The role of phosphorylation of the hepatitis C virus NS5A protein revealed by a combination of biochemistry and reverse genetics

Douglas Ross-Thriepland (BSc (Hons))

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
Faculty of Biological Sciences
School of Molecular and Cellular Biology
September 2013

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The chapter within this thesis that has been based on work from jointly-authored publication is chapter 3 and 4:

Ross-Thriepland, D. Harris, M. (2013). Insights into the complexity and functionality of hepatitis C virus NS5A phosphorylation. J Virol. Doi:10.1128/JVI.03017-13:

- Douglas Ross-Thriepland preformed all experiments and co-authored the paper
- Prof M. Harris provided supervision and co-authored the paper.

Ross-Thriepland, D., Amako, Y., Harris, M. (2013). The C terminus of NS5A domain II is a key determinant of hepatitis C virus genome replication, but is not required for virion assembly and release. J Gen Virol **94**(Pt5) 1009-18:

- Douglas Ross-Thriepland preformed all experiments and co-authored the paper
- Dr Y. Amako developed qRT-PCR primers.
- Prof M. Harris provided supervision and co-authored the paper.

Acknowledgements

I would like to thank Prof. Mark Harris for his invaluable supervision of this project. Without his advice and encouragement over the past four years this project would have never have even got off the ground.

A special thanks to Mair Hughes, Jamel Mankouri and Barnabas King for teaching me the ropes and keeping my science on track, no easy task. Further thanks to all my colleagues at the University of Leeds and specifically in Lab 8.61, didn't we all have fun?

Thank you to my family and to my friends for their never-ending support through some of the less enjoyable times. It is only now with hind sight that I really see how important they have been and how lucky I am. I am eternally grateful, and I am sure they are as relieved as I am that this ordeal is (nearly) all over.

I finally would like to acknowledge the Biotechnology and Biological Sciences Research Council (BBSRC) and my Cooperative Awards in Science and Engineering (CASE) sponsor AstraZeneca for funding this project. I would also like to thank the Society for General Microbiology (SGM) for helping me present my work to the wider scientific community

Abstract

Hepatitis C virus causes a chronic infection that affects 2-3 % of the world population and is a significant causative agent of liver cirrhosis and hepatocellular carcinoma. It has a positive-sense, single-stranded genome that encodes for a single polyprotein that is cleaved by both host and viral proteases into 10 mature viral proteins. The non-structural 5A (NS5A) protein is a pleiotropic protein with essential functions in the replication and production of virus, as well as in perturbing host pathways in favour of virus persistence. NS5A is extensively phosphorylated, forming two distinct species termed the basally and hyperphosphorylated; however the location and function of phosphorylation remains poorly defined.

To address this, a significant quantity of NS5A was purified from an actively replicating context and subjected to extensive phospho-mapping by mass spectrometry. Through this approach 12 phosphorylation sites were identified. Subsequent phenotyping of phosphorylation sites by reverse genetics generated evidence for a sequential phosphorylation cascade within the low complexity region I, and that this was responsible for the formation of the hyperphosphorylated species. Further analysis identified the phosphorylation of the distant serine 146 as negatively regulating formation of the hyperphosphorylated species. To investigate phosphorylation events in greater detail the SNAP-tag technology (NEB) was applied to NS5A. The insertion of the SNAP-tag into domain III of NS5A was shown to be well tolerated with respect to virus replication and function of the SNAP-tag.

In parallel this study also set out to establish whether domain II of NS5A had a hitherto unknown function in the production of infectious virus, however no such effects were observed. In combination with existing data it is clear that no residue within NS5A domain II is essential for the production of infectious virus. Furthermore, this work highlights significant differences in the requirement for domain II in virus replication between different HCV genotypes, despite high conservation within this region.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

1ST	One Strep-Tag
5A1ST	NS5A-One Step Tag
A	Ala, alanine
aa	Amino Acid
ADRP	Adipose differentiation-related protein
Ago-2	Argonaute-2
AH	Amphipathic Helix
AH	Amphipathic helix
AP2M1	Adapter protein-2 M1
Apo	Apolipoprotein
ATP	Adenosine triphosphate
BC	Benzylcytosine
BCA	Bicinchoninic acid
BG	Benzylguanine
BMS	Bristol Myers-Squibb
BODIPY	Boron-dipyrromethene
bp	Base Pairs
BSA	Bovine serum albumin
BSL3	Biosafety laboratory level 3
BVDV	Bovine virus diarrhoea virus
C	Core protein
C	Cys, cysteine
cAMP	Cyclic adenosine monophosphate
CC50	Cytotoxicity concentration, 50%
CD4+	Cluster of differentiation 4 positive
CD8+	Cluster of differentiation 8 positive
CD81	Cluster of differentiation 81
CID	Collision-induced dissociation
CKI, CKII	Casein kinase I, II
CMV	Cytomegalovirus
CsA	Cyclosporin A
CTD	C-terminal Domain
CV	Column Volume
Cyp	Cyclophilin
Cys	Cysteine
D	Asp, aspartic acid
Da	Dalton
DAA	Direct Acting Antivirals
DAGT1	Diglyceride acyltransferase
DAPI	4',6-diamidino-2-phenylindole
DENV	Dengue Fever Virus
DEPC	Diethyl pyrocarbonate
DMEM	Dulbecco's modified eagles medium
DMSO	Dimethyl sulfoxide
DMV	Double membrane vesicle

DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
E	Glu, glutamic acid
E1	Envelope 1 protein
E2	Envelope 2 protein
EC50	Effective concentration, 50%
EDTA	Ethylenediamine tetraacetic acid
EM	Electron Microscopy
EMCV	Encephalomyocarditis virus
emGFP	Emerald green fluorescence protein
ER	Endoplasmic Reticulium
ESAL	European association for the study of the liver
ESI	Electrospray ionisation
ETR	Early Treatment Responder
EVR	Early Virological Response
F	Phe, phenylalanine
FBS	foetal bovine serum
ffu	Focus forming unit
g	Gravitational force
G	Gly, glycine
GAG	Glycosaminoglycans
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green fluorescence protein
GLB	Glasgow lysis buffer
GRASP	GRP-1 associated scaffold protein
GRP1	General receptor for phosphoinositides-1
GSK-3	Glycogen synthase kinase-3
GST	Glutathione S-transferases
gt	Genotype
H	His, histidine
H77	Hutchinson HCV isolate 77, genotype 1a
HBV	Hepatitis C Virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C Virus
HDL	High density lipoprotein
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonate
HepG2	Hepatocellular carcinoma cell line G2
HIV	Human immunodeficiency virus
hpe	Hours post-electroporation
hpi	Hours Post Infection
hpt	Hours Post Transfection
hrs	Hour
hrs	Hours
Huh7	Human hepatocellular carcinoma cell line-7
I	Ile, isoleucine
IDU	Intravenous drug use
IF	immunofluorescence

IFNα2	Interferon-2α
IL28B	Interleukin 28-B
IRES	Internal Ribosome Entry Site
ISDR	interferon sensitivity determining region
JcI	J6/JFH-1 chimeric virus
JFH-1	Japanese fulminant HCV isolate 1, genotype 2a
K	Lys, lysine
kb	kilobase
kDa	Kilodalton
L	Leu, leucine
LarI	Luciferase reagent I
LB	Luria-Bertani
LC-MS	Liquid chromatography-mass spectrometry
LCS	Low complexity sequence
LD	Lipid Droplet
LDL	Low density lipoprotein
LDL	Low Density Lipoprotein
LDL-R	Low density lipoprotein-receptor
LLQ	Lower level of quantification
M	Met, methionine
M.Wt	Molecular weight
m/z	mass to charge
MALDI	Matrix-assisted laser desorption/ionization
MCS	Multiple Cloning Site
MeOH	Methanol
min	Minutes
mins	minutes
miR-122	micro-RNA-122
MMV	Multi-Membrane Vesicle
MS	Mass Spectrometry
MTP	Microsomal transfer protein
MTT	Thiazolyl Blue Tetrazolium Bromide
MWCO	Molecular weight cut-off
N	Asn, asparagine
n.d.	no data/not determined
NCE	New Chemical Entity
NMR	Nuclear magnetic resonance
npt	Neomycin phosphotransferase
NS	Non-structural
NS2	Non-structural 2
NS3	Non-structural 3
NS4A	Non-structural 4A
NS4B	Non-structural 4B
NS5A	Non-structural 5A
NS5B	Non-structural 5B
nt	nucleotide
o/n	Overnight

ORF	Open Reading Frame
OSBP	Oxysterol binding protein
OST	One strep tag, see 1ST
P	Pro, proline
PAGE	Polyacrylamide gel electrophoresis
PBD	Protein Binding Domain
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein database
PDB	Protein Data Base
peg	Pegylated
PFA	Paraformaldehyde
PI4KIIIα	Phosphatidylinositol-4-kinase-α, type III
PI4P	Phosphatidylinositol-4-phosphate
PIPES	Piperazine-N,N'-bis (2-ethanesulfonic acid)
PIV5	Parainfluenza virus 5
PKR	Protein Kinase R
PLB	Passive lysis buffer
PlkI	Polo-Like Kinase 1
PP2.1	Polyproline 2.1
PP2.2	Polyproline 2.2
PPIase	Peptidyl prolyl isomerase
PSTCD	P. shermanii transcarboxylase domain
PTEN	Phosphatase and tensin homolog
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
Q	Gln, glutamine
qRT-PCR	Quantative, Real Time PCR
R	Arg, arginine
RBV	Ribavirin
RdRp	RNA-dependent RNA-polymerase
RFP	Red fluorescent protein
RNA	Ribonucleic acid
rpm	Revolutions Per Minute
RT	Room Temperature
S	Serine
S	Ser, serine
SARS	Severe acute respiratory syndrome
SDS	Sodium dodecyl sulphate
sec	Seconds
SEM	Standard error of the mean
SGR	Sub-Genomic Replicon
shRNA	short-hairpin RNA
siRNA	Small interfering RNA
SMV	Single membrane vesicle
SNP	Single Nucleotide Polymorphism
SR-BI	scavenger receptor b1

SVR	sustained virological response
T	Thr, threonine
TAE	Tris-Acetate-EDTA buffer
TC	Tissue culture
TIP47	Tail-interacting protein of 47 kDa
TJ	Tight Junction
TMR	Tetramethylrhodamine
TOF	Time of flight
UTR	Untranslated Region
V	Val, valine
v/v	Volume by Volume
VAMP	Vesicle-associated membrane protein
VAP	VAMP-associated protein
VLDL	Very-Low Density Lipoprotein
VLDL	Very-Low Density Lipoprotein
W	Trp, tryptophan
w/v	weight by volume
WHO	World Health Organisation
WNV	West Nile Virus
WT	Wildtype
Y	Tyr, tyrosine
YFV	Yellow Fever Virus

Chapter 1: Introduction

1.1 The Hepatitis C Virus

1.1.1 Background to the HCV virus

The Hepatitis C virus (HCV) was first identified in 1989 as the causative agent of non-A non-B hepatitis (NANBH) (van der Poel et al., 1989, Kuo et al., 1989). It is a blood-borne pathogen and is typically transmitted from exposure to blood products of an infected individual (Chen and Morgan, 2006). Prior to the identification of HCV the major route of infection was through medical transfusion of unscreened infected blood products. After the implementation of screening the incidence of transfusion acquired HCV dramatically reduced to around 1:2,000,000 (Schreiber et al., 1996). As a result in the US there was a decline in new cases of HCV from an estimated 180,000 in 1988 to around 30,000 by 1995 (Alter et al., 1999). However increased awareness of disease transmission on the back the HIV epidemic is likely also a factor in this reduction. Despite these dramatic reductions in transmission, due to the asymptomatic nature and long duration until clinical presentation of symptoms, it is likely that the peak of patients with transfusion acquired HCV is only just being identified. Major routes of transmission are now mostly confined to intravenous drug use (IDU), sexual contact and parenteral transmission.

Current estimates are a global prevalence of around 170 million infected individuals (Chen and Morgan, 2006). Whilst HCV is endemic in all countries for which there are data, the prevalence varies greatly between countries. The highest prevalence is found in African and Eastern Mediterranean regions (2.5 to 10 %) while the vast majority of other countries have a lower prevalence of 1-2.5 % (Shepard et al., 2005). However in some countries this has been reported as significantly higher, principally in Egypt where reports are of a prevalence of up to 28% of the population (Frank et al., 2000).

HCV is classified into 6 major genotypes (1-6), with the recent proposal of a seventh genotype (Murphy et al., 2007). Each genotype is divided into multiple subtypes (a, b, c etc.), illustrated in Figure 1.1 (Simmonds, 2004). The genetic variability between genotypes is around 31-35 %, and between subtypes 20-25 %, and within an infected individual there are multiple quasispecies of virus present. The presence of quasispecies is likely the product of a long duration of infection and the error prone nature of the polymerase. The presence within these quasispecies of naturally occurring resistance mutations to the new direct acting antivirals (DAA) is currently under surveillance (Pawlotsky, 2006, Newman et al., 2013, Thomas et al., 2012). Genetic recombination of HCV is considered a rare event, with less

than 1 % of genomes in a cohort of over 1000 viruses being identified as recombinants (Shi et al., 2012).

The distribution of genotypes throughout the world is not uniform, Figure 1.1. In Western Europe, the Americas, Russia and Australia there is a dominance of genotype 1 infections, with either genotype 2 or 3 being the second most prevalent. However in Africa, the Far East and the Middle East there are much higher rates of genotype 4, 5 and 3 infections, with genotype 1 infections contributing only a small number of cases (WHO, 2009). Interestingly, the sequence homology between the most common 1b genotype and the predominantly research 2a genotype is as low as 52-55 %. As these two genotypes respond very differently to current treatments, it does raise the question as to how discoveries in genotype 2a will translate to the more clinically relevant, and harder to treat, genotype 1b.

HCV is a member of the *Hepacivirus* genus in the *Flaviviridae* family, Figure 1.1. While this genus is host to other animal viruses, HCV is currently the sole human pathogen found in this genus. Other non-human viruses within the *Hepacivirus* genus are GB virus-B (GBV-B) and non-primate *Hepacivirus* (NPHV) that have been identified in dogs, horses and rodents (Kapoor et al., 2011, Kapoor et al., 2013). The recent identification that bats are a natural reservoir for both *Hepacivirus* and *Pegivirus* could support a model where by an ancient zoonotic leap of *Hepacivirus* from bats to humans was the origins of HCV (Quan et al., 2013). The *Flaviviridae* family contains many well-known pathogens, the *Flavivirus* genus is host to Dengue virus (DENV), West Nile virus (WNV) and Yellow fever virus (YFV); the *Pestivirus* genus, host to bovine viral diarrhoea virus (BVDV); the *Pegivirus* genus, host to the George Barker virus-D (GBV-D). Similarities between the function of proteins from this family of viruses is reviewed in (Murray et al., 2008).

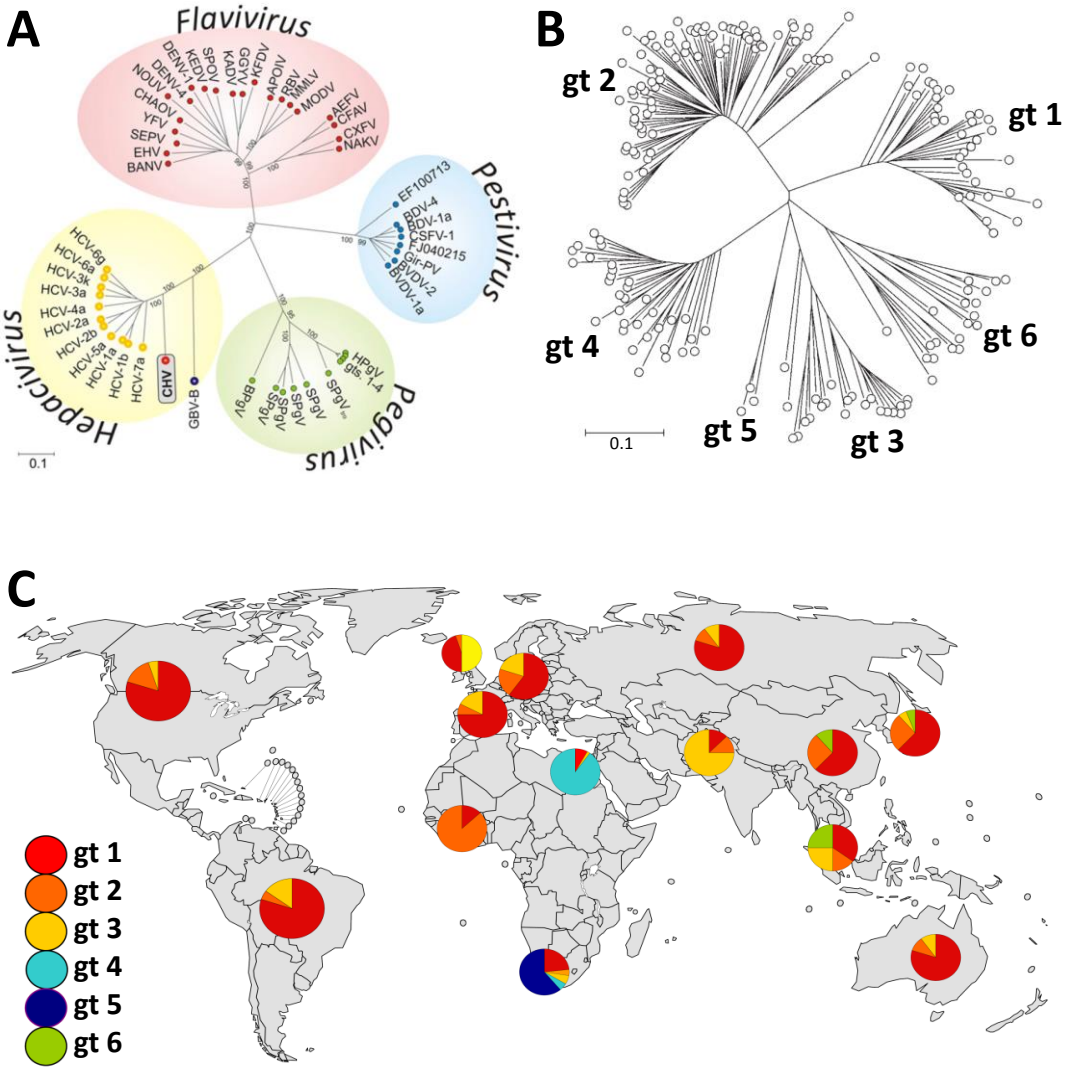


Figure 1.1. Phylogenetic analysis of the *Flaviviridae* family and HCV genotypes and their worldwide distribution.

A. Homology of *Flaviviridae* based on sequence comparison of the viral polymerases, reproduced from (Kapoor et al., 2011). **B.** Sequence homology of HCV genotypes based on the NS5B sequence, adapted from (Simmonds, 2004). **C.** Distribution of HCV by genotype across all major continents and sub-continents, reproduced from (WHO, 2009).

1.1.2 Disease pathology

The natural progression of HCV, in the absence of treatment, has been well documented and is summarised in Figure 1.2, reviewed in (Chen and Morgan, 2006, Mauss et al., 2011b). After exposure to the Hepatitis C virus the initial acute phase of infection is often asymptomatic, however clinical markers such as aminotransferases become elevated around 6-12 weeks post-exposure and seroconversion often occurs between 8 weeks to several months post-exposure (Mauss et al., 2011a). Where symptoms are present they are often general malaise and nausea. The self-clearance of HCV RNA within this acute infection phase occurs in about 20 % of infected individuals (Chen and Morgan, 2006). A very rare progression of disease is the onset of fulminant hepatitis soon after the acute infection stage, typically progressing to complete hepatic failure (Farci et al., 1996).

Individuals who have detectable HCV RNA in the blood and/or liver beyond 6 months are considered to be chronically infected, with approximately 80% of infected individuals progressing to this stage (Chen and Morgan, 2006). The risk factors for progression into chronic infection includes age, gender, ethnic origin and co-infections, and once a chronic infection has been established the spontaneous resolution of HCV is very rare. During this chronic phase, prior to the onset of cirrhosis, patients often remain asymptomatic but symptoms can include fatigue, nausea and weakness (Lauer and Walker, 2001).

It has been estimated that between 10-50 % of patients will progress to cirrhosis within 20 years of infection, with approximately 30 % of patients not progressing for at least 50 years, if at all (Poynard et al., 1997, Seeff, 2002, Mauss et al., 2011a). The risk factors associated with the rate at which a patient will progress from chronic HCV to the onset of cirrhosis is not fully understood, but known to include excessive consumption of alcohol, gender, age at time of infection as well as co-infections (Chen and Morgan, 2006, Seeff, 2002).

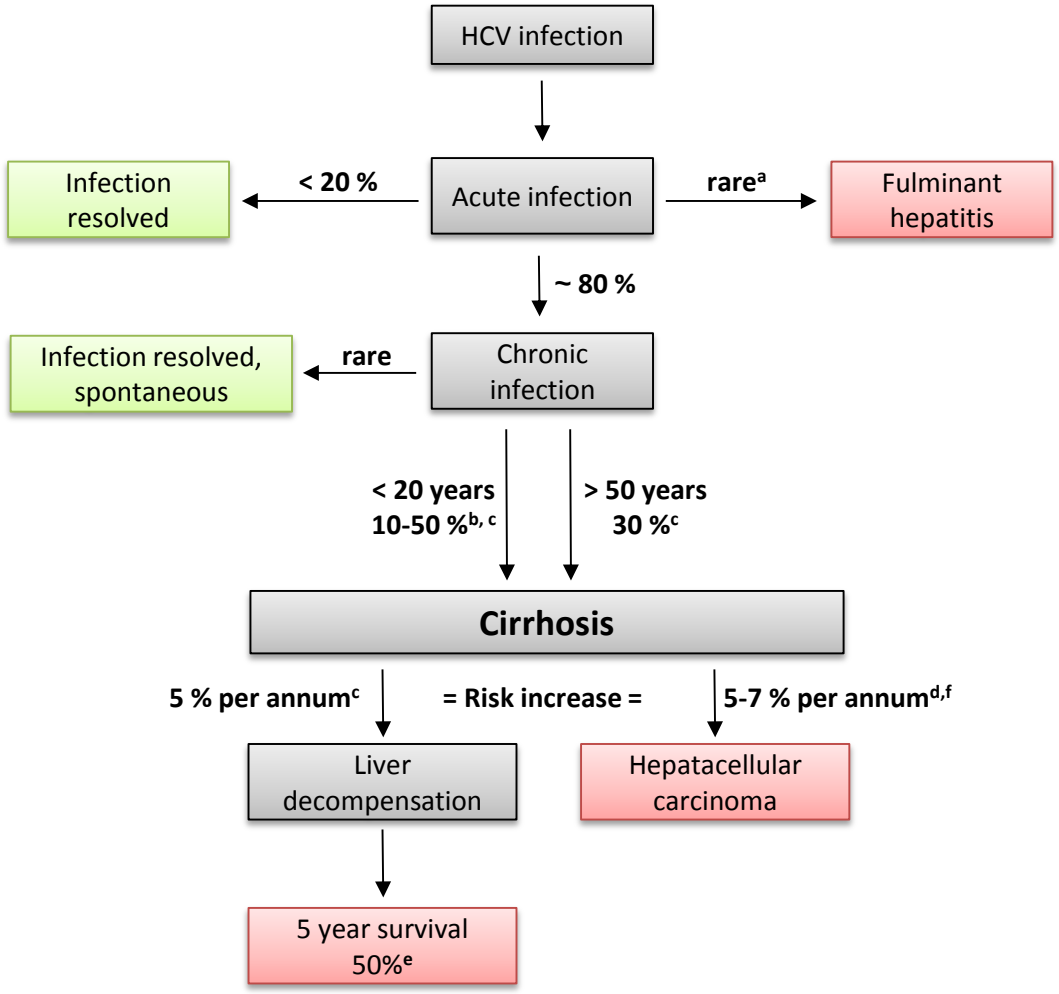


Figure 1.2. Natural disease progression of hepatitis C virus infection.
Summary of disease progression in the absence, or failure from, therapeutic treatment. Boxes denote permanent self-resolution of infection (green) or end stage liver disease followed by death (red). [a](Farci et al., 1996); [b](Mauss et al., 2011a); [c](Poynard et al., 1997); [d](Tang and Grise, 2009); [e](Planas et al., 2004); [f](Fattovich et al., 1997a)

Cirrhosis is a result of the progressive fibrosis of hepatic tissue, a product of the continual deposition of collagen by activated stellate cells in a complex, immune mediated response to HCV infection (Guidotti and Chisari, 2006). Initially fibrosis is limited to the portal region but eventually progresses to irreversible bridging fibrosis between portal regions. The standard in assessing the extent of fibrosis is histological staining of liver biopsies and observational grading the extent of fibrosis, the presence of mononuclease inflammatory cells and extent of hepatocyte death (Theise, 2007). While the use of liver biopsy is currently considered the standard for grading liver fibrosis, new and emerging techniques allow for the non-invasive grading of liver disease by measuring liver elasticity and viscosity through a variety of methods including ultrasound and magnetic resonance (Faria et al., 2009, Huwart and van Beers, 2008). These techniques are more suited to regular screening, monitoring and follow-up than traditional liver biopsy sampling.

After the diagnosis of HCV induced hepatic cirrhosis the risk for progression to decompensated cirrhosis is estimated to increase at 5 % per annum (Poynard et al., 1997). After which the 5 year survival rate is around 50 % following decompensation of hepatic function (Planas et al., 2004).

Alongside the eventual decompensation of hepatic function there is an increase prevalence of hepatocellular carcinoma (HCC) amongst patients with HCV induced cirrhosis. Note that this increased prevalence is primarily only in cirrhotic HCV patients, contrary to HBV infections where the increased prevalence of HCC is observed in both chronic HBV and cirrhotic patients (Tang and Grise, 2009). The primary link between HCV and HCC was established through epidemiological observations, with about 70% of HCC cases occurring in either HCV or HBV infected patients, however insight into the possible mechanistic links is starting to be investigated (Levrero, 2006, Tang and Grise, 2009). The risk of HCC developing was shown to be between 1 to 4 % per annum post diagnosis of cirrhosis, although in certain populations this has been reported in the range of 5 to 7 % per annum (Fattovich et al., 1997b, Chen and Morgan, 2006).

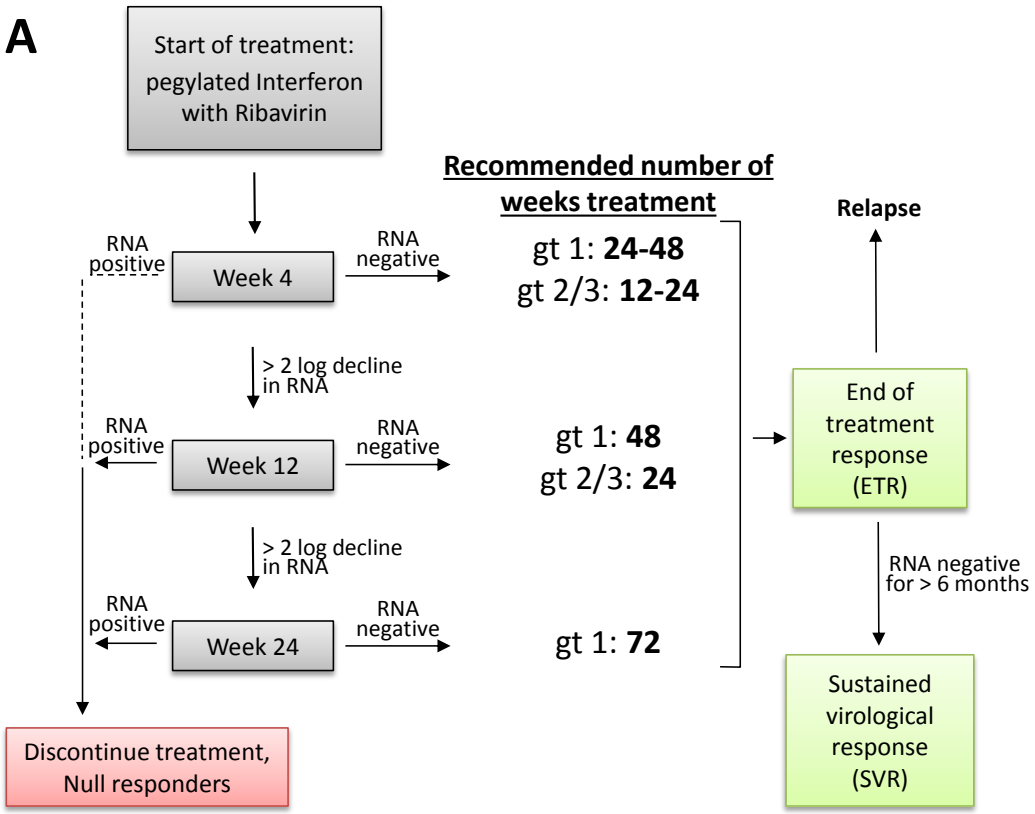
The exact role of the innate and adaptive immune response in viral clearance is not fully understood, but self-clearance of virus infection does require a rapid response from both of these elements of the immune response. Concomitant with an elevation of CD4+, CD8+ T-cells and well as intrahepatic interferon-γ levels there is rapid decline in viral titre (Feinstone et al., 2012). Interestingly seroconversion is not a requirement for the self-clearance of virus,

and does not confer a protective immunity upon future re-exposure. However, viral kinetics and duration are generally shorter in reinfected individuals, with a reduced risk of a chronic infection establishing (Bowen and Walker, 2005, Feinstone et al., 2012, Post et al., 2004).

1.1.3 Current and future therapeutics

With the exception of the emerging direct acting antivirals (DAAs) discussed shortly, the established and most successful treatment for HCV infection is a combination of pegylated Interferon-α2 (peg-IFNα2) and ribavirin (RBV) for varying durations. The treatment cascade is summarised in Figure 1.3. A sustained virological response (SVR), where HCV RNA is undetectable 6 months after completing treatment, is used to define the efficiency of anti—HCV treatments. Dual therapy of peg-IFNα2/ RBV has reported efficacy of between 46 - 82 % depending on the genotype treated, with genotype 2/3 giving the more favourable response and genotype 1 being harder to treat.

This is a great improvement from original SVR rates of 22 % seen following monotherapy of IFNα2 for 78 weeks in the past, but still leaves a huge number of patients with untreatable HCV infections. Note that the pegylation of IFNα2 was to improve the bioavailability, and so improve the dosing regimen, of IFNα2 treatment.



B

Drug name	Drug ID	Company	Target, activity	Stage of development	Expected licensed
Telaprevir	VX-950	Janssen	NS3/4A, protease	Licensed , 2012	-
Boceprevir	SCH503034	Merck	NS3/4A, protease	Licensed , 2012	-
Sofosbuvir	GS-7977	Gilead	NS5B, polymerase	Phase III	2014
Daclatasvir	BMS-7900052	Bristol-Myers Squibb	NS5A, n/a	Phase III	2014/5
Alisporivir	DEB025	Novartis	Cyclophilin A, PPIase	On hold (phase III)	n/a

Figure 1.3. Current and emerging treatments for HCV infection.
A. Clinical treatment cascade for HCV infection with peg-IFNα2/RBV. **B.** Direct acting antivirals (DAAs) currently under investigation. Alisporivir was put on a partial clinical hold by the FDA due to a small number of cases of pancreatitis, phase II studies have recently recommenced.

The correlation between an early virological response (EVR) and obtaining an SVR at the end of treatment is well established. Indeed it serves as a key guide for the early discontinuation of treatment should EVR not be achieved (Figure 1.3). Currently following combination therapy of peg-IFNα2 and RBV for 12 weeks, if there is less than a 2-log drop in blood HCV RNA levels then there is grounds for the discontinuation of treatment. As is the case at 24 weeks, were any detectable HCV RNA gives grounds to discontinue treatment. Following a genotype and virological response guided duration of peg-IFNα2 and RBV treatment (12, 24, 28 or 72 weeks), the absence of detectable HCV RNA is termed an end of treatment response (ETR). From this point patients will either stay HCV RNA negative, and so achieve a SVR at 6 months and be considered 'cured', or will relapse back into viremia. The treatment cascade for relapse is similar to that of treatment naïve patients, with genotype and virological response guiding duration of treatment.

It is worth noting that the side effects from the current dual therapy are sufficient to cause 10 - 20 % of patients to prematurely withdraw from treatment prior to achieving an ETR (Manns et al., 2006). The most commonly report symptoms are "flu-like", including headaches, pyrexia, nausea and fatigue, and occur in approximately 50 % of patients (Mauss et al., 2011a). Other side effects include severe depression, neutropenia and negative effects on bone marrow function (Manns et al., 2006). These side effects are predominantly associated with the peg-IFNα2, but adverse reactions to RBV are also reported to included haemolytic anaemia (Reddy et al., 2007).

Recently the identification of a variety of single nucleotide polymorphisms (SNPs) in the IL28B locus have been shown to be a strong indicator of both response to treatment and spontaneous clearance of HCV. Specifically the CC genotype at SNP rs12979860 has been shown to be most favourable for treatment response (Ge et al., 2009). The current understanding and clinical applications of SNP analysis in the treatment of HCV is reviewed extensively in (Rau et al., 2012).

The pipelines of pharmaceutical companies are brimming with new chemical entireties (NCE) currently under investigation for their activity towards different elements of the HCV virus. There are currently in excess of 1,600 clinical trials listed on clinicaltrial.gov evaluating the efficacy of new anti-HCV therapies. This equates to approximately 38 NCE in the late stage of clinical development which target a broad spectrum of virus and host protein (phase II-III) (Manns and von Hahn, 2013).

The first two direct acting antiviral drugs to be licensed in 2012 were the orally available Telaprevir (Janssen) and Boceprevir (Merck), reviewed in (Klibanov et al., 2011, Klibanov et al., 2012, Chen and Njoroge, 2010). These peptidomimetics target the NS3/4A protease complex in a covalent, but reversible manner (Manns and von Hahn, 2013). The addition of protease inhibitors to the current triple therapy regime has brought SVR rates for the typically harder to treat genotype 1a/b from around 46 % to 70 %, in line with the SVR rates observed previously with a peg-IFNα2/RBV regime against genotype 2/3 (Manns and von Hahn, 2013). Clinical trials are currently underway for the licensing of interferon free combinations of DAAs.

Several other key drugs rapidly approaching expected licensing dates are highlighted in Figure 1.3B. Sofosbuvir (Gilead), a nucleotide analogue targeting NS5B was reported at ESAL 2013 to have made excellent phase II progress, with certain trials reporting a SVR of 90% (pan-genotype) after 12 or 24 weeks treatment in combination with peg-IFNα/Rib. While other trials did report a lower SVR rate for an interferon free dual therapy of Sofosbuvir/RBV, those patients did all reach undetectability, with relapse being the primary cause of non-SVR. Interestingly in a manner that correlated with IL28B alleles (Barreiro et al., 2013).

The front running NS5A inhibitor Daclatasvir, (Bristol Myers-Squibb, BMS) has also made excellent clinical progress (Lemm et al., 2010). The most exciting clinical trial to date showed that a combination of Daclatasvir and Sofosbuvir - dosed once daily, with or without RBV and over a 12 or 24 week course - achieved virtually 100 % SVR in patients with genotype 1 infections, including difficult-to-treat patients (Sulkowski et al., 2012). The current state of new and emerging therapeutics is reviewed further in (Manns and von Hahn, 2013). The reduced viral kinetics and broader T-cell response observed in reinfected individuals is evidence of an adaptive immune response. This, along with the high cost of emerging DAAs and occurrence of drug resistance mutation, has encouraged the development of prophylactic and therapeutic vaccines (Feinstone et al., 2012). Currently approaches have focused on stimulating the adaptive immune response, as a weak HCV-specific T-cell response is typical in chronically infected patients. However, to data, vaccines have only achieved a transient reduction in viral load (Feinstone et al., 2012).

1.2 Molecular biology of HCV

1.2.1 Genomic organisation

HCV has a single stranded, 9,600 nucleotide positive-sense RNA genome containing 5' and 3' untranslated regions (UTRs) and two open reading frames (ORF). The primary ORF codes for a ~3000 aa polyprotein which is cleaved co- and post-translationally by both host and viral proteases to produce 10 mature viral proteins, Figure 1.4. The second ORF overlaps with the start of the primary ORF and codes for a single protein, the 'F' or alternative reading frame protein, ARFP protein. The reference genome for HCV is the genotype 1a H77 isolate (accession number AF009606) and the amino acid number in the polyprotein is typically used to describe the position of residues (Kuiken et al., 2006). However as this thesis focuses solely on the genotype 2a JFH-1 isolate, numbering herein refers to the NS5A amino acid position and not the H77 polyprotein numbering.

The first three proteins to be cleaved from the polyprotein by host proteases are the Core, Envelope 1 and 2 (E1 and E2) proteins and these are the principal components of the virus particle. Core forms the nucleocapsid while E1 and E2 decorate the phospho-lipid membrane envelope. The following 7 proteins are denoted as non-structural (NS) protein due to their absence from the virion; p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B. However, the minimal genetic requirement for autonomous replication of the RNA genome is the 5' and 3' UTRs and NS3-5B proteins. In fact inclusions of the p7 and NS2 coding region have been shown to impair replication in cell culture systems.

The 5' UTR is highly structured and composed of five domains (I, II, III, IV and V) of which domains II, III and IV and along with the first 12-30 nt of the core coding region form an internal ribosome entry site (IRES) (Honda et al., 1996). The translation of both ORFs is driven from this IRES in a cap-independent manner to produce both the polyprotein and the F protein. The 5' UTR also contains two binding sites at the very N-terminus for the micro RNA miR-122, the function of which is discussed shortly.

The 3' UTR is a ~230 nt stretch following from the end of the polyprotein coding sequence. It is composed of a variable region, a polyuridine and polypyrimidine tract (polyU/UC) and a highly conserved 3' X-tail composed of three stem loops (Blight and Rice, 1997, Ito and Lai, 1997, Kolykhalov et al., 1996). Whilst deletion of the variable region only impaired replication, the deletion of either the polyU/UC or the 3' X-tail completely abrogated replication (Friebe and Bartenschlager, 2002, Yanagi et al., 1999).

1.2.2 Structural proteins

Core is a dimeric protein composed of two domains (Boulant et al., 2005). The first domain is hydrophilic and proposed to bind viral RNA during the nucleocapsid formation, although precise packaging signals have yet to be identified. The second domain is a hydrophobic domain that is responsible for anchoring core to lipid membranes such as the ER and lipid droplets (Boulant et al., 2006). The core protein requires an N-terminal cleavage by signal peptide peptidase protease in order to traffic from the ER to lipid droplets (LDs) (McLauchlan et al., 2002). A disruption of the interaction between core and lipid droplets is sufficient to block virus release (Boulant et al., 2007).

The envelope proteins 1 and 2 (E1, E2) are the key components of the virus envelope and are essential for viral entry and fusion. They are highly glycosylated type I transmembrane proteins that, when in matured virions, have been shown to form large complexes stabilised by disulphide bridges (Vieyres et al., 2010). Conversely, intracellular E1/E2 is observed as non-covalent heterodimer and this shift to large complexes could be indicative of virion maturation post-assembly (Lavie et al., 2007).

1.2.3 Non-structural proteins

The structural proteins are followed by p7, a viroporin that is an essential factor for the release of infectious virions (Griffin, 2010, Jones et al., 2007a). The p7 protein in concert with NS2 recruits core from lipid droplets to the proposed site of virus assembly on ER-derived membranes in a manner essential for virus production (Boson et al., 2011). Moreover it was recently shown that the ion channel activity of p7, and not the formation of protein-protein interactions, is required for the release of infectious virus at a late stage (Bentham et al., 2013). Structurally, p7 is composed of two transmembrane domains that are believed to oligomerise into either a hexamer or heptamer forming an ion channel (Bentham et al., 2013, Griffin, 2010, Foster et al., 2013). The ion channel activity is cation-selective and is thought to be gated by the lumen residue His17 (Chew et al., 2009).

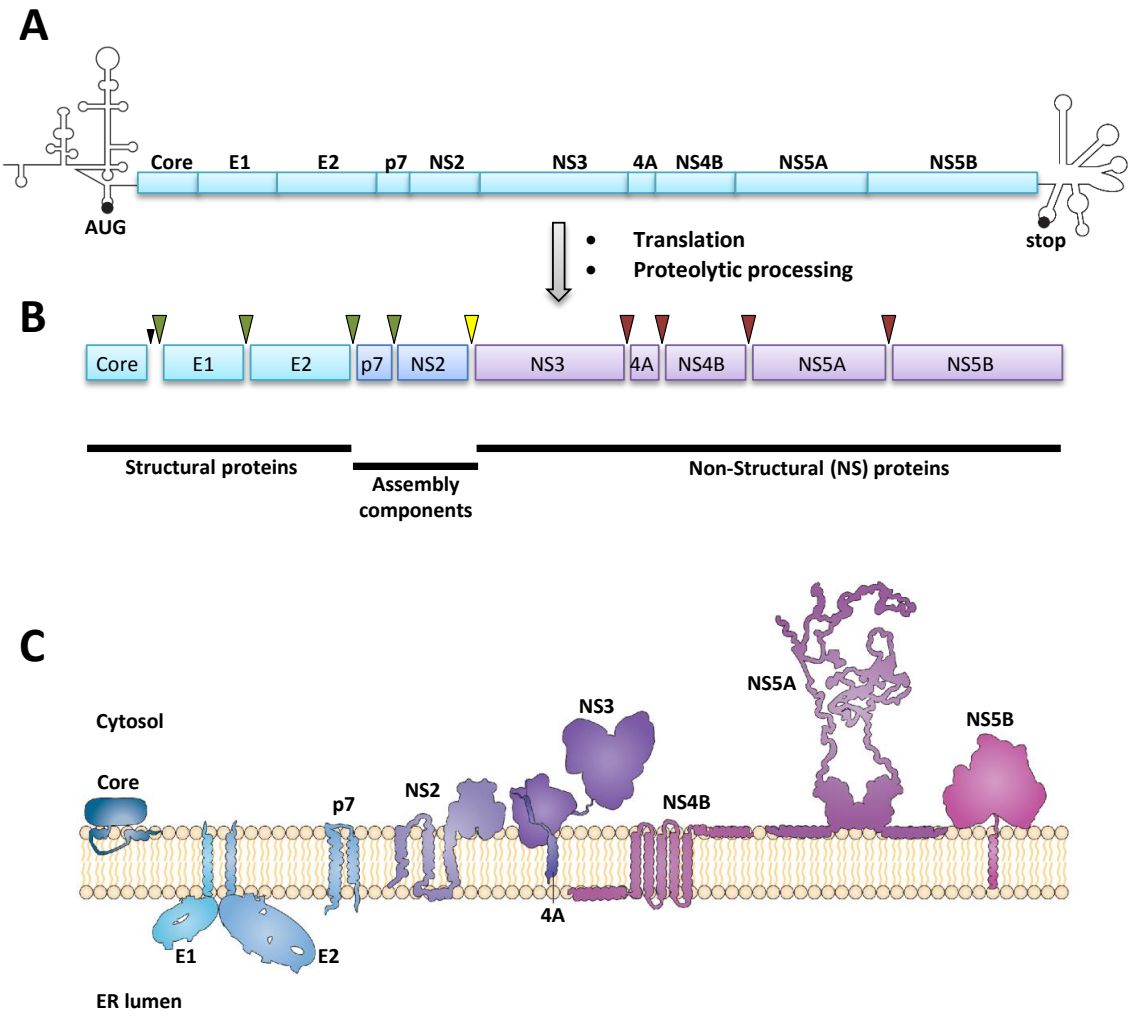


Figure 1.4. Genome organisation and polyprotein processing of the HCV virus.
A. HCV 9.6 kb genome, 5' and 3' UTRs are indicated with start (AUG) and stop codons indicated. **B.** The HCV polyprotein is cleaved by host proteases (green), the NS2 protease (yellow) and the NS3/4A protease complex (red). Following cleavage from the polyprotein the core protein is cleaved further by the signal peptide peptidase (black, small) to produce the mature core. **C.** Following proteolytic processing of polyprotein HCV proteins are anchored to ER-derived membranes with the indicated orientation. Part C reproduced from (Bartenschlager et al., 2013).

The NS2 protein is a membrane anchored, 24 kDa protein composed of up to three transmembrane domains and a catalytic C-terminal protease domain. The protease is responsible for cleaving on the NS2-NS3 protein junction in the polyprotein, however as commented earlier, the replication of sub-genomic replicons is increased in the absence of the NS2 protein. In recent years a key role for NS2 in the assembly of infectious virus has been shown, while transmembrane and catalytic domain are both required for this function, the catalytic activity of the protease is not (Jirasko et al., 2008, Jones et al., 2007a). Mutations which disrupt NS2 interactions with p7, E2, NS3, and to a reduced degree NS5A, were shown to block virus production (Jirasko et al., 2010). As such it has been proposed that NS2, in concert with p7, is involved in orchestrating virus assembly at an early stage (Boson et al., 2011, Jirasko et al., 2010).

The NS3 protein is a large (69 kDa) multifunctional protein composed of two domains, an N-terminal serine protease (residues 1 to ~180) linked by a flexible domain to a C-terminal helicase domain with NTPase activity, reviewed in (Lin, 2006). The protein is anchored to ER-derived membranes by both an N-terminal amphipathic helix and the NS4A protein. The NS4A protein is a small (6 kDa) co-factor essential for the serine protease activity of NS3, contributing an essential β-sheet to the protease domain. NS3 is a type I membrane protein that has a short N-terminal transmembrane domain that inserts into membranes post-NS3 driven cleavage from the polyprotein. The NS3 protease is responsible for cleavage of both itself and the remaining non-structural proteins from the polyprotein. The helicase activity of NS3 has been shown to unwind dsRNA and RNA-DNA hetero-duplexes in a 3'-5' direction, in an ATP-dependent manner (Serebrov and Pyle, 2004, Tai et al., 1996). Like many of the HCV viral proteins, NS3 and NS4A proteins have been shown to interact with host and viral proteins in other stages of the virus lifecycle including virus assembly (de Chassey et al., 2008, Ma et al., 2008).

The NS4B is a 27 kDa integral membrane protein consisting of a two N-terminal amphipathic helices, four central transmembrane domains and a cytosolic C-terminal domain (CTD) (Lundin et al., 2003). It is proposed that the second of N-terminal amphipathic helices acts as a transmembrane helix upon oligomerisation, and so giving the first helices a dual membrane topology, luminal and cytosolic (Lundin et al., 2006). Primarily, the NS4B protein is critical in driving the rearrangement of cellular membranes to form the membranous web upon which replication machinery forms, discussed in detail later (Section 1.3.2), but it is also involved

with binding other viral proteins (Aligo et al., 2009, Dimitrova et al., 2003), binding the RNA genome (Einav et al., 2008) and recently in the production of infectious virus (Han et al., 2013). Reviewed further in (Gu and Rice, 2013).

The NS5A protein is a highly phosphorylated, multi-functional protein that has essential roles throughout many stages of the virus life cycle. It is introduced in more detail in (Section 1.6)

The NS5B protein is the RNA-dependent RNA-polymerase (RdRp) responsible for the synthesis of both intermediary negative-sense and positive-sense RNA genomes. It has a C-terminal transmembrane domain anchoring it to the cytosolic face of the ER-derived membranes and the three classic 'fingers, palm and thumb' subdomains homologous to other viruses (Choi and Rossmann, 2009). The catalytic 'palm' domain contains a GDD motif responsible for coordinating two magnesium ions. Mutation of this motif, typically GDD->GND, disrupt this coordination and render the polymerase inactive, and this mutation is regularly used as a 'non-replicating' negative controls in experiments.

1.3 The HCV life cycle

1.3.1 Binding and Entry of virus

The current understanding is that the circulating "infectious particle" of HCV is composed of an enveloped viral particle in tight association with VLDL/LDL particles, Figure 1.5A (Nielsen et al., 2006, Vieyres and Pietschmann, 2013). A recent model suggest that in fact the virus envelope and the VLDL particle might actually be a single unit, with lipidation occurring in between the phospho-lipid membrane of the virus envelope (Bartenschlager et al., 2011). This infectious particle is known to enter hepatocytes from the sinusoidal blood in a complex receptor-mediated, clathrin-dependent endocytosis process involving at a minimum the following receptors; heparin sulphate glycosaminoglycans (GAGs) (Barth et al., 2003); the low density lipoprotein receptor (LDL-R) (Agnello et al., 1999, Molina et al., 2007, Owen et al., 2009); the scavenger receptor class B type I (SR-BI) (Evans et al., 2007, Scarselli et al., 2002); the tetraspanin CD81 (Pileri et al., 1998); and the two tight junction (TJ) proteins, claudin-I (Evans et al., 2007) and occludin (Ploss et al., 2009).

It is not clear whether claudin-I and occludin interact directly with incoming virus particles, as hepatocyte tight junctions are directly inaccessible to the circulating viral particles. It is thought that binding of viral particles to the aforementioned receptors could result in translocation of receptor : virus complexes to these tight junctions where endocytosis could

subsequently occur, a mechanism paralleled in the Coxsackievirus (Coyne and Bergelson, 2006). Indeed CD81-complex translocation to cell junctions after receptor engagement (by E2 or αCD81) has been previously observed (Brazzoli et al., 2008).

The binding of the virus to the LDL-R is mediated through virus associated apolipoprotein E (ApoE), an essential component of the infectious particle discussed shortly (Owen et al., 2009). This is consistent with observations that antibodies against both envelope proteins and ApoE proteins can neutralise the infectivity of viral particles (Owen et al., 2009, Jiang and Luo, 2009). The current understanding is that the envelope proteins are shielded from immune surveillance through a combination of extensive glycosylation and being intimately associated in the VLDL particles (Helle et al., 2007, Bartenschlager et al., 2011). This would be compatible with a model whereby the initial attachment to hepatocytes would occur via the glycosylated envelope proteins interacting with GAGs, and ApoE binding the LDL-R and SR-BI receptors. These initial binding events in turn could promote the interaction of envelope proteins with the CD81 receptor, and subsequent translocation to tight junctions for claudin-I and occluding mediated endocytosis (Bartenschlager et al., 2011, Popescu et al., 2011b). Such a model would allow for minimal exposure of envelope proteins to the circulating immune system. It is worth noting that the majority of our understanding of the HCV entry process has been derived from the non-polarised, hepatocellular carcinoma cell lines, Huh7 and HepG2. Extrapolation of these mechanisms to polarised hepatocytes like those observed in hepatic tissue needs to be carefully measured.

After the highly orchestrated attachment of the infectious particle, internalisation is known to occur through clathrin-mediated endocytosis (Blanchard et al., 2006, Coller et al., 2009). Following endocytosis the infectious particle is trafficked to early endosomes, thought to be in a Rab5a dependent manner (Coller et al., 2009). In support of this i*et al.* the co-endocytosis of CD81 and claudin-1 followed by fusion with Rab5 positive endosomes was recently visualised (Farquhar et al., 2012). The subsequent acidification of the endosome results in envelop directed membrane fusion and release of RNA into the cytoplasm.

The targeting of HCV entry is an attractive therapeutic option, and several monoclonal antibodies targeting viral and host entry factors are currently being developed. However only two antibodies, targeting the E1 and E2, have so far progress into phased clinical trials (Zeisel et al., 2011).

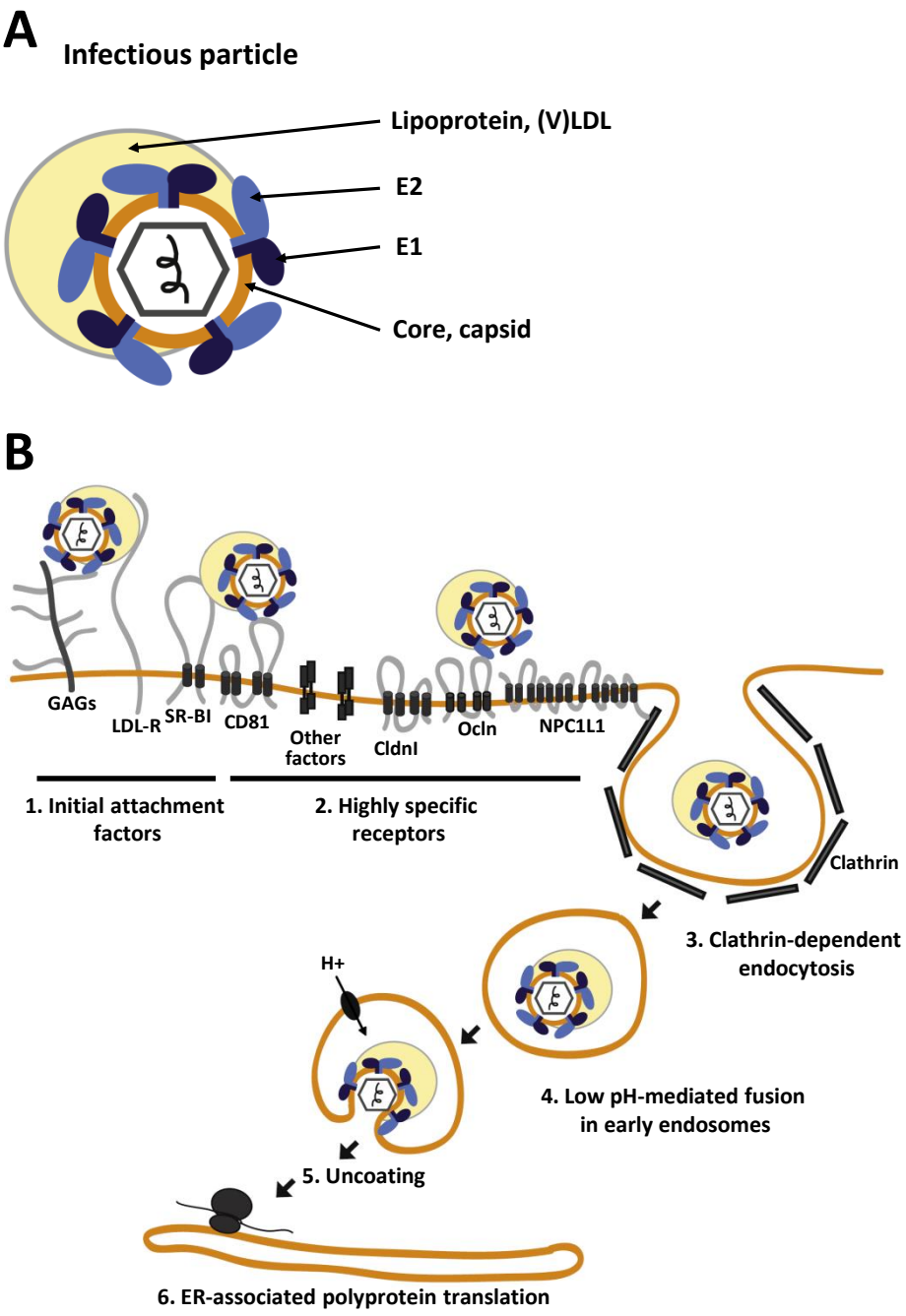


Figure 1.5. Structure and entry process of the HCV virus.
A. Cartoon structure of the HCV virion with E1 and E2 proteins (dark and light blue respectively), double membrane envelop (orange), capsid (black) and virus associated lipoprotein particle, (V)LDL (light orange). **B.** Illustration of the known elements of HCV entry process. Figures adapted from (Vieyres and Pietschmann, 2013).

1.3.2 Replication of viral RNA

Once in the cytoplasm the 9.6kb genome is rapidly translated into a ~3000 aa polyprotein that is cleaved co and post-translationally by host and viral proteases. This translation occurs rapidly, and within 8 hours of HCV RNA entering the cytoplasm the virus replication and translation of RNA exceeds that of degradation. From this point the intracellular levels of both viral proteins and genomes, positive and negative, exponentially increase until a steady state is reached at around 24 hours (Binder et al., 2013).

A key feature of many positive-sense RNA virus is the rapid remodelling of cellular membranes into what is termed, a membranous web, Figure 1.6A. These structures are often derived from ER-membranes, are rich in phosphatidylinositol-4-phosphate lipids (PI4P) and are the proposed sites of viral replication and virus assembly, as is the case in HCV (Delang et al., 2012). There are many benefits that this membrane remodelling is thought to confer, there is an aggregation of protein synthesis machinery and the environment is tightly excluded from the surveillance of the host cell innate immunity, Figure 1.6B.

In the case of HCV, the NS4B protein is the major viral protein required for the remodelling of ER-derived membranes. The host factor phosphatidylinositol-4 kinase-III alpha/beta (PI4KIIIα/β) is responsible for the phosphorylation of inositol rings to generate PI4P lipids and has also been shown to be essential in remodelling the ER-derived membranes; with knockdown or inhibition abrogating HCV replication, and in certain HCV genotypes entry and release (Berger et al., 2009, Borawski et al., 2009, Reiss et al., 2011, Tai et al., 2009). More recently it was shown that NS5A interacts with and stimulates the kinase activity of PI4KIIIα in a manner critical for the correct formation of the membranous web (Reiss et al., 2011, Reiss et al., 2013). The interaction of NS5A and PI4PKIIIα is discussed in further detail later (Section 1.6.2.1). Other host factors involved in the membranous web formation and virus replication are reviewed in (Berger and Randall, 2009).

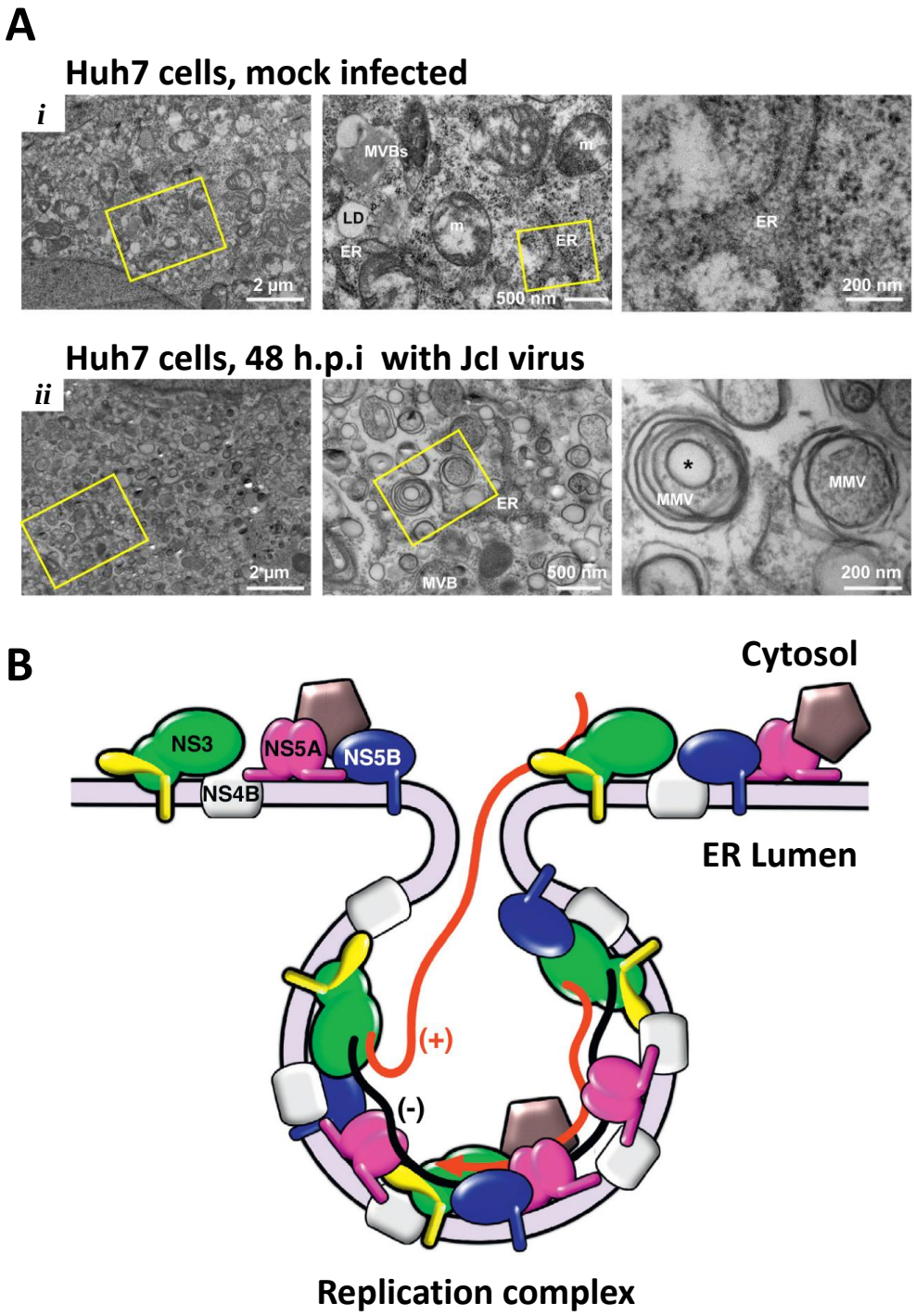


Figure 1.6. The membranous web and the structure of the replication complex.
A. Comparison of the perinuclear region of mock infected (i) and HCV infected (ii) Huh7.5 cells showing the formation of multi-membrane vesicles (MMVs). Reproduced from (Romero-Brey et al., 2012) **B.** Cartoon schematic of the replication complex showing NS3 (green) tethered to the ER-derived membrane by NS4A (yellow), in close association with NS4B (grey), NS5A (purple) and NS5B (blue). Other host factors are denoted by pentagon (violet). Adapted from (Gu and Rice, 2013).

The sub-structure of the membranous web has been shown to be composed of single and double membrane vesicles tightly packed together (SMV and DMV) and it is accepted that these are the structures upon which replication machinery is tightly associated (Romero-Brey et al., 2012). These structures were also recently shown to be significantly enriched in proteins like NS3, NS5A and NS5B as well has host factors like VAP-A, VAP-B and cholesterol (Paul et al., 2013). The VAP-A/B proteins are essential host factors discussed in detail later (Section 1.6.2). An illustration of these ER-membranes forming deep pockets containing the replication machinery is shown in Figure 1.6B.

Whether these DMVs are completely encircled or have small - possibly gated - pores has yet to be conclusively demonstrated, but recent EM-tomography is highly suggestive of such a model. As is the consistent finding that HCV RNA, HCV proteins and even *in vivo* replicase activity in isolated DMVs is both protease and nuclease resistant until the addition of detergent (Paul et al., 2013, Quinkert et al., 2005). Recent work has highlighted the involvement of elements from the nuclear pore complex (NPC), showing they were relocalised to the membranous web in HCV infected cells, and that knockdown was inhibitory to replication and release (Neufeldt et al., 2012). This is thought to be a potential mechanism for regulating the influx and efflux from either DMVs or the membranous web in general, perhaps now hinting at the function of putative nuclear localisation sequences found within NS5A.

