverse complement of HCV Internal Ribosome Entry Site followed by a partial HCV Core-coding sequence (325 nt)	6880-7544
HDV	Hepatitis D Virus genomic ribozyme	7545-7631
T7T	T7 RNA Polymerase terminator	7632-7678
Amp R	Ampicillin resistance gene (RC)	8855-9715

6.1.12 pUC18_M.s_SCR_hp_10334



Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
Stem-loop	40 nt GC-rich stem-loop sequence	25-64
S1 Aptamer	S1 Streptavidin aptamer sequence	72-116
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	163-260
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	300-855
NS3_SCR	Scrambled NS3 coding sequence	862-2754
NS4A_SCR	Scrambled NS4A coding sequence	2755-2916
NS4B_SCR	Scrambled NS54B coding sequence	2917-3699
NS5A_SCR	Scrambled NS5A coding sequence	3700-5097
NS5B*_SCR_1636 nt	Partial scrambled NS5B sequence (1636 nt)	5098-6721
NS5B*_WT_152 nt	Partial wild-type NS5B sequence (152 nt)	6722-6873
IRES+Core*_RC	Reverse complement of HCV Internal Ribosome Entry Site followed by a partial HCV Core-coding sequence (325 nt)	6880-7544
HDV	Hepatitis D Virus genomic ribozyme	7545-7631
T7T	T7 RNA Polymerase terminator	7632-7678
Amp R	Ampicillin resistance gene (RC)	8855-9715

6.1.13-14 pUC18_P.s_WT_SL I-II_wt_9968 / pUC18_P.s_WT_SL I-II_S1mS2m_9968

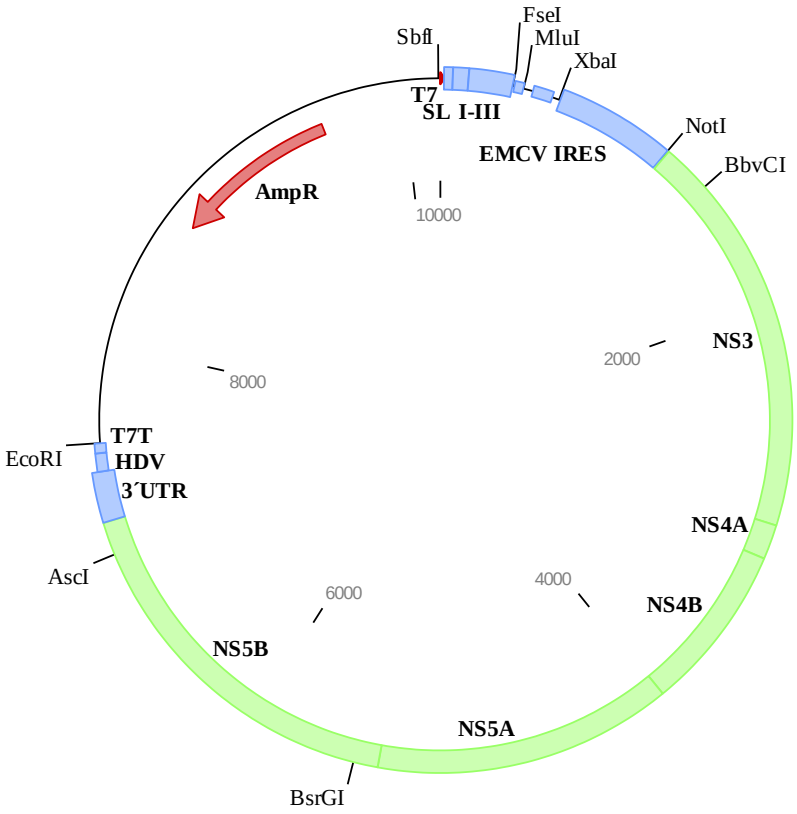


Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt or S1mS2m miR-122 sites)	HCV SL I sequence (with wild-type or mutated miR-122 binding sites S1 and S2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
S1 Aptamer	S1 Streptavidin aptamer sequence	141-185
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	232-329
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	369-924
NS3	Wild-type NS3 coding sequence	931-2823
NS4A	Wild-type NS4A coding sequence	2824-2985
NS4B	Wild-type NS4B coding sequence	2986-3768
NS5A	Wild-type NS5A coding sequence	3769-5166
NS5B	Wild-type NS5B coding sequence	5167-6942
3'UTR	HCV 3'-untranslated region	6943-7178
HDV	Hepatitis D Virus genomic ribozyme	7179-7265
T7T	T7 RNA Polymerase terminator	7266-7312
Amp R	Ampicillin resistance gene (RC)	8489-9349

miR-122 binding site S1:
GCCTGCCCCTAATAGGGGCGACACTCCGCCATGAATCACTCCCCT… mutated to S1m:
GACTGCCCCTAATAGGGGCGAGACTACGCCATGAATCACTCCCCT…

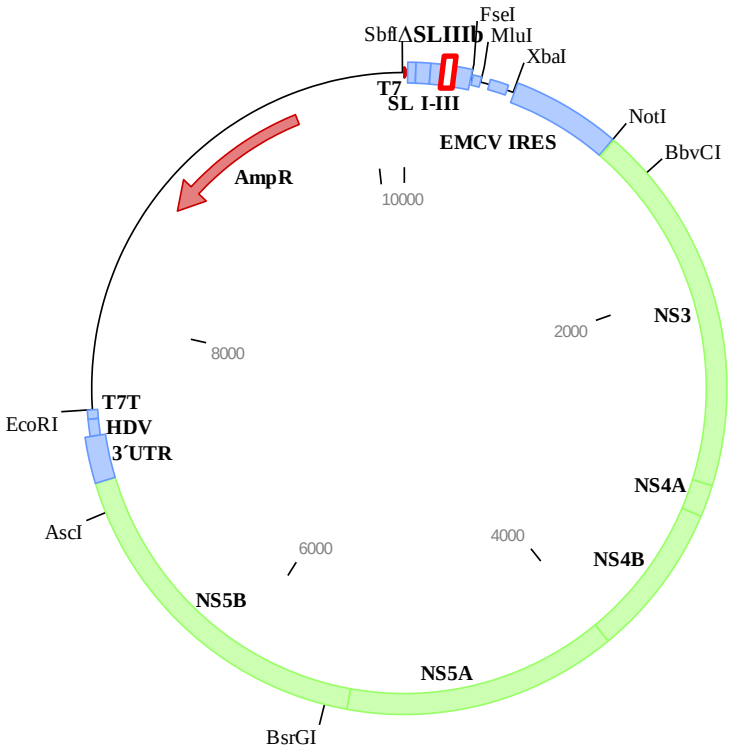
miR-122 binding site S2:
GCCTGCCCCTAATAGGGGCGACACTCCGCCATGAATCACTCCCCT… mutated to S2m:
GACTGCCCCTAATAGGGGCGAGACTACGCAGTGAATCTATCCCCT…

6.1.15-16 pUC18_P.s_WT_SL I-III_10182 / pUC18_P.s_WT_SL I-III_IIId mut_10182



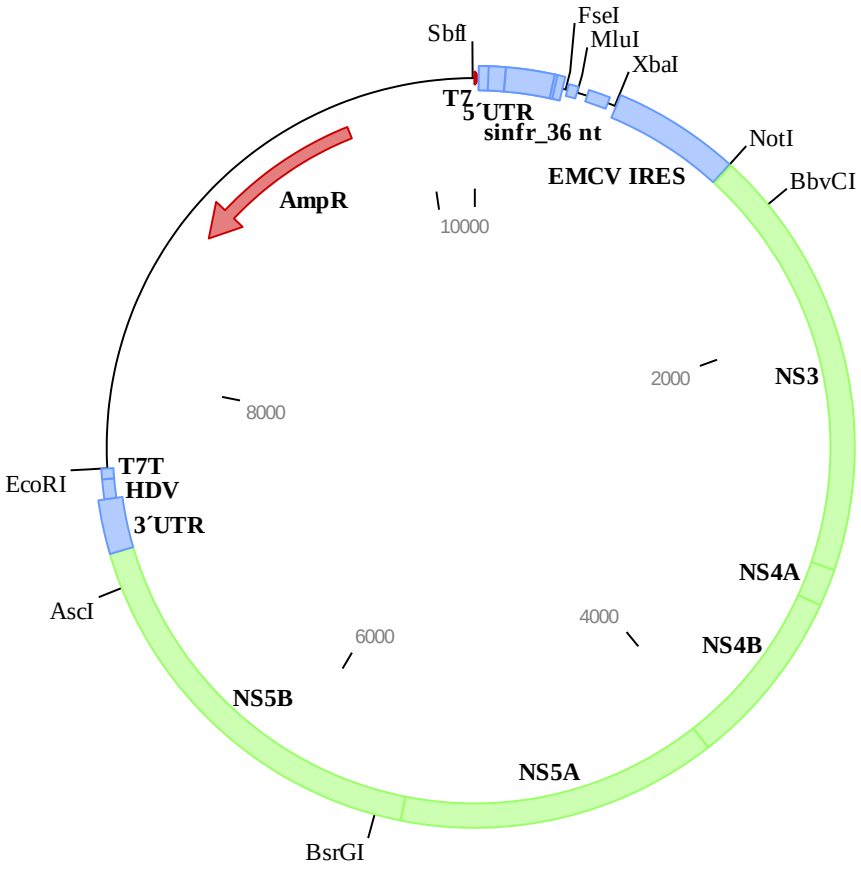
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt miR-122 sites)	HCV SL I sequence (with wild-type miR-122 binding sites 1 and 2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
SL III (wt or SL IIId mut)	HCV SL III sequence (with wild-type or GGG→CCC mutation in SL IIId)	135-346 (282-284)
S1 Aptamer	S1 Streptavidin aptamer sequence	355-399
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	446-543
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	583-1138
NS3	Wild-type NS3 coding sequence	1145-3037
NS4A	Wild-type NS4A coding sequence	3038-3199
NS4B	Wild-type NS4B coding sequence	3200-3982
NS5A	Wild-type NS5A coding sequence	3983-5380
NS5B	Wild-type NS5B coding sequence	5381-7156
3’UTR	HCV 3’-untranslated region	7157-7392
HDV	Hepatitis D Virus genomic ribozyme	7393-7479
T7T	T7 RNA Polymerase terminator	7480-7526
Amp R	Ampicillin resistance gene (RC)	8703-9563

6.1.17 pUC18_P.s_WT_SL I-III_IIIb del_10136



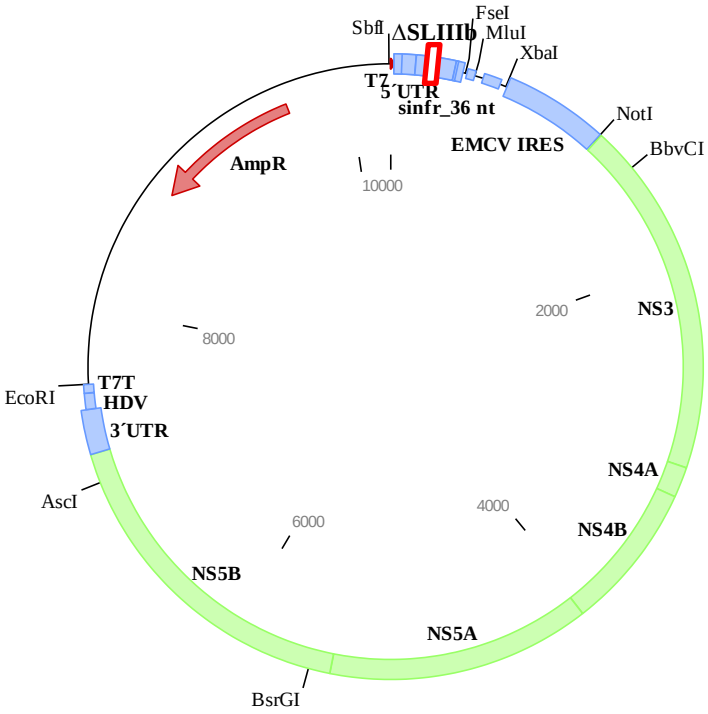
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt miR-122 sites)	HCV SL I sequence (with wild-type miR-122 binding sites S1 and S2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
SL III_IIIb del	HCV SL III sequence a deletion of SL IIIb (46 nt; 193-238 in wt)	135-300
S1 Aptamer	S1 Streptavidin aptamer sequence	309-353
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	400-497
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	537-1092
NS3	Wild-type NS3 coding sequence	1099-2991
NS4A	Wild-type NS4A coding sequence	2992-3153
NS4B	Wild-type NS4B coding sequence	3154-3936
NS5A	Wild-type NS5A coding sequence	3937-5334
NS5B	Wild-type NS5B coding sequence	5335-7110
3’UTR	HCV 3’-untranslated region	7111-7346
HDV	Hepatitis D Virus genomic ribozyme	7347-7433
T7T	T7 RNA Polymerase terminator	7434-7480
Amp R	Ampicillin resistance gene (RC)	8657-9517

6.1.18-19 pUC18_P.s_WT_5'UTR_sinfr_stop_10242 and pUC18_P.s_WT_5'UTR_sinfr_stop_IIId mut_10242



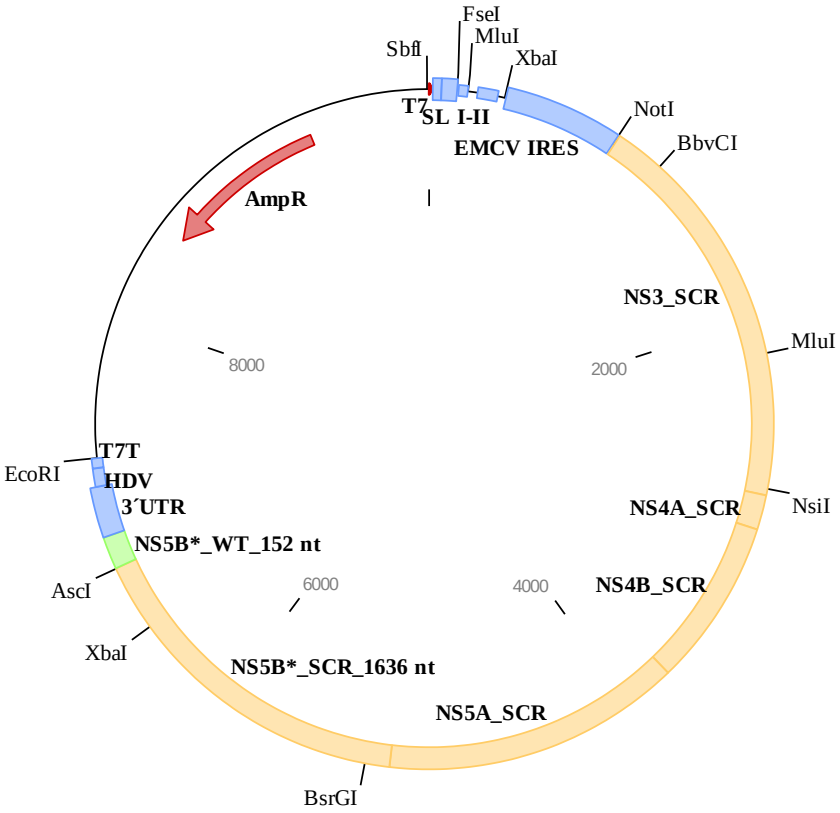
Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt miR-122 sites)	HCV SL I sequence (with wild-type miR-122 binding sites S1 and S2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
SL III (wt or SL IIId mut)	HCV SL III sequence (with wild-type or GGG→CCC mutation in SL IIId)	135-346 (282-284)
SL IV	HCV SL IV sequence	347-370
sinfr_36 nt	Short in-frame HCV Core-coding sequence (36 nt) followed by UGA stop codon	358-393
S1 Aptamer	S1 Streptavidin aptamer sequence	415-459
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	506-603
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	643-1198
NS3	Wild-type NS3 coding sequence	1205-3097
NS4A	Wild-type NS4A coding sequence	3098-3259
NS4B	Wild-type NS4B coding sequence	3260-4042
NS5A	Wild-type NS5A coding sequence	4043-5440
NS5B	Wild-type NS5B coding sequence	5441-7216
3'UTR	HCV 3'-untranslated region	7217-7452
HDV	Hepatitis D Virus genomic ribozyme	7453-7539
T7T	T7 RNA Polymerase terminator	7540-7586
Amp R	Ampicillin resistance gene (RC)	8763-9623

6.1.20 pUC18_P.s_WT_5'UTR_sinfr_stop_IIIb del_10196



Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt miR-122 sites)	HCV SL I sequence (with wild-type miR-122 binding sites S1 and S2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
SL III_IIIb del	HCV SL III sequence a deletion of SL IIIb (46 nt; 193-238 in wt)	135-300
SL IV	HCV SL IV sequence	301-324
sinfr_36 nt	Short in-frame HCV Core-coding sequence (36 nt) followed by UGA stop codon	312-347
S1 Aptamer	S1 Streptavidin aptamer sequence	369-413
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	460-557
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	597-1152
NS3	Wild-type NS3 coding sequence	1159-3051
NS4A	Wild-type NS4A coding sequence	3052-3213
NS4B	Wild-type NS4B coding sequence	3214-3996
NS5A	Wild-type NS5A coding sequence	3997-5394
NS5B	Wild-type NS5B coding sequence	5395-7170
3'UTR	HCV 3'-untranslated region	7171-7406
HDV	Hepatitis D Virus genomic ribozyme	7407-7493
T7T	T7 RNA Polymerase terminator	7494-7540
Amp R	Ampicillin resistance gene (RC)	8717-9577

6.1.21-22 pUC18_P.s_SCR_SL I-II_wt_9968 and pUC18_P.s_SCR_SL I-II_ S1mS2m_9968



Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt or S1mS2m miR-122 sites)	HCV SL I sequence (with wild-type or mutated miR-122 binding sites S1 and S2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
S1 Aptamer	S1 Streptavidin aptamer sequence	141-185
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	232-329
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	369-924
NS3_SCR	Scrambled NS3 coding sequence	931-2823
NS4A_SCR	Scrambled NS4A coding sequence	2824-2985
NS4B_SCR	Scrambled NS54B coding sequence	2986-3768
NS5A_SCR	Scrambled NS5A coding sequence	3769-5166
NS5B*_SCR_1636 nt	Partial scrambled NS5B sequence (1636 nt)	5167-6790
NS5B*_WT_152 nt	Partial wild-type NS5B sequence (152 nt)	6791-6942
3’UTR	HCV 3’-untranslated region	6943-7178
HDV	Hepatitis D Virus genomic ribozyme	7179-7265
T7T	T7 RNA Polymerase terminator	7266-7312
Amp R	Ampicillin resistance gene (RC)	8489-9349

6.1.23 pUC18_P.s_WT_hp_GND_9899

For a plasmid map see 6.1.7.

A replicase-deficient plasmid variant of pUC18_P.s_WT_hp_9899 (6.1.7) harboring a GND mutation (318D→N; Lohmann et al. 1997) within the NS5B RdRp gene (6049G→A).

6.1.24 pUC18_P.s_WT_SL I-II_wt_GND_9968

For a plasmid map see 6.1.13.

A replicase-deficient plasmid variant of pUC18_P.s_WT_SL I-II_wt_9968 (6.1.13) harboring a GND mutation (318D→N; Lohmann et al. 1997) within the NS5B RdRp gene (6118G→A).

6.1.25-27 pUC18_P.s_WT_SL I-II_wt_5B.2m_9968, pUC18_P.s_WT_SL I-II_ wt_5B.3m_9968 and pUC18_P.s_WT_SL I-II_wt_S3m_9968

For a plasmid map see 6.1.13. The plasmids are generated by Nadia Dünnes.

miR-122 binding site 5B.2: ACACTCC (6369-6375) mutated to 5B.2m: ACAC**AG**C (6373-6374 TC→AG);

miR-122 binding site 5B.3: CACTCC (6774-6779) mutated to 5B.3m: **G**AC**A**CC (6774C→G and 6777T→A);

miR-122 binding site S3: ACACTCC (6960-6966) mutated to S3m: AC**T**CT**G**C (6962A→T and 6965C→G).

6.1.28-30 pUC18_P.s_WT_SL I-II_S1mS2m_5B.2m_9968, pUC18_P.s_WT_SL I-II_ S1mS2m_5B.3m_9968 and pUC18_P.s_WT_SL I-II_S1mS2m_S3m_9968

For a plasmid map see 6.1.14. The plasmids are generated by Nadia Dünnes.

miR-122 binding site 5B.2: ACACTCC (6369-6375) mutated to 5B.2m: ACAC**AG**C (6373-6374 TC→AG);

miR-122 binding site 5B.3: CACTCC (6774-6779) mutated to 5B.3m: **G**AC**A**CC (6774C→G and 6777T→A);

miR-122 binding site S3: ACACTCC (6960-6966) mutated to S3m: AC**T**CT**G**C (6962A→T and 6965C→G).

6.1.31 pUC18_P.s_WT_SL I-III_GND_10182

For a plasmid map see 6.1.15.

A replicase-deficient plasmid variant of pUC18_P.s_WT_SL I-III_GND_10182 (6.1.15) harboring a GND mutation (318D→N; Lohmann et al. 1997) within the NS5B RdRp gene (6332G→A).

6.1.32 pUC18_P.s_WT_5'UTR_sinfr_stop_10242

For a plasmid map see 6.1.18.

A replicase-deficient plasmid variant of pUC18_P.s_WT_5'UTR_sinfr_stop_10242 (6.1.18) harboring a GND mutation (318D→N; Lohmann et al. 1997) within the NS5B RdRp gene (6392G→A).

6.1.33-34 pUC18_P.s_WT_SL I-II_wt_8680mut_9968 and pUC18_P.s_WT_SL I-II_ wt_9170mut_9968

For a plasmid map see 6.1.13.

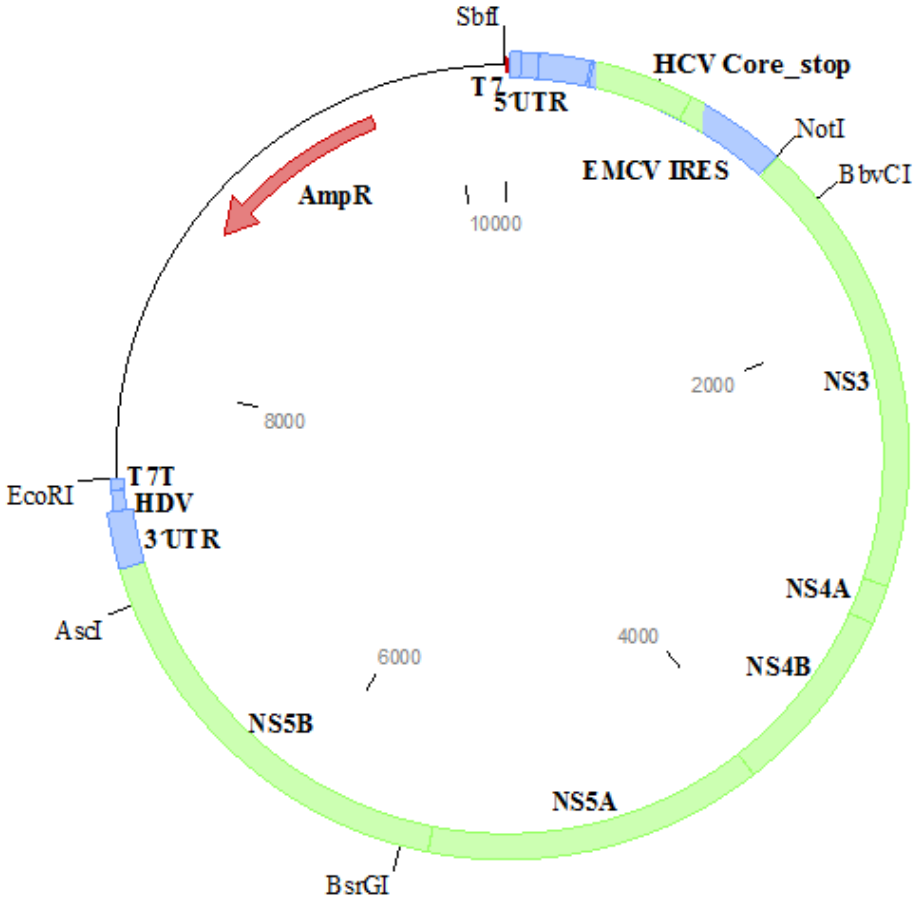
Positions 6180-6187 in the apical loop of the NS5B *cis*-acting RNA element 8680 (in JFH-1 sequence; designated as J8640 in Mauger et al. 2015) in 6.1.13:

…CTTCACGG… mutated to 8680mut: …ATTTACCG…

Positions 6670-6677 for the *cis*-acting RNA element 9170 (in JFH-1 sequence; designated as "9110" in Diviney et al. 2008) in 6.1.13:

…AGTCGGGC… mutated to 9170mut: … UCCAGAGC…

6.1.35 pUC18_P.s_WT_5'UTR_Core_10779



Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt miR-122 sites)	HCV SL I sequence (with wild-type miR-122 binding sites S1 and S2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
SL III	HCV SL III sequence	135-346
SL IV	HCV SL IV sequence	347-370
HCV Core_stop	HCV Core-coding sequence followed by a stop codon	358-930
S1 Aptamer	S1 Streptavidin aptamer sequence	952-996
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	1043-1140
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	1180-1735
NS3	Wild-type NS3 coding sequence	1742-3634
NS4A	Wild-type NS4A coding sequence	3635-3796
NS4B	Wild-type NS4B coding sequence	3797-4579
NS5A	Wild-type NS5A coding sequence	4580-5977
NS5B	Wild-type NS5B coding sequence	5978-7753
3'UTR	HCV 3'-untranslated region	7754-7989
HDV	Hepatitis D Virus genomic ribozyme	7990-8076
T7T	T7 RNA Polymerase terminator	8077-8123
Amp R	Ampicillin resistance gene (RC)	9300-10160

The plasmid and its derivatives 6.1.36-37 are generated by a Gesche Gerresheim.