Studies by Wolk *et al.* showed that the membranous web was composed of structures of different mobility. The fluorescent labelling of NS5A highlighted large, static structures that were stable and showed minimal NS5A exchange post bleaching. They also observed smaller, highly mobile structures that trafficked in a dynein dependent manner over long distances (Wolk et al., 2008). It is feasible that the larger static structures represent the aforementioned RC DMVs, while the small mobile structures represent either the recruitment of factors into replication complexes or NS5A fulfilling the numerous other functions it is involved with outside of replication.

It was recently shown that micro-RNA 122 (miR-122) is an essential host factor for HCV replication as well as virus production (Jangra et al., 2010). Exactly how miR-122 functions in the virus lifecycle is not fully understood but it is known to bind the 5' UTR at two binding sites in a non-canonical manner. Recent studies show that at a minimum miR-122 recruits Argonaute-2 (Ago-2) to the HCV genome and prevents degradation by the cellular

exonuclease Xrn1 (Li et al., 2013). But Li *et al.* concede that this is likely not the sole function of miR-122 that is critical for virus replication.

A paradox for positive-sense RNA viruses like HCV is their genome must fulfil multiple, mutually exclusive roles. The genome must serve as template for negative-strand synthesis by the RdRp, as well as a template for protein synthesis by the ribosomal complex and also being available for packaging into assembling virions. What regulates the fate of RNA genomes is complex and not fully understood. The quantification of genomes in RCs has shown a 10-1000 fold excess of positive-sense genomes (Komurian-Pradel et al., 2004, Quinkert et al., 2005). This hints that perhaps once RCs are established, negative-sense genomes are used for the repeated synthesis of an excess of positive-sense genomes. These can then be trafficked out of tightly excluded replication complex into the cytoplasm, from where they are directed to either protein synthesis or encapsidation.

1.3.3 Assembly and release of virions

The exact mechanics of how RNA is packaged into the virion and subsequently associates with VLDL to form the circulating infectious particle is still poorly defined, reviewed in (Bartenschlager et al., 2011, Popescu et al., 2011b). The association of core with lipid droplets has been well validated as critical for the production of infectious particle (Boulant et al., 2007, Miyanari et al., 2007, Targett-Adams et al., 2008). As the core protein is recruited (by domain II) to the lipid droplet (LD) there is a concomitant displacement of adipocyte differential related protein (ADRP) from the surface of LDs, and a redistribution of LD to the membranous web region (Boulant et al., 2008). The exact function of this association and redistributing is not fully understood, but it appears more complex than a simple “increase core on LD = increase in virus assembly” model. Indeed a comparison of the sub-cellular distribution of core from the high titre JcI chimera with that of the low titre JFH-1 virus, showed virus titre inversely correlated with the extent of core accumulation on lipid droplets. Furthermore, it was shown that this was a result of the difference in mobility of core domain II (Boson et al., 2011, Shavinskaya et al., 2007).

The association of core on LDs is likely to fulfil multiple roles. The LD could be utilised as a cellular trafficking mechanism, transporting core from sites of protein synthesis to sites of virus assembly (Welte, 2009); the accumulation of core on LD could also act as a reservoir for core, as significantly more core per RNA molecule is required for nucleocapsid formation; finally the LD itself could be the *de novo* site of RNA encapsidation. Although the

aforementioned inverse relationship between core accumulation on LD and virus titre suggests this is less likely (Shavinskaya et al., 2007). Interestingly in naïve Huh7 cells the LDs are only rarely closely associated with other organelles, however in HCV infected cells a significant percentage of the LD are in close association with the re-arranged ER-derived membranes (Miyanari et al., 2007).

It has been well established that the maturation of the virion into an infectious particle is highly dependent on the VLDL pathway (Chang et al., 2007, Gastaminza et al., 2008, Huang et al., 2007, Jiang and Luo, 2009). Typically, VLDL particles are synthesised at the ER and secreted from the cell via the Golgi (Bamberger and Lane, 1990). However, at what stage the core-encapsidated RNA acquires the E1/2 containing lipid envelope and subsequently associates with VLDL to form the infectious particle is poorly understood. The identification of oxysterol binding protein (OSBP) as a host factor that interacts with NS5A in a manner essential for virus maturation could shed some light on this (Amako et al., 2009). It is known that the HCV envelope proteins traffic through the cisternae of the Golgi compartment for glycosylation (Op De Beeck et al., 2004).

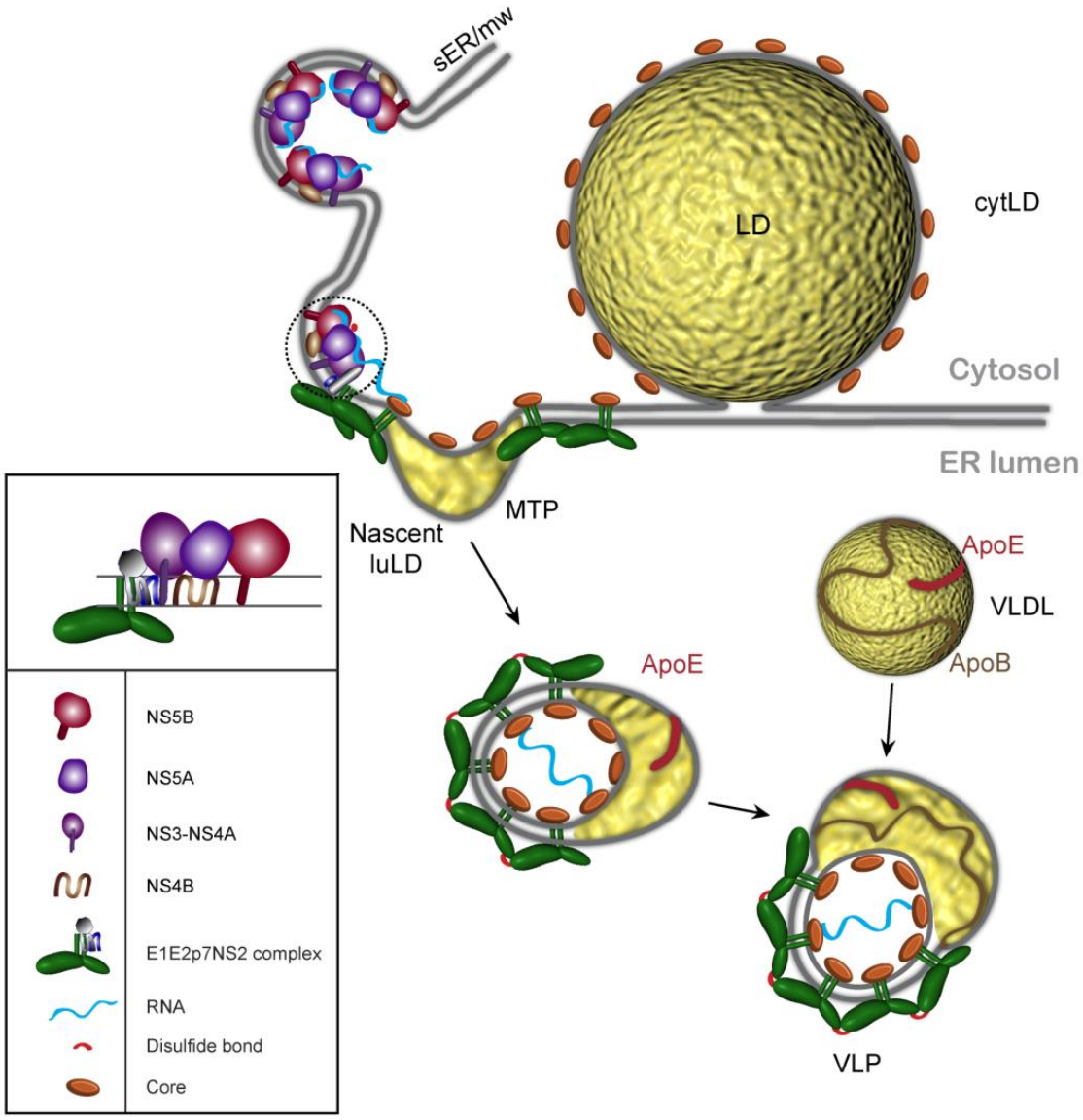


Figure 1.7. A proposed model for the assembly of infectious virus.

Illustration showing the replication of RNA in membrane invaginations, then subsequent encapsidation of viral RNA in the close proximity to both ER-derived membranes and lipid droplets (LD). At what point the lipoviralparticle associates with VLDL components is unsure. Reproduced from (Popescu et al., 2011b).

Recently several other virus proteins have been shown to be essential for the assembly of infectious virus. As mentioned earlier the p7 and NS2 are both critical factors for the assembly and maturation of infectious virus (Section 1.2.3). The interaction of domain I of core (residues 64-66) with the helicase domain of NS3 (Jones et al., 2011) and elements within the C-terminal domain of NS4B are all required for virus assembly and egress (Han et al., 2013). Exactly how these proteins function in these processes is not fully understood. NS2 was shown to be able to interact with both structural (E2) and non-structural elements (p7, NS3 and NS5A) in a manner required for assembly of virus, suggesting a role in coordinating viral proteins in virus assembly (Jirasko et al., 2010, Ma et al., 2011). It has also been shown that specifically the ion channel activity of p7 is critical for release (Bentham et al., 2013) and evidence suggests that this might be important in protecting the immature E2 glycoprotein from degradation during the process of either assembly or egress (Atoom et al., 2013).

The importance of apolipoproteins (Apo) and the VLDL pathway in the production of virus have been well studies. However, it is disputed as to which Apo proteins are required for virus production, and which are incorporated into the infectious particle, reviewed in (Bartenschlager et al., 2011). The ApoE protein is known to be essential for the production of infectious virus but was only recently shown biochemically as a component of the infectious particle when produced from Huh7.5 cells (Chang et al., 2007, Merz et al., 2011). The siRNA knockdown ApoE dramatically reduced virus titres but the same was not observed for ApoB knockdown (Benga et al., 2010, Jiang and Luo, 2009). Interestingly, both Merz *et al.* and Icard *et al.* did not identify ApoB as a component of Huh7 cell derived infectious particles, but ApoB containing infectious particles have been identified in patients (Andre et al., 2002, Icard et al., 2009).

It was noted by Bartenschlager *et al.* that while there is no clear mechanism behind these discrepancies in ApoE and ApoB, this difference could in part be a result of the different cell lines utilised (Bartenschlager et al., 2011). As mentioned VLDL is the key pathway in virus production, and the ApoB100 protein is an essential component in this production of VLDL. While the two common hepatocellular carcinoma cell lines utilised, Huh7 and HepG2, are both able to secrete ApoB100 lipoparticles, it is relatively dense and lipid poor when compared to VLDL secretion in vivo (Meex et al., 2011). Suggesting that in the context of Huh7/HepG2 cells, HCV production might not utilise ApoB to the same extent as *in vivo.*

Consistent with observations that antibodies targeting ApoE, but not ApoB, neutralised viral particles produced from Huh7.5 cells (Owen et al., 2009).

This could correlate with differences observed by EM between HCV particles derived from cell culture systems and from patients. Infectious particles isolated from HCV-positive patients (treatment naïve) were shown to be VLDL like, being rich in triglycerides, containing ApoB and were large spherical particles of > 100 nm. Following delipidation smaller, core-positive and capsid like particles were observed (Andre et al., 2002). Characterisation of virus produced from cell culture observed enveloped particles that were significantly smaller, ~60 nm, but with a small sub-population of particles >100 nm smaller (Gastaminza et al., 2010). A comparison between infectious particles generated in a cell culture system verses those isolated from patient blood showed key differences. Patient derived particles were less dense (~1.05 g/ml), had a higher ApoB content and had a higher specific infectivity (Bartenschlager et al., 2011).

Other host proteins such as diacylglycerol acyltransferase-1 (DAGT1), microsomal triglyceride transfer protein (MTP) and Annexin A2 have all been shown to be essential for virus assembly (Backes et al., 2010, Herker et al., 2010, Jiang and Luo, 2009). DAGT1 is involved in triglyceride synthesis, lipid droplet formation and is thought to be involved in the recruitment of core to LD while MTP is important for the formation of lipoprotein particles. How Annexin A2 functions in assembly is not well understood, but the interaction it makes with NS5A is discussed later (Section 1.6.2). Production of infectious virus was also recently shown to be dependent on core trafficking via the adapter protein-2-M (AP2M1) (Neveu et al., 2012).

1.4 Systems for the study of HCV

1.4.1 Cell culture systems

At this point it is worth briefly discussing the different types of systems available for investigating HCV. The development of the first sub-genomic replicon (SGR) systems in 1999 dramatically increased our ability to study the HCV lifecycle (Lohmann et al., 1999). This was a bicistronic system that composed the 5' UTR IRES of HCV driving expression of a reporter gene, followed by the EMCV IRES driving the expression of the non-structural genes either NS2-5B or NS3-5B, Figure 1.8. Initially a neomycin phosphotransferase (npt) reporter gene was utilised as it allowed for positive selection of replication. The minimal polyprotein requirement for the replication of HCV RNA is NS3-5B and from an early stage it was

established that the incorporation of the NS2 protein into subgenomic replicons elicited a negative effect on replication and so was frequently omitted (Bartenschlager and Sparacio, 2007, Lohmann et al., 1999). The majority of genotypes required culture adaptive mutations in order replicate to detectable levels.

This was the case until 2003 when Kato *et al.* identified a genotype 2a isolate, termed JFH-1 whose NS3-5B replicon was capable of high level of replication in the absence of culture adaptations (Kato et al., 2003). The insertion of a firefly luciferase (luc) gene in place of the npt gene allowed for easy measurement of replication kinetics (Lohmann et al., 2001, Targett-Adams and McLauchlan, 2005) which itself facilitated for the first time the high throughput screening of anti-HCV compounds.

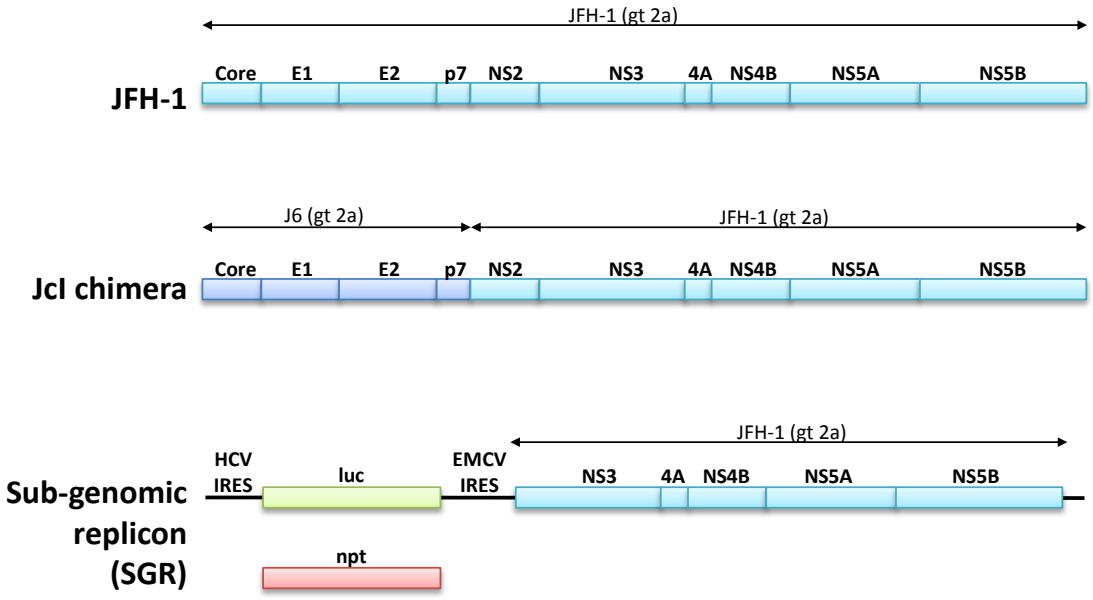


Figure 1.8. Structure of virus and replicon constructs.
The genotype 2a JFH-1 isolate (Wakita et al., 2005), the chimeric JcI virus (Pietschmann et al., 2006) and the two common subgenomic replicon (SGR) system (Kato et al., 2003, Targett-Adams and McLauchlan, 2005) with reporter genes firefly luciferase (luc) or neomycin phosphotransferase (npt)

The next leap forward in the ability to investigate the HCV lifecycle came in 2005 when Wakita *et al.* and Lindenbach *et al.* showed that the JFH-1 isolate was capable of undergoing the entire virus life cycle in the hepatocellular carcinoma cell line Huh7 (Lindenbach et al., 2005, Wakita et al., 2005). A chimeric virus, JcI composed of structural genes from the genotype 2a J6CF isolate (5' UTR to p7) and non-structural genes from the genotype 2a JFH-1 isolate (NS2 to 3' UTR) gave a 100 - 1000 fold increase in virus release over the WT JFH-1 isolate, Figure 1.8 (Pietschmann et al., 2006).

Based on the same approach as for JcI there are now chimeric viruses with respectable levels of virus release for all major genotypes, and these are all based around JFH-1 non-structural elements (Gottwein et al., 2009).. For a complete review on the current state of molecular clones, culture adaptive derivatives and their amenity to either sub-genomic (i.e. RNA replication) or full length virus (whole virus life cycle) see review (Bartenschlager and Sparacio, 2007, Vieyres and Pietschmann, 2013).

Along with advances in virus chimera, two groups recently demonstrated the ability of HCV to efficiently propagate in primary hepatocytes (Podevin et al., 2010, Ploss et al., 2010). Such systems have the advantage of being more physiologically relevant than current tissue culture systems, such as in the formation of polarised monolayers. Indeed, it was recently reported that the use of culture media where FBS was substituted for human serum, was sufficient to induce a hepatocyte-like morphology of Huh7 cells and a 1000-fold increase in the virus production of the JFH-1 isolate (Steenbergen et al., 2012). A combination of these systems will help give insight into some of the differences observed between current tissue culture and in patients, such as the coordination of receptors in polarised hepatocytes, and the use of different apolipoproteins in virus release.

1.4.2 Animal models

Chimpanzees are the only other natural host capable of supporting HCV infections, but the clinical course in chimpanzees is generally milder and with less chance of establishing a chronic infection (Bassett et al., 1998, Boonstra et al., 2009). When chimpanzees were inoculated with the JFH-1 isolate it caused only limited viraemia and was cleared by day 3 (Kato et al., 2008, Wakita et al., 2005). Similar observations of culture adaptive mutations showed that while enhancing replication in cell culture systems, they dramatically reduce the infectivity in chimpanzees (Bukh et al., 2002). This highlights an as yet unknown difference in the requirements of HCV in cell culture verse *in vivo*.

Much work has tried to establish mice as a host for the HCV lifecycle as they are not naturally susceptible to HCV infection. While mouse hepatocytes have been shown to support very limited replication, they are naturally completely refractory to entry, reviewed in (Boonstra et al., 2009). That was until very recently when the complete virus lifecycle was efficiently recapitulated over several weeks in humanised mice through the expression of CD81, occluding and in combination with an inhibited innate immune response (Dorner et al., 2013). A pre-clinical model such as this, especially if it is applicable to different genotypes, will be invaluable in developing the second and third generation DAAs.

1.5 Phosphorylation

A large part of this investigation focuses on the phosphorylation of the viral protein NS5A. As such the following section will briefly introduce the key concepts in phosphorylation and draw parallels with other known examples of viral phosphorylation.

Phosphorylation is the post-translational modification (PTM) of a protein by the enzyme-catalysed (kinases) addition of an inorganic phosphate molecule (PO_4^{-3}). Typically this occurs on the side chain hydroxyl group of serines, threonine and tyrosine, but phosphorylation of the basic amines of arginine, lysine, histidine as well as the carboxylic acid groups of aspartate and glutamate are known (Klumpp and Krieglstein, 2002, Besant et al., 2009, Attwood et al., 2011). Indeed due to the technical challenges in investigating the phosphorylation of basic amino acids their reported occurrence in the literature is most likely an under representation. It is estimated that 1.7 % of the human genome can be ascribed to the kinases family.

Typically a kinase will recognise a short consensus protein motif flanking the residue to be phosphorylated. The phosphorylation of a residue has both steric and charge effects on the neighbouring protein environment; as such they are predominantly involved with regulating a the function of a protein through conformation changes. Phosphorylations can also be readily removed by the phosphatase family of enzymes, making them a reversible PTM. It is estimated that up to 30 % of the human proteome is phosphorylated by kinases, highlighting how extensively utilised this regulatory mechanism is (Manning et al., 2002).

1.5.1 Methods for investigating phosphorylation

The gold standard in the identification of phosphorylation sites is mass spectrometry, reviewed in (Witze et al., 2007). Typically proteins are resolved by SDS-PAGE, enzymatically cleaved into smaller peptides and then analysed by a variety of mass spectrometry (MS) techniques. This technique can definitively identify which amino acids in a given peptide have the addition of a phosphate group. However, the amount of protein required for such an approach often precludes the analysis of low abundance proteins, or where the phosphorylated form is only a minor species. Furthermore, the requirement for a high abundance of naturally occurring enzyme cleavage sites to produce small enough peptide fragments for MS analysis is also a constraint.

Further limitations come from the suppressive effect that phosphorylation can have on the detection of peptide fragments by MS. The detection of molecules by MS requires they have a net charge and typically in the analysis of peptides this is a positive charge (due to the high density of protonatable amine/amide groups). The addition of negatively charged phosphate groups can therefore suppress the detection of highly phosphorylated peptide fragments. As a result of these effects, not observing a phosphorylation event by MS is insufficient evidence to conclude that the residue is not phosphorylated.

When the use of MS is not applicable, such as for low abundance proteins, then genetic disruption of phosphorylation sites is a common approach. Typically putative phosphorylation sites are predicted based upon homology to kinase recognition sequences or when there is conservation of serine rich regions. Mutagenesis is then used to disrupt phosphorylation sites and either phenotypic events or the ability of the protein to be labelled with radioactive inorganic phosphate, are subsequently observed. Such mutations are termed phosphoablatant or phosphomimetic and often involve substitution to alanine or aspartic/glutamic acid respectively. The lack of hydroxyl and amine groups means that alanine cannot be phosphorylated, while aspartic/glutamic acid mimics the steric and negative charge of a phosphorylated residue. It is worth bearing in mind that kinases inhibitors are used extensively as ways of probing the function of kinases, however due to the high sequence and structural homology within the kinome, these inhibitors are very rarely 100 % specific towards only a single kinase.

In this investigation a combination of these approaches are utilised; the MS analysis of purified protein to identify phosphorylation sites and the mutagenesis of these sites to investigate the functional role.

1.5.2 Phosphorylation of viral proteins

A recent review of medically relevant virus showed that phosphorylation was ubiquitous across all classes of viruses (RNA, DNA and retroviruses) and that many DNA viruses code for their own viral kinases (Keating and Striker, 2012). The *Flaviviridae* family is not exempt from this, non-structural proteins from Yellow fever virus (YFV), Dengue virus (DENV) and bovine virus diarrhoea virus (BVDV) have all been shown to be phosphorylated by host kinases (Jakubiec and Jupin, 2007).

1.5.3 Examples of phosphorylation events

A phosphorylation mechanism, termed sequential phosphorylation cascade, is of relevance to later findings and therefore introduced here. In such a mechanism a serine rich region within a protein has the sequential addition of several phosphate groups typically involving one or more kinases. Key to this process is the requirement by the kinase that a neighbouring residue is phosphorylated, typically in the -4, -3 or -1 position. This proximal phosphorylation completes the kinase recognition motif and thereby allows efficient kinase activity. Several proteins such as glycogen synthase kinase-3 (GSK3) and casein kinase I and II (CKI and CKII) have such recognition motifs and are involved in the sequential phosphorylation cascades of cellular proteins like PTEN, GRASP65 and β-catenin (Tang et al., 2012, Cordier et al., 2012, Liu et al., 2002).

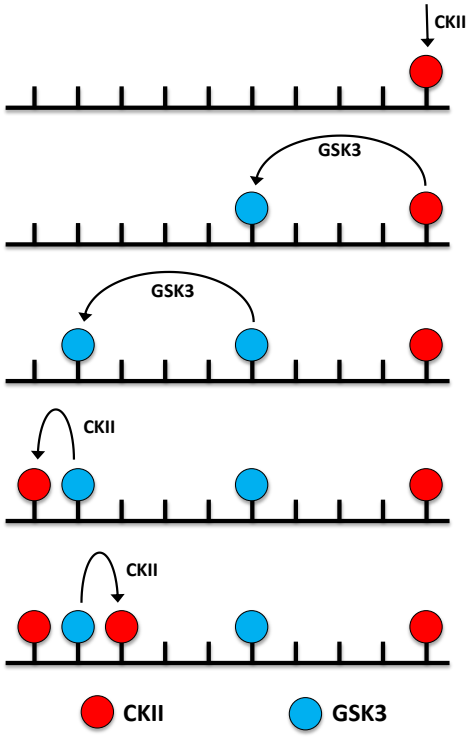


Figure 1.9. Sequential phosphorylation cascade of PTEN protein.
Illustration of residues 361 – 370 of the phosphatase and tensin homolog (PTEN). After an initial priming phosphorylation by CKI the peptide is sequentially phosphorylated by GSK3 and CKII to produce an extensively phosphorylated region. Figure adapted from (Cordier et al., 2012).

An example of a sequential phosphorylation cascade in the PTEN protein is illustrated in Figure 1.9. Here the initial phosphorylation by CKII completes the recognition motif for GSK3 thereby priming phosphorylation in the -4 position, this phosphorylation then primes the next phosphorylation by GSK3 in the -4 position and so on (Cordier et al., 2012). It required sophisticated temporal NMR in order to decipher this complex mechanism of phosphorylation and also required a preliminary understanding of the kinases involved and the sites of phosphorylation.

Another phosphorylation mechanism that later has parallels to NS5A phosphorylation is that of the polo like kinase-I (PlkI) protein, a key regulator of cell cycle progress. This protein has also been shown to be important for virus replication and is discussed in that respect later. PlkI is composed of a polo-box domain (PBD) at the C-terminus and a kinase domain at the N-terminus. The PBD has been shown to bind the phosphorylated recognition motif SpS(P/T) and subsequently direct the activity of the kinase domain either towards the same protein or a third protein that it is associated with (Park et al., 2010). The result of having the PBD and kinase domain in different regions of the protein is that the priming phosphorylation recognised by the PBD can be in a spatially distant region to the residue that PlkI phosphorylates.

This mechanism of distant phosphorylation has been demonstrated in the P protein of the Parainfluenza virus-5 (PIV5), the prototype member of the *Paxamxyovirus* genus. Here a phosphorylation at serine 157 of the P protein facilitates the binding of PlkI via the PBD and subsequently directs the PlkI phosphorylation of serine 308 (Sun et al., 2009). A comparison between this long range priming phosphorylation of the P protein and recent discoveries about the NS5A protein is discussed later (Section 3.3).

1.6 The Non-structural 5A protein

1.6.1 Structure and function of NS5A

The Non-Structural 5A protein (NS5A) is a large (56-58 kDa), multi-functioning phosphoprotein that has an essential role throughout the virus life cycle. It is composed of three major domains (I, II and III) linked by low complexity sequences, although in recent years domains II and III have been increasingly defined as a single, unstructured domain, discussed later (Figure 1.10A). The protein is anchored to phospholipid membranes by an N-terminal amphipathic helix (residues 1-33) in a manner essential for replication (Brass et al., 2002, Penin et al., 2004).

The crystallographic structure of domain I was solved using x-ray diffraction and gave numerous insights into potential mechanisms of action (Tellinghuisen et al., 2005). This structure (residues 36-198) had an asymmetric unit composed of two identical domain I monomers in a dimeric conformation that formed a large, positively charged groove, proposed as a putative RNA binding groove. When this structure was computationally modelled with the N-terminal amphipathic helix on a phospholipid membrane, it was predicted that NS5A would be anchored to the membrane in a way that the putative RNA binding groove was projected outwards in a manner that could facilitate the binding of RNA, illustrated in (Figure 1.10B).

The Tellinghuisen *et al.* crystal structure of NS5A showed that domain I could be divided into two sub-domains, IA and IB, the former confirmed the presence of a coordinated zinc ion between four highly conserved cysteines (C39 C57, C59 and C80); disruption of which was previously shown to abrogate replication (Tellinghuisen et al., 2004). Of relevance later, a disulphide bond was observed between Cys142 and Cys190 in each monomer subunit (Tellinghuisen et al., 2005). However, the biochemical disruption of it by a C190G mutation

was shown to only cause a minor impairment to replication, despite computational analysis that suggested the absence of such a bond was structurally unfavourable (Tellinghuisen et al., 2004).

A second domain I crystal structure was solved in 2009 and showed a near identical monomer structure, but in a dramatically different dimer conformation (Love et al., 2009). In this new dimer the monomers were stacked tightly together with residues from both sub-domains IA and IB involved in the dimer interface, unlike the Tellinghuisen *et al.* structure where dimer interface residues were predominantly from sub-domain IA (Tellinghuisen et al., 2005). Interestingly the disulphide bond present in the Tellinghuisen *et al.* structure, between Cys142 and Cys190, was not present in this structure although the thiol groups remained in close proximity. Love *et al.* modelled multiple dimer units together to theorise the presence of a superhelical structure composed of repeating Tellinghuisen *et al.* NS5A dimers that would no longer be membrane associated. However there has yet to be *in vivo* evidence to suggest either the presence, or possible functions, or such a species.

Domains II and III, unlike domain I, have been shown to be inherently unstructured (Hanoulle et al., 2009b, Liang et al., 2007), forming long flexible domains as illustrated in Figure 1.10B. Undoubtedly the unstructured nature of these domains is central to the ability of NS5A to interact with such a diverse range of proteins, with 130 host protein interactions being identified to date, including all HCV non-structural proteins (de Chassey et al., 2008, Macdonald and Harris, 2004, Dimitrova et al., 2003, Tripathi et al., 2013). It could be readily imagined that with the amphipathic helix in domain I anchoring NS5A to lipid membranes, the flexible nature of domains II and III leaves them free to interact with other proteins in a 'fly-casting' type mechanism (Shoemaker et al., 2000). Several examples of such a mechanism are discussed in (Gsponer and Babu, 2009). This is supported by data showing that several interaction partners of NS5A, notably CypA, NS5B, VAP-B and SH3 domains, have all been shown to have short linear binding motifs that would be compatible with this type of model.

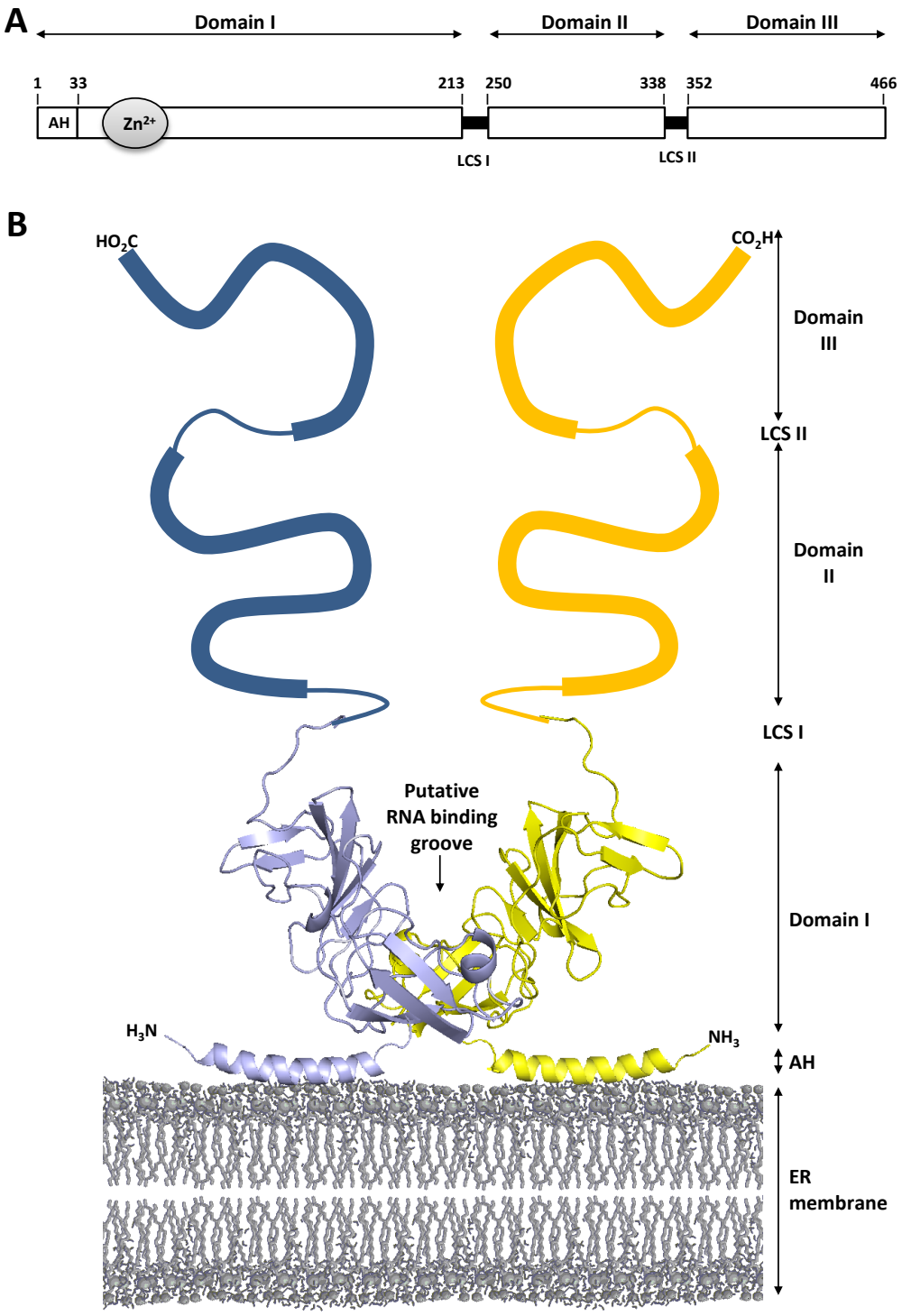


Figure 1.10. Structure of the NS5A protein.

A. NS5A is composed of three domains (I, II and III) linked by two low complexity sequences (LCS I and LCS II). The N-terminus of NS5A contains a 33 aa amphipathic helix (AH) and domain I coordinates a zinc ion in a highly ordered manner. **B.** Illustration of the two NS5A domain I monomers (blue and yellow) in the dimer conformation observed by (Tellinghuisen et al., 2005), anchored to the membrane (grey) by the AH (Penin et al., 2004). The unstructured nature of domains II and III are illustrated.

Furthermore, it was recently discovered that domain II forms several transient α-helices that were unobservable with previous NMR techniques. These transient helices had low-affinity SH3-binding properties and were shown to interact with the host protein BinI (Feuerstein et al., 2012). It is therefore entirely plausible that this combination of transient secondary structure and linear binding epitopes is what facilitates the diverse range of protein interactions formed by NS5A (Liang et al., 2007).

The low complexity sequence I (LCS I) is a 36 aa, serine rich region of NS5A linking domains I and domain II, Figure 1.10. This region has been shown to be essential for NS5A phosphorylation (Appel et al., 2005, Lemay et al., 2013, Tanji et al., 1995) but through what mechanism remains to be eluded. Based on the Tellinghuisen *et al.* crystal structure it is conceivable that an extensively phosphorylated LCS I could compete with RNA for binding within the NS5A dimer.

The low complexity sequence II (LCS II) is a short, 13 aa proline rich region of NS5A linking domains II and III. Within this region are two class II polyproline motifs, PP2.1 and PP2.2, which have been characterised extensively for the ability to bind cellular SH3 domains (Macdonald et al., 2004, Tan et al., 1999, Zech et al., 2003). The function of these motifs in the virus life cycle has been investigated and showed that PP2.1 but not PP2.2, is a requirement for replication in a genotype 1b replicon. The common genotype 2a isolate JFH-1 does not contain the PP2.1 motif, but in comparison to the 1b replicon the PP2.2 motif was also not required for replication (Hughes et al., 2009a). While the PP2.2 motif may not be required for virus replication, roles in ion channel modulation (Mankouri et al., 2009) and inhibition of apoptosis (Amako et al., 2013) suggests a role in virus persistence. This is supported by a lack of infectivity in chimpanzees of a genotype 1b virus harbouring a mutation of the first PP2.1 motif (Nanda et al., 2006).

The involvement of the NS5A protein in perturbing host pathways has been extensively documented and will be only briefly introduced here as it is not central to this study. The NS5A protein has been shown to inhibit apoptosis, both intrinsic and extrinsic, through several different pathways including the TNF-α pathway by an interaction with the TNF-α-responsive adaptor protein (TRADD) (Ghosh et al., 2000); through direct sequestering of p53 in the cytoplasm (Lan et al., 2002); and inhibition of the pro-apoptotic potassium ion channel Kv2.1 (Mankouri et al., 2009, Amako et al., 2013). As well as being anti-apoptotic, NS5A also

modulates the three major mitogenic signalling pathways in favour or virus persistence. Both the extracellular signal-regulated kinase (ERK) pathway and p38-Mitogen-Activated Protein Kinase (p38MAPK) pathway have been shown to be inhibited by NS5A, reviewed in (Macdonald and Harris, 2004); while the third pathway, the stress-activated protein kinase/c-Jun N-terminal kinase (SAPK/JNK) pathway, is activated by NS5A (Park et al., 2003). The net effect on these MAPK pathways is a perturbation of host cell cycle control; this would correlate with observation that there was an increase proportion of hepatocytes in the G1 phase in HCV infected patients when compared to patients with alcohol induced liver disease (Freeman et al., 2003).

Many of these interactions that modify host signalling pathways are dispensable for the replication of the virus in an acute cell culture setting. It is thought that many of these pathways are involved in the chronic persistence of infection that is observed *in vivo*, supported by their conservation across evolutionary diverse genotypes (Macdonald and Harris, 2004). While the NS5A protein has key functions throughout the virus life cycle, the following sections will focus solely on those interactions that have been shown to be important in either the replication or production of virus.

1.6.2 The role of NS5A in replication

NS5A has long been established as a critical viral protein for replication and as such has been extensively studied. As a non-enzymatic protein it achieves these functions through a plethora of protein-protein interactions with virus and host proteins. Through deletion analysis of NS5A it has been shown that domain I, LCS I and resides 244-302 at the C-terminal end of domain II, form the minimum requirement for virus replication (Appel et al., 2008). It is worth noting that the majority of studies investigating which elements of NS5A are required for replication focus on short term, < 96 hrs, replication assays.

An interaction between NS5A and NS5B has been well established with recent NMR mapping some element of this interaction to the C-terminus of NS5A domain II (Coelmont et al., 2010, Shimakami et al., 2004, Shirota et al., 2002). Interestingly it was recently shown that the phosphorylation state of NS5A directs whether it has an inhibitory or enhancing effect on NS5B activity and could point towards a potential role in regulation (Ivanov et al., 2009). NS5A has been shown to interact with NS3 in bacterial expression and yeast two hybrid systems (Dimitrova et al., 2003) and it was shown that the catalytic activity of NS3 is required

for the hyperphosphorylation of NS5A (Neddermann et al., 1999). However, whether there is a functional interaction between NS5A and NS3 has not been shown.

The ability of HCV to bind RNA was first predicated from the Tellinghuisen *et al.* crystal structure in 2005. Since then it has been shown biochemically that NS5A binds uridylate and guanylate-rich RNA, as well as the 3'UTR (positive- and negative-sense) (Huang et al., 2005). More recently it was shown that all three domains of NS5A can bind RNA individually, albeit with a lower affinity than full length NS5A (Foster et al., 2010, Huang et al., 2005). Interestingly Huang *et al.* also demonstrated that extensive phosphorylation of NS5A was not a requirement for the *in vitro* RNA binding that they reported, hinting at a potential regulatory mechanism if it is later shown that extensive phosphorylation of NS5A is inhibitory to RNA binding.

The functional importance of NS5A binding RNA is not known and hard to establish. Specific mutations or deletions that block NS5A from binding RNA have not yet been established and therefore it can only be speculated as to what function the RNA binding properties have. It has been proposed that NS5A binding RNA is essential in replication; this being supported by the tight and critical associations it forms with the replication machinery NS5B and NS3. Similarly it is thought that RNA binding is key in virus assembly as deletions discussed later that block assembly also block the NS5A:core interaction and reduce the interaction of core with viral RNA (Masaki et al., 2008).

1.6.2.1 Host Factors

Numerous host proteins have been shown to interact with NS5A in a manner critical for the replication of virus, only those relevant to this thesis are discussed below. Other interactions are reviewed in (de Chassey et al., 2008, Macdonald and Harris, 2004, Tan, 2006). The Cyclophilins (Cyp) are a family of peptidyl-prolyl isomerases (PPIase) enzymes that catalyse the *cis/trans* isomerisation of the peptide bond preceding a proline residue, typically resulting in major conformational shifts. Much work has investigated the importance of the Cyp family in the HCV lifecycle. They have been shown as critical host factors for several viruses including severe acute respiratory syndrome (SARS)-Coronovirus and Human Immunodeficiency Virus (HIV) (Watashi and Shimotohno, 2007) and most recently HCV. Two isoforms have been shown to be important in the HCV life cycle, CypA and CypB, through interactions with viral factors NS5A and NS5B respectively. Disruption of these interactions is

deleterious to replication (Chatterji et al., 2009, Coelmont et al., 2010, Heck et al., 2009, Yang et al., 2010).

The majority of recent work has focused on understanding the molecular details of the NS5A-CypA interaction, as several CypA inhibitors are progressing through clinical trials (Paeshuyse et al., 2006). Structural evidence now shows that the active site of CypA interacts with the C-terminus of NS5A domain II, and likely catalyses an isomerisation of a prolyl bond within this region (Coelmont et al., 2010, Grise et al., 2012, Hanoulle et al., 2009a). Interestingly this binding site was recently shown to overlap with the binding site of NS5B (Rosnoblet et al., 2012). Further details of these interactions and requirement are discussed later (Section 4.1). Moreover it has been shown that CypA stimulates the RNA binding of NS5A domain II in an isomerase dependent manner; while NS5A domain II harbouring two CypA independent mutations (D316E and Y317N) did not exhibit this simulation (Foster et al., 2011).

The vesicle associated membrane protein (VAMP)-associated proteins (VAP) are type II membrane proteins localised to the ER and involved in numerous host processes including membrane trafficking and lipid transport (Lev et al., 2008). The importance of VAP proteins and HCV was initially demonstrated by Tu *et al.* and later interactions of viral proteins with VAP-A (hVAP-33) and VAP-B were identified (Gao et al., 2004, Hamamoto et al., 2005, Tu et al., 1999). VAP-A was shown to interact with both NS5A and NS5B, and knockdown of VAP-A resulted in a relocalisation of NS5B and a small reduction in RNA replication (Gao et al., 2004). VAP-B has recently been shown by NMR to bind domain III of NS5A in a region previously proposed by Masaki *et al.* as important for basal-phosphorylation of NS5A (Gupta et al., 2012, Masaki et al., 2008).

The phosphorylation state of NS5A and subsequent interaction with VAP-A is considered controversial. It was putatively shown that the phosphorylation of NS5A regulated binding to VAP-A in an inverse manner, with hyperphosphorylation of NS5A apparently being inhibitory to this interaction (Evans et al., 2004). This model has not been subsequently validated, but nevertheless is used as one of the main arguments for a more general model where hyperphosphorylation of NS5A negatively regulates RNA replication. However Evans *et al.* did suggest the region of NS5A that interacts with VAP-A is in the region of the LCS I.

That the VAP proteins play a key role in replication is supported by the antiviral activity of Viperin, which was shown to act by interacting with VAP-A and displacing NS5A (Helbig et al., 2011, Wang et al., 2012). Paul *et al.* recently isolated both VAP-A and VAP-B as significantly enriched components of HCV RNA positive DMVs isolated from infected cells (Paul et al., 2013). Interestingly VAP-A also interacts with two other host factors important for the HCV lifecycle, OSBP (Wyles et al., 2002, Wyles and Ridgway, 2004) and occludin (Lapierre et al., 1999). Whether the interaction of these proteins with VAP-A is relevant to HCV infection is yet to be suggested.

The importance of PI4KIIIα/ β in the synthesis of the PI4P lipids, the formation of the membranous web and the virus life cycle have been discussed previously (Section 1.3.2). An interaction between NS5A and PI4KIIIα was shown to be essential for these effects and the binding site on NS5A was mapped to the C-terminus of domain I, residues 187-214 (JFH-1 numbering). This was refined further to show that residues 203-209 formed the functional interaction site of PI4KIIIα, termed PFIS (Reiss et al., 2013). Mutation of the PFIS was sufficient to block the NS5A-PI4KIIIα interaction and cause the reduction in virus replication, PI4P synthesis and membranous web formation. The interaction of NS5A with PI4KIIIα has also been observed (Lim and Hwang, 2011). Interesting observation of effects on the phosphorylation state of NS5A and the interaction with PI4KIIIα have been made and are discussed shortly.

The role of the lipid droplet organelle in the production of infectious virus is well established, discussed previously (Section 1.3.3) and reviewed (Miyanari et al., 2007). It has been well established that the NS5A protein interacts and localises with LDs, but whether this interaction plays a critical role in virus replication has not been well defined. The lipid droplet associated tail-interacting protein 47 (TIP47) was shown to interact with the N-terminal amphipathic helix of NS5A and shRNA knockdown subsequently showed a loss of virus replication (Vogt et al., 2013). However triple alanine mutations in domain I (99-101 and 102-105) dramatically reduced the localisation of NS5A and other non-structural with lipid droplets, but with no reported effect on virus replication (Miyanari et al., 2007). Implying that while the LD-associated TIP47 is essential for virus replication, the recruitment of NS5A to LD is not.

Another interesting host factor that was recently shown to interact with NS5A is the lipid droplet associated Rab18 (Salloum et al., 2013). Rab proteins are GTPases of the Ras

superfamily involved in membrane trafficking, vesicle formation and membrane fusion. Salloum *et al.* showed that the NS5A-Rab18 interaction was critical for replication, observing a significant drop in replication upon Rab18 silencing. Interestingly upon ectopic expression of Rab18 they also observed an increase in virus release but not replication. As Rab proteins are involved in vesicle trafficking it is plausible that ectopic expression of Rab18 might facilitate the increased export of maturing virus particles or increase the recruitment of resources to the membranous web.

Protein name		Role in virus life cycle	Region of NS5A	Mechanism	Reference
VAMP-associated protein A	VAP -A	Replication	unknown	Possible scaffolding replication complex.	Tu e al, 1999. Gao et al, 2004 and Evans et al, 2004
Cyclophilin A	CypA	Replication	DII	Unknown	Hanoulle *et al*,2009
			D317, Y317 (DII)	Unknown	Yang *et al*, 2010
			306-333 (DII)	Isomerisation of DII	Rosnoblet *et al*, 2012
			310-330 (DII)	Cis/trans isomerisation	Coelmont *et al*, 2010
Cyclophilin B	CypB	Replication	DII	poss. NS5B interaction	Hanoulle *et al*, 2009
Phosphotidylinositol-4 kinase III alpha	PI4KIII α	Replication	DI	Membranous web formation	Reiss *et al*, 2011
			187-214 (DI)	Stimulation of kinase activity	Reiss *et al*, 2013
Tail-interacting protein-47	TIP47	Replication	amphipathic helix	Unknown, LD localisation?	Yogt *et al*, 2013
Protein kinase R	PKR	Persistence	237-298 (LCSI/DII)	Inhibition of PKR	Gale *et al*, 1998

Table 1.1. Host interacting partners of NS5A that are essential to virus replication.

The NS5A protein also makes a key interaction with the interferon-inducible protein kinase R (PKR). This kinase is an element of the innate immune response which in response to dsRNA halts cellular translation (Balachandran et al., 2000). The NS5A protein has been shown to bind and inhibit the kinase activity of PKR, but in a genotype dependent manner. NS5A from

genotype 1 but not genotype 2a/3a is able to inhibit PKR, partly explaining the improved outcome for genotype 2a/3a patients on interferon treatment. The interaction of NS5A with PKR is the mechanism behind how the previously identified interferon sensitivity determining region (ISDR), residues 237-272 (JFH-1 numbering), effectively predicted the outcome of interferon treatment (Enomoto et al., 1995, Gale et al., 1997, Pascu et al., 2004). However deletion of the PKR binding site, residues 237-298 (JFH-1 numbering), did not block HCV replication, suggesting that the interaction of NS5A with PKR is not a requirement for replication in a cell culture system, at least over a 72 hour infection (Appel et al., 2008). While the PKR binding site in NS5A is adjacent to regions of NS5A that are highly phosphorylated, there has been no evidence that NS5A is a substrate of the PKR kinase; indeed bacterially expressed NS5A has been shown to not be phosphorylated by PKR (Chen et al., 2010, Gale et al., 1997). Further a serine to alanine mutation the phosphorylation site 222 (JFH numbering) did not disrupt the binding of PKR (Katze et al., 2000). Several kinases have also been identified as essential in the virus lifecycle, casein kinase I (CKI) and polo-like kinase I (PlkI), discussed in further detail later (Section 1.6.4.3).

1.6.3 The role of NS5A in assembly

Domain III is the primary region of NS5A that is critical for the assembly of infectious virus. Whether domain I and regions of domain II are also required for virus assembly is challenging to investigate as they both contain elements essential for virus replication. However, In chapter 4 results are presented that argue against a role of the domain II in virus assembly.

The entirety of domain III can be deleted with no observable effect on virus replication, but a near complete block on virus assembly (Appel et al., 2008). Further deletions identified residues 427 - 457 (JFH-1 numbering) as the major epitope responsible for this phenotype, with concurrent deletion of resides 376 - 426 increasing the impairment to virus release. This correlates with Tellinghuisen *et al.* who identified serine 457 (JFH-1 numbering) as the major requirement, with an alanine mutation at this point blocking virus release. This however is contrary to observations by Masaki *et al.* who showed that mutation of serine 457 alone was insufficient to block assembly, but required additional mutations at serine 452 or 454 to elicit the same phenotype (Masaki et al., 2008, Tellinghuisen et al., 2008a). Furthermore, Masaki *et al.* demonstrated that mutation of serine 452, 454 and 457 also disrupted the interaction with core and reduced the amount of RNA that co-immunoprecipitates with core. These mutations also shifted the localisation of NS5A from fractions enriched in LD proteins

towards fractions enriched in ER markers (Masaki et al., 2008). In line with Appel *et al.* these mutations had no effect on virus replication.

Protein name		Role in virus life cycle	Region of NS5A	Mechanism	Reference
Apolipoprotein E	ApoE	Assembly	205-280 and/or DIII	Unknown	Benga *et al*, 2010. Cun et al, 2003
Annexin A2	ANXA2	Assembly	Domain III	Unknown	Backes *et al*, 2010
Phosphotidylinositol-4 kinase III beta	PI4KIIIβ	Secretion	Unknown	Unknown	Coller *et al*, 2012
Rab18	Rab18	Possibly assembly	Unknown	Unknown	Salloum *et al*, 2013
Oxysterol binding protein	OSBP	Assembly	Unknown	Unknown	Amako *et al*, 2009

Table 1.2. Host interacting partners of NS5A that are essential to virus production.

More recently this was reiterated in the genotype 1a virus H77S, where it was shown that mutation of the serine cluster 452-457 (JFH-1 numbering) again had no effect on replication but a complete block of virus release. Interestingly an extra serine at position 458 (glutamic acid in JFH-1) and threonine at 462 (both JFH-1 numbering) also elicited a reduction in virus assembly that were not previously investigated (Kim et al., 2011).

There was a concurrent change in the phosphorylation state of NS5A when mutations were inserted into domain III. While Masaki *et al.* observed a reduction in basal-phosphorylation with mutation of these serine clusters (Figure 1.11), Tellinghuisen *et al.* observed a reduction in hyperphosphorylation and this discrepancy has yet to be resolved.

More supporting evidence of the importance of domain III in the assembly of infectious virion was shown by the generation of a culture adapted JFH-1 virus with enhanced virus production, giving a 3 log increase in released virus titres. When sequenced the adaptive mutations clustered in domain III of NS5A and when engineered back into WT JFH were shown to account for the majority of increase in virus release (Liu et al., 2012).

1.6.3.1 Host Factors

The importance of interaction between core and lipid droplets for the production of infectious virus has already been mentioned (Section 1.3.3). The recruitment of NS5A to LDs is determined by residues 99-105 in domain I, but for the co-localisation of NS5A with core on LDs there is the additional requirement for domain III of NS5A (Appel et al., 2008, Miyanari et al., 2007). The region of domain III required for this co-localisation is likely to correspond to the serine rich region 452-457 (JFH-1 numbering) identified by Masaki *et al.* to be essential for an interaction with core (Masaki et al., 2008). Recently Ploen *et al.* presented on the lipid droplet associated protein TIP47 as being a component of the lipoviralparticle, and observed that overexpression of TIP47 with a mutated Rab9 binding site was sufficient to block virus production (Ploen et al., 2012). Whether this is involved with the aforementioned interaction that NS5A makes with TIP47 is as yet unknown.

The NS5A protein has been shown to interact with apolipoproteins, discussed earlier for their importance in lipoparticles and virus assembly/maturation (Section 1.3.3). It was shown that siRNA knockdown of ApoE was refractory to virus production in Huh7.5 cells, and subsequently attributed this to an interaction with NS5A, an interaction that was reduced when domain III was deleted (Benga et al., 2010). Later Cun *et al.* confirmed an interaction with NS5A but mapped it to residues 205-280 (Cun et al., 2010). The ApoA1 protein, a component of the high-density lipoprotein (HDL), was also shown to interact with NS5A, and to co-localise with it on the Golgi apparatus, but no functional information was derived from the interaction (Shi et al., 2002).

The interaction of NS5A with annexin A2 was identified by proteomics and shown to be dispensable for replication, but critical for the release of infectious particles (Backes et al., 2010). Deletion studies identified domain III as the critical element required for the co-localisation of annexin A2 to with NS5A. Exactly whether this interaction is involved with assembly or release of infectious virus is still uncertain, and whether or not this interaction is modulated by NS5A phosphorylation would be interesting to investigate.

The chemical inhibition or siRNA knockdown of the host kinase CKII was also shown to abrogate virus production, while not affecting replication (Section 1.6.4.3) (Tellinghuisen et al., 2008a). Another host factor whose knock down results in a reduction in infectious virus production is the oxysterol binding protein (OSBP) (Amako et al., 2009). Deletion studies identified domain I of NS5A as the site of OSBP binding, but it is once again worth noting that

as domain I is critical for the localisation and trafficking of NS5A, that domains II and III don't bind OSBP does not exclude an *in vivo* interaction. Furthermore, it is worth noting that the OSBP protein has both a VAP-A and PI4P binding domain (Amako et al., 2009).

1.6.4 Phosphorylation of NS5A

1.6.4.1 Background of NS5A phosphorylation

NS5A is a phospho-protein that exists in two forms with slightly different mobility on SDS-PAGE; these have been termed a basally phosphorylated (apparent molecular weight of 56 kDa) and a hyperphosphorylated (58 kDa) form. These two forms are commonly referred to as p56 and p58; however, in JFH-1 there is a 19 aa insertion in domain III compared to genotype 1b, they migrate with apparent molecular weights of 63 and 65 kDa (Figure 3.1A). This increase in an apparent molecular weight of 2 kDa, in a manner sensitive to phosphatase treatment, is indicative of an extensively phosphorylated protein. This is supported by observations that the *in vitro* incorporation of 5 phosphates per molecule of NS5A was still insufficient to recapitulate this shift in molecular weight (Huang et al., 2004). Interestingly these two phosphorylation species are always observed as two discrete bands by SDS-PAGE, without the presence of intermediary species.

As has been introduced previously, domains II and III of NS5A are inherently unstructured; with only small, transient elements of secondary structure being recently identified (Gupta et al., 2012, Feuerstein et al., 2012). As reviewed in (Gsponer and Babu, 2009, Dyson and Wright, 2005) extensive post-translational modifications (PTM) is a common feature of intrinsically unstructured proteins (IUPs). Indeed phosphorylations appear to predominantly occur in unstructured regions of proteins (Iakoucheva et al., 2004). Through altering the local charge and hydrophobicity of a protein these modifications elicit effects on the proteins conformation both locally and globally (Wright and Dyson, 1999). The unstructured domains of NS5A are thought to act in a similar fashion, with extensive phosphorylation modulating protein-protein interaction sites through these local effects on protein conformation. Through this mechanism NS5A is able to interact with a diverse network of proteins and enabling it to function in spatially and temporally distant events like replication, virus production and modulating host cell pathways.

Much work has previously tried to conclusively map the location of phosphorylation sites by a combination of mutagenesis or biochemical analysis (MS, Edman degradation etc.). Previously, with the absence of highly efficient replicon and virus systems, much of this work was constrained to ectopic expression of NS5A, either as a single protein or as part of polyprotein. It is well established that phosphorylation of NS5A is dependent on, at a minimum, other non-structural proteins and most likely active replication complexes (Koch and Bartenschlager, 1999, Neddermann et al., 1999). As NS5A is a serine/threonine rich

protein it brings into question the functional relevance of phosphorylation sites identified by *in vitro* kinase assays, where the ratio of kinase to substrate is vastly higher than would occur *in vivo*. With this in mind it is worth considering that, with a recent exception (Lemay et al., 2013), the identification of NS5A phosphorylation sites by mass spectrometry, on protein purified from cells harbouring either replicons or virus, has not been carried out.

The serine rich LCS I was shown to be important for phosphorylation by Tanji *et al.* who showed that mutation of serines at residues 225, 229 and 232 in NS5A (ConI) resulted in a loss of hyperphosphorylation on NS5A when ectopically expressed from a NS2-5B polyprotein in COS-1 cells (Tanji et al., 1995). It was later shown by MS that NS5A expressed as a single protein in Sf9 cells was phosphorylated in the region of 222-232, and after additional *in vitro* kinase treatment identified phosphorylation at serine 222 (S2194 ConI) (Katze et al., 2000). Interestingly when Tanji *et al.* mutated serine 222 they saw no effect on phosphorylation despite all the serines within this LCS I being absolutely conserved across HCV genotypes.

More recently it was demonstrated by mass spectrometry that serine 222 was a major phosphorylation site in a JFH-1 subgenomic replicon, and mutation of this residue to aspartic acid (as a phosphomimetic) had a modest inhibitory effect on RNA replication. Contrary to Tanji *et al*, a mutation of serine 222 to alanine had a modest inhibitory effect on hyperphosphorylation (Lemay et al., 2013).

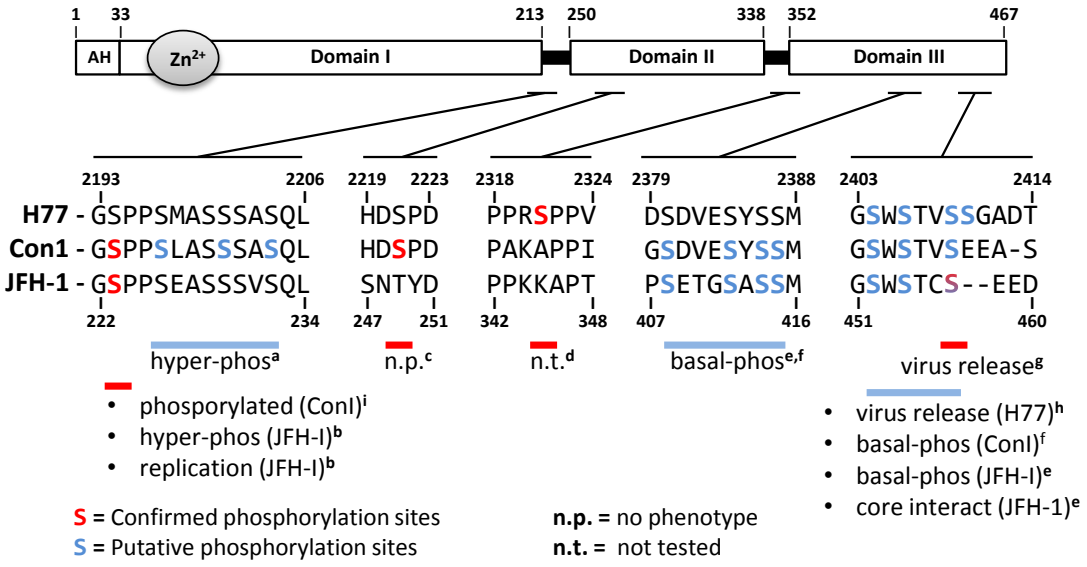


Figure 1.11. Summary of regions important for NS5A phosphorylation.
Numbering of highlighted region refers to either H77 polyprotein (2193-2414) or JFH-1 NS5A (217-460). [a] (Tanji et al., 1995); [b] (Lemay et al., 2013); [c] (Nordle Gilliver et al., 2010); [d] (Reed and Rice, 1999); [e] (Masaki et al., 2008); [f] (Appel et al., 2005); [g] (Tellinghuisen et al., 2008a); [h] (Kim et al., 2011); [i] (Katze et al., 2000)

Another residue, serine 232, was shown to be phosphorylated *in vitro* after casein kinase I (CKI) treatment of a synthetic mimetic of the LCS I. In line with the requirement for a priming phosphorylated serine in the -3 position for efficient CKI phosphorylation, a synthetic phosphorylation at serine 229 greatly enhanced phosphorylation of the LCS I peptide. Interestingly the phosphomimetic, glutamic acid, did not recapitulate this increase in efficiency suggesting here glutamic acid does not sufficiently mimic a phosphorylated serine (Quintavalle et al., 2007). Two major culture adaptive mutations in a ConI replicon, S2197P and S2204I (serines 225 and 232, respectively), were both associated with significant enhancements in replication and a concurrent reduction in hyperphosphorylation (Blight et al., 2000, Krieger et al., 2001). However, as mentioned earlier these adaptive mutations do not establish a robust infection in chimpanzees (Bukh et al., 2002)

To investigate the phenotype of other putative phosphorylation sites in the LCS I Appel *et al.* introduced phosphoablatant or phosphomimetic mutations across this region in a Con I subgenomic replicon. A broad enhancement in replication was observed when phosphoablatant mutations were introduced (225, 229, 232 and 235, JFH-1 numbering) and

one phosphomimetic (serine 229, JFH-1 numbering) (Appel et al., 2005). These are discussed in further detail in (Section 3.3) due to the relevance to results generated herein.