6.1.36 pUC18_P.s_WT_5'UTR_Core_GND_10779

For a plasmid map see 6.1.35.

A replicase-deficient plasmid variant of pUC18_P.s_WT_5'UTR_Core_10779 (6.1.35) harboring a GND mutation (318D→N; Lohmann et al. 1997) within the NS5B RdRp gene (6929G→A).

6.1.37 pUC18_P.s_WT_5'UTR_Core_IIIb del _10733

For a plasmid map see 6.1.35.

Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	1-18
SL I (wt miR-122 sites)	HCV SL I sequence (with wild-type miR-122 binding sites S1 and S2)	18-59 (38-44 and 54-59)
SL II	HCV SL II sequence	60-134
SL III_IIIb del	HCV SL III sequence a deletion of SL IIIb (46 nt; 193-238 in wt)	135-300
SL IV	HCV SL IV sequence	301-324
Core	HCV Core-coding sequence	312-884
HCV Core_stop	HCV Core-coding sequence followed by a stop codon	906-950
Spinach aptamer_RC	Reverse complement of Spinach RNA aptamer sequence	997-1094
EMCV IRES	Encephalomyocarditis Virus Internal Ribosome Entry Site	1134-1689
NS3	Wild-type NS3 coding sequence	1696-3588
NS4A	Wild-type NS4A coding sequence	3589-3750
NS4B	Wild-type NS4B coding sequence	3751-4533
NS5A	Wild-type NS5A coding sequence	4534-5931
NS5B	Wild-type NS5B coding sequence	5932-7707
3'UTR	HCV 3'-untranslated region	7708-7943
HDV	Hepatitis D Virus genomic ribozyme	7944-8030
T7T	T7 RNA Polymerase terminator	8031-8077
Amp R	Ampicillin resistance gene (RC)	9254-10114

6.1.38 pHCV-SIN_3235

Feature	Description	Start – Stop (nt)
T7	T7 RNA Polymerase promoter	3219-3235
5’UTR	HCV 5’-untranslated region	1-341
Core	Partial HCV Core-coding sequence (61 nt)	342-402
Linker	Partial firefly luciferase ORF	403-437
3’UTR	HCV 3’-untranslated region	438-658
Amp R	Ampicillin resistance gene (RC)	1877-2737

6.1.39 pHCV-3’UTR only_3571

Feature	Description	Start – Stop (nt)
Sp6	SP6 RNA Polymerase promoter	1-21
Linker	Partial firefly luciferase ORF	22-62
3’UTR	HCV 3’-untranslated region	63-283
Amp R	Ampicillin resistance gene (RC)	1510-2370

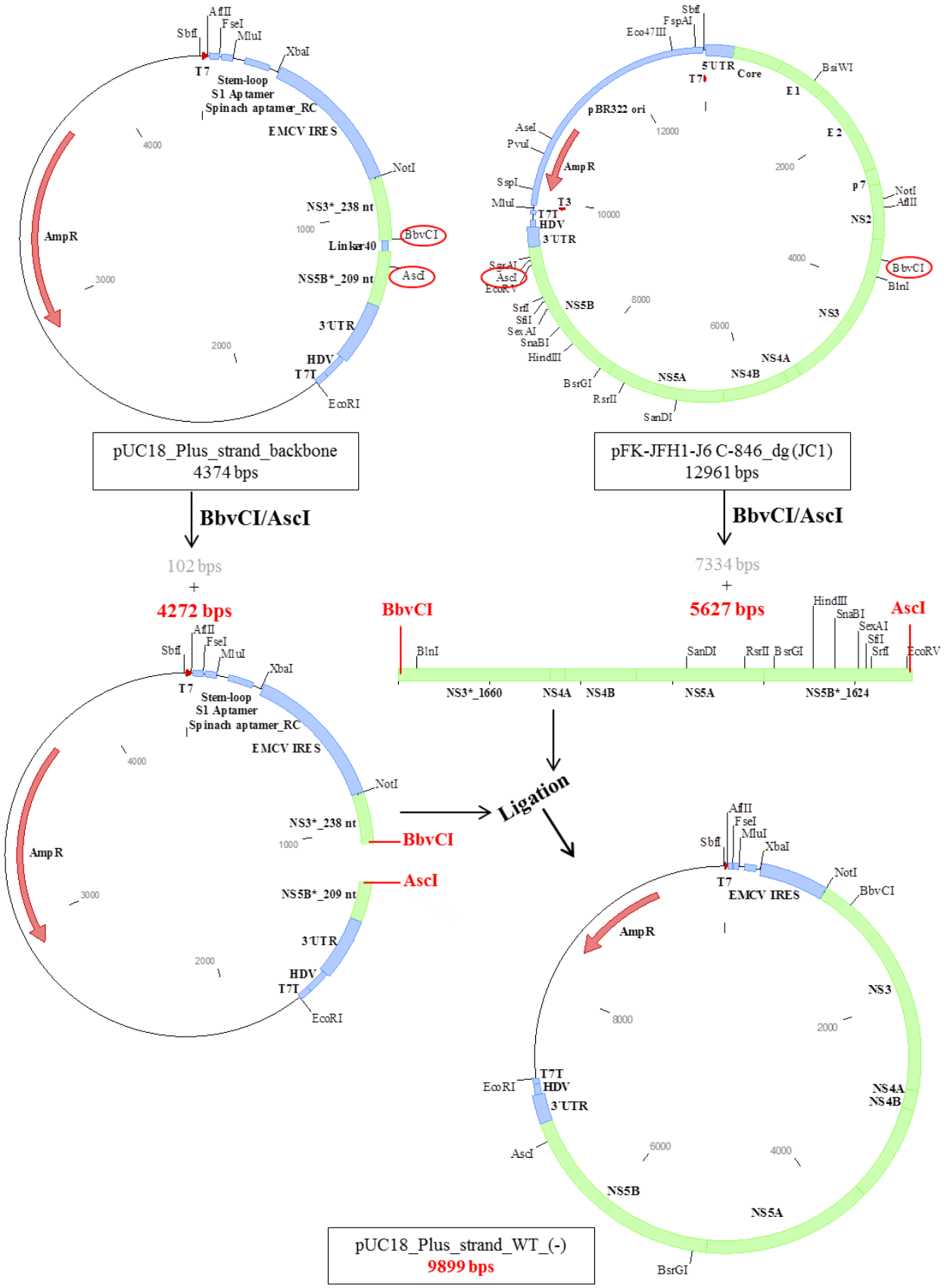
6.2 Supplementary materials

6.2.1 Assembly of DNA templates for minus and plus strand synthesis initiation

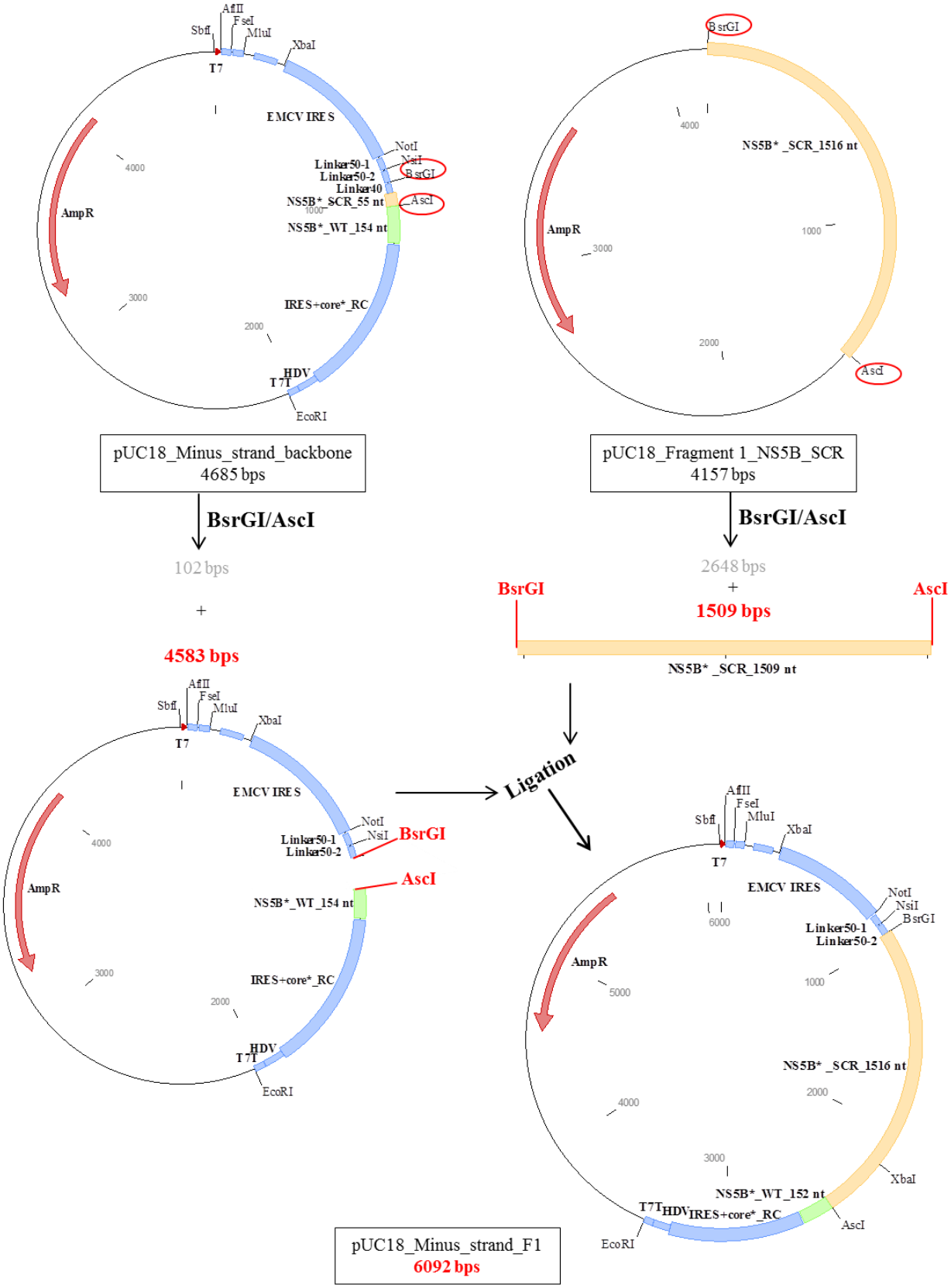
The underlying concept for design of the reduced systems for HCV minus and plus strand replication initiation is described in detail in 4.1. This section illustrates the routine cloning procedures that were carried out for the system assembly. The products of each decisive restriction digest are depicted elucidating all the subsequent cloning steps.

Suppl. Fig. 6.1 provides an overview on the assembly of a DNA template for the synthesis of the HCV minus strands. A backbone construct that mimics plus strand HCV RNA – plasmid pUC18_Plus_strand_backbone_4374 (plasmid map 6.1.2) – was chemically synthesized and inserted into pUC18 vector. It was designed to contain a short random linker sequence flanked by unique restriction sites - *Bbv*CI and *Asc*I. Short stretches of the HCV coding sequences were retained upstream and downstream the linker according to the cloning strategy. However, the majority of HCV NS3-NS5B coding sequence was derived from the plasmid pFK-JFH1-J6 C-846_dg_12961 (plasmid map 6.1.1) that encodes full-length HCV genome. Therefore, following the restriction digest of both plasmids by *Bbv*CI/*Asc*I endonucleases, the NS3-NS5B cassette was transferred into the backbone plasmid, resulting in the complete construct for minus strand synthesis initiation - pUC18_P.s_WT_hp_9899 – minimal “hp” construct (plasmid map 6.1.7).

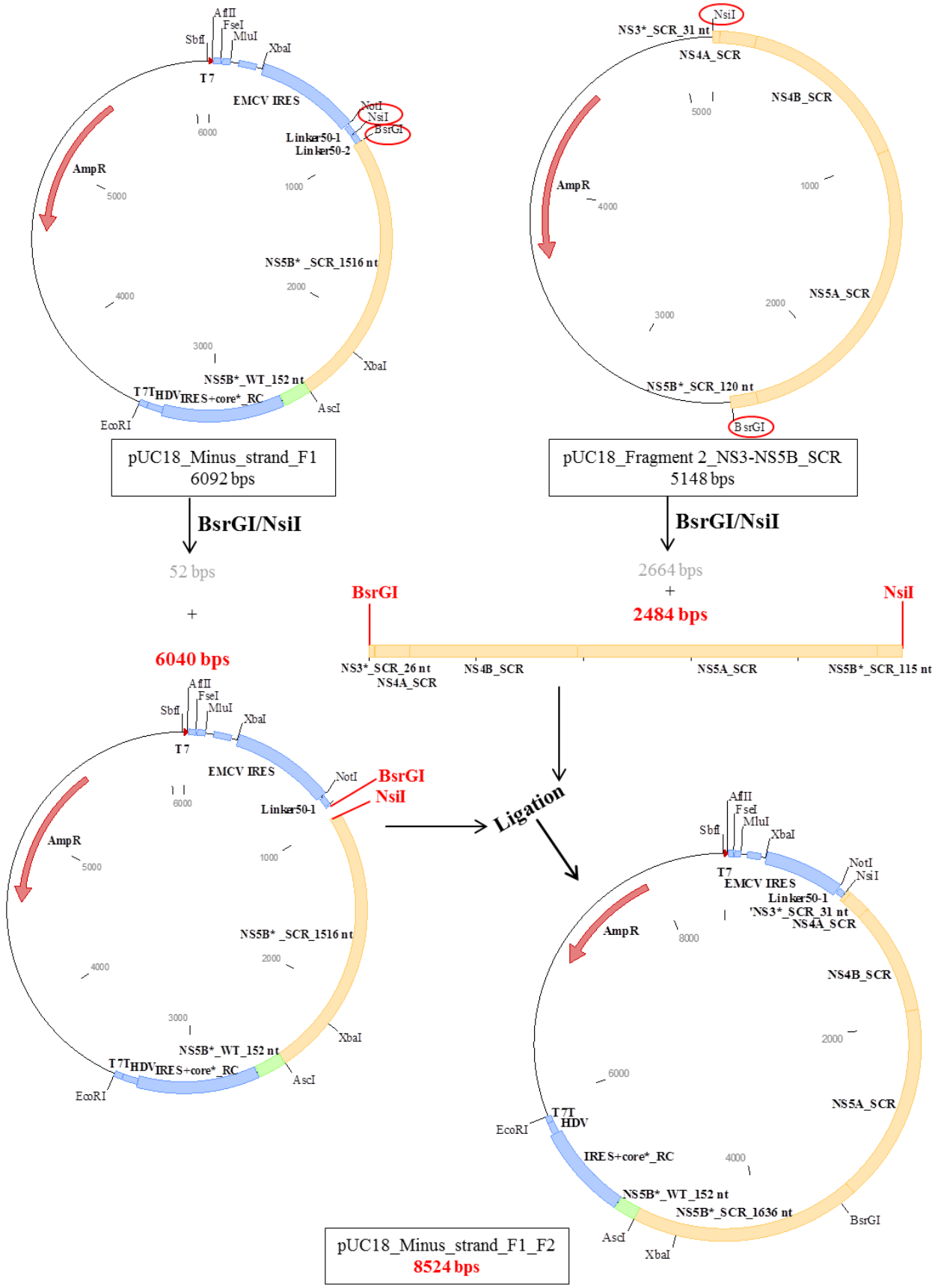
An assembly of a DNA template for the HCV plus strand synthesis initiation required three cloning steps that are illustrated by Suppl. Fig. 6.2-6.4. In analogy to the construct described above, the backbone of the system was chemically synthesized. Similarly, this backbone was incorporated into a pUC18 vector, resulting in pUC18_Minus_strand_backbone_4685 (plasmid map 6.1.3), and designed to harbor three linkers separated by unique restriction sites: *Not*I, *Nsi*I, *Bsr*GI and *Asc*I. To minimize costs, the plasmid was destined to contain scrambled NS3-NS5B cassette, in contrast to pUC18_P.s_WT_hp_9899 (plasmid map 6.1.7), and an exchange of the cassettes between the backbones eventually allowed deriving the basic constructs in all four combinations. Only downstream 152 nt region of the NS5B coding sequence was maintained wild-type in both types of the cassette. Since the scrambled NS3-NS5B sequence was developed specifically for the project, it had to be newly synthesized. The length of the cassette required devision of the sequence into three parts that were subsequently inserted into the backbone. The first cloning step is depicted in Suppl. Fig. 6.2 and was accomplished via substitution of Linker40 with the Fragment 1 (restriction sites used are *Bsr*GI and *Asc*I) (plasmid map 6.1.4), resulting in an intermediate plasmid pUC18_Minus_strand_F1_6092 (plasmid map 6.1.9). Next, after the digest with *Bsr*GI and *Nsi*I, the Fragment 2 (plasmid map 6.1.5) was introduced into the intermediate plasmid replacing Linker50-2, and the second intermediate plasmid - pUC18_Minus_strand_F1_F2_8524 (plasmid map 6.1.10) - was obtained (Suppl. Fig. 6.3). The final cloning step completed assembly of the minus strand mimicking template: the last linker replacement with the Fragment 3 (plasmid map 6.1.6) was mediated by a digest with *Not*I and *Nsi*I endonucleases, resulting in pUC18_M.s_SCR_hp_10334 (plasmid map 6.1.12) (Suppl. Fig. 6.4). Each DNA template – for minus and plus strand replication – was finally derived in two variants: encoding either wild-type or scrambled NS3-NS5B cassette. The initial design allowed an exchange of the cassettes between the backbones due to unique *Not*I/*Asc*I restriction sites flanking the NS3-NS5B sequences. Suppl. Fig. 6.5 illustrates a production of plasmids pUC18_P.s_SCR_hp_9899 (plasmid map 6.1.8) and pUC18_M.s_WT_hp_10334 (plasmid map 6.1.11).



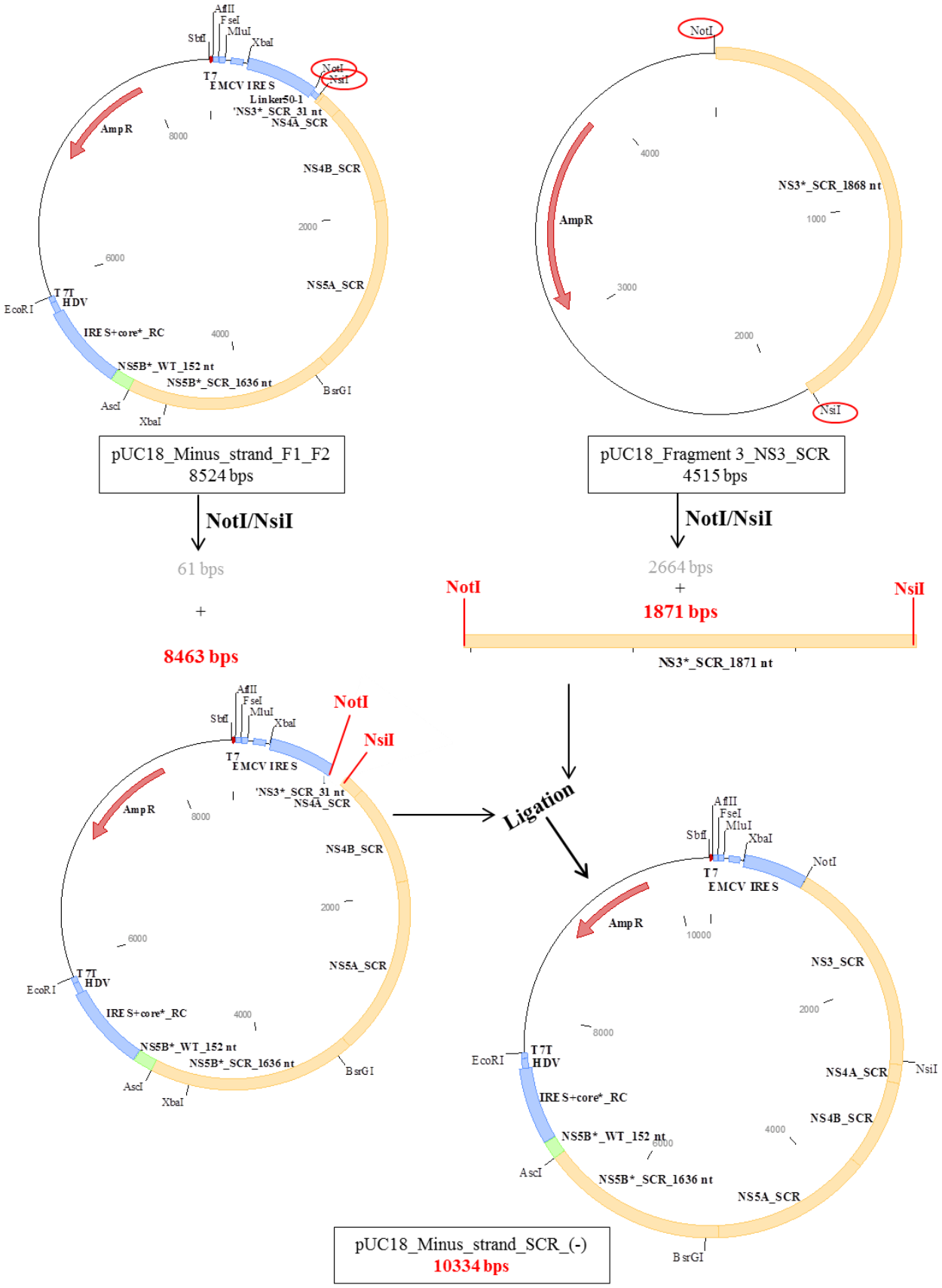
Supplementary Figure 6.1: Assembly of a DNA template for minus strand synthesis initiation.



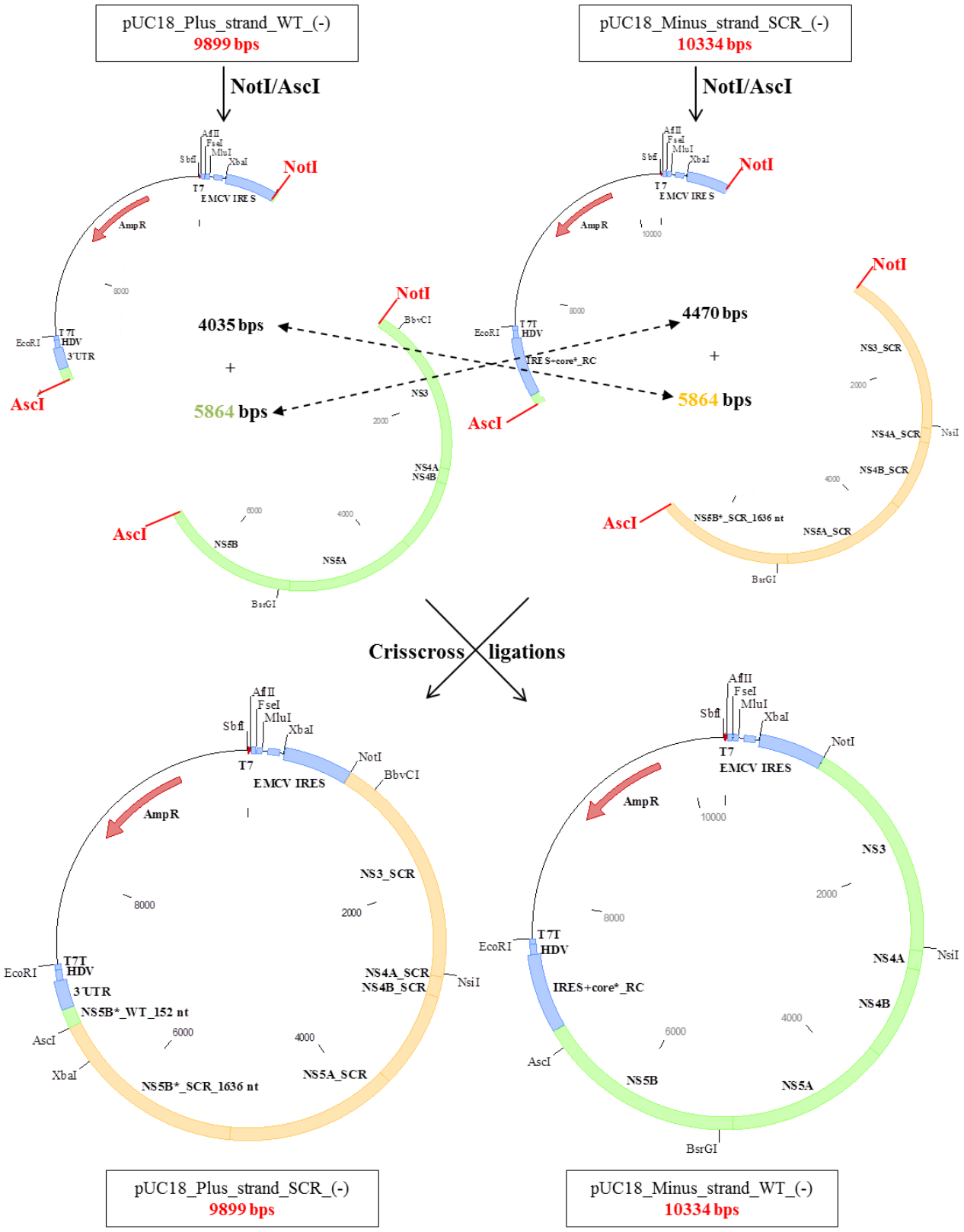
Supplementary Figure 6.2: Assembly of an intermediate plasmid pUC18_Minus_strand_F1.



Supplementary Figure 6.3: Assembly of an intermediate plasmid pUC18_Minus_strand_F1_F2.



Supplementary Figure 6.4: Assembly of a DNA template for plus strand synthesis initiation.



Supplementary Figure 6.5: A scheme of the backbones/cassettes exchange.

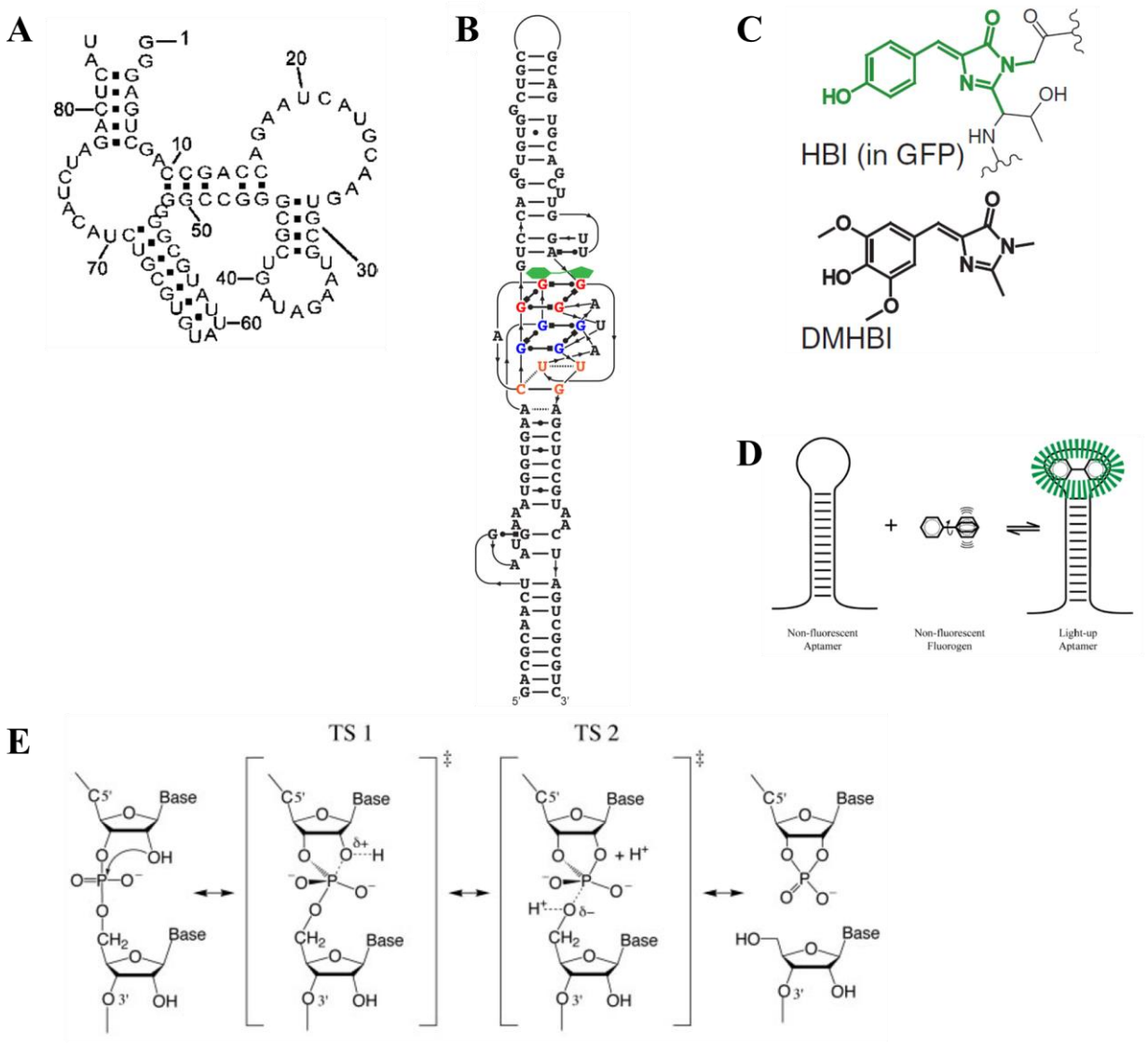
6.2.2 Structural elements of the replication system

The design of the replication system in the current work required introduction of several complex structural elements serving functional purposes for the constructs production and detection. Both – the template for minus and plus strand synthesis initiation – share the same functional units (Fig. 3.1.1, B and C).

The Streptavidin aptamer (Suppl. Fig. 6.6, A) was included in the system design to allow isolation and analysis of the HCV replication complex acting at the functional replication template. This RNA affinity tag enables a selective recovery of an input construct together with all associated proteins under nondenaturing conditions: the aptamer binds specifically to streptavidin and can be gently eluted by competition with biotin. Using streptavidin as a target of the RNA affinity tag is very advantageous due its availability and extraordinary high affinity to biotin. The optimal sequence of the Streptavidin aptamer was derived via a SELEX approach and the corresponding aptamers were easily removable from streptavidin with biotin (Srisawat and Engelke 2001). The aptamers are classified in accordance to their consensus sequence, and in the current work the full-length S1 aptamer (group 1) was included for the system construction.

Following the Streptavidin aptamer, the antisense Spinach aptamer sequence is incorporated in the designed constructs ("as-spinach"; Fig. 3.1.1, B and C). This aptamer is provided on input replication templates as a reverse complementary in order to express its functional secondary structure at the newly synthesized RNA strand and to enable a detection of the synthesis initiation event. The unique spatial configuration of this RNA aptamer was developed to bind fluorophores and emit a green fluorescence demonstrating a GFP-like functionality (Suppl. Fig. 6.6, B and D) (Paige et al. 2011). The Spinach aptamer allows a protein-free direct post-transcriptional live-cell imaging that, in contrast to the GFP fluorescence, is resistant to photobleaching. To enhance fluorescence provided with the use of 4-hydroxybenzlidene imidazolinone (HBI; fluorogen in GFP system) its derivative – DMHBI (3,5-dimethoxy-4-hydroxybenzylidene imidazolinone) – is applied to induce fluorescence when combined with the Spinach aptamer (Suppl. Fig. 6.6, C). DMHBI appeared to be the most suitable due to the lack of nonspecific cellular fluorescence as well as cyto- and photo- toxicity. The role of the aptamer itself is to stabilize the planar structure of the fluorogen directing dissipation of its energy predominantly through radiative decay pathways (Ouellet 2016) (Suppl. Fig. 6.6, D). Although both aptamers have not been in use in the present work, they will certainly advantage further applications of the developed replication system.

A structural component of the system that plays a vital role in replication RNA template production is a Hepatitis Delta Virus (HDV) ribozyme (Fig. 3.1.1, B and C). Insertion of a ribozyme downstream the HCV coding sequence aims at processing of *in vitro* transcribed RNA to generate the exact HCV 3'-end. The very same nucleotide at the 3'-end of replication templates is an indispensable prerequisite for the function of the HCV replication complex in the first place, in addition to valid comparison of synthesis initiation efficiency from different templates. The HDV ribozyme undertakes non-enzyme-catalyzed RNA cleavage mechanism in which making and breaking of phosphorus-oxygen bonds take place: the adjacent 2'-hydroxyl group is the attacking nucleophile in the transesterification reaction and a 2',3'-cyclic phosphate group is generated upon cleavage (Suppl. Fig. 6.6, E). The genomic and antigenomic HDV ribozyme sequences form a similar secondary structure of a nested double pseudoknot containing base-paired regions (P1-4), joining sequences, and hairpin loops. The folding into two helical stacks composed of the duplex regions stabilizes the ribozyme's structure and brings the catalytic residue (cytosine C75) to a proximity of the cleavage site (Suppl. Fig. 6.6, F) (Shih and Been 2002).



Supplementary Figure 6.6: Structure and function of the aptamers included in the replication system.
(A) The predicted secondary structure of the full-length S1 Streptavidin aptamer (Srisawat and Engelke 2001).
(B) Secondary structure of the Spinach aptamer determined from its crystal structure; the 3,5-dimethoxy-4-hydroxybenzylidene imidazolinone (DMHBI) fluorogen is depicted in green.
(C) Structures of 4-hydroxybenzlidene imidazolinone (HBI) in the context of GFP (in green) and DMHBI.
(D) A mechanism of a light-up RNA aptamer function: a fluorogen alone dissipates its energy on non-radiative decay pathways (e.g. molecular vibrations); binding to an aptamer promotes a radiative decay pathway resulting in fluorescence enhancement.
(E) Chemical mechanism of 2'-hydroxyl displacement for ribozyme RNA cleavage. Transition states (TS) 1 and 2 illustrate the formation of the new P–O bond and the breaking of the original P–O bond, respectively.
(F) Structure and sequence of the HDV genomic and antigenomic ribozymes. P1-P4 are the base-paired regions; the catalytic cytosine is labeled in blue.
(A is from Srisawat and Engelke 2001; B and D are from Ouellet 2016; C is from Paige et al. 2011; E and F are from Shih and Been 2002)