The mass spectrometry analysis of NS5A expressed as a single protein by a baculovirus system identified serine 247 as phosphorylated. The functional relevance of this phosphorylation was investigated by mutagenesis in both replicon and virus systems, however it was shown to have no phenotype (Nordle Gilliver et al., 2010). While the deletion of this entire region (residues 244-302, JFH-1 numbering) was shown to elicit no effect on the replication or release of the JcI chimeric virus, there was however a reduction in NS5A hyperphosphorylation (Appel et al., 2008).

Similarly a phosphorylation site in the proline rich LCS II was identified using Edman degradation analysis of NS5A ectopically expressed from a polyprotein in a vaccinia-T7 system in BHK21 cells. Such analysis identified serine 2321 (lysine 345 in JFH-1) as phosphorylated, but the functional relevance of this phosphorylation could not be investigated because, at the time, both replicon and virus systems were still in development (Reed and Rice, 1999). To our knowledge, the phenotype of S2321 phosphorylation has not been retrospectively investigated.

In recent years the importance of domain III in the assembly of infectious virus has been well established (Section 1.6.3). As part of many of those studies it was shown that mutations which disrupted virus production also had effects on the phosphorylation of NS5A. In the J6/JFH (genotype 2a) chimeric virus the mutation to alanine of two serine clusters in domain III, 408-415 and 452-457 (JFH-1 numbering) dramatically reduced the basal phosphorylation of NS5A, Figure 1.11 (Masaki et al., 2008). As mentioned previously these mutations had no effect on virus replication but the second cluster abrogated virus assembly. In a ConI (genotype 1b) replicon mutation of the same two serine clusters also significantly reduced the basal phosphorylation of NS5A with no apparent effect on replication (Appel et al., 2008)

A marked reduction in hyperphosphorylation of NS5A was also observed in the culture adapted virus JFH-AM120, previously discussed, where culture adaptive mutations predominantly clustered in domain III of NS5A. It was shown that this increase in virus production was not the result of enhanced replication, implying that extensive hyperphosphorylation is not a requirement for virus production (Liu et al., 2012).

Contrary to the observations that domain III is important for basal phosphorylation Tellinghuisen *et al.* convincingly showed that alanine mutation of serine 457 (in the second serine cluster of domain III) caused a reduction in hyperphosphorylation. However in line with other studies, virus assembly was abrogated and virus replication was not affected as a result. When a phosphomimetic mutation was made, to aspartic acid, a wild type phenotype was restored (Tellinghuisen et al., 2008a). Interestingly the phosphomimetic mutation neither enhanced above WT the levels of virus release nor impaired replication, arguing that if phosphorylation at serine 457 was the "switch" to virus release, it does not switch off virus replication. The phenotype of serine 457 was also later investigated in the context of NS2 and virus assembly. It was shown that phosphoablatant mutations resulted in disruption of the punctate localisation of NS2 and that this was restored by phosphomimetic mutations (Popescu et al., 2011a).

It is worth noting that to date no putative phosphorylation sites in domain III have been confirmed biochemically by techniques such as mass spectrometry. Nevertheless the current mutagenesis data is very suggestive that phosphorylation of serines within these two clusters in domain III is requirement for the assembly of virus, Figure 1.11. Furthermore, while it is disputed in the literature as to whether phosphorylated domain III is a component of basal- or hyperphosphorylated NS5A, the consensus is that it is involved with the former.

1.6.4.2 Factors effecting NS5A phosphorylation

While the identification of NS5A phosphorylation sites is far from complete, many different virus and host factors have been shown to dramatically effect NS5A phosphorylation, summarised in Table 1.3. The formation of hyperphosphorylated NS5A is dependent on the presence of the NS3-5A polyprotein in *cis*, with the addition in *trans* of each protein being insufficient for hyperphosphorylation (Koch and Bartenschlager, 1999, Asabe et al., 1997). Furthermore, the protease activity of NS3 in *cis* was also shown as a requirement for hyperphosphorylation, together arguing that the phosphorylation of NS5A is dependent on translation (Neddermann et al., 1999).

As mentioned, several culture adaptive mutations within the LCS I result in a reduction in hyperphosphorylation and a corresponding enhancement of replication (Blight et al., 2000, Krieger et al., 2001, Neddermann et al., 2004). Specifically Blight *et al.* showed that the replication enhancing mutation S2204I (serine 232 in JFH-1) reduced hyperphosphorylation, while the equally replication enhancing mutation A2199T/S (alanine 227 in JFH-1) did not

reduce hyperphosphorylation (Evans *et al.*, 2004). However, more recent analysis of a ConI polyprotein expressed by a Vaccinia/T7 system showed that the mutation A2199S had significantly increased levels of NS5A hyperphosphorylation (Qiu et al., 2011). Moreover, it was shown that NS5A with a glycine instead of an isoleucine mutation at S2204, NS5A was still hyperphosphorylated, but to a lesser extent than WT. Suggesting that it is not simply the absence of a phosphorylation site at position 2204, but the steric/conformation influence of the bulkier isoleucine reside that reduces hyperphosphorylation (Fridell et al., 2011).

Similar effects have been observed with culture adaptive mutations in the NS4B protein, were the K1846T mutation showed a reduction in hyperphosphorylation and an enhancement to virus replication (Evans et al., 2004). Other work showed that swapping the C-terminal domain of NS4B with that of ConI or H77 isolates resulted in a specific reduction in NS5A hyperphosphorylation, no effect on replication and a reduction in virus production. (Han et al., 2013). These phenotypes were also partially rescued by adaptive mutations in NS4B and NS5A. Interestingly the culture adaptive mutations that reduce NS5A basal or hyperphosphorylation do not appear to confer a different sub-cellular localisation of NS5A, in both cases having WT localisation to lipid droplets (Han et al., 2013, Liu et al., 2012). The application of modern high resolution confocal microscopy would undoubtedly shed more light on this.

Several point mutations in NS4A, between residues 1697-1710 (polyprotein numbering), were also shown to reduce hyperphosphorylation of NS5A in a manner that was recovered by culture adaptive mutations within NS3 (Lindenbach et al., 2007). Most interestingly was the direct linear relationship between the reduction in NS5A phosphorylation and the reduction in RNA replication that was observed. That this relationship is the opposite of the many other studies previously discussed highlights that our understanding of NS5A phosphorylation is insufficient to be able to rationalise these different experimental observations, Table 1.3.

Alongside the effects of virus factors on NS5A phosphorylation many studies have identified several host cellular proteins and chemical inhibitors that also influence NS5A phosphorylation. For example, upon siRNA knockdown or chemical inhibition of the phosphatidylinositol 4-kinases-IIIα (PI4KIIIα) there is a consistent drop in NS5A basal-phosphorylation and replication (Reiss et al., 2013, Bianco et al., 2012).

	Agonist	Effect on phosphorylation	Other effects	Comments	References
Inhibitor	AL-9, PI4K inhibitor	Reduced **basal**-phosp.	Impaired replication	Same effects with siRNA knockdown	Reiss *et al.*, 2013, Bianco *et al.,* 2012.
	Undisclosed, CKI inhibitor	Reduced **hyper**-phosp.	Impaired replication		Neddermann *et al.*, 2004. Quintavalle *et al.*, 2006
	Daclatasvir, NS5A inhibitor	Reduced **hyper**-phosp.	Impaired replication	Block on NS5A phosphorylation not the mechanism of inhibition	Qiu *et al.*, 2011
	BI2536, PlkI inhibitor	Poss. reduced **hyper**-phosp.	Impaired replication	Same effects with siRNA knockdown	Chen *et al.*, 2010.
Mutations	S2197P and S2204I (serine 225 and 232 in JFH-1)	Reduced **hyper**-phosp.	Increased replicon (ConI).		Blight *et al.*, 2000. Krieger *et al.*, 2001
	Multiple LCSI mutations	Reduced **hyper**-phosp.	Impairment or enhancement to replication	Genotype dependent effect on replication. (1b enhance, 2a inhib.)	Appel *et al.*, 2005, Fridell *et al*., 2013
	S2204A (serine 232 in JFH-1)	Reduced **hyper**-phosp.	Increased replicon (ConI)		Neddermann *et al.*, 2004
	S222A	Reduced **hyper**-phosp.	Slightly reduced replication (JFH)		LeMay *et al.*, 2013
	Serine cluster domain III, (serines 408-415, 452-457 in JFH-1)	Reduced **basal**-phosp.	Wildtype replication. Impaired release		Masaki *et al.*, 2008. Appel *et al*, 2008.
	Culture adapted virus	Reduced **hyper**-phosp.	Increased virus production	Wt RNA replication. No change in LD association	Liu *et al.*, 2012
Other proteins	NS3 protease activity	Reduced **hyper**-phosp.		Wt NS3 in trans does not recover phos	Neddermann *et al.*, 1999
	Multiple NS4A point mutations	Reduced **hyper**-phosp.	Impaired replication	Linear relationship between replication and phosphorylation	Lindenbach *et al.,* 2007
	NS4B culture adaptive mutations. K1846T	Reduced **hyper**-phosp.	increase replication		Evans *et al.,* 2004.
	NS3/4A/4B in cis	Reduced **hyper**-phosp.	n/a		Koch *et al.*, 1999. Nedderman *et al* 1999
	Absence of NS4A	Reduced **hyper**-phosp.	n/a		Kaneko *et al.,* 1994. Asabe *et al.*, 1997
	Regions of NS4B	Reduced **hyper**-phosp.	Wildtype replication, reduced virus production	Chimeras of C-terminal domain from ConI and H77 isolates	Han *et al.*, 2013
	Amphiphysin II	Reduced **hyper**-phosp.	n/a		Masumi *et al.,* 2005

Table 1.3. Factors effecting phosphorylation

While it is tempting/possible that PI4KIIIα is involved with NS5A phosphorylation, there has been no reported evidence that phosphatidylinositol 4-kinases-IIIα (PI4KIIIα) directly phosphorylates NS5A, more over there is only very limited evidence of this lipid kinase having protein kinase activity (Suer et al., 2001). Nevertheless, it is also worth noting that the PI4PKIIIα binding site is between residues 187-214, just N-terminal of the phosphorylated LCS I. Observations with a PI4KIIIα inhibitor, AL-9, also showed an alteration in the basal phosphorylation state of NS5A (Bianco et al., 2012, Reghellin et al., 2012). Furthermore, both basally and hyperphosphorylated NS5A could be co-immunoprecipitated with PI4KIIIα, suggesting that phosphorylation of NS5A might not regulate this interaction (Reiss et al., 2013).

The NS5A inhibitor Daclatasvir has also been shown to have dramatic effects on the phosphorylation of NS5A, causing a rapid and significant reduction in hyperphosphorylated NS5A prior to having an effect on inhibiting replication (Qiu et al., 2011, Fridell et al., 2011).
As well as blocking phosphorylation and replication, Daclatasvir also causes a dramatic relocalisation of NS5A from the ER towards lipid droplets, and that the Daclatasvir resistance mutation Y93H is refractory to this relocalisation (Targett-Adams et al., 2011, Qiu et al., 2011). While the mechanism of action of Daclatasvir is not fully understood, it is not believed to inhibit HCV replication by blocking NS5A phosphorylation.

1.6.4.3 Kinases involved in NS5A phosphorylation

The identification of kinases that phosphorylate NS5A and the precise sites that they phosphorylate remain to be unambiguously determined. A range of kinases have been reported to phosphorylate NS5A either *in vitro* or *in vivo* and these are discussed in turn, Table 1.4.

Casein kinase I and II (CKI and CKII) have both been shown to phosphorylate NS5A in a variety of systems, and knockdown or inhibition of these kinases abolishes replication or virus assembly respectively. The recognition motifs for these two kinases are pSXXS/T and SXXE/D respectively. Ectopic expression of CKI-α, -δ and -ε with the NS3-5B polyprotein resulted in an increase in NS5A hyperphosphorylation, while the siRNA knockdown or chemical inhibition of endogenous CKI-α resulted in a reduction in hyperphosphorylation (Neddermann et al., 2004, Quintavalle et al., 2006). It was subsequently shown that CKI-α could phosphorylate a synthetic mimic of the LCS I at position serine 232 (Quintavalle et al., 2007). Interestingly kinase inhibitors with activity towards CKI enhanced the replication of a

non-culture adapted ConI SGR to levels comparable with those containing the S2204A adaptive mutation. The S2204A mutation, like the S2204I culture adapted mutation, enhances replication and results in a reduction in NS5A hyperphosphorylation (Neddermann et al., 2004).

Casein kinase II (CKII) has also been shown to phosphorylate NS5A, initially identified in an *in vitro* kinase screen (Kim et al., 1999) and subsequently shown to phosphorylate S457 in domain III in the context of Huh7 cells infected with the JcI virus (Tellinghuisen et al., 2008a). Disruption of CKII phosphorylation of NS5A - either by site mutation, drug inhibition or siRNA knockdown - was inhibitory to the release of infectious virus, but not viral RNA replication. Controversially, and discussed earlier, Tellinghuisen *et al.* implicated phosphorylated serine 457 to exist in the hyperphosphorylated species, but this is not the consensus in the literature (Appel et al., 2005, Liu et al., 2012, Masaki et al., 2008).

Kinase involved		**Region of NS5A involved**	**Effect on phosphorylation**	**System**	**Reference**
Casein Kinase I-α	CKIα	Hyper-phosphorylation	Silencing CKIα inhibited replication	*in vitro* kinase, overexpression in cell	Quintavalle *et al*, 2006
Casein Kinase I-α	CKIα	S232 (JFH-1)	-	*in vitro* kinase assay	Quintavalle *et al*, 2007
Casein Kinase II	CKII	Phosphorylated	-	*in vitro* kinase assay	Kim et al, 1999
Casein Kinase II	CKII	S457 (JFH-1), hyper-phosphorylation	Inhibition reduced virus release	Replicon (siRNA and inhib) and *in vitro* kinase assay	Tellinghuisen et al, 2008
Casein Kinase II	CKII	-	Regulation of NS5B polymerase activity	*in vitro* kinase/polymerase assay	Ivanov *et al*, 2008
Polo-like Kinase I	PLKI	DI, LCSI-DII, LCSII-DIII	Silencing PLKI inhibited replication	Replicon (siRNA and inhib) and *in vitro* kinase assay	Chen *et al*, 2010
cyclic adenosine monophosphate dependent protein kinase A-α	PKAα	phosphorylated	-	*in vitro* kinase assay	Ide *et al*, 1997
Phosphotidylinositol-4 kinase IIIα	PI4KIIIα	Possible NS5A kinase? Basal phosphorylation	-	Replicon (siRNA, chemical inhibition and mutagenesis)	Reiss *et al*, 2013

Table 1.4. Summary of known kinases involved with the phosphorylation of NS5A

Interesting observations in an *in vitro* polymerase system showed that the phosphorylation state of NS5A directed whether or not it inhibited the RdRp activity of NS5B (Ivanov et al., 2009). Specifically, CKII phosphorylated NS5A blocked the polymerase activity of NS5B, whereas CKI phosphorylated NS5A did not, and the converse was seen when a poly-A was used instead of viral RNA as RdRp template.

The polo-like kinase I (PlkI) is another kinase that was shown to *in vitro* phosphorylated NS5A (Chen et al., 2010). Each individual domain of bacterially expressed NS5A was capable of being phosphorylated by PlkI, note that domain II and III were inclusive of the LCSI and LCSII respectively. The shRNA knockdown or chemical inhibition of PlkI was sufficient to specifically block viral RNA replication, and there was indications of a concurrent reduction in hyperphosphorylated NS5A (Chen et al., 2010). The exact residue phosphorylated by PlkI *in vitro* was not determined, but the consensus recognition motif for PlkI does correlate to serines 225, 229 and 232 in the LCS I.

The kinase cAMP-dependent protein kinase A-α subunit (PKAα) has also been shown to phosphorylate NS5A in an *in vitro* kinase system, but what region of NS5A was phosphorylated or whether there was an *in vivo* function was not investigated (Coito et al., 2004, Ide et al., 1997). Similarly a yeast kinase screen identified number of kinases capable of *in vitro* phosphorylation of NS5A that have yet to be investigated further, these yeast kinases had high homology to the following human kinases: glycogen synthase kinase-3 (GSK3), protein kinase C (PKC), AKT3, serum/glucocorticoid-regulated kinase 2 (SGK2), AKT1, MEK, MKK, (Coito et al., 2004).

Chapter 2: Materials and Methods

2.1 Mammalian cell line

The human hepatocellular carcinoma cell line, Huh7, is permissive to the complete HCV lifecycle in tissue culture and was used throughout.

2.2 Plasmid and virus sequences

DNA constructs of either the full length pJFH-1 virus (Wakita *et al*., 2005), luciferase reporter sub-genomic replicon (SGR-luc JFH-1) (Targett-Adams & McLauchlan 2005) or neomycin reporter SGR with NS5A containing a One-Strep-tag (1ST) (pSGR-5A1ST) (Amako et al., 2009) were used throughout. The Litmus28i (New England Biolabs, NEB) was used as a sub-cloning vector; previously the *NsiI*/*HindIII* fragment from the JFH-1 was cloned into the multiple cloning site (MCS) of Litmus 28i.

H77	-	genotype 1a, accession number:	AF011753
ConI	-	genotype 1b, accession number:	AJ238799
JFH-1	-	genotype 2a, accession number:	AB047639
HC-J6	-	genotype 2a, accession number:	D00944

2.3 Bacteria stains

Escherichia coli (E. coli) DH5α genotype: F- ϕ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(rk-, mk+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ- (Invitrogen) were used for cloning.

2.4 Antibodies

	Target (species)	Source	Comments	Dilution
Primary	NS5A (sheep)	MacDonald et al, 2003.	Polyclonal serum	Various
	NS5A (mouse)	Lindenbach *et al*., 2005	Monoclonal, 9E10	Various
	Tubulin (rat)	Abcam	Monoclonal	Various
	GAPDH (mouse)	Abcam	Monoclonal, 6C5	1 : 20,000
	Core (sheep)	In house	Polyclonal serum	Various
Secondary	Donkey α-Sheep	Sigma-Aldrich	HRP conjugate	1 : 10,000
	Goat α-Rat	Sigma-Aldrich	HRP conjugate	1 : 10,000
	Goat α-Mouse	Sigma-Aldrich	HRP conjugate	1 : 10,000
	Donkey α-Goat/Sheep	LiCor	Fluorescence 800 nm	1 : 10,000
	Goat α-Mouse	LiCor	Fluorescence 700 nm	1 : 10,000
	Donkey α-Sheep	Invitrogen	Fluorescence, various	1 : 1,000
	Goat α-Rat	Invitrogen	Fluorescence, various	1 : 1,000
	Goat α-Mouse	Invitrogen	Fluorescence, various	1 : 1,000

Table 2.1. Summary of antibodies and dilutions.

2.5 Basic molecular biology techniques

2.5.1 Preparation of plasmid DNA from bacterial cultures

Chemical competent bacteria were prepared following the Inoue *et al.* method and transformed with plasmid DNA at a volume ratio of 50:1 by heat shock at 42 °C for 45 seconds, followed by incubation at 30 °C for 60 minutes (Inoue et al., 1990). After incubation transformed bacteria were plated onto agar plates containing the appropriate antibiotic and incubated o/n at 30 °C. Single colonies were subsequently picked from agar plates into LB medium again containing the appropriate antibiotic, cultures were grown at 30 °C, 200 rpm agitation until saturated, typically overnight. Bacterial broths were pelleted at 4000 x g, 60 mins at 4 °C, supernatants removed and bacteria cell pellet either stored at -20 °C or processed immediately. DNA was extracted by alkaline lysis and purified by anion-exchange, typically using the commercial kit Maxiprep (Qiagen). In brief, bacterial pellets were resuspended in resuspension solution (50 mM Tris.Cl pH 8.0, 10 mM EDTA, 100 µg/ml RNase A), before the addition of lysis buffer (200 mM NaOH, 1 % (w/v) SDS) and gently mixed before incubation for 5 mins, RT and the addition of neutralisation buffer (3 M potassium acetate pH 5.5). Bacterial debris was pelleted at 4000 x g, 60 min at 4 °C and the clarified supernatant containing DNA loaded onto an anion-exchange column. Long term storage of plasmid DNA was as glycerol preparations of bacterial cultures, prepared by resuspending a pelleted 1 ml bacterial culture in 30 % (v/v) glycerol : 70 % (v/v) LB medium and freezing to -80°C.

2.5.2 Polymerase chain reaction (PCR)

The polymerase chain reaction was utilised for site-directed mutagenesis (see below), PCR cloning and screening bacterial colonies. For the latter, a stab culture was generated by picking bacterial colonies into 10 µl of dH_2O and incubated for 10 mins at RT before progressing into the PCR. Reactions were composed of either 50 ng template or 10 µl of stab culture, 1 µM forward and reverse primers, 1 mM dNTP, 1 % (v/v) glycerol, 6 % (v/v) DMSO. PCR reactions were cycled depending on experiment, but in general, 30 sec at 95 °C followed by, 40 cycles of 30 sec at 95 °C, 60 sec at 55 °C, 10 min at 68 °C.

2.5.3 Site-directed (Quikchange) mutagenesis

Complementary site-directed mutagenesis primers were designed using PrimerX software (Lapid and Gao, 2003) following the constraints 25 to 45 bp, GC content 40-100 %, melting temp between 60-90 °C and 15 to 21 bp flanking sequence, Appendix figure 8.15. PCR reactions composed of 50ng template, 1 µM forward and reverse primers, 1 mM dNTP, 1 % (v/v) glycerol, 6 % (v/v) DMSO. PCR reactions were cycled as follows, 30 sec at 95 °C followed by, 5 cycles of 30 sec at 95 °C, 60 sec at 52 °C, 25 min at 68 °C then 13 cycles of 30 sec at 95 °C, 60 sec at 55 °C, 25 min at 68 °C. After PCR cycling the input template was degraded by the addition of 1 µl *DpnI* and incubation for 60 mins at 37 °C, before transformation into bacteria as described (Section 2.5.1). Mutagenesis of NS5A was performed directly on a Litmus28i (NEB) sub-clone containing an *NsiI/HindIII* fragment of the mJFH-1 cDNA, and then cloned into the relevant vector as described.

2.5.4 Protein quantification

Protein solutions, e.g. cell lysates, were quantified using bicinchoninic acid (BCA) assay (Thermo Scientific). Bovine serum albumin (BSA) was prepared to 10 mg/ml in dH_20 and serially diluted 10-fold over 12 concentrations. To each reaction well (96-well format) either 2 µl cell lysate or 10 µl BSA standards was added, along with either 10 µl dH20 or 2 µl of the buffer used for the cell lysate followed by the addition of 88 µl BCA reaction mixture as directed. The 96-well plate was incubated at 37 °C for 20 minutes before reading absorbance at 570 nm and manual calculation protein concentrations using a linear regression.

2.5.5 Gel electrophoresis, SDS-PAGE and Western blot

Cells were washed twice with PBS, lysed by resuspension in Glasgow lysis buffer (GLB – 1 % (v/v) Triton X-100, 120 mM KCl, 30 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 10 mM PIPES-NaOH, pH 7.2) containing protease and phosphatase inhibitors, and incubated on ice for 15 mins. Insoluble material was pelleted by centrifugation at 500 x g for 5 mins at 4 °C. Following separation by SDS-PAGE proteins were transferred to a polyvinylidene fluoride (PVDF) membrane and blocked in 50 % (v/v) Odyssey blocking buffer (LiCor) in PBS. The membrane was incubated with primary antibody in 50 % (v/v) Odyssey blocking buffer, then incubated with fluorescently labelled α-Goat (800nm) or α-Mouse (700 nm) before imaging on a LiCor Odyssey Sa fluorescent. Primary antibodies α-NS5A (9E10) (Lindenbach et al., 2005); α-NS5A (sheep) and α-Core (sheep), in house.

2.5.6 Bacterial expression of GST-SH3 domains

GST tagged fusion proteins of SH3 domains were created previously and expressed as described in (Macdonald et al., 2004). In brief, bacteria were transformed with DNA (Section 2.5.1) and a single colony used to inoculate 10 ml LB and grown o/n at 30 °C. This saturated starter culture was then used to inoculate 100 ml of LB which was subsequently grown at 30 °C until it reached an OD of 0.6, and induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). After 4 hrs of induction bacteria were cooled to 4 °C, pelleted at 2,000 x g for 30 mins at 4 °C. Bacterial pellet was resuspended in lysis buffer (1 % (v/v) Triton X-100 in PBS), sonicated for 10 cycles of 45 sec and cell debris pelleted at 4,000 x g. GST fusion proteins were subsequently immobilised to glutathione resin (GE Healthcare) o/n at 4 °C following manufacturers guidelines. Decorated glutathione resin was then used in NS5A interaction experiments.

2.6 Nucleic acid techniques

2.6.1 Phenol/chloroform purification

Nucleic acid was purified from solutions containing proteins (typically following an endonuclease or *I.V.T.* reaction) by phenol/chloroform extraction. An equal volume of phenol:chloroform:isoamylalcohol (25:24:1, pH 8 for DNA or pH 5.2 for RNA) (Sigma-Aldrich) was added to the reaction, vortexed and phases separated by centrifugation at 16,000 x g for 5 mins at RT. Aqueous phase was removed and added to an equal volume of chloroform, vortex and phases separated as described. DNA was purified from second aqueous phase by ethanol precipitation, see below.

2.6.2 Ethanol precipitation

Nucleic acid was desalted and concentrated, typically after phenol/chloroform purification, by ethanol precipitation. To nucleic acid the addition of 2 volumes 100 % (v/v) ethanol, and 0.1 volumes 3M sodium acetate, pH 5.2, before cooling to -20 °C for 20 mins and pelleting at 16,000 x g for 30 mins. DNA pellet was washed in 70 % (v/v) ethanol and resuspended in dH_2O, typically 20-100 µl.

2.6.3 Quantification

Nucleic acid was quantified by absorbance at 260 nm by NanoDrop (Thermo Scientific) or direct comparison with markers after agarose electrophoresis, Hyperladder I (Bioline) or Millenium Markers (Millipore) for DNA and RNA respectively.

2.6.4 Endonuclease digestion

Typically 3-10 µg of DNA was digested with 10 U of commercial endonucleases (*BamHI*, *RsrII*, *AfeI* etc.) in the relevant commercial buffers for 4 – 8 hrs. A standard temperature of 37 °C was used unless denoted otherwise by the manufactures. Digested DNA fragments were purified by agarose gel electrophoresis and extracted from the gel using the commercial QIAquick gel extraction kit (Qiagen).

2.6.5 Agarose gel electrophoresis

Gels were prepared at between 0.8 - 1.0 % (w/v) agarose in 1 x TAE (40 mM Tris, 0.11 % (v/v) acetic acid, 1 mM EDTA with the addition of SYBR safe at 1:10,000 dilution. Samples were prepared in loading buffer (1 x TAE, 5 % (v/v) glycerol and containing bromophenol blue) before loading onto gel and electrophoresing at 100 V for between 15-60 mins. Gels were imaged with blue light and orange filter.

2.6.6 DNA sequencing and analysis

DNA was sequenced by commercial company using Sanger sequencing techniques with custom primers designed against the relevant sequence, see Appendix figure 8.16.

2.6.7 In vitro transcription of RNA

DNA constructs, e.g. mSGR-luc JFH-1 or mJFH-1, were linearised by *XbaI* and overhanging ends degraded by mung bean nuclease treatment. DNA was phenol/chloroform extracted and 1 µg of linearised DNA used as template in a 20 µl RiboMAX reaction (Promega). RNA transcripts were purified by phenol/chloroform extraction and quantified by absorbance at 260 nm. RNA was also analysed by agarose gel electrophoresis to confirm integrity and quantification.

2.7 Mammalian tissue culture

2.7.1 Mammalian tissue culture

Huh7 cells were cultured in Dulbecco's Modified Eagles Medium (DMEM; Sigma) supplemented with 10 % foetal bovine serum (FBS), 100 IU penicillin/ml, 100 µg/ml streptomycin and 1 % non-essential amino acids (Lonza) in a humidified incubator at 37 °C, 5 % CO_2. Cells were passaged every three days by washing in phosphate buffered saline (PBS) prior to incubating with trypsin-EDTA solution (Sigma-Aldrich) for 5 mins to detach cells, then resuspending in complete culture media before either reseeding for further passage or using for experiments. For virus propagation media was supplemented with 25 mM HEPES.

Replicons containing a neomycin resistance gene were passaged 1:5 every 3 days under 0.5 mg/ml G418 (Sigma-Aldrich).

2.7.2 Transfection of DNA

DNA was transfected into cells using FuGene 6 (Promega) transfection reagent according to the manufacturer's directions. For example, Huh7 cells were seeded at $2x10^5$ cells / 6-well 24 hrs prior to transfection. Transfection reaction was prepared by adding 1 µg plasmid DNA to 98 µl Opti-MEM (Life technologies) followed by the addition of 3 µl FuGene 6, gentle mixing and incubation at RT for 30 mins. Transfection reaction was then added drop-wise to seed cells, without exchange of media. Typically at 48 hours post-transfection (hpt) cells were washed twice in PBS and lysed in GLB.

2.7.3 Immunofluorescence

In experiments utilising confocal microscopy based immunofluorescence (IF) cells were seeded onto 20 mm sterile glass coverslips housed in normal 12-well tissue culture plates. Cells were fixed by washing twice in PBS before the addition of 4 % (w/v) paraformaldehyde (4 % PFA) in PBS and incubation at RT for 20 minutes. Cells were then washed in PBS twice before either storing at 4 °C or permeabilising by the addition of 0.1 % (v/v) Triton-X100 in PBS for 7 mins followed by washing in PBS. After cells were fixed and permeabilised they were ready for immunostaining. Primary and secondary staining was carried out in 5 % (v/v) FBS in PBS, in the case of primary for 1 hr at RT or o/n at 4 °C and for secondary staining 1 hr at RT. Antibody dilutions denoted in Table 2.1. After staining coverslips were air dried to near complete dryness before mounting onto glass slides with Prolong gold (Invitrogen).

2.7.4 SNAP/CLIP-tag labelling of proteins

Labelling solution was prepared by dissolving SNAP/CLIP-substrates (e.g. SNAP-505, NEB) in DMSO to 1 mM, before diluting in complete DMEM culture medium to a final concentration of to 5 µM. Huh7 cells were subsequently incubated with labelling solution for 1 hr under normal tissue culture conditions before washing twice with complete media, and incubating for a further 30 mins with fresh complete media to allow unreacted dye to wash out. Cells were then either fixed in 4 % PFA as described earlier or imaged directly.

2.8 Sub-genomic replicon assays

2.8.1 Transient replication assay

Huh7 cells were washed twice in diethylpyrocarbonate (DEPC)-treated PBS before electroporating $4x10^6$ cells in DEPC-PBS with 2 µg of RNA at 950 µF and 270 V. Cells were resuspended in complete media before being seeded into both 96-well plates (n=6) at $3x10^4$ cells / well, and 6-well plates (n=2) at $3x10^5$ cells / well, both plates incubated under cell culture conditions. After 4, 24, 48 and 72 hours post-electroporation (hpe) cells were harvested by lysis with either 30 µl or 200 µl passive lysis buffer (PLB; Promega), 96- and 6-well respectively. Luciferase activity was determined from 96-well samples on a BMG plate reader by automated addition of 50 µl LarI reagent (Promega) and total light emission monitored over 12 seconds

2.8.2 Drug efficacy assays (EC50)

Huh7 cells were electroporated and seeded as described for luciferase based replication assay in 96-well plates (Section 2.8.1). The drug to be tested was serially diluted 12 times with 5-fold dilutions in DMSO, followed by further dilutions in complete DMEM media until DMSO concentration was < 0.5 % (v/v). At 4 hpe the cells were treated with the serially diluted drug, incubated under normal conditions and harvested at 48 hpe by washing the cells in PBS followed by the addition of 30 µl PLB. Luciferase activity was then determined by addition of 50 µl luciferase reagent (LAR I, Promega). Data was modelled using a log[agonist] vs. response model (Appendix Figure 8.10A) and EC50 calculated using Prism (Graphpad).

2.8.3 Drug toxicity assay (CC50)

Drug toxicity assays were carried out in the same manner (typically in parallel) as the drug efficacy assays. However at 48 hpe there was an addition of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT, Sigma) to cells at a final concentration of 1 µM, followed by incubated for 2 hrs under normal TC conditions. Following this media was removed from cells and the MTT precipitate dissolved by the addition of 100 µl DMSO and absorbance measured at 570 nm. Data was modelled and CC50 calculated as for EC50.

2.9 Full length virus assay

2.9.1 Electroporation of cells

Huh7 cells were washed twice in DEPC-treated PBS before electroporating $2x10^6$ cells in DEPC-PBS with 1 µg of RNA at 950 µF and 270 V. Cells were resuspended in complete media and seeded at $1x10^6$ cells / well into 6-well plates. At 72 hpe cells were split at 1:5 before

incubating for a further 72 hrs. At 144 hpe cells were harvested in 400 µl TRIzol for qRT-PCR analysis while supernatants were removed and clarified at 2800 x g for 5 mins at RT before storing at -80 °C.

2.9.2 Titration of virus by focus forming assay

Virus supernatants were clarified at 2,800 x g for 5 minutes at RT before being titred on Huh7 cells as follows. In 96-well format, clarified virus supernatants were serially diluted 5-fold in complete DMEM media + HEPES before the addition of 100 µl diluted virus to Huh7 cells that were seeded 24 hrs before into 96-well plates at $8x10^3$ cells / well, final volume 200 µl. Cells incubated under normal cell culture conditions for 72 hrs before washing in PBS and fixation in 4 % PFA for 20 min. Cells were permeabilised in 0.1 % (v/v) Triton X-100, 10 % (v/v) FBS, 1 X PBS for 7 mins followed by staining with α-NS5A (sheep) (Macdonald *et al*., 2003) at 1:5000 followed by a corresponding secondary fluorescent antibody. Foci were counted manually to determine virus titres.

2.9.3 Genome quantification by qRT-PCR

To quantify the number of HCV genomes total cell RNA was extracted by TRIzol following manufacturer's instructions (Invitrogen). Total extracted cellular RNA, 100 ng, was analysed by qRT-PCR using a one-step qRT-PCR TaqMan based kit as directed (Eurogentec), primers and probe designed against the 5' UTR described previously (Takeuchi *et al*., 1999).

2.10 Expression and purification of NS5A-1ST

2.10.1 Generating replicon stable cell line

A stable Huh7 cell line harbouring the SGR-5A1ST subgenomic replicon (Amako et al., 2009) was passaged 1:5 every 3 days under 0.5 mg/ml G418 selection (Sigma) and cell passage number was maintained low to avoid culture adaptations. At each passage excess cells were washed twice in PBS, pelleting at 500 x g, 5 min at RT, before the resuspension in GLB, 2 ml per T175, and storing at -20 °C until NS5A-OST purification.

2.10.2 Purification of NS5A-OST by affinity column

In batch, lysates were clarified by centrifugation at 4000 x g, 5 min at 4 °C, filtered through 0.45 µm cellulose filter before affinity purification of NS5A-OST in batch on Strep-Tactin Sepharose columns, 5 ml CV (IBA) following the manufacturer's instructions. Bound protein washed with 1 CV of wash buffer supplemented with low salt (0.5 M NaCl) or high salt (1 M NaCl, 1 % (v/v) Triton X-100) for three and four washes respectively before two standard

washes and competitive elution with 2.5 mM desthiobiotin, 0.5 CV. Peak elution fractions, 3 and 4, were identified by Western blot analysis and concentrated, first by centrifugation filtration (10 kDa MCO) followed by methanol/chloroform precipitation. Final precipitate was resuspended in 20 µl of 1X SDS-LB and separated further on a 12 % Tris-Glycine SDS-PAGE. After staining with Coomassie R250 followed by de-staining in 50 % (v/v) water, 40 % (v/v) MeOH, 10 % (v/v) acetic acid the band corresponding to NS5A was excised for mass spectrometry analysis.

2.11 Phospho-mapping by Mass spectrometry (MS)

2.11.1 Enzymatic digestion of protein

In-gel proteolytic digestions with endoproteinases trypsin and Glu-C were performed after reduction with DTE and S-carbamidomethylation with iodoacetamide. Gel pieces were washed two times with 50 % (v/v) aqueous acetonitrile containing 25 mM ammonium bicarbonate, then once with acetonitrile and dried in a vacuum concentrator for 20 min. Endoproteinase Glu-C from *Staphylococcus aureus* strain V8 (Sigma-Aldrich) and sequencing grade modified porcine trypsin (Promega) were reconstituted in 25 mM ammonium bicarbonate to give a final enzyme concentration of 0.004 µg/µl. Gel pieces were rehydrated by adding 50 µl of either endoproteinase solution and allowed to digest overnight at 37 ^{o}C.

2.11.2 Enrichment by Titanium dioxide affinity

Peptide solutions were dried down in a vacuum concentrator to approximately 10 µl and diluted to 100 µl with a solution of 1 M glycolic acid in 80 % (v/v) aqueous acetonitrile, 5 % (v/v) trifluoroacetic acid. Phosphopeptides were selectively enriched using manually packed titanium microcolumns as described by (Larsen et al., 2005). Enriched Phosphopeptides were purified by Poros Oligo R3 reversed-phase a material and eluted directly onto a ground steel MALDI target plate with freshly-prepared 10 mg/ml solution of 2,5-dihydroxybenzoic acid (Sigma-Aldrich) in 50 % (v/v) aqueous acetonitrile, 1 % (v/v) phosphoric acid.

2.11.3 Separation by HPLC

The nanoLC system was interfaced with a maXis LC-MS/MS System (Bruker Daltonics) with a nano-electrospray source fitted with a steel emitter needle (180 µm O.D. x 30 µm I.D., Proxeon). Positive ESI- MS & MS/MS spectra were acquired using AutoMSMS mode. Instrument control, data acquisition and processing were performed using Compass 1.3 SP1 software (microTOF control, Hystar and DataAnalysis, Bruker Daltonics). Instrument settings were: ion spray voltage: 1,500 V, dry gas: 6 L/min, dry gas temperature 160 °C, ion

acquisition range: *m/z* 50-2,200. AutoMSMS settings were: MS: 0.5 s (acquisition of survey spectrum), MS/MS (CID with N_2 as collision gas): ion acquisition range: *m/z* 350-1,400, 0.1 sec acquisition for precursor intensities above 100,000 counts, for signals of lower intensities down to 1,000 counts acquisition time increased linear to 1.5 sec, the collision energy and isolation width settings were automatically calculated using the AutoMSMS fragmentation table:, 3 precursor ions, absolute threshold 1,000 counts, preferred charge states: 2 – 4, singly charged ions excluded. 2 MS/MS spectra were acquired for each precursor and former target ions were excluded for 60 sec.

2.11.4 ESI MS analysis

Samples were loaded onto a nanoAcquity UPLC system (Waters) equipped with a nanoAcquity Symmetry C_{18}, 5 µm trap (180 µm x 20 mm Waters) and a nanoAcquity BEH130 1.7 µm C_{18} capillary column (75 µm x 250 mm, Waters). The trap wash solvent was 0.1 % (v/v) aqueous formic acid and the trapping flow rate was 10 µl/min. The trap was washed for 5 min before switching flow to the capillary column. The separation used a gradient elution of two solvents (solvent A: 0.1 % (v/v) formic acid; solvent B: acetonitrile containing 0.1 % (v/v) formic acid). The flow rate for the capillary column was 300 nl/min Column temperature was 60 °C and the gradient profile was as follows: initial conditions 5 % (v/v) solvent B (2 min), followed by a linear gradient to 35 % (v/v) solvent B over 20 min, then a wash with 95 % (v/v) solvent B for 2.5 min. The column was returned to initial conditions and re-equilibrated for 25 min before subsequent injections

2.11.5 MALDI MS analysis

Positive-ion MALDI mass spectra were obtained using an ultraflex III mass spectrometer (Bruker, Bramen, Germany) in reflectron mode, equipped with a Nd:YAG smart beam laser. MS spectra were acquired over a mass range of 800-4000 *m/z*. Final mass spectra were externally calibrated against an adjacent spot containing 6 peptides of known *m/z* (des-Arg^1-Bradykinin, 904.681; Angiotensin I, 1296.685; Glu^1-Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929.). Peptides were manually selected for MS/MS fragmentation, which was performed in LIFT mode without the introduction of a collision gas. Fragmentation spectra were acquired manually with laser intensity and number of summed spectra optimised for each acquisition. The default calibration was used for MS/MS spectra, which were baseline-subtracted and smoothed (Savitsky-Golay, width 0.15 m/z, cycles 4); monoisotopic peak detection used a SNAP averaging algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) with a minimum S/N

of 6. Bruker flex. Analysis software (version 3.3) was used to perform the spectral processing and peak list generation for both the MS and MS/MS spectra.

2.11.6 Computational analysis of spectra

Tandem mass spectral data were submitted to database searching using a locally-running copy of the Mascot program (Matrix Science Ltd., version 2.3), through the Bruker ProteinScape interface (version 2.1). All spectral data were searched against an in-house database containing the expected protein sequence. Search criteria specified: Enzyme, either trypsin or Glu-C; Fixed modifications, Carbamidomethyl (C); Variable modifications, Oxidation (M) Phosphorylation (S,T); Peptide tolerance, 250 ppm for MALDI-MS/MS and 10 ppm for LC-MS/MS; MS/MS tolerance, 0.5 Da for MALDI-MS/MS and 0.1 Da for LC-MS/MS; Instrument, MALDI-TOF-TOF or maXis.

2.12 Statistical analysis of data

Statistical analysis of data was carried out using a Students t-test assuming a two tailed distribution with a unequal variance. Error bars presented on all graphs illustrate the Standard Error of the Mean.

Chapter 3: The phosphorylation of NS5A

3.1 Introduction

The NS5A protein is a highly phosphorylated protein that migrates as two different bands on a SDS PAGE gel, termed the basally phosphorylated and hyperphosphorylated species. These have an apparent molecular weight of 56 and 58 kDa respectively, however in genotype 2a the NS5A protein has a 19 aa insertion in domain III meaning it migrates as a 63 and 65 kDa species, Figure 3.1A.

Our current understanding about regions of NS5A important for phosphorylation, and what the functional role of phosphorylation is, has been reviewed extensively in (Section 1.6.4). Much of this literature has focused on the importance of LCS I and domain III in the phosphorylation of NS5A. The dogma is that basal-phosphorylation is important for RNA replication and hyperphosphorylation is important for virus production, however this dogma is no longer supported by the literature. For example, the LCS I has been shown to be important for hyperphosphorylation (Lemay et al., 2013, Tanji et al., 1995) but mutations within this region have been shown to both enhance as well as impair replication in a genotype dependent manner (Appel et al., 2005, Fridell et al., 2013, Lemay et al., 2013). Similarly domain III is known to be essential for virus production, but mutations of serine clusters with those regions results in a reduction of NS5A basal-phosphorylation with no effect on hyperphosphorylation (Appel et al., 2005, Masaki et al., 2008). There is however evidence that phosphorylated serine 457 is a component of the hyperphosphorylated species (Tellinghuisen et al., 2008a).

Much of the established literature on the phosphorylation on NS5A has been derived from either reverse genetic approach (Appel et al., 2005, Masaki et al., 2008, Tanji et al., 1995, Tellinghuisen et al., 2008a) or biochemical analysis of ectopically expressed NS5A (Katze et al., 2000, Nordle Gilliver et al., 2010, Reed and Rice, 1999). The use of reverse genetics has strongly implicated several serines as being sites of phosphorylation, however it is possible that these serines are not phosphorylated but are instead involved in kinase recognition or protein structure. As such biochemical observation, typically by MS, is required to confirm which serines are actual phospho-acceptors.

The ectopic expression of NS5A has facilitated the MS identification of two serines (Katze et al., 2000, Nordle Gilliver et al., 2010). However, it is well established that the correct phosphorylation of NS5A is dependent on - at a minimum - the presence of other virus proteins as well as non-kinase host proteins (Koch and Bartenschlager, 1999, Reiss et al.,

2013). Therefore whether these serines are phosphorylated in the context of a replicon or virus remains to be shown, especially as mutation of these sites elicited no phenotype in replicon (Nordle Gilliver et al., 2010).

Indeed with the exception of a very recent study (Lemay et al., 2013) the mass spectrometry identification of phosphorylation sites on NS5A purified from infected cells has not been done. The technical challenges associated with this type of approach meant that Lemay *et al.* only managed to identify a single phosphorylation at position S222. When a virus harbouring an ablation of the S222 site (S222A) was labelled with ^{32}P they observed only a slight reduction in intensity of the hyperphosphorylated species and no change in the basally phosphorylated species (Lemay et al., 2013). This again supporting the idea that NS5A has significantly more than one phosphorylation site, however these were likely not observed by LeMay *et al.* for reasons surrounding the detection of phosphopeptides by MS discussed previously (section 1.5.1). If more phosphorylation sites are to be identified by MS then significantly more purified protein is required, alongside new enzymatic digestion and phosphopeptide enrichment approaches.

Taken together this highlights how poor our current understanding is of the complexities of NS5A phosphorylation and how it is involved in the virus lifecycle. The aim of this study was therefore to identify and characterise novel phosphorylation sites in the context of an actively replicating system. It was expected to shed light on exactly which residues in the serine rich clusters of the LCS I and domain III are actual sites of phosphorylation. This more detailed phosphorylation map would then be able to facilitate the production of phosphospecific antibodies, which could in turn be used to probe the role of each phosphorylation site in greater detail.

To facilitate maximum expression levels of NS5A a Huh7 cell line a subgenomic replicon with a neomycin resistance gene was utilised, Figure 1.8. This stable replicon cell line could be sequentially passaged on a large scale and under selectional pressure, allowing for the purification of NS5A at significantly higher levels than previously and from Huh7 cells actively replicating HCV RNA.

3.2 Results

3.2.1 Identification of NS5A phosphorylation sites by mass spectrometry.

The human hepatocellular carcinoma cell line, Huh7, containing a HCV subgenomic replicon was used to express NS5A in a functionally relevant context, and on a large enough scale so as to facilitate the MS identification of phosphorylated residues. Advantage was taken of the ability of NS5A domain III to tolerate in-frame insertions with no effect on the function of the protein in virus RNA replication, and only limited impairment to virus production, discussed in (Section 5.1). The JFH-1 NS3-5B subgenomic replicon in which NS5A was tagged with the One-Strep affinity tag (OST) was therefore utilised. This tag had been previously shown not to affect RNA replication (Amako et al., 2009) and provided a highly efficient and selective purification strategy. This allowed the purification of NS5A with a higher efficiency and greater stringency than a traditional IP approach.

A stable cell line harbouring the above replicon was cultured on a large scale and over several passages. At each passage surplus cells were pelleted at 1,200 x g for 5 mins and resuspended in Glasgow lysis buffer (Section 2.5.52.5.5). It is important to note that during the propagation of replicon harbouring cells the passage number was kept low, as a loss in hyperphosphorylated NS5A following long term passage of replicon harbouring cells is routinely observed (see Figure 3.1A). The sequence analysis of replicons following long term passage did not reveal any mutations in NS5A (data not shown) but mutations elsewhere in the genome were not investigated.

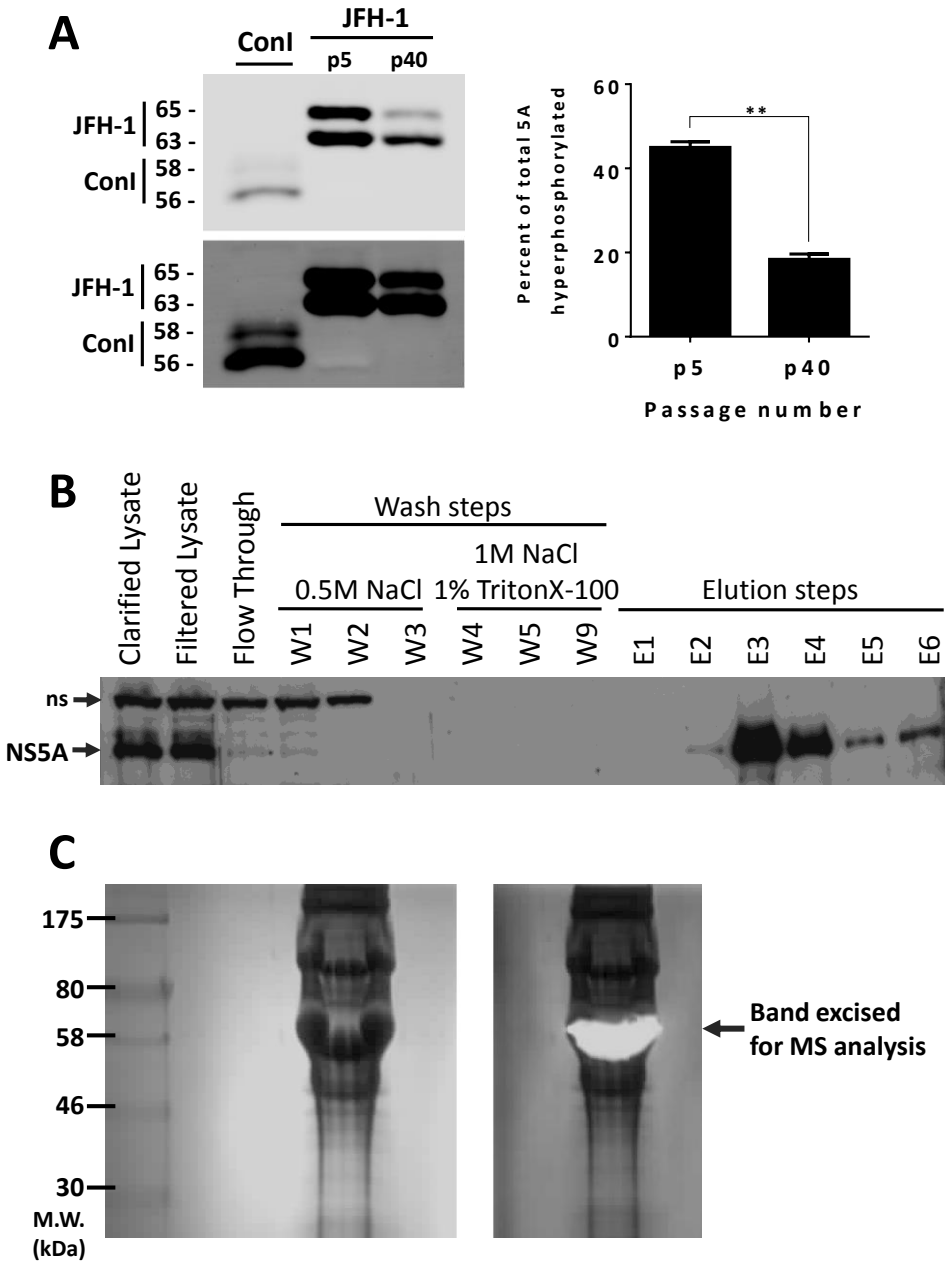


Figure 3.1. Summary of NS5A and NS5A-OST purification.

A Huh7 cells harbouring ConI or JFH-1 NS3-5B replicons were lysed at the indicated passage (p) in GLB and analysed by SDS-PAGE, western blot probed with α-NS5A (sheep) and fluorescent secondary, followed by quantification of different phosphorylation species. ** $p<0.01$. **B.** Purification of NS5A-OST by Strep-Tactin affinity column following (Section 2.10.2). ns = non-specific band **C.** Concentrated NS5A-OST was separated further by 12 % SDS PAGE, Coomassie stained and the band corresponding to NS5A excised for MS phospho-mapping.

The NS5A-OST was affinity purified from cell lysates using Strep-Tactin beads from cytoplasmic lysates as described (Section 2.10.2). Purification was conducted in batch, and an example western blot illustrating the different steps is shown in Figure 3.1B. In total NS5A-OST was purified from cell lysates corresponded to approximately 5 x 10^9 cells. Purified NS5A-OST was concentrated by centrifugation filtration (10 kDa MWCO), followed by methanol/chloroform precipitation prior to separation by SDS-PAGE. After staining of the gel with Coomassie Brilliant Blue R250, the band corresponding to NS5A was excised (Figure 3.1C). The excised gel slice was subjected to trypsin cleavage, TiO_2 enrichment of phosphorylated peptides and mass spectrometry analysis by MALDI, ESI and tandem MS-MS to identify the phosphorylated species (Section 2.11). In both systems ions were automatically selected for collision-induced dissociation (CID) to generate MS/MS spectra from which the phosphopeptide sequence and location of phosphorylation could be assigned.

A total of 10 phosphopeptides were identified that corresponded to six different regions of NS5A and contained a total of 12 phosphorylation sites. The acquisition of MS/MS spectra on all 10 phosphopeptides allowed for sequence assignment. Firstly the phosphopeptide with m/z 1120 corresponded to NS5A domain I and showed that serine 146 was phosphorylated. Secondly, the two phosphopeptides with m/z 2023 and 2103 corresponded to the LCS I and showed a single phosphorylation at serine 222 and a double phosphorylation of serine 222 and 225. The fourth phosphopeptide with m/z 943 corresponded to the LCS II and showed that threonine 348 was phosphorylated. The MS spectra for the four assigned phosphopeptides serine 146, 222, 222/225 and threonine 348 are shown in Figure 3.2 and corresponding MS/MS spectra along with b- and y-ion assignment is shown in Appendix figures 8.1, 8.2, 8.3 and 8.4.

A further three phosphopeptides with a m/z of 2183, 2263 and 2343 also corresponded to the LCS I region with a m/z consistent with the addition of 3, 4 and 5 phosphates respectively, Figure 3.2C. While the MS/MS spectra allowed sequence assignment of the phosphopeptide, it was not possible to assign the exact phosphorylation sites, appendix figures 8.5, 8.6 and 8.7. In the case of the tri- and tetra-phosphorylated peptides the MS/MS data supports the presence of multiple isomers based on observation of four and five phosphorylated serines in the tri- and tetra-phosphorylated peptides respectively. In the case of the penta-phosphorylated species the ion abundance was too low for complete

assignment. The partial assignment of the tri- tetra- and penta-phosphorylated peptides was however sufficient to confirm that serines 229, 230, 232 and 235 were amongst the phosphorylated residues.

A further two phosphopeptides with m/z 2423 and 2503 were identified by MS and had a m/z consistent with the addition 6 and 7 phosphorylations respectively on the same LCS I peptide fragment described. These putative hexa- and hepta-phosphorylated peptides co-eluted from the HPLC with the other LCS I phosphopeptides and so are observed in the same MS spectrum, figure 3.C. However, ion suppression effects caused by such extensive phosphorylations resulted in insufficient ion abundance for MS/MS data to be collected, and therefore not even a partial assignment could be carried out. For similar reasons, the presence of an octa-phosphorylated species of this peptide fragment 221-240 cannot be excluded. Our data point to a complex pattern of phosphorylation in this region of NS5A, discussed later.

Finally three other phosphopeptides with m/z 3724, 1931 and 1349 corresponded to NS5A residues 241-273, 375-395 and 396-409 respectively. The MS/MS allowed for complete sequence assignment but only partial assignment of the phosphorylation site. The region of partial assignment is denoted in Table 3.1.

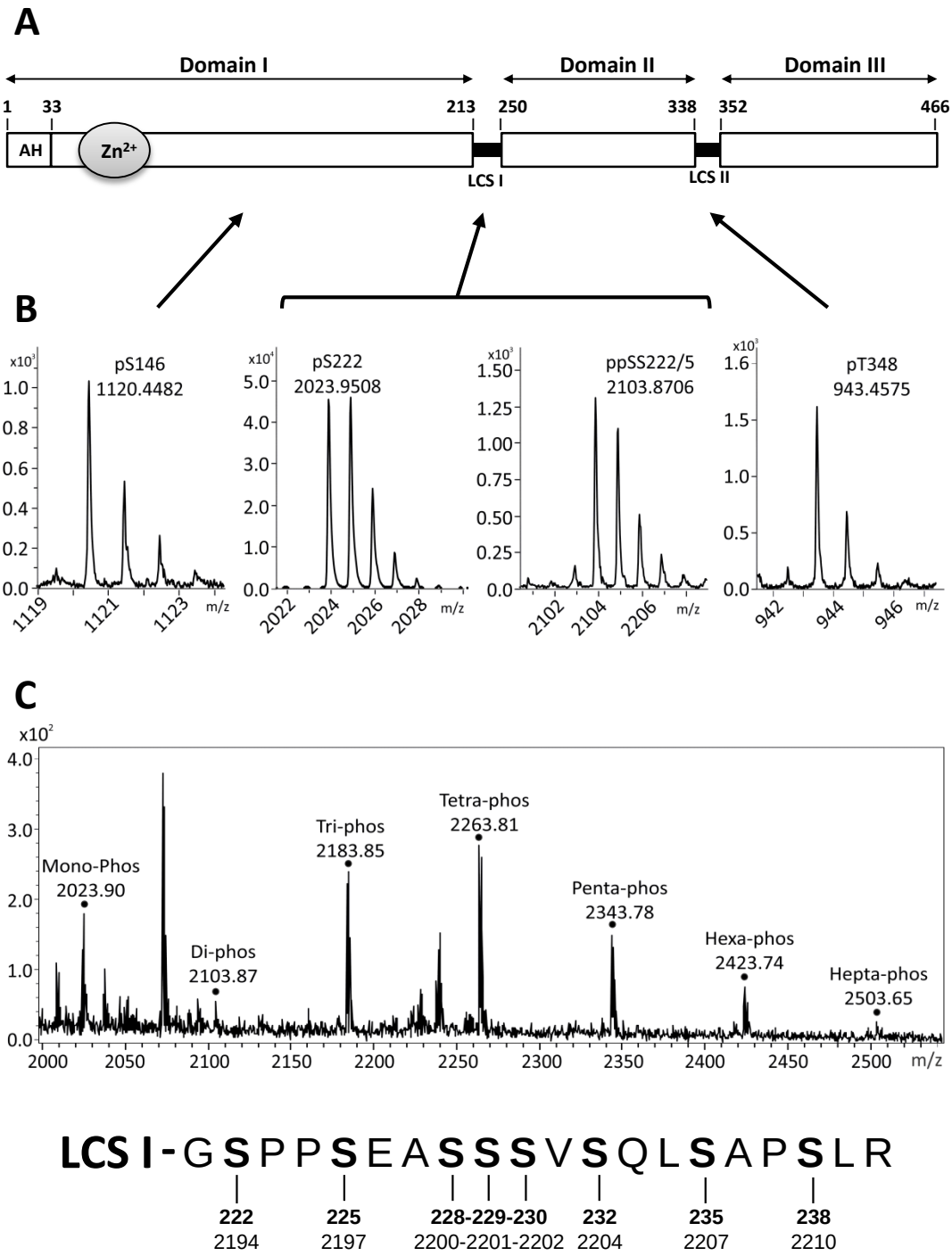


Figure 3.2. Mass spectrometric analysis of NS5A phosphorylation.

A. Schematic of NS5A structure showing amino acid residue numbers for JFH-1, amphipathic helix (AH), zinc binding motif (Zn^{2+}), and low complexity sequences (LCS). **B.** The mass spectra of NS5A phosphopeptides after analysis by HPLC-MS/M, the main monoisotopic peaks were selected for collision induced dissociation to allow the precise assignment of phosphorylation sites, for MS/MS spectra see appendix figures 8.1-8.4. **C.** Mass spectra of phosphopeptide fragment corresponding to the LCS I region (sequence shown), for MS/MS see appendix figure 8.5-8.7.

Peptide fragment observed by mass spectrometry	Phosphorylation number and location	Calculated $[M+H]^+$	Observed $[M+H]^+$	Type of MS analysis		
				MALDI	ESI	MS/MS
KAP**pT**PPPR	1 - T348	943.4761	943.4575	Y	N	Y
IPCQLP**pS**PE	1 - S146	1120.4743	1120.4482	Y	Y	Y
G**pS**PPSEASSSVSQLSAPSLR	1 - S222	2023.9332	2023.9508	Y	Y	Y
G**pS**PP**pS**EASSSVSQLSAPSLR	2 - SS222/225	2103.8995	2103.8706	Y	Y	Y
GSPPSEASSSVSQLSAPSLR	3 - multiple isomers	2183.8659	2183.8477	Y	N	Y
GSPPSEASSSVSQLSAPSLR	4 - multiple isomers	2263.8322	2263.811	Y	N	Y
GSPPSEASSSVSQLSAPSLR	5 - multiple isomers	2343.7985	2343.7756	Y	N	Y
GSPPSEASSSVSQLSAPSLR	6 - unassigned	2423.7648	2423.7419	Y	N	N
GSPPSEASSSVSQLSAPSLR	7 - unassigned	2503.7311	2503.7082	Y	N	N
A**T**C**TT**H**S**N**T**YDVDMVDANLLM EGGVAQTEPESR	1 - T242, T244, T245, S247, T249	3724.5449	3724.5441	N	Y	Y
TFGQPPSSGDAG**SS**TGAGAAE	1 - S387, S388	1931.7655	1931.7629	N	Y	Y
SGGP**TS**PGEPAPSE	1 - S396, T400, S401	1349.5257	1349.5414	Y	N	Y

Table 3.1. Summary of identified phosphopeptides of NS5A
Summary of observed phosphorylated peptides and, where possible, assignment of phosphorylated residue (red) and partial assignment of phosphorylated residues (underlined).

3.2.2 Role of phosphorylation sites in the HCV lifecycle.

To investigate the role of the four identified phosphorylation sites (146, 222, 225 and 346) in the virus life cycle each residue was mutated to either alanine to block phosphorylation (phosphoablatant), or aspartic acid to introduce a negative charge, thereby acting as a phosphomimetic. These mutations were introduced into the luciferase based subgenomic replicon (mSGR-luc) and the full length JFH-1 virus construct (Figure 4.2), allowing investigation of the role of phosphorylation at different stages of the virus lifecycle (Hughes et al., 2009b). In addition to the four individual mutations, the two serines at 222 and 225 were mutated simultaneously. Details regarding the manner in which the mutations were inserted into the above constructs is explained in further detail in Section 4.2.1. A NS5B polymerase inactive mutant, where the catalytic GDD motif had been mutated to GND, was included as a negative control in all assays.

The effect of these mutations on replication was determined by a luciferase based replication assay. *In vitro* transcripts of phosphomutant replicons were electroporated into Huh7 cells and luciferase activity measured at 4 and 72 hpe. It has been previously

established that luciferase activity correlates with the number of intracellular HCV genomes and so is used as a measure of replication competency (Targett-Adams and McLauchlan, 2005).