6.2.3 HCV NS3-NS5B wild-type and scrambled sequence

Comparison of the nucleotide HCV NS3-NS5B coding sequences within wild-type (WT) and scrambled (SCR) RNA constructs' versions. The last 152 nt of the NS5B coding-sequence were retained unmodified during cloning process to maintain a wild-type sequence of the 5BSL3.2 element. The alignment is CLUSTAL multiple sequence alignment by MUSCLE (3.8).

```
NS3-NS5B_WT       GCTCCCATCACTGCTTATGCCCAGCAAACACGAGGCCTCCTGGGCGCCATAGTGGTGAGT
NS3-NS5B_SCR      GCTCCTATCACGGCCTACGCACAACAGACGAGGGGACTTCTTGGGGCTATAGTTGTCTCC
                  ***** ***** ** ** ** ** ** **  * ** ** ** ** ** ***** **

NS3-NS5B_WT       ATGACGGGGCGTGACAGGACAGAACAGGCCGGGGAAGTCCAAATCCTGTCCACAGTCTCT
NS3-NS5B_SCR      ATGACAGGAAGAGACCGAACTGAGCAGGCTGGCGAAGTCCAAATCTTATCGACCGTGAGC
                  ***** **  * *** * ** ** ***** ** ************ * ** ** **

NS3-NS5B_WT       CAGTCCTTCCTCGGAACAACCATCTCGGGGGTTTTGTGGACTGTTTACCACGGAGCTGGC
NS3-NS5B_SCR      CAATCCTTTCTGGGCACCACCATCTCCGGTGTCTTGTGGACAGTCTACCATGGGGCGGGG
                  ** ***** ** ** ** ******** ** ** ******** ** ***** ** ** **

NS3-NS5B_WT       AACAAGACTCTAGCCGGCTTACGGGGTCCGGTCACGCAGATGTACTCGAGTGCTGAGGGG
NS3-NS5B_SCR      AATAAGACCCTGGCGGGGCTCAGAGGTCCGGTCACCCAAATGTATTCGTCAGCTGAGGGG
                  ** ***** ** ** **  *  * ********** ** ***** ***   *********

NS3-NS5B_WT       GACTTGGTAGGCTGGCCCAGCCCCCCTGGGACCAAGTCTTTGGAGCCGTGCAAGTGTGGA
NS3-NS5B_SCR      GACCTCGTTGGGTGGCCCTCCCCGCCCGGGACCAAAAGTCTGGAACCCTGCAAGTGCGGA
                  *** * ** ** ******  *** ** ********   * **** ** ******** ***

NS3-NS5B_WT       GCCGTCGACCTATATCTGGTCACGCGGAACGCTGATGTCATCCCGGCTCGGAGACGCGGG
NS3-NS5B_SCR      GCCGTGGATCTGTATCTCGTCACCCGCAATGCCGATGTGATCCCAGCCCGCCGGCGTGGT
                  ***** ** ** ***** ***** ** ** ** ***** ***** ** **  * ** **

NS3-NS5B_WT       GACAAGCGGGGAGCATTGCTCTCCCCGAGACCCATTTCGACCTTGAAGGGGTCCTCGGGG
NS3-NS5B_SCR      GACAAACGGGGGGCTCTTCTCTCGCCCAGACCTATTTCAACACTAAAAGGGAGCTCAGGT
                  ***** ***** **  * ***** ** ***** ***** **  * ** ***  *** **

NS3-NS5B_WT       GGGCCGGTGCTCTGCCCTAGGGGCCACGTCGTTGGGCTCTTCCGAGCAGCTGTGTGCTCT
NS3-NS5B_SCR      GGCCCAGTGCTTTGCCCCCGAGGGCATGTGGTGGGCTTGTTCAGGGCTGCGGTCTGCAGC
                  ** ** ***** *****  * ** ** ** ** **  * *** * ** ** ** ***

NS3-NS5B_WT       CGGGGCGTGGCCAAATCCATCGATTTCATCCCCGTTGAGACACTCGACGTTGTTACAAGG
NS3-NS5B_SCR      CGCGGCGTAGCAAAGTCTATCGATTTTATCCCCGTGGAGACCTTGGATGTGGTGACTCGG
                  ** ***** ** ** ** ******** ******** *****  * ** ** ** **  **

NS3-NS5B_WT       TCTCCCACTTTCAGTGACAACAGCACGCCACCGGCTGTGCCCCAGACCTATCAGGTCGGG
NS3-NS5B_SCR      TCTCCCACTTTCAGCGATAACTCCACACCTCCCGCAGTTCCCCAAACTTACCAAGTAGGC
                  ************** ** ***  *** ** ** ** ** ** ***** ** ** ** ** **

NS3-NS5B_WT       TACTTGCATGCTCCAACTGGCAGTGGAAAGAGCACCAAGGTCCCTGTCGCGTATGCCGCC
NS3-NS5B_SCR      TATCTCCACGCGCCGACGGGGAGCGGTAAGTCGACAAAAGTTCCAGTCGCTTATGCGGCT
                  **  * ** ** ** ** ** ** ** ***   ** ** ** ** ***** ***** **

NS3-NS5B_WT       CAGGGGTACAAAGTACTAGTGCTTAACCCCTCGGTAGCTGCCACCCTGGGGTTTGGGGCG
NS3-NS5B_SCR      CAGGGCTACAAGGTCTTGGTTCTAAACCCCTCTGTCGCGGCGACGCTGGGGTTCGGGGCC
                  ***** ***** **  * ** ** ******** ** ** ** ** ******** *****

NS3-NS5B_WT       TACCTATCCAAGGCACATGGCATCAATCCCAACATTAGGACTGGAGTCAGGACCGTGATG
NS3-NS5B_SCR      TATTTGAGCAAAGCCCACGGGATCAACCCGAATATTAGGACCGGTGTTCGCACAGTCATG
                  **  *   *** ** ** ** ***** ** ** ******** ** **  * ** ** ***

NS3-NS5B_WT       ACCGGGGAGGCCATCACGTACTCCACATATGGCAAATTTCTCGCCGATGGGGGCTGCGCT
NS3-NS5B_SCR      ACCGGCGAAGCAATTACCTACTCAACCTACGGGAAGTTCCTAGCCGACGGCGGCTGCGCC
                  ***** ** ** ** ** ***** ** ** ** ** ** ** ***** ** ********

NS3-NS5B_WT       AGCGGCGCCTATGACATCATCATATGCGATGAATGCCACGCTGTGGATGCTACCTCCATT
NS3-NS5B_SCR      TCAGGGGCCTACGACATCATCATCTGCGACGAGTGTCATGCCGTCGACGCCACGTCCATC
                     ** ***** ********** ***** ** ** ** ** ** ** ** ** *****

NS3-NS5B_WT       CTCGGCATCGGAACGGTCCTTGATCAAGCAGAGACAGCCGGGGTCAGACTAACTGTGCTG
NS3-NS5B_SCR      CTCGGGATCGGCACTGTCCTCGACCAGGCGGAGACAGCGGGCGTCCGGCTCACCGTCCTT
                  ***** ***** ** ***** ** ** ** ******** ** *** * ** ** ** **
```

```
NS3-NS5B_WT       GCTACGGCCACACCCCCCGGGTCAGTGACAACCCCCCATCCCGATATAGAAGAGGTAGGC
NS3-NS5B_SCR      GCCACAGCCACTCCCCCCGGCAGCGTCACCACGCCCCACCCAGACATCGAGGAGGTGGGT
                  ** ** ***** ********    ** ** ** ***** ** ** ** ** ***** **

NS3-NS5B_WT       CTCGGGCGGGAGGGTGAGATCCCCTTCTATGGGAGGGCGATTCCCCTATCCTGCATCAAG
NS3-NS5B_SCR      TTGGGCAGAGAAGGAGAAATCCCTTTTTATGGACGCGCAATCCCTCTCAGCTGCATCAAA
                   * **  * ** ** ** ***** ** *****  * ** ** ** **   *********

NS3-NS5B_WT       GGAGGGAGACACCTGATTTTCTGCCACTCAAAGAAAAAGTGTGACGAGCTCGCGGCGGCC
NS3-NS5B_SCR      GGGGGCAGGCACCTCATTTTCTGCCACTCTAAGAAGAAGTGCGACGAGTTGGCTGCCGCT
                  ** ** ** ***** ************** ***** ***** ****** * ** ** **

NS3-NS5B_WT       CTTCGGGGCATGGGCTTGAATGCCGTGGCATACTATAGAGGGTTGGACGTCTCCATAATA
NS3-NS5B_SCR      CTAAGGGGAATGGGACTCAATGCGGTAGCTTATTACCGGGGACTGGACGTTTCGATAATA
                  **  **** *****  * ***** ** ** ** **  * **  ******* ** ******

NS3-NS5B_WT       CCAGCTCAGGGAGATGTGGTGGTCGTCGCCACCGACGCCCTCATGACGGGGTACACTGGA
NS3-NS5B_SCR      CCAGCCCAGGGGGATGTGGTCGTGGTGGCCACGGACGCGTTAATGACTGGATATACGGGG
                  ***** ***** ******** ** ** ***** *****  * ***** ** ** ** **

NS3-NS5B_WT       GACTTTGACTCCGTGATCGACTGCAATGTAGCGGTCACCCAAGCTGTCGACTTCAGCCTG
NS3-NS5B_SCR      GACTTCGACTCCGTCATCGACTGCAACGTGGCCGTGACCCAGGCCGTGGACTTTTCCCTA
                  ***** ******** ********** ** ** ** ***** ** ** *****   ***

NS3-NS5B_WT       GACCCCACCTTCACTATAACCACACAGACTGTCCCACAAGACGCTGTCTCACGCAGT
NS3-NS5B_SCR      GATCCTACATTCACAATAACAACCCAAACGGTGCCGCAGGACGCCGTCAGCAGGTCT
                  ** ** ** ***** ***** ** ** ** ** ** ** ***** ***    *   *

NS3-NS5B_WT       CAGCGCCGCGGGCGCACAGGTAGAGGAAGACAGGGCACTTATAGGTATGTTTCCACTGGT
NS3-NS5B_SCR      CAGAGAAGAGGGAGAACTGGCAGAGGCCGGCAAGGGACCTATCGCTACGTGTCCACGGGT
                  *** *  * *** * ** ** *****  * ** ** ** *** * ** ** ***** ***

NS3-NS5B_WT       GAACGAGCCTCAGGAATGTTTGACAGTGTAGTGCTTTGTGAGTGCTACGACGCAGGGGCT
NS3-NS5B_SCR      GAGAGGGCGTCAGGCATGTTCGACTCCGTGGTCCTGTGCGAGTGCTACGATGCCGGCGCT
                  **  * ** ***** ***** ***   ** ** ** ** ********** ** ** ***

NS3-NS5B_WT       GCGTGGTACGATCTCACACCAGCGGAGACCACCGTCAGGCTTAGAGCGTATTTCAACACG
NS3-NS5B_SCR      GCCTGGTACGACTTGACTCCGGCTGAGACCACGGTGAGGCTGCGGGCCTACTTCAACACC
                  ** ********  * ** ** ** ******** ** *****  * ** ** ********

NS3-NS5B_WT       CCCGGCCTACCCGTGTGTCAAGACCATCTTGAATTTTGGGAGGCAGTTTTCACCGGCCTC
NS3-NS5B_SCR      CCCGGGCTGCCCGTGTGCCAGGACCACCTAGAGTTCTGGGAGGCCGTCTTCACTGGCCTG
                  ***** ** ******** ** ***** ** ** ** ******** ** ***** *****

NS3-NS5B_WT       ACACACATAGACGCCCACTTCCTCTCCCAAACAAAGCAAGCGGGGGAGAACTTCGCGTAC
NS3-NS5B_SCR      ACCCACATTGATGCTCATTTTCTGTCACAGACGAAGCAGGCTGGAGAAAATTTTGCCTAC
                  ** ***** ** ** ** ** ** ** ** ** ***** ** ** ** ** ** ** ***

NS3-NS5B_WT       CTAGTAGCCTACCAAGCTACGGTGTGCGCCAGAGCCAAGGCCCCTCCCCCGTCCTGGGAC
NS3-NS5B_SCR      CTTGTGGCCTACCAGGCCACCGTCTGCGCGCGGGCCAAGGCGCCTCCACCATCGTGGGAT
                  ** ** ******** ** ** ** *****  * ******** ***** ** ** *****

NS3-NS5B_WT       GCCATGTGGAAGTGCCTGGCCCGACTCAAGCCTACGCTTGCGGGCCCCACACCTCTCCTG
NS3-NS5B_SCR      GCAATGTGGAAGTGCCTGGCCAGGCTGAAACCTACCCTGGCCGGGCCCACCCCGCTTCTC
                  ** ****************** * ** ** ***** ** ** ** ***** ** ** **

NS3-NS5B_WT       TACCGTTTGGGCCCTATTACCAATGAGGTCACCCTCACACACCCTGGGACGAAGTACATC
NS3-NS5B_SCR      TACAGGCTCGGGCCCATCACAAACGAGGTAACGTTGACTCACCCTGGCACCAAGTATATC
                  *** *  * ** ** ** ** ** ***** **  * ** ******** ** ***** ***

NS3-NS5B_WT       GCCACATGCATGCAAGCTGACCTTGAGGTCATGACCAGCACGTGGGTCCTAGCTGGAGGA
NS3-NS5B_SCR      GCTACATGCATGCAGGCAGACTTGGAGGTGATGACATCGACCTGGGTGCTGGCTGGGGGA
                  ** ********** ** *** * ***** *****    ** ***** ** ***** ***

NS3-NS5B_WT       GTCCTGGCAGCCGTCGCCGCATATTGCCTGGCGACTGGATGCGTTTCCATCATCGGCCGC
NS3-NS5B_SCR      GTTCTGGCTGCTGTGGCGGCTTACTGCCTGGCCACAGGATGCGTATCCATCATCGGACGG
                  ** ***** ** ** ** ** ** ******** ** ******** ********** **
```

```
NS3-NS5B_WT       TTGCACGTCAACCAGCGAGTCGTCGTTGCGCCGGATAAGGAGGTCCTGTATGAGGCTTTT
NS3-NS5B_SCR      CTTCATGTGAATCAAAGGGTCGTCGTTGCCCCGGACAAGGAGGTGCTCTACGAGGCCTTC
                   * ** ** ** **  * ********** ***** ******** ** ** ***** **

NS3-NS5B_WT       GATGAGATGGAGGAATGCGCCTCTAGGGCGGCTCTCATCGAAGAGGGGCAGCGGATAGCC
NS3-NS5B_SCR      GACGAGATGGAGGAGTGTGCTTCGCGTGCCGCCTTGATCGAGGAGGGGCAGAGAATCGCG
                  ** ********** ** ** **  * ** **  * ***** ********* * ** **

NS3-NS5B_WT       GAGATGTTGAAGTCCAAGATCCAAGGCTTGCTGCAGCAGGCCTCTAAGCAGGCCCAGGAC
NS3-NS5B_SCR      GAAATGTTGAAGTCTAAGATACAGGGGCTACTACAGCAAGCTTCGAAGCAGGCTCAAGAC
                  ** ********** ***** ** **  * ** ***** ** ** ******** ** ***

NS3-NS5B_WT       ATACAACCCGCTATGCAGGCTTCATGGCCCAAAGTGGAACAATTTTGGGCCAGACACATG
NS3-NS5B_SCR      ATACAACCCGCCATGCAAGCCAGCTGGCCGAAGGTCGAGCAGTTCTGGGCGCGGCACATG
                  ********** ***** **    ***** ** ** ** ** ** *****  * ******

NS3-NS5B_WT       TGGAACTTCATTAGCGGCATCCAATACCTCGCAGGATTGTCAACACTGCCAGGGAACCCC
NS3-NS5B_SCR      TGGAACTTTATTTCAGGGATACAATATTTGGCTGGTCTCTCCACCCTCCCAGGAAACCCC
                  ******** ***   ** ** *****  * ** **  * ** ** ** ***** ******

NS3-NS5B_WT       GCGGTGGCTTCCATGATGGCATTCAGTGCCGCCCTCACCAGTCCGTTGTCGACCAGTACC
NS3-NS5B_SCR      GCTGTGGCGTCTATGATGGCCTTCAGTGCGGCTCTTACGTCCCCACTATCCACATCGACT
                  ** ***** ** ******** ******** ** ** **    **  * ** **    **

NS3-NS5B_WT       ACCATCCTTCTCAACATCATGGGAGGCTGGTTAGCGTCCCAGATCGCACCACCCGCGGGG
NS3-NS5B_SCR      ACTATCTTGCTAAATATCATGGGGGGCTGGCTGGCCTCCCAAATCGCCCCTCCTGCTGGA
                  ** *** * ** ** ******** ****** * ** ***** ***** ** ** ** **

NS3-NS5B_WT       GCCACCGGCTTTGTCGTCAGTGGCCTGGTGGGGGCTGCCGTGGGCAGCATAGGCCTGGGT
NS3-NS5B_SCR      GCAACCGGATTTGTTGTATCAGGGCTCGTTGGCGCGGCCGTTGGTTCCATCGGTCTCGGC
                  ** ***** ***** **    ** ** ** ** ** ***** **   *** ** ** **

NS3-NS5B_WT       AAGGTGCTGGTGGACATCCTGGCAGGATATGGTGCGGGCATTTCGGGGGCCCTCGTCGCA
NS3-NS5B_SCR      AAAGTTTTAGTAGATATACTGGCTGGGTACGGAGCTGGAATCTCGGGAGCTCTCGTGGCC
                  ** **  * ** ** ** ***** ** ** ** ** ** ** ***** ** ***** **

NS3-NS5B_WT       TTCAAGATCATGTCTGGCGAGAAGCCCTCTATGGAAGATGTCATCAATCTACTGCCTGGG
NS3-NS5B_SCR      TTTAAAATTATGAGCGGGGAGAAACCCTCTATGGAGGACGTCATTAACCTGCTCCCGGGC
                  ** ** ** ***   ** ***** ********** ** ***** ** ** ** ** **

NS3-NS5B_WT       ATCCTGTCTCCGGGAGCCCTGGTGGTGGGGGTCATCTGCGCGGCCATTCTGCGCCGCCAC
NS3-NS5B_SCR      ATTCTCTCTCCTGGGGCACTTGTGGTTGGTGTCATATGTGCAGCCATTCTGAGGAGGCAT
                  ** ** ***** ** ** ** ***** ** ***** ** ** ********* *  * **

NS3-NS5B_WT       GTGGGACCGGGGGAGGGCGCGGTCCAATGGATGAACAGGCTTATTGCCTTTGCTTCCAGA
NS3-NS5B_SCR      GTGGGTCCAGGAGAGGGAGCTGTTCAGTGGATGAACCGCCTGATAGCTTTCGCCTCGCGG
                  ***** ** ** ***** ** ** ** ********* * ** ** ** ** ** ** **  *

NS3-NS5B_WT       GGAAACCACGTCGCCCCTACTCACTACGTGACGGAGTCGGATGCGTCGCAGCGTGTGACC
NS3-NS5B_SCR      GGCAACCACGTGGCGCCCACCCACTATGTCACCGAATCAGATGCCTCCCAAAGGGTCACT
                  ** ******** ** ** ** ***** ** ** ** ** ***** ** **  * ** **

NS3-NS5B_WT       CAACTACTTGGCTCTCTTACTATAACCAGCCTACTCAGAAGACTCCACAATTGGATAACT
NS3-NS5B_SCR      CAGCTGCTAGGGTCCCTCACCATCACTTCTCTCCTGAGACGGCTTCACAACTGGATAACA
                  ** ** ** ** ** ** ** ** **    ** ** *** * ** ***** ********

NS3-NS5B_WT       GAGGACTGCCCCATCCCATGCTCCGGATCCTGGCTCCGCGACGTGTGGGACTGGGTTTGC
NS3-NS5B_SCR      GAGGACTGCCCAATTCCTTGCAGTGGGAGCTGGCTAAGAGATGTCTGGGACTGGGTCTGC
                  *********** ** ** ***   **   ******  * ** ** ********** ***

NS3-NS5B_WT       ACCATCTTGACAGACTTCAAAAATTGGCTGACCTCTAAATTGTTCCCCAAGCTGCCCGGC
NS3-NS5B_SCR      ACCATCCTCACGGACTTCAAAAACTGGCTGACCAGTAAGCTCTTCCCCAAACTCCCGGGG
                  ****** * ** ********** *********  ***  * ******** ** ** **

NS3-NS5B_WT       CTCCCCTTCATCTCTTGTCAAAAGGGGTACAAGGGTGTGTGGGCCGGCACTGGCATCATG
NS3-NS5B_SCR      CTTCCCTTTATCTCTTGTCAGAAGGGCTATAAAGGGGTGTGGGCGGGGACAGGTATCATG
                  ** ***** *********** ***** ** ** ** ******** ** ** ** ******

NS3-NS5B_WT       ACCACGCGCTGCCCTTGCGGCGCCAACATCTCTGGCAATGTCCGCCTGGGCTCTATGAGG
NS3-NS5B_SCR      ACAACCCGGTGCCCCTGTGGGGCGAACATCTCTGGGAACGTGAGGCTGGGGAGCATGCGA
                  ** ** ** ***** ** ** ** *********** ** **  * *****    *** *
```

```
NS3-NS5B_WT       ATCACAGGGCCTAAAACCTGCATGAACACCTGGCAGGGGACCTTTCCTATCAATTGCTAC
NS3-NS5B_SCR      ATTACTGGCCCTAAGACTTGCATGAACACGTGGCAAGGCACGTTCCCAATCAACTGCTAC
                  ** ** ** ***** ** ********** ***** ** ** ** ** ***** ******

NS3-NS5B_WT       ACGGAGGGCCAGTGCGCGCCGAAACCCCCCACGAACTACAAGACCGCCATCTGGAGGGTG
NS3-NS5B_SCR      ACCGAAGGCCAATGCGCCCCCAAGCCTCCCACCAACTACAAAACCGCAATATGGCGAGTA
                  ** ** ***** ***** ** ** ** ***** ******** ***** ** *** * **

NS3-NS5B_WT       GCGGCCTCGGAGTACGCGGAGGTGACGCAGCATGGGTCGTACTCCTATGTAACAGGACTG
NS3-NS5B_SCR      GCAGCTTCAGAATATGCAGAGGTGACCCAGCACGGCTCATATTCCTACGTTACGGGCTTG
                  ** ** ** ** ** ** ******** ***** ** ** ** ***** ** ** **  **

NS3-NS5B_WT       ACCACTGACAATCTGAAAATTCCTTGCCAACTACCTTCTCCAGAGTTTTTCTCCTGGGTG
NS3-NS5B_SCR      ACAACCGACAACCTAAAGATCCCCTGCCAACTGCCCTCGCCCGAGTTTTTCTCCTGGGTA
                  ** ** ***** ** ** ** ** ******** ** ** ** ****************

NS3-NS5B_WT       GACGGTGTGCAGATCCATAGGTTTGCACCCACACCAAAGCCGTTTTTCCGGGATGAGGTC
NS3-NS5B_SCR      GATGGCGTGCAAATACACCGCTTCGCGCCCACGCCCAAGCCTTTCTTTAGAGACGAGGTG
                  ** ** ***** ** **  * ** ** ***** ** ***** ** **  * ** *****

NS3-NS5B_WT       TCGTTCTGCGTTGGGCTTAATTCCTATGCTGTCGGGTCCCAGCTTCCCTGTGAACCTGAG
NS3-NS5B_SCR      TCTTTCTGCGTCGGCCTCAATTCCTACGCGGTAGGCAGTCAGCTCCCATGCGAGCCCGAG
                  ** ******** ** ** ******** ** ** **    ***** ** ** ** ** ***

NS3-NS5B_WT       CCCGACGCAGACGTATTGAGGTCCATGCTAACAGATCCGCCCCACATCACGGCGGAGACT
NS3-NS5B_SCR      CCCGACGCCGACGTGCTGCGATCCATGCTGACCGACCCCCCGCACATAACAGCAGAGACA
                  ******** *****  ** * ******** ** ** ** ** ***** ** ** *****

NS3-NS5B_WT       GCGGCGCGGCGCTTGGCACGGGGATCACCTCCATCTGAGGCGAGCTCCTCAGTGAGCCAG
NS3-NS5B_SCR      GCAGCGCGCAGACTAGCGAGAGGCTCCCCTCCCTCCGAGGCCAGTAGCTCAGTCTCTCAA
                  ** *****  *  * **  * ** ** ***** ** ***** **   ******    **

NS3-NS5B_WT       CTATCAGCACCGTCGCTGCGGGCCACCTGCACCACCCACAGCAACACCTATGACGTGGAC
NS3-NS5B_SCR      TTGTCAGCACCGTCTCTCAGAGCCACGTGCACAACCCATTCTAATACATACGACGTCGAT
                   * ************ **  * ***** ***** *****    ** ** ** ***** **

NS3-NS5B_WT       ATGGTCGATGCCAACCTGCTCATGGAGGGCGGTGTGGCTCAGACAGAGCCTGAGTCCAGG
NS3-NS5B_SCR      ATGGTCGACGCTAACCTCCTTATGGAAGGAGGAGTTGCCCAAACTGAACCAGAATCTCGG
                  ******** ** ***** ** ***** ** ** ** ** ** ** ** ** ** **  **

NS3-NS5B_WT       GTGCCCGTTCTGGACTTTCTCGAGCCAATGGCCGAGGAAGAGAGCGACCTTGAGCCCTCA
NS3-NS5B_SCR      GTCCCCGTCCTAGACTTTCTCGAGCCAATGGCAGAGGAGGAGTCAGACCTGGAGCCTTCC
                  ** ***** ** ******************** ***** ***   ***** ***** **

NS3-NS5B_WT       ATACCATCGGAGTGCATGCTCCCCAGGAGCGGGTTTCCACGGGCCTTACCGGCTTGGGCA
NS3-NS5B_SCR      ATCCCTAGTGAGTGCATGTTGCCGCGCTCAGGCTTCCCCAGAGCACTTCCCGCTTGGGCC
                  ** **    ********* * **  *    ** ** **  * **  * ** ********

NS3-NS5B_WT       CGGCCTGACTACAACCCGCCGCTCGTGGAATCGTGGAGGAGGCCAGATTACCAACCGCCC
NS3-NS5B_SCR      AGACCGGATTATAACCCCCCTCTGGTCGAGTCCTGGCGCCGCCCAGACTACCAGCCTCCC
                   * ** ** ** ***** ** ** ** ** ** *** *  * ***** ***** ** ***

NS3-NS5B_WT       ACCGTTGCTGGTTGTGCTCTCCCCCCCCCCAAGAAGGCCCCGACGCCTCCCCCAAGGAGA
NS3-NS5B_SCR      ACTGTGGCGGGCTGCGCCCTGCCCCCGCCTAAGAAGGCGCCTACCCCGCCACCTCGCCGG
                  ** ** ** ** ** ** ** ***** ** *******      ***** * **  *  *

NS3-NS5B_WT       CGCCGGACAGTGGGTCTGAGCGAGAGCACCATATCAGAAGCCCTCCAGCAACTG
NS3-NS5B_SCR      AGGAGAACTGTGGGGCTGTCAGAAAGCACCATATCCGAGGCGCTGCAACAGCTT
                   *  * ** ***** ***   ** ********** ** ** ** ** ** **

NS3-NS5B_WT       GCCATCAAGACCTTTGGCCAGCCCCCCTCGAGCGGTGATGCAGGCTCGTCCACGGGGGCG
NS3-NS5B_SCR      GCAATAAAAACATTTGGCCAACCTCCCTCCAGTGGGGATGCGGGCTCCTCCACAGGAGCT
                  ** ** ** ** ******** ** ***** ** ** ***** ***** ***** ** **

NS3-NS5B_WT       GGCGCCGCCGAATCCGGCGGTCCGACGTCCCCTGGTGAGCCGGCCCCCTCAGAGACAGGT
NS3-NS5B_SCR      GGCGCGGCCGAATCCGGCGGTCCGACGTCCCCGGGGGAGCCCGCACCTTCCGAGACTGGT
                  ***** ************************** ** ***** ** ** ** ***** ***
```

```
NS3-NS5B_WT       TCCGCCTCCTCTATGCCCCCCCTCGAGGGGGAGCCTGGAGATCCGGACCTGGAGTCTGAT
NS3-NS5B_SCR      TCCGCCTCGTCCATGCCGCCACTGGAAGGTGAGCCAGGCGACCCAGATTTAGAATCGGAC
                  ******** ** ***** ** ** ** ** ***** ** ** ** **  * ** ** **

NS3-NS5B_WT       CAGGTAGAGCTTCAACCTCCCCCCCAGGGGGGGGGGGTAGCTCCCGGTTCGGGCTCGGGG
NS3-NS5B_SCR      CAAGTGGAGCTGCAGCCTCCACCTCAGGGTGGTGGTGTCGCCCCCGGGTCCGGGAGCGGC
                  ** ** ***** ** ***** ** ***** ** ** ** ** ***** ** **    **

NS3-NS5B_WT       TCTTGGTCTACTTGCTCCGAGGAGGACGATACCACCGTGTGCTGCTCCATGTCATACTCC
NS3-NS5B_SCR      TCATGGTCTACTTGCTCGGAGGAAGACGACACAACCGTCTGCTGCAGCATGTCTTACTCC
                  ** ************** ***** ***** ** ***** ******  ****** ******

NS3-NS5B_WT       TGGACCGGGGCTCTAATAACTCCCTGTAGCCCCGAAGAGGAAAAGTTGCCAATCAACCCT
NS3-NS5B_SCR      TGGACCGGTGCCCTTATCACCCCATGCTCTCCTGAGGAGGAGAAATTGCCGATCAACCCG
                  ******** ** ** ** ** ** **    ** ** ***** ** ***** ********

NS3-NS5B_WT       TTGAGTAACTCGCTGTTGCGATACCATAACAAGGTGTACTGTACAACATCAAAGAGCGCC
NS3-NS5B_SCR      CTCTCCAACTCGCTCCTCAGGTATCACAATAAGGTGTATTGTACAACCTCCAAGTCCGCG
                   *    ********  *  * ** ** ** ******** ******** ** ***  ***

NS3-NS5B_WT       TCACAGAGGGCTAAAAAGGTAACTTTTGACAGGACGCAAGTGCTCGACGCCCATTATGAC
NS3-NS5B_SCR      TCCCAGCGGGCCAAGAAGGTCACCTTCGACCGTACCCAGGTCTTGGACGCTCACTACGAC
                  ** *** **** ** ***** ** ** *** * ** ** **  * ***** ** ** ***

NS3-NS5B_WT       TCAGTCTTAAAGGACATCAAGCTAGCGGCTTCCAAGGTCAGCGCAAGGCTCCTCACCTTG
NS3-NS5B_SCR      AGCGTGTTGAAGGACATAAAGCTGGCCGCCTCTAAGGTGTCCGCCCGCCTACTTACTCTC
                     ** ** ******** ***** ** ** ** *****   ***  * ** ** **  *

NS3-NS5B_WT       GAGGAGGCGTGCCAGTTGACTCCACCCCATTCTGCAAGATCCAAGTATGGATTCGGGGCC
NS3-NS5B_SCR      GAGGAGGCCTGCCAGCTCACTCCACCCCACTCCGCCCGGTCGAAGTACGGATTTGGAGCG
                  ******** ****** * ********** ** **  * ** ***** ***** ** **

NS3-NS5B_WT       AAGGAGGTCCGCAGCTTGTCCGGGAGGGCCGTTAACCACATCAAGTCCGTGTGGAAGG
NS3-NS5B_SCR      AAGGAGGTGAGGTCACTTAGTGGCCGCGCCGTCAACCATATCAAATCGGTCTGGAAGG
                  ********  *     *    **  * ***** ***** ***** ** ** *******

NS3-NS5B_WT       ACCTCCTGGAAGACCCACAAACACCAATTCCCACAACCATCATGGCCAAAAATGAGGTGT
NS3-NS5B_SCR      ACTTGTTGGAGGACCCCCAGACCCCCATCCCGACTACAATCATGGCAAAGAATGAGGTAT
                  ** *  **** ***** ** ** ** ** ** ** ** ******** ** ******** *

NS3-NS5B_WT       TCTGCGTGGACCCCGCCAAGGGGGGTAAGAAACCAGCTCGCCTCATCGTTTACCCTGACC
NS3-NS5B_SCR      TCTGCGTAGATCCGGCTAAGGGAGGCAAGAAGCCCGCGAGGCTAATCGTGTATCCCGATT
                  ******* ** ** ** ***** ** ***** ** **  * ** ***** ** ** **

NS3-NS5B_WT       TCGGCGTCCGGGTCTGCGAGAAAATGGCCCTCTATGACATTACACAAAAGCTTCCTCAGG
NS3-NS5B_SCR      TGGGGGTAAGAGTCTGTGAGAAGATGGCATTGTATGACATCACCCAGAAGCTCCCCCAGG
                  * ** **  * ***** ***** *****  * ******** ** ** ***** ** ****

NS3-NS5B_WT       CGGTAATGGGAGCTTCCTATGGCTTCCAGTACTCCCCTGCCCAACGGGTGGAGTATCTCT
NS3-NS5B_SCR      CCGTCATGGGGGCCTCGTATGGATTTCAGTATTCTCCCGCCCAGAGAGTGGAATACCTCT
                  * ** ***** ** ** ***** ** ***** ** ** *****  * ***** ** ****

NS3-NS5B_WT       TGAAAGCATGGGCGGAAAAGAAGGACCCCATGGGTTTTTCGTATGATACCCGATGCTTCG
NS3-NS5B_SCR      TGAAGGCCTGGGCCGAGAAGAAGGACCCCATGGGCTTTAGTTACGACACAAGGTGCTTTG
                  **** ** ***** ** ***************** ***   ** ** **  * ***** *

NS3-NS5B_WT       ACTCAACCGTCACTGAGAGAGACATCAGGACCGAGGAGTCCATATACCAGGCCTGCTCCC
NS3-NS5B_SCR      ACAGCACGGTAACTGAGCGGGACATACGTACGGAGGAGTCTATATATCAAGCCTGCTCCC
                  **   ** ** ****** * *****  * ** ******** ***** ** **********

NS3-NS5B_WT       TGCCCGAGGAGGCCCGCACTGCCATACACTCGCTGACTGAGAGACTTTACGTAGGAGGGC
NS3-NS5B_SCR      TACCCGAGGAGGCTCGCACAGCCATTCACTCCCTCACAGAGAGGCTCTATGTGGGGGGGC
                  * *********** ***** ***** ***** ** ** ***** ** ** ** ** ****

NS3-NS5B_WT       CCATGTTCAACAGCAAGGGTCAAACCTGCGGTTACAGACGTTGCCGCGCCAGCGGGGTGC
NS3-NS5B_SCR      CCATGTTTAATTCCAAAGGCCAGACGTGTGGCTACCGACGTTGCAGGGCTAGCGGGGTGC
                  ******* **   *** ** ** ** ** ** *** ******** * ** **********

NS3-NS5B_WT       TAACCACTAGCATGGGTAACACCATCACATGCTATGTGAAAGCCCTAGCGGCCTGCAAGG
NS3-NS5B_SCR      TCACCACTAGCATGGGGAACACAATTACCTGTTACGTGAAGGCGCTGGCCGCCTGCAAAG
                  * *************** ***** ** ** ** ** ***** ** ** ** ******** *
```

```
NS3-NS5B_WT       CTGCGGGGATAGTTGCGCCCACAATGCTGGTATGCGGCGATGACCTAGTAGTCATCTCAG
NS3-NS5B_SCR      CCGCGGGAATCGTTGCCCCAACGATGCTAGTCTGTGGGGACGACCTCGTGGTAATCTCTG
                  * ***** ** ***** ** ** ***** ** ** ** ** ***** ** ** ***** *

NS3-NS5B_WT       AAAGCCAGGGGACTGAGGAGGACGAGCGGAACCTGAGAGCCTTCACGGAGGCCATGACCA
NS3-NS5B_SCR      AGTCCCAAGGAACAGAAGAAGATGAAAGAAATCTCCGGGCCTTCACAGAAGCGATGACTC
                  *   *** ** ** ** ** ** **  * ** **  * ******** ** ** *****

NS3-NS5B_WT       GGTACTCTGCCCCTCCTGGTGATCCCCCCAGACCGGAATATGACCTGGAGCTAATAACAT
NS3-NS5B_SCR      GCTACTCCGCCCCACCCGGGGACCCCCCCAGGCCTGAGTACGATTTGGAGCTGATCACCT
                  * ***** ***** ** ** ** ******** ** ** ** **  ******* ** ** *

NS3-NS5B_WT       CCTGTTCCTCAAATGTGTCTGTGGCGTTGGGCCCGCGGGGCCGCCGCAGATACTACCTGA
NS3-NS5B_SCR      CATGCTCATCTAACGTTTCGGTAGCCCTGGGGCCTAGAGGACGGAGGCGCTATTATTTAA
                  * ** ** ** ** ** ** ** **  **** **  * ** **  *  * ** **  * *

NS3-NS5B_WT       CCAGAGACCCAACCACTCCACTCGCCCGGGCTGCCTGGGAAACAGTTAGACACTCCCCTA
NS3-NS5B_SCR      CACGGGACCCGACCACTCCACTGGCACGCGCGGCGTGGGAGACCGTCCGGCATAGCCCCA
                  *  * ***** ********** ** ** ** ** ***** ** **  * **   *** *

NS3-NS5B_WT       TCAATTCATGGCTGGGAAACATCATCCAGTATGCTCCAACCATATGGGTTCGCATGGTCC
NS3-NS5B_SCR      TCAACAGCTGGCTAGGGAATATAATCCAGTACGCGCCGACCATCTGGGTGAGGATGGTGC
                  ****    ***** ** ** ** ******** ** ** ***** *****  * ***** *

NS3-NS5B_WT       TAATGACACACTTCTTCTCCATTCTCATGGTCCAAGACACCCTGGACCAGAACCTCAACT
NS3-NS5B_SCR      TCATGACCCACTTCTTTTCAATTCTGATGGTACAGGACACTCTAGATCAGAACCTCAATT
                  * ***** ******** ** ***** ***** ** ***** ** ** ********** *

NS3-NS5B_WT       TTGAGATGTATGGATCAGTATACTCCGTGAATCCTTTGGACCTTCCAGCCATAATTGAGA
NS3-NS5B_SCR      TCGAGATGTACGGCTCGGTCTACTCCGTGAACCCCCTTGATCTCCCAGCCATCATAGAGC
                  * ******** ** ** ** ********** **  * ** ** ******** ** ***

NS3-NS5B_WT       GGTTACACGGGCTTGACGCCTTTTCTATGCACACATACTCTCACCACGAACTGACGCGGG
NS3-NS5B_SCR      GCCTTCACGGCTTGGACGCCTTCAGCATGCACACCTATTCCCACCACGAGCTCACCAGAG
                  *  * *****  * ********    ******** ** ** ******** ** **  * *

NS3-NS5B_WT       TGGCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGGGTGTGGAAGAGTCGGGCTC
NS3-NS5B_SCR      TTGCGAGTGCTCTGCGCAAGCTCGGTGCCCCCCCCCTGAGGGTCTGGAAGTCCAGAGCTA
                  * **    ** **  * ** ** ** ** ** ***** ***** ******    * ***

NS3-NS5B_WT       GCGCAGTCAGGGCGTCCCTCATCTCCCGTGGAGGGAAAGCGGCCGTTTGCGGCCGATATC
NS3-NS5B_SCR      GGGCTGTCCGAGCCAGCCTCATCAGTAGGGGGGGCAAGGCAGCTGTGTGCGGGCGATATC
                  * ** *** * **   *******    * ** ** ** ** ** ** ***** *******

NS3-NS5B_WT       TCTTCAATTGGGCGGTGAAGACCAAGCTCAAACTCACTCCATTGCCGGAGGCGCGCCTAC
NS3-NS5B_SCR      TCTTCAATTGGGCCGTGAAGACAAAGTTGAAGCTAACACCCCTTCCCGAGGCGCGCCTAC
                  ************* ******** *** * ** ** ** **  * ** *************

NS3-NS5B_WT       TGGACTTATCCAGTTGGTTCACCGTCGGCGCCGGCGGGGGCGACATTTTTCACAGCGTGT
NS3-NS5B_SCR      TGGACTTATCCAGTTGGTTCACCGTCGGCGCCGGCGGGGGCGACATTTTTCACAGCGTGT
                  ************************************************************

NS3-NS5B_WT       CGCGCGCCCGACCCCGCTCATTACTCTTCGGCCTACTCCTACTTTTCGTAGGGGTAGGCC
NS3-NS5B_SCR      CGCGCGCCCGACCCCGCTCATTACTCTTCGGCCTACTCCTACTTTTCGTAGGGGTAGGCC
                  ************************************************************

NS3-NS5B_WT       TCTTCCTACTCCCCGCTCGGTAG
NS3-NS5B_SCR      TCTTCCTACTCCCCGCTCGGTAG
                  ***********************
```

6.2.4 Establishment of RT-qPCR for HCV plus and minus strand detection

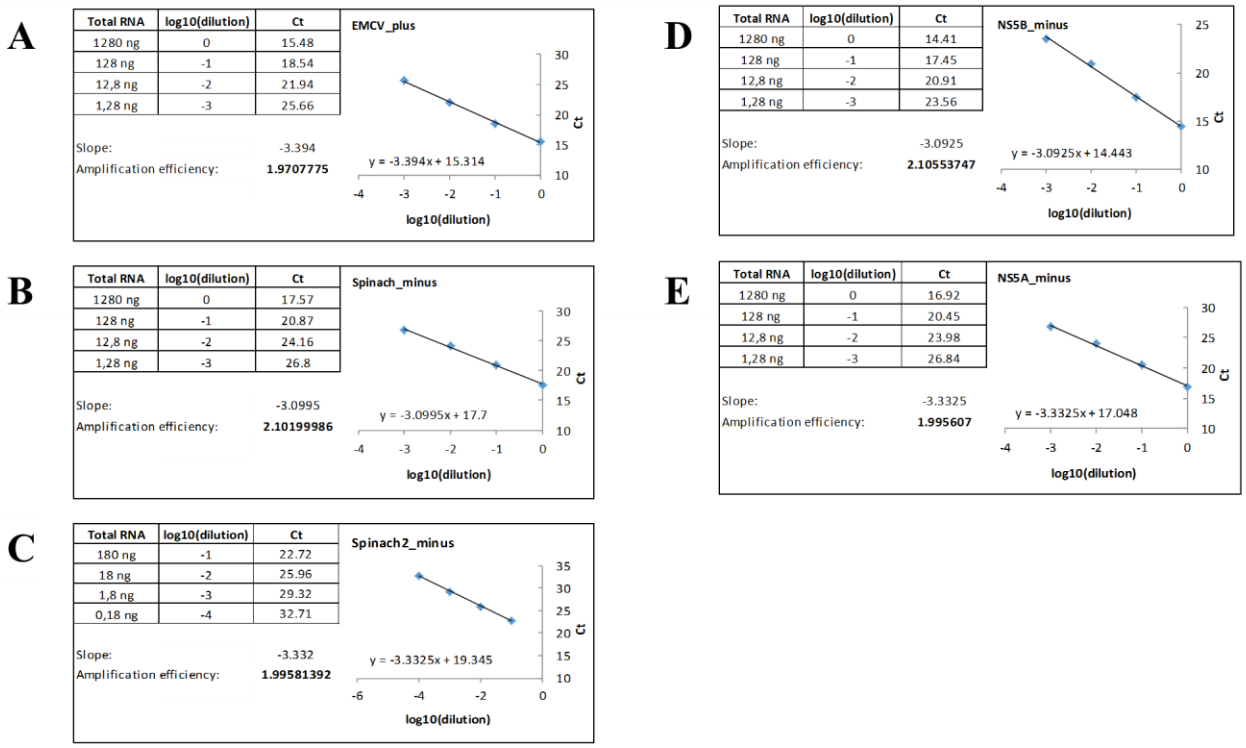
Canonical end-point PCR method appears poorly quantitative since nucleotides and primers are gradually becoming limited providing a nonlinear correlation between the starting copy number and the final yield of the amplified product. Recently developed approaches for simultaneous DNA amplification and concentration tracking enabled a reliable and reproducible quantification of nucleic acids. Quantitative PCR (qPCR) method is based on detection and quantification of a fluorescent reported signal which linearly increases in relation to the amount of DNA in the PCR reaction. Using a fluorescence detecting thermocycler, fluorescence emission is monitored at each cycle during the exponential phase when none of the components are limited and initial amounts of template correlate with an increase of the PCR product amounts. Moreover, advantageous high sensitivity of fluorescent detection empowers DNA quantification in a wide dynamic range (up to nine orders of magnitude). Combined with reverse transcription, qPCR is applied for quantification of mRNA for evaluation of gene expression and even for microRNA. The detection in qPCR may exploit different chemistries: using DNA-binding dyes, probe-based chemistry or quenched dye primers. In the present work all assays were conducted using DNA-binding SYBR Green I dye.

All fluorogenic DNA-binding dyes bind to any double-stranded DNA in reaction mixture in a sequence-independent manner. Unbound dye demonstrates little fluorescence (1000-fold less) when in a solution and the emission intensity increases proportionally to the amplified DNA. SYBR Green I is currently one of the most optimal DNA-binding dyes due to its high sensitivity and reliability as well as relatively low cost that allows large number of samples. The lack of specificity to a target sequence provides some shortcomings such as detection of nonspecific reaction products or primer dimers. For this reason at the end of amplification a melting curve of the amplified DNA is generated to determine the melting temperature of product(s). Melting curves provide the information on homogeneity of the product and allows validating it by the specific melting temperature (Tm) (from Sambrook, vol. 2, fourth edition).

Fluorescence detection during amplification is represented by a sigmoidal amplification plot where for the first 10-20 cycles the curve is flat until the fluorescent signal becomes detectable; then the curve becomes linear for several cycles representing exponential phase of the reaction. Eventually, reagents become limited and the curve plateaus. The initial concentration of a target is expressed as a cycle number required to reach a certain threshold of amplification (Ct). The threshold reflects a statistically significant level of fluorescence over the background and usually automatically programmed. Two methods are available for the analysis of derived Ct values: the ΔΔCt method and the standard curve method (used in the present work).

For the standard curve method the amounts of target and reference genes in the calibrator and experimental sample are first determined using a standard curve, followed by the target gene normalization to the reference gene in both samples. When only the relative quantification is conducted, the target amount in the experimental sample is compared to that in the calibrator sample; while for the absolute quantification the Ct value of the test is compared to values of standards with known concentration plotted on a standard curve. Each target/reference gene requires construction of a separate standard curve. To ensure reliability of quantification, the standard curve should be constructed in a dynamic range wider than the range of expected concentrations.

The present research deals only with relative quantification; therefore the plot of a standard curve is represented by Ct value versus log of a dilution factor. The standard curves were generated for each of the primer pairs targeting HCV minus strand (Spinach_minus, Spinach2_minus, NS5A_minus and NS5B_minus) as well as for the HCV plus strand (EMCV_plus) (Suppl. Fig. 6.7). According to the main focus of the work on the HCV minus strand synthesis initiation, input plus strands during transfection served as a reference to normalize the relative abundance of minus strands targeted at either of the above mentioned regions. Normalization to the plus strand content is more reasonable than to one of the cellular genes since it permits to eliminate discrepancies related to lipofection efficiency in individual samples and to an ability of selected replication templates to undergo continuous replication. The standard curves for each



Supplementary Figure 6.7: Standard curves and calculation of qPCR primers amplification efficiency.
(A-E) Standard curves were constructed for dilution series of total RNA fraction from HuH-7.5 cells transfected with the HCV SL I-III replication construct (2 days post transfection). The reference region is the EMCV IRES sequence on input plus strands (A); the target regions are reverse complementary sequences of the Spinach region (two variants, B and C) and the HCV NS5A/B coding regions (D and E, respectively) – on newly synthesized minus strands. In tables total RNA amounts in each serial dilution are provided. X-axis is a log dilution factor, Y-axis is a mean Ct value (also given in tables). Amplification efficiencies are calculated according to obtained "slope" values for each primer pair as explained in the text. Each data point represents a mean value from technical triplicates.

target/reference gene provide important information on amplification efficiency (E) for each of the qPCR primer pairs, which is calculated according to the equation: $E = 10^{(-1/slope)}$. All derived curves (Suppl. Fig. 6.7, A-E) were generated using a total RNA sample from HuH-7.5 cells transfected with the SL I-III construct (2 days post transfection). In scope of the current work this construct displayed the highest ability for minus strand synthesis. Serial 10-fold dilutions of this total RNA sample were subjected to reverse transcription and qPCR in a manner similar to all experimental samples. Importantly, following a dilution of the total RNA sample, cellular (mock-transfected) total RNA was utilized to equalize total RNA content in each downstream dilution sample. The linearity of plots within the present dynamic range indicates the reliable quantification of all targets with designed primers and enables determination of corresponding amplification efficiencies:

EMCV_plus: ≈ 1.97 (dynamic range 15.48 to 25.66);

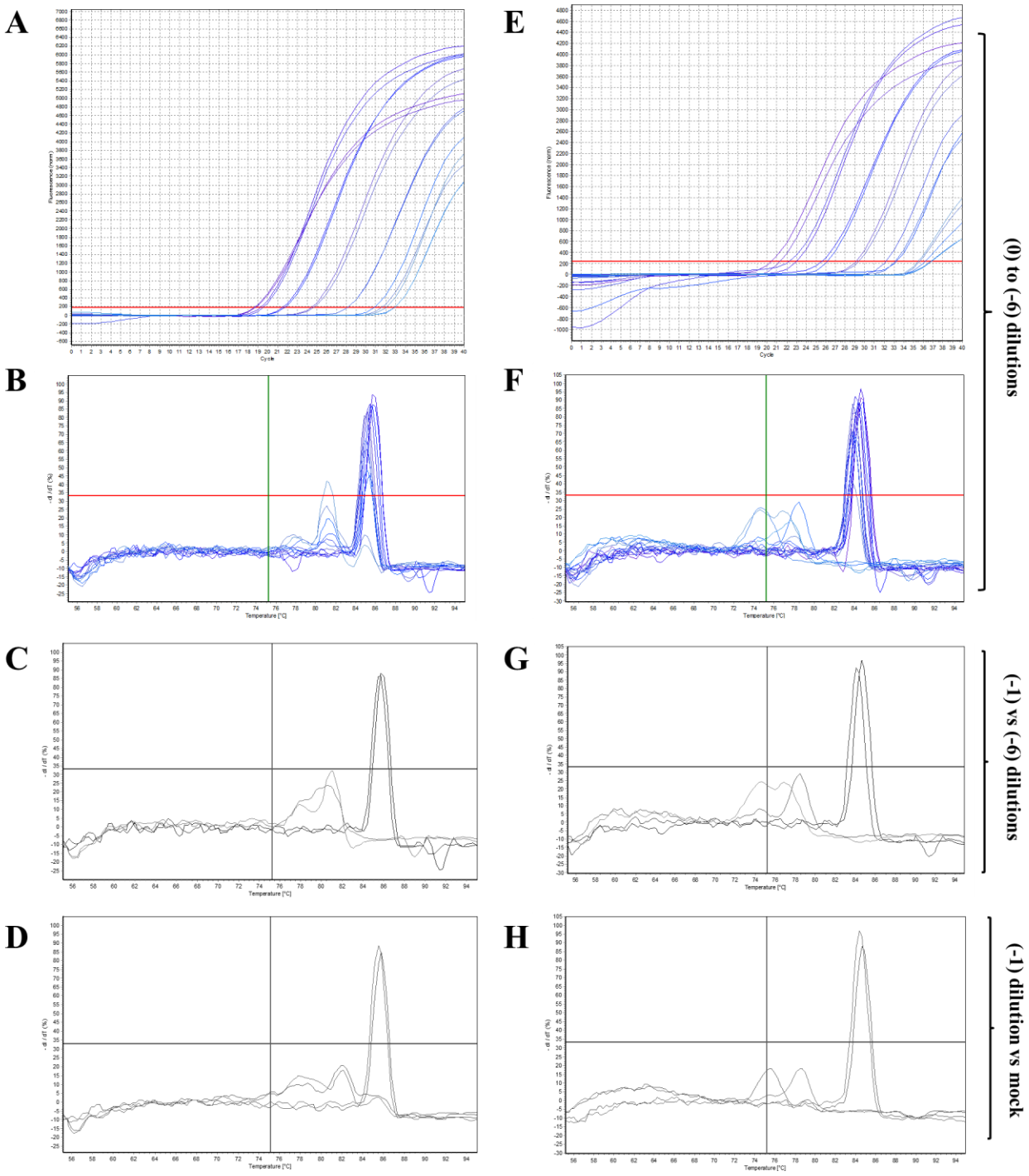
Spinach_minus: ≈ 2.10 (dynamic range 17.57 to 26.8);

Spinach2_minus: ≈ 1.99 (dynamic range 22.72 to 32.71);

NS5A_minus: ≈ 1.99 (dynamic range 16.92 to 26.84);

NS5B_minus: ≈ 2.10 (dynamic range 14.41 to 23.56).

An illustration of amplification plots derived in a course of standard curves generation is presented on Suppl. Fig. 6.8. Serial 10-fold dilutions (up to (-6)) of experimental total RNA, as well as the mock-transfected total RNA in equal amounts, were subjected to RT-qPCR targeting the EMCV region on the genome strand (EMCV_plus; Suppl. Fig. 6.8, A) and the Spinach region on the complementary minus strand (Spinach2_minus; Suppl. Fig. 6.8, E). As evident from the plots, linearity of standard curves becomes



Supplementary Figure 6.8: Amplification plots and melting curve analysis.
(A, B) Amplification plots illustrating quantification of target EMCV (A) and Spinach (E) regions on plus and minus strands of the HCV SL I-III construct, respectively, upon RT-qPCR of total RNA dilution series (as in Suppl. Fig. 6.7 and detailed in 6.2.4).
(B, F) Corresponding melting curve plots are illustrated for the amplification plots from A and E, respectively.
(C, G) Melting curve plots comparing the serial dilutions from A and E: melting curves for dilutions (-1) and (-6). Technical duplicates are provided for each cDNA species.
(D, H) Melting curve plots comparing the serial dilution (-1) (from A and E, respectively) to mock-transfected sample. Technical duplicates are provided for each cDNA species.
X-axis is temperature in °C, Y-axis is a change in fluorescence with increasing temperature.

inaccurate for undiluted total RNA (corresponds to around 1000 ng total RNA per RT, which is an amount recommended by the manufacturer), and is not mediated by generally high total RNA content, since it was maintained for all dilutions by cellular total RNA. For both target regions, starting from (-1) dilutions (corresponds to around 100 ng target-containing total RNA per RT) the linearity is fulfilled. On the other hand, (-6) (and to a smaller extent (-5)) dilution does not contain enough target HCV RNAs: the Ct values

reach the ones for cellular total RNA indicating the background level, which represents the border of a dynamic range. The latter can in turn be visualized by a melting curve analysis: high specificity is indicated by the presence of one single, sharp peak. Indeed, for the two lowest dilutions – (-5) and (-6) – profiles of melting curves significantly alter and point out an absence of a specific target for both, EMCV and Spinach, regions (Suppl. Fig. 6.8, B and F, respectively). Separated comparison of melting curves, derived for the (-1) dilution of experimental RNA sample with its (-6) dilution (Suppl. Fig. 6.8, C and G) and with a mock-transfected cellular total RNA (Suppl. Fig. 6.8, D and H) verifies the unspecific nature of Ct values originating from samples with no or very low content of target molecules.

Selection of the primers for RT-qPCR with optimal target-specificity is generally a challenge in establishing quantitative PCR protocols. Newly-designed primers should be characterized by low specificity to cellular genome and transcriptome and especially to a viral RNA strand complementary to a target strand. The latter ensures the strand-specificity of RT-qPCR assays aiming to distinguish between genome and antigenome strand synthesis initiation. All primers for detection of either of HCV strands lack targets on opposite strands (Suppl. Fig. 6.9): each primer has minimum 5 mismatches and for most of the primers – 8 to 10 mismatches (analysis was conducted in Clustal Omega). In turn, BlastN searches against the human genome and transcriptome using the primer sequences and their reverse complements yielded smaller amount of mismatches (minimum 3 to 4). However, in all assays mock-transfected cellular total RNA exclusively resulted in very high Ct values and melting curve plots implicating the absence of any cross-reactivity (Suppl. Fig. 6.8, D and H).

Notably, the background signals in minus strand quantification often appear higher within samples corresponding to early time-points (up to 24 h) than at later time-points (around 96 h) (Fig. 3.6.1, B and C). This phenomenon is likely explained by an aberrant activity of the T7 RNA polymerase in initiation of transcription from a promoter-less end of template DNA (Mu et al. 2018). Such activity results in production of an antisense complementary RNA that, unlike another potential source of background signals – DNA transcription templates, is insensitive to DNase treatment. Upon transfection of *in vitro* transcribed RNA constructs this small fraction of antigenome strands is detected by minus strand specific primers. An elimination of these artifacts is related to subsequent degradation of "transfected" minus strands and detection of genuine minus strands, if they can be synthesized from a construct. Another possibility – a cross-reactivity with HCV genomic RNA, which is 10-100 times more abundant than minus strands, unlikely to take place, since for replicating constructs dynamics of plus strands at later time-points does not correlate with dynamics of minus strands; on the contrary, plus and minus strands peak at dissimilar time during time-courses.

An attempt to further improve a strand-specificity of primers targeting HCV minus strands was made concerning the Spinach primers, as the most often utilized. In a fashion of "molecular beacon" a new RT primer was designed to self-hybridize at its 5'- and 3'- ends to form a short stem (see 2.1.8.2). The basic idea was to exploit this self-hybridization of the primer's ends at temperatures lower than Tm in order to impair the primers' hybridization to off-target sequences. Once this primer binds to a genuine target, the stem is distracted and a cDNA product is synthesized. Additional changes were introduced at a step of total RNA/primers hybridization (prior an addition of reverse transcriptase). Incubation under condition of gradual decrease from temperatures higher to temperatures slightly lower than Tm aimed at hybridization to the most sequence-specific targets. This approach enabled a reduction of background signals up to 10 times (data not shown).

```
>EMCV_plus_RT - 10 mismatches
plasmid_Ps_WT    832 CTTCTTGAAGAC    843
                     |||.|||||.||
EMCV_plus_RT       4 CTTGTTGAATAC     15

>RC_EMCV_plus_RT
plasmid_Ps_WT   1009 CAAGCGTATTCAACAAGGGG  1028
                     ||||||||||||||||||||
RC_EMCV_plus_      1 CAAGCGTATTCAACAAGGGG    20


>EMCV_plus_qPCR_for - 7 mismatches
plasmid_Ps_WT   2629 GAGCCTCAGGAATGTTTG   2646
                     ||.||..|||||||.|.|
EMCV_plus_qPC      2 GACCCCTAGGAATGCTCG     19

>RC_EMCV_plus_qPCR_for
plasmid_Ps_WT    747 ACGAGCATTCCTAGGGGTCT  766
                     ||||||||||||||||||||
RC_EMCV_plus_      1 ACGAGCATTCCTAGGGGTCT   20


>EMCV_plus_qPCR_rev
plasmid_Ps_WT    557 CCGTCCTTCACCATTTCATT  576
                     ||||||||||||||||||||
EMCV_plus_qPC      1 CCGTCCTTCACCATTTCATT   20

>RC_EMCV_plus_qPCR_rev - 7 mismatches
plasmid_Ps_WT   4990 TGGAATCGTGGAGGAGG   5006
                     ||.|||.|||.||||.|
RC_EMCV_plus_      3 TGAAATGGTGAAGGACG     19


>Spinach_minus_RT
plasmid_Ps_WT    557 CCGTCCTTCACCATTTCATT  576
                     ||||||||||||||||||||
Spinach_minus      1 CCGTCCTTCACCATTTCATT   20

>RC_Spinach_minus_RT - 7 mismatches
plasmid_Ps_WT   4990 TGGAATCGTGGAGGAGG   5006
                     ||.|||.|||.||||.|
RC_Spinach_mi      3 TGAAATGGTGAAGGACG     19


>Spinach_minus_qPCR_for - 5 mismatches
plasmid_Ps_WT     50 GGAACTACTGTCTTCACG     67
                     ||||||.|||..|||||||
Spinach_minus      1 GGAACTGCTTCCTTCACG     18

>RC_Spinach_minus_qPCR_for
plasmid_Ps_WT    804 GTCGTGAAGGAAGCAGTTCC  823
                     ||||||||||||||||||||
RC_Spinach_mi      1 GTCGTGAAGGAAGCAGTTCC   20


>Spinach_minus_qPCR_rev
plasmid_Ps_WT    690 ACCATATTGCCGTCTTTTGG  709
                     ||||||||||||||||||||
Spinach_minus      1 ACCATATTGCCGTCTTTTGG   20

>RC_Spinach_minus_qPCR_rev - 6 mismatches
plasmid_Ps_WT   4340 CTACAAGACCGCCATCTGG  4358
                     |.|.||||||.||.||.|||
RC_Spinach_mi      1 CCAAAAGACGGCAATATGG    19


>NS5A_minus_qPCR_RT/for:
plasmid_Ps_WT   4123 GCCTCCCCTTCATCTCTTGT 4142
                     ||||||||||||||||||||
NS5A_minus_qP      1 GCCTCCCCTTCATCTCTTGT   20

>RC_NS5A_minus_qPCR_RT/for - 6 mismatches
plasmid_Ps_WT   1962 GTGATGACCGGGGAGGC   1978
                     |.||||||. ||||||||
RC_NS5A_minus      5 GAGATGAA-GGGGAGGC     20
```

```
>NS5A_minus_qPCR_rev - 5 mismatches
plasmid_Ps_WT   4361 GGTGGCGGCCTCGGAGTACG 4380
                     |.|||||.||.|.|||.||
NS5A_minus_qP      1 GATGGCGGTCTTGTAGTTCG   20


>RC_NS5A_minus_qPCR_rev:
plasmid_Ps_WT   4336 CGAACTACAAGACCGCCATC 4355
                     ||||||||||||||||||||
RC_NS5A_minus      1 CGAACTACAAGACCGCCATC   20


>NS5B_minus_RT:
plasmid_Ps_WT   5849 TGAGGTGTTCTGCGTGGAC  5867
                     ||||||||||||||||||||
NS5B_minus_RT      1 TGAGGTGTTCTGCGTGGAC    19

>RC_NS5B_minus_RT - 8 mismatches
plasmid_Ps_WT   4681 CACGGGGATCACCTC   4695
                     ||||..||.||||||
RC_NS5B_minus      4 CACGCAGAACACCTC     18


>NS5B_minus_qPCR_for:
plasmid_Ps_WT   7099 ACATTTTTCACAGCGTGTCG 7118
                     ||||||||||||||||||||
NS5B_minus_qP      1 ACATTTTTCACAGCGTGTCG   20

>RC_NS5B_minus_qPCR_for - 8 mismatches
plasmid_Ps_WT   7363 CGGCTAGCTGTGAAA   7377
                     ||.|..|||||||||
RC_NS5B_minus      1 CGACACGCTGTGAAA     15


>NS5B_minus_qPCR_rev - 5 mismatches
plasmid_Ps_WT   5408 TACC-ACCGTGTGCTGCTC 5425
                     |||| |..||||||.||||
NS5B_minus_qP      2 TACCTAGTGTGTGCCGCTC   20

>RC_NS5B_minus_qPCR_rev:

plasmid_Ps_WT   7198 AGAGCGGCACACACTAGGTAC 7218
                     ||||||||||||||||||||
RC_NS5B_minus      1 AGAGCGGCACACACTAGGTAC   21
```

Supplementary Figure 6.9: Analysis of RT-qPCR strand specificity.
Primers are matched to their target strand and to the opposite strand by using the reverse complement of each primer. Within a non-target opposite strand each primer has minimum 5 mismatches, some of the primers have up to 8-10 mismatches. Matching was conducted by Clustal Omega.