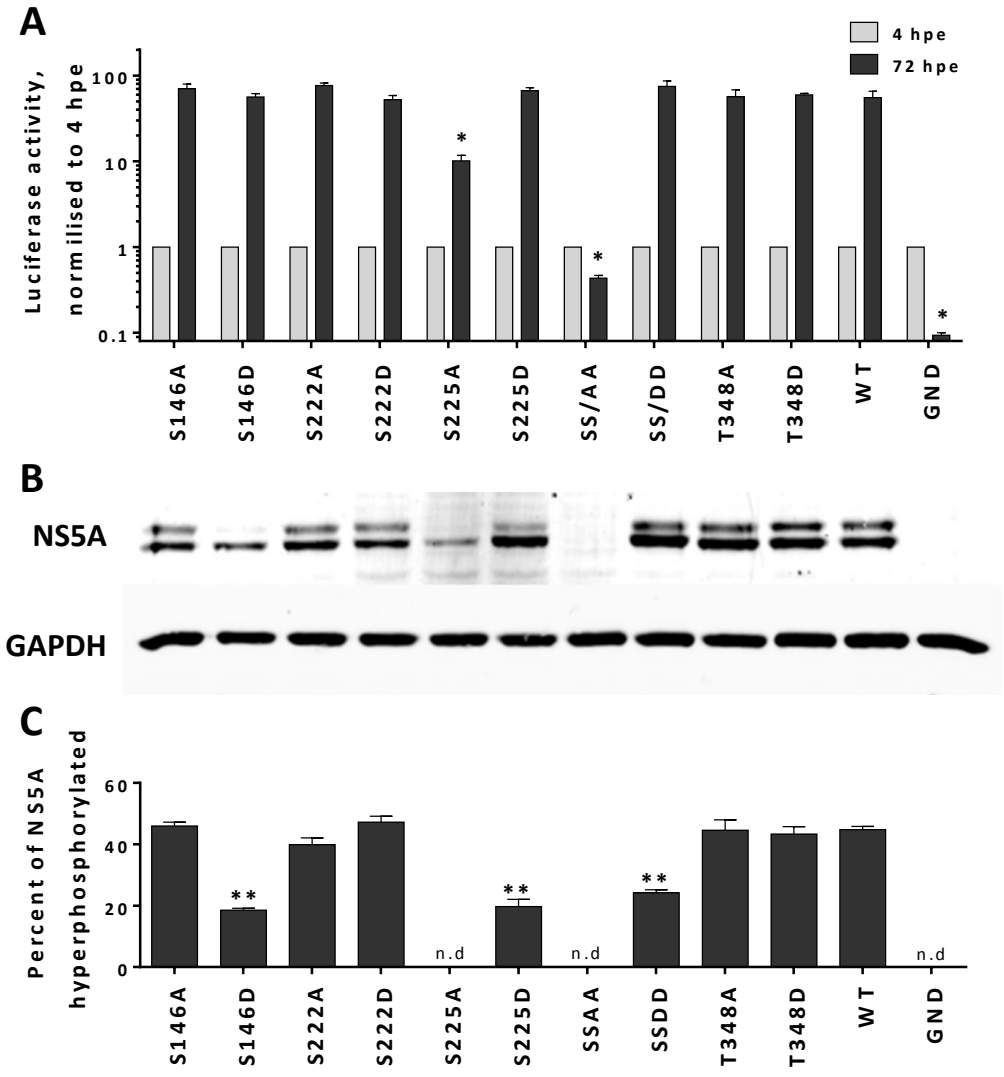


Figure 3.4. Mutagenic analysis of phosphorylation in the context of the HCV subgenomic replicon.

In vitro transcribed RNA was electroporated into Huh7 cells and luciferase activity determined at 4 and 72 hpe (Section 2.8.1). **A.** Luciferase activity at 72 hpe is shown (n=3) normalised to activity at 4 hours. * $p<0.05$ significance from wildtype. **B.** Western blot analysis of cell lysates at 72 hpe probed with NS5A (9E10) or GAPDH (mouse) antibodies. Use of 7.5% SDS-PAGE allowed the separation of the basally (lower band) and hyperphosphorylated (upper band) forms of NS5A. **C.** Quantification of the percentage of total NS5A that is hyperphosphorylated from western blot. Western blots were imaged by fluorescence using a LiCor Odyssey 5a, enabling highly accurate quantification (n=3). ** $p<0.01$ significance from wildtype. n.d. – no data due to low levels of expression.

The effect of phospho-mutations at residues 146, 222, 225 and 348 on the replication capacity of the NS3-5B replicon is shown in Figure 3.4A. Neither the phosphoablatant nor phosphomimetic mutations at serines 146 and 222 (S146A/D, S222A/D) or threonine 348 (T348A/D) had any significant effect on RNA replication; all of the resulting subgenomic replicons were able to replicate with equal efficiency (~100 fold increase in luciferase expression over 72 hpe). In contrast the S225A phosphoablatant mutation resulted in a one-log reduction in replication, whereas S225D had no phenotype, suggesting that phosphorylation of this residue is important for HCV genome replication. Consistent with this, the double alanine substitution (S222/225A) almost completely abolished the replicative ability of the subgenomic replicon, with levels of luciferase activity at 72 hpe only slightly above the GND negative control (an inactivating mutation in the NS5B polymerase). Again, a double aspartic acid substitution (S222/225D) had no phenotype. Thus it appears that S222 phosphorylation does play a subtle role in genome replication but that this is only manifest in the context of a further phosphorylation event at S225.

When analysed by western blot wildtype, NS5A appears as a doublet of what have been referred to as basally and hyperphosphorylated forms. To assess the potential effects of these mutations on NS5A expression and the relative abundance of these two forms, lysates from cells electroporated with subgenomic replicon RNAs were analysed by western blotting (Figure 3.4B). In most cases there was no difference in the expression profile of NS5A compared to wildtype. Where possible, the percentage of hyperphosphorylated NS5A was determined by quantification of western blots (Figure 3.4C), although the low levels of NS5A expression in the poorly replicating mutants S225A and S222/225A did not allow quantitation. Note that the use of a fluorescent based western blot greatly improved the accuracy of quantification over traditional chemiluminescence techniques. Surprisingly, S225D and S222/225D exhibited a reduction in hyperphosphorylation, although these residues (along with S229 and S232) have previously been proposed as phosphoacceptor sites in the hyperphosphorylated form of NS5A. This may indicate that an aspartate residue may mimic the charge, but not the structural effects, of phosphorylation. However, the most intriguing result was the reduction in hyperphosphorylation exhibited by the S146D mutation. As S146A still exhibits wildtype levels of the hyperphosphorylated species, this suggests that S146 is not a phosphoacceptor within this form, but that phosphorylation of this residue negatively regulates hyperphosphorylation.

To investigate whether this series of mutations elicited an effect at later stages of the virus lifecycle like virus production the series of phospho-mutations were inserted into the mJFH-1 virus. Huh7 cells were electroporated with *in vitro* transcribed RNA and cultured for 144 hrs, including a 1:5 passage at 72 hpe (Section 2.9.1). At 72 and 144 hpe supernatants were removed and virus titre determined by focus forming assay (Section 2.9.2) and total RNA was extracted and intracellular HCV genomes quantified by qRT-PCR using primers targeted to the 5'UTR (Section 2.9.3).

A similar set of data was obtained when this series of mutations was evaluated in the context of the full length JFH-1 infectious clone (Figure 3.5). The majority of the mutants showed wildtype levels of both intracellular genome abundance measured by qRT-PCR (Figure 3.5A) and released virus titre measured by focus forming assay (Figure 3.5B). Again S225A exhibited an approximately one-log reduction in both assays and S222/225A was almost completely deficient in virus replication and assembly. The western blot analysis (Figure 3.5C) also closely mirrored that observed for the subgenomic replicons, although in this case a low level of hyperphosphorylation could be seen in the context of the S146D mutant, this was clearly reduced compared to both wildtype and S146A.

It was therefore concluded that that the phenotype of these mutants is entirely manifest at the level of virus genome replication and there are no additional effects on virion assembly or release. The reductions in virus titre that are observed correlate with a reduction in replication, this corresponds to observations made previously that replication is the rate limiting step with respect to virus production (Binder et al., 2013).

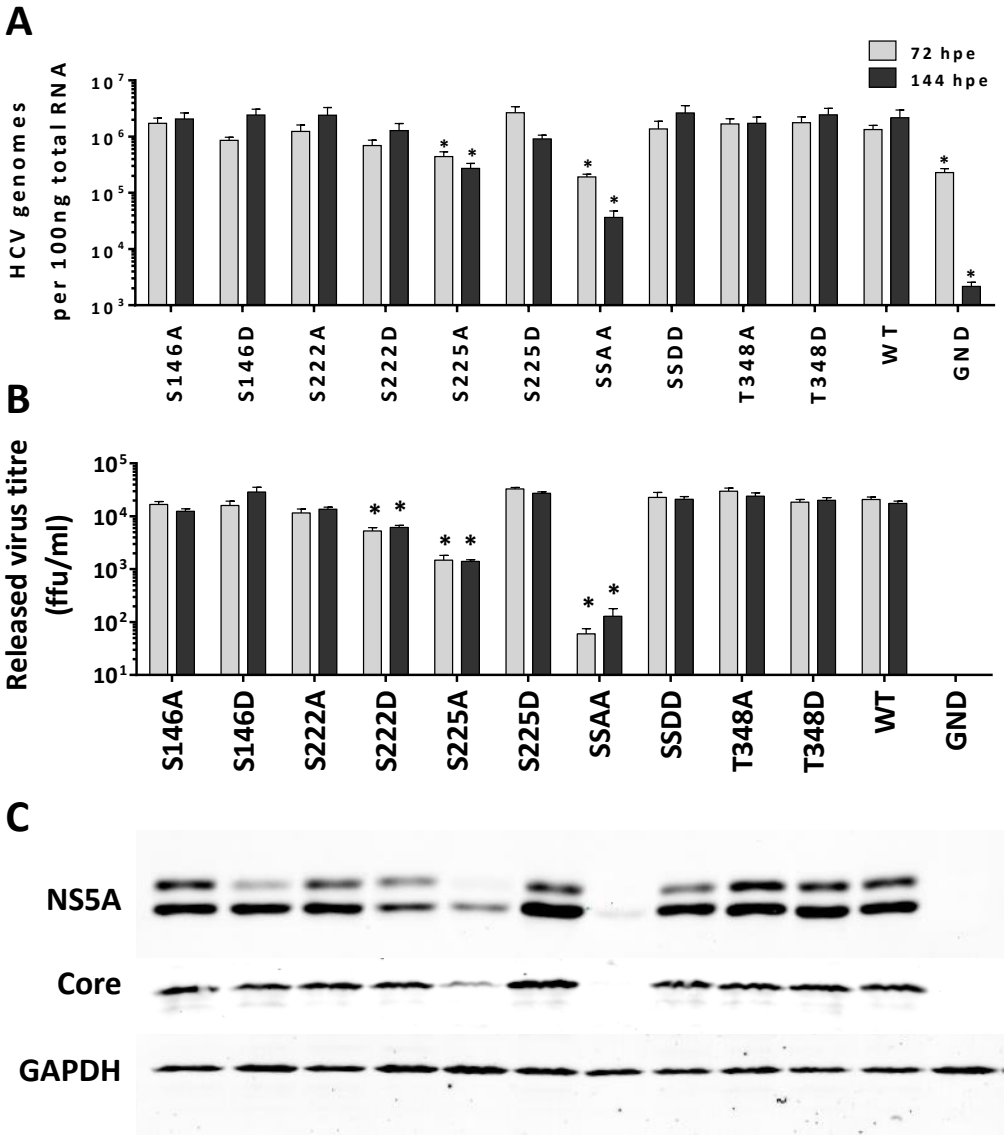


Figure 3.5. Mutagenic analysis of phosphorylation in the context of JFH-1 virus.
In vitro transcribed RNA was electroporated into Huh7 cells. Cells were cultured for 144 hrs, including a 1:5 passage at 72 hpe. **A.** Total RNA was extracted from cells at 72 and 144 hpe and HCV genomes quantified by qRT-PCR using primers targeted to the 5'UTR (Section 2.9.3) (n=3). * $p<0.05$ from wildtype. HCV genomes for the non-replicating control, GND, were still detectable, most likely due to the stability of the 5'UTR. **B.** Supernatants were assayed for released infectious virus titre by focus forming assay at 72 and 144 hpe. **C.** Lysates were prepared from cells at 144 hpe and probed for NS5A (9E10), Core (sheep) and GAPDH (mouse).

3.2.3 The effect of S225A on NS5A localisation

Whilst determining the titre of the released virus of S225A it was noted that the distribution of NS5A-S225A in infected cells was dramatically different from that of WT. To investigate this in greater detail, cells harbouring the denoted replicons were grown on coverslips and at 72 hpe were fixed and immunostained with α-NS5A (sheep) and the lipid droplet associating fluorophore BODIPY-C12. When imaged by confocal microscopy the re-distribution of NS5A-S225A into large, bright puncta was clear (Figure 3.6). While the mutation of the phosphorylation site serine S225 did cause a 1 log drop in replication and correspondingly virus titre, it was still capable of undergoing the complete virus lifecycle. Indicating that while the subcellular distribution of NS5A is dramatically altered, the protein can still function.

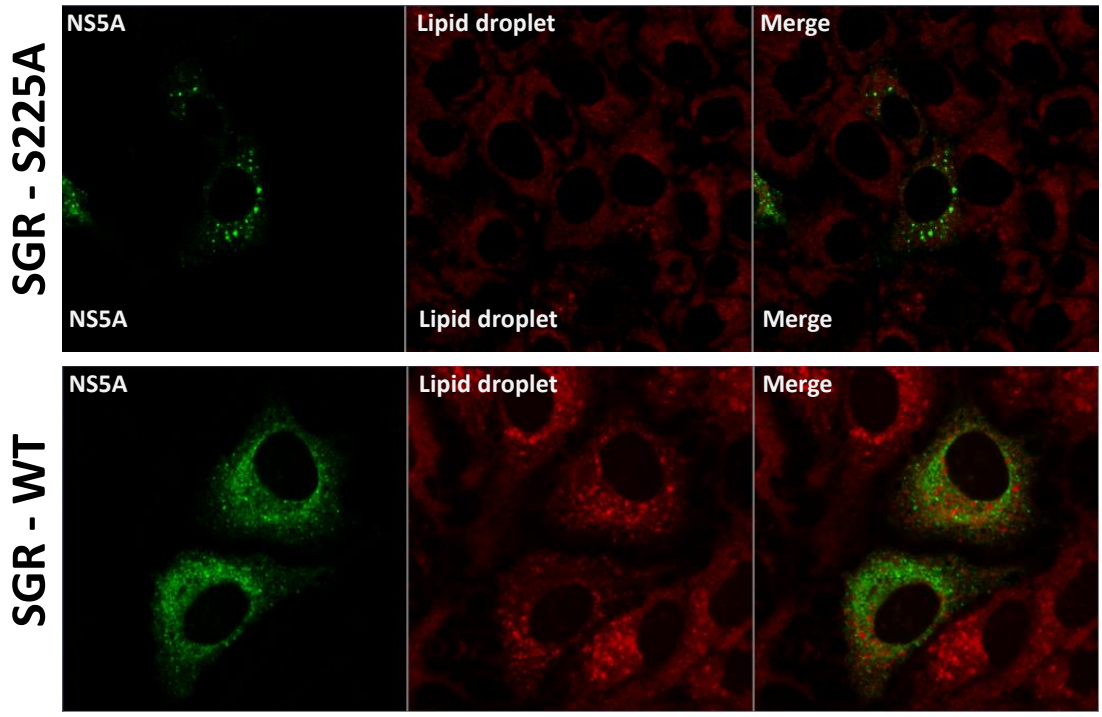


Figure 3.6. Differences in the subcellular localisation of S225A
Huh 7 cells were electroporated with SGR-WT or SGR-S225A, seeded onto coverslips and at 48 hpe fixed and immunostained for α-NS5A (sheep) in green and lipid droplets (BODIPY-555-C12) in red.

It was decided to pursue the different sub-cellular localisation of NS5A-S225A in the context of fluorescently labelled full length virus and there was the recent capability to undertake live-cell confocal imaging in a biosafety level 3 (BSL3) environment.

Two different approaches were utilised, firstly the S225A/D mutations were introduced into the JcI virus where the NS5A protein had been previously tagged with emerald GFP (emGFP) and RFP (Schaller et al., 2007). The insertion of the emGFP and RFP had been previously shown to have no effect on replication and only a modest reduction in virus titre. Secondly a novel NS5A tag system was developed for application to both live- and fixed-cell imaging and is the focus of Chapter 4. While the insertion of the S225A/D mutation into the JcI-emGFP/RFP virus was successful, due to time constraints on the project, it was not possible to generate any data with the system. As such, only anecdotal observations were made and have not been included. What can be noted was that the JcI-emGFP/RFP viruses were still viable after the introduction of S225A/D mutations and were still able to produce infectious virus. Furthermore, the sub-cellular distribution phenotype of S225A was also observed in this system.

3.2.4 Threonine 348 and the binding of SH3 domains.

The analysis identified threonine 348 within LCS II as a phosphoacceptor, but neither a phosphoablatant nor a phosphomimetic mutation at this site had any phenotype in replication and/or virus production. However, it was intriguing that this residue was located in the middle of a well characterised class II polyproline motif PP2.2 (PxxPxR) that binds to a range of SH3 domains, discussed previously (Section 1.6.1). It was therefore thought that the phosphorylation of threonine 348 might be involved in regulating the binding of SH3 domains to the PP2.2 motif (Figure 3.7A). Interestingly previous studies have shown that mutation of this motif to alanine (PA2.2) blocks SH3 binding but had no effect on virus replication and only minimal effects on virus release (Hughes et al., 2009a). It was proposed by Hughes *et al.* that the PP2.2 motif is only required for virus persistence, if this is the case it could explain why mutations of the phosphorylation site T348 also have no effect on virus replication and release.

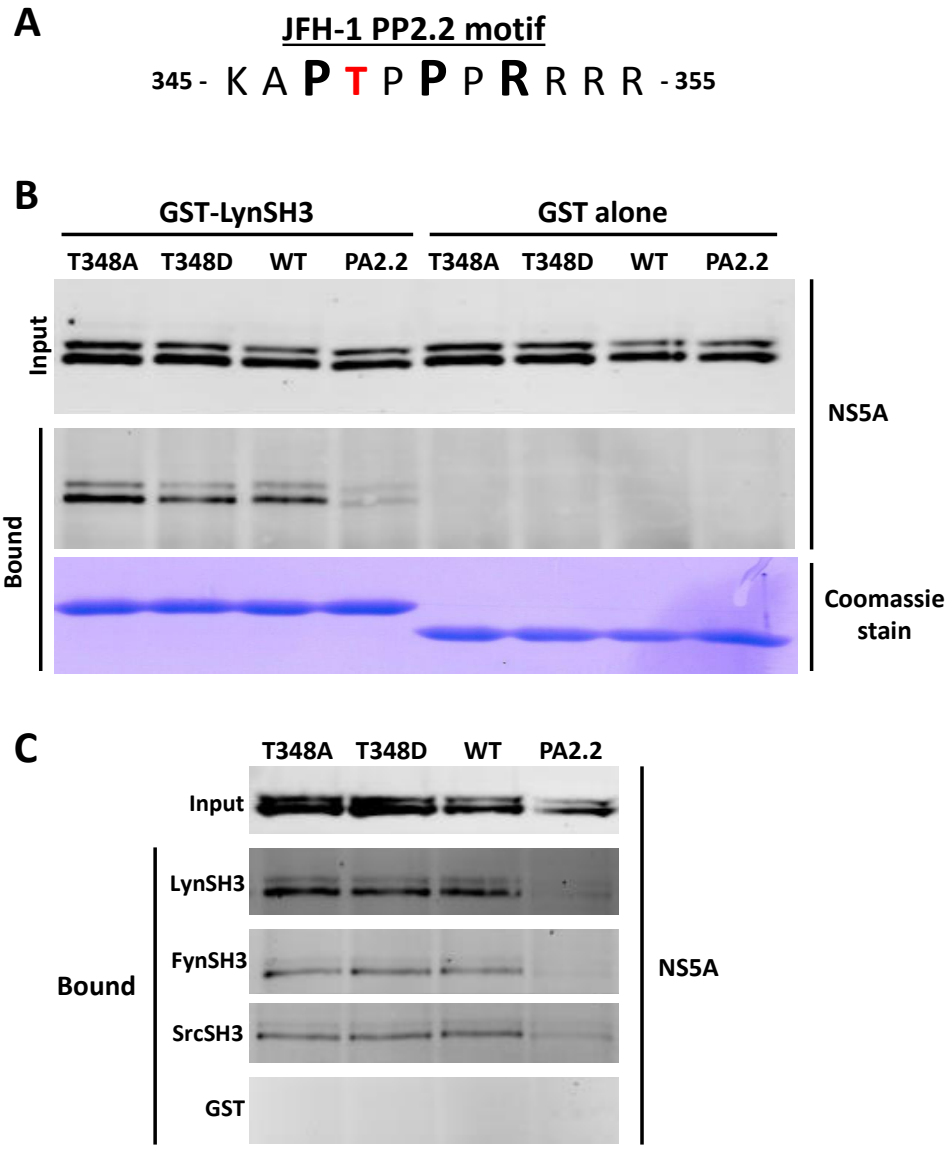


Figure 3.7. The effect of T348 phosphorylation on SH3 binding.

A. Sequence of the phosphopeptide from LCS II showing the conserved prolines (bold) and the phosphorylated threonine residue (red). **B.** Lysates from Huh7 cells electroporated with the indicated subgenomic replicons were subjected to GST pull-down using either GST-LynSH3 (lanes 1-4) or GST alone (lanes 5-8). 5% of the input lysate and the total bound fraction were analysed by western blot with an anti-NS5A antiserum (sheep). Equivalent loading of the GST fusions was verified by Coomassie staining (lower panel). PA2: mutation of the two prolines highlighted in **A** to alanine (Hughes 2009b). **C.** As for **B.** but lysates were subjected to GST pull-down using GST-LynSH3, FynSH3 or SrcSH3.

To investigate whether the phosphorylation of threonine 348 was affecting SH3 binding a glutathione pull down of GST-SH3 domains with infected cell lysates was carried out (Section 2.5.6). In brief, GST tagged SH3 domains were expressed and purified in bacteria as described and used to decorate glutathione-Sepharose resin. The decorated resin was incubated with lysates from cells infected with JFH-1 WT or harbouring T348A/D mutations. The GST-SH3 pull-down was then analysed by SDS-PAGE and western blot for α-NS5A (9E10). As a negative control for the interaction between the PP2.2 motif and SH3 domains the NS5A mutation PA2.2 was used.

Both the T348A and T348D mutants bound with equal efficiency as wildtype to the GST-LynSH3 domain fusion protein, while the PA2.2 mutant showed a dramatic impairment, Figure 3.7B. The Coomassie stain of the GST-LynSH3, and the western blot of NS5A input, showed that the same amount of bait and prey were present in each pull down (Figure 3.7B). Two other SH3 domains, Fyn and Src, were also investigated to see if the regulation of SH3 binding to the PP2.2 motif by phosphorylation at T348 was substrate specific, but in case both T348A and T348D bound to the SH3 domains with comparable efficiency, Figure 3.7C.

It is still feasible that the phosphorylation of threonine 348 is involved with regulating the binding of different proteins to the PP2.2 motif; it is also possible that the phosphomimetic aspartic acid is insufficient to recapitulate the effect of a phosphate group. Alternatively it is possible that this phosphorylation site regulates the activity of the putative nuclear localisation single (NLS) that is one residue downstream of threonine 348 (Ide et al., 1996), while the localisation of NS5A to the nucleus is not been convincingly demonstrated the recruitment of nuclear pore complex proteins into the membranous web has been (Neufeldt et al., 2012). It is therefore a possibility that this NLS at the start of domain III is important in trafficking NS5A out of the tightly enclosed membranous web, and that the negative charge of the phosphorylation at threonine 348 could regulate this by masking the positive charge of the NLS. Furthermore the four arginines (residues 353-355) at the centre of the putative NLS have previously been shown not to be important for replication in either ConI (Tellinghuisen et al., 2008b) or JFH-1 (data not show).

3.2.5 Analysis of phosphorylation within LCS I

In addition to the mono- and di-phosphorylated peptides mapping to LCS I (S222 and S222/225), the mass spectrometric analysis revealed evidence for the existence of multiply phosphorylated species containing between three and seven phosphorylation events (Table

3.1). In order to investigate the potential function of these phosphorylation events a further set of phosphoablatant and phosphomimetic mutants were generated and their phenotypes were evaluated in the context of the SGR-luc-JFH-1 subgenomic replicon as described previously (Section 3.2.2).

As illustrated in Figure 3.8A, S228A/D and S230A/D mutations appeared to have no phenotype; however, S229 was a key residue as both S229A and S229D mutations completely abrogated RNA replication. Consistent with this a triple mutation of S228/229/230 to either alanine or aspartic acid was also completely defective. S232 resembled S225 in that the S232A mutation reduced replication by approximately one-log, whereas S232D exhibited wildtype activity. These data points to important roles for S229 and S232 in HCV genome replication, at this stage it is pertinent to note that these residues were identified as phosphoacceptors in the tri- and tetra-phosphorylated peptides, but they were present in combination with other phosphorylated residues in multiple isomers that could not be assigned.

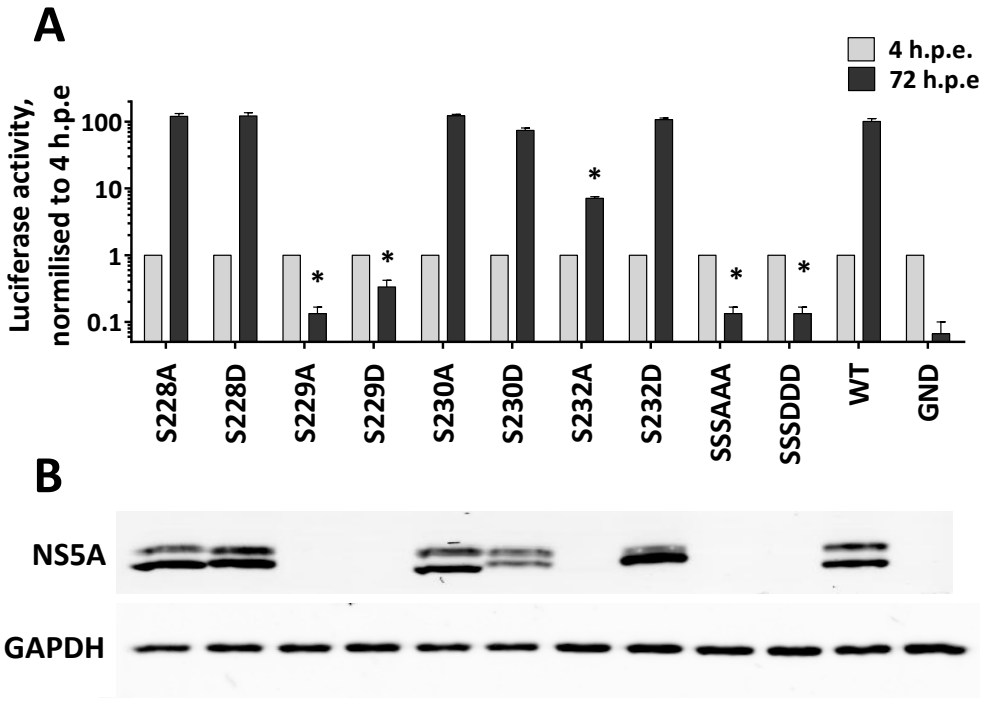


Figure 3.8. Mutagenic analysis of putative phosphorylation sites within LCS I.
In vitro transcribed RNA was electroporated into Huh7 cells and luciferase activity measured at 4 and 72 hpe. **A.** Luciferase activity at 72 hpe is shown (n=3) normalised to activity at 4 hours. * $p<0.05$ from wildtype. **B.** Western blot analysis of cell lysates at 72 hpe probed with α-NS5A (9E10) or α-GAPDH (mouse) antibodies.

Western blot analysis of NS5A expression in cells harbouring the subgenomic replicons reflected the levels of replication (Figure 3.8B), however this analysis led to an intriguing observation regarding the mobility of the two phosphorylated forms. Although the mobility of the hyperphosphorylated form did not vary, there were differences in the mobility of the basally phosphorylated forms between the various mutants. Specifically, although alanine substitutions did not change the mobility of the basally phosphorylated form compared to wildtype, aspartic acid substitutions resulted in a subtle retardation of its mobility. This could also be seen to a lesser extent for mutations at S222, S225 and the combination of S222/225, but interestingly not for S146 or T348 (Figure 3.4B), suggesting that this was not simply due to the substitution of an acidic residue into the protein.

The differences in the migration of the basal-phosphorylation species as a result of phosphomimetic mutations in the LCS I was investigated further. For completeness the two remaining serines within this region, serine 235 and 238, were also mutated to aspartic acid. When lysates from cells harbouring these phosphomimetic mutations were analysed by SDS-PAGE a clear trend emerged. As the position of the aspartic acid mutation progressed towards the C-terminus of this serine cluster there was a concomitant retardation of mobility, Figure 3.9.

This shift in the apparent molecular weight of the basal-phosphorylation species was clearly not a result of the aspartic acid mutation, as the mutation of serine 146, 222 or threonine 348 to aspartic acid does not elicit the same effect (Figure 3.4A). Due to the serine-rich and extensively phosphorylated nature of the LCS I, this change in molecular weight is most likely the result of differing degrees of phosphorylation. The presence of these phosphomimetic mutations could therefore be influencing different patterns of phosphorylation within this region. One explanation for this result is that phosphorylation of one residue completes the kinase recognition motif for a neighbouring residue to be phosphorylated, and so on, resulting in a phosphorylation cascade. This is discussed in further detail (Section 3.3.2)

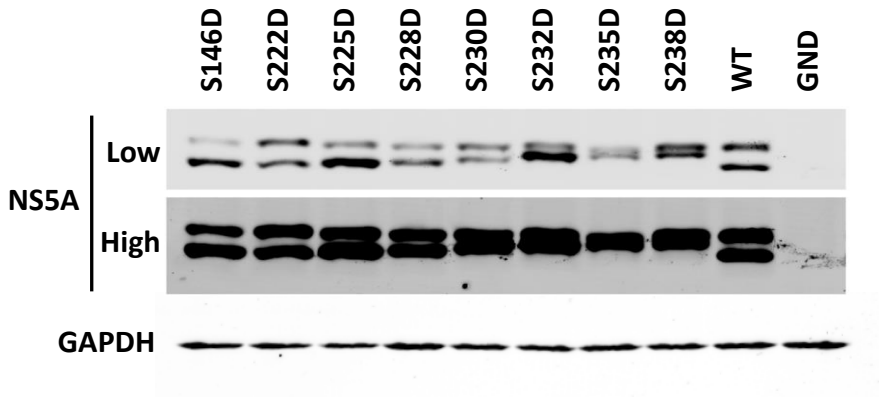


Figure 3.9. Phosphomimetic mutations in LCS I increase the apparent molecular weight of the basally phosphorylated NS5A form.

Lysates from electroporated cells were prepared at 72 hpe and analysed by western blotting with antibodies to α-NS5A (9E10) or α-GAPDH (mouse). Lysates were separated by 7.5% SDS-PAGE at low voltage for an extended time (110V, 120 min) allowing for optimal resolution of basally and hyperphosphorylated forms. The NS5A western blot is shown at both low and high exposures.

3.2.6 Regulation of hyperphosphorylation by S146.

As shown in Figure 3.4B, 3.5C and 3.9, the introduction of a phosphomimetic at serine 146 resulted in a reduction of the level of hyperphosphorylation, suggesting that phosphorylation of this residue in some way negatively regulated phosphorylation at other sites (presumably within the serine rich LCS I). It was reasoned that if the sequential phosphorylation with LCS I contributed to hyperphosphorylation, then the introduction of mutation S146D, which reduces hyperphosphorylation, might influence this process. To investigate this effect the mutation S146D was introduced into the panel of LCS I phosphomimetic mutations discussed previously, Figure 3.9. This generated a panel of mSGR-luc constructs where phosphomimetic mutations of residues 222-238 were present either with or without the S146D mutation.

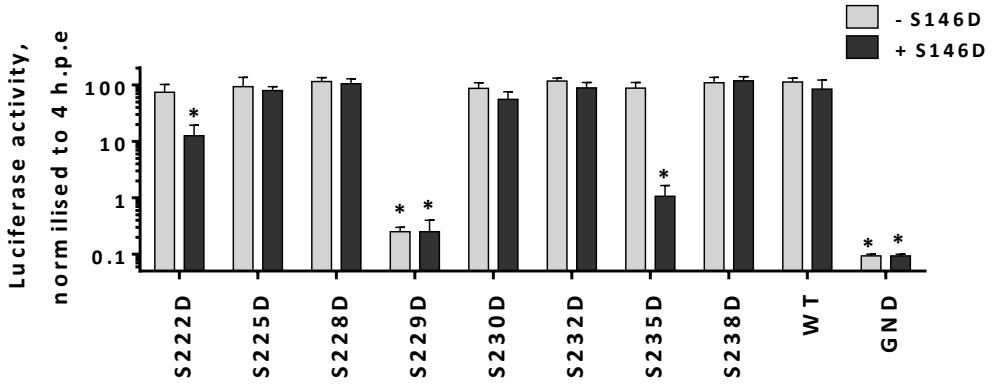


Figure 3.10. Analysis of LCS I phosphomimetic mutations in the context of S146D.
In vitro transcribed RNA was electroporated into Huh7 cells and luciferase activity measured at 4 and 72 hpe. A polymerase inactive mutant, GND, was included as a negative control. Luciferase activity at 72 hpe is shown (n=3) normalised to activity at 4 hours. For each mutant the results obtained in either a wildtype (grey bars) or S146D mutant (black bars) background are shown side by side. * $p<0.05$ from wildtype.

This panel of phosphomutants was first investigated for the effect on replication by electroporating *in vitro* transcripts into Huh7 cells and determining luciferase activity at 72 hpe. Figure 3.10 demonstrates that for S225D, S228D, S230D, S232D and S238D the addition of the S146D mutation had no effect on replication, as all of these subgenomic replicons were indistinguishable from wildtype. As previously shown (Figure 3.8), S229D was completely replication defective and was not rescued by the S146D mutation. Interestingly two mutations, S222D and S235D, showed an altered phenotype, whereby in the wildtype background these mutations had no effect on replication, but in the context of S146D they both exhibited impaired or ablation of replicative capacity, S222D and S235D respectively.

Western blot analysis of cell lysates showed two distinct features. Firstly, the introduction of S146D alongside LCS I mutations (S228D, S230D, S232D and S238D) did not block their ability to increase the M.Wt of basally phosphorylated NS5A (Figure 3.11A). Secondly, the presence of any S->D mutation in LCS I (222-238) did not block the ability of S146D to reduce hyperphosphorylation of NS5A (Figure 3.11B). These data shows that phenotype of both S146 and LCS I phosphomimetic mutations cannot elicit a dominant effect over the other.

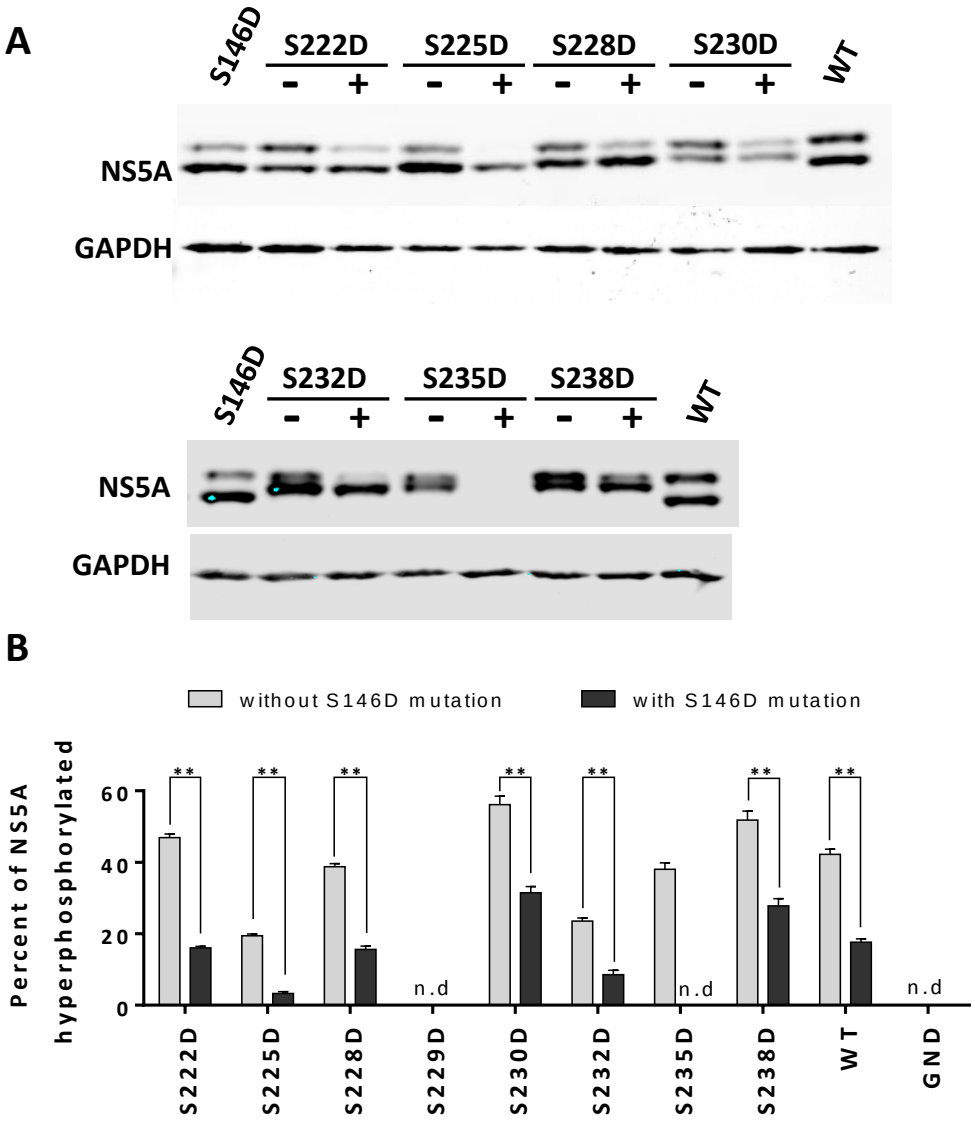


Figure 3.11. Effect of S146D on LCS I phosphorylation.

A. Lysates from electroporated cells were prepared at 72 hpe and analysed by western blotting with antibodies to α-NS5A (9E10) or α-GAPDH (mouse). The presence or absence of the S146D mutation is indicated by +/- signs. **B.** Quantification of the percentage of total NS5A that is hyperphosphorylated from western blot. Western blots were imaged by fluorescence using a LiCor Odyssey 5a, enabling highly accurate quantification (n=3). ** p <0.01 . n.d. – no data due to low levels of expression.

It has been discussed previously that the NS5A inhibitor Daclatasvir is able to reduce the hyperphosphorylation of NS5A within 4 hrs of treatment, while the block on replication does not occur for 24 hrs (Section 1.6.4.2). This effect on NS5A phosphorylation is not thought to be the mechanism of action of Daclatasvir. However, it has been shown that in patients

treated with Daclatasvir there is a rapid drop in the number of circulating virus particles (6 - 9 hrs), long before inhibition of replication (Guedj et al., 2013). With the phosphorylation of NS5A known to be important in virus production, could the effect of Daclatasvir on NS5A phosphorylation be the mechanism behind this rapid drop in virus production seen *in vivo*?

To investigate this further the effective concentration, 50% (EC50) of Daclatasvir on replication was determined for replicon constructs containing the phosphomutations in the LCS I, with and without the S146D mutation. Huh7 cells were electroporated with mSGR-luc *in vitro* transcripts and at 4 hpe cells were treated with Daclatasvir at a range of concentrations (1 fM to 10 µM), for details see (Section 2.8.2). At 48 hpe luciferase activity was determined and used to calculate EC50s. As a positive control the Daclatasvir resistance mutation Y93H was used to demonstrate a shift in EC50. If phosphorylation of these serines was involved in the mechanism of action of Daclatasvir then a change in EC50 would be expected. However, as can be seen in Table 3.2, no combination of any phosphorylation mutation elicited a significant effect on the inhibitory concentration of Daclatasvir.

LCSI mutation	With or without mutation S146D (+/-)	EC50, pM
S222D	-	27.6
	+	n/a
S225D	-	32.2
	+	9.8
S228D	-	28.1
	-	10.9
S230D	+	16.8
	-	7.4
S232D	+	31.3
	-	11.5
S235D	-	21.7
	+	17.0
S238D	-	33.7
	+	12.7
WT		31.9
S146D		18.5
Y93H		47 030.0

Table 3.2. The effect of LCSI phosphomutants in combination with S146D on the inhibitory concentration of Daclatasvir.
In vitro transcripts were electroporated into Huh7 cells and at 4 hpe were treated with Daclatasvir (1 fM to 10 µM). Luciferase was determined at 48 hpe and EC50 concentrations by analysis in Prism (Graphpad).

3.2.7 S222 phosphorylation is a hallmark of the hyperphosphorylated NS5A species.

It has been previously reported that S222 is a critical phosphoacceptor in the formation of hyperphosphorylated NS5A, although the evidence for this is indirect and it was possible that pS222 was present in both phosphorylated species. In order to address this question a sheep polyclonal phosphospecific antiserum was raised against phosphorylate S222 using an appropriate phosphopeptide. Figure 3.12A presents a validation of this reagent and shows that the α-pS222 antiserum did indeed only react with the hyperphosphorylated NS5A form. A very faint band could be seen corresponding to basally phosphorylated NS5A, however this is likely due to recognition of the surrounding peptide sequence. Reassuringly the serum showed only very limited cross-reactivity with the S222A or S222D mutants, again the faint bands are likely due to recognition of the surrounding sequence.

Furthermore, western blot analysis using this reagent revealed that, as expected, pS222 reactivity was much reduced in the S146D mutant (Figure 3.12B) and that the pS222 antibody was more reactive towards the basal phosphorylated band in mutations S225D, S228D and S232D (Figure 11C). This is the first biochemical evidence that S222 phosphorylation is enriched in the hyperphosphorylated NS5A form and that the S146 is involved with regulating this phosphorylation. It cannot however be rule out that pS222 is also present at very low levels in the basally phosphorylated form.

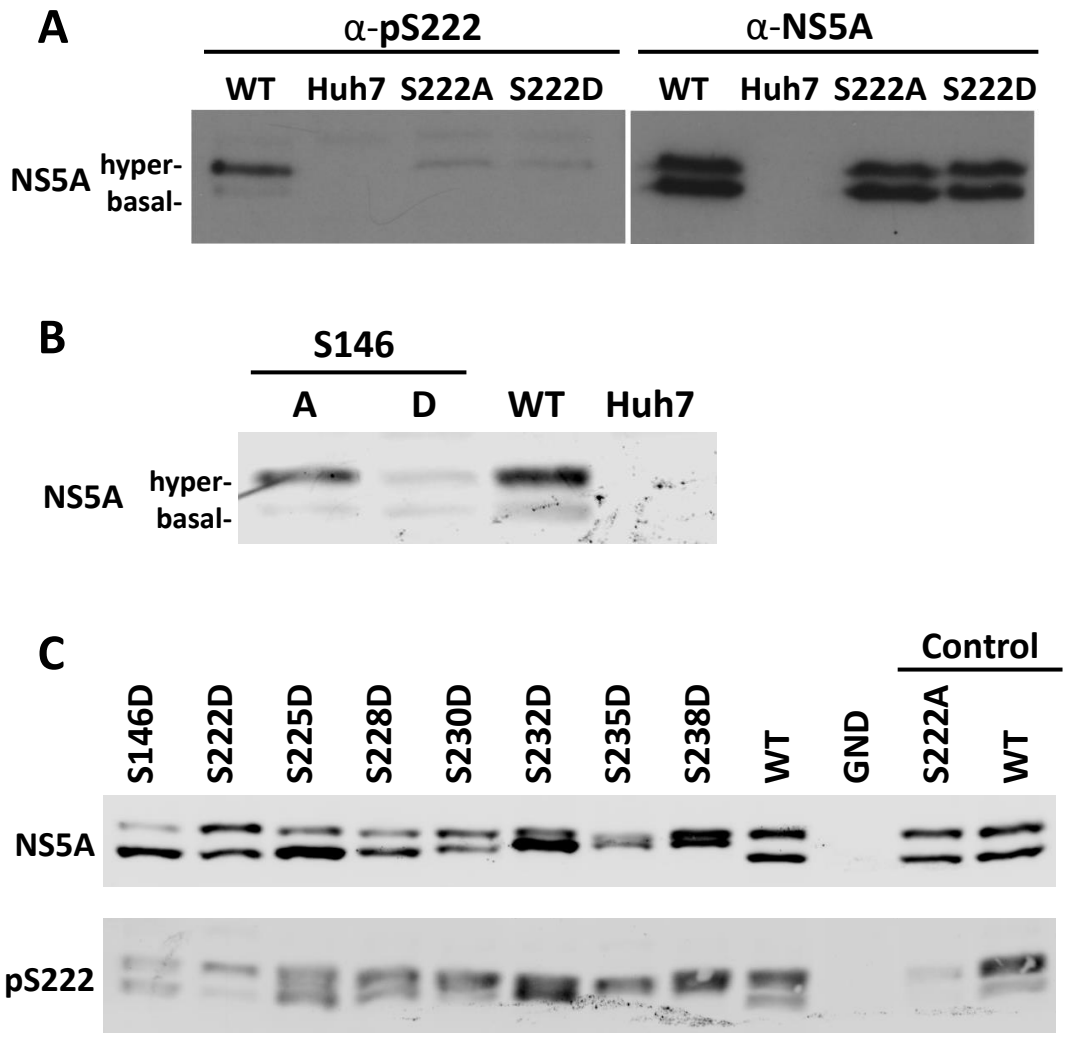


Figure 3.12. Phosphorylated S222 is predominately in the hyperphosphorylated form. A phosphospecific antibody was raised in sheep against pS222 using an appropriate 13mer peptide and purified by IgG enrichment. **A.** Specificity of the pS222 antibody. Lysates from Huh7 cells electroporated with the indicated subgenomic replicons were analysed by SDS-PAGE/western blot and probed with either α-pS222 or α-NS5A (sheep). **B.** and **C.** lysates processed as in **A.**

3.3 Discussion

This investigation has attempted to make significant progress towards a comprehensive analysis of the phosphorylation of NS5A by analysing protein purified from a physiologically relevant context, i.e. from Huh7.5 cells harbouring a JFH-1 derived subgenomic replicon. With the exception of the recent publication (Lemay et al., 2013), this study therefore represents the only biochemical analysis of NS5A phosphorylation expressed in human cells in the context of the other non-structural proteins and an active RNA replication complex.

Both studies identified S222 within LCS I as a major site of phosphorylation (Lemay et al., 2013). However, whereas LeMay *et al.* only identified S222 as a phosphoacceptor, this study identified a large number of other sites of phosphorylation, not only in LCS I but also in domains I, III and LCS II (Table 3.1). More phosphorylation sites were identified than previous due simply to the amount of starting material – LeMay *et al.* used 7.5 x 10^7 cells, this study used approximately 100-fold more. This allowed for the detection of phosphopeptides that are detected with a lower efficiency by MS (peptides that are large and/or highly phosphorylated) as well as those that are physiologically less abundant.

NS5A contains no trypsin cleavage sites within the C-terminal 121 residues that form domain III. This has hampered previous phospho-mapping attempts because large peptide fragments are poorly detected by MS analysis. To overcome the absence of trypsin cleavage sites in domain III the purified protein was also digested with Glu-C endoproteinase. This protease cleaves C-terminal to glutamic residues and results in a further 11 cleavage sites throughout domain III, resulting in peptide fragments that are significantly more amenable to MS analysis. Despite this approach the putative phosphorylation sites in the C-terminal serine cluster identified by Masaki *et al.* and Tellinghuisen *et al.* were not observed (Figure1.11). This could be a reflection of poor protease activity in the locality of phosphorylation sites. For example, the putative phosphorylation site serine 457 is adjacent to a Glu-C cleavage site (-SEED-), if phosphorylation of this serine were to block Glu-C digestion then the peptide fragment containing pS457 would remain 41 aa in length and would not be detectable by MS. Similarly if these phosphorylated peptides were in low abundance and extensively phosphorylated, this too would limit their detection by MS. Therefore the data generated here does not preclude the presence of phosphorylation sites within this region. Neither this study nor LeMay *et al.* were able to discriminate between phosphoacceptors present on the basal and/or hyperphosphorylated NS5A forms, due to the loss of resolution when separating large amounts of protein by SDS-PAGE.

3.3.1 LCS I is the predominant location for hyperphosphorylation of NS5A

The majority of phosphoacceptor residues identified in this study were in LCS I, between serines 222 and 238. Remarkably, this serine rich cluster is absolutely conserved across all genotypes of HCV highlighting the importance of this region in the virus lifecycle. A 19 aa phosphopeptide corresponding to this region was shown to be extensively phosphorylated, containing combination of between 1 and 7 phosphorylated residues, Figure 3.2C and Table 3.1. While the assignment of phosphorylation at serine 222 and the double serine 222/225

was possible, the presence of different isomers precluded the definitive assignment of the tri, tetra-, penta-, hexa- and hepta-phosphorylated species. It was however possible to partially assign serines 229, 230, 232 and 235 as amongst the phosphorylated resides in these species, Appendix figure 8.5-8.7. Note that the phosphorylation of serines 228 and 238 is not excluded on the bases they were not observed in this analysis.

By generating an antiserum specific to phospho-S222 it was possible to provide the first definitive evidence that phosphorylation of S222 is indeed a hallmark of hyperphosphorylated NS5A. The antiserum showed very little reactivity towards the basally phosphorylated NS5A species, or towards NS5A containing the S222A or S222D mutants (Figure 3.12A). However, despite the fact that S222 phosphorylation was only detected in the hyperphosphorylated species, in this study mutation at this residue did not affect the ratio of hyper- to basally phosphorylated NS5A (Figure 3.4). This suggests that phosphorylation of S222 alone is neither the sole component of, nor sufficient to produce the hyperphosphorylated form. This is consistent with previous studies that have implicated several serines in the LCS I as required for hyperphosphorylation (Tanji et al., 1995). LeMay *et al.* have also provided evidence for phosphorylation of S222 being present in the hyperphosphorylated NS5A form, shown by a 36% reduction in hyperphosphorylation in the context of the S222A mutant.

To complement the biochemical analysis of phosphorylation an extensive mutagenesis of the identified phosphoacceptor residues within the LCS I was conducted. Many of these mutants had no phenotype but this analysis pointed to a critical role for S229 – both phosphoablatant and phosphomimetic mutations at this site abrogated replication of the subgenomic replicon, in agreement with the recent data of (Fridell et al., 2013).

Interesting the *in vitro* phosphorylation by CKI of a short synthetic peptide equivalent of the LCS I required the presence of a priming phosphorylation event at position S229. However, neither aspartic acid nor glutamic acid when introduced at this position could recapitulate this effect (Quintavalle et al., 2007). This suggests that in this context, aspartic acid does not successfully act as a phosphomimetic, potentially explaining why both alanine and aspartic acid mutation at position S229 are lethal for replication. In contrast the mutation of phosphorylation sites at serine 225 and 222/225, which resulted in only a partial or complete defect in replication respectively, was successfully restored by the corresponding phosphomimetic mutations, consistent with a role for phosphorylation at these residues

(Figure 3.4). A drop in virus production was observed, but this correlated to the impairment to replication, in line with recent reports of replication being the rate limiting step in virus production (Binder et al., 2013).

Interestingly no mutations within the LCS I stimulated RNA replication, this was in contrast to previous reports in the genotype 1b isolate ConI (Appel et al., 2005). When a comparison is drawn between data generated here, the ConI data and the recently published data by Fridell *et al.* an interesting observation is made, Table 3.3 (Fridell et al., 2013). Where the introduction of phosphomimetic or phosphoablatant mutations results in a change in RNA replication, there is an inverse effect in the JFH-1 than in the ConI isolate. These differences are hard to interpret, but presumably are indicative of fundamental genotype-specific differences in the interactions between NS5A and other viral (or cellular) proteins during the process of virus genome replication. It is worth noting that the ConI isolate used by Appel *et al.* contained two culture adaptive mutations in the NS3 protein (E1202G and T1280I).

NS5A residue number	Mutation	Gt 1b (Appel *et al*)	Gt 2a (this study)	Gt 2a (Fridell *et al*)
S222 (2194)	Ablatant	Wildtype	Wildtype	Wildtype
	Mimic	Wildtype	Wildtype	Not tested
S225 (2197)	Ablatant	Up 6 fold	Down 6 fold	Down 3 fold
	Mimic	Wildtype	Wildtype	Wildtype
S228 (2200)	Ablatant	Wildtype	Wildtype	Wildtype
	Mimic	Wildtype	Wildtype	Not tested
S229 (2201)	Ablatant	Up > 10 fold	Down > 1000 fold	Lethal
	Mimic	Up 50 fold	Down > 1000 fold	Lethal
S230 (2202)	Ablatant	Up 5 fold	Wildtype	Wildtype
	Mimic	Wildtype	Wildtype	Not tested
S232 (2204)	Ablatant	Up > 10 fold	Down 10 fold	Down 10 fold
	Mimic	Wildtype	Wildtype	Wildtype
S235 (2207)	Ablatant	Up > 10 fold	Not tested	Lethal
	Mimic	Wildtype	Wildtype	Wildtype
S238 (2210)	Ablatant	Wildtype	Not tested	Wildtype
	Mimic	Wildtype	Wildtype	Not tested

Table 3.3. Comparison of the mutations in LCS I between genotypes.
genotype 1b data from the ConI NS3-5B replicon (Appel et al., 2005), genotype 2a data from the JFH-1 isolate, this study or (Fridell et al., 2013). NS5A numbering (S222-238) and H77 reference polyprotein numbering (2194-2210).

3.3.2 Evidence for sequential phosphorylation across LCS I.

During this analysis it was observed that the phosphomimetic mutations within the LCS I affected the mobility of the basally phosphorylated species, but not that of the hyperphosphorylated species. Moreover, other phosphomimetic mutations in different regions of NS5A did not elicit the same effect (S146D and T348D), suggesting this is specific to the basally phosphorylated species, and specific to the serine rich LCS I. A clear trend emerged whereby, as the position of the phosphomimetic mutation was shifted towards the C-terminus of LCS I there was a concomitant decrease in the mobility of the basally phosphorylated species, indicative of additional phosphorylations (Figure 3.9).

One interpretation of this data is that the serines with the LCS I region are part of a regulated sequential phosphorylation cascade that generates the hyperphosphorylated species. The introduction of phosphomimetic mutations into the LCS I triggered this cascade to occur unregulated on the basal-phosphorylation species, giving rise to an increase in apparent molecular weight. The extent of extra phosphorylation on the basal-phosphorylation species could then indicate at what point within the sequential phosphorylation cascade it was triggered, i.e. near the start or end. A rapid sequential phosphorylation cascade over the LCS I forming the hyperphosphorylated species would explain why intermediary phosphorylation states of NS5A are not observed by SDS-PAGE. Indeed Quintavalle *et al.* have already demonstrated that on a synthetic peptide mimetic of the LCS I a priming phosphorylation at S229 is required for the CKI directed phosphorylation of S232 (Quintavalle et al., 2007).

This model would be supported by observations that the ConI culture adaptive mutation S2204I (S232 in JFH-1) resulted in a complete loss of the hyperphosphorylated species and not a slight decrease in molecular weight. Hence why in the JFH-I isolate a phosphomimetic mutation at this position results in a significant increase in the apparent molecular weight of the basal-phosphorylation species.

However in the case of the phosphomimetic S238D, the basally phosphorylated species (where the largest increase in apparent molecular weight was observed) could still be distinguished from the hyperphosphorylated species (Figure 3.9). This suggests that the mechanism for generating hyperphosphorylation cannot be fully recapitulated solely by the introduction of single phosphomimetic mutations in the LCS I.

There are many several host proteins that have been shown to undergo these sequential phosphorylation cascades and they were discussed in detail previously (Section 1.5.3). Typically more than one kinase is involved in orchestrating these sequential phosphorylation cascades, as is seen with β-catenin where a phosphorylation by CKI primers the three subsequent phosphorylations by glycogen synthase kinase 3 (GSK-3). Also, these sequential phosphorylation cascades do not have to proceed in a straight linear fashion, as is the case in the PTEN protein where phosphorylation proceeds from S385 → S380 → T383 → T382 (Cordier et al., 2012). To investigate the details of this kind of sequential phosphorylation cascade would first require the identification of the phosphorylating kinases, this would then allow sophisticated temporal NMR analysis as to what order phosphorylation was occurring, such as in (Cordier et al., 2012).

3.3.3 S146 regulates hyperphosphorylation.

S146 is located within domain I and has not previously been identified as a phosphoacceptor. Mutation of this site had no effect on either genome replication or virus assembly, however, it was noteworthy that S146D exhibited a significant decrease in the abundance of hyperphosphorylated NS5A – decreasing from 42% (+/- 0.6 SEM) of total NS5A in wildtype, to 17% (+/- 1.0 SEM) in S146D. S146D also exhibited a loss in reactivity to the phospho-S222 antiserum, consistent with the fact that S222 phosphorylation is a hallmark of the hyperphosphorylated species (Figure 3.12B). Given that phosphomimetic mutations in LCS I appeared to at least partially drive the conversion of the basally phosphorylated species into the hyperphosphorylated form, it was investigated whether these mutations could override the inhibitory effect of S146D on hyperphosphorylation. However, this was not the case with the same shift in molecular weight observed for LCS I phosphomimetic mutations both in the presence and absence of S146D. Conversely no phosphomutation in the LCS I prevent the mutation S146D from reducing the amount of NS5A that was hyperphosphorylated. Taken together these data might be misinterpreted as suggesting that phosphorylation at serine 146 and the LCS I are mutually exclusive events. However, the mutation S146D in combination with either S222D or S235D resulted in an impairment or ablation to replication that was not observed when the mutations were individually introduced.

It is far from certain what complex interplay is occurring between serine 146 and the LCS I in the formation of hyperphosphorylation, but the data generated here suggest the following model. Firstly the sequential phosphorylation of the LCS I region drives the formation of the hyperphosphorylated species, and that elements of this are essential for virus replication.

Secondly that hyperphosphorylation is negatively regulated by phosphorylation at S146 (Figure 3.13). Further, based on the observation that concurrent mutation of S146D and S238D results in NS5A that still exhibits two species of distinct mobility, additional (as yet undefined) phosphorylation events are required to produce the hyperphosphorylated species of NS5A.

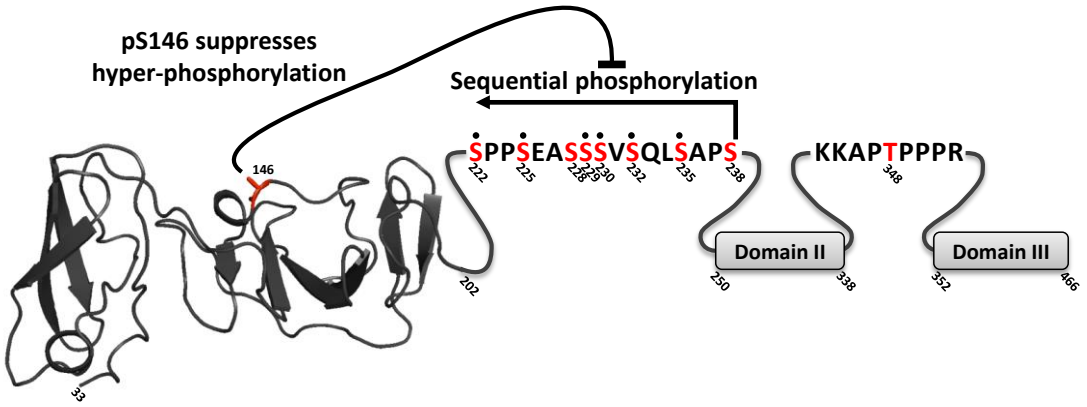


Figure 3.13. Summary NS5A phosphorylation identified.
Schematic of NS5A showing the location and relationship of phosphorylation events identified. Domain I denoted by the structure of the Con1 monomer (Tellinghuisen et al., 2005), sequences of LCS I and LCS II are shown, and domain II and III denoted by grey boxes. Highlighted in red are those residues identified as phosphorylation sites. Black dots over residues within the LCS I identify serines that were observed by MS as phosphorylated.

There is evidence of a similar mechanism occurring on the P protein of the Parainfluenza virus 5 (PIV5), discussed previously (Section 1.5.3). Here a phosphorylation of the P protein at serine 157 allows the binding of Polo-like kinase I via the C-terminal polo-box domain (PBD) and a subsequent phosphorylation by the N-terminal kinase domain at position serine 308 (Sun et al., 2009). Taken together with recent findings that PlkI can phosphorylate NS5A *in vitro* this hints at a potential mechanism by which phosphorylation of NS5A at serine 146 might regulate hyperphosphorylation in a distal region (Chen et al., 2010).

Interestingly, a serine at residue 146 is only present in genotype 1a and 2a isolates of HCV, in the majority of other isolates across all genotypes and subtypes (including 1b and 2b) this residue is an alanine, although the surrounding sequence is highly conserved. It is possible therefore that the presence of a phosphorylatable residue at this position might contribute

to the enhanced replicative capacity of JFH-1 by allowing for an additional level of control of hyperphosphorylation.

Further, albeit circumstantial, evidence supporting this hypothesis comes from an examination of the two published structures of domain I. As shown in Figure 3.14 there are two differing dimeric conformations of NS5A domain I, the “closed” conformation (1ZH1) (Love et al., 2009) and the “open” conformation (3QFM) (Tellinghuisen et al., 2005). Both structures utilised the genotype 1b Con1 sequence in which an alanine is present at position 146. In the closed conformation A146 is located close to the dimer interface, whereas in the open conformation it is on the opposite face of each monomer. In the closed conformation, a proximal residue E148 (conserved in both JFH-1 and Con1) formed an intramolecular contact via a hydrogen bond with R112 on the other monomer. It is therefore possible that phosphorylation at residue 146 would either disrupt or stabilise the dimer interaction thereby regulating a switch between the closed and open conformations.