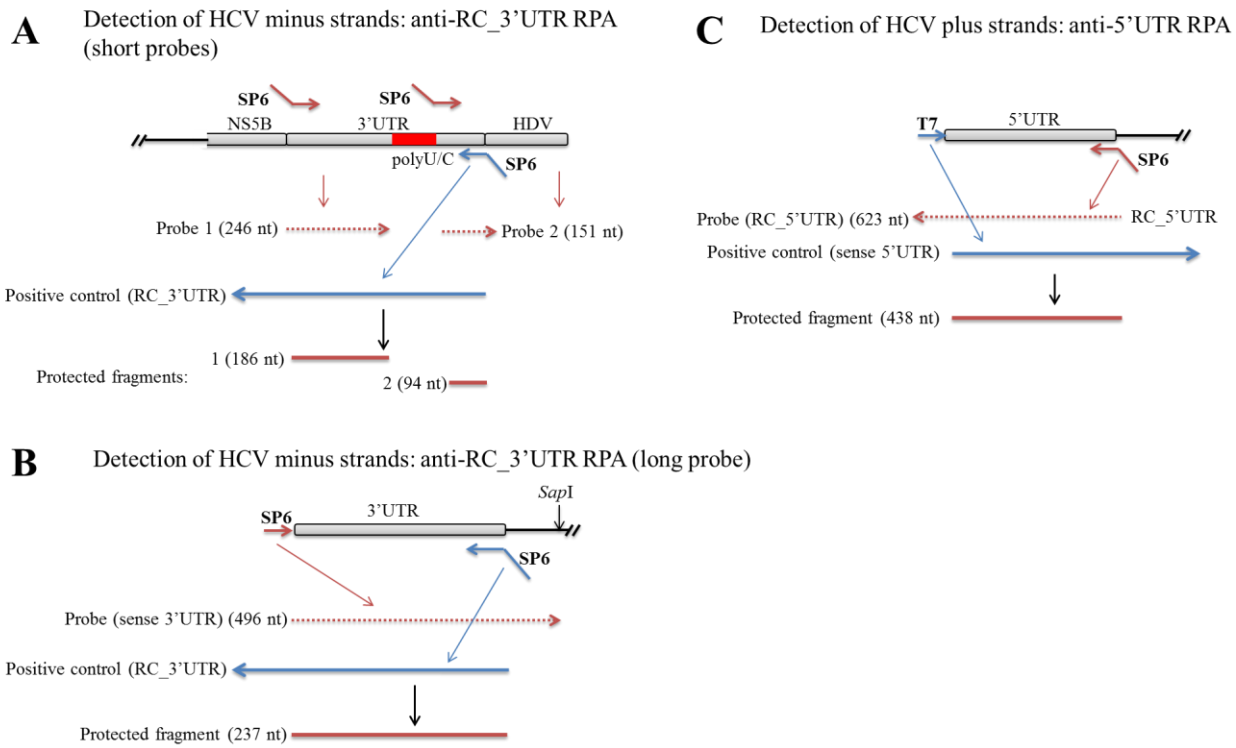
6.3 Supplementary results

6.3.1 Detection of HCV plus and minus strands by Ribonuclease Protection Assay

Prior an establishment of quantitative RT-PCR (see 6.2.4) for detection of HCV plus and minus strands upon transfection of studied replication RNA templates, a possibility of utilization of a classic ribonuclease protection assay (RPA) was challenged. RPA is a sensitive and target-specific technique for visualization and quantification of RNA of interest that allows detection of a target in a total cellular RNA pool (see 2.2.8). Principally, the technique is based on a hybridization of radioactively labeled probe to a target RNA and therefore on protection of a duplex from hydrolytic activity of single-strand specific ribonucleases (RNase), whereas all non-target RNA is being degraded. Visualization of the protected dsRNA fragment is conducted via a denaturing polyacrylamide gel electrophoresis followed by autoradiography. Commonly, a probe is designed to be not completely complementary to its target, providing a ssRNA stretch at one of the ends. The latter facilitates discrimination of genuine probe-target hybrids when compared to initial probe and prevents a misinterpretation of signals originating from possible probe self-hybridization or hybridization to other targets.

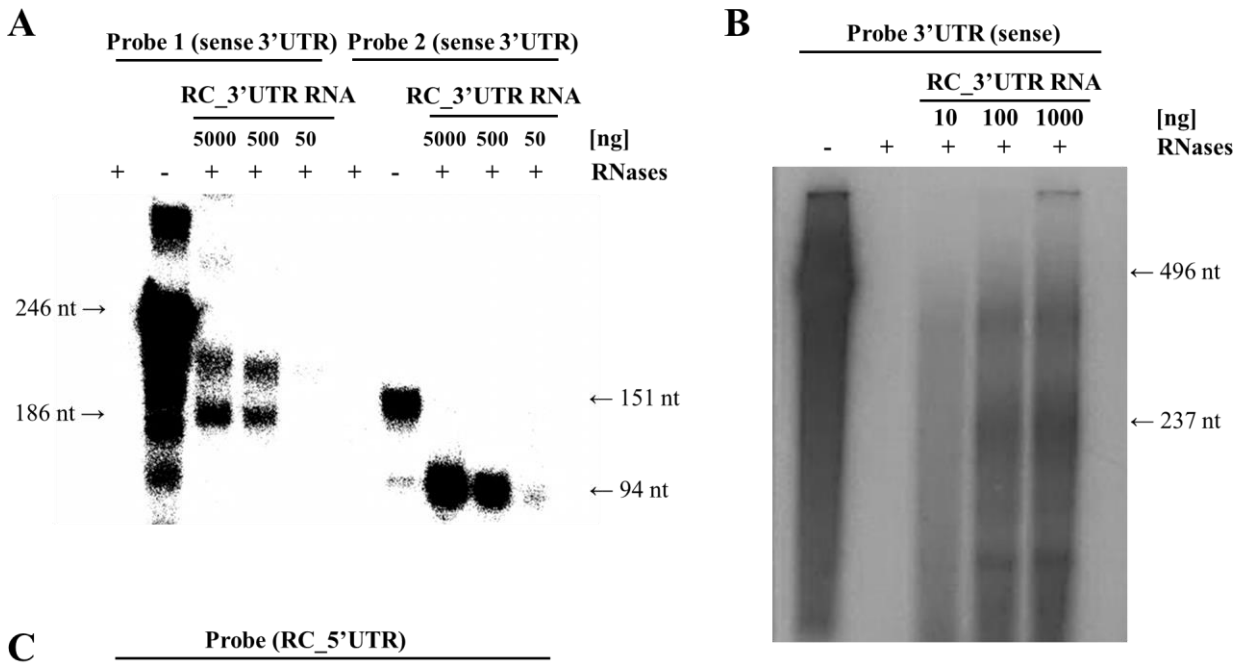
In scope of the presented in the current work experimental setup, it was initially attempted to develop a set of probes – specific to either of the HCV constructs' strands – that would be universal for all investigated constructs. Taking into consideration a series of constructs for minus strand synthesis initiation where most of an NS3-NS5B open reading frame was scrambled, probes targeting the reverse complementary (RC) of HCV 3'UTR were considered. Following a primary goal of the project in detection of the HCV minus strand synthesis from designed minimal (hp) and extended RNA constructs (e.g. 5'UTR), several probes were examined, at first, using control *in vitro* transcribed templates (as positive control) and then using total RNA isolated from cells transfected with experimental constructs (see below). Two approaches of targeting the HCV 3'UTR in RPA were considered. Regarding the presence of a long polyU/C tract within the HCV 3'UTR it was thought to potentially give rise to problems in full-length probe synthesis and therefore target-hybridization. Hence, two strategies were applied: a production of shorter probes that hybridize up- or down- stream the antisense 3'UTR sequence (Suppl. Fig. 6.10, A) and a utilization of a long 3'UTR-containing probe (Suppl. Fig. 6.10, B). All shorter probes, as well as control templates applied in preliminary RPA experiments, were *in vitro* transcribed from PCR templates generated using SP6 promoter provided at one of the primers (Suppl. Fig. 6.10, A and B). The long probe was generated from a plasmid encoding only the HCV 3'UTR from T7 promoter. In all pairs of probe/target RNA a part of the probe's sequence had no complementarity to a control template. Concerning input plus strands detection, a probe complementary to the sense HCV 5'UTR was designed (generated from using SP6 promoter) and a control template was *in vitro* transcribed from T7 promoter providing a sense HCV 5'UTR sequence (Suppl. Fig. 6.10, C). However, a detection of input RNA templates in minus strand initiation assays requires a distinct probe universal for all investigated constructs, since the complete HCV 5'UTR is not present in most of experimental constructs.

Applicability of designed probes for HCV minus and plus strand detection was challenged in RPA using a suitable (antisense or sense, respectively) *in vitro* generated RNA transcripts. In order to estimate a sensitivity of an assay for each of the probes, serial 10-fold dilutions of template RNAs were subjected to detection. In more detail, in an assay utilizing the shorter sense-3'UTR probes (Suppl. Fig. 6.10, A), dilutions of *in vitro* transcribed antisense HCV 3'UTR control template - 5000 ng to 50 ng – were applied (Suppl. Fig. 6.11, A). The probe hybridizing upstream the polyU/C tract appeared sensitive up to 50 ng of target RNA, whereas the probe hybridizing downstream the polyU/C tract remained sensitive even at 50 ng of control target. The latter probe also demonstrated preferable specificity resulting in a single protected RNA signal that corresponded to predicted length of shared by probe/target sequence; this probe was further applied in downstream assays. On the contrary, utilization of the long sense-3'UTR probe (Suppl. Fig. 6.10, B) did not result in specific binding and protection when combined with the control antisense HCV 3'UTR template (Suppl. Fig. 6.11, B). It can be speculated, that *in vitro* transcription across the polyU/C tract in HCV 3'UTR



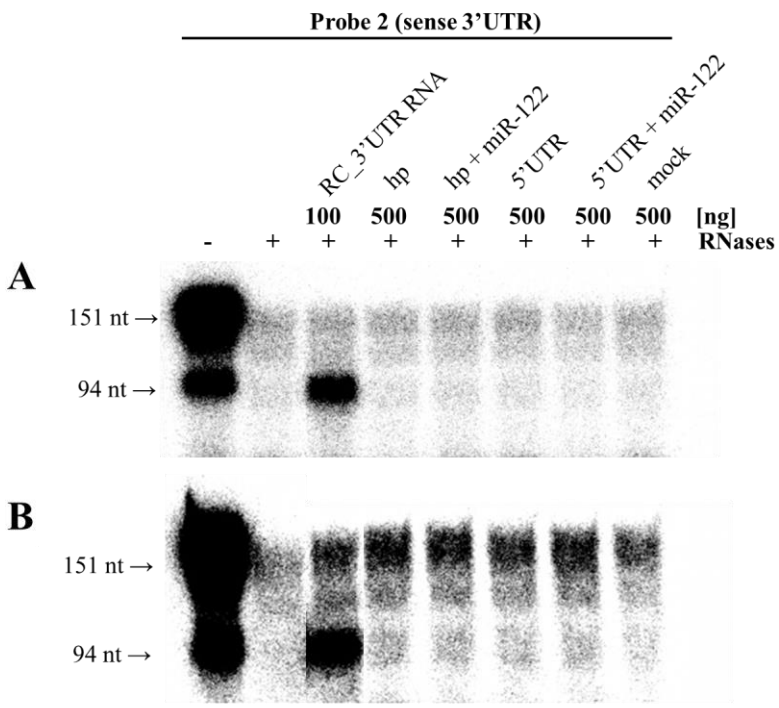
Supplementary Figure 6.10: Schematic representation of RNA probes and control templates engaged in ribonuclease protection assay (RPA).
(A) Detection of HCV minus strands using two short probes targeting the reverse complementary 3'UTR region (RC_3'UTR) up- and down- stream of the polyU/C tract. Probes and RC_3'UTR RNA (serving as a positive control) are synthesised by *in vitro* transcription from different DNA templates, derived by PCR in which one of the primers contains SP6 promoter. Protected fragments obtained in RPA with either of the probes are shorter than corresponding probes that allows to distinguish between a protected RNA and potentially undigested probe RNA.
(B) Detection of HCV minus strands using a long probe targeting the reverse complementary 3'UTR region (RC_3'UTR) across the polyU/C tract. The probe RNA is derived by *in vitro* transcription from SP6 promoter located directly upstream the HCV 3'UTR at the plasmid DNA (cleaved prior *in vitro* transcription at *Sap*I restriction site). The RC_3'UTR RNA (serving as a positive control) is generated by *in vitro* transcription from PCR DNA template and driven from SP6 promoter. An expected protected RNA fragment is shorter than the probe RNA.
(C) Detection of HCV plus strands using a long probe targeting the sense HCV 5'UTR region. The probe represent a corresponding antisense sequence (RC_5'UTR) and generated from SP6 promoter within a PCR DNA template. The positive control template is transcribed directly from T7 promoter situated on the 5'UTR-containing plasmid. It does not completely overlap with the probe RNA resulting in a shorter protected fragment.

results in imprecise transcripts of both probe and control template. This fact in turn does not affect detection by shorter probes, avoiding the 3'UTR and protected by the control template either up- or down- stream. At last, challenging of the antisense HCV 5'UTR probe in RPA with serial dilutions of a control sense 5'UTR template (Suppl. Fig. 6.11, C) confirmed specificity and sensitivity of the probe. It allowed detection of up to 1 ng of target, inclusively, however was not quantitative in a range 100-1000 ng of target RNA, possibly due to limiting probe molecules. It is virtually important to predetermine a protected RNA fragment to be shorter than a probe, since residual amounts of probe RNA may remain undigested after RNAses treatment (Suppl. Fig. 6.11, C) and result in false-positive results. This artifact may originate from retention of DNA transcription template during a probe synthesis and can be eliminated by usage of longer enough templates to be further reliably separated from a probe by gel-electrophoresis followed by probe RNA purification from a gel. Another possible origin of ssRNase resistant products in a "probe only" experimental sample is self-hybridization of probe RNA. Here, the assay for sense HCV 5'UTR detection (Suppl. Fig. 6.10, C) can be improved by a careful isolation of a probe RNA from its *in vitro* transcription reaction by extraction from a preparative gel.



Supplementary Figure 6.11: Ribonuclease protection assay establishment for HCV minus and plus strand detection.
Separation and visualization was conducted via a denaturing polyacrylamide gel electrophoresis followed by autoradiography. Suitable templates are generated by *in vitro* transcription and subjected to RPA in serial 10-fold dilutions. Addition of RNases is indicated by "+", amount of control RNA template is given in ng, identical amount of probe RNA is utilized in all samples within individual assays. Predicted size of radio-labeled probe and protected RNA fragments is noted and was estimated from migration of xylene cyanol and bromophenol blue dyes fronts in denaturing polyacrylamide gels.
(A) Detection of a control RC_3'UTR RNA template by short probes, as in Suppl. Fig. 6.10, A.
(B) Detection of a control RC_3'UTR RNA template by a long probe, as in Suppl. Fig. 6.10, B.
(C) Detection of a control sense HCV 5'UTR RNA template in RPA, as in Suppl. Fig. 6.10, C.

Providing that one of the shorter probes for HCV minus strand detection was functional in a preliminary assay with *in vitro* transcribed control template (Suppl. Fig. 6.11, B, right), it was utilized in attempt to detect HCV minus strand synthesis initiation in a replication assay. As in the majority of experiments presented in the current research, HuH-7.5 cells were transfected for two days with either a minimal "hp" replication template or with one of its extended versions, harboring a complete HCV 5'UTR at the construct's 5'-end (see Fig. 3.2.1, A). Additionally, each of the experimental constructs was alternatively supplemented with miR-122 in order to potentially facilitate RNA constructs' functions (Suppl. Fig. 6.12, A and B). An *in vitro* synthesized control template RNA used at the step of assay establishment (Suppl. Fig. 6.10, A) served as a positive control to indicate an expected length of the protected fragment. The Suppl. Fig. 6.12 illustrates two representative experiments, as described above. In contrast to the control template, experimental total RNA from transfected cells did not give rise to a protected fragment, indicating an absence of target RNA – reverse complementary HCV 3'UTR sequence – at a detection limit ensured by the assay. Taking into consideration that an increase of sample total RNA leads to elevation of background and also that HCV minus strands (unlike plus strands) are generally present in cells at very low abundance, the ribonuclease protection assay was found insufficiently sensitive for this application. As an alternative to



Supplementary Figure 6.12: Detection of HCV minus strand synthesis by RPA.
Ribonuclease protection assay was conducted using a short probe targeting the reverse complementary 3'UTR region (RC_3'UTR) downstream of the polyU/C tract (as in Suppl. Fig. 6.10, A, right). As a positive control, *in vitro* transcribed RNA containing the reverse complementary sequence of HCV 3'UTR was utilized. The experimental samples are represented by total cellular RNA isolated from HuH-7.5 cells transfected with hairpin (hp) and 5'UTR constructs (with or without additional miR-122) for two days.
(A) and **(B)** are representative biological replicates of the experiment. Probe RNA (before "-" and after "+" treatment with RNases) and protected fragments are visualized via a denaturing polyacrylamide gel electrophoresis followed by autoradiography. Control and total RNA amounts are shown in ng, identical amount of probe RNA is utilized in all samples. Predicted size of radio-labeled probe and protected RNA fragments is verified by migration of xylene cyanol and bromophenol blue dyes fronts in denaturing polyacrylamide gels.

RPA, quantitative RT-PCR (see 6.2.4) was established for evaluation of minus strand synthesis efficiency from different templates as well as of input plus strand construct RNA. Indeed, providing relatively small variations in minus strand abundance from replication-competent templates, it would have been challenging to obtain quantity data, similarly to another possible detection method – northern blot.

6.3.2 Aminoacyl-tRNA synthetases binding at the HCV 3'UTR

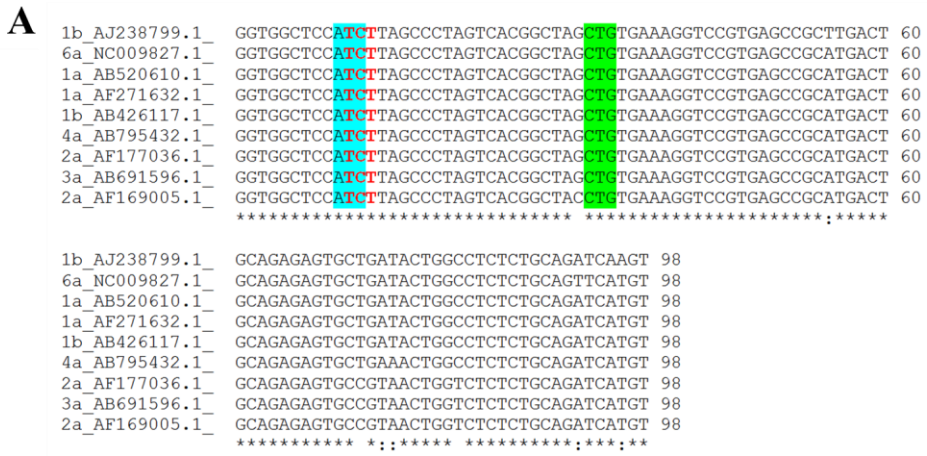
RNA viruses widely exploit host cell proteins in regulation of their translation and replication. Such proteins, when acting directly at the viral RNA, contain multiple RNA-binding motifs and commonly interact at highly structured 5'- and/or 3'- UTR of the genome. Some cellular proteins, e.g. La protein (Spangberg et al. 1999), PCPB2 (Spangberg and Schwartz 1999; Wang et al. 2011), PTB and hnRNP C (Gontarek et al. 1999), were proven to fulfill key functions at the HCV UTRs, however, other potential factors may remain undiscovered. Several proteomic studies provide extensive data collections on cellular proteins that can associate with certain regions of the HCV RNA and potentially play a functional role. In the recent study applying a specific affinity capture system in combination with LC/MS/MS, 83 cellular factors were identified to associate with the HCV genome (Upadhyay et al. 2013). Another study based on biotinylated RNA pull-down assay followed by 2DE/MALDI-TOF MS and 1DE/LC/MS methods, revealed 10 proteins binding particularly at the HCV 3'UTR (Tingting et al. 2006). Apart from well-investigated factors, both studies demonstrated the binding of selected cellular aminoacyl-tRNA synthetases (ARSs). A peculiar phenomenon of the ARS acting at viral genome was illustrated for Poliovirus: recruitment of glycyl-tRNA synthetase (GARS) to the apical part of the IRES domain V facilitates a correct positioning of the 40S ribosomal subunit at the initiation region of the Poliovirus IRES therefore elevating the translation initiation. Remarkably, an association of GARS is promoted by the anticodon stem-loop mimicry of $tRNA^{Gly}$ (anticodon ACC) in the apical part of the IRES domain V (Andreev et al. 2012).

The above mentioned allowed to put forward a hypothesis on a possible function of ARSs at the HCV RNA. Taken together, the proteomics studies suggested a direct interaction of the Arginyl-, Aspartyl-, Glutaminyl- and Lysyl- tRNA synthetases (RARS, DARS, QARS and KARS, respectively) with the HCV RNA (Tingting et al. 2006; Upadhyay et al. 2013); notably, all these are also parts of multi-tRNA synthetase complex (Norcum and Warrington 1998). Most of the complex-forming ARS are known to have

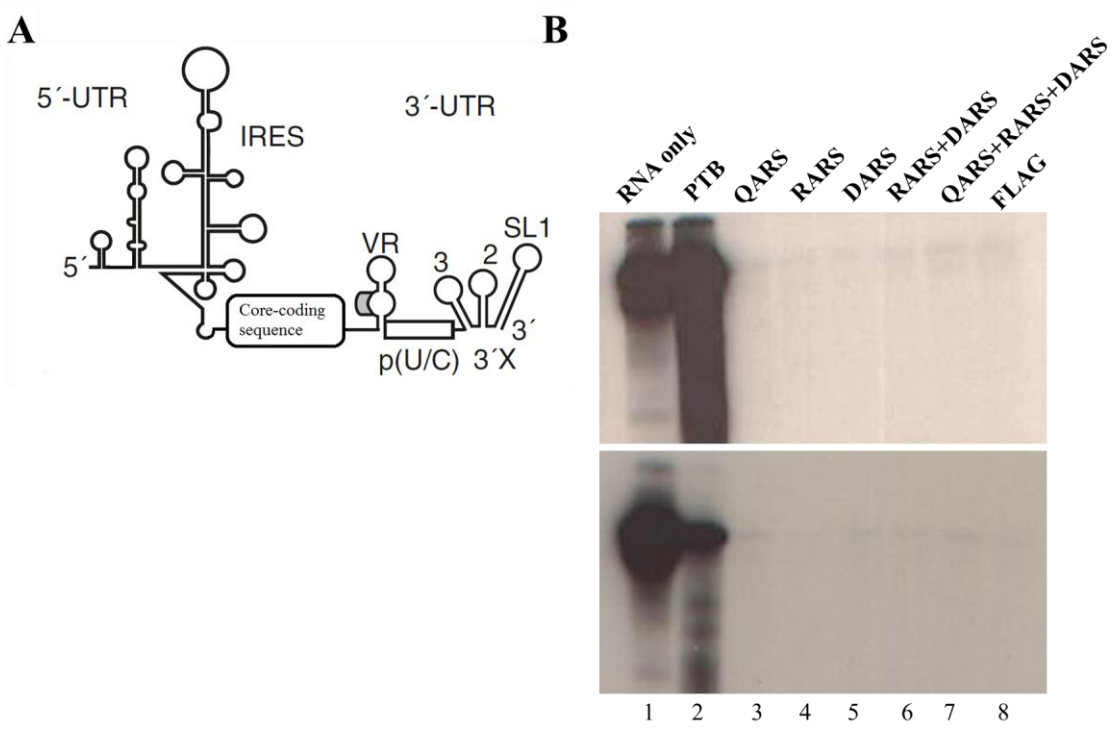
non-canonical activities in cell apart from the protein synthesis, therefore may assumingly act at viral RNA. The multi-tRNA synthetase complex is formed by nine different ARSs and three non-enzymatic factors, associated within 3 subdomains (DARS and QARS are included in domain 1, KARS and RARS – in domain 2) (Park et al. 2010).

Following the fashion of GARS binding to the Poliovirus IRES, anticodon stem-loop mimicry was hypothesized to represent the mechanism of the ARSs binding to the HCV 3'UTR. Since the HCV 3'UTR variable region reflects a low degree of conservation and the polyU/C stretch is not structured, the X-tail was found to be the most promising platform for the binding. Primary sequence of the X-tail is nearly invariant (Suppl. Fig. 6.13, A), and its secondary structure has three stable stem-loops (SL1-3) harboring 6, 10 and 7 nucleotides within their apical loops, respectively (Suppl. Fig. 6.13, B). An alignment of X-tail sequences from different HCV isolates illustrates the presence of the anticodon sequences corresponding to the ARSs with proposed binding to the 3'UTR located within loop sequences of the SL2 and SL3 (Suppl. Fig. 6.13, B). The following "anticodon" nucleotides are absolutely conserved among studied HCV subtypes: DARS "anticodon" 5'-AUC-3' (codons are GAPy), RARS "anticodon" 5'-UCU-3' (codons are CGN, AGPu), QARS "anticodon" 5'-CUG-3' (codons are CAPu) (Suppl. Fig. 6.13, A, blue, red and green labeling, respectively). The "anticodon" for KARS - 5'-PyUU-3' (codons are AAPu) was not found within single-stranded sequences of the X-tail. Thus, it was hypothesized that the SL2 and SL3 may mimic anticodon stem-loops for the Arginyl-, Aspartyl- and Glutaminyl- tRNA synthetases (RARS, DARS and QARS, respectively) mediating their recruitment to the HCV 3'UTR.

Prior investigation of the RARS/DARS/QARS functions at the HCV 3'UTR, the binding of these ARSs was challenged in RNA immunoprecipitation (RIP) assay. Following validation of antibodies specific for the studied ARSs (data not shown) immunoprecipitation of the HCV RNA was performed with each of the ARSs separately and in combination. As a positive control of binding to the HCV 3'UTR, the PTB protein was pulled down. Radioactively labeled RNA construct harboring HCV 5'UTR, Core-coding sequence and the 3'UTR (Suppl. Fig. 6.14, A) was transfected into HeLa cells. After 6 hours, cell lysates were prepared and combined with the protein A magnetic beads hybridized to an antibody or a combination



Supplementary Figure 6.13: Potential aminoacyl-tRNA synthetase (ARS) anticodon stem-loop mimicry within the HCV X-tail sequence.
(A) Alignment of the X-tail sequences from different HCV isolates illustrates high primary sequence conservation in this region, also for the potential "anticodon" sequences: DARS "anticodon" 5'-AUC-3' (blue), RARS "anticodon" 5'-UCU-3' (red), QARS "anticodon" 5'-CUG-3' (green).
(B) The sequence and secondary structure of the HCV X-tail containing three stem-loops (SL1-3). Possible recruitment of RARS/DARS/QARS may occur via recognition of anticodon stem-loop exposing a suitable anticodon (encircled).



Supplementary Figure 6.14: Pull-down of the selected aminoacyl-tRNA synthetases (ARSs) with the HCV RNA.
(A) The RNA construct used for RNA immunoprecipitation contains HCV 5'UTR, the Core-coding region representing an ORF and the 3'UTR. Variable region (VR), polyU/C tract (p(U/C)) and the X-tail (3'X) with stem-loops 1-3 (SL1-3) are annotated.
(B) RNA immunoprecipitation assay was applied to pull down a selected ARS or their combination from HeLa cell lysate after transfection with the radioactively labeled construct from A. Pull-down of PTB protein serves as a control of immunoprecipitation, since PTB strongly binds to the HCV 3'UTR; FLAG serves as a negative control due to the lack of binding affinity to HCV sequences. Complexes' formation is visualized by PAGE followed by autoradiography. Two representative experiments are illustarted; the lane 1 is the RNA transcript serving as a length marker.

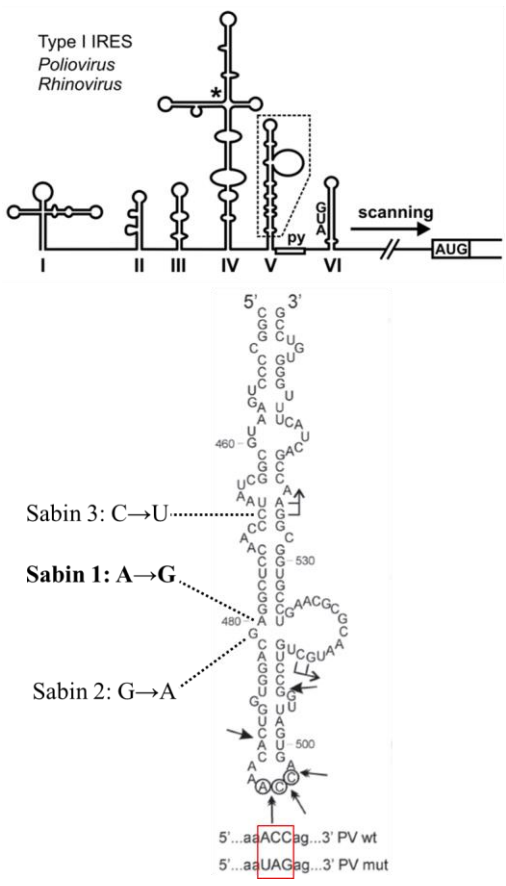
of antibodies specific to selected proteins (RARS, DARS, QARS or PTB). Following the binding for 6 hours (at 4°C under a constant mixing), the RNA was eluted from beads and visualized by PAGE. In multiple assays binding of neither of studied ARSs was demonstrated in contrast to the binding of PTB (Suppl. Fig. 6.14, B) that has a strong affinity to the HCV 3'UTR. Once altering of the pull-down conditions did not affect the result, it was concluded that the binding of the ARSs may be temporal occurring with a low affinity and/or relaying on other unknown components.

6.3.3 Generation of Poliovirus IRES domain V mutants

As already briefly discussed in 6.3.2, Poliovirus (PV) sets an example of exceptional adaptation to the host cell environment and of its exploitation for fulfillment of the viral life cycle. The Poliovirus IRES element is located at the genomic 5'UTR and is known to recruit a variety of cellular RNA-binding proteins that modulate its activity and mediate cap-independent translation initiation (Belsham and Jackson 2000). The Poliovirus IRES is classified as a Type 1 (according to its secondary structure; see 1.2.1 and Fig. 1.2.2) and unlike Type 3 and 4 IRESes utilizes some auxiliary proteins in addition to certain canonical initiation factors. Many of these non-canonical factors (IRES *Trans*-Acting Factors, ITAFs) are also known to stimulate translation from the HCV IRES (Type 3 IRES), for instance, La protein, PTB or PCPB2 (Niepmann 1999). Domain V of the Poliovirus IRES appears to be the most important structural element, it binds the complex of eIF4G/eIF4A (that mediates the entry of the 40S ribosomal subunit on the IRES) as well as eIF4B factor and almost all mutations affecting its structure abrogate the IRES activity (Ochs et al. 2003). In particular, several single point mutations within this domain are presented in well-known attenuated Sabin vaccine strains (Minor 1993). These are A(480)→G, G(481)→A and C(472)→U mutations in Sabin 1, 2 and 3 vaccine strains, respectively (Suppl. Fig. 6.15).

Moreover, an unexpected ITAF - cellular glycyl-tRNA synthetase (GARS) - was discovered to be an essential player in PV translation initiation (Andreev et al. 2012). It was demonstrated that GARS specifically binds to the apical part of the IRES domain V and mediates a correct positioning of the 40S ribosomal subunit at the PV IRES. Strikingly, the structure and sequence of the apical stem-loop at the domain V mimics the anticodon stem-loop of tRNAGly (anticodon ACC) and ensures an exceptional specificity of GARS binding. Mutations affecting the anticodon and therefore preventing GARS recruitment were shown to results in a dramatic drop of the Poliovirus IRES activity. In a fashion of tRNAGly recognition by GARS, the two C residues appeared to be essential for GARS binding to the domain V, whereas the first residue A is redundant. Peculiarly, Poliovirus appears to evolve a specific mechanism to exploit a cytoplasmic house-keeping enzyme, probably shared with other representatives of Type 1 IRES elements (Belsham and Jackson 2000).

As a side project, we aimed to investigate an impact of the above-mentioned mutations within the PV IRES domain V on the viral life cycle. Any of the Sabin mutations are well-known to crucially impair Poliovirus infectivity and neurovirulence due to significant impairment of the translation initiation factors eIF4B and eIF4G binding to the domain V and therefore of an association of ribosomes at the viral genome. In contrast, the effect of mutations within the apical loop of the domain V, which is proposed to be responsible for binding of GARS, on Poliovirus translation and/or life cycle has not been tested *in vivo*. Hence, Gly-anticodon mutation (ACC→TAG), which fully abrogates tRNAGly binding by GARS, was introduced into full-length Poliovirus genome by site-directed mutagenesis approach. In order to control a degree of attenuation of Poliovirus propagation by the latter, a Sabin 1 mutation was introduced into a plasmid encoding a wild-type PV genomic RNA; additionally, a double mutant was generated (Suppl. Fig. 6.15). Transfection of *in vitro* transcribed Poliovirus genomic RNA (derived in the same fashion as presented for HCV in the current work) into HeLa cells and isolation and characterization of virus stock still remains to be done. Upon titration of the virus stock on susceptible cells, a multiplicity of infection (MOI) is to be determined. At last, a growth kinetics of the wild-type and mutated Poliovirus is to be characterized via infection of HeLa cells with either of the virus type at a high MOI, followed by a plaque assay.



Supplementary Figure 6.15: Poliovirus IRES domain V mutations affecting viral translation initiation.
A schematic representation of the Poliovirus IRES; the domain V of the IRES (dashed region) is detailed below. Predicted secondary structures of domains I to VI are illustrated together with an oligopyrimidine tract (py) and a silent AUG codon within domain VI; the Poliovirus translation initiation AUG is located at the position 743. The single nucleotide substitutions corresponding to poliovirus Sabin vaccine strains (Sabin 1-3) are indicated. The apical loop of the domain V is proposed to mimic tRNAGly anti-codon stem-loop; nucleotides forming a Gly-anticodon ACC are encircled. Conversion of this sequence into UAG abrogated binding of GARS to the domain V. This mutation – alone or in combination with the attenuating Sabin 1 mutation – was generated, as detailed in 6.3.3.
(modified from Ochs et al. 2003)

The described experiment aims to support the data obtained *in vitro* and to provide vital information on attenuating ability of GARS binding impairment and on its potential application for vaccine strains generation. An additive attenuating effect, if demonstrated by the double mutant, may be applied for creation of safer vaccines without affecting a degree of immune response.

6.4 Functional sequestration of microRNA-122 from Hepatitis C Virus by circular RNA sponges

RNA BIOLOGY, 2018
VOL. 0, NO. 0, 1–8
https://doi.org/10.1080/15476286.2018.1435248

Taylor & Francis
Taylor & Francis Group

BRIEF COMMUNICATION

OPEN ACCESS Check for updates

Functional sequestration of microRNA-122 from Hepatitis C Virus by circular RNA sponges

Isabelle Jost[a,†], Lyudmila A. Shalamova[a,b,†], Gesche K. Gerresheim[a,b], Michael Niepmann[b], Albrecht Bindereif[a] and Oliver Rossbach[a]

[a]Institute of Biochemistry, Faculty of Biology and Chemistry, University of Giessen, Heinrich-Buff-Ring 17, Giessen, Germany; [b]Institute of Biochemistry, Faculty of Medicine, University of Giessen, Friedrichstrasse 24, Giessen, Germany

ABSTRACT

Circular RNAs (circRNAs) were recently described as a novel class of cellular RNAs. Two circRNAs were reported to function as molecular sponges, sequestering specific microRNAs, thereby de-repressing target mRNAs. Due to their elevated stability in comparison to linear RNA, circRNAs may be an interesting tool in molecular medicine and biology. In this study, we provide a proof-of-principle that circRNAs can be engineered as microRNA sponges. As a model system, we used the Hepatitis C Virus (HCV), which requires cellular microRNA-122 for its life cycle. We produced artificial circRNA sponges *in vitro* that efficiently sequester microRNA-122, thereby inhibiting viral protein production in an HCV cell culture system. These circRNAs are more stable than their linear counterparts, and localize both to the cytoplasm and to the nucleus, opening up a wide range of potential applications.

ARTICLE HISTORY
Received 18 July 2017
Revised 30 November 2017
Accepted 26 January 2018

KEYWORDS
Circular RNA; Hepatitis C Virus; microRNA-122; molecular sponge; RNA therapy

Recent advantages in deep sequencing technologies enabled a discovery of non-conventional splicing of primary transcripts in human genome. A substantial number of RNAs were identified with circularly permuted exon orders and hypothesized to result from head-to-tail splicing of exons (also known as back-splicing), when a donor splice site splicing to an upstream instead of a downstream acceptor site (Salzman et al. 2012). In the following years thousands of circular RNAs (circRNAs) were identified in different species, leading to a creation of a database "circBase" (Glažar et al. 2014). Biological functions of eukaryotic exonic circRNAs, however, largely remain to be understood.

One of the well-investigated examples of circRNA function in human is CDR1as (also known as ciRS-7 in Hansen et al. 2013), an antisense to the cerebellar degeneration-related protein 1 transcript. Containing 73 conserved binding sites for miR-7, CDR1as functionally suppresses miR-7 and miR-671 activity in neuronal tissues and de-represses their direct targets, which appears critical for normal brain function (Memczak et al. 2013, Hansen et al. 2013, Piwecka et al. 2017). Such way of action was defined as miRNA sponge.

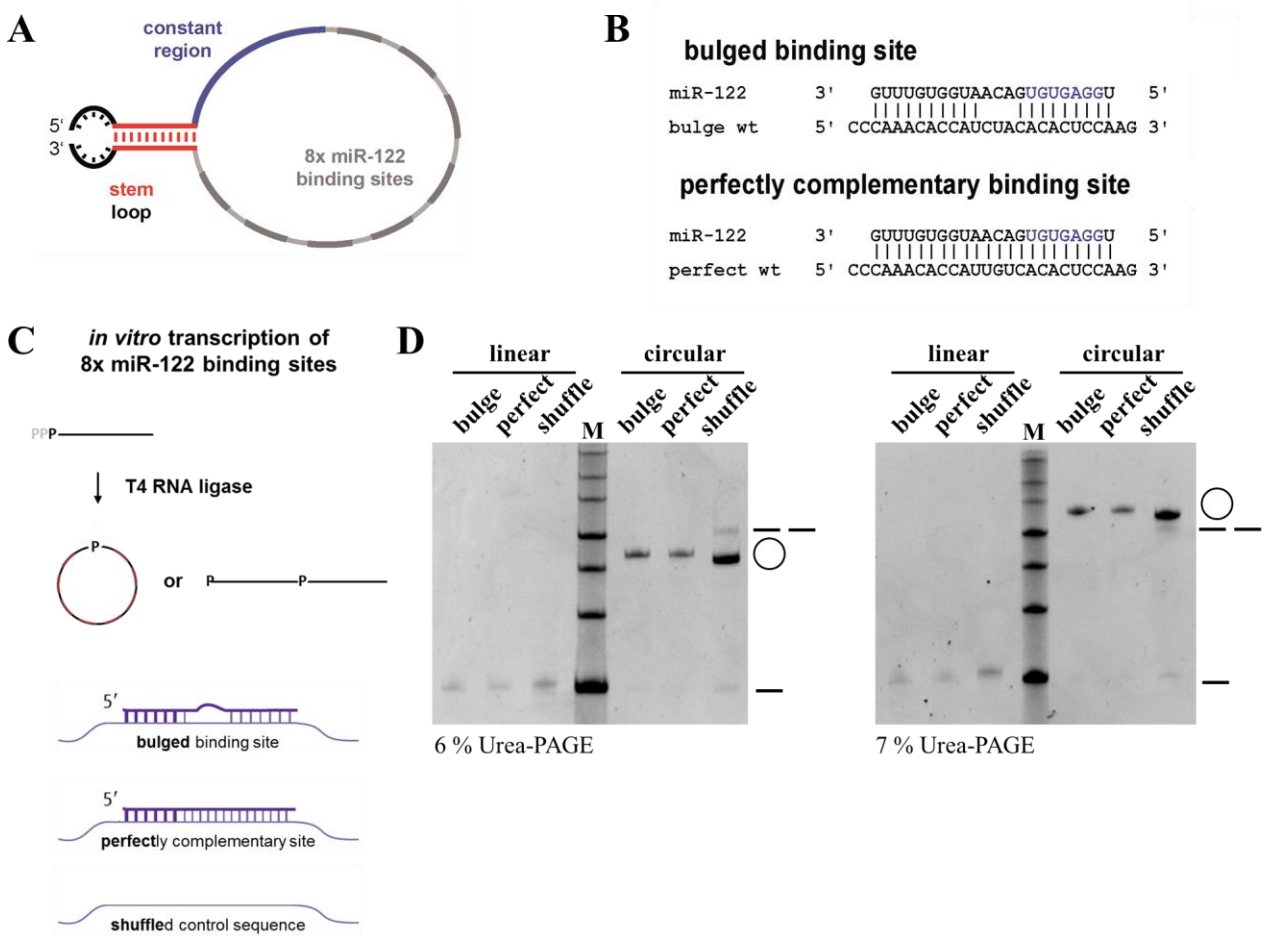
In the published research (Jost et al. 2018), conducted prevalently by Isabelle Jost, an artificial miRNA sponge was designed, synthesized and validated. Developed in analogy to natural circRNA sponge, as a proof-of-principle, an anti-miR-122 sponge was created and applied to HCV cell culture system. The latter enabled verification of a functional sequestration of miRNA-122 by the engineered circular RNA sponge. Upon efficient binding of cellular miR-122 by the sponge, HCV protein production was dramatically reduced, comparable to an effect of the antimiR-122 drug Miravirsen (Janssen et al. 2013). The study provides a detailed protocol for engineering and generation *in vitro* of circRNA sponges, also harboring sequence of repetitive nature. Importantly, the circRNA sponges were testified to be more stable than their linear counterparts, suggesting an extended sequestration during a possible application.

The majority of the work was performed by Isabelle Jost and detailed in the completed PhD thesis. During my PhD thesis, in parallel with the main project, I have conducted the part that required work with an infectious HCV full-length system under BSL3** conditions, since it was unavailable for Isabelle. Additionally, with a help of Gesche Gerresheim, I performed all experiments requested during further revision process. Exclusively these results are summarized below.

6.4.1 Design and biogenesis of circular miR-122 sponge RNAs

Development and optimization of artificial anti-miR-122 sponge production procedure comprises a substantial part of the PhD thesis of Isabelle Jost. In the development of sponges, high-affinity miRNA binding sites were designed and experimentally validated *in vitro*, since the binding affinity of artificial miR-122 binding sites needs to be comparable of the ones situated on HCV RNA.

Hereby an application of the optimized protocol illustrates the key steps and intermediates of circRNA *in vitro* synthesis (Suppl. Fig. 6.16). An RNA template for circularization is derived by *in vitro* transcription

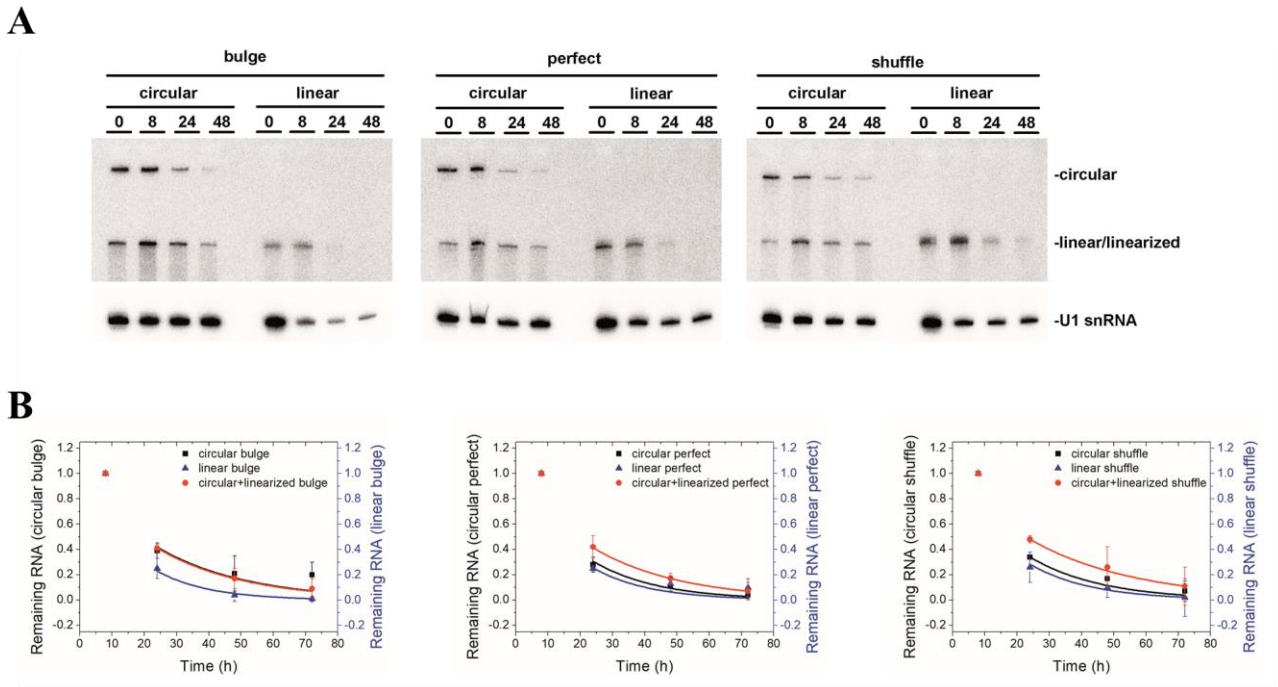


Supplementary Figure 6.16: Preparation and analysis of circRNA sponges *in vitro*.
(A) A general representation of the circularization templates derived by *in vitro* transcription: 8 tandem repeats for a designed miR-122 binding site (in grey), a constant region serving as a probe/primer binding platform for analytical purposes (in blue), an 11 nt reverse-complementary repeat sequence forms a stem (in red) that brings the 5'- and 3' - ends in close proximity to enhance ligation efficiency.
(B) Sequence of wild-type bulged and perfectly complementary miR-122 binding sites; a bulged binding site lacks complementarity at nucleotides 10–12.
(C) A scheme of *in vitro* ligation reaction. RNA transcription is performed with 4-fold excess of GMP over GTP in order to enrich for transcripts (as in A) with a 5'-monophosphate that is required for efficient RNA ligation. T4 RNA ligase 1 activity results in the formation of circular and linear dimer forms.
(D) Analytical denaturing polyacrylamide gel electrophoresis (PAGE) of purified circular and linear constructs (containing bulged, perfectly complementary or shuffled binding sites) illustrating a variable mobility of circular products in gels of different acrylamide concentration (here, 6 and 7 %); linear monomers and dimers retain their electrophoretic mobility. M, low range RNA ladder.
(modified from Jost et al. 2018)

from a linearized DNA plasmid encoding corresponding tandem binding sites as well as shared by all constructs' sequences (Suppl. Fig. 6.16, A). Due to the repetitive nature of these tandem binding sites, theirnumber was limited to 8 in the course of optimization preventing dramatic impairment in the efficiency of major biochemical techniques. Three types of miR-122 binding sites were evaluated in the study: a bulged binding site with a mismatch at nucleotides 10-12, a perfectly complementary binding site and a shuffled sequence serving as a negative control (Suppl. Fig. 6.16, B and C). Linear RNA transcripts harboring a monophosphate at the 5'-end can be circularized by RNA ligase 1 or form linear dimers (Suppl. Fig. 6.16, C). The ratio between products varies in accordance to a degree of sequence repetitiveness, resulting in the highest outcome for the shuffled circRNA construct. Omitting linear dimer forms, both monomeric forms – linear and circular – were separated on a gel in order to gel-purify and concentrate the resulting constructs. A simple diagnostic approach allows distinguishing between circular and linear forms by differential electrophoretic mobility of circular molecules in polyacrylamide gels of different acrylamide concentration (Tabak et al. 1988; Suppl. Fig. 6.16, D). To sum up, three types of circular and linear sponges containing miR-122 or control binding sites can be generated by *in vitro* approaches; for more details and optimizations see the PhD thesis of Isabelle Jost.