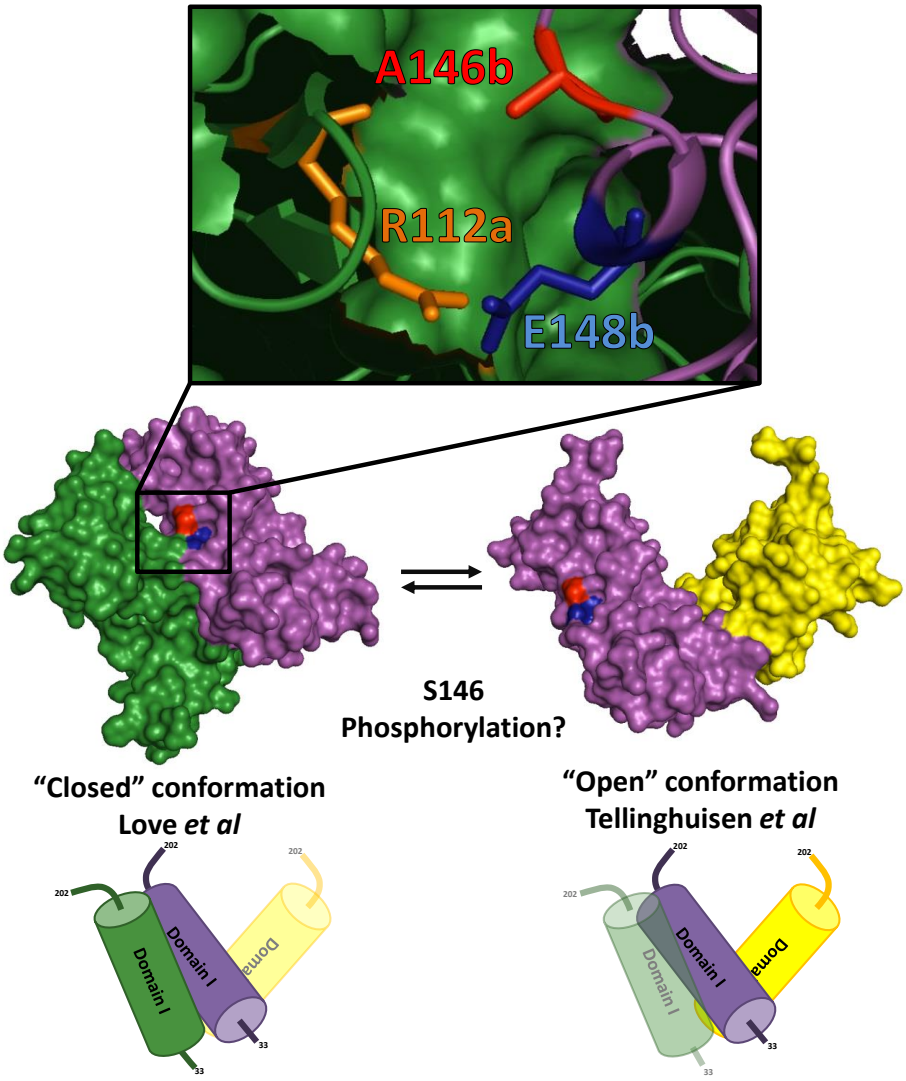


Figure 3.14. Possible mechanism by which S146 phosphorylation might function. The two differing dimeric conformations of NS5A domain I are shown, the "closed" (1ZH1) (Love et al., 2009) and the "open" (3QFM) (Tellinghuisen et al., 2005), green/purple and purple/yellow respectively. The structural studies utilised a Con1 sequence where an alanine is present at position 146. Residue E148b (conserved in JFH-1 and Con1) forms an intramolecular bond with R112a on the opposite monomer, blue and orange respectively (Love et al., 2009). It is likely that phosphorylation at 146 would exert an effect on this key dimer interaction of the "closed" conformation. A schematic illustrating the relative orientations of "open" and "closed" conformations is shown below.

To further support this as a potential mechanism the structural consequences of a glutamic acid substitution at position A146 was modelled using the Robetta full-chain protein structure prediction server (http://robetta.bakerlab.org) (Kim et al., 2004). The generated *de novo* structure of ConI-A146E was then aligned with the Tellinghuisen *et al.* crystal structure using PyMol, Figure 3.15. This computational modelling suggests the introduction of a negative charge at position 146 can be tolerated with respect to monomer structure. This suggests that the mechanism by which phosphorylation of serine 146 negative regulates phosphorylation is not through conformational changes in the domain I monomer.

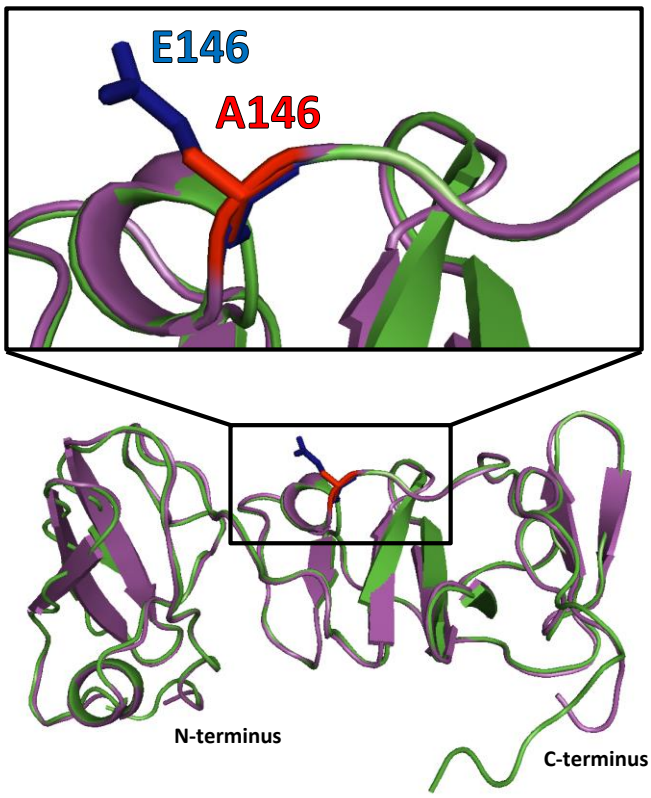


Figure 3.15. Computational modelling of a glutamic acid substitution at position 146.
A glutamic acid was substituted into the position 146 in the ConI sequence to mimic phosphorylation. *De novo* tertiary structure of domain I with the phosphomimetic was predicted using the Robetta full-chain protein structure prediction server (green) and aligned to the NS5A domain I crystal structure (purple) (Tellinghuisen et al., 2005) using PyMol (Schrodinger, 2010)

3.4 Summary

This analysis revealed multiple phosphorylation sites and complemented data with extensive mutagenesis and phenotype profiling in both sub-genomic replicon and virus systems. These data reveals a complex pattern of phosphorylation within LCS I and provide evidence for a sequential cascade of phosphorylation across this sequence and regulation by a distal phosphorylation event within domain I. These data provides novel insights into the complexity and regulation of NS5A phosphorylation and provides a framework for future studies, including identification of the kinases involved. Furthermore, NS5A phosphorylation

is likely to be a dynamic process, therefore also requiring the action of protein phosphatases, such as PP2A, which has been shown to be up regulated by HCV infection (Bernsmeier et al., 2008), and to interact with NS5A (Georgopoulou et al., 2006).

In summary these data does not support the previously proposed hypothesis that hyperphosphorylation is required for virus release and inhibits genome replication. To the contrary, the disruption of certain phosphorylation sites in LCS I, the proposed site of hyperphosphorylation, was inhibitory to replication (S222, S225 and S229) and that phosphomimetic at certain positions restored replication (S222 and S225). Mutations were also identified that significantly reduce hyperphosphorylation yet have no effect on either genome replication or virus production (S146D and S225D).

While it is still likely that phosphorylation of NS5A regulates the different roles it plays in the virus lifecycle, the data here shows clearly that these roles cannot be partitioned between one or either phosphorylation state. Indeed upon review of the literature the consensus is already emerging that elements of both phosphorylated species are important for both replication and virus release. Key to a unifying model on the phosphorylation of NS5A is further attempts to completely map and characterise the regions of phosphorylation. Only then can the complexities of how function follows phosphorylation be fully eluded.

Chapter 4: The role of domain II in virus replication and release

4.1 Introduction

The domain II of NS5A spans resides 250-338 (JFH-1 numbering) of the protein and has been shown to be intrinsically unstructured. Early work identified a correlation between the sequence of the N-terminus of domain II (residues 237-272) and how well a patient responded to interferon treatment, and it was termed the interferon sensitivity determining region (ISDR) (Gale et al., 1997). The ISDR was later shown to act by binding the interferon stimulated gene PKR and inhibiting its kinase activity in a genotype dependent manner (Gale et al., 1998). Interestingly the deletion of this entire region (residues 244-302) was shown to have no effect on virus replication or virus production (Appel et al., 2008).

However, the C-terminus of domain II has previously been shown to be an absolute requirement for virus replication (Appel et al., 2008). It was shown that in the genotype 2a J6/JFH-1 chimeric virus (JcI) a deletion of the C-terminal 36 aa of domain II (residues 303-338) resulted in a complete abrogation of virus replication. Conversely the entire preceding N-terminus of domain II (residues 244-302) could be deleted with no detrimental effect on replication or virus production (Appel et al., 2008, Tellinghuisen et al., 2008b) (Figure 4.1A). At the same time utilising the genotype 1b SGR Tellinghuisen *et al.* found that small deletions within the C-terminal 56 aa of domain II completely abrogated replication. Further analysis by site-directed mutagenesis identified 16 individual residues within this region that were important for replication (Tellinghuisen et al., 2008b).

Neither of these studies investigated whether the C-terminus of domain II contained residues which were important for the production of virus. Tellinghuisen *et al.* were constrained by using a genotype 1b replicon system that was not able to undergo virus production while Appel *et al.* did not introduce small enough deletions so as to generate non-lethal mutations, again precluding the investigating of potential effects on virus production.

It was therefore of interest to investigated whether the highly conserved C-terminus of domain II contained residues important for the production of virus, as well as virus replication. From such data it would also be possible to draw comparisons between two different genotypes as to the requirement of this domain in HCV replication.

Furthermore, observations by Tellinghuisen *et al.* found that in genotype 1b the presence of different culture adaptive mutations influenced whether or not an alanine point mutation in

domain II affected replication. For example, an alanine substitution at proline 324 (P320 in JFH-1) was lethal to replication in the presence of the culture adaptive mutations E1202G, T1289I and K1846T. However, in presence of only the culture adaptive mutation S2204I, the same alanine substitution resulted in only a slight impairment to replication (Tellinghuisen et al., 2008b). As culture adaptive mutations are a requirement for detectable levels of ConI replication, it is therefore unknown as to what residues in wildtype ConI would be important for genome replication.

Whilst this investigation was in progress work by other groups began to shed light on the role of domain II in virus replication. This work highlighted the host factor cyclophilin A (CypA) and its corresponding peptidyl-prolyl isomerase (PPIase) activity as essential for virus replication. Interestingly CypA has also been identified as a critical host factor in a diverse group of other viruses including HIV, influenza and recently here, HCV (Fischer et al., 2010). Disruption of the PPIase activity of CypA through a variety of mechanisms, including chemical inhibition, active site mutations and siRNA knockdown have all been shown to abrogate to HCV replication (Chatterji et al., 2009, Yang et al., 2010).

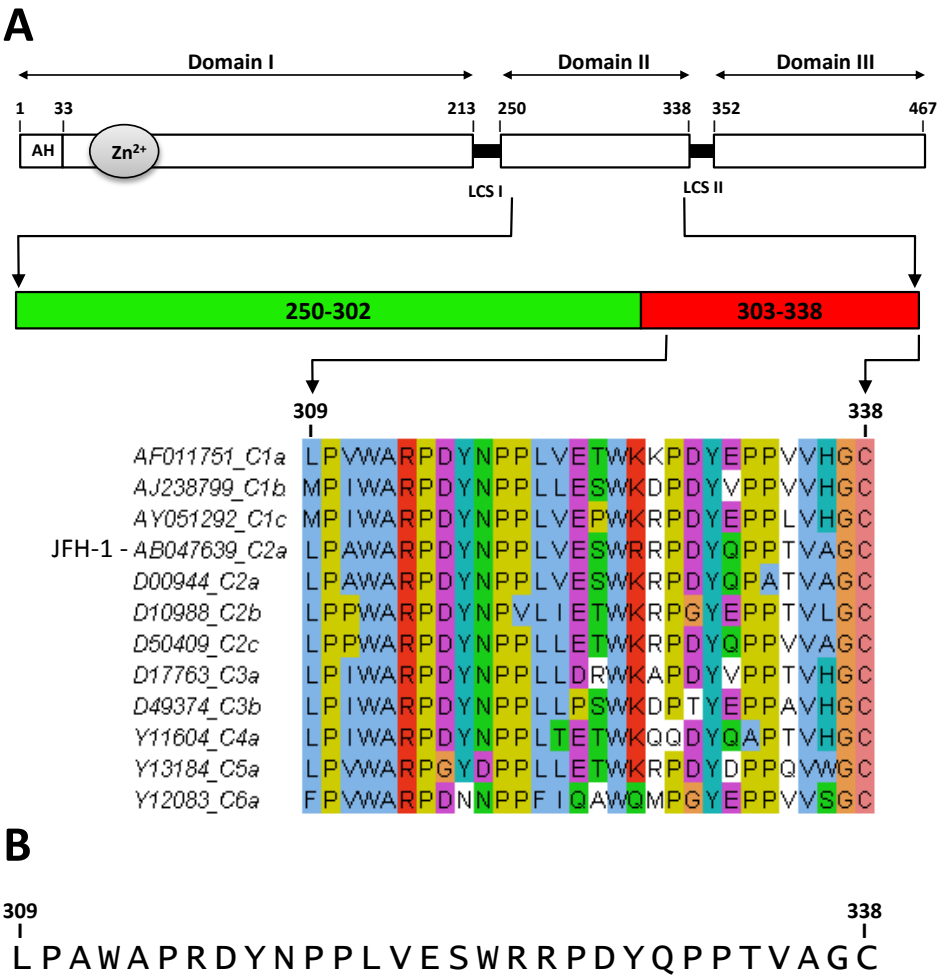


Figure 4.1. NS5A domain schematic and conservation across C-terminus of domain II.
A. Three major domains linked by LCSI and II. Previously a 35 aa deletion at the C-terminus of NS5A was shown to be lethal to viral replication (red), while deletion of the remaining N-terminus of domain II was tolerable to deletion (green) (Appel et al., 2008). Conservation of amino acids over this region in major isolates (Simmonds et al., 2005) were aligned by ClustalW2 and graphic generated by Jalview. Colour in sequence alignment denotes conservation of residue properties (hydrophobic=blue, hydrophilic=green, positive=red, negative=purple and proline=yellow). Numbering refers to JFH-1 NS5A aa position. **B.** The highly conserved C-terminal 30 aa of domain II were targeted for alanine scanning mutagenesis.

When cells harbouring HCV are treated with CypA inhibitors, like cyclosporine A (CsA) and Alisporivir, several drug resistance mutations readily arise that cluster in domain II of NS5A. When these resistance mutations are engineered back into the HCV virus they confer a consistent, and additive, increase in the concentration of inhibitors required to block replication (Coelmont et al., 2010, Yang et al., 2010). Recent NMR studies have also generated convincing evidence that CypA interacts with several residues within the C-terminus of domain II (Coelmont et al., 2010, Hanoulle et al., 2009a). This interaction motif was shown to overlap with the NS5B motif and while CypA inhibitors blocked the domain II:CypA interaction, they did not disrupt the domain II:NS5B interaction, thereby hinting at a potential mechanism for this interaction (Rosnoblet et al., 2012). While cyclosporine A is known to elicit its immunosuppressive effect via the calcineurin pathway, the development of a CsA structural analogue without this activity demonstrated that it is specifically the inhibition of CypA that is detrimental to HCV replication (Hopkins et al., 2010).

It was therefore not only the aim of this project to investigate whether the C-terminus of domain II has a role in virus replication and release, but also to see whether mutation of residues shown to interact with CypA altered the requirement of CypA for HCV replication. This was achieved by utilising the fully infectious genotype 2a isolate JFH-1 that is able to undergo the complete virus lifecycle in tissue culture. Furthermore, this isolate does not require culture adaptive mutations in order to replicate in such systems, thereby avoiding some of the conflicting findings of Tellinghuisen *et al.* discussed previously.

4.2 Results

4.2.1 Generating a panel of alanine mutations in the C-terminus of domain II

To examine the role of the C-terminus of domain II in the virus lifecycle a reverse genetics approach was adopted. The original deletion that Appel *et al.* demonstrated was lethal to virus replication spanned residues 303-338; however through smaller deletions it was shown by Tellinghuisen *et al.* that residues 304-308 could also be deleted with no effect on replication (Appel et al., 2008, Tellinghuisen et al., 2008b). The sequence alignment of residues 309-338 showed a high degree of conservation across all genotypes (Figure 4.1A). It was therefore decided to target the entire 30 aa at the C-terminus of domain II (309-338) for alanine scanning mutagenesis and subsequent profiling in the JFH-1 replicon and virus system. Where alanine occurred in the wild type sequence a mutation to glycine was introduced.

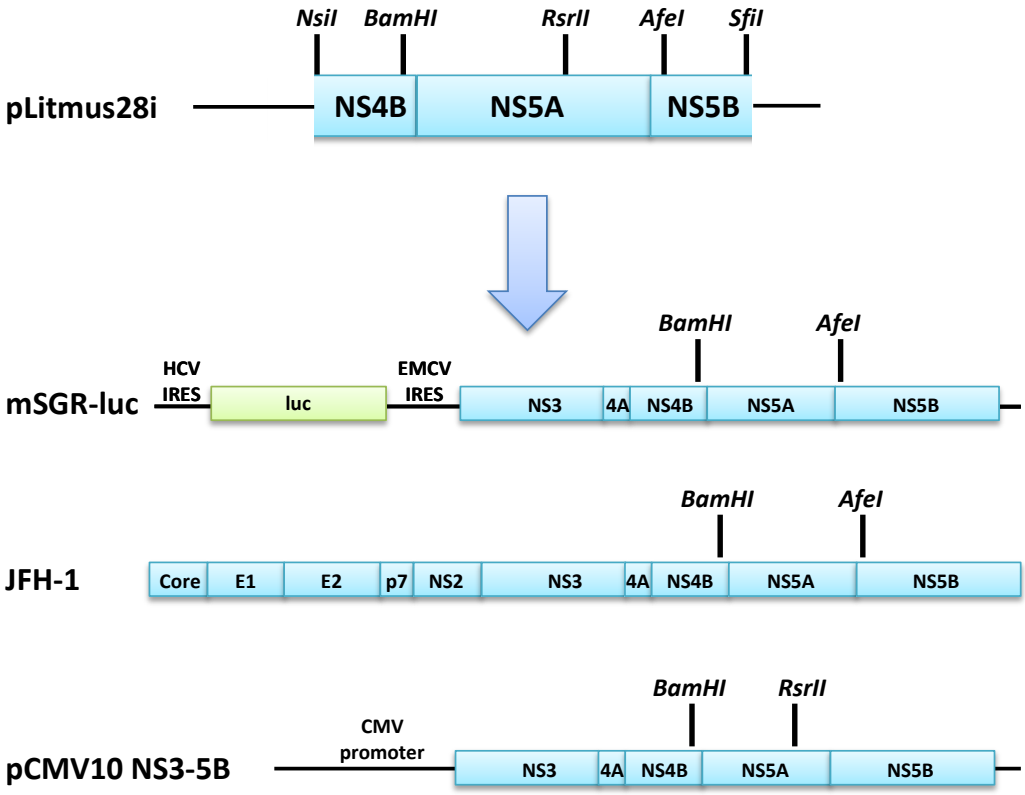


Figure 4.2. Introduction of point mutations into HCV genomes.
A Litmus28i sub-vector containing the *NsiI*/*HindIII* fragment from mJFH-1 was previously generated (Hughes et al., 2009b). Point mutations were inserted by site directed mutagenesis then cloned into replicon (mSGR-Luc JFH), full length virus (mJFH-1) or mammalian cell expression vector (pCMV10-NS3-5B) (Jones et al., 2007b) via *BamHI*, *RsrII* and *AfeI* restriction sites as indicated. All constructs were confirmed by DNA sequencing.

Previously unique *BamHI/AfeI* restriction sites that flank NS5A were introduced into both replicon and virus constructs to increase cloning efficiency. These synonymous mutations were shown to have no effect on virus replication, assembly or release, and these constructs are prefixed with "m", luc-mSGR, mJFH (Hughes et al., 2009b). The domain II mutations were first introduced by site directed mutagenesis into the Litmus28i vector containing the *NsiI*/*HindIII* fragment from the mJFH-1 genome (Section 2.5.3). The NS5A gene was subsequently cloned into the mSGR-luc and the JFH-1 infectious clone construct (mJFH-1) via *BamHI/AfeI* sites. The mutant NS5A genes were also cloned into the pCMV10-NS3-5B expression vector via unique *BamHI/RsrII* restriction sites. This vector was created by cloning the *NsiI/RsrII* fragment from mJFH-1 into the pCMV10-NS3-5B[NS5A(GFP)] to allow for the

ectopic expression of NS5A in the cases where mutations blocked replication (Jones et al., 2007b). All newly synthesised constructs were confirmed by DNA sequence analysis. A summary of the cloning routes used is shown in Figure 4.2.

4.2.2 The role of domain II in replication of the SGR

To investigate the effect of disrupting residues in the C-terminus of domain II the panel of alanine mutations in the mSGR-luc JFH-1 was electroporated into Huh7 cells as described (Section 2.8.1) and luciferase activity measured at 4, 48 and 72 hpe. The luciferase activity at 4 hpe correlates with translation of input transcripts, prior to onset of replication or degradation, as such all subsequent time points are normalised to the signal at 4 hpe to account for electroporation efficiency (Figure 4.3A). Raw luciferase data for each time point is shown in appendix (Appendix figure 8.8). As a negative control the NS5B polymerase inactive mutation, GDD->GND, was used throughout (SGR-luc GND) (Targett-Adams and McLauchlan, 2005).

Of the thirty mutations, twelve (L309A, P310A, W312A, A313G, N318A, W325A, R326A, Y330A, V335A, A336G, G337A and C338A) were shown to completely disrupt the ability of the mSGR-luc-JFH-1 to replicate in Huh7 cells. A further eight mutations (A311G, R314A, P315A, P319A, P320A, D329A, P332A and T334A) showed a significant reduction ($p>0.05$) from WT replication. The remaining ten mutations (D316A, Y317A, L321A, V322A, E323A, S324A, R327A, P328A, Q331A and P333A) showed no statistically significant difference ($p<0.05$) from WT mSGR-luc-JFH-1 replication (Figure 4.3). All mutant transcripts showed a broadly comparable luciferase activity at 4 hpe, demonstrating that the absence of replication is not a result of poor electroporation efficiency (Appendix figure 8.8)

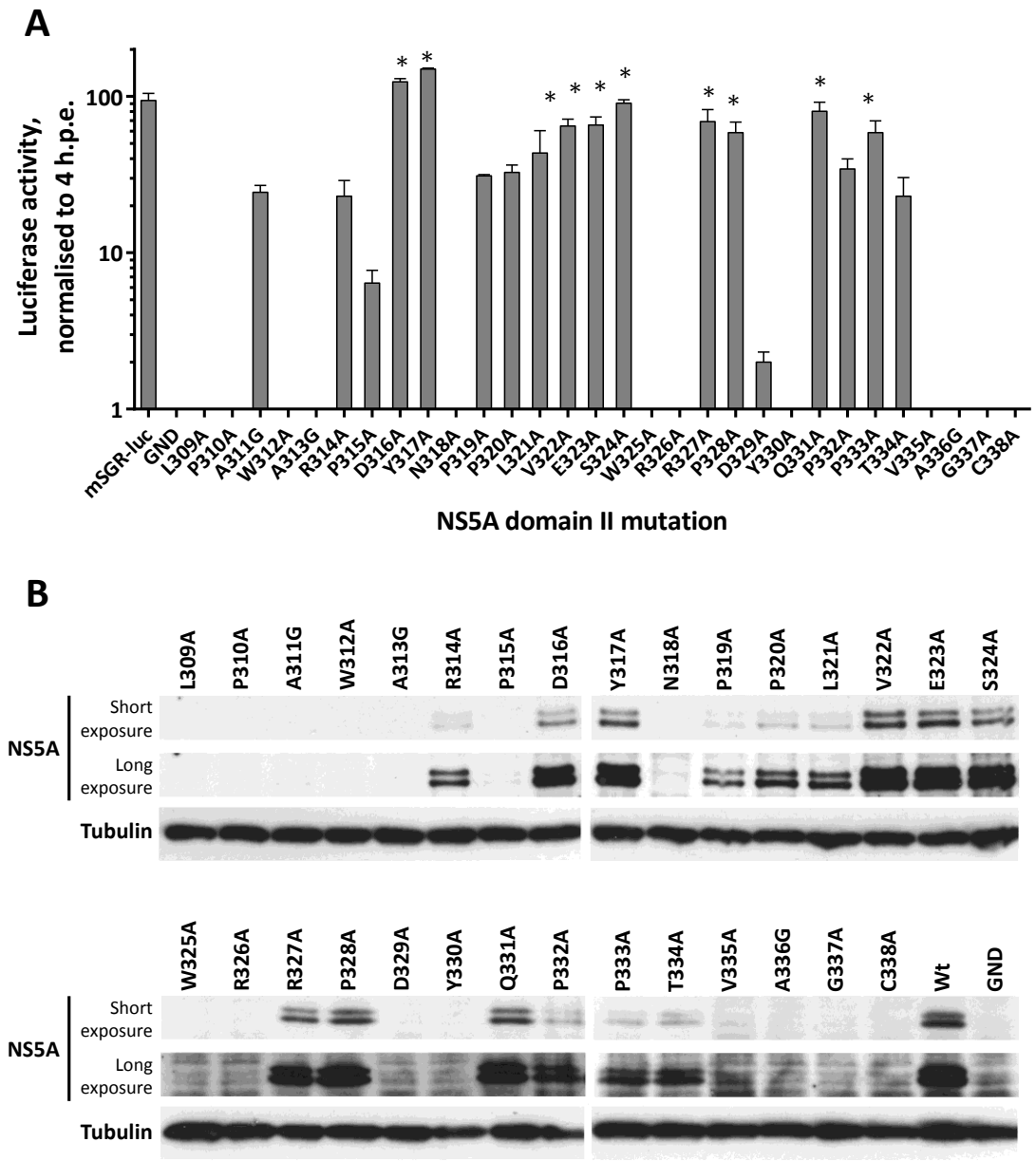


Figure 4.3. The NS5A domain II in replication.

In vitro transcripts of mSGR-luc containing alanine mutations were electroporated into Huh7 cells and followed over 72 hours. **A**. Luciferase activity at 72 hpe normalised to 4 hpe. **B**. Cell lysates at 72 hpe were analysed by SDS-PAGE/Western blot and probed for α-NS5A (Sheep) and α-Tubulin (rat). Short and long exposures of the same NS5A blot are shown. Levels of NS5A broadly correlate with luciferase replication levels, with the exception of A311G where NS5A levels are dramatically reduced compared to similarly replicating mutants. * $p>0.05$ from WT, $n > 3$.

To further support the luciferase data, total protein at 72 hpe was analysed by SDS-PAGE/Western blotting and probed for both NS5A and tubulin (Figure 4.3B). In the case of 15 of the 18 replicating mutants, both phosphorylated forms of NS5A could be detected at levels broadly correlating with the observed level of replication measured by luciferase assay. However for mutants A311G, P315A and D329A, the NS5A protein could not be detected, even at higher exposures of the western blot. For P315A and D329A the absence of NS5A was expected due to the very low levels of replication observed, but A311G replicated to levels comparable to R314A and T334A, both of which had detectable levels of NS5A in cell lysates. Therefore it was possible that the A311G mutation might result in the destabilization of NS5A, or disrupt a major antibody epitope, however when expressed from a NS3-5B expression vector (discussed below) NS5A was detectable in a manner indistinguishable from wild type, both with respect to molecular weight and intensity.

4.2.3 Lethal mutations do not disrupt polyprotein processing.

In the case of non-replicating mutations it was important to establish that loss of replication had resulted from a loss or disruption of a specific function of NS5A, rather than disruption of polyprotein translation or proteolytic processing as a whole. To investigate this, an ectopic expression vector containing the NS3-NS5B coding region under the control of a CMV promoter, pCMV10-NS3-5B, was used (Jones et al., 2007b). This would allow for replication-independent expression and proteolytic processing of the polyprotein. The 12 non-replicating mutants, together with the three mutants for which NS5A expression could not be observed (A311G, P315A and D329A) and R314A as a representative mutant with an intermediate phenotype, were cloned into pCMV10-NS3-5B as described previously.

These plasmids were transfected into Huh7 cells as detailed in the methods (Section 2.7.2), and cell lysates were analysed for protein expression by Western blot at 72 hours post transfection (hpt). All 16 mutants expressed levels of NS5A comparable to WT and detectable levels of both basally and hyperphosphorylated NS5A (Figure 4.4). This confirmed that the lethal replication phenotype observed for these alanine mutations results from a direct loss of NS5A function, and was not due to global effects on NS5A translation, cleavage from the polyprotein or phosphorylation. Intriguingly A311G, which was undetectable in the context of a replicating SGR, could be detected at wildtype levels in this assay. This suggests that additional factors may influence the stability of NS5A when incorporated into active RNA replication complexes.

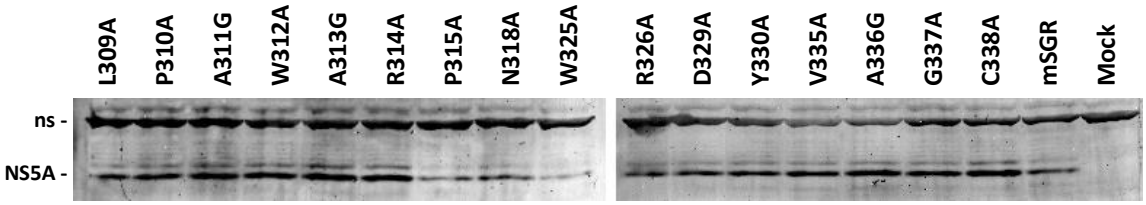


Figure 4.4. Expression of NS5A from pCMV-NS3-5B expression vector in Huh7 cells.
Huh7 cells were transfected with the pCMV10-NS3-5B expression vector containing the denoted mutation. At 72 hpt cell lysates were harvested in GLB and analysed by SDS-PAGE/Western blot and probed with α-NS5A (sheep). The non-specific band has previously been show to correlate well with total protein and so is included as a loading control. Both hyper and basally phosphorylated NS5A can be detected all mutations.

4.2.4 The role of domain II in virus release.

Having established what elements of domain II are required for replication it was important to next establish whether there were residues within this region that were critical to the release of infectious virus. Previous work in this lab and others, (Tellinghuisen et al., 2008b), have identified residues in LCS II and domain III of NS5A that have no phenotype in genome replication but play roles in later stages of the virus lifecycle. For example P342A mutation has no effect on genome replication but reduced infectious virus titres by 1 log (Hughes et al., 2009a).

To investigate whether domain II mutations were disrupting virus production the full length infectious clone JFH-1 (mJFH-1) was use. The panel of domain II alanine point mutation were cloned into the mJFH-1 virus as described earlier (Section 4.2.1). *In vitro* transcripts were electroporated into Huh7 cells as described and incubated for 144 hrs, with a 1:5 passage at 72 hpe. At 144 hpe the supernatant was removed, clarified and virus titre determined using a focus forming assay (Section 2.9.2). In order to be able to correlate virus release with replication in a virus system, which may differ from the mSGR-luc system, total RNA was also extracted from cells and HCV genomes quantified using qRT-PCR (Section 2.9.3). This combination of data meant that a direct comparison between virus replication and virus release could be made from the same sample.

A 144 hrs incubation with 1:5 passage at 72 hpe was required in order to allow for the degradation of input RNA in non-replicating mutants, despite this the negative control JFH-1

GND maintained levels of HCV genomes (or fragments thereof) at significant levels at 144 hpe. The background level of HCV RNA at 144 hpe suggested that the 5' UTR (the target of the qRT-PCR primers) was highly resistant to degradation by cellular RNases; possibly the result of the highly structured nature of the 5'UTR and the stabilising effect of miR-122 (Shimakami et al., 2012). There was however a sufficient window between WT and the GND negative control to identify a genome replication phenotype for the panel of mutations.

11 mutations (L309A, P310A, W312A, A313G, N318A, W325A, R326A, D329A, Y330A, V335A, A336G, G337A and C338A) had level of intracellular genomes at or below that of the negative control GND, indicating that they are not capable of RNA replication. These mutations also released no detectable virus particles, with the exception of D329A which had detectable virus particles at the very limit of quantification. 9 other mutations (A311G, R314A, P315A, P319A, P320A, L321A, R327A, P328A and T334A) had levels of intracellular genomes and released infectious virus that was statistically reduced from that of WT. The remaining 8 mutations (D316A, Y317A, V322A, E323A, S324A, Q331A, P332A and P333A) had levels of intracellular genomes and released virus that was not statistically different from that of WT ($p > 0.05$). The replication of these domain II mutations and the extent to which certain mutations impair or block replication correlates entirely with the observations made in the subgenomic replicon system.

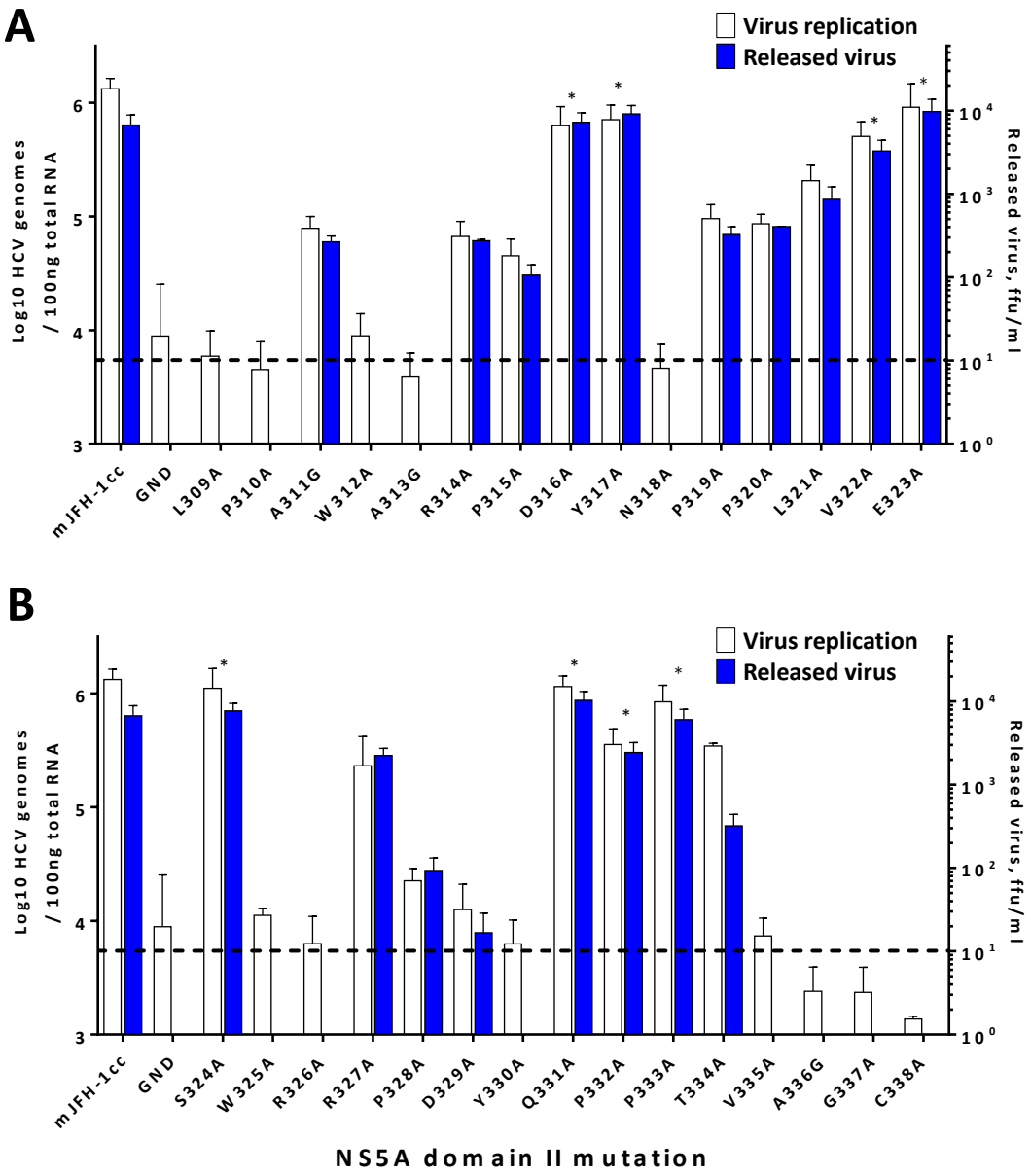


Figure 4.5. The effect of domain II alanine mutations of virus replication and release.
Huh7 cells were electroporated with *in vitro* transcripts of mJFH-1 containing the denoted alanine mutations. At 72 hpe cells were passaged 1:5 before culturing for a further 72 hrs. At 144 hpe supernatants were removed for virus titre determination by focus forming assay (blue), and cells harvested in TRIzol and for analysis of intracellular genomes by qRT-PCR (white). **A.** Residues 309-323 of domain II. **B.** Residues 324-338 of domain II. * $p>0.05$ from WT, $n > 3$. Lower level of quantification (LLQ) for released virus denoted by dotted line.

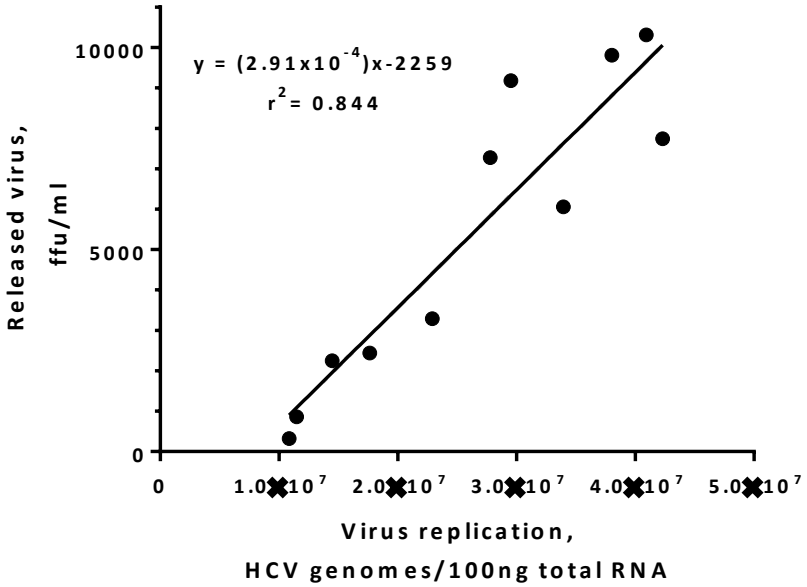


Figure 4.6. Relationship between replication and the release of infectious virus.
Virus replication and release data (Figure 4.5) were averaged the relationship modelled by linear regression (Prism, Graphpad). Mutations which completely blocked virus replication and release were excluded (L309A, P310A, W312A, A313G, N318A, W325A, R326A, D329A, Y330A, V335A, A336G, G337A, C338A).

A relationship was observed between the extent to which a mutation would impair replication and the corresponding reduction in virus release. When analysed by linear regression there was a direct linear relationship, R^2 = 0.844, between virus replication and release was present (Figure 4.6). As replication precedes virus release these data are suggestive that genome replication – specifically the number of intracellular genomes – is a rate limiting step to virus release. This is in agreement with Binder *et al*, who through kinetic analysis of HCV infection also established that replication was the rate limiting factor on virus production (Binder et al., 2013).

Therefore, the conclusion of these data is that the reduction in virus titre observed for domain II mutations stemmed from impairment to replication and not as a result of an additional impairment to the virus assembly/release pathways. These data provide further support for the conclusion that the C-terminus of domain II plays a pivotal role in HCV genome replication but does not contain residues that are essential for the process of infectious virus assembly and release.

4.2.5 The role of domain II in the assembly of infectious virus

It has now been established that there are no residues in the C-terminal 30 aa of domain II that caused a drop in released virus without a corresponding drop in replication. However, to fully discount a role of domain II in virus production it was important to establish if any residues in this region were able to modulate the assembly but not necessarily the release of infectious virus. For example, if mutations were to enhance virus assembly, it might not be observed by assaying only released virus if the release pathway had then become a rate limiting step. While there are not many examples of such a mechanism, to discount domain II playing such a role the effect of alanine mutations on the assembly of infectious virus was investigated.

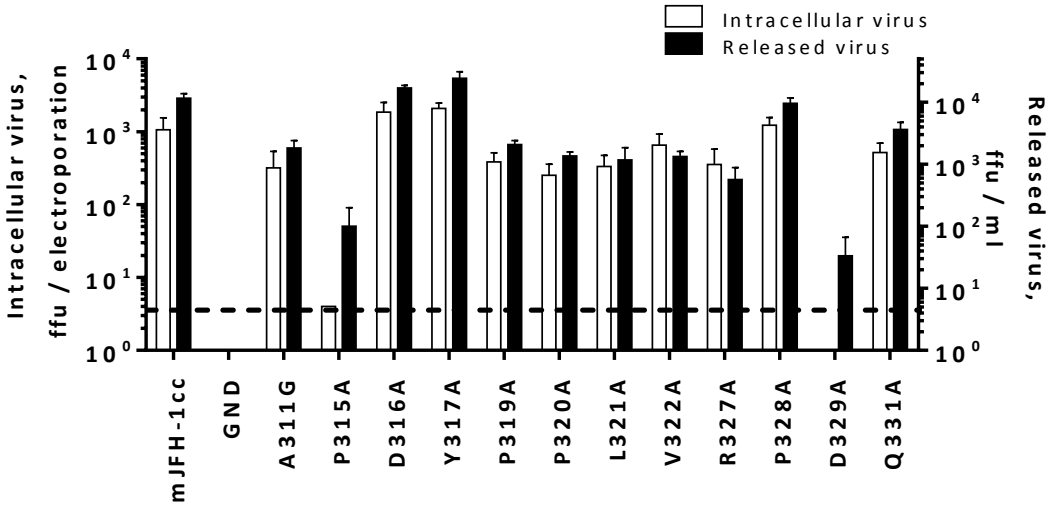


Figure 4.7. A comparison of intracellular and extracellular infectious virus for a selection of domain II alanine mutations.

In vitro transcripts were electroporated into Huh7 cells, at 72 hpe supernatants were removed for virus titre and cells were harvested for intracellular virus. Both extracellular and intracellular virus was determined by focus forming assay, n > 3.

A representative panel of alanine mutations were selected and possible effects on virus assembly investigated by assaying intracellular virus titres. The panel of mutations (A311G, P315A, D316A, Y317A, P319A, P320A, L321A, V322A, R327A, P328A, D239A and Q331A) were selected on the basis they are highly conserved and result in a range of replication phenotypes, from WT to severely impaired. Huh7 cells were electroporated with *in vitro*

transcripts as described previously and at 72 hpe intracellular virus harvested by freeze/thaw lysis (Section 2.9). Released virus was collected in parallel and both intracellular and released virus titre was determined by focus forming assay (Section 2.9.2). Without exception, in the panel of domain II mutants investigated, there was no significant difference between intracellular virus and extracellular virus (Figure 4.7). It is formally possible that there is duality in function amongst certain amino acids shown to be essential for replication, but such a scenario is unlikely and holds many technical difficulties to investigate

From these data it can therefore be concluded that the C-terminal 30 aa of NS5A do not play a significant role in the assembly or release of infectious virus, and instead only plays a critical role at the level of genome replication (Table 4.1).

4.2.6 Effect of domain II mutations Cyclophilin A inhibition

During the course of this research much work was published establishing that an interaction between the host protein CypA and the NS5A protein is critical for virus replication, discussed previously (Sections 1.6.2 and 4.1). It was shown that the chemical inhibition or siRNA knockdown of CypA would readily select for resistance mutations that clustered in the C-terminus of domain II, and that these mutations conferred CypA independence to the virus. In the JFH-1 isolate the double mutant D316E and Y317N exhibited 20-fold resistance to CsA and unlike WT JFH-1 was able to replicate efficiently in CypA silenced cells (Yang et al., 2010). Similarly in the ConI isolate the double mutation R318W and D320E (corresponding to R314W and D320E in JFH) inferred a 10-fold resistance to CsA (Coelmont et al., 2010).

	Effect on virus lifecycle		
NS5A residue	**Virus replication**	**Virus assembly**	**Virus release**
L309	Lethal	n/a	Lethal
P310	Lethal	n/a	Lethal
A311	Impaired	Impaired	Impaired
W312	Lethal	n/a	Lethal
A313	Lethal	n/a	Lethal
P314	Impaired	n/a	Impaired
R315	Impaired	Impaired	Impaired
D316	WT	WT	WT
Y317	WT	WT	WT
N318	Lethal	n/a	Lethal
P319	Impaired	Impaired	Impaired
P320	Impaired	Impaired	Impaired
L321	WT	WT	WT
V322	WT	WT	WT
E323	WT	n/a	WT
S324	Lethal	n/a	Lethal
W325	Lethal	n/a	Lethal
R326	Lethal	n/a	Lethal
R327	WT	WT	WT
P328	WT	WT	WT
D329	Impaired	n/a	Impaired
Y330	Lethal	n/a	Lethal
Q331	WT	WT	WT
P332	Impaired	n/a	Impaired
P333	WT	n/a	WT
T334	Impaired	n/a	Impaired
V335	Lethal	n/a	Lethal
A336	Lethal	n/a	Lethal
G337	Lethal	n/a	Lethal
C338	Lethal	n/a	Lethal

Table 4.1. Phenotype of C-terminal domain II mutations
Summary of alanine mutagenesis (or glycine in the case of WT alanine) on the replication, assembly and release of JFH-I virus.

These resistant residues are located within a motif (RPDY) that is present twice within this region of domain II (R314-Y317 and R327-Y330, Figure 4.1). As residues in both of these motifs showed a reduction in genome replication (R314, P315 and D329, Y330) it was

considered that this might be explained through an alteration in the dependency of NS5A on CypA. It was therefore investigated whether alanine mutations in either of these RDPY motifs were able to modulate the dependency of NS5A on CypA for HCV replication. This was achieved by determining at what concentration the CypA inhibitor cyclosporine A (CsA) was able to block HCV replication. The rationale was that, if a mutation were to reduce the dependence of NS5A on CypA for replication, then a concomitant increase in the concentration of CsA that is inhibitory to HCV replication would be observed. Conversely, if mutations increased the dependence on CypA then a decrease in the concentration of CsA required to block HCV replication would be observed.

The concentration of CsA that is inhibitory to HCV replication was determined by treating Huh7 cells harbouring the mSGR-luc with a range of CsA concentrations and determining the 50 % effective concentration (EC50). Note that an EC50 is the functionally equivalent of a 50 % inhibitory concentration (IC50). The term EC50 is used in place of IC50 where it is feasible that the drug induced response could occur through a pathway other than inhibition of the target molecule.

Huh7 cells were electroporated with *in vitro* transcripts of mutant and WT mSGR-luc-JFH-1. At 4 hpe cells were treated with CsA at concentrations ranging from 0.01 to 100 µM, then followed for a further 44 hrs before determining luciferase activity at 48 hpe (Section 2.8.2). Data were normalised to the 0 % inhibition plateau before modelling by a standard concentration of agonist vs. response model (Appendix figure 8.10A) to determine the EC50 for each mutant replicon. As positive control for a modulation in CypA dependence the major resistance mutation D316E was introduced into the mSGR-luc and determined its EC50 in parallel. Conversely, as negative control the EC50 of BMS-790052 against both WT and D316E replicons was determined.

The EC50 curves are shown for mutants 314-317, 327-329 and WT, Figure 4.8. In the case of Y330 the mutation to alanine completely abrogated replicon, so prohibiting EC50 calculations. With the exception of Y317A, all alanine mutations in both -RPDY- motifs (314-317, 327-329) resulted in modest increases in the sensitivity of the mSGR-luc to CsA treatment. Conversely Y317A showed an equally modest decrease in sensitivity to CsA treatment. As expected D316E resulted in a 1-log increase in resistance to CsA, while remaining comparable to WT in its sensitivity to BMS-790052 inhibition, summary in Table 4.2 and raw data in Appendix figures 8.10B, C and 8.11B, C.

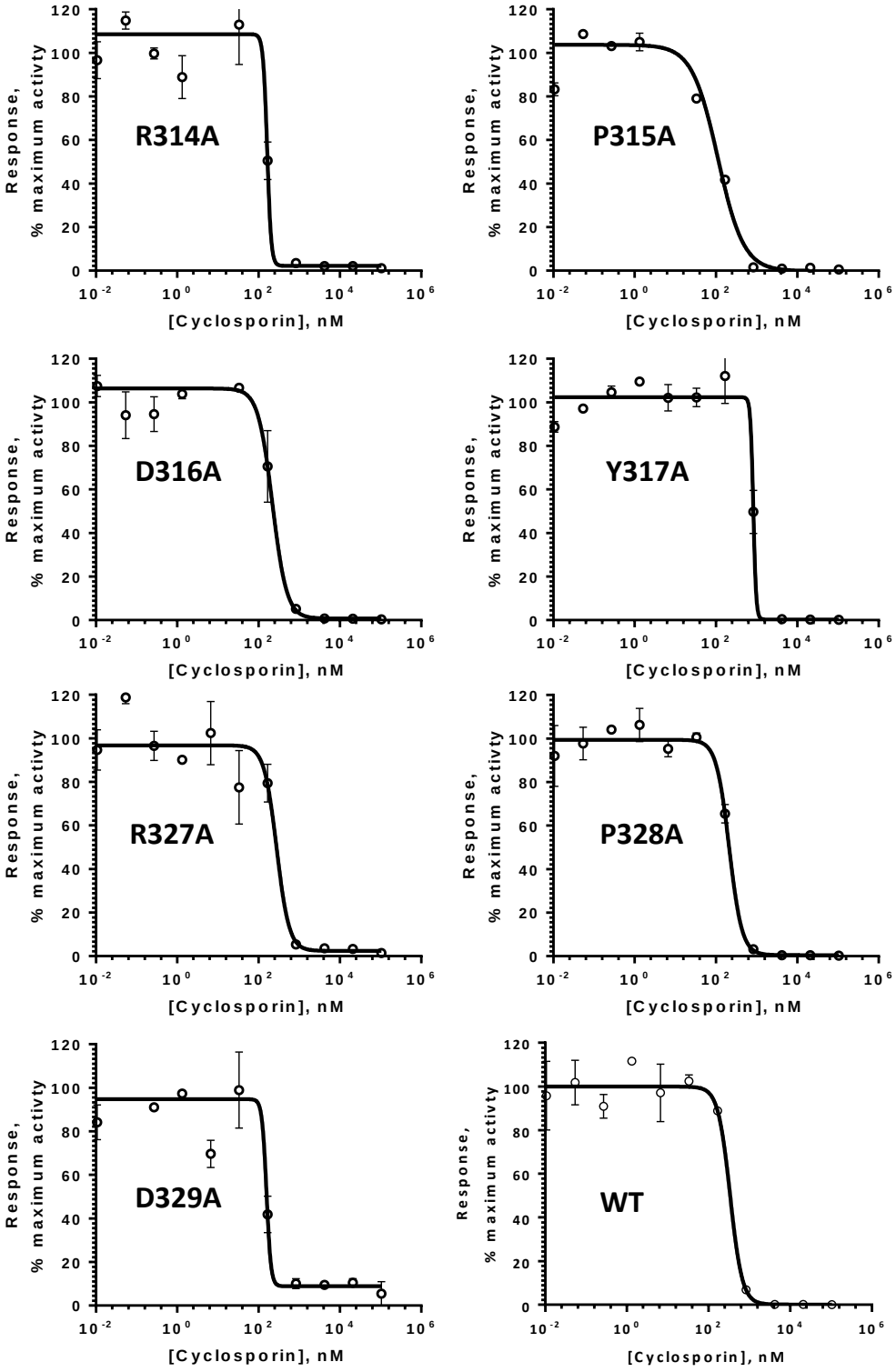


Figure 4.8. Sensitivity of -RPDY- motif mutations to CsA treatment.

Huh7 cells electroporated with the denoted mSGR-luc were treated at 4 hpe with serial dilutions of CsA (0.2% DMSO final) and followed over 48 hrs. After which luciferase activity was determined and EC50 curves calculated using Prism 6 (Graphpad). The CsA EC50 for Y330A could not be determined with confidence due to its low level of replication.

In order to correlate the loss of luciferase signal to a specific loss of the HCV replicon, and not a result of drug induced cell toxicity, the concentration at which CsA was toxic to cells was determined by a MTT assay (Section 2.8.3). In brief, at 48 hpe drug treated cells were incubated with 1mM MTT for 2 hrs before washing, dissolving the MTT metabolite in DMSO and determining absorbance at 570 nm. The cytotoxicity concentration, 50% activity (CC50) was calculated for WT and D316E by the same process as EC50. Both WT and D316E replicons showed comparable sensitivity to CsA induced toxicity (Table 4.2 and Appendix figure 8.10B, C). This confirmed that the calculated CsA EC50 correlated to inhibition of the mSGR-luc and not a cell toxicity effect.

These data show that within this region of domain II there are several mutations capable of modulating the dependence of NS5A on the activity of cyclophilin A, but that this modulation is not as significant as for the major CsA resistance mutation D316E.

Drug treatment	**Domain II mutation**	**EC50 +/- SEM**		**CC50**
Cyclosporin A, nM	R314A	124	+/- 20.1	16,452
	P315A	108	n/a	n.d
	D316A	262	+/- 20.3	n.d
	Y317A	802	+/- 22.1	n.d
	R327A	264	+/- 411.5	n.d
	P328A	228	+/- 17.8	n.d
	D329A	153	n/a	n.d
	D316E	3,199	+/- 22.1	17,303
	mSGR	378	+/- 43.2	17,586
BMS-790052, pM	D316E	24	+/- 2.1	n.d
	mSGR	17	+/- 5.0	n.d

Table 4.2. Summary of CsA sensitivity of DII mutant replicons.
CsA EC50 for domain II alanine mutations in RPDY motif. As positive control the major resistance mutation, D316E, showed a 1-log increase in CsA resistance. Conversely the D316E mutation showed no change in BMS-790052 sensitivity compared to WT. n > 2.

4.3 Discussion

4.3.1 Involvement of domain II in virus replication and release

This study has utilised the HCV genotype 2a isolate JFH-1 to investigate the role of the C-terminal 30 residues of NS5A domain II within the virus life cycle. A total of 12 residues within this region were identified as essential for viral genome replication. The substitutions of these amino acids to alanine (or to glycine in the case of alanine residues within NS5A) abrogated genome replication in both full length virus and subgenomic replicons. A further 8 mutations resulted in a significant inhibition of replication implying that, although these residues are not essential, they do contribute to genome replication. There was however, no residue within this region that when mutated resulted in a reduction in released virus titre without a corresponding reduction in viral replication, implying that this region is exclusively involved in genome replication.

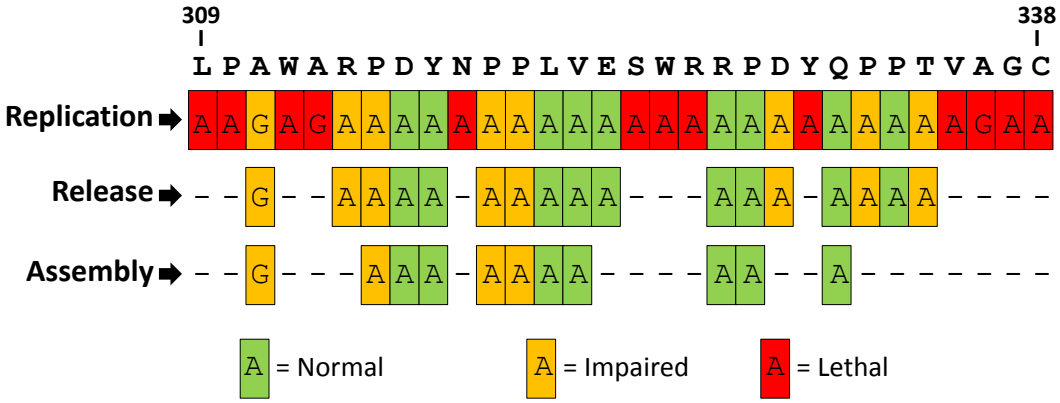


Figure 4.9. Summary of the effect of alanine mutations in the C-terminal of domain II.
Mutations either had no statistically significant effect (green), impaired (orange) or completely abrogated (red) virus function when compared to WT and non-replicating (GND). Impairment was defined as being statistically different ($p<0.05$) from WT, but still with detectable levels of replication and/or virus release.

These data reinforces the conclusion that the primary role of domain II is in genome replication, and that those residues not required for replication played no role in the release of infectious virions. However, the possibility that residues essential for replication are also required for virus assembly and/or release cannot be ruled out. Although such a duality in function is unlikely and technically challenging to address experimentally. Two previous studies have assessed the requirements for genome replication of the corresponding region in the genotype 1b subgenomic replicon (Shimakami et al., 2004, Tellinghuisen et al., 2008b);

but as the genotype 1b isolates were not able to undergo infection in cell culture investigation of the later stages of the virus life cycle was precluded. The data generated in that study are presented here allow a comparison of residues required for genome replication between the two genotypes.

4.3.2 Genotype specific requirements within domain II for RNA replication.

As illustrated in Figure 4.1 this region exhibits a high level of sequence conservation across the genotypes in this region of domain II. Ten of the 30 residues are absolutely conserved, and a further 10 show a very high level of sequence conservation (at least 90 %). One prediction would be that those residues that are absolutely conserved would also play critical roles, however this is not entirely the case. Although P310, W325, G337 and C338 are essential for genome replication in both genotype 1b and 2a, the data for other conserved residues is less clear cut. For example, in this study the mutation of residues W312, A313, Y330 and V336 are lethal, however, in genotype 1b the corresponding mutations have either a WT or partial phenotype.

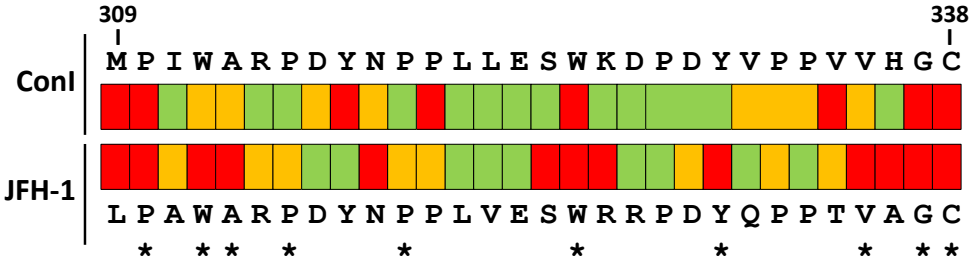


Figure 4.10. Comparing the requirements of domain II for replication from different genotypes.

Replication data generated previously (Tellinghuisen et al) in the genotype 1b Con I replicon is compared with data from this investigation using the genotype 2a JFH-1 isolate. Mutation to alanine (or glycine) either had no statistical effect on replication when compared to WT (green) or completely abrogated replication (red). Where mutation resulted in an intermediary impairment is denoted by orange.

Focussing on the differences between Con1 and JFH-1, 22 residues are conserved between the two genotypes, yet in only a minority of these residues do the requirements match (Figure 4.10). In fact only five residues appeared to be dispensable for both genotypes. Overall within this region, 77 % of residues in the JFH-1 isolate are required for genome

replication compared to only 50 % in the Con1 isolate, indicating that this region appears to have greater importance for genome replication in the JFH-1 isolate.

These results argue strongly that it is difficult to extrapolate between data sets derived from genotype 1b and 2a, suggesting that there may well be significant functional differences between the NS5A proteins of these two genotypes. An additional confounding factor is the observation that the phenotype of selected mutants in the genotype 1b replicon are influenced by the combination of culture adaptive mutations there were present within the remainder of the replicon. For example P324A (corresponding to JFH-1 P320A) is lethal in the context of the GIT replicon (mutations E1202G and T1280I in NS3, and K1846T in NS4A) but only displays a slight impairment in the NS5A mutant S2201I replicon (Tellinghuisen et al., 2008b).

4.3.3 Involvement of domain II in cyclophilin A dependence

How can single point mutations, in some cases as subtly as alanine to glycine, elicit such deleterious effects on a region of protein that has only limited elements of secondary structure? One possibility might be in the context of protein-protein interactions as this region of NS5A has been reported to interact with the NS5B polymerase (Shimakami et al., 2004), and the cellular factor, CypA (Hanoulle et al., 2009a, Yang et al., 2010).

The involvement of this region with CypA was focused on because, as discussed, the inhibition of CypA activity by various methods results in the major resistance mutations R314W, D316E, Y317N (Coelmont et al., 2010, Yang et al., 2010). More recently NMR evidence has implicated A311 - N318, T334, G337 and C338 as sites of CypA interaction and preliminary evidence that the P315-D316 bond as the prolyl bond that is cis-trans isomerised by CypA (Coelmont et al., 2010, Hanoulle et al., 2009a, Yang et al., 2010). These residues are located within a motif, RPDY, that repeats twice in the C-terminus of NS5A, Figure 4.11.

During the course of this investigation it was realised that residues within this repeating motif, -RDPY-, did not have corresponding phenotypes. In the first occurrence of the motif the mutation of the RP (314, 315) impaired replication, while the mutation of DY (316, 317) showed no difference from WT. In the second occurrence of this motif the reverse was observed, with RP (327, 328) showing no difference from WT, and DY (329, 330) impairing replication - indeed Y330A completely abrogated replication.