6.4.2 Stability of circRNA sponges

As a part of my contribution to the study, stability of the presented sponge RNAs was evaluated. Formerly, Isabelle Jost had performed time-course experiments aiming to quantify half-life times of the three circRNA constructs and their linear counterparts in HuH-7.5 cells that are routinely used for the functional assays. However, the former analyses were biased due to the use of Lipofectamine transfection reagent, which can artificially alter an uptake of circular versus linear RNAs and/or shield RNAs from degradation on



Supplementary Figure 6.17: Stability of circular and linear RNA sponges in HuH-7.5 cells.
(A) Circular and linear RNA fractions detected by ^{32}P-northern blot 0, 8, 24 and 48 hours post electroporation into HuH-7.5 cells. The northern blot probe targets a conserved region shared by all constructs (containing bulged, perfectly complementary or shuffled binding sites). Circular RNA constructs undergo decay through a linear form (an open circle) that positively contributes to their overall half-life values (also see B).
(B) Quantification of the results from (A) represented in decay curves; the graphs illustrate differences between stability of circular RNA (in black), linear RNA (in blue) and "circular + linearized" RNA forms (in red) - for bulge, perfect and shuffle sponges (left to right). Normalization was conducted to 8 h values within three independent experiments and mean fractions of remaining RNA were fitted to exponential decay function to calculate half-lives of each sponge RNA.
(from Jost et al. 2018)

the outer cellular surface. Therefore, to gain more reliable comparison of stability of circular and linear constructs, electroporation method of RNA transfection was applied on HuH-7.5 cells, followed by isolation of total cellular RNA according to a selected time-course (input, 8, 24, 48 and 72 hours post electroporation; three biological replicates). Once equal amounts of circRNA sponges or their linear counterparts are transfected, fractions of intact RNAs remaining in the course of an experiment allow mathematical prediction of each construct's half-life time. For each RNA species, mean fractions of remaining RNA were normalized to 8 h values. Decay curves for the analyzed constructs were derived by plotting to an exponential decay function (ExpDec1) using OriginPro8 software. CircRNA constructs were characterized by stability of either a circular form alone or summed up with a corresponding linearized fraction (an opened cycle) that still functions as a sponge.

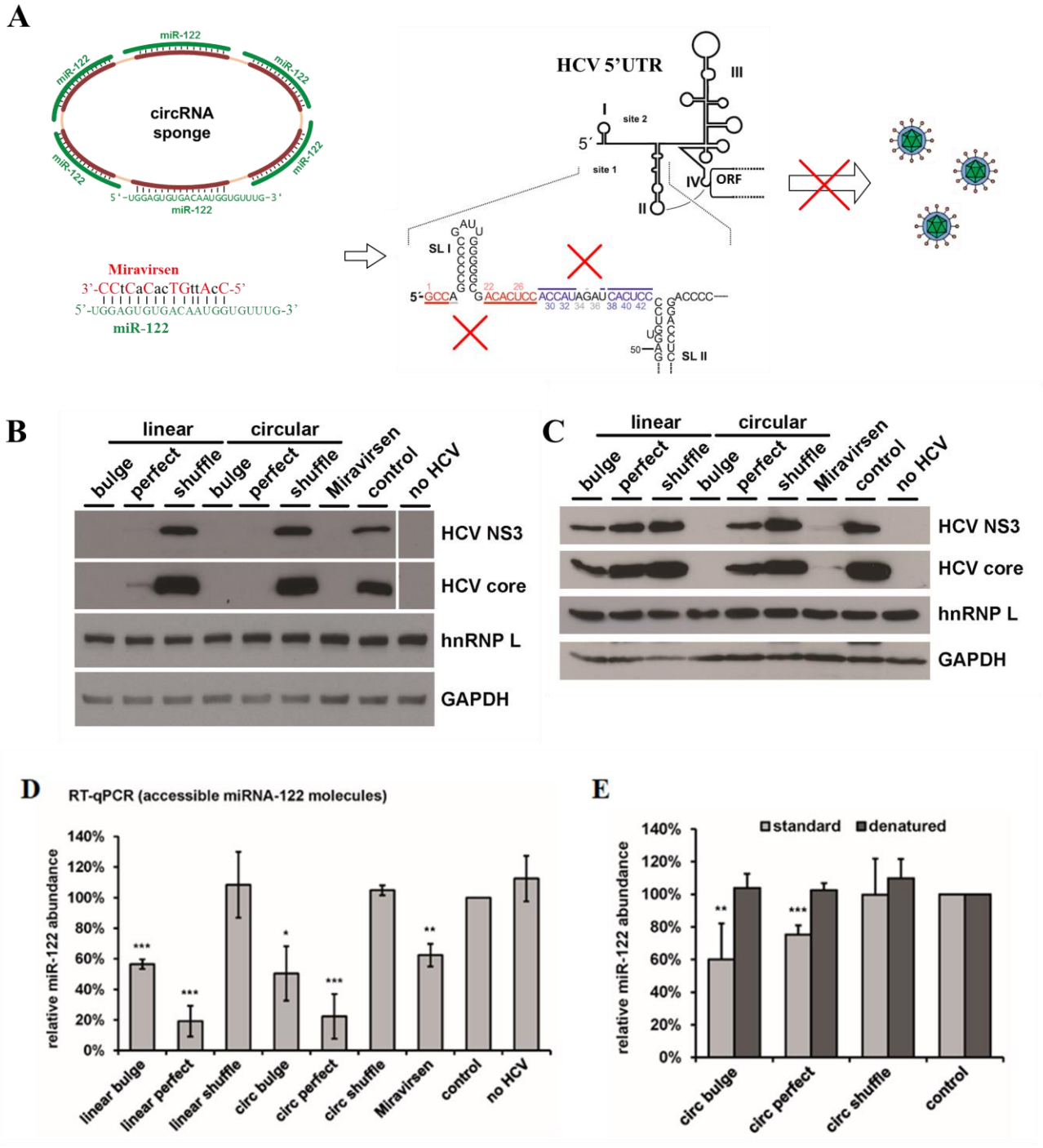
Visualization of intact fractions of circular and linear forms was fulfilled by ^{32}P-northern blot analysis using identical probes in the constant region for all constructs and U1 snRNA as a loading control. The representative images (Suppl. Fig. 6.17, A) testify for an increased stability of circular RNA sponges over their linear counterparts, which originate from both inaccessibility for exonucleolytic degradation and from the two-step decay. For the downstream calculations (Suppl. Fig. 6.17, B), the autoradiography bands' intensity was quantified using ImageQuantTL software. Taking results for bulge, perfect and shuffle constructs together, half-life values varied from 11.3 to 13.2 h for linear RNAs and from 14.1 to 19.1 h – for a circular form, whereas the overall value for "circular + linearized" forms varied from 18.7 to 22.7 h. Thus, the analyzed circular constructs appear to be functional nearly twice as long as the linear constructs providing an advantageous feature for potential therapeutic applications.

6.4.3 Application of circRNA sponges to infectious HCV system

The major part of my contribution to the study was a validation of the designed miR-122 sponges' functionality using the full-length HCV cell culture system. Prior referring to their effects on infectious virus, meticulous selection and verification of binding sites was conducted by Isabelle Jost utilizing two different HCV-reporter systems. As long as the artificial circRNAs demonstrated a comparable or even a higher activity as of Miravirsen-like antisense oligonucleotide in impairment of the miR-122 dependent reporter gene translation, the sponges' functionality was trialed in HCV infection model.

MiR-122 is beyond doubt one of the most vital host factors for HCV propagation. It acts directly on the viral genomic RNA and modulates almost every step of the HCV life cycle. One well-investigated mechanism of miR-122 action is fulfilled through its binding to two tandem binding sites at the very 5'-end of the HCV genome. In association with Ago2 miR-122 was demonstrated to increase HCV RNA stability as well as to promote its replication and translation (for more details see 1.3.4.3). In a full-length HCV system, in contrast to HCV reporter systems, miR-122 may also contribute via binding to an array of binding sites located at the 3'-end of genomic RNA and additionally affecting the stages downstream of translation/replication. Although the use of a full-length HCV system does not reveal which exact stage of the viral life cycle is primarily inhibited upon sponging of miR-122 from HCV by the designed constructs, it provides an insight into their overall effect on viral propagation in cell culture (schematically illustrated by Suppl. Fig. 6.18, A).

Taking advantage of the developed fully permissive HuH-7.5 cell culture system (Lindenbach et al. 2005; Wakita et al. 2005; Zhong et al. 2005), the effect of the circRNA miR-122 sponges was compared to as of linear constricts and of Miravirsen upon co-transfection together with *in vitro* transcribed Jc1 HCV RNA. It should be noted that the introduction of equal absolute mass quantities of all sponge RNAs and Miravirsen in the experiments below reflects a three-fold higher molarity of miRNA-122 binding sites for the transfection of Miravirsen. Following the co-transfection, HuH-7.5 cells were incubated for 5 days to allow a full-blown HCV infection and further harvested to isolate total protein extract and total RNA. Usually 50 ng of a sequestration construct was used for co-transfection in 12-well format and resulted in a reduction of HCV proteins translation to undetectable level for both circular bulge and perfect constructs as well as for the Miravirsen-like antisense oligonucleotide (Suppl. Fig. 6.18, B). Reproduced multiple times such



Supplementary Figure 6.18: Functional sequestration of miR-122 by circular and linear RNA sponges affects HCV translation.

(A) Schematic illustration of sequestration of miR-122 by circRNA sponges or Miravirsen (Janssen et al. 2013) preventing binding of miR-122 to the HCV 5'UTR and leading to impairment of HCV production. Two well-characterized miR-122 binding sites at the HCV genomic 5'UTR upstream the IRES are colored.

(B) Detection of the HCV NS3 and Core proteins by western blot in HuH-7.5 cells lysates (5 days post transfection). Cells were co-transfected with *in vitro* transcribed HCV RNA and circular or linear sponge RNAs (containing bulged, perfectly complementary or shuffled binding sites). Cellular hnRNP L and GAPDH proteins serve as loading control.

(C) Western blot of the HCV NS3 and Core proteins in the experiment similar to (B), but using a limiting amount of sponge RNA ($1/5^{th}$).

(D) TaqMan-RT-qPCR quantification of accessible miR-122 molecules after circular or linear sponge RNA application (as in B).

(E) TaqMan-RT-qPCR quantification of accessible miR-122 molecules under standard reaction conditions ("standard") and of total amount of miR-122 molecules – under denaturing conditions recreated by pre-heating of sample RNA with primers prior reaction ("denatured").

(modified from Jost et al. 2018)

experimental setup only rarely demonstrated noticeable differences in sponging efficiencies of circular constructs and their linear counterparts. The latter are characterized by nearly two times lower stability in HuH-7.5 cells, as presented above, and are expected to be less efficient than circRNAs. Indeed, when 10 ng of each sponge was applied in the similar format, significant advantages in circular constructs' functionality, especially for bulge circRNA, were demonstrated over their linear counterparts (Suppl. Fig. 6.18, C). Ultimately, the bulge circRNA sponge manifested the more profound functional sequestration of miR-122 than bulge linear RNA or the construct containing perfectly complementary binding sites, or Miravirsen. As expected, shuffle linear and circular sponges had no effect on HCV translation: viral infection development was as in control experiment (HCV RNA only transfected). Moreover, efficient interference with production of both non-structural (NS3) and structural (Core) HCV proteins indirectly implies that both - intracellular stages of the HCV life cycle (e.g. polyprotein processing, RNA replication) and budding – would be impaired by functional RNA sponges.

In order to provide a prove that observed effects on HCV production are indeed mediated by sponging of miR-122, the total intracellular level of miR-122 was compared to amount of unbound, accessible molecules. While northern blot detection of miR-122 content reveals all molecules in the cell (including those sequestered by sponges or situated on its cellular target and/or HCV RNA), reverse transcription followed by TaqMan qPCR quantifies only functionally available molecules. Carried out under denaturing conditions, quantification of miR-122 by northern blot within the samples corresponding to Suppl. Fig. 6.18, B, indicated identical total levels of the targeted microRNA (data not shown here). In contrast, quantification of miR-122 by TaqMan RT-qPCR resulted in remarkable reduction of accessible molecules, supposedly upon sponging (Suppl. Fig. 6.18, D). Interestingly, the degree of reduction was in agreement with the stability of circular versus linear RNA forms and with observed effects on HCV translation. Aiming to support an assumption that TaqMan RT-qPCR approach virtually quantifies only unbound microRNAs, recreation of denaturing conditions prior reverse transcription was attempted. Pre-heating of analyzed total RNA together with RT primers to 95°C for 2 minutes denatures secondary structures liberating bound miR-122 molecules for hybridization to primers. As expected, a reversion to quantification of the total pool of miR-122 was observed (Suppl. Fig. 6.18, E).

In summary, circRNA sponges appeared to be more efficient for miR-122 sequestration than their linear analogs suggesting the more prominent potential for clinical applications. Due to their extended stability alone and functionality both as circular and linearized sponges, they were shown to sequester more miR-122 molecules than linear RNA in scope of the presented experiments. Unlike Miravirsen, the developed circRNA sponges are metabolized when applied *in vivo*, arguing against side effects mediated by accumulation of a therapeutic agent (Swayze et al. 2006). This proof-of-principle study has demonstrated how an efficient targeting of pathogens' essential factor by artificial circular RNA sponges can effectively inhibit pathogens' production opening a wide range of possible applications in molecular therapy and medicine.

6.5 List of abbreviations and symbols

3D - three-dimensional
AA - acrylamide
Ago2 - Argonaute protein 2
APS - ammonium persulfate
ARS - aminoacyl-tRNA synthetase
ATP - adenosine triphosphate
bp - base pair
BR - broad range
BVDV - Bovine Viral Diarrhea Virus
C - Cytidine
CD - cluster of differentiation
cDNA - complementary DNA
circRNA – circular RNA
CLDN1 - Claudin-1
cm^2 - square centimeter
CRE - *cis*-acting replication element
CrPV - Cricket Paralysis Virus
CSB3 - Coxsackievirus B3
Ct - threshold cycle
CTP - cytosine triphosphate
CypA - Cyclophilin A
DAA - direct-acting antiviral
DARS – Aspartyl-tRNA synthetase
DENV - Dengue Virus
DLS - dimerization linkage sequence
DMEM - Dulbecco's Modified Eagle's Medium
DMHBI - 3,5-dimethoxy-4-hydroxybenzylidene imidazolinone
DMSO - dimethyl sulfoxide
DMV/MMV - double/multi membrane vesicle
DNA - deoxyribonucleic acid
DNase - deoxyribonuclease
dsRNA - double-stranded RNA
DTT - dithiothreitol
E - amplification efficiency
EBV - Epstein-Barr Virus
E. coli - Escherichia coli
EDTA - ethylenediaminetetraacetic acid
EGTA - ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid
eIF - eukaryotic initiation factor
EMCV - Encephalomyocarditis Virus
ER - endoplasmic reticulum
FBS - fetal bovine serum
FMDV - Foot-and-Mouth Disease Virus
for - forward
g - gram
G - Guanine
G-418 - Geneticin 418
GAPDH - glyceraldehyde-3-phosphate dehydrogenase

GARS - Glycyl-tRNA synthetase
GBV-B - GB Virus B
GDP - guanosine diphosphate
GFP - green fluorescent protein
Gly – Glycine
GSH - glutathione
GTP - guanosine triphosphate
h - hour
HAV - Hepatitis A Virus
HBI - 4-hydroxybenzlidene imidazolinone
HCMV - Human Cytomegalovirus
HCV - Hepatitis C Virus
HDV - Hepatitis Delta Virus
HF - high fidelity
HIV-1 - Human Immunodeficiency Virus type 1
hnRNP - heterogeneous nuclear ribonucleoprotein
hp - hairpin
h.p.t. - hours post transfection
HRP - horseradish peroxidase
HSPG - heparan sulfate proteoglycan
HSV - Herpes Simplex Virus
IFN(-α) - Interferon(-alpha)
Ile - Isoleucine
IRES - internal ribosome entry site
IRF3 - interferon regulatory factor 3
ISG - IFN-stimulated gene
ITAF - IRES *trans*-acting factor
JFH - Japanese fulminant hepatitis
kb - kilobase
kDa - kilodalton
l - liter
La - Lupus antigen
LB - Luria-Bertani
LD - lipid droplet
LDL - low-density lipoprotein
LDLR - LDL receptor
LNA - locked nucleic acid
Met - Methionine
mg - milligram
min - minute
min. - minimum
miR(NA) - microRNA
miRISC - miRNA-induced silencing complex
ml - milliliter
mM - millimolar
mm - millimeter
mm^2 - square millimeter
MOPS - 3-(N-morpholino) propanesulfonic acid
mRNA - messenger RNA
ms - millisecond

M.s - minus strand
MS – mass-spectrometry
mut - mutated
NANB(H) - non-A/non-B (hepatitis)
NEB - New England Biolabs
neo - neomycin
nm - nanometer
NPHV - Non-Primate Hepacivirus
NS - non-structural
NTP - nucleoside triphosphate
nt - nucleotide(s)
ORF - open reading frame
PABP - poly(A)-binding protein
PAGE - polyacrylamide gel electrophoresis
PBS - phosphate buffered saline
PCBP2 - Poly(rC) binding protein 2
PCR - polymerase chain reaction
p.e.p – post electroporation
PI4KIIIα - phosphatidylinositol-4-kinase IIIα
PI4P - phosphatidylinositol-4-phosphate
PNA - peptide nucleic acid
RER - relative expression ratio
RPA - ribonuclease protection assay
pp - pseudoparticle
pre-miRNA - precursor miRNA
pri-miRNA - primary miRNA
P.s - plus strand
PTB - polypyrimidine tract-binding protein
Pu - purine
PV - Poliovirus
PVDF - polyvinylidene difluoride
Py – pyrimidine
QARS - Glutaminyl-tRNA synthetase
qPCR - quantitative PCR
RARS - Arginyl-tRNA synthetase
RC - reverse complement(ary)
RdRp - RNA-dependent RNA polymerase
REM - replication-enhancing mutation
rev - reverse
RIG-I - Retinoic Acid Inducible Gene I
RIP - RNA immunoprecipitation
RNA - ribonucleic acid
RNAi - RNA interference
RNase - ribonuclease
rpm - revolutions per minute
RT - reverse transcription
RT-qPCR - quantitative RT-PCR
Rxn - reaction
s – second
scr - scrambled

SD - standard deviation
SDS - sodium dodecyl disulfate
Ser - Serine
SELEX - systematic evolution of ligands by exponential enrichment
siRNA - small interfering RNA
SHAPE - selective 2'-hydroxyl acylation and primer extension
SL - stem-loop
SPP - signal peptide peptidase
SR-BI - scavenger receptor class B type I
ssRNA – single-stranded RNA
T7T - T7 Terminator
TAP1 - tocopherol-associated protein 1
TEMED - N,N,N',N'-tetramethylethylenediamine
T_m - melting temperature
tRNA - transfer RNA
U - Uridine
U - unit
UTP - uridine triphosphate
UTR - untranslated region
VAPA/VAPB - vesicle-associated membrane proteins A/B
VLDL - very low-density lipoprotein
WB - western blot
WHO - World Health Organization
WNV - West Nile Virus
wt - wild-type
Xrn - exoribonuclease
YFV - Yellow Fever Virus
µg - microgram
µl - microliter
µM - micromolar

6.6 Publications and conferences

6.6.1 Publications

Niepmann M, **Shalamova LA**, Gerresheim GK, Rossbach O. Signals Involved in Regulation of Hepatitis C Virus RNA Genome Translation and Replication. *Front Microbiol*. 2018;9:395. doi: 10.3389/fmicb.2018.00395.

Jost I*, **Shalamova LA***, Gerresheim GK, Niepmann M, Bindereif A, Rossbach O. Functional sequestration of microRNA-122 from Hepatitis C Virus by circular RNA sponges. *RNA Biol*. 2018;15(8):1032-1039. doi: 10.1080/15476286.2018.1435248. *These authors contributed equally to this work.

Gerresheim GK, Dünnes N, Nieder-Röhrmann A, **Shalamova LA**, Fricke M, Hofacker I, Höner Zu Siederdissen C, Marz M, Niepmann M. microRNA-122 target sites in the hepatitis C virus RNA NS5B coding region and 3' untranslated region: function in replication and influence of RNA secondary structure. *Cell Mol Life Sci*. 2017 Feb;74(4):747-760. doi: 10.1007/s00018-016-2377-9.

Nieder-Röhrmann A, Dünnes N, Gerresheim GK, **Shalamova LA**, Herchenröther A, Niepmann M. Cooperative enhancement of translation by two adjacent microRNA-122/Argonaute 2 complexes binding to the 5' untranslated region of hepatitis C virus RNA. *J Gen Virol*. 2017 Feb;98(2):212-224. doi: 10.1099/jgv.0.000697.

6.6.2 Manuscripts in preparation

Shalamova LA, Dünnes N, Nieder-Röhrmann A, Gerresheim GK, Mueller S, Song Y, Wimmer E, Rossbach O, Niepmann M. The Hepatitis C Virus RNA genome 5'-end regulates minus strand RNA synthesis initiation at the genome 3'-end.

Hu P, Wilhelm J, Gerresheim GK, **Shalamova LA**, Niepmann M. Lnc-ITM2C-1 is a Proviral Host Factor for Hepatitis C Virus.

6.6.3 Presentations

Poster presentations:

27th Annual Meeting of the Society for Virology
March 22nd to 25th 2017, Marburg, Germany
Lyudmila A. Shalamova, Nadia Dünnes, Anika Nieder-Röhrmann, Gesche K. Gerresheim, Dennis Metzger, Steffen Mueller, Yutong Song, Eckard Wimmer, Oliver Rossbach and Michael Niepmann
Poster title: Dissecting the requirements for Hepatitis C virus RNA synthesis using a split replication system

IRTG Symposium
September 14th to 15th 2015, Gießen, Germany
Lyudmila A. Shalamova, Nadia Dünnes, Anika Nieder-Röhrmann, Gesche K. Gerresheim, Steffen Mueller, Yutong Song, Eckard Wimmer, Oliver Rossbach and Michael Niepmann
Poster title: Dissecting the requirements for Hepatitis C virus RNA synthesis using a split replication system

Oral presentation:

23rd International Symposium on Hepatitis C Virus and Related Viruses (HCV2016)
October 11th to 15th 2016, Kyoto, Japan
Abstract title: Dissecting the requirements for Hepatitis C virus RNA synthesis using a split replication system

6.7 Acknowledgements

At first, I would like to thank Prof. Michael Niepmann for providing me the opportunity to work on this project, for the immense scientific and personal support, for providing the work space, collaboration and travelling possibilities. Big thanks for the help and understanding in the process of writing this dissertation and for his valuable advice.

I would like to especially thank Prof. Albrecht Bindereif for being my first supervisor in the Faculty of biology and chemistry at Justus Liebig University Gießen as well as for providing me an opportunity to conduct qPCR in his laboratory. Here I also want to thank members of the AG Bindereif for having me in the lab and in particular Oliver Rossbach for sharing a working bench during my stay and for indispensable consult on the methods and lab equipment.

Next, I am thankful to my former colleagues Anika Nieder-Röhrmann, Nadia Dünnes, Gesche K. Gerresheim and Pan Hu for both their valuable contribution to the project and their work attitude that created a very special and friendly atmosphere in the lab, for support on bad days in the lab and for all the laughter outside work.

I would like to thank the contributors to a planning of my project - Eckard Wimmer, Yutong Song and Steffen Mueller (Stony Brook University, NY, USA) - for the advisory on the constructs' design as well as for generation of a scrambled sequence.

Big thanks to Dieter Glebe for an opportunity to perform experiments using HCV infectious system at the S3** facility in the BFS.

Thanks to Ivan Shatsky and Dmitri Andreev for our memorable visits to Moscow, for their valuable consult on the Ribosome Profiling methodology and for scientific discussions.

I would like to acknowledge Thomas Pietschmann (Hannover, Germany) and Ralf Bartenschlager (Heidelberg, Germany) for sharing of the Jc1 clone and Charles M. Rice (New York, USA) for HuH-7.5 cells, which made this work possible.

My great thanks to the SFB 1021 "RNA viruses: RNA metabolism, host response and pathogenesis" for an opportunity to join with our project, for the funding organization, for priceless seminars and symposiums allowing to develop presenter's skills and improve scientific thinking. Certainly, I express big thanks to the Deutsche Forschungsgemeinschaft (DFG) for funding of the SFB 1021 and our project.

Additionally, I want to thank the IRTG "Enzymes and Multienzyme complexes acting on Nucleic Acids" for all the valuable workshops, seminars and symposiums, for the travelling opportunities and for incredible contribution to my scientific maturation.

6.8 Eidesstaatliche Erklärung

Ich erkläre:

Ich habe die vorgelegte Dissertation selbstständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe.

Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Gießen, Oktober 2018

Lyudmila Shalamova