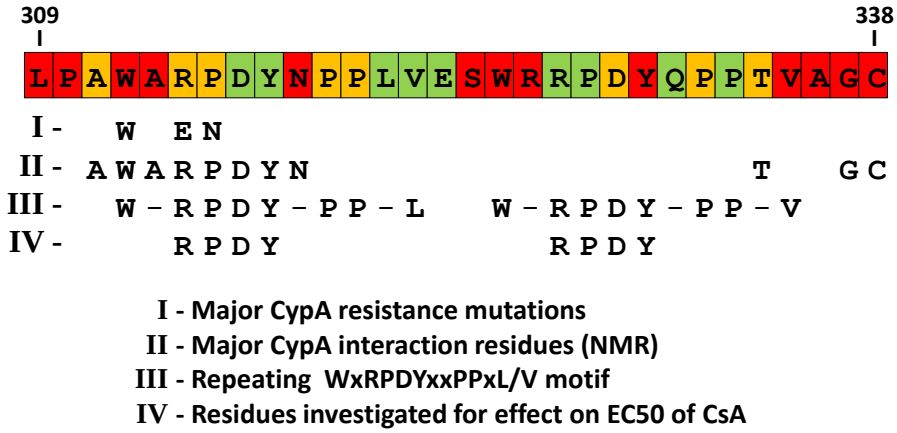


Figure 4.11 Summary of CypA involvement with residues 309-338 of NS5A.
Sequence of JFH-1 NS5A with replication phenotype denoted by colour. WT (green), impaired (orange) or abrogated (red) replication. Major resistance mutations identified in (Coelmont et al., 2010, Yang et al., 2010) and major CypA interaction residues identified in (Hanoulle et al., 2009a, Rosnoblet et al., 2012).

The effect of alanine mutation within this repeating RDPY motif on the susceptibility of HCV to CypA inhibition was investigated. Some modulation of the inhibitory concentration of CsA was observed for several of these mutations, but not as dramatic as the 1-log shift observed for the resistance mutations D316E. In contrast to the NMR discussed above, both Foster *et al.* and Chatterji *et al.* have shown that D316 and Y317 are not required for CypA binding in vitro (Foster et al., 2010, Foster et al., 2011, Chatterji et al., 2010). The data presented here are consistent with the latter observation, as although D316A and Y317A had no apparent phenotype, mutation of the surrounding residues to alanine (or to glycine in the case of A311 and A313) was sufficient to either strongly impair or completely abolish virus replication.

Taken together, these data are therefore consistent with the previously proposed hypothesis that binding of CypA to the motif PAWARP between residues 310-315 is important for HCV genome replication (Hanoulle et al., 2009a). Intriguingly the RPDY motif discussed overlaps with a longer sequence that is present twice within the region of interest WxRPDYxPP (x=any amino acid) (Figure 4.11, III) and that has a high level of conservation across all major genotype isolates (Figure 4.1).

It would be interesting to analyse the phenotype of combinatorial mutations to determine if there is any redundancy resulting from the duplication. A further twist to this story is the observation that in Con1 (Tellinghuisen et al., 2008b), the phenotypes of mutations of the two DY motifs are opposite to the JFH-1 phenotype. D320 and Y321 (corresponding to JFH-1 D316 and Y317) are required for genome replication whereas D333 and Y334 (JFH-1 D329 and Y330) are not required.

4.4 Summary

In conclusion, the C-terminal 30 residues of NS5A domain II has no role in either the assembly or release of infectious virus in the genotype 2a isolate JFH-1, and that the function of this region is restricted to that of the replication of viral genomic RNA. Although this role in genomic replication is consistent with data obtained for genotype 1b, this data highlights phenotypic variations between genotypes irrespective of the high sequence conservation and caution against extrapolation of datasets between genotypes, Figure 4.10 and Appendix figure 8.9. A key challenge for the future will be to determine the mechanisms underpinning how single amino acid changes can completely ablate the function of a protein which holds no intrinsic enzymatic function. In particular it will be intriguing to determine if any of these changes influence the ability of domain II to interact with either cellular or viral proteins, or indeed viral RNA.

Chapter 5: Novel protein tag system for NS5A live-imaging

5.1 Introduction

The subcellular distribution of wildtype NS5A in either replicon or infect cells is that of discrete puncta, present throughout the cytoplasm but enriched in the perinuclear region where the membranous web forms. NS5A is known to co-localise with other viral proteins like NS3 and NS5B as well as with cellular compartments like lipid droplets and ER-derived membranes. Work in recent years has shown that NS5A is actively transported around the cell in a dynein dependent manner and two different sub-populations of NS5A have been observed (Wolk et al., 2008). The NS5A protein is localised in large, predominantly static structures that co-localise with membranous web as well as much small puncta that are highly mobile and are trafficked rapidly around the cell. Initially it was thought that only the static structures represented sites of replication, but recent findings presented by Eyre *et al.* illustrated that both static and mobile structures co-localised with HCV RNA, VAP-A and Rab5A (Eyre et al., 2012). While all these proteins are enriched in replication complexes, whether or not the small mobile structures are actively replicating remains to be shown.

As has been discussed at length the NS5A protein is essential in roles that are spatially and temporally separate events within the cell; replication, assembly and influencing signalling pathways. While the consensus is that phosphorylation of NS5A is regulating these different roles, there is real need for functional data to support this hypothesis.

The presence of functionally separate populations of NS5A has already been illustrated (Fridell et al., 2011, Fridell et al., 2013) and it is possibly that these functional differences correspond to spatially different populations of NS5A as well. All of which could be driven by phosphorylation. Indeed previous data herein has already illustrated that mutation of the phosphorylation site serine 225 is sufficient to dramatically alter the localisation of NS5A in replicon cells while only slightly impairing replication or virus release (Section 3.2.3).

It was therefore of interest to further investigate how phosphorylation events might be affecting the localisation and trafficking of NS5A within a cell through the application of confocal-microscopy on both fixed- and live-cells. In order to distinguish between the different phosphorylated forms of NS5A in confocal microscopy there is the requirement to be able to differentially label them. Typically in fixed-cell IF this is achieved by either generating phosphospecific antibodies or introducing different epitope tags (FLAG, HA etc.) into phosphomutants and WT. Similarly in live-cell confocal imaging this is achieved by introducing different wavelength fluorescent protein into phosphomutants and into WT.

These approaches require time in cloning and in validating the insertion of epitope tags and fluorescent proteins. A bottle neck forms when experiments require that either WT or a phosphomutant be tagged with another tag, as this process of cloning and validation must be repeated. As such, a novel tag system was sought that allowed for much greater interchangeability of variables like fluorophore wavelengths, as well as opening up new experimental possibilities.

The primary aim of this project was the direct comparison of the subcellular localisation of a NS5A harbouring phosphomutants and with that of wildtype. While this could be achieved by screening for differences in co-localisation with other proteins, such as NS3, subtle differences in localisation could easily be missed. However, by differentially labelling the phosphomutant and wildtype NS5A within the same co-electroporated cell, small differences in localisation would become very apparent. This could for example highlight whether or not the phosphomutant S146D, which lacks hyperphosphorylation, localises in the same manner as the culture adaptive mutation S2204I (serine 232 in JFH-1 numbering), which also lacks hyperphosphorylation.

The secondary aim was to investigate temporal differences in the distribution of NS5A in a pseudo pulse-chase experiment. NS5A could be initially labelled with a fluorophore and then, following a defined period of time or drug treatment, could be labelled again with a different wavelength fluorophore. Such a system applied to live confocal imaging it could answer whether newly synthesis NS5A is responsible for the highly mobile puncta while ‘older’ NS5A is more confined to replication complexes, or vice versa.

There is a diverse family of fluorescent proteins with different properties, such as those that are photoactivatable, pH sensitive or shift wavelength as a function of time. However, the modern tag systems that are emerging offer greater flexibility and access to new experimental space that cannot be achieved with even these sophisticated fluorescent proteins. These new tag systems typically involve a genetically modified protein domain that is able to conjugate or chelate to specific substrates. These substrates are themselves linked to a diverse variety of functional groups including chemical fluorophores, quantum dots, biotin or immobilised resins. As these fluorophores are synthetic they generally have a better quantum yield and are less susceptible to photobleaching than their fluorescent protein counter parts.

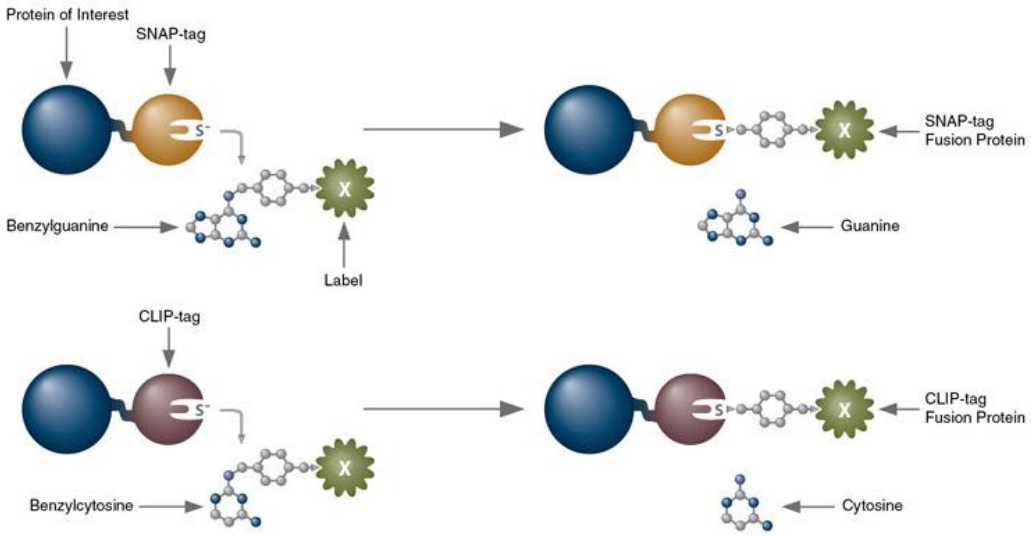


Figure 5.1. Illustration of the SNAP/CLIP technology.
The SNAP or CLIP domain is traditionally fused to the N or C terminus of the protein of interest, when expressed it is able to covalently conjugate to a specific substrate, BG or BC respectively. This system was developed by NEB specifically for the confocal microscope imaging of cells containing labelled proteins. As such there is a wide spectrum of different colour fluorophores already conjugated to both benzylguanine/cytosine that are commercial available and are both non-toxic and cell-permeable. Figure reproduced from (NEB, 2010).

The modern tag system adopted was based on the SNAP-tag technology developed by NEB, a highly versatile protein tag composed of a 20 kDa domain originally derived from the O^6-alkylguanine-DNA alkyltransferase (AGT) (Provost and Sun, 2010). This domain has been modified to react specifically and rapidly with benzylguanine (BG) derivatives, resulting in covalent linkage of the functional group to the protein of interest (Figure 5.1). To develop the possibility for co-labelling the SNAP domain was engineered further to alter its substrate specificity such that it reacted with O^6-benzylcytosine (BC) derivatives. This domain, termed CLIP, would allow for co-labelling within a cell of two differentially tagged proteins.

There is a wide variety of different commercial available substrates for both SNAP- and CLIP-tag domains. These include cell permeable fluorescent substrates that are applicable to both live- and fixed- cell imaging, cell impermeable fluorescent substrates for specific labelling of surface markers as well as a range of affinity capture substrates. The SNAP-technology has also been applied to more complex experiments probing in-cell protein:protein interactions, these are reviewed in (Barth, 2013, Haruki et al., 2012). Most interesting is the use of a SNAP

and CLIP substrate joined by a long chemical linker, such a probe can be used to identify, in a quantitative manner, whether a SNAP-tag and CLIP-tag are in very close proximity.

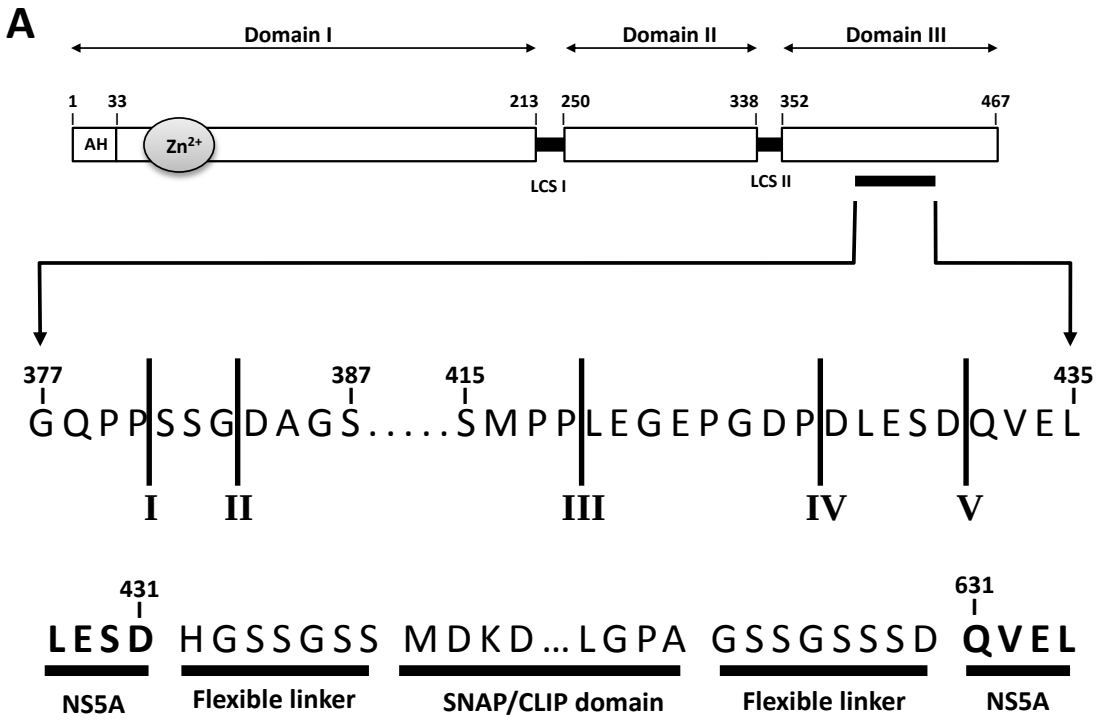
This project aimed to investigate NS5A in the context of active replication/infection not from ectopically expressed vectors. As such the first technical challenge was linking the SNAP/CLIP domain with the NS5A protein in a manner that did not disrupt the virus lifecycle. In the context of both replicon and virus the NS5A protein is translated as part of the viral polyprotein from which it is subsequently cleaved by the NS3/4A polymerase. The insertion of the SNAP/CLIP domain in either an N- or C-terminal position could possibly disrupt this cleavage event and was therefore excluded as a tag location. It has previously been shown by several groups that regions within domain III are tolerable to large insertions with no effect on replication and only marginal effects on virus release, Figure 5.2. As such it was decided to introduce the SNAP/CLIP domain as an internal insertion in domain III at one of these previously validated sites.

The SNAP/CLIP technology has previously only been deployed as either an N- or C-terminal tag, with expression typically driven from a ectopic vector. It was therefore unknown whether the insertion of the SNAP/CLIP tag at an internal position would disrupt the folding of the tag, and subsequently block the ability to conjugate to the specific substrate. It was necessary to validate not only whether the replicon was still viable with a SNAP/CLIP-tag, but also whether the SNAP/CLIP-tag had retained catalytic activity and specificity towards the BG and BC substrate.

5.2 Results

5.2.1 Insertion of SNAP/CLIP domain into NS5A domain III

As discussed domain III of NS5A has been shown to tolerate the insertion of small and large tags between residues 380-432 aa (JFH-1 NS5A numbering) (Figure 5.2). The insertion site V, between residues 430 and 431, was chosen for the insertion of the SNAP- and CLIP-tag as it had been previously validated in the JFH-1 isolate and shown to have no significant effect on replication or virus production (Igloi and Harris, in preparation). The location of insertion site V is within a 19 aa naturally occurring insertion that is present in several genotype 2a isolates, including JFH-1. This 19 aa insertion is not present in other genotypes, implying that it cannot be a key requirement for virus replication or production.



B

Site	Isolate, system	Insertion	Effect on replication	Effect of release	NS5A phosphorylation	Reference
I	ConI, SGR	GFP-FLAG	Significantly impaired	not tested	not tested	Moradpour *et al* 2004
II	JcI, virus	emGFP, RFP	WT replication over 72 hrs	Present, but impaired	not tested	Schaller *et al* 2007
III	ConI, SGR	GFP-FLAG	Impaired	not tested	not tested	Moradpour *et al* 2004
III	JFH-I, virus	GFP	WT replication over 72 hrs	Present, but impaired	not tested	Masaki *et al* 2008
III	JFH-1, SGR	GFP and PAGFP	WT replication over 72 hrs	not tested	Two bands, equal intensity	Jones *et al* 2006
III	JFH-1, SGR	One-Strep-tag	WT replication over 72 hrs	not tested	not tested	Amako *et al* 2009
IV	ConI, SGR	GFP	Impaired	not tested	Two bands, equal intensity	McCormick *et al* 2006
IV	ConI, SGR	PSTCD	WT replication over 72 hrs	not tested	Two bands, equal intensity	McCormick *et al* 2006
V	JFHI, virus	PSTCD	WT replication over 72 hrs	WT virus titres	Two bands, equal intensity	Igoli *et al* in preparation

Figure 5.2. Insertion sites in NS5A domain III and cloning design of SNAP/CLIP insertion.
A. Regions of NS5A where either fluorescent proteins or affinity tags have previously been inserted. Sequence of inserted SNAP/CLIP domain into site V. **B.** Summary of the different virus and phenotypes of previous domain III insertions.

To increase the steric freedom and allow for the correct folding of the SNAP/CLIP-tags, essential in order to retain catalytic activity, the domain was flanked with serine rich flexible linkers, similar to those utilised previously (Igloi and Harris, in preparation). The insertion is shown in Figure 5.3B with an abridged protein sequence of the SNAP/CLIP-tag, a complete protein and DNA sequence is shown in Appendix figure 8.12.

A simple cloning strategy was utilised for the insertion of the SNAP/CLIP domain based around the *BclI* restriction site (Figure 5.3A), a site that is unique in the previously generated pLitmus28i sub-clone containing the JFH-1 *NsiI-HindIII* fragment (Figure 4.2). The DNA coding sequence for the SNAP/CLIP-tag was amplified by PCR using either pSNAPf or pCLIPf vector as template. Custom PCR primers were designed that contained regions complementary to the SNAP/CLIP coding sequence as well as the containing the flexible linker sequence and *BclI* restriction sites. The PCR amplification thereby created a gene block containing the SNAP/CLIP with flexible linkers with flanking *BclI* restriction sites at either ends, Figure 5.3B. The mSGR-luc-JFH vector and gene block were digested with *BclI* and the gene block ligated into the parent vector. This cloning strategy obviously result in both forward and reverse insertions of the gene block so colonies were initially screened for the correct orientation by directional colony PCR and positive colonies confirmed by DNA sequencing.

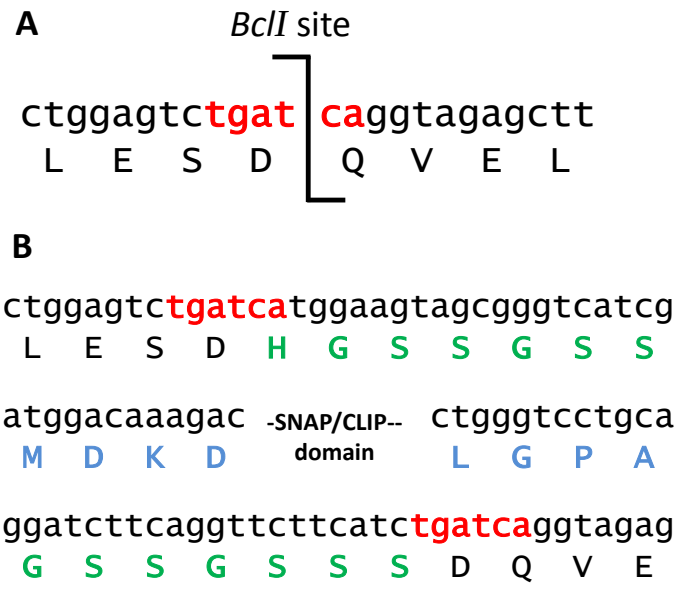


Figure 5.3. Cloning strategy for insertion of SNAP/CLIP domain.
A. Region of NS5A domain III where the SNAP/CLIP domains were inserted, aa 428-435. **B.** Sequence of NS5A domain III showing the insertion of the SNAP/CLIP domain (blue) and flexible linkers (green). The unique *BclI* site used to insert the SNAP/CLIP gene block is highlighted in red. Complete sequence of SNAP- and CLIP-tag is shown in Appendix figure

5.2.2 The effect of inserted SNAP/CLIP domains on HCV replication

While there is now much precedence for the insertion of small to medium sized protein tags into domain III as being well tolerated, there have been occasions whereby insertions have impaired or abrogated replication, Figure 5.2B. As such it was important to first establish whether the insertion of the SNAP and CLIP domains into NS5A had an effect of the replication. To do this a luciferase based replication assay was utilised as described previously (Section 3.2.2 and 4.2.2).

In brief, Huh7 cells were electroporated with *in vitro* transcripts of either a WT mSGR-luc-JFH-1 or derivatives containing the NS5A-SNAP or NS5A-CLIP and replication followed over a 72 hr period by measuring luciferase activity. As shown in Figure 5.4, the replication kinetics of the SNAP and CLIP containing replicons are indistinguishable from that of WT, confirming that the insertion of SNAP and CLIP domains do not disrupt the role the NS5A plays in replication.

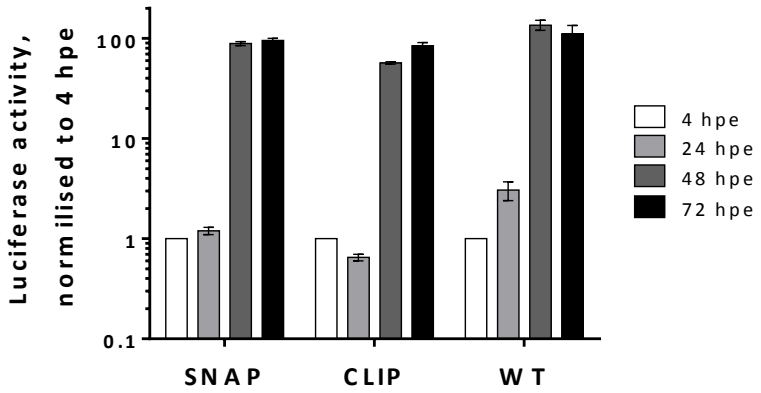


Figure 5.4. Replication competence of the mSGR-luc JFH-1 containing the NS5A-SNAP/CLIP. Indicated replicon genomes were electroporated into Huh7 cells and luciferase activity followed over 72 hrs. Luciferase activity is expressed as the fold increase over the signal at 4 hrs.

5.2.3 Specific in-cell labelling of SNAP/CLIP-NS5A

It was next critical to establish whether the inserted SNAP/CLIP domain was able to fold correctly and conjugate specifically to the appropriate substrate. As mentioned the SNAP and CLIP domains have only been previously utilised as either an N- or C-terminal tag. To investigate whether the SNAP and CLIP domains were still able to conjugate to their substrates they were labelled with substrates in which the benzylguanine/cytosine groups

were conjugated to the fluorophore TMR-Star, termed SNAP-TMR and CLIP-TMR respectively. The TMR-Star fluorophore is a derivative of 6-carboxytetramethylrhodamine and has an excitation/emission maxima at 554/580nm.

Huh7 cells were electroporated with mSGR-luc genomes containing either NS5A-SNAP or NS5A-CLIP, seeded into an 8-chambered microscope slide, $1cm^2$ tissue culture area (Ibidi) and incubated for 48 hrs under normal TC conditions. These 8-chambered slides allow for small volumes of cell labelling media to be used, as well as being suitable for mounting directly onto an inverted confocal microscope in subsequent live-cell imaging. The SNAP/CLIP cell labelling media was prepared by dissolving the SNAP-TMR and CLIP-TMR dye in DMSO to 1 mM before diluting to 5 µM in complete media without phenol red. At 48 hpe media was removed from cells which were washed twice in PBS before the addition of either the SNAP-TMR or CLIP-TMR labelling media. Cells were incubated in the presence of labelling media for 1 hr at 37 °C, 5 % CO_2 before a dye washout step where the labelling media was removed, cells washed with PBS and incubated for a further 30 mins in fresh complete media without phenol red. The dye washout step is required to allow any unconjugated dye to diffuse out of the cells and so increasing the signal to background. At this stage cells were able to be imaged directly on a confocal microscope, but to allow for co-labelling by immunofluorescence (IF) cells were fixed in 4 % (w/v) PFA and immunostained, described previously. In brief, cells were permeabilised with 0.1 % (v/v) Triton X-100, PBS for 7 mins at RT before washing and then incubating with α-NS5A (sheep) at 1:1000 for 1 hr at RT. Cells then washed and incubated with a Donkey α-Sheep (594 nm) fluorescent antibody followed by staining of the nucleus with DAPI via the addition of ProLong gold mounting media (Life Technologies).

Prepared slides of both the mSGR-luc NS5A-SNAP and NS5A-CLIP replicons were imaged on an LSM 700 using the same laser, optic and acquisition settings, allowing for broad comparison between images to be made. To examine whether the SNAP-TMR or CLIP-TMR were non-specifically labelling cellular material Huh7 cells harbouring the WT SGR-luc were labelled with both SNAP- and CLIP-TMR and were imaged as described. As can be seen in both Figure 5.5 and 5.6 the replicons harbouring the tagged NS5A-SNAP and NS5A-CLIP were both able to successfully be labelled by the appropriate dye. This confirms that the mSGR-luc replicon is still viable with presence of these tags, as shown by luciferase assay, and that the SNAP and CLIP domains are still able to conjugate with the relevant substrates.

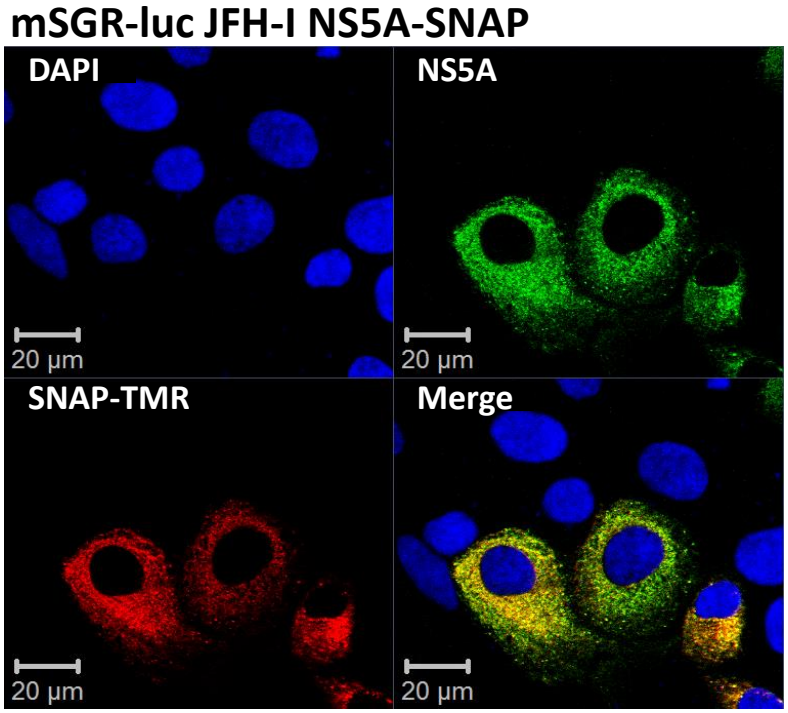


Figure 5.5. Immuno-fluorescence analysis of mSGR-luc NS5A-SNAP replicon.
Huh7 cells were electroporated with mSGR-luc NS5A-SNAP RNA and at 48 hpe labelled with the SNAP-TMR (red) before fixing in PFA and immunostaining for NS5A (green), nucleus labelled with DAPI (blue).

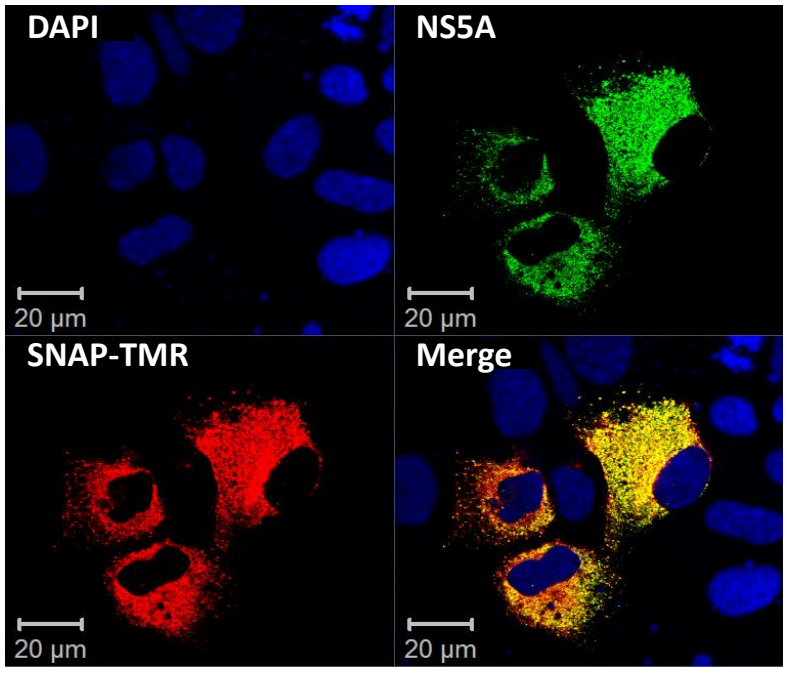


Figure 5.6. Immuno-fluorescence analysis of mSGR-luc NS5A-CLIP replicon.
Huh7 cells were electroporated with mSGR-luc NS5A-CLIP RNA and at 48 hpe labelled with the CLIP-TMR (red) before fixing and immunostaining for NS5A (green), nucleus labelled with DAPI (blue).

It was important to confirm that the SNAP/CLIP fluorophores were not non-specifically labelling Huh7 cells, for example through accumulating in lipid rich compartments. As such Huh7 cells were electroporated with WT mSGR-luc replicons and subsequently labelled with both SNAP-TMR and CLIP-TMR dyes simultaneously and imaged as described previously. From Figure 5.7 it is clear that Huh7 cells, neither with or without the replicon (indicated by NS5A staining), were non-specifically labelled by either the SNAP-TMR or CLIP-TMR to a detectable degree.

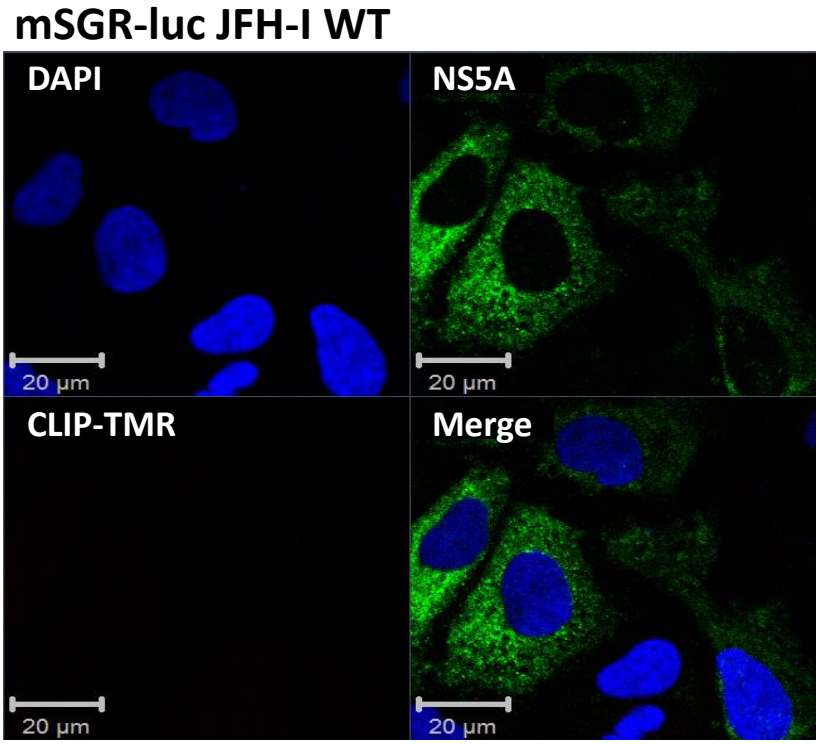


Figure 5.7. WT mSGR-luc replicon labelled with both SNAP-TMR and CLIP-TMR.
Huh7 cells were electroporated with WT mSGR-luc genomes and at 48 hpe labelled with both SNAP-TMR and CLIP-TMR (red) before fixing in PFA and immune-staining for NS5A (green), nucleus labelled with DAPI (blue)

While the insertion of the SNAP/CLIP domain was not detrimental to the function of NS5A it was important to establish whether the presence of the tag altered the sub-cellular distribution of NS5A. Comparison of the NS5A distribution in the tagged replicons (Figures 5.5 and 5.6) with that of the WT replicon (Figure 5.7) showed a broadly comparable localisation. However, to demonstrate this in a more comparably manner, Huh7 cells were electroporated with either the WT or the NS5A-SNAP containing replicon separately, then post-electroporation cells were co-seeded 1:1 into 8-chamered slides. Cells were then incubated for 48 hrs, labelled with SNAP-TMR, fixed and immunostained with α-NS5A (sheep) followed by a donkey α-sheep (594nm) secondary as described previously (Section 2.7.3). Cells were then stained with DAPI by the addition of ProLong Gold and imaged on a Zeiss LSM 700.

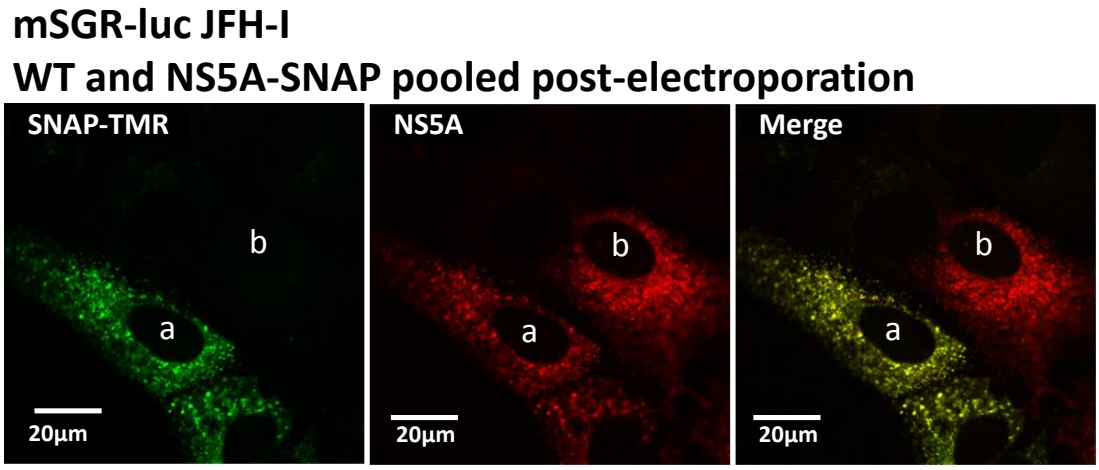


Figure 5.8. Comparison of sub-cellular localisation of NS5A-SNAP vs WT NS5A.
Huh7 cells were electroporated with either mSGR-luc NS5A-SNAP or WT mSGR-luc RNA separately, then co-seeded and incubated for 48 hrs. Cells were labelled with SNAP-TMR (red) before fixing and immunostaining for α-NS5A (green), nucleus labelled with DAPI (blue). Field of view shows cells harbouring with the NS5A-SNAP containing (a) or WT (b) replicons.

Figure 5.8 shows Huh7 cells harbouring either the NS5A-SNAP containing or WT replicons, cell (a) and (b) respectively. The identification was based on the reactivity towards both α-NS5A and the SNAP-TMR dye, or reactivity towards only the α-NS5A staining. In the cell harbouring the NS5A-SNAP containing replicon mSGR-luc NS5A-SNAP there is a high degree of co-localisation between the SNAP-TMR dye and the α-NS5A IF staining, cell (a) in Figure 5.8. This shows that there was a high efficiency of SNAP-TMR labelling, and that NS5A-SNAP is not localised to compartments and/or protein complexes that preclude the SNAP-tag from being labelling with the SNAP-TMR substrate. Comparison between the sub-cellular distribution of NS5A in the NS5A-SNAP or WT replicon, cells (a) and (b), shows a very comparable sub-cellular localisation. This suggests the presence of the SNAP/CLIP domain is not dramatically altering the distribution of NS5A. However, it is acknowledged that further co-labelling with virus proteins like NS3, NS5B and host compartments like lipid droplets and ER markers would be required to confirm that WT NS5A and NS5A-SNAP/CLIP have the same sub-cellular localisation.

5.2.4 Inserting SNAP/CLIP tags into replicons containing phosphorylation site mutations.

The next step in the project was the introduction of the SNAP/CLIP tag into replicons where the major phosphorylation sites, residues 146, 222, 225 and 348, had previously been mutated. This would allow further investigation as described in the introduction. The SNAP-tag domain III was introduced into the phosphoablatant replicons, while the CLIP-tagged domain III was introduced into the phosphomimetic replicons. In combination with the WT replicon tagged with both SNAP and CLIP-tags, this gave the flexibility to differentially label WT from phosphoablatant or phosphomimetic, as well as being able to differentially label between phosphoablatant from phosphomimetic.

The SNAP/CLIP-tag containing domain III was cloned into mSGR-luc JFH-1 constructs containing the above phospho-mutants by the flanking restriction sites *RsrII* and *AfeI*. The *RsrII* restriction site is located at position 398 (JFH-1 numbering), as such it is located inbetween the phosphorylation sites (146-348) and the site of the SNAP/CLIP insertion (residue 430), therefore allowing for this simple cloning route. Correct insertion of the SNAP/CLIP tag in each construct was confirmed by DNA sequencing. In the case of the SS222/225AA mutation a SNAP-tag was not inserted as the double serine mutation had been previously shown to be lethal to replication.

Once these constructs were created a replication assay was utilised to confirm that the phenotype of the various phosphorylation mutants now containing SNAP/CLIP-tags were the same as had been previously observed (Section 3.2.2). Huh7 cells were electroporated with the denoted tagged mutant constructs and luciferase activity determined at 72 hpe. Phenotypes of the S->A SNAP-tag and S->D CLIP-tag replicons were broadly comparable to those observed previously (Figure 5.9). Due to time constraints on this project the experiment was only conducted to an n=1.

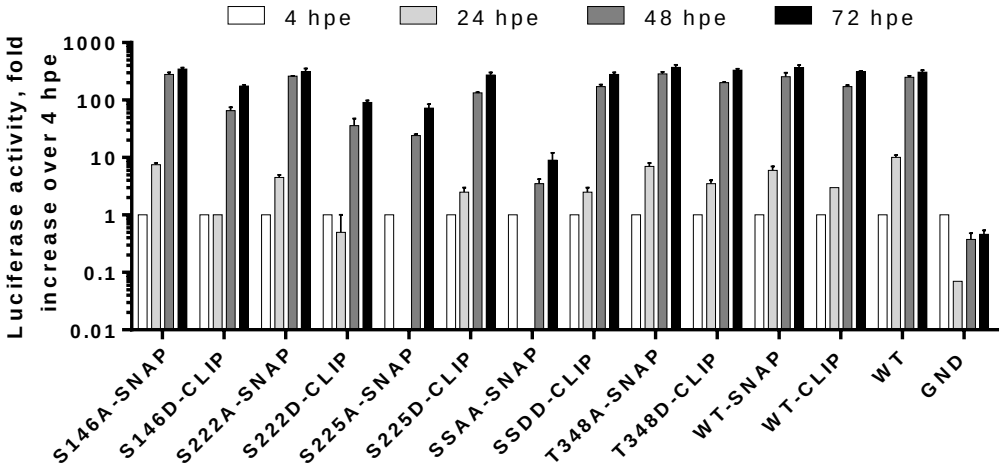


Figure 5.9. Replication analysis of major phosphomutants containing the NS5A-SNAP/CLIP domain.

Huh7 cells were electroporated with the denoted genomes before determining luciferase activity at 4, 24, 48 and 72 hpe, as described previously

5.3 Discussion

This project aimed to develop a novel way of visualising the NS5A protein by confocal microscopy that would open up new experimental space. This system also aimed to allow greater flexibility in the live-imaging of NS5A by allowing for interchangeable fluorophores without the need for lengthy cloning. This was achieved by adopting the SNAP-tag technology (NEB), a small (20 kDa) protein domain that can covalently conjugate a fluorescent substrate under physiological conditions. Such a system allows for a tagged protein to be labelled at a variety of different wavelengths by simply incubating live cells with fluorescent substrates of different wavelengths. This system also facilitates pseudo pulse-chase imaging approaches discussed shortly.

Previously the SNAP-tag technology has only been deployed as an N- or C-terminal protein tag. However, due to the nature of NS5A discussed earlier this was not possible, meaning the SNAP-tag had to be inserted internally to the protein. While the NS5A protein has previously been shown to tolerate large insertions, Figure 5.2, the SNAP-tag has not previously been tested in this kind of application. It was hoped that the unstructured nature of NS5A domain III and the introduction of flexible linker sequences either side, would allow for correct folding of the SNAP-tag and subsequent of catalytic activity. The SNAP and CLIP-tag were inserted into domain III of NS5A at residue 430 (JFH-1 numbering), described as site

V in figure 5.2, which is within the naturally occurring 19 aa insertion found in many genotype 2 isolates. This site was chosen as it has previously been shown to tolerate the insertion of a biotinylated motif with no observable effect on the replication or production of the JFH-1 virus (Igloi and Harris, in preparation).

The insertion of SNAP/CLIP-tag was well tolerated with respect to replication, with both NS5A-SNAP and NS5A-CLIP being indistinguishable from WT (Figure 5.4). It was also demonstrated that the SNAP and CLIP domains are still catalytically active when inserted internally to a protein, as they were able to conjugate to the relevant substrate with a high efficiency (Figure 5.5 and 5.6).

This validation has also shown that in the NS5A-SNAP/CLIP replicons there was near complete co-localisation between the SNAP-TMR and α-NS5A staining, showing that the entire NS5A population within a cell is accessible to the SNAP-TMR dye. This was expected as the SNAP and CLIP series of dyes utilised were designed for live-cell imaging and as such are small, lipophilic and cell permeable. When a cell containing the NS5A-SNAP replicon was compared to an adjacent cell containing the WT replicon the sub-cellular localisation of NS5A was highly similar, Figure 5.8. However further co-localisation with viral and host proteins would be required to say definitively that the presence of the SNAP/CLIP-tag does not disrupts the localisation of NS5A.

Taken together these data supports the use of SNAP/CLIP tagged NS5A as a novel model for WT NS5A in the replicon system. This system was applied to investigate the role of phosphorylation in the sub-cellular localisation of NS5A. The SNAP/CLIP-tagged NS5A was inserted into replicons harbouring major phosphorylation mutations, and it was confirmed that those replicons had the same phenotype as in the absence of the tag, Figure 5.9. However due to times constraints on this project it was not progressed further.

Therefore this project currently stands as proof of principle for the application of the SNAP-tag technology to NS5A and has created the reagents for future projects. Two key experiments that it is felt should be prioritised are the following. Firstly, the co-electroporation of phospho-ablatant/mimetic replicons containing the SNAP-tag with WT replicon containing the CLIP-tag. The subsequent co-labelling with different wavelength SNAP and CLIP substrates would allow for subtle differences in both localisation and live cell trafficking of NS5A to be determined. Such an approach could highlight whether a particular

phosphorylation site is associated with the static or highly mobile populations of NS5A discussed earlier. Secondly, a pseudo pulse-chase experiment, whereby electroporation of a NS5A-SNAP containing replicon, either with or without other mutations, could be labelled with different wavelength SNAP substrates at different time points. This could help identify different temporal populations of NS5A as well as being combined with drug treatments that are known to affect the NS5A, such as the PI4KIIIα inhibitor like AL-9 or the NS5A inhibitor Daclatasvir (BMS-790025) (Bianco et al., 2012, Targett-Adams et al., 2011).

Chapter 6: Conclusion and future perspectives

The NS5A protein has been well established as a pleiotropic protein with functions throughout the virus lifecycle, but despite 20 years of research the mechanism behind these functions remains enigmatic. In achieving these spatially and temporally distant functions in the virus lifecycle NS5A has been documented to interact with some 130 proteins, viral and cellular. It is therefore a clear requirement that such a promiscuous protein be tightly regulated and, in common with other virus and cellular proteins, this regulation is in the form of extensive and complex phosphorylation events.

Previous efforts to investigate this have been hampered by a lack of understanding about the location of phosphorylation sites. Although much work has implicated several clusters of serines as important for phosphorylation, the exact location of NS5A phosphorylation sites had yet to be defined. This study set out to obtain biochemical evidence in the form of mass spectrometry as to the location of these phosphorylation sites. As each phosphorylation site in itself probably represents only a small percentage of the total NS5A population, similar attempts to achieve this previously have been met with only limited success. Therefore in this study significant quantities of NS5A were purified from replicon harbouring cell lines and analysed by several orthogonal MS approaches. As a result a total of 12 functional relevant phosphorylation sites were identified spanning almost every domain of the protein. The subsequent functional analysis through reverse genetics highlighted several phosphorylation events as critical for virus replication and shed light on a complex mechanism of phosphorylation regulation.

The data presented here shows good evidence for a rapid phosphorylation cascade occurring in the LCS I of NS5A. This was observed as a step-wise increase in the apparent molecular weight of the basal phosphorylated species whereby the further C-terminal the phosphomimetic the greater the increase in apparent molecular weight. It is thought to be a rapidly occurring cascade as no intermediary species are observed and while the data suggested a linear cascade, a more complex mechanism like that observed in PTEN is at this stage equally feasible (Cordier et al., 2012). Exactly what mechanism in the virus life cycle an extensively phosphorylated LCS I is important for remains unclear, and that the requirement for phosphorylation is genotype dependent inevitable adds further complexities. It is plausible that a highly phosphorylated LCS I, with the resulting clustering of negative charge, could compete with viral RNA for binding to the NS5A, or other, protein. Indeed the proximity of the LCS I to the proposed RNA binding groove is a prime candidate for such a mechanism (Tellinghuisen et al., 2005).

The observation that the phosphorylation of serine 146 in domain I significantly reduces NS5A hyperphosphorylation points towards a new and exciting mechanism of regulation. While it is almost certain that there are phosphorylation sites outside of the LCS I that are important for hyperphosphorylation, the LCS I is at a minimum the major component. Through the use of a phosphospecific antibody it has been clearly demonstrate here that the phosphomimetic S146D specifically inhibits the phosphorylation of serine 222, and that phosphorylation of S222 is a hallmark of LCS I hyperphosphorylation. A combination of phosphomimetic mutations in the LCS I with the S146D mutation further support a relationship between these distant phosphorylation sites as serines in the LCS I not important for replication where, when in combination with S146D, significantly impairing replication. While circumstantial, the close proximity of a phosphorylation at residue 146 to the NS5A dimer interface does hint towards a possible involvement in switching between the two documented dimer conformations.

In pursuit of a greater insight into how phosphorylation might regulate the function of NS5A this study also took the SNAP-tag technology and applied it to NS5A. The insertion of the SNAP- and CLIP- tag into domain III of NS5A was shown to be tolerable with respect to both HCV replication and importantly the activity of the SNAP/CLIP-tag activity. This now serves as a versatile platform for both the live-cell imaging of cells with differentially labelled replicon proteins, as well as allowing interesting pulse-chase experiments. The demonstration by Fridell *et al.* that NS5A can be segregated into two different functional populations within a cell encourages a model whereby NS5A might also be spatially localised into different populations as well (Fridell et al., 2011, Fridell et al., 2013).

In parallel to investigating the phosphorylation of NS5A this study also set out to fully establish whether domain II of NS5A had a hitherto unknown function in the production of infectious virus, however no such effects were observed. In combination with existing data it is now clear that there is no residue with NS5A domain II that is essential for the production of infectious virus, and that any reduction in released virus correlates to an impairment of replication. Furthermore, the alanine mutations in the C-terminus of domain II impaired replication to varying degree, when this range of replication phenotypes was correlated with impairment to virus release there was a strong linear relationship. This supports recent modelling which identified replication as the rate limiting step in the production of infectious virus (Binder et al., 2013)

When the requirement of residues for replication in genotype 2a was compared with genotype 1b there was, despite high sequence conservation, very little correlation between phenotypes from the different genotypes. In fact one of the very few conserved features was the absolute requirement in both genotypes of the conserved GC337/338 for replication. These residues overlap with a previously identified VAP-B binding site (residues 337-340) and this could point towards how mutation of these residues disrupts replication (Hamamoto et al., 2005). A logical progression for this project would be to screen lethal domain II mutations for the loss of interaction with proteins already known to interact with this region. The precedence in the literature would suggest CypA, NS5B, VAP-B and HCV RNA as prime candidates, but through modern mammalian 2-hybrid system a much larger panel of proteins could be rapidly screened.

Chapter 7: References

AGNELLO, V., ABEL, G., ELFAHAL, M., KNIGHT, G. B. & ZHANG, Q. X. 1999. Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proc Natl Acad Sci U S A,* 96**,** 12766-71.

ALIGO, J., JIA, S., MANNA, D. & KONAN, K. V. 2009. Formation and function of hepatitis C virus replication complexes require residues in the carboxy-terminal domain of NS4B protein. *Virology,* 393**,** 68-83.

ALTER, M. J., KRUSZON-MORAN, D., NAINAN, O. V., MCQUILLAN, G. M., GAO, F., MOYER, L. A., KASLOW, R. A. & MARGOLIS, H. S. 1999. The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. *N Engl J Med,* 341**,** 556-62.

AMAKO, Y., IGLOI, Z., MANKOURI, J., KAZLAUSKAS, A., SAKSELA, K., DALLAS, M., PEERS, C. & HARRIS, M. 2013. Hepatitis C Virus NS5A Inhibits Mixed Lineage Kinase 3 to Block Apoptosis. *J Biol Chem,* 288**,** 24753-63.

AMAKO, Y., SARKESHIK, A., HOTTA, H., YATES, J., 3RD & SIDDIQUI, A. 2009. Role of oxysterol binding protein in hepatitis C virus infection. *J Virol,* 83**,** 9237-46.

ANDRE, P., KOMURIAN-PRADEL, F., DEFORGES, S., PERRET, M., BERLAND, J. L., SODOYER, M., POL, S., BRECHOT, C., PARANHOS-BACCALA, G. & LOTTEAU, V. 2002. Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. *J Virol,* 76**,** 6919-28.

APPEL, N., PIETSCHMANN, T. & BARTENSCHLAGER, R. 2005. Mutational analysis of hepatitis C virus nonstructural protein 5A: potential role of differential phosphorylation in RNA replication and identification of a genetically flexible domain. *J Virol,* 79**,** 3187-94.

APPEL, N., ZAYAS, M., MILLER, S., KRIJNSE-LOCKER, J., SCHALLER, T., FRIEBE, P., KALLIS, S., ENGEL, U. & BARTENSCHLAGER, R. 2008. Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. *PLoS Pathog,* 4**,** e1000035.

ASABE, S. I., TANJI, Y., SATOH, S., KANEKO, T., KIMURA, K. & SHIMOTOHNO, K. 1997. The N-terminal region of hepatitis C virus-encoded NS5A is important for NS4A-dependent phosphorylation. *J Virol,* 71**,** 790-6.

ATOOM, A. M., JONES, D. M. & RUSSELL, R. S. 2013. Evidence suggesting that HCV p7 protects E2 glycoprotein from premature degradation during virus production. *Virus Res,* 176**,** 199-210.

ATTWOOD, P. V., BESANT, P. G. & PIGGOTT, M. J. 2011. Focus on phosphoaspartate and phosphoglutamate. *Amino Acids,* 40**,** 1035-51.

BACKES, P., QUINKERT, D., REISS, S., BINDER, M., ZAYAS, M., RESCHER, U., GERKE, V., BARTENSCHLAGER, R. & LOHMANN, V. 2010. Role of annexin A2 in the production of infectious hepatitis C virus particles. *J Virol,* 84**,** 5775-89.

BALACHANDRAN, S., ROBERTS, P. C., BROWN, L. E., TRUONG, H., PATTNAIK, A. K., ARCHER, D. R. & BARBER, G. N. 2000. Essential role for the dsRNA-dependent protein kinase PKR in innate immunity to viral infection. *Immunity,* 13**,** 129-41.

BAMBERGER, M. J. & LANE, M. D. 1990. Possible role of the Golgi apparatus in the assembly of very low density lipoprotein. *Proc Natl Acad Sci U S A,* 87**,** 2390-4.

BARREIRO, P., VISPO, E., POVEDA, E., FERNANDEZ-MONTERO, J. V. & SORIANO, V. 2013. Hepatitis C therapy: highlights from the 2012 annual meeting of the European Association for the Study of the Liver. *Clin Infect Dis,* 56**,** 560-6.

BARTENSCHLAGER, R., LOHMANN, V. & PENIN, F. 2013. The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. *Nat Rev Microbiol,* 11**,** 482-96.

BARTENSCHLAGER, R., PENIN, F., LOHMANN, V. & ANDRE, P. 2011. Assembly of infectious hepatitis C virus particles. *Trends Microbiol,* 19**,** 95-103.

BARTENSCHLAGER, R. & SPARACIO, S. 2007. Hepatitis C virus molecular clones and their replication capacity in vivo and in cell culture. *Virus Res,* 127**,** 195-207.

BARTH, H., SCHAFER, C., ADAH, M. I., ZHANG, F., LINHARDT, R. J., TOYODA, H., KINOSHITA-TOYODA, A., TOIDA, T., VAN KUPPEVELT, T. H., DEPLA, E., VON WEIZSACKER, F., BLUM, H. E. & BAUMERT, T. F. 2003. Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J Biol Chem,* 278**,** 41003-12.

BARTH, S. 2013. Editorial (The SNAP-tag Technology - A Versatile Tool with many Applications). *Curr Pharm Des,* 19**,** 5404-5.

BASSETT, S. E., BRASKY, K. M. & LANFORD, R. E. 1998. Analysis of hepatitis C virus-inoculated chimpanzees reveals unexpected clinical profiles. *J Virol,* 72**,** 2589-99.

BENGA, W. J., KRIEGER, S. E., DIMITROVA, M., ZEISEL, M. B., PARNOT, M., LUPBERGER, J., HILDT, E., LUO, G., MCLAUCHLAN, J., BAUMERT, T. F. & SCHUSTER, C. 2010. Apolipoprotein E interacts with hepatitis C virus nonstructural protein 5A and determines assembly of infectious particles. *Hepatology,* 51**,** 43-53.

BENTHAM, M. J., FOSTER, T. L., MCCORMICK, C. & GRIFFIN, S. 2013. Mutations in Hepatitis C virus p7 reduce both the egress and infectivity of assembled particles via impaired proton channel function. *J Gen Virol*, 94, 2236-48.

BERGER, K. L., COOPER, J. D., HEATON, N. S., YOON, R., OAKLAND, T. E., JORDAN, T. X., MATEU, G., GRAKOUI, A. & RANDALL, G. 2009. Roles for endocytic trafficking and phosphatidylinositol 4-kinase III alpha in hepatitis C virus replication. *Proc Natl Acad Sci U S A,* 106**,** 7577-82.

BERGER, K. L. & RANDALL, G. 2009. Potential roles for cellular cofactors in hepatitis C virus replication complex formation. *Commun Integr Biol,* 2**,** 471-3.

BERNSMEIER, C., DUONG, F. H., CHRISTEN, V., PUGNALE, P., NEGRO, F., TERRACCIANO, L. & HEIM, M. H. 2008. Virus-induced over-expression of protein phosphatase 2A inhibits insulin signalling in chronic hepatitis C. *J Hepatol,* 49**,** 429-40.

BESANT, P. G., ATTWOOD, P. V. & PIGGOTT, M. J. 2009. Focus on phosphoarginine and phospholysine. *Curr Protein Pept Sci,* 10**,** 536-50.

BIANCO, A., REGHELLIN, V., DONNICI, L., FENU, S., ALVAREZ, R., BARUFFA, C., PERI, F., PAGANI, M., ABRIGNANI, S., NEDDERMANN, P. & DE FRANCESCO, R. 2012. Metabolism of phosphatidylinositol 4-kinase IIIalpha-dependent PI4P Is subverted by HCV and is targeted by a 4-anilino quinazoline with antiviral activity. *PLoS Pathog,* 8**,** e1002576.

BINDER, M., SULAIMANOV, N., CLAUSZNITZER, D., SCHULZE, M., HUBER, C. M., LENZ, S. M., SCHLODER, J. P., TRIPPLER, M., BARTENSCHLAGER, R., LOHMANN, V. & KADERALI, L. 2013. Replication Vesicles are Load- and Choke-Points in the Hepatitis C Virus Lifecycle. *PLoS Pathog,* 9**,** e1003561.

BLANCHARD, E., BELOUZARD, S., GOUESLAIN, L., WAKITA, T., DUBUISSON, J., WYCHOWSKI, C. & ROUILLE, Y. 2006. Hepatitis C virus entry depends on clathrin-mediated endocytosis. *J Virol,* 80**,** 6964-72.

BLIGHT, K. J., KOLYKHALOV, A. A. & RICE, C. M. 2000. Efficient initiation of HCV RNA replication in cell culture. *Science,* 290**,** 1972-4.

BLIGHT, K. J. & RICE, C. M. 1997. Secondary structure determination of the conserved 98-base sequence at the 3' terminus of hepatitis C virus genome RNA. *J Virol,* 71**,** 7345-52.

BOONSTRA, A., VAN DER LAAN, L. J., VANWOLLEGHEM, T. & JANSSEN, H. L. 2009. Experimental models for hepatitis C viral infection. *Hepatology,* 50**,** 1646-55.

BORAWSKI, J., TROKE, P., PUYANG, X., GIBAJA, V., ZHAO, S., MICKANIN, C., LEIGHTON-DAVIES, J., WILSON, C. J., MYER, V., CORNELLATARACIDO, I., BARYZA, J., TALLARICO, J., JOBERTY, G., BANTSCHEFF, M., SCHIRLE, M., BOUWMEESTER, T., MATHY, J. E., LIN, K., COMPTON, T., LABOW, M., WIEDMANN, B. & GAITHER, L. A. 2009. Class III phosphatidylinositol 4-kinase alpha and beta are novel host factor regulators of hepatitis C virus replication. *J Virol,* 83**,** 10058-74.

BOSON, B., GRANIO, O., BARTENSCHLAGER, R. & COSSET, F. L. 2011. A concerted action of hepatitis C virus p7 and nonstructural protein 2 regulates core localization at the endoplasmic reticulum and virus assembly. *PLoS Pathog,* 7**,** e1002144.

BOULANT, S., DOUGLAS, M. W., MOODY, L., BUDKOWSKA, A., TARGETT-ADAMS, P. & MCLAUCHLAN, J. 2008. Hepatitis C virus core protein induces lipid droplet redistribution in a microtubule- and dynein-dependent manner. *Traffic,* 9**,** 1268-82.

BOULANT, S., MONTSERRET, R., HOPE, R. G., RATINIER, M., TARGETT-ADAMS, P., LAVERGNE, J. P., PENIN, F. & MCLAUCHLAN, J. 2006. Structural determinants that target the hepatitis C virus core protein to lipid droplets. *J Biol Chem,* 281**,** 22236-47.

BOULANT, S., TARGETT-ADAMS, P. & MCLAUCHLAN, J. 2007. Disrupting the association of hepatitis C virus core protein with lipid droplets correlates with a loss in production of infectious virus. *J Gen Virol,* 88**,** 2204-13.

BOULANT, S., VANBELLE, C., EBEL, C., PENIN, F. & LAVERGNE, J. P. 2005. Hepatitis C virus core protein is a dimeric alpha-helical protein exhibiting membrane protein features. *J Virol,* 79**,** 11353-65.

BOWEN, D. G. & WALKER, C. M. 2005. Adaptive immune responses in acute and chronic hepatitis C virus infection. *Nature,* 436**,** 946-52.

BRASS, V., BIECK, E., MONTSERRET, R., WOLK, B., HELLINGS, J. A., BLUM, H. E., PENIN, F. & MORADPOUR, D. 2002. An amino-terminal amphipathic alpha-helix mediates membrane association of the hepatitis C virus nonstructural protein 5A. *J Biol Chem,* 277**,** 8130-9.

BRAZZOLI, M., BIANCHI, A., FILIPPINI, S., WEINER, A., ZHU, Q., PIZZA, M. & CROTTA, S. 2008. CD81 is a central regulator of cellular events required for hepatitis C virus infection of human hepatocytes. *J Virol,* 82**,** 8316-29.

BUKH, J., PIETSCHMANN, T., LOHMANN, V., KRIEGER, N., FAULK, K., ENGLE, R. E., GOVINDARAJAN, S., SHAPIRO, M., ST CLAIRE, M. & BARTENSCHLAGER, R. 2002. Mutations that permit efficient replication of hepatitis C virus RNA in Huh-7 cells prevent productive replication in chimpanzees. *Proc Natl Acad Sci U S A,* 99**,** 14416-21.

CHANG, K. S., JIANG, J., CAI, Z. & LUO, G. 2007. Human apolipoprotein e is required for infectivity and production of hepatitis C virus in cell culture. *J Virol,* 81**,** 13783-93.

CHATTERJI, U., BOBARDT, M., SELVARAJAH, S., YANG, F., TANG, H., SAKAMOTO, N., VUAGNIAUX, G., PARKINSON, T. & GALLAY, P. 2009. The isomerase active site of cyclophilin A is critical for hepatitis C virus replication. *J Biol Chem,* 284**,** 16998-7005.

CHATTERJI, U., LIM, P., BOBARDT, M. D., WIELAND, S., CORDEK, D. G., VUAGNIAUX, G., CHISARI, F., CAMERON, C. E., TARGETT-ADAMS, P., PARKINSON, T. & GALLAY, P. A. 2010. HCV resistance to cyclosporin A does not correlate with a resistance of the NS5A-cyclophilin A interaction to cyclophilin inhibitors. *J Hepatol,* 53**,** 50-6.

CHEN, K. X. & NJOROGE, F. G. 2010. The Journey to the Discovery of Boceprevir: an NS3-NS4 HCV protease inhibitor for the treatment of chronic hepatitis C. *Prog Med Chem,* 49**,** 1-36.

CHEN, S. L. & MORGAN, T. R. 2006. The natural history of hepatitis C virus (HCV) infection. *Int J Med Sci,* 3**,** 47-52.

CHEN, Y. C., SU, W. C., HUANG, J. Y., CHAO, T. C., JENG, K. S., MACHIDA, K. & LAI, M. M. 2010. Polo-like kinase 1 is involved in hepatitis C virus replication by hyperphosphorylating NS5A. *J Virol,* 84**,** 7983-93.

CHEW, C. F., VIJAYAN, R., CHANG, J., ZITZMANN, N. & BIGGIN, P. C. 2009. Determination of pore-lining residues in the hepatitis C virus p7 protein. *Biophys J,* 96**,** L10-2.

CHOI, K. H. & ROSSMANN, M. G. 2009. RNA-dependent RNA polymerases from Flaviviridae. *Curr Opin Struct Biol,* 19**,** 746-51.

COELMONT, L., HANOULLE, X., CHATTERJI, U., BERGER, C., SNOECK, J., BOBARDT, M., LIM, P., VLIEGEN, I., PAESHUYSE, J., VUAGNIAUX, G., VANDAMME, A. M., BARTENSCHLAGER,

R., GALLAY, P., LIPPENS, G. & NEYTS, J. 2010. DEB025 (Alisporivir) inhibits hepatitis C virus replication by preventing a cyclophilin A induced cis-trans isomerisation in domain II of NS5A. *PLoS One,* 5**,** e13687.

COITO, C., DIAMOND, D. L., NEDDERMANN, P., KORTH, M. J. & KATZE, M. G. 2004. High-throughput screening of the yeast kinome: identification of human serine/threonine protein kinases that phosphorylate the hepatitis C virus NS5A protein. *J Virol,* 78**,** 3502-13.

COLLER, K. E., BERGER, K. L., HEATON, N. S., COOPER, J. D., YOON, R. & RANDALL, G. 2009. RNA interference and single particle tracking analysis of hepatitis C virus endocytosis. *PLoS Pathog,* 5**,** e1000702.

CORDIER, F., CHAFFOTTE, A., TERRIEN, E., PREHAUD, C., THEILLET, F. X., DELEPIERRE, M., LAFON, M., BUC, H. & WOLFF, N. 2012. Ordered phosphorylation events in two independent cascades of the PTEN C-tail revealed by NMR. *J Am Chem Soc,* 134**,** 20533-43.

COYNE, C. B. & BERGELSON, J. M. 2006. Virus-induced Abl and Fyn kinase signals permit coxsackievirus entry through epithelial tight junctions. *Cell,* 124**,** 119-31.

CUN, W., JIANG, J. & LUO, G. 2010. The C-terminal alpha-helix domain of apolipoprotein E is required for interaction with nonstructural protein 5A and assembly of hepatitis C virus. *J Virol,* 84**,** 11532-41.

DE CHASSEY, B., NAVRATIL, V., TAFFOREAU, L., HIET, M. S., AUBLIN-GEX, A., AGAUGUE, S., MEIFFREN, G., PRADEZYNSKI, F., FARIA, B. F., CHANTIER, T., LE BRETON, M., PELLET, J., DAVOUST, N., MANGEOT, P. E., CHABOUD, A., PENIN, F., JACOB, Y., VIDALAIN, P. O., VIDAL, M., ANDRE, P., RABOURDIN-COMBE, C. & LOTTEAU, V. 2008. Hepatitis C virus infection protein network. *Mol Syst Biol,* 4**,** 230.

DELANG, L., PAESHUYSE, J. & NEYTS, J. 2012. The role of phosphatidylinositol 4-kinases and phosphatidylinositol 4-phosphate during viral replication. *Biochem Pharmacol,* 84**,** 1400-8.

DIMITROVA, M., IMBERT, I., KIENY, M. P. & SCHUSTER, C. 2003. Protein-protein interactions between hepatitis C virus nonstructural proteins. *J Virol,* 77**,** 5401-14.

DORNER, M., HORWITZ, J. A., DONOVAN, B. M., LABITT, R. N., BUDELL, W. C., FRILING, T., VOGT, A., CATANESE, M. T., SATOH, T., KAWAI, T., AKIRA, S., LAW, M., RICE, C. M. & PLOSS, A. 2013. Completion of the entire hepatitis C virus life cycle in genetically humanized mice. *Nature*, 501(7466), 237-41.

DYSON, H. J. & WRIGHT, P. E. 2005. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol,* 6**,** 197-208.

EINAV, S., GERBER, D., BRYSON, P. D., SKLAN, E. H., ELAZAR, M., MAERKL, S. J., GLENN, J. S. & QUAKE, S. R. 2008. Discovery of a hepatitis C target and its pharmacological inhibitors by microfluidic affinity analysis. *Nat Biotechnol,* 26**,** 1019-27.

ENOMOTO, N., SAKUMA, I., ASAHINA, Y., KUROSAKI, M., MURAKAMI, T., YAMAMOTO, C., IZUMI, N., MARUMO, F. & SATO, C. 1995. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *J Clin Invest,* 96**,** 224-30.

EVANS, M. J., RICE, C. M. & GOFF, S. P. 2004. Phosphorylation of hepatitis C virus nonstructural protein 5A modulates its protein interactions and viral RNA replication. *Proc Natl Acad Sci U S A,* 101**,** 13038-43.

EVANS, M. J., VON HAHN, T., TSCHERNE, D. M., SYDER, A. J., PANIS, M., WOLK, B., HATZIIOANNOU, T., MCKEATING, J. A., BIENIASZ, P. D. & RICE, C. M. 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature,* 446**,** 801-5.

EYRE, N. S., FICHES, G., ALOIA, A. L., HELBIG, K., MCERLEAN, C., TURVILLE, S. & BEARD, M. 2012. Dynamic trafficking of the Hepatitis C virus NS5A protein during an infectious life cycle. *19th International Symposium on Hepatitis C Virus and Related Viruses.* Venice, Italy.

FARCI, P., ALTER, H. J., SHIMODA, A., GOVINDARAJAN, S., CHEUNG, L. C., MELPOLDER, J. C., SACHER, R. A., SHIH, J. W. & PURCELL, R. H. 1996. Hepatitis C virus-associated fulminant hepatic failure. *N Engl J Med,* 335**,** 631-4.

FARIA, S. C., GANESAN, K., MWANGI, I., SHIEHMORTEZA, M., VIAMONTE, B., MAZHAR, S., PETERSON, M., KONO, Y., SANTILLAN, C., CASOLA, G. & SIRLIN, C. B. 2009. MR imaging of liver fibrosis: current state of the art. *Radiographics,* 29**,** 1615-35.

FARQUHAR, M. J., HU, K., HARRIS, H. J., DAVIS, C., BRIMACOMBE, C. L., FLETCHER, S. J., BAUMERT, T. F., RAPPOPORT, J. Z., BALFE, P. & MCKEATING, J. A. 2012. Hepatitis C virus induces CD81 and claudin-1 endocytosis. *J Virol,* 86**,** 4305-16.

FATTOVICH, G., GIUSTINA, G., DEGOS, F., DIODATI, G., TREMOLADA, F., NEVENS, F., ALMASIO, P., SOLINAS, A., BROUWER, J. T., THOMAS, H., REALDI, G., CORROCHER, R. & SCHALM, S. W. 1997a. Effectiveness of interferon alfa on incidence of hepatocellular carcinoma and decompensation in cirrhosis type C. European Concerted Action on Viral Hepatitis (EUROHEP). *J Hepatol,* 27**,** 201-5.

FATTOVICH, G., GIUSTINA, G., DEGOS, F., TREMOLADA, F., DIODATI, G., ALMASIO, P., NEVENS, F., SOLINAS, A., MURA, D., BROUWER, J. T., THOMAS, H., NJAPOUM, C., CASARIN, C., BONETTI, P., FUSCHI, P., BASHO, J., TOCCO, A., BHALLA, A., GALASSINI, R., NOVENTA, F., SCHALM, S. W. & REALDI, G. 1997b. Morbidity and mortality in compensated cirrhosis type C: a retrospective follow-up study of 384 patients. *Gastroenterology,* 112**,** 463-72.

FEINSTONE, S. M., HU, D. J. & MAJOR, M. E. 2012. Prospects for prophylactic and therapeutic vaccines against hepatitis C virus. *Clin Infect Dis,* 55 Suppl 1**,** S25-32.

FEUERSTEIN, S., SOLYOM, Z., ALADAG, A., FAVIER, A., SCHWARTEN, M., HOFFMANN, S., WILLBOLD, D. & BRUTSCHER, B. 2012. Transient structure and SH3 interaction sites in an intrinsically disordered fragment of the hepatitis C virus protein NS5A. *J Mol Biol,* 420**,** 310-23.

FISCHER, G., GALLAY, P. & HOPKINS, S. 2010. Cyclophilin inhibitors for the treatment of HCV infection. *Curr Opin Investig Drugs,* 11**,** 911-8.

FOSTER, T. L., BELYAEVA, T., STONEHOUSE, N. J., PEARSON, A. R. & HARRIS, M. 2010. All three domains of the hepatitis C virus nonstructural NS5A protein contribute to RNA binding. *J Virol,* 84**,** 9267-77.

FOSTER, T. L., GALLAY, P., STONEHOUSE, N. J. & HARRIS, M. 2011. Cyclophilin A interacts with domain II of hepatitis C virus NS5A and stimulates RNA binding in an isomerase-dependent manner. *J Virol,* 85**,** 7460-4.

FOSTER, T. L., THOMPSON, G. S., KALVERDA, A. P., KANKANALA, J., BENTHAM, M., WETHERILL, L. F., THOMPSON, J., BARKER, A. M., CLARKE, D., NOERENBERG, M., PEARSON, A. R., ROWLANDS, D. J., HOMANS, S. W., HARRIS, M., FOSTER, R. & GRIFFIN, S. 2013. Structure-guided design affirms inhibitors of hepatitis C virus p7 as a viable class of antivirals targeting virion release. *Hepatology*, doi 10.1002/hep.26685

FRANK, C., MOHAMED, M. K., STRICKLAND, G. T., LAVANCHY, D., ARTHUR, R. R., MAGDER, L. S., EL KHOBY, T., ABDEL-WAHAB, Y., ALY OHN, E. S., ANWAR, W. & SALLAM, I. 2000. The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet,* 355**,** 887-91.

FREEMAN, A., HAMID, S., MORRIS, L., VOWLER, S., RUSHBROOK, S., WIGHT, D. G., COLEMAN, N. & ALEXANDER, G. J. 2003. Improved detection of hepatocyte proliferation using antibody to the pre-replication complex: an association with hepatic fibrosis and viral replication in chronic hepatitis C virus infection. *J Viral Hepat,* 10**,** 345-50.

FRIDELL, R. A., QIU, D., VALERA, L., WANG, C., ROSE, R. E. & GAO, M. 2011. Distinct functions of NS5A in hepatitis C virus RNA replication uncovered by studies with the NS5A inhibitor BMS-790052. *J Virol,* 85**,** 7312-20.

FRIDELL, R. A., VALERA, L., QIU, D., KIRK, M. J., WANG, C. & GAO, M. 2013. Intragenic complementation of hepatitis C virus NS5A RNA replication-defective alleles. *J Virol,* 87**,** 2320-9.

FRIEBE, P. & BARTENSCHLAGER, R. 2002. Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. *J Virol,* 76**,** 5326-38.

GALE, M., JR., BLAKELY, C. M., KWIECISZEWSKI, B., TAN, S. L., DOSSETT, M., TANG, N. M., KORTH, M. J., POLYAK, S. J., GRETCH, D. R. & KATZE, M. G. 1998. Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. *Mol Cell Biol,* 18**,** 5208-18.

GALE, M. J., JR., KORTH, M. J., TANG, N. M., TAN, S. L., HOPKINS, D. A., DEVER, T. E., POLYAK, S. J., GRETCH, D. R. & KATZE, M. G. 1997. Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. *Virology,* 230**,** 217-27.

GAO, L., AIZAKI, H., HE, J. W. & LAI, M. M. 2004. Interactions between viral nonstructural proteins and host protein hVAP-33 mediate the formation of hepatitis C virus RNA replication complex on lipid raft. *J Virol,* 78**,** 3480-8.

GASTAMINZA, P., CHENG, G., WIELAND, S., ZHONG, J., LIAO, W. & CHISARI, F. V. 2008. Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *J Virol,* 82**,** 2120-9.

GASTAMINZA, P., DRYDEN, K. A., BOYD, B., WOOD, M. R., LAW, M., YEAGER, M. & CHISARI, F. V. 2010. Ultrastructural and biophysical characterization of hepatitis C virus particles produced in cell culture. *J Virol,* 84**,** 10999-1009.

GE, D., FELLAY, J., THOMPSON, A. J., SIMON, J. S., SHIANNA, K. V., URBAN, T. J., HEINZEN, E. L., QIU, P., BERTELSEN, A. H., MUIR, A. J., SULKOWSKI, M., MCHUTCHISON, J. G. & GOLDSTEIN, D. B. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature,* 461**,** 399-401.

GEORGOPOULOU, U., TSITOURA, P., KALAMVOKI, M. & MAVROMARA, P. 2006. The protein phosphatase 2A represents a novel cellular target for hepatitis C virus NS5A protein. *Biochimie,* 88**,** 651-62.

GHOSH, A. K., MAJUMDER, M., STEELE, R., MEYER, K., RAY, R. & RAY, R. B. 2000. Hepatitis C virus NS5A protein protects against TNF-alpha mediated apoptotic cell death. *Virus Res,* 67**,** 173-8.

GOTTWEIN, J. M., SCHEEL, T. K., JENSEN, T. B., LADEMANN, J. B., PRENTOE, J. C., KNUDSEN, M. L., HOEGH, A. M. & BUKH, J. 2009. 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,* 49**,** 364-77.

GRIFFIN, S. 2010. Inhibition of HCV p7 as a therapeutic target. *Curr Opin Investig Drugs,* 11**,** 175-81.

GRISE, H., FRAUSTO, S., LOGAN, T. & TANG, H. 2012. A conserved tandem cyclophilin-binding site in hepatitis C virus nonstructural protein 5A regulates Alisporivir susceptibility. *J Virol,* 86**,** 4811-22.

GSPONER, J. & BABU, M. M. 2009. The rules of disorder or why disorder rules. *Prog Biophys Mol Biol,* 99**,** 94-103.

GU, M. & RICE, C. M. 2013. Structures of hepatitis C virus nonstructural proteins required for replicase assembly and function. *Curr Opin Virol,* 3**,** 129-36.

GUEDJ, J., DAHARI, H., RONG, L., SANSONE, N. D., NETTLES, R. E., COTLER, S. J., LAYDEN, T. J., UPRICHARD, S. L. & PERELSON, A. S. 2013. Modeling shows that the NS5A inhibitor daclatasvir has two modes of action and yields a shorter estimate of the hepatitis C virus half-life. *Proc Natl Acad Sci U S A,* 110**,** 3991-6.

GUIDOTTI, L. G. & CHISARI, F. V. 2006. Immunobiology and pathogenesis of viral hepatitis. *Annu Rev Pathol,* 1**,** 23-61.

GUPTA, G., QIN, H. & SONG, J. 2012. Intrinsically unstructured domain 3 of hepatitis C Virus NS5A forms a "fuzzy complex" with VAPB-MSP domain which carries ALS-causing mutations. *PLoS One,* 7**,** e39261.

HAMAMOTO, I., NISHIMURA, Y., OKAMOTO, T., AIZAKI, H., LIU, M., MORI, Y., ABE, T., SUZUKI, T., LAI, M. M., MIYAMURA, T., MORIISHI, K. & MATSUURA, Y. 2005. Human VAP-B is involved in hepatitis C virus replication through interaction with NS5A and NS5B. *J Virol,* 79**,** 13473-82.

HAN, Q., MANNA, D., BELTON, K., COLE, R. & KONAN, K. V. 2013. Modulation of hepatitis C virus genome encapsidation by nonstructural protein 4B. *J Virol,* 87**,** 7409-22.

HANOULLE, X., BADILLO, A., WIERUSZESKI, J. M., VERDEGEM, D., LANDRIEU, I., BARTENSCHLAGER, R., PENIN, F. & LIPPENS, G. 2009a. Hepatitis C virus NS5A protein is a substrate for the peptidyl-prolyl cis/trans isomerase activity of cyclophilins A and B. *J Biol Chem,* 284**,** 13589-601.

HANOULLE, X., VERDEGEM, D., BADILLO, A., WIERUSZESKI, J. M., PENIN, F. & LIPPENS, G. 2009b. Domain 3 of non-structural protein 5A from hepatitis C virus is natively unfolded. *Biochem Biophys Res Commun,* 381**,** 634-8.

HARUKI, H., GONZALEZ, M. R. & JOHNSSON, K. 2012. Exploiting ligand-protein conjugates to monitor ligand-receptor interactions. *PLoS One,* 7**,** e37598.

HECK, J. A., MENG, X. & FRICK, D. N. 2009. Cyclophilin B stimulates RNA synthesis by the HCV RNA dependent RNA polymerase. *Biochem Pharmacol,* 77**,** 1173-80.

HELBIG, K. J., EYRE, N. S., YIP, E., NARAYANA, S., LI, K., FICHES, G., MCCARTNEY, E. M., JANGRA, R. K., LEMON, S. M. & BEARD, M. R. 2011. The antiviral protein viperin inhibits hepatitis C virus replication via interaction with nonstructural protein 5A. *Hepatology,* 54**,** 1506-17.

HELLE, F., GOFFARD, A., MOREL, V., DUVERLIE, G., MCKEATING, J., KECK, Z. Y., FOUNG, S., PENIN, F., DUBUISSON, J. & VOISSET, C. 2007. The neutralizing activity of anti-hepatitis C virus antibodies is modulated by specific glycans on the E2 envelope protein. *J Virol,* 81**,** 8101-11.

HERKER, E., HARRIS, C., HERNANDEZ, C., CARPENTIER, A., KAEHLCKE, K., ROSENBERG, A. R., FARESE, R. V., JR. & OTT, M. 2010. Efficient hepatitis C virus particle formation requires diacylglycerol acyltransferase-1. *Nat Med,* 16**,** 1295-8.

HONDA, M., PING, L. H., RIJNBRAND, R. C., AMPHLETT, E., CLARKE, B., ROWLANDS, D. & LEMON, S. M. 1996. Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA. *Virology,* 222**,** 31-42.

HOPKINS, S., SCORNEAUX, B., HUANG, Z., MURRAY, M. G., WRING, S., SMITLEY, C., HARRIS, R., ERDMANN, F., FISCHER, G. & RIBEILL, Y. 2010. SCY-635, a novel nonimmunosuppressive analog of cyclosporine that exhibits potent inhibition of hepatitis C virus RNA replication in vitro. *Antimicrob Agents Chemother,* 54**,** 660-72.

HUANG, H., SUN, F., OWEN, D. M., LI, W., CHEN, Y., GALE, M., JR. & YE, J. 2007. Hepatitis C virus production by human hepatocytes dependent on assembly and secretion of very low-density lipoproteins. *Proc Natl Acad Sci U S A,* 104**,** 5848-53.

HUANG, L., HWANG, J., SHARMA, S. D., HARGITTAI, M. R., CHEN, Y., ARNOLD, J. J., RANEY, K. D. & CAMERON, C. E. 2005. Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. *J Biol Chem,* 280**,** 36417-28.

HUANG, L., SINEVA, E. V., HARGITTAI, M. R., SHARMA, S. D., SUTHAR, M., RANEY, K. D. & CAMERON, C. E. 2004. Purification and characterization of hepatitis C virus non-structural protein 5A expressed in Escherichia coli. *Protein Expr Purif,* 37**,** 144-53.

HUGHES, M., GRETTON, S., SHELTON, H., BROWN, D. D., MCCORMICK, C. J., ANGUS, A. G., PATEL, A. H., GRIFFIN, S. & HARRIS, M. 2009a. A conserved proline between domains II and III of hepatitis C virus NS5A influences both RNA replication and virus assembly. *J Virol,* 83**,** 10788-96.

HUGHES, M., GRIFFIN, S. & HARRIS, M. 2009b. Domain III of NS5A contributes to both RNA replication and assembly of hepatitis C virus particles. *J Gen Virol,* 90**,** 1329-34.

HUWART, L. & VAN BEERS, B. E. 2008. MR elastography. *Gastroenterol Clin Biol,* 32**,** 68-72.

IAKOUCHEVA, L. M., RADIVOJAC, P., BROWN, C. J., O'CONNOR, T. R., SIKES, J. G., OBRADOVIC, Z. & DUNKER, A. K. 2004. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res,* 32**,** 1037-49.

ICARD, V., DIAZ, O., SCHOLTES, C., PERRIN-COCON, L., RAMIERE, C., BARTENSCHLAGER, R., PENIN, F., LOTTEAU, V. & ANDRE, P. 2009. Secretion of hepatitis C virus envelope glycoproteins depends on assembly of apolipoprotein B positive lipoproteins. *PLoS One,* 4**,** e4233.

IDE, Y., TANIMOTO, A., SASAGURI, Y. & PADMANABHAN, R. 1997. Hepatitis C virus NS5A protein is phosphorylated in vitro by a stably bound protein kinase from HeLa cells and by cAMP-dependent protein kinase A-alpha catalytic subunit. *Gene,* 201**,** 151-8.

IDE, Y., ZHANG, L., CHEN, M., INCHAUSPE, G., BAHL, C., SASAGURI, Y. & PADMANABHAN, R. 1996. Characterization of the nuclear localization signal and subcellular distribution of hepatitis C virus nonstructural protein NS5A. *Gene,* 182**,** 203-11.

IGLOI, Z. & HARRIS, M. in preparation.

INOUE, H., NOJIMA, H. & OKAYAMA, H. 1990. High efficiency transformation of Escherichia coli with plasmids. *Gene,* 96**,** 23-8.

ITO, T. & LAI, M. M. 1997. Determination of the secondary structure of and cellular protein binding to the 3'-untranslated region of the hepatitis C virus RNA genome. *J Virol,* 71**,** 8698-706.

IVANOV, A. V., TUNITSKAYA, V. L., IVANOVA, O. N., MITKEVICH, V. A., PRASSOLOV, V. S., MAKAROV, A. A., KUKHANOVA, M. K. & KOCHETKOV, S. N. 2009. Hepatitis C virus NS5A protein modulates template selection by the RNA polymerase in in vitro system. *FEBS Lett,* 583**,** 277-80.

JAKUBIEC, A. & JUPIN, I. 2007. Regulation of positive-strand RNA virus replication: the emerging role of phosphorylation. *Virus Res,* 129**,** 73-9.

JANGRA, R. K., YI, M. & LEMON, S. M. 2010. Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122. *J Virol,* 84**,** 6615-25.

JIANG, J. & LUO, G. 2009. Apolipoprotein E but not B is required for the formation of infectious hepatitis C virus particles. *J Virol,* 83**,** 12680-91.

JIRASKO, V., MONTSERRET, R., APPEL, N., JANVIER, A., EUSTACHI, L., BROHM, C., STEINMANN, E., PIETSCHMANN, T., PENIN, F. & BARTENSCHLAGER, R. 2008. Structural and functional characterization of nonstructural protein 2 for its role in hepatitis C virus assembly. *J Biol Chem,* 283**,** 28546-62.

JIRASKO, V., MONTSERRET, R., LEE, J. Y., GOUTTENOIRE, J., MORADPOUR, D., PENIN, F. & BARTENSCHLAGER, R. 2010. Structural and functional studies of nonstructural protein 2 of the hepatitis C virus reveal its key role as organizer of virion assembly. *PLoS Pathog,* 6**,** e1001233.

JONES, C. T., MURRAY, C. L., EASTMAN, D. K., TASSELLO, J. & RICE, C. M. 2007a. Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *J Virol,* 81**,** 8374-83.

JONES, D. M., ATOOM, A. M., ZHANG, X., KOTTILIL, S. & RUSSELL, R. S. 2011. A genetic interaction between the core and NS3 proteins of hepatitis C virus is essential for production of infectious virus. *J Virol,* 85**,** 12351-61.

JONES, D. M., GRETTON, S. N., MCLAUCHLAN, J. & TARGETT-ADAMS, P. 2007b. Mobility analysis of an NS5A-GFP fusion protein in cells actively replicating hepatitis C virus subgenomic RNA. *J Gen Virol,* 88**,** 470-5.

KAPOOR, A., SIMMONDS, P., GEROLD, G., QAISAR, N., JAIN, K., HENRIQUEZ, J. A., FIRTH, C., HIRSCHBERG, D. L., RICE, C. M., SHIELDS, S. & LIPKIN, W. I. 2011. Characterization of a canine homolog of hepatitis C virus. *Proc Natl Acad Sci U S A,* 108**,** 11608-13.

KAPOOR, A., SIMMONDS, P., SCHEEL, T. K., HJELLE, B., CULLEN, J. M., BURBELO, P. D., CHAUHAN, L. V., DURAISAMY, R., SANCHEZ LEON, M., JAIN, K., VANDEGRIFT, K. J., CALISHER, C. H., RICE, C. M. & LIPKIN, W. I. 2013. Identification of rodent homologs of hepatitis C virus and pegiviruses. *MBio,* 4**,** e00216-13.

KATO, T., CHOI, Y., ELMOWALID, G., SAPP, R. K., BARTH, H., FURUSAKA, A., MISHIRO, S., WAKITA, T., KRAWCZYNSKI, K. & LIANG, T. J. 2008. Hepatitis C virus JFH-1 strain infection in chimpanzees is associated with low pathogenicity and emergence of an adaptive mutation. *Hepatology,* 48**,** 732-40.

KATO, T., DATE, T., MIYAMOTO, M., FURUSAKA, A., TOKUSHIGE, K., MIZOKAMI, M. & WAKITA, T. 2003. Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon. *Gastroenterology,* 125**,** 1808-17.

KATZE, M. G., KWIECISZEWSKI, B., GOODLETT, D. R., BLAKELY, C. M., NEDDERMANN, P., TAN, S. L. & AEBERSOLD, R. 2000. Ser(2194) is a highly conserved major phosphorylation site of the hepatitis C virus nonstructural protein NS5A. *Virology,* 278**,** 501-13.

KEATING, J. A. & STRIKER, R. 2012. Phosphorylation events during viral infections provide potential therapeutic targets. *Rev Med Virol,* 22**,** 166-81.

KIM, D. E., CHIVIAN, D. & BAKER, D. 2004. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res,* 32**,** W526-31.

KIM, J., LEE, D. & CHOE, J. 1999. Hepatitis C virus NS5A protein is phosphorylated by casein kinase II. *Biochem Biophys Res Commun,* 257**,** 777-81.

KIM, S., WELSCH, C., YI, M. & LEMON, S. M. 2011. Regulation of the production of infectious genotype 1a hepatitis C virus by NS5A domain III. *J Virol,* 85**,** 6645-56.

KLIBANOV, O. M., VICKERY, S. B., OLIN, J. L., SMITH, L. S. & WILLIAMS, S. H. 2012. Boceprevir: a novel NS3/4 protease inhibitor for the treatment of hepatitis C. *Pharmacotherapy,* 32**,** 173-90.

KLIBANOV, O. M., WILLIAMS, S. H., SMITH, L. S., OLIN, J. L. & VICKERY, S. B. 2011. Telaprevir: a novel NS3/4 protease inhibitor for the treatment of hepatitis C. *Pharmacotherapy,* 31**,** 951-74.

KLUMPP, S. & KRIEGLSTEIN, J. 2002. Phosphorylation and dephosphorylation of histidine residues in proteins. *Eur J Biochem,* 269**,** 1067-71.

KOCH, J. O. & BARTENSCHLAGER, R. 1999. Modulation of hepatitis C virus NS5A hyperphosphorylation by nonstructural proteins NS3, NS4A, and NS4B. *J Virol,* 73**,** 7138-46.

KOLYKHALOV, A. A., FEINSTONE, S. M. & RICE, C. M. 1996. Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. *J Virol,* 70**,** 3363-71.

KOMURIAN-PRADEL, F., PERRET, M., DEIMAN, B., SODOYER, M., LOTTEAU, V., PARANHOS-BACCALA, G. & ANDRE, P. 2004. Strand specific quantitative real-time PCR to study replication of hepatitis C virus genome. *J Virol Methods,* 116**,** 103-6.

KRIEGER, N., LOHMANN, V. & BARTENSCHLAGER, R. 2001. Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations. *J Virol,* 75**,** 4614-24.

KUIKEN, C., COMBET, C., BUKH, J., SHIN, I. T., DELEAGE, G., MIZOKAMI, M., RICHARDSON, R., SABLON, E., YUSIM, K., PAWLOTSKY, J. M., SIMMONDS, P. & LOS ALAMOS, H. I. V. D. G. 2006. A comprehensive system for consistent numbering of HCV sequences, proteins and epitopes. *Hepatology,* 44**,** 1355-61.

KUO, G., CHOO, Q. L., ALTER, H. J., GITNICK, G. L., REDEKER, A. G., PURCELL, R. H., MIYAMURA, T., DIENSTAG, J. L., ALTER, M. J., STEVENS, C. E. & ET AL. 1989. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science,* 244**,** 362-4.

LAN, K. H., SHEU, M. L., HWANG, S. J., YEN, S. H., CHEN, S. Y., WU, J. C., WANG, Y. J., KATO, N., OMATA, M., CHANG, F. Y. & LEE, S. D. 2002. HCV NS5A interacts with p53 and inhibits p53-mediated apoptosis. *Oncogene,* 21**,** 4801-11.

LAPID, C. & GAO, Y. 2003. *PrimerX: Automated design of mutagenic primers for site-directed mutagenesis* [Online]. Available: http://www.bioinformatics.org/primerx/ 2012].

LAPIERRE, L. A., TUMA, P. L., NAVARRE, J., GOLDENRING, J. R. & ANDERSON, J. M. 1999. VAP-33 localizes to both an intracellular vesicle population and with occludin at the tight junction. *J Cell Sci,* 112 (Pt 21)**,** 3723-32.

LARSEN, M. R., THINGHOLM, T. E., JENSEN, O. N., ROEPSTORFF, P. & JORGENSEN, T. J. 2005. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics,* 4**,** 873-86.

LAUER, G. M. & WALKER, B. D. 2001. Hepatitis C virus infection. *N Engl J Med,* 345**,** 41-52.

LAVIE, M., GOFFARD, A. & DUBUISSON, J. 2007. Assembly of a functional HCV glycoprotein heterodimer. *Curr Issues Mol Biol,* 9**,** 71-86.

LEMAY, K. L., TREADAWAY, J., ANGULO, I. & TELLINGHUISEN, T. L. 2013. A hepatitis C virus NS5A phosphorylation site that regulates RNA replication. *J Virol,* 87**,** 1255-60.

LEMM, J. A., O'BOYLE, D., 2ND, LIU, M., NOWER, P. T., COLONNO, R., DESHPANDE, M. S., SNYDER, L. B., MARTIN, S. W., ST LAURENT, D. R., SERRANO-WU, M. H., ROMINE, J. L., MEANWELL, N. A. & GAO, M. 2010. Identification of hepatitis C virus NS5A inhibitors. *J Virol,* 84**,** 482-91.

LEV, S., BEN HALEVY, D., PERETTI, D. & DAHAN, N. 2008. The VAP protein family: from cellular functions to motor neuron disease. *Trends Cell Biol,* 18**,** 282-90.

LEVRERO, M. 2006. Viral hepatitis and liver cancer: the case of hepatitis C. *Oncogene,* 25**,** 3834-47.

LI, Y., MASAKI, T. & LEMON, S. M. 2013. miR-122 and the Hepatitis C RNA genome: More than just stability. *RNA Biol,* 10**,** 919-24.

LIANG, Y., YE, H., KANG, C. B. & YOON, H. S. 2007. Domain 2 of nonstructural protein 5A (NS5A) of hepatitis C virus is natively unfolded. *Biochemistry,* 46**,** 11550-8.

LIM, Y. S. & HWANG, S. B. 2011. Hepatitis C virus NS5A protein interacts with phosphatidylinositol 4-kinase type IIIalpha and regulates viral propagation. *J Biol Chem,* 286**,** 11290-8.

LIN, C. 2006. HCV NS3-4A Serine Protease. *In:* TAN, S. L. (ed.) *Hepatitis C Viruses: Genomes and Molecular Biology.* Norfolk (UK): Horizon Bioscience.

LINDENBACH, B. D., EVANS, M. J., SYDER, A. J., WOLK, B., TELLINGHUISEN, T. L., LIU, C. C., MARUYAMA, T., HYNES, R. O., BURTON, D. R., MCKEATING, J. A. & RICE, C. M. 2005. Complete replication of hepatitis C virus in cell culture. *Science,* 309**,** 623-6.

LINDENBACH, B. D., PRAGAI, B. M., MONTSERRET, R., BERAN, R. K., PYLE, A. M., PENIN, F. & RICE, C. M. 2007. The C terminus of hepatitis C virus NS4A encodes an electrostatic switch that regulates NS5A hyperphosphorylation and viral replication. *J Virol,* 81**,** 8905-18.

LIU, C., LI, Y., SEMENOV, M., HAN, C., BAEG, G. H., TAN, Y., ZHANG, Z., LIN, X. & HE, X. 2002. Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell,* 108**,** 837-47.

LIU, S., XIAO, L., NELSON, C. & HAGEDORN, C. H. 2012. A cell culture adapted HCV JFH1 variant that increases viral titers and permits the production of high titer infectious chimeric reporter viruses. *PLoS One,* 7**,** e44965.

LOHMANN, V., KORNER, F., DOBIERZEWSKA, A. & BARTENSCHLAGER, R. 2001. Mutations in hepatitis C virus RNAs conferring cell culture adaptation. *J Virol,* 75**,** 1437-49.

LOHMANN, V., KORNER, F., KOCH, J., HERIAN, U., THEILMANN, L. & BARTENSCHLAGER, R. 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science,* 285**,** 110-3.

LOVE, R. A., BRODSKY, O., HICKEY, M. J., WELLS, P. A. & CRONIN, C. N. 2009. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. *J Virol,* 83**,** 4395-403.

LUNDIN, M., LINDSTROM, H., GRONWALL, C. & PERSSON, M. A. 2006. Dual topology of the processed hepatitis C virus protein NS4B is influenced by the NS5A protein. *J Gen Virol,* 87**,** 3263-72.

LUNDIN, M., MONNE, M., WIDELL, A., VON HEIJNE, G. & PERSSON, M. A. 2003. Topology of the membrane-associated hepatitis C virus protein NS4B. *J Virol,* 77**,** 5428-38.

MA, Y., ANANTPADMA, M., TIMPE, J. M., SHANMUGAM, S., SINGH, S. M., LEMON, S. M. & YI, M. 2011. Hepatitis C virus NS2 protein serves as a scaffold for virus assembly by interacting with both structural and nonstructural proteins. *J Virol,* 85**,** 86-97.

MA, Y., YATES, J., LIANG, Y., LEMON, S. M. & YI, M. 2008. NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. *J Virol,* 82**,** 7624-39.

MACDONALD, A., CROWDER, K., STREET, A., MCCORMICK, C. & HARRIS, M. 2004. The hepatitis C virus NS5A protein binds to members of the Src family of tyrosine kinases and regulates kinase activity. *J Gen Virol,* 85**,** 721-9.

MACDONALD, A. & HARRIS, M. 2004. Hepatitis C virus NS5A: tales of a promiscuous protein. *J Gen Virol,* 85**,** 2485-502.

MANKOURI, J., DALLAS, M. L., HUGHES, M. E., GRIFFIN, S. D., MACDONALD, A., PEERS, C. & HARRIS, M. 2009. Suppression of a pro-apoptotic K+ channel as a mechanism for hepatitis C virus persistence. *Proc Natl Acad Sci U S A,* 106**,** 15903-8.

MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T. & SUDARSANAM, S. 2002. The protein kinase complement of the human genome. *Science,* 298**,** 1912-34.

MANNS, M. P. & VON HAHN, T. 2013. Novel therapies for hepatitis C - one pill fits all? *Nat Rev Drug Discov,* 12**,** 595-610.

MANNS, M. P., WEDEMEYER, H. & CORNBERG, M. 2006. Treating viral hepatitis C: efficacy, side effects, and complications. *Gut,* 55**,** 1350-9.

MASAKI, T., SUZUKI, R., MURAKAMI, K., AIZAKI, H., ISHII, K., MURAYAMA, A., DATE, T., MATSUURA, Y., MIYAMURA, T., WAKITA, T. & SUZUKI, T. 2008. Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. *J Virol,* 82**,** 7964-76.

MAUSS, S., BERG, T., ROCKSTROH, J., SARRAZIN, C. & WEDEMEYER, H. 2011a. *Short guide to Hepatitis C*, Flying Publisher.

MAUSS, S., HUEPPE, D., JOHN, C., GOELZ, J., HEYNE, R., MOELLER, B., LINK, R., TEUBER, G., HERRMANN, A., SPELTER, M., WOLLSCHLAEGER, S., BAUMGARTEN, A., SIMON, K. G., DIKOPOULOS, N. & WITTHOEFT, T. 2011b. Estimating the likelihood of sustained virological response in chronic hepatitis C therapy. *J Viral Hepat,* 18**,** e81-90.

MCLAUCHLAN, J., LEMBERG, M. K., HOPE, G. & MARTOGLIO, B. 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *Embo j,* 21**,** 3980-8.

MEEX, S. J., ANDREO, U., SPARKS, J. D. & FISHER, E. A. 2011. Huh-7 or HepG2 cells: which is the better model for studying human apolipoprotein-B100 assembly and secretion? *J Lipid Res,* 52**,** 152-8.

MERZ, A., LONG, G., HIET, M. S., BRUGGER, B., CHLANDA, P., ANDRE, P., WIELAND, F., KRIJNSE-LOCKER, J. & BARTENSCHLAGER, R. 2011. Biochemical and morphological properties of hepatitis C virus particles and determination of their lipidome. *J Biol Chem,* 286**,** 3018-32.

MIYANARI, Y., ATSUZAWA, K., USUDA, N., WATASHI, K., HISHIKI, T., ZAYAS, M., BARTENSCHLAGER, R., WAKITA, T., HIJIKATA, M. & SHIMOTOHNO, K. 2007. The lipid droplet is an important organelle for hepatitis C virus production. *Nat Cell Biol,* 9**,** 1089-97.

MOLINA, S., CASTET, V., FOURNIER-WIRTH, C., PICHARD-GARCIA, L., AVNER, R., HARATS, D., ROITELMAN, J., BARBARAS, R., GRABER, P., GHERSA, P., SMOLARSKY, M., FUNARO, A., MALAVASI, F., LARREY, D., COSTE, J., FABRE, J. M., SA-CUNHA, A. & MAUREL, P.

2007. The low-density lipoprotein receptor plays a role in the infection of primary human hepatocytes by hepatitis C virus. *J Hepatol,* 46**,** 411-9.

MURPHY, D. G., WILLEMS, B., DESCHENES, M., HILZENRAT, N., MOUSSEAU, R. & SABBAH, S. 2007. Use of sequence analysis of the NS5B region for routine genotyping of hepatitis C virus with reference to C/E1 and 5' untranslated region sequences. *J Clin Microbiol,* 45**,** 1102-12.

MURRAY, C. L., JONES, C. T. & RICE, C. M. 2008. Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. *Nat Rev Microbiol,* 6**,** 699-708.

NANDA, S. K., HERION, D. & LIANG, T. J. 2006. The SH3 binding motif of HCV [corrected] NS5A protein interacts with Bin1 and is important for apoptosis and infectivity. *Gastroenterology,* 130**,** 794-809.

NEB. 2010. *Protein Labeling with SNAP-tag and CLIP-tag* [Online]. Available: http://international.neb.com/applications/protein-analysis-and-tools/protein-labeling/protein-labeling-snap-clip#tabselect4 2013].

NEDDERMANN, P., CLEMENTI, A. & DE FRANCESCO, R. 1999. Hyperphosphorylation of the hepatitis C virus NS5A protein requires an active NS3 protease, NS4A, NS4B, and NS5A encoded on the same polyprotein. *J Virol,* 73**,** 9984-91.

NEDDERMANN, P., QUINTAVALLE, M., DI PIETRO, C., CLEMENTI, A., CERRETANI, M., ALTAMURA, S., BARTHOLOMEW, L. & DE FRANCESCO, R. 2004. Reduction of hepatitis C virus NS5A hyperphosphorylation by selective inhibition of cellular kinases activates viral RNA replication in cell culture. *J Virol,* 78**,** 13306-14.

NEUFELDT, C. J., JOYCE, M. A., LEVIN, A., STEENBERGEN, R. H., TYRRELL, L. D. J. & WOZNIAK, R. W. 2012. A role for cytoplasmic nuclear pore complexes in Hepatitis C virus replication. *19th International Symposium on Hepatitis C Virus an Related Viruses.* Venice, Italk.

NEVEU, G., BAROUCH-BENTOV, R., ZIV-AV, A., GERBER, D., JACOB, Y. & EINAV, S. 2012. Identification and targeting of an interaction between a tyrosine motif within hepatitis C virus core protein and AP2M1 essential for viral assembly. *PLoS Pathog,* 8**,** e1002845.

NEWMAN, R. M., KUNTZEN, T., WEINER, B., BERICAL, A., CHARLEBOIS, P., KUIKEN, C., MURPHY, D. G., SIMMONDS, P., BENNETT, P., LENNON, N. J., BIRREN, B. W., ZODY, M. C., ALLEN, T. M. & HENN, M. R. 2013. Whole genome pyrosequencing of rare hepatitis C virus genotypes enhances subtype classification and identification of naturally occurring drug resistance variants. *J Infect Dis,* 208**,** 17-31.

NIELSEN, S. U., BASSENDINE, M. F., BURT, A. D., MARTIN, C., PUMEECHOCKCHAI, W. & TOMS, G. L. 2006. Association between hepatitis C virus and very-low-density lipoprotein (VLDL)/LDL analyzed in iodixanol density gradients. *J Virol,* 80**,** 2418-28.

NORDLE GILLIVER, A., GRIFFIN, S. & HARRIS, M. 2010. Identification of a novel phosphorylation site in hepatitis C virus NS5A. *J Gen Virol,* 91**,** 2428-32.

OP DE BEECK, A., VOISSET, C., BARTOSCH, B., CICZORA, Y., COCQUEREL, L., KECK, Z., FOUNG, S., COSSET, F. L. & DUBUISSON, J. 2004. Characterization of functional hepatitis C virus envelope glycoproteins. *J Virol,* 78**,** 2994-3002.

OWEN, D. M., HUANG, H., YE, J. & GALE, M., JR. 2009. Apolipoprotein E on hepatitis C virion facilitates infection through interaction with low-density lipoprotein receptor. *Virology,* 394**,** 99-108.

PAESHUYSE, J., KAUL, A., DE CLERCQ, E., ROSENWIRTH, B., DUMONT, J. M., SCALFARO, P., BARTENSCHLAGER, R. & NEYTS, J. 2006. The non-immunosuppressive cyclosporin DEBIO-025 is a potent inhibitor of hepatitis C virus replication in vitro. *Hepatology,* 43**,** 761-70.

PARK, J. E., SOUNG, N. K., JOHMURA, Y., KANG, Y. H., LIAO, C., LEE, K. H., PARK, C. H., NICKLAUS, M. C. & LEE, K. S. 2010. Polo-box domain: a versatile mediator of polo-like kinase function. *Cell Mol Life Sci,* 67**,** 1957-70.

PARK, K. J., CHOI, S. H., CHOI, D. H., PARK, J. M., YIE, S. W., LEE, S. Y. & HWANG, S. B. 2003. Hepatitis C virus NS5A protein modulates c-Jun N-terminal kinase through interaction with tumor necrosis factor receptor-associated factor 2. *J Biol Chem,* 278**,** 30711-8.

PASCU, M., MARTUS, P., HOHNE, M., WIEDENMANN, B., HOPF, U., SCHREIER, E. & BERG, T. 2004. Sustained virological response in hepatitis C virus type 1b infected patients is predicted by the number of mutations within the NS5A-ISDR: a meta-analysis focused on geographical differences. *Gut,* 53**,** 1345-51.

PAUL, D., HOPPE, S., SAHER, G., KRIJNSE-LOCKER, J. & BARTENSCHLAGER, R. 2013. Morphological and biochemical characterization of the membranous hepatitis C virus replication compartment. *J Virol*, 87(19), 10612-27.

PAWLOTSKY, J. M. 2006. Hepatitis C virus population dynamics during infection. *Curr Top Microbiol Immunol,* 299**,** 261-84.

PENIN, F., BRASS, V., APPEL, N., RAMBOARINA, S., MONTSERRET, R., FICHEUX, D., BLUM, H. E., BARTENSCHLAGER, R. & MORADPOUR, D. 2004. Structure and function of the membrane anchor domain of hepatitis C virus nonstructural protein 5A. *J Biol Chem,* 279**,** 40835-43.

PIETSCHMANN, T., KAUL, A., KOUTSOUDAKIS, G., SHAVINSKAYA, A., KALLIS, S., STEINMANN, E., ABID, K., NEGRO, F., DREUX, M., COSSET, F. L. & BARTENSCHLAGER, R. 2006. Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. *Proc Natl Acad Sci U S A,* 103**,** 7408-13.

PILERI, P., UEMATSU, Y., CAMPAGNOLI, S., GALLI, G., FALUGI, F., PETRACCA, R., WEINER, A. J., HOUGHTON, M., ROSA, D., GRANDI, G. & ABRIGNANI, S. 1998. Binding of hepatitis C virus to CD81. *Science,* 282**,** 938-41.

PLANAS, R., BALLESTE, B., ALVAREZ, M. A., RIVERA, M., MONTOLIU, S., GALERAS, J. A., SANTOS, J., COLL, S., MORILLAS, R. M. & SOLA, R. 2004. Natural history of decompensated hepatitis C virus-related cirrhosis. A study of 200 patients. *J Hepatol,* 40**,** 823-30.

PLOEN, D., HAFIRASSOU, M. L., HIMMELSBACH, K., SAUTER, D., BINIOSSEK, M. L., BAUMERT, T., SCHUSTER, C. & HILDT, E. 2012. Tail interacting protein of 47 kD (TIP47) acts as an essential sorting factor for HCV morphogenesis and release of viral particles. *19th International Symposium on Hepatitis C Virus and Related Viruses.* Venice, Italy.

PLOSS, A., EVANS, M. J., GAYSINSKAYA, V. A., PANIS, M., YOU, H., DE JONG, Y. P. & RICE, C. M. 2009. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature,* 457**,** 882-6.

PLOSS, A., KHETANI, S. R., JONES, C. T., SYDER, A. J., TREHAN, K., GAYSINSKAYA, V. A., MU, K., RITOLA, K., RICE, C. M. & BHATIA, S. N. 2010. Persistent hepatitis C virus infection in microscale primary human hepatocyte cultures. *Proc Natl Acad Sci U S A,* 107**,** 3141-5.

PODEVIN, P., CARPENTIER, A., PENE, V., AOUDJEHANE, L., CARRIERE, M., ZAIDI, S., HERNANDEZ, C., CALLE, V., MERITET, J. F., SCATTON, O., DREUX, M., COSSET, F. L., WAKITA, T., BARTENSCHLAGER, R., DEMIGNOT, S., CONTI, F., ROSENBERG, A. R. & CALMUS, Y. 2010. Production of infectious hepatitis C virus in primary cultures of human adult hepatocytes. *Gastroenterology,* 139**,** 1355-64.

POPESCU, C. I., CALLENS, N., TRINEL, D., ROINGEARD, P., MORADPOUR, D., DESCAMPS, V., DUVERLIE, G., PENIN, F., HELIOT, L., ROUILLE, Y. & DUBUISSON, J. 2011a. NS2 protein of hepatitis C virus interacts with structural and non-structural proteins towards virus assembly. *PLoS Pathog,* 7**,** e1001278.

POPESCU, C. I., ROUILLE, Y. & DUBUISSON, J. 2011b. Hepatitis C virus assembly imaging. *Viruses,* 3**,** 2238-54.

POST, J. J., PAN, Y., FREEMAN, A. J., HARVEY, C. E., WHITE, P. A., PALLADINETTI, P., HABER, P. S., MARINOS, G., LEVY, M. H., KALDOR, J. M., DOLAN, K. A., FFRENCH, R. A., LLOYD, A.

R., RAWLINSON, W. D., HEPATITIS, C. I. & TRANSMISSION IN PRISONS STUDY, G. 2004. Clearance of hepatitis C viremia associated with cellular immunity in the absence of seroconversion in the hepatitis C incidence and transmission in prisons study cohort. *J Infect Dis,* 189**,** 1846-55.

POYNARD, T., BEDOSSA, P. & OPOLON, P. 1997. Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet,* 349**,** 825-32.

PROVOST, C. R. & SUN, L. 2010. Fluorescent labeling of COS-7 expressing SNAP-tag fusion proteins for live cell imaging. *J Vis Exp*, 39, 1879.

QIU, D., LEMM, J. A., O'BOYLE, D. R., 2ND, SUN, J. H., NOWER, P. T., NGUYEN, V., HAMANN, L. G., SNYDER, L. B., DEON, D. H., RUEDIGER, E., MEANWELL, N. A., BELEMA, M., GAO, M. & FRIDELL, R. A. 2011. The effects of NS5A inhibitors on NS5A phosphorylation, polyprotein processing and localization. *J Gen Virol,* 92**,** 2502-11.

QUAN, P. L., FIRTH, C., CONTE, J. M., WILLIAMS, S. H., ZAMBRANA-TORRELIO, C. M., ANTHONY, S. J., ELLISON, J. A., GILBERT, A. T., KUZMIN, I. V., NIEZGODA, M., OSINUBI, M. O., RECUENCO, S., MARKOTTER, W., BREIMAN, R. F., KALEMBA, L., MALEKANI, J., LINDBLADE, K. A., ROSTAL, M. K., OJEDA-FLORES, R., SUZAN, G., DAVIS, L. B., BLAU, D. M., OGUNKOYA, A. B., ALVAREZ CASTILLO, D. A., MORAN, D., NGAM, S., AKAIBE, D., AGWANDA, B., BRIESE, T., EPSTEIN, J. H., DASZAK, P., RUPPRECHT, C. E., HOLMES, E. C. & LIPKIN, W. I. 2013. Bats are a major natural reservoir for hepaciviruses and pegiviruses. *Proc Natl Acad Sci U S A,* 110**,** 8194-9.

QUINKERT, D., BARTENSCHLAGER, R. & LOHMANN, V. 2005. Quantitative analysis of the hepatitis C virus replication complex. *J Virol,* 79**,** 13594-605.

QUINTAVALLE, M., SAMBUCINI, S., DI PIETRO, C., DE FRANCESCO, R. & NEDDERMANN, P. 2006. The alpha isoform of protein kinase CKI is responsible for hepatitis C virus NS5A hyperphosphorylation. *J Virol,* 80**,** 11305-12.

QUINTAVALLE, M., SAMBUCINI, S., SUMMA, V., ORSATTI, L., TALAMO, F., DE FRANCESCO, R. & NEDDERMANN, P. 2007. Hepatitis C virus NS5A is a direct substrate of casein kinase I-alpha, a cellular kinase identified by inhibitor affinity chromatography using specific NS5A hyperphosphorylation inhibitors. *J Biol Chem,* 282**,** 5536-44.

RAU, M., BAUR, K. & GEIER, A. 2012. Host genetic variants in the pathogenesis of hepatitis C. *Viruses,* 4**,** 3281-302.

REDDY, K. R., SHIFFMAN, M. L., MORGAN, T. R., ZEUZEM, S., HADZIYANNIS, S., HAMZEH, F. M., WRIGHT, T. L. & FRIED, M. 2007. Impact of ribavirin dose reductions in hepatitis C virus genotype 1 patients completing peginterferon alfa-2a/ribavirin treatment. *Clin Gastroenterol Hepatol,* 5**,** 124-9.

REED, K. E. & RICE, C. M. 1999. Identification of the major phosphorylation site of the hepatitis C virus H strain NS5A protein as serine 2321. *J Biol Chem,* 274**,** 28011-8.

REGHELLIN, V., FENU, S., BIANCO, A., ABRIGNANI, S., DE FRANCESCO, R. & NEDDERMANN, P. 2012. HCV NS5A inhibitor BMS-790052 exerts its antiviral effects by reducing phosphatidylinositol 4-phosphate levels in the virus-induced membranous web. *19th International Symposium on Hepatitis C Virus and Related Viruses.* Venice, Italy.

REISS, S., HARAK, C., ROMERO-BREY, I., RADUJKOVIC, D., KLEIN, R., RUGGIERI, A., REBHAN, I., BARTENSCHLAGER, R. & LOHMANN, V. 2013. The lipid kinase phosphatidylinositol-4 kinase III alpha regulates the phosphorylation status of hepatitis C virus NS5A. *PLoS Pathog,* 9**,** e1003359.

REISS, S., REBHAN, I., BACKES, P., ROMERO-BREY, I., ERFLE, H., MATULA, P., KADERALI, L., POENISCH, M., BLANKENBURG, H., HIET, M. S., LONGERICH, T., DIEHL, S., RAMIREZ, F., BALLA, T., ROHR, K., KAUL, A., BUHLER, S., PEPPERKOK, R., LENGAUER, T., ALBRECHT, M., EILS, R., SCHIRMACHER, P., LOHMANN, V. & BARTENSCHLAGER, R. 2011. Recruitment and activation of a lipid kinase by hepatitis C virus NS5A is

essential for integrity of the membranous replication compartment. *Cell Host Microbe,* 9**,** 32-45.

ROMERO-BREY, I., MERZ, A., CHIRAMEL, A., LEE, J. Y., CHLANDA, P., HASELMAN, U., SANTARELLA-MELLWIG, R., HABERMANN, A., HOPPE, S., KALLIS, S., WALTHER, P., ANTONY, C., KRIJNSE-LOCKER, J. & BARTENSCHLAGER, R. 2012. Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. *PLoS Pathog,* 8**,** e1003056.

ROSNOBLET, C., FRITZINGER, B., LEGRAND, D., LAUNAY, H., WIERUSZESKI, J. M., LIPPENS, G. & HANOULLE, X. 2012. Hepatitis C virus NS5B and host cyclophilin A share a common binding site on NS5A. *J Biol Chem,* 287**,** 44249-60.

SALLOUM, S., WANG, H., FERGUSON, C., PARTON, R. G. & TAI, A. W. 2013. Rab18 Binds to Hepatitis C Virus NS5A and Promotes Interaction between Sites of Viral Replication and Lipid Droplets. *PLoS Pathog,* 9**,** e1003513.

SCARSELLI, E., ANSUINI, H., CERINO, R., ROCCASECCA, R. M., ACALI, S., FILOCAMO, G., TRABONI, C., NICOSIA, A., CORTESE, R. & VITELLI, A. 2002. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J,* 21**,** 5017-25.

SCHALLER, T., APPEL, N., KOUTSOUDAKIS, G., KALLIS, S., LOHMANN, V., PIETSCHMANN, T. & BARTENSCHLAGER, R. 2007. Analysis of hepatitis C virus superinfection exclusion by using novel fluorochrome gene-tagged viral genomes. *J Virol,* 81**,** 4591-603.

SCHREIBER, G. B., BUSCH, M. P., KLEINMAN, S. H. & KORELITZ, J. J. 1996. The risk of transfusion-transmitted viral infections. The Retrovirus Epidemiology Donor Study. *N Engl J Med,* 334**,** 1685-90.

SCHRODINGER, L. 2010. *The PyMol Molecular Graphics system, Version 1.5.0.4* [Online]. Available: http://pymol.org/.

SEEFF, L. B. 2002. Natural history of chronic hepatitis C. *Hepatology,* 36**,** S35-46.

SEREBROV, V. & PYLE, A. M. 2004. Periodic cycles of RNA unwinding and pausing by hepatitis C virus NS3 helicase. *Nature,* 430**,** 476-80.

SHAVINSKAYA, A., BOULANT, S., PENIN, F., MCLAUCHLAN, J. & BARTENSCHLAGER, R. 2007. The lipid droplet binding domain of hepatitis C virus core protein is a major determinant for efficient virus assembly. *J Biol Chem,* 282**,** 37158-69.

SHEPARD, C. W., FINELLI, L. & ALTER, M. J. 2005. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis,* 5**,** 558-67.

SHI, S. T., POLYAK, S. J., TU, H., TAYLOR, D. R., GRETCH, D. R. & LAI, M. M. 2002. Hepatitis C virus NS5A colocalizes with the core protein on lipid droplets and interacts with apolipoproteins. *Virology,* 292**,** 198-210.

SHI, W., FREITAS, I. T., ZHU, C., ZHENG, W., HALL, W. W. & HIGGINS, D. G. 2012. Recombination in hepatitis C virus: identification of four novel naturally occurring inter-subtype recombinants. *PLoS One,* 7**,** e41997.

SHIMAKAMI, T., HIJIKATA, M., LUO, H., MA, Y. Y., KANEKO, S., SHIMOTOHNO, K. & MURAKAMI, S. 2004. Effect of interaction between hepatitis C virus NS5A and NS5B on hepatitis C virus RNA replication with the hepatitis C virus replicon. *J Virol,* 78**,** 2738-48.

SHIMAKAMI, T., YAMANE, D., JANGRA, R. K., KEMPF, B. J., SPANIEL, C., BARTON, D. J. & LEMON, S. M. 2012. Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex. *Proc Natl Acad Sci U S A,* 109**,** 941-6.

SHIROTA, Y., LUO, H., QIN, W., KANEKO, S., YAMASHITA, T., KOBAYASHI, K. & MURAKAMI, S. 2002. Hepatitis C virus (HCV) NS5A binds RNA-dependent RNA polymerase (RdRP) NS5B and modulates RNA-dependent RNA polymerase activity. *J Biol Chem,* 277**,** 11149-55.

SHOEMAKER, B. A., PORTMAN, J. J. & WOLYNES, P. G. 2000. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci U S A,* 97**,** 8868-73.

SIMMONDS, P. 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. *J Gen Virol,* 85**,** 3173-88.

SIMMONDS, P., BUKH, J., COMBET, C., DELEAGE, G., ENOMOTO, N., FEINSTONE, S., HALFON, P., INCHAUSPE, G., KUIKEN, C., MAERTENS, G., MIZOKAMI, M., MURPHY, D. G., OKAMOTO, H., PAWLOTSKY, J. M., PENIN, F., SABLON, E., SHIN, I. T., STUYVER, L. J., THIEL, H. J., VIAZOV, S., WEINER, A. J. & WIDELL, A. 2005. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology,* 42**,** 962-73.

STEENBERGEN, R., JOYCE, M., THOMAS, B., LAW, J., HOUGHTON, M. & TYRRELL, L. 2012. Human serum leads to differentiation of human hepatoma cells and a 1000-fold increase in JFH-1 titers. *19th International Symposium on Hepatitis C virus and Related Viruses.* Venice, Italy.

SUER, S., SICKMANN, A., MEYER, H. E., HERBERG, F. W. & HEILMEYER, L. M., JR. 2001. Human phosphatidylinositol 4-kinase isoform PI4K92. Expression of the recombinant enzyme and determination of multiple phosphorylation sites. *Eur J Biochem,* 268**,** 2099-106.

SULKOWSKI, M., GARDINER, D. & RODRIGUEZ-TORRES, M. 2012. High rate of sustained virologic response with the all-oral combination of daclatasvir (NS5A inhibitor) plus sofosbuvir (nucleotide NS5B inhibitor), with or without ribavirin, in treatment-naive patients chronically infected with HCG genotype 1, 2, or 3.

. *63rd Annual Meeting of the American Association for the Study of Liver Diseases.* Boston, Massachusetts.

SUN, D., LUTHRA, P., LI, Z. & HE, B. 2009. PLK1 down-regulates parainfluenza virus 5 gene expression. *PLoS Pathog,* 5**,** e1000525.

TAI, A. W., BENITA, Y., PENG, L. F., KIM, S. S., SAKAMOTO, N., XAVIER, R. J. & CHUNG, R. T. 2009. A functional genomic screen identifies cellular cofactors of hepatitis C virus replication. *Cell Host Microbe,* 5**,** 298-307.

TAI, C. L., CHI, W. K., CHEN, D. S. & HWANG, L. H. 1996. The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). *J Virol,* 70**,** 8477-84.

TAN, S. 2006. *In:* TAN, S. L. (ed.) *Hepatitis C Viruses: Genomes and Molecular Biology.* Norfolk (UK): Horizon Bioscience.

TAN, S. L., NAKAO, H., HE, Y., VIJAYSRI, S., NEDDERMANN, P., JACOBS, B. L., MAYER, B. J. & KATZE, M. G. 1999. NS5A, a nonstructural protein of hepatitis C virus, binds growth factor receptor-bound protein 2 adaptor protein in a Src homology 3 domain/ligand-dependent manner and perturbs mitogenic signaling. *Proc Natl Acad Sci U S A,* 96**,** 5533-8.

TANG, D., YUAN, H., VIELEMEYER, O., PEREZ, F. & WANG, Y. 2012. Sequential phosphorylation of GRASP65 during mitotic Golgi disassembly. *Biol Open,* 1**,** 1204-14.

TANG, H. & GRISE, H. 2009. Cellular and molecular biology of HCV infection and hepatitis. *Clin Sci (Lond),* 117**,** 49-65.

TANJI, Y., KANEKO, T., SATOH, S. & SHIMOTOHNO, K. 1995. Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. *J Virol,* 69**,** 3980-6.

TARGETT-ADAMS, P., GRAHAM, E. J., MIDDLETON, J., PALMER, A., SHAW, S. M., LAVENDER, H., BRAIN, P., TRAN, T. D., JONES, L. H., WAKENHUT, F., STAMMEN, B., PRYDE, D., PICKFORD, C. & WESTBY, M. 2011. Small molecules targeting hepatitis C virus-encoded NS5A cause subcellular redistribution of their target: insights into compound modes of action. *J Virol,* 85**,** 6353-68.

TARGETT-ADAMS, P., HOPE, G., BOULANT, S. & MCLAUCHLAN, J. 2008. Maturation of hepatitis C virus core protein by signal peptide peptidase is required for virus production. *J Biol Chem,* 283**,** 16850-9.

TARGETT-ADAMS, P. & MCLAUCHLAN, J. 2005. Development and characterization of a transient-replication assay for the genotype 2a hepatitis C virus subgenomic replicon. *J Gen Virol,* 86**,** 3075-80.

TELLINGHUISEN, T. L., FOSS, K. L. & TREADAWAY, J. 2008a. Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. *PLoS Pathog,* 4**,** e1000032.

TELLINGHUISEN, T. L., FOSS, K. L., TREADAWAY, J. C. & RICE, C. M. 2008b. Identification of residues required for RNA replication in domains II and III of the hepatitis C virus NS5A protein. *J Virol,* 82**,** 1073-83.

TELLINGHUISEN, T. L., MARCOTRIGIANO, J., GORBALENYA, A. E. & RICE, C. M. 2004. The NS5A protein of hepatitis C virus is a zinc metalloprotein. *J Biol Chem,* 279**,** 48576-87.

TELLINGHUISEN, T. L., MARCOTRIGIANO, J. & RICE, C. M. 2005. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. *Nature,* 435**,** 374-9.

THEISE, N. D. 2007. Liver biopsy assessment in chronic viral hepatitis: a personal, practical approach. *Mod Pathol,* 20 Suppl 1**,** S3-14.

THOMAS, X. V., DE BRUIJNE, J., SULLIVAN, J. C., KIEFFER, T. L., HO, C. K., REBERS, S. P., DE VRIES, M., REESINK, H. W., WEEGINK, C. J., MOLENKAMP, R. & SCHINKEL, J. 2012. Evaluation of persistence of resistant variants with ultra-deep pyrosequencing in chronic hepatitis C patients treated with telaprevir. *PLoS One,* 7**,** e41191.

TRIPATHI, L. P., KAMBARA, H., CHEN, Y. A., NISHIMURA, Y., MORIISHI, K., OKAMOTO, T., MORITA, E., ABE, T., MORI, Y., MATSUURA, Y. & MIZUGUCHI, K. 2013. Understanding the biological context of NS5A-host interactions in HCV infection: a network-based approach. *J Proteome Res,* 12**,** 2537-51.

TU, H., GAO, L., SHI, S. T., TAYLOR, D. R., YANG, T., MIRCHEFF, A. K., WEN, Y., GORBALENYA, A. E., HWANG, S. B. & LAI, M. M. 1999. Hepatitis C virus RNA polymerase and NS5A complex with a SNARE-like protein. *Virology,* 263**,** 30-41.

VAN DER POEL, C. L., REESINK, H. W., LELIE, P. N., LEENTVAAR-KUYPERS, A., CHOO, Q. L., KUO, G. & HOUGHTON, M. 1989. Anti-hepatitis C antibodies and non-A, non-B post-transfusion hepatitis in The Netherlands. *Lancet,* 2**,** 297-8.

VIEYRES, G. & PIETSCHMANN, T. 2013. Entry and replication of recombinant hepatitis C viruses in cell culture. *Methods,* 59**,** 233-48.

VIEYRES, G., THOMAS, X., DESCAMPS, V., DUVERLIE, G., PATEL, A. H. & DUBUISSON, J. 2010. Characterization of the envelope glycoproteins associated with infectious hepatitis C virus. *J Virol,* 84**,** 10159-68.

VOGT, D. A., CAMUS, G., HERKER, E., WEBSTER, B. R., TSOU, C. L., GREENE, W. C., YEN, T. S. & OTT, M. 2013. Lipid droplet-binding protein TIP47 regulates hepatitis C Virus RNA replication through interaction with the viral NS5A protein. *PLoS Pathog,* 9**,** e1003302.

WAKITA, T., PIETSCHMANN, T., KATO, T., DATE, T., MIYAMOTO, M., ZHAO, Z., MURTHY, K., HABERMANN, A., KRAUSSLICH, H. G., MIZOKAMI, M., BARTENSCHLAGER, R. & LIANG, T. J. 2005. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med,* 11**,** 791-6.

WANG, S., WU, X., PAN, T., SONG, W., WANG, Y., ZHANG, F. & YUAN, Z. 2012. Viperin inhibits hepatitis C virus replication by interfering with binding of NS5A to host protein hVAP-33. *J Gen Virol,* 93**,** 83-92.

WATASHI, K. & SHIMOTOHNO, K. 2007. Cyclophilin and viruses: cyclophilin as a cofactor for viral infection and possible anti-viral target. *Drug Target Insights,* 2**,** 9-18.

WELTE, M. A. 2009. Fat on the move: intracellular motion of lipid droplets. *Biochem Soc Trans,* 37**,** 991-6.

WHO 2009. Global distribution of HCV genotypes.

WITZE, E. S., OLD, W. M., RESING, K. A. & AHN, N. G. 2007. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods,* 4**,** 798-806.

WOLK, B., BUCHELE, B., MORADPOUR, D. & RICE, C. M. 2008. A dynamic view of hepatitis C virus replication complexes. *J Virol,* 82**,** 10519-31.

WRIGHT, P. E. & DYSON, H. J. 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol,* 293**,** 321-31.

WYLES, J. P., MCMASTER, C. R. & RIDGWAY, N. D. 2002. Vesicle-associated membrane protein-associated protein-A (VAP-A) interacts with the oxysterol-binding protein to modify export from the endoplasmic reticulum. *J Biol Chem,* 277**,** 29908-18.

WYLES, J. P. & RIDGWAY, N. D. 2004. VAMP-associated protein-A regulates partitioning of oxysterol-binding protein-related protein-9 between the endoplasmic reticulum and Golgi apparatus. *Exp Cell Res,* 297**,** 533-47.

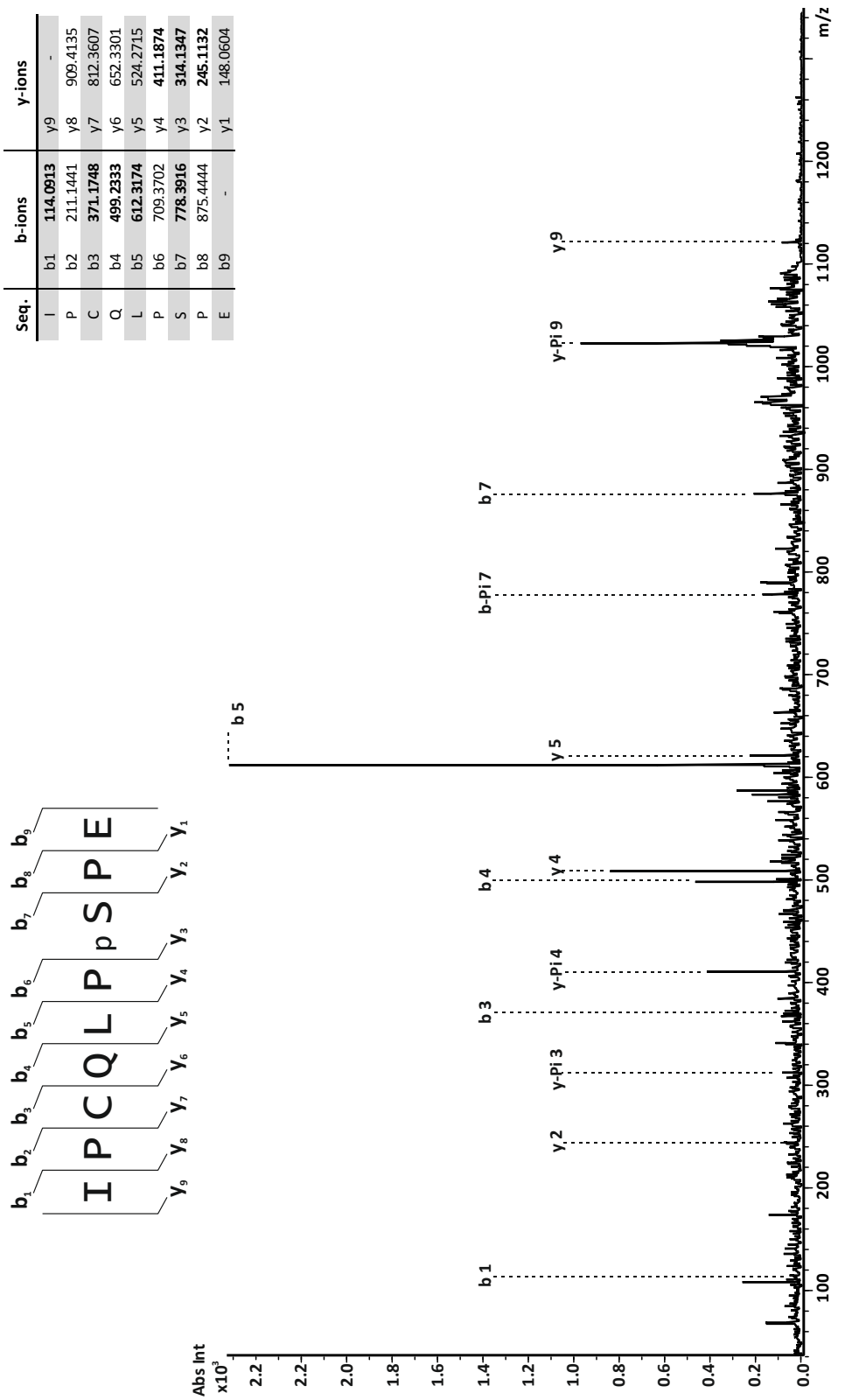
YANAGI, M., ST CLAIRE, M., EMERSON, S. U., PURCELL, R. H. & BUKH, J. 1999. In vivo analysis of the 3' untranslated region of the hepatitis C virus after in vitro mutagenesis of an infectious cDNA clone. *Proc Natl Acad Sci U S A,* 96**,** 2291-5.

YANG, F., ROBOTHAM, J. M., GRISE, H., FRAUSTO, S., MADAN, V., ZAYAS, M., BARTENSCHLAGER, R., ROBINSON, M., GREENSTEIN, A. E., NAG, A., LOGAN, T. M., BIENKIEWICZ, E. & TANG, H. 2010. A major determinant of cyclophilin dependence and cyclosporine susceptibility of hepatitis C virus identified by a genetic approach. *PLoS Pathog,* 6**,** e1001118.

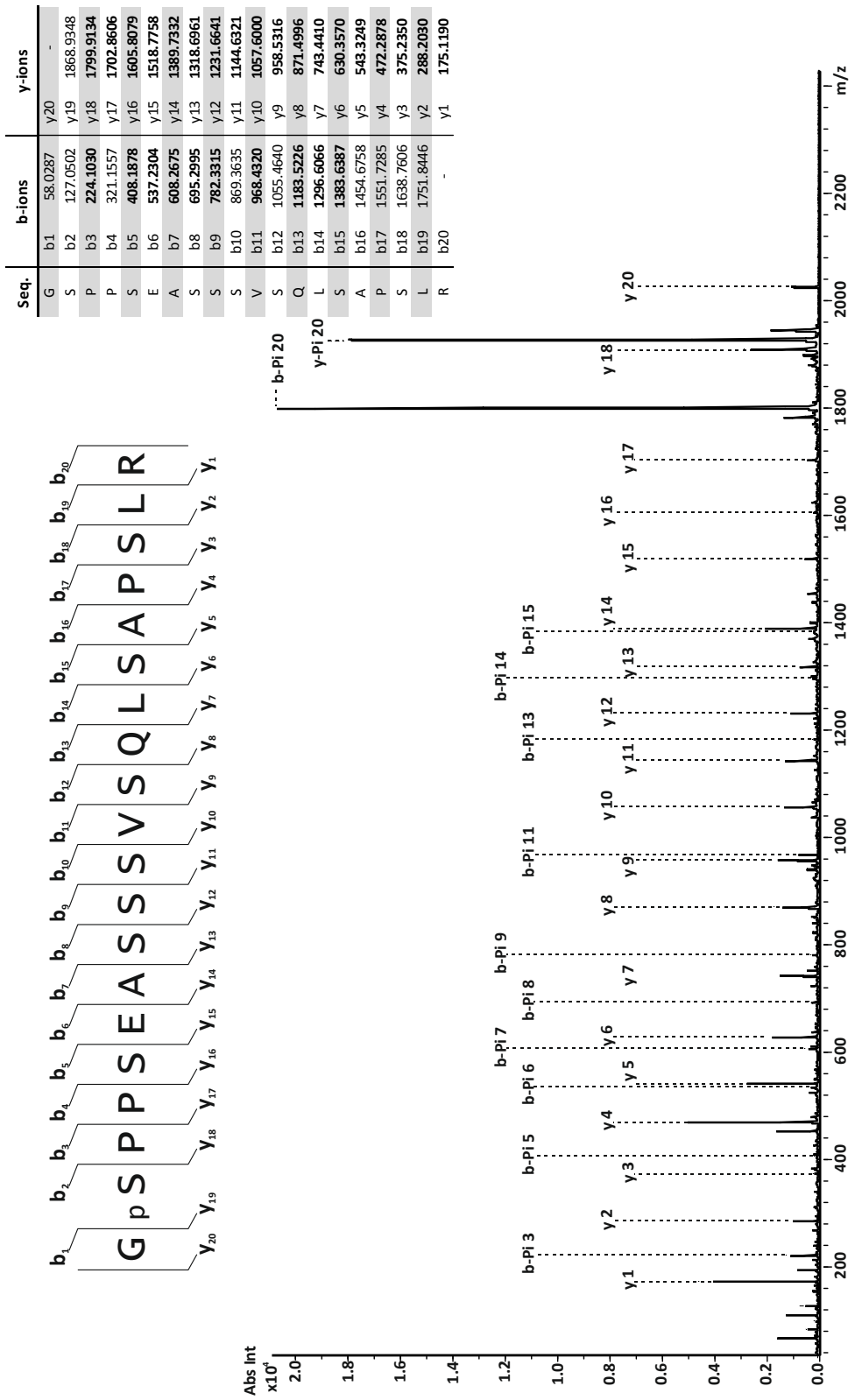
ZECH, B., KURTENBACH, A., KRIEGER, N., STRAND, D., BLENCKE, S., MORBITZER, M., SALASSIDIS, K., COTTEN, M., WISSING, J., OBERT, S., BARTENSCHLAGER, R., HERGET, T. & DAUB, H. 2003. Identification and characterization of amphiphysin II as a novel cellular interaction partner of the hepatitis C virus NS5A protein. *J Gen Virol,* 84**,** 555-60.

ZEISEL, M. B., FOFANA, I., FAFI-KREMER, S. & BAUMERT, T. F. 2011. Hepatitis C virus entry into hepatocytes: molecular mechanisms and targets for antiviral therapies. *J Hepatol,* 54**,** 566-76.

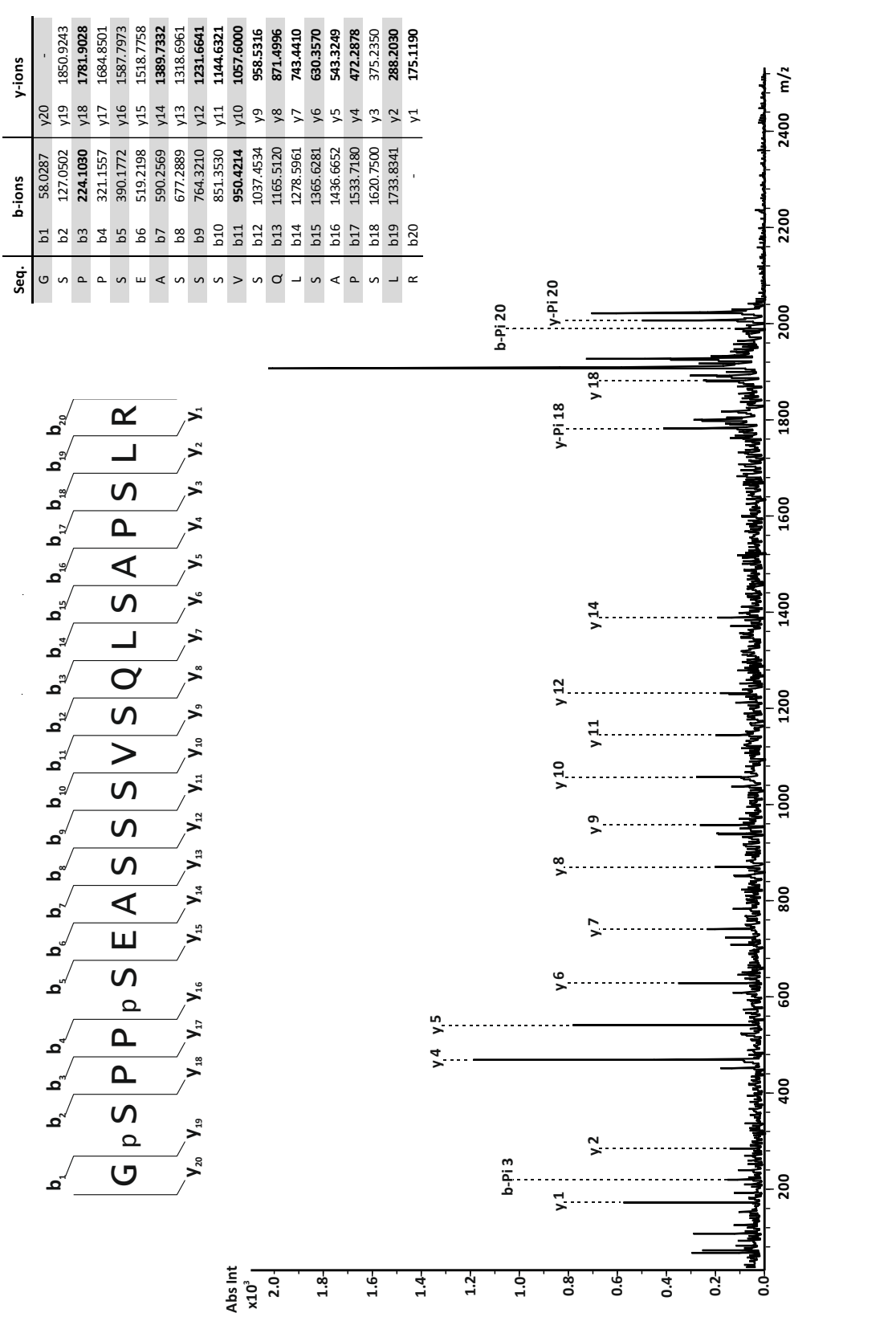
Chapter 8: Appendix



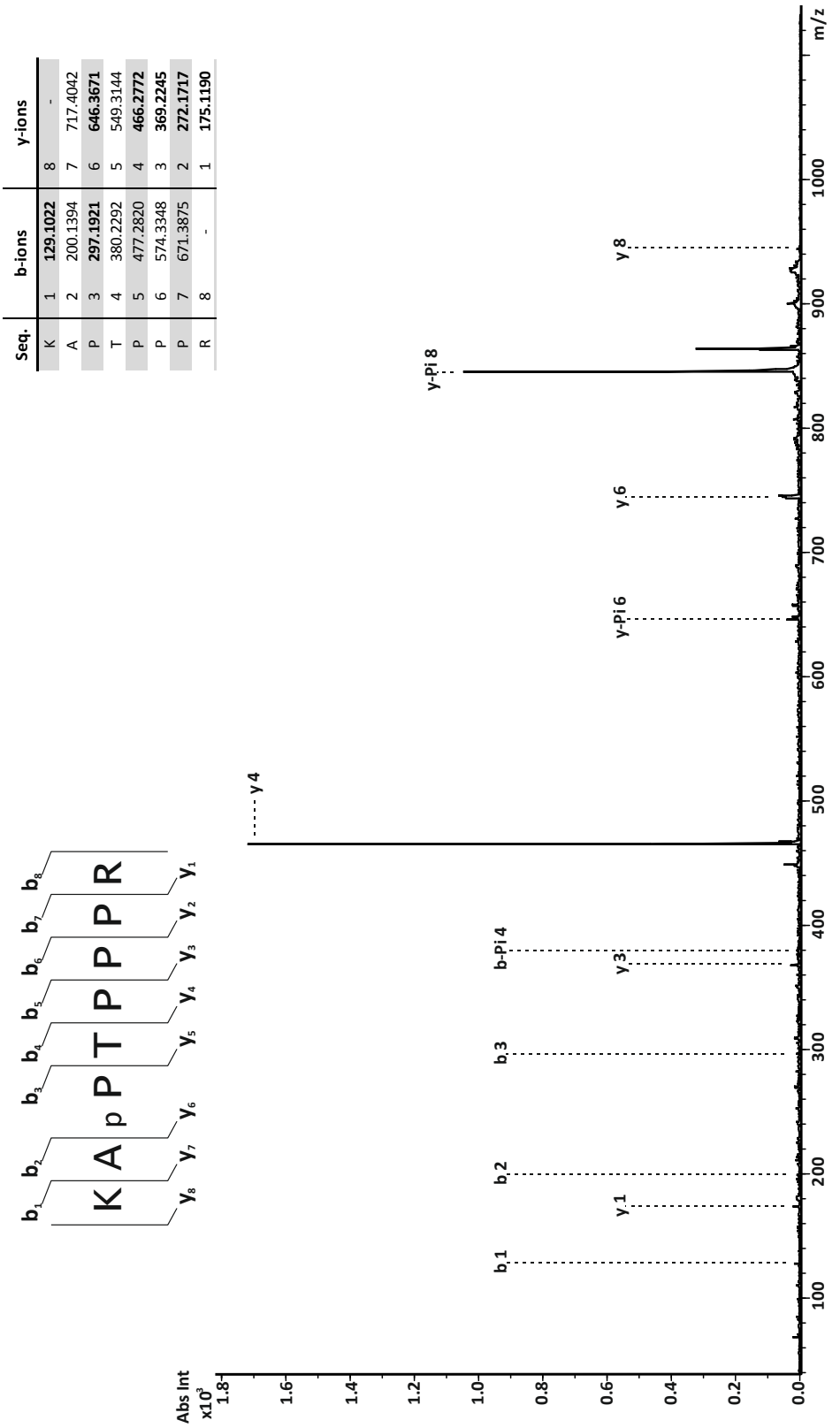
Appendix figure 8.1. MS/MS spectrum of phosphorylated serine 146.



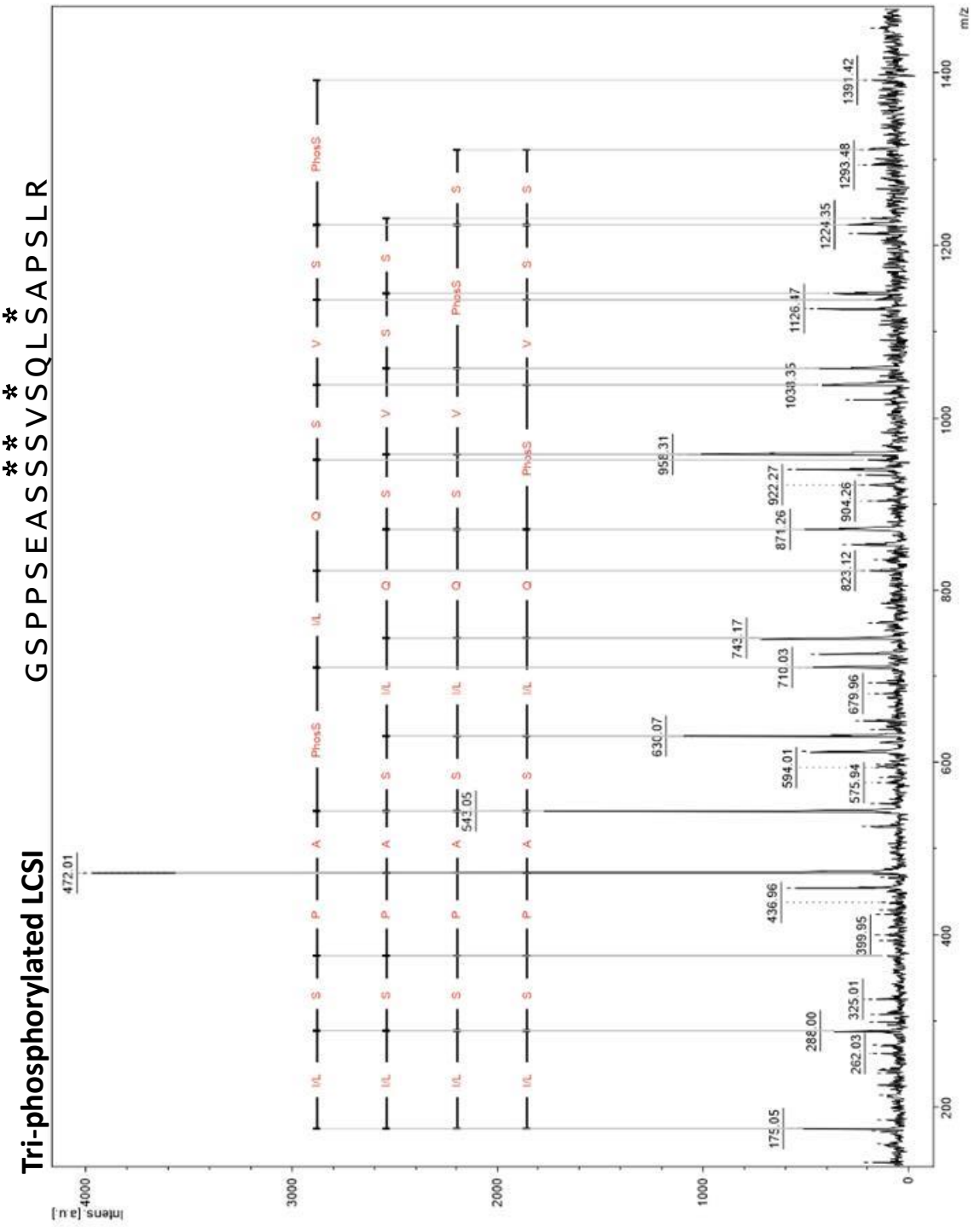
Appendix figure 8.2. MS/MS spectrum of phosphorylated serine 222.



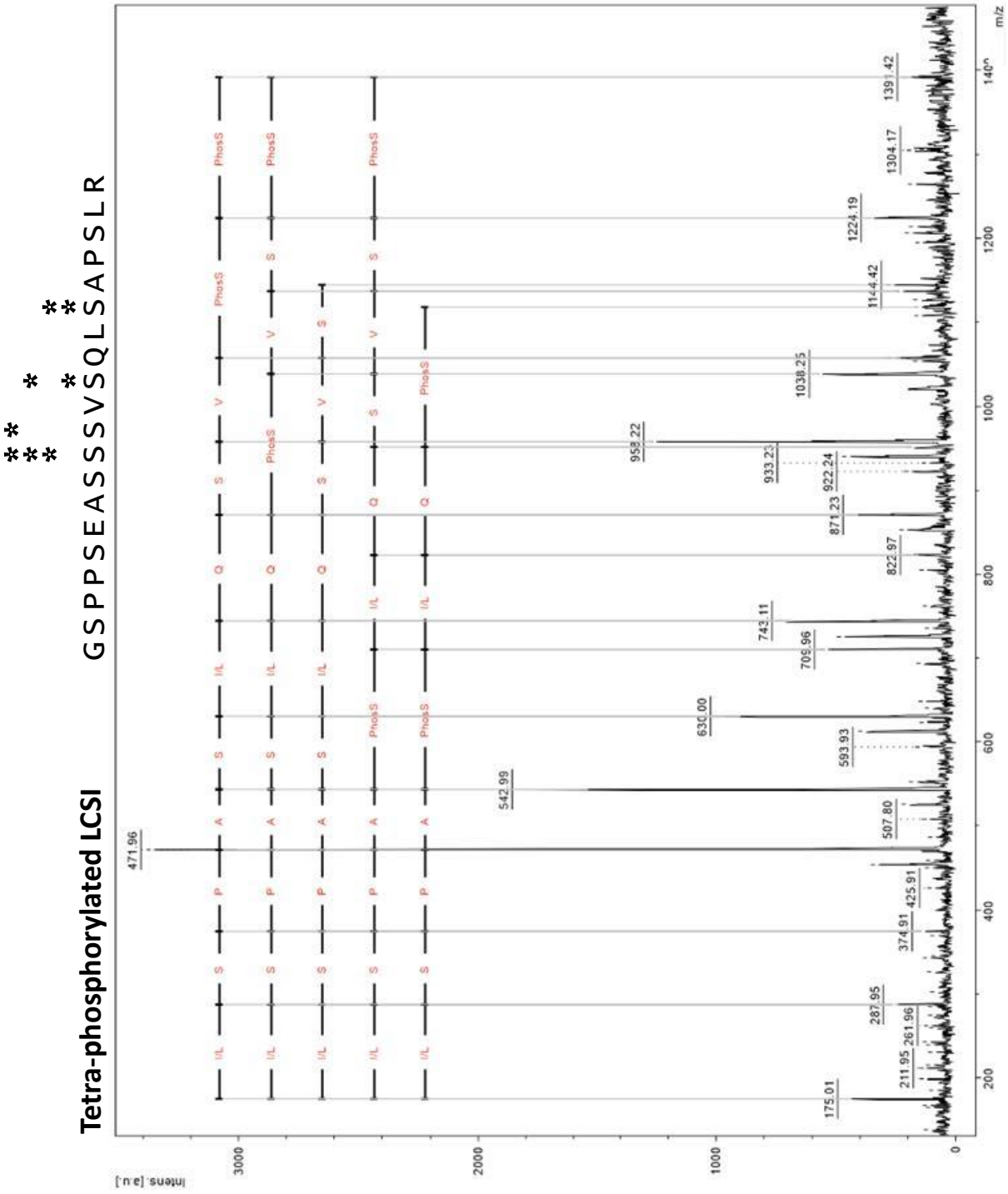
Appendix figure 8.3. MS/MS spectrum of the double phosphorylated serine 222 and 225.



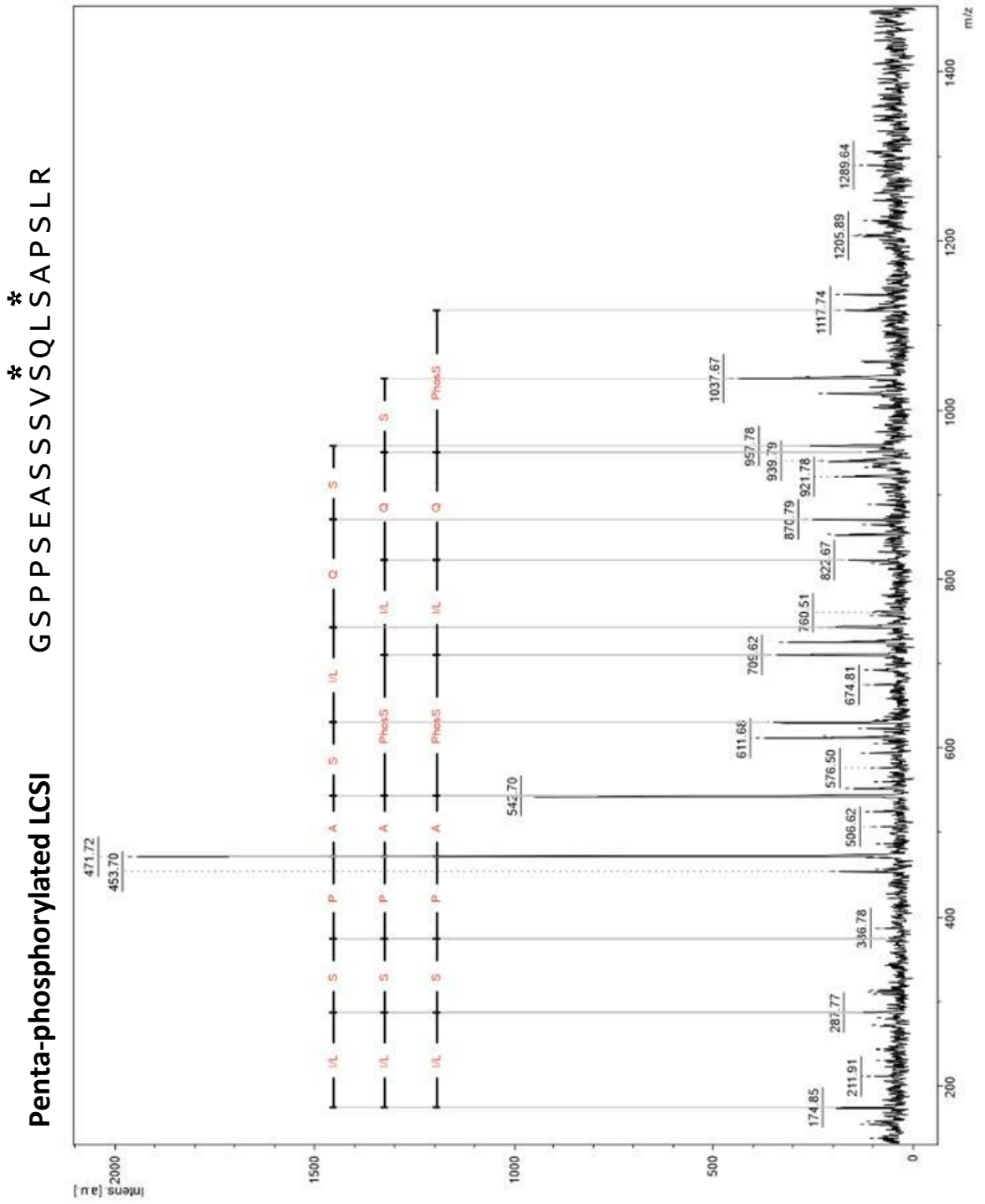
Appendix figure 8.4. MS/MS spectrum of phosphorylated threonine 348.



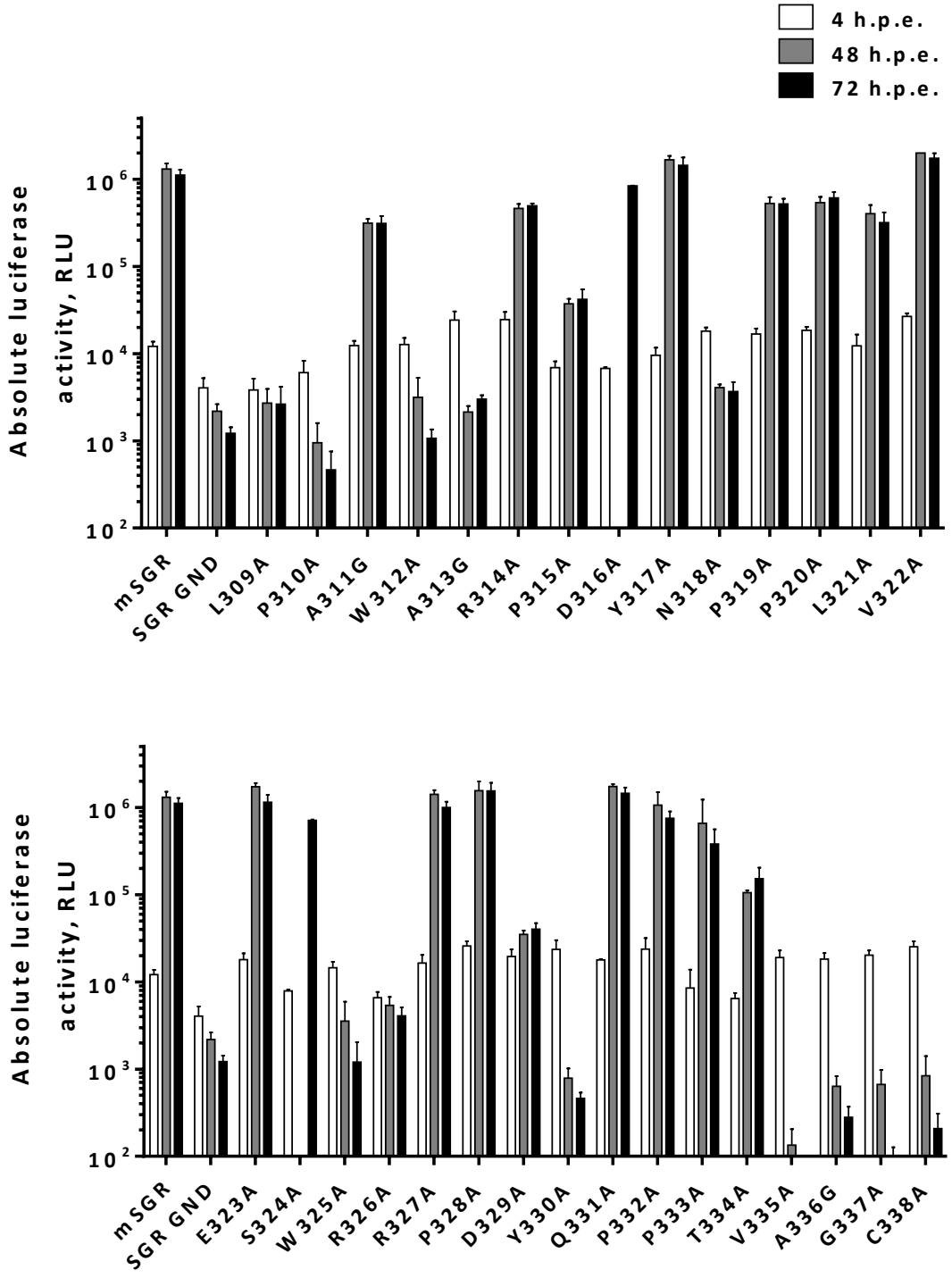
Appendix figure 8.5. MS/MS spectrum of tri-phosphorylated LCS I.



Appendix figure 8.6. MS/MS spectrum of tetra-phosphorylated LCS I.



Appendix figure 8.7. MS/MS spectrum of penta-phosphorylated LCS I.



Appendix figure 8.8. Absolute luciferase data for domain II alanine mutations in mSGR-luc.

GT 1b, ConI			GT 2a, JFH-1			
Replication		**Residue number**		**Virus replication**	**Virus assembly**	**Virus release**
Lethal	M	309	L	Lethal	n/a	Lethal
Lethal	P	310	P	Lethal	n/a	Lethal
WT	I	311	A	Impaired	Impaired	Impaired
Impaired	W	312	W	Lethal	n/a	Lethal
Impaired	A	313	A	Lethal	n/a	Lethal
WT	R	314	P	Impaired	n/a	Impaired
WT	P	315	R	Impaired	Impaired	Impaired
Impaired	D	316	D	WT	WT	WT
Lethal	Y	317	Y	WT	WT	WT
Impaired	N	318	N	Lethal	n/a	Lethal
WT	P	319	P	Impaired	Impaired	Impaired
Lethal	P	320	P	Impaired	Impaired	Impaired
WT	L	321	L	WT	WT	WT
WT	L	322	V	WT	WT	WT
WT	E	323	E	WT	n/a	WT
WT	S	324	S	Lethal	n/a	Lethal
Lethal	W	325	W	Lethal	n/a	Lethal
WT	K	326	R	Lethal	n/a	Lethal
WT	D	327	R	WT	WT	WT
WT	P	328	P	WT	WT	WT
	D	329	D	Impaired	n/a	Impaired
	Y	330	Y	Lethal	n/a	Lethal
Impaired	V	331	Q	WT	WT	WT
	P	332	P	Impaired	n/a	Impaired
	P	333	P	WT	n/a	WT
Lethal	V	334	T	Impaired	n/a	Impaired
Impaired	V	335	V	Lethal	n/a	Lethal
WT	H	336	A	Lethal	n/a	Lethal
Lethal	G	337	G	Lethal	n/a	Lethal
Lethal	C	338	C	Lethal	n/a	Lethal

Appendix figure 8.9. Comparison of domain II replication data from genotype 1b and genotype 2a.

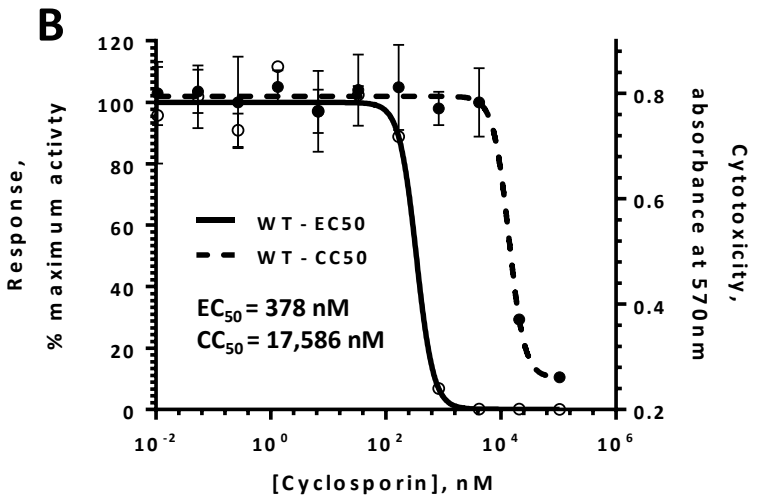
The genotype 2a data generated from the mSGR JFH-1 replicon herein, the genotype 1b data generated from the SGR-neo ConI (Tellinghuisen et al., 2008b).

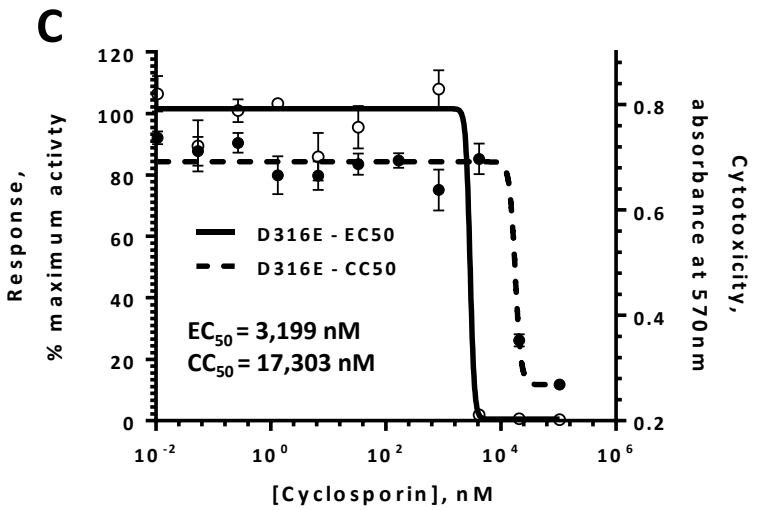
A

Agonist vs. response model

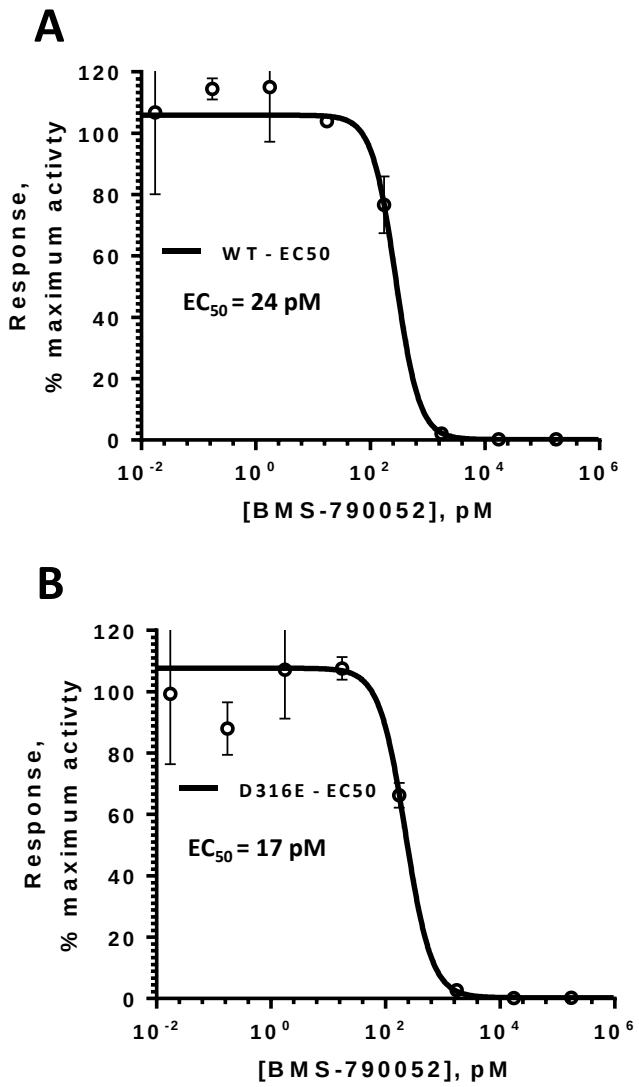
logEC50=logEC50 - (1/HillSlope)*log(50/(100-50))
Y=B + (T-B)/(1+10^((LogEC50-X)*HillSlope))

B = 100% inhibition plateau
T = 0% inhibition plateau





Appendix figure 8.10. Cyclosporin A EC50 and CC50 curves for WT and D316E mutant.
A. EC50 and CC50 mathematical model. **B.** EC50 and CC50 curves for WT mSGR treated with CsA. **C.** EC50 and CC50 curves for D316E mSGR treated with CsA.



Appendix figure 8.11. BMS-790052 (Daclatasvir) EC50 curves for WT and D316E mutant.
A. EC50 curve for WT mSGR treated with BMS-790052. **C.** EC50 curves for D316E mSGR treated with BMS-790052.

A

```
>SNAP-tag from vector pSNAPf 5,849bp (NEB, #N9183S)
atggacaaagactgcgaaatgaagcgcaccaccctggatagccctctgggcaagctggaa
M  D  K  D  C  E  M  K  R  T  T  L  D  S  P  L  G  K  L  E
ctgtctgggtgcgaacagggcctgcaccgtatcatcttcctgggcaaaggaacatctgcc
L  S  G  C  E  Q  G  L  H  R  I  I  F  L  G  K  G  T  S  A
gccgacgccgtggaagtgcctgccccagccgccgtgctgggcggaccagagccactgatg
A  D  A  V  E  V  P  A  P  A  A  V  L  G  G  P  E  P  L  M
caggccaccgcctggctcaacgcctactttcaccagcctgaggccatcgaggagttccct
Q  A  T  A  W  L  N  A  Y  F  H  Q  P  E  A  I  E  E  F  P
gtgccagccctgcaccacccagtgttccagcaggagagctttacccgccaggtgctgtgg
V  P  A  L  H  H  P  V  F  Q  Q  E  S  F  T  R  Q  V  L  W
aaactgctgaaagtggtgaagttcggagaggtcatcagctacagccacctggccgccctg
K  L  L  K  V  V  K  F  G  E  V  I  S  Y  S  H  L  A  A  L
gccggcaatcccgccgccaccgccgccgtgaaaaccgccctgagcggaaatcccgtgccc
A  G  N  P  A  A  T  A  A  V  K  T  A  L  S  G  N  P  V  P
attctgatcccctgccaccgggtggtgcagggcgacctggacgtgggggggctacgagggc
I  L  I  P  C  H  R  V  V  Q  G  D  L  D  V  G  G  Y  E  G
gggctcgccgtgaaagagtggctgctggcccacgagggccacagactgggcaagcctggg
G  L  A  V  K  E  W  L  L  A  H  E  G  H  R  L  G  K  P  G
ctgggtcctgca L  G  P  A
```

B

```
>CLIP-tag from vector pCLIPf 5,862bp NEB, (#N9215S)
atggacaaagactgcgaaatgaagcgcaccaccctggatagccctctgggcaagctggaa
 M  D  K  D  C  E  M  K  R  T  T  L  D  S  P  L  G  K  L  E
ctgtctgggtgcgaacagggcctgcaccgtatcatcttcctgggcaaaggaacatctgcc
 L  S  G  C  E  Q  G  L  H  R  I  I  F  L  G  K  G  T  S  A
gccgacgccgtggaagtgcctgccccagccgccgtgctgggcggaccagagccactgatc
 A  D  A  V  E  V  P  A  P  A  A  V  L  G  G  P  E  P  L  I
caggccaccgcctggctcaacgcctactttcaccagcctgaggccatcgaggagttccct
 Q  A  T  A  W  L  N  A  Y  F  H  Q  P  E  A  I  E  E  F  P
gtgccagccctgcaccacccagtgttccagcaggagagctttacccgccaggtgctgtgg
 V  P  A  L  H  H  P  V  F  Q  Q  E  S  F  T  R  Q  V  L  W
aaactgctgaaagtggtgaagttcggagaggtcatcagcgagagccacctggccgccctg
 K  L  L  K  V  V  K  F  G  E  V  I  S  E  S  H  L  A  A  L
gtgggcaatcccgccgccaccgccgccgtgaacaccgccctggacggaaatcccgtgccc
 V  G  N  P  A  A  T  A  A  V  N  T  A  L  D  G  N  P  V  P
attctgatcccctgccaccgggtggtgcagggcgacagcgacgtggggccctacctgggc
 I  L  I  P  C  H  R  V  V  Q  G  D  S  D  V  G  P  Y  L  G
gggctcgccgtgaaagagtggctgctggcccacgagggccacagactgggcaagcctggg
 G  L  A  V  K  E  W  L  L  A  H  E  G  H  R  L  G  K  P  G
ctgggtcctgca
 L  G  P  A
```

Appendix figure 8.12. SNAP/CLIP-tag nucleotide and protein sequence

A. Sequence of SNAP-tag from the pSNAPf vector (NEB, #N9183S). **B.** Sequence of the CLIP-tag from the pCLIPf vector (NEB, #N9215S).

```
SNAP-tag          1 ATGGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGG     50
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
CLIP-tag          1 ATGGACAAAGACTGCGAAATGAAGCGCACCACCCTGGATAGCCCTCTGGG     50

SNAP-tag         51 CAAGCTGGAACTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCTTCC    100
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
CLIP-tag         51 CAAGCTGGAACTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCTTCC    100

SNAP-tag        101 TGGGCAAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCC    150
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
CLIP-tag        101 TGGGCAAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCC    150

SNAP-tag        151 GCCGTGCTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCTGGCTCAA    200
                    ||||||||||||||||||||||||||||||.||||||||||||||||||||
CLIP-tag        151 GCCGTGCTGGGCGGACCAGAGCCACTGATCCAGGCCACCGCCTGGCTCAA    200

SNAP-tag        201 CGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCC    250
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
CLIP-tag        201 CGCCTACTTTCACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCC    250

SNAP-tag        251 TGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGG    300
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
CLIP-tag        251 TGCACCACCCAGTGTTCCAGCAGGAGAGCTTTACCCGCCAGGTGCTGTGG    300

SNAP-tag        301 AAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACAGCCACCT    350
                    ||||||||||||||||||||||||||||||||||||||.|.|||||||||
CLIP-tag        301 AAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCGAGAGCCACCT    350

SNAP-tag        351 GGCCGCCCTGGCCGGCAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCC    400
                    ||||||||||..|||||||||||||||||||||||||.|||||||||
CLIP-tag        351 GGCCGCCCTGGTGGGCAATCCCGCCGCCACCGCCGCCGTGAACACCGCCC    400

SNAP-tag        401 TGAGCGGAAATCCCGTGCCCATTCTGATCCCCTGCCACCGGGTGGTGCAG    450
                    ||..||||||||||||||||||||||||||||||||||||||||||||||
CLIP-tag        401 TGGACGGAAATCCCGTGCCCATTCTGATCCCCTGCCACCGGGTGGTGCAG    450

SNAP-tag        451 GGCGACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGCCGTGAAAGAGTG    500
                    ||||||...|||||||||..|||||..|||||||||||||||||||||||
CLIP-tag        451 GGCGACAGCGACGTGGGGCCCTACCTGGGCGGGCTCGCCGTGAAAGAGTG    500

SNAP-tag        501 GCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCCTG    550
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
CLIP-tag        501 GCTGCTGGCCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCCTG    550

SNAP-tag        551 CA    552
                    ||
CLIP-tag        551 CA    552
```

Appendix figure 8.13. Alignment of SNAP-tag and CLIP-tag nucleotide sequence.
Highlighting the 15 nucleotide changes responsible for the alteration in substrate specificity of the SNAP- and CLIP-tags.

NS5A mutation	Backbone
S146A	mSGR-luc and mJFH
S146D	mSGR-luc and mJFH
S222A	mSGR-luc and mJFH
S222D	mSGR-luc and mJFH
S225A	mSGR-luc and mJFH
S225D	mSGR-luc and mJFH
SS222/225AA	mSGR-luc and mJFH
SS222/225DD	mSGR-luc and mJFH
T348A	mSGR-luc and mJFH
T348D	mSGR-luc
S228A	mSGR-luc
S228D	mSGR-luc
S229A	mSGR-luc
S229D	mSGR-luc
S230A	mSGR-luc
S230D	mSGR-luc
SSS228/229/230AAA	mSGR-luc
SSS228/229/230DDD	mSGR-luc
S232A	mSGR-luc
S232D	mSGR-luc
S235D	mSGR-luc
S238D	mSGR-luc
NS5A-SNAP/CLIP	mSGR-luc and mJFH
S146A-SNAP	mSGR-luc
S146D-CLIP	mSGR-luc
S222A-SNAP	mSGR-luc
S222D-CLIP	mSGR-luc
S225A-SNAP	mSGR-luc
S225D-CLIP	mSGR-luc
SS222/225AA-SNAP	mSGR-luc
SS222/225DD-CLIP	mSGR-luc
T348A-SNAP	mSGR-luc
T348D-CLIP	mSGR-luc

NS5A mutation	Backbone
L309A	mSGR-luc, mJFH and pCMV10-NS3-5B
P310A	mSGR-luc, mJFH and pCMV10-NS3-5B
A311G	mSGR-luc, mJFH and pCMV10-NS3-5B
W312A	mSGR-luc, mJFH and pCMV10-NS3-5B
A313G	mSGR-luc, mJFH and pCMV10-NS3-5B
R314A	mSGR-luc, mJFH and pCMV10-NS3-5B
P315A	mSGR-luc, mJFH and pCMV10-NS3-5B
D316A	mSGR-luc and mJFH
Y317A	mSGR-luc and mJFH
N318A	mSGR-luc, mJFH and pCMV10-NS3-5B
P319A	mSGR-luc and mJFH
P320A	mSGR-luc and mJFH
L321A	mSGR-luc and mJFH
V322A	mSGR-luc and mJFH
E323A	mSGR-luc and mJFH
S324A	mSGR-luc and mJFH
W325A	mSGR-luc, mJFH and pCMV10-NS3-5B
R326A	mSGR-luc, mJFH and pCMV10-NS3-5B
R327A	mSGR-luc and mJFH
P328A	mSGR-luc and mJFH
D329A	mSGR-luc, mJFH and pCMV10-NS3-5B
Y330A	mSGR-luc, mJFH and pCMV10-NS3-5B
Q331A	mSGR-luc and mJFH
P332A	mSGR-luc and mJFH
P333A	mSGR-luc and mJFH
T334A	mSGR-luc and mJFH
V335A	mSGR-luc, mJFH and pCMV10-NS3-5B
A336G	mSGR-luc, mJFH and pCMV10-NS3-5B
G337A	mSGR-luc, mJFH and pCMV10-NS3-5B
C338A	mSGR-luc, mJFH and pCMV10-NS3-5B
D316E	mSGR-luc

Appendix figure 8.14. List of mutations created in the course of this study.

NS5A mutation	Quikchange primer, Forward	Quikchange primer, Reverse
S146A	CTTGCCAACTACCTGCTCCAGAGTTTTTC	GAAAAACTCTGGAGCAGGTAGTTGGCAAG
S146D	CCTTGCCAACTACCTGATCCAGAGTTTTTCTC	GAGAAAAACTCTGGATCAGGTAGTTGGCAAGG
S222A	CGCTTGGCACGGGGAGCACCTCCATCTGAGG	CCTCAGATGGAGGTGCTCCCCGTGCCAAGCG
S222D	CGCTTGGCACGGGGAGATCCTCCATCTGAGGCG	CGCCTCAGATGGAGGATCTCCCCGTGCCAAGCG
S225A	CGGGGATCACCTCCAGCTGAGGCGAGCTCCTC	GAGGAGCTCGCCTCAGCTGGAGGTGATCCCCG
S225D	CGGGGATCACCTCCAGATGAGGCGAGCTCCTC	GAGGAGCTCGCCTCATCTGGAGGTGATCCCCG
SS222/225AA	CGCTTGGCACGGGGAGCACCTCCAGCTGAGGCGAGCTCCTC	GAGGAGCTCGCCTCAGCTGGAGGTGCTCCCCGTGCCAAGCG
SS222/225DD	CGCTTGGCACGGGGAGATCCTCCAGATGAGGCGAGCTCCTC	GAGGAGCTCGCCTCATCTGGAGGATCTCCCCGTGCCAAGCG
T348A	CCCAAGAAGGCCCCGGCGCCTCCCCCAAGGAG	CTCCTTGGGGGAGGCGCCGGGGCCTTCTTGGG
T348D	CCCAAGAAGGCCCCGGATCCTCCCCCAAGGAGAC	CCCAAGAAGGCCCCGGATCCTCCCCCAAGGAGAC
S228A	CCTCCATCTGAGGCGGCATCCTCAGTGAGCC	GGCTCACTGAGGATGCCGCCTCAGATGGAGG
S228D	CCTCCATCTGAGGCGGACTCCTCAGTGAGCC	GGCTCACTGAGGAGTCCGCCTCAGATGGAGG
S229A	CCATCTGAGGCGAGCGCTTCAGTGAGCCAGC	GCTGGCTCACTGAAGCGCTCGCCTCAGATGG
S229D	CCATCTGAGGCGAGCGACTCAGTGAGCCAGC	GCTGGCTCACTGAGTCGCTCGCCTCAGATGG
S230A	CTGAGGCGAGCTCCGCAGTGAGCCAGCTATC	GATAGCTGGCTCACTGCGGAGCTCGCCTCAG
S230D	CTGAGGCGAGCTCCGATGTGAGCCAGCTATC	GATAGCTGGCTCACATCGGAGCTCGCCTCAG
S228/229/230A	CCTCCATCTGAGGCGGCAGCTGCAGTGAGCCAGCTATCAG	CTGATAGCTGGCTCACTGCAGCTGCCGCCTCAGATGGAGG
S228/229/230D	CCTCCATCTGAGGCGGACGACGATGTGAGCCAGCTATCAG	CTGATAGCTGGCTCACATCGTCGTCCGCCTCAGATGGAGG
S232A	GCGAGCTCCTCAGTGGCACAGCTATCAGCACC	GGTGCTGATAGCTGTGCCACTGAGGAGCTCGC
S232D	GCGAGCTCCTCAGTGGACCAGCTATCAGCACC	GGTGCTGATAGCTGGTCCACTGAGGAGCTCGC
S235D	CAGTGAGCCAGCTATCAGCACCGTCGCTGCG	CGCAGCGACGGTGCTGATAGCTGGCTCACTG
S238D	GCTATCAGCACCGTCGCTGCGGGCCACCTGC	GCAGGTGGCCCGCAGCGACGGTGCTGATAGC
NS5A-SNAP/ CLIP	GAGTCTGATCATGGAAGTAGCGGGTCATCGATGGACAAAGA CTGCGAAATGAAGCGC	CTACCTGATCAGATGAAGAACCTGAAGATCCTGCAGGACCC AGCCCAGGCTTGCCCA
L309A	GGGTTTCCACGGGCCGCACCGGCTTGGGCACG	CGTGCCCAAGCCGGTGCGGCCCGTGGAAACCC
P310A	n/a	n/a
A311G	CACGGGCCTTACCGGGTTGGGCACGGCCTGAC	GTCAGGCCGTGCCCAACCCGGTAAGGCCCGTG
W312A	n/a	n/a
A313G	CCTTACCGGCTTGGGGACGGCCTGACTACAAC	GTTGTAGTCAGGCCGTCCCCAAGCCGGTAAGG
R314A	CTTACCGGCTTGGGCAGCGCCTGACTACAACCC	GGGTTGTAGTCAGGCGCTGCCCAAGCCGGTAAG
P315A	CCGGCTTGGGCACGGGCTGACTACAACCCGC	GCGGGTTGTAGTCAGCCCGTGCCCAAGCCGG
D316A	n/a	n/a
Y317A	n/a	n/a
N318A	n/a	n/a
P319A	n/a	n/a
P320A	n/a	n/a
L321A	GACTACAACCCGCCGGCCGTGGAATCGTGGAG	CTCCACGATTCCACGGCCGGCGGGTTGTAGTC
V322A	CTACAACCCGCCGCTCGCGGAATCGTGGAGGAGG	CCTCCTCCACGATTCCGCGAGCGGCGGGTTGTAG
E323A	CAACCCGCCGCTCGTGGCATCGTGGAGGAGGCCAG	CTGGCCTCCTCCACGATGCCACGAGCGGCGGGTTG
S324A	CCGCCGCTCGTGGAAGCGTGGAGGAGGCCAG	CTGGCCTCCTCCACGCTTCCACGAGCGGCGG
W325A	n/a	n/a
R326A	GCTCGTGGAATCGTGGGCGAGGCCAGATTACCAAC	GTTGGTAATCTGGCCTCGCCCACGATTCCACGAGC
R327A	GTGGAATCGTGGAGGGCGCCAGATTACCAACC	GGTTGGTAATCTGGCGCCCTCCACGATTCCAC
P328A	GAATCGTGGAGGAGGGCAGATTACCAACCGC	GCGGTTGGTAATCTGCCCTCCTCCACGATTC
D329A	CGTGGAGGAGGCCAGCTTACCAACCGCCCAC	GTGGGCGGTTGGTAAGCTGGCCTCCTCCACG
Y330A	GTGGAGGAGGCCAGATGCCCAACCGCCCACCGTTG	CAACGGTGGGCGGTTGGGCATCTGGCCTCCTCCAC
Q331A	GAGGAGGCCAGATTACGCACCGCCCACCGTTGC	GCAACGGTGGGCGGTGCGTAATCTGGCCTCCTC
P332A	n/a	n/a
P333A	n/a	n/a
T334A	GATTACCAACCGCCCGCCGTTGCTGGTTGTG	CACAACCAGCAACGGCGGGCGGTTGGTAATC
V335A	GATTACCAACCGCCCACCGCTGCTGGTTGTGCTCTCCC	GGGAGAGCACAACCAGCAGCGGTGGGCGGTTGGTAATC
A336G	CAACCGCCCACCGTTGGTGGTTGTGCTCTCCC	GGGAGAGCACAACCACCAACGGTGGGCGGTTG
G337A	CGCCCACCGTTGCTGCTTGTGCTCTCCCCCC	GGGGGGAGAGCACAAGCAGCAACGGTGGGCG
C338A	CCCACCGTTGCTGGTGCTGCTCTCCCCCCCCC	GGGGGGGGGAGAGCAGCACCAGCAACGGTGGG
D316E	CTTGGGCACGGCCTGAGTACAACCCGCCGCTC	GAGCGGCGGGTTGTACTCAGGCCGTGCCCAAG

Appendix figure 8.15. List of oligonucleotide primers used to generate mutations.

Primer name	Primer sequence	Region sequenced
SeqPrimer1	ACATGGACTTCCCGTGTCC	Start of NS3
SeqPrimer2	ACTGGCAGTGGAAAGAGCA	Mid NS3
SeqPrimer3	CTGGACCCCACCTTCACTA	End of NS3
SeqPrimer4	TAAGGAGGTCCTGTATGAG	NS4
SeqPrimer5	GGATGCGTCGCAGCGTGTGA	Start of NS5A
SeqPrimer6	CACGGCGGAGACTGCGGCG	Mid NS5A
SeqPrimer7	GGGTAGCTCCCGGTTCGGG	End NS5A
SeqPrimer8	AACGGGTGGAGTATCTCTT	Mid NS5B
SeqPrimer9	TGGGTTCGCATGGTCCTAA	End NS5B

Appendix figure 8.16. List of sequencing primers

And they all lived happily ever after……

….the end