Structural and electrophysiological analysis of Hepatitis C Virus p7

Benjamin P Oestringer
St John's College

Prof Zitzmann laboratory, Oxford Glycobiology Institute, Department of Biochemistry
University of Oxford

A thesis submitted in partial fulfillment of the requirements for the degree of *Doctor Philisophiaæ(DPhil)* at the University of Oxford

Trinity Term 2013

Structural and electrophysiological analysis of Hepatitis C Virus p7

Benjamin P Oestringer, St John's College

A thesis submitted in partial fulfillment of the requirements for the degree of *Doctor Philisophiaæ(DPhil)* at the University of Oxford, Trinity Term 2013

Abstract

Infection with the hepatitis C virus (HCV) has a big impact on global health. It is estimated that approximately 3 % of the world's population carry HCV, putting more than 200 million people at risk of developing severe liver disease, including chronic hepatitis, liver cirrhosis and hepatocellular carcinoma. The HCV encoded viroporin p7 forms ion channels that are crucial for the assembly and secretion of infectious viruses, making it a potential drug target. Its hydrophobic nature makes p7 notoriously difficult to investigate in an untagged native form. A previously determined 16 Å electron microscopy single-particle reconstruction in detergent showed a hexameric, flower-shaped p7 protein. In conjunction with one hexameric and several monomeric p7 solution state NMR structures published, this constitutes the currently available structural information framework. An *E. coli* expression system is introduced, which is especially adapted to express isotopically labeled p7. For the first time, suitable solution-state NMR conditions at physiological pH and temperature were identified that gave rise to high quality spectra suitable to interrogate iminosugar drug interactions with untagged isotopically labeled J4 p7 (C27S) solubilised in Cyclofos-6. A novel secondary structure topology was observed and preliminary iminosugar binding sites were determined. Further, a DIB (droplet interface bilayer) system to analyse p7 ion channel function was established, which is suitable to elucidate how inhibitors act on p7 genotypes and how different lipids influence the ion channel function of p7. The p7 oligomeric state was further investigated using native gel analysis, showing that isolates representing HCV genotypes 1 - 6 form oligomeric complexes. An ion channel defective dibasic mutant implicated in severely compromising viral fitness is also shown for the first time to form an oligomer, implicating that it is not an assembly problem that leads to the abrogated function.

Acknowledgements

I want to thank my supervisor Prof Nicole Zitzmann for giving me the opportunity to work in her lab on such a challenging project and the continuous support, discussions and patience. I also want to thank Prof Raymond Dwek for his continuous support and suggestions and especially for making the Baruch-Blumberg Scholarship possible.
I especially want to thank Dr Jason Schnell for his support and help with the NMR project and comments for this manuscript.
I also want to thank Dr Mark Wallace and everyone in his lab for help with the DIB system. Special thanks go to Bryd, Lydia, Lauriane and Jason.
I am grateful for the help of Jacob Brady with the Rosetta modeling. Further, I would like to thank Dr David Staunton and Prof David Harvey for help with MS and Dr Robert Gilbert for help with the AUC. I am very grateful to everyone in the Zitzmann laboratory for creating such a special and nice working atmosphere, especially to Stephen Woodhouse, Dina Fotinou, Pietro Roversi, Jo Miller, Bevin Gangadharan and Stephanie Pollock.
A thanks goes also to all the other DPhil students past and present in the Zitzmann lab, including Emma, JL, Drew, Mel, Simon, Phillip and Kelly. A special thanks goes to Alex, who greatly contributed to the optimisation of the p7 purification while he was a Wellcome-Trust rotation student under my supervision. A very special dankeschön goes to Barbara.
And finally I want to thank my family who always supported me throughout my scientific endeavors.

Contents

List of Figures

List of Tables

Abbreviations

aa	**A**mino **a**cid
ARV	**A**vian **reov**irus
ARFP	**A**lternative **r**ibosomal **f**rameshift **p**rotein
ANN	**A**rtificial **n**eural **n**etwork
AUC	**A**nalytical **u**ltra **c**entrifugation
apo	**A**polipoprotein
BIT225	*N*-[5-(1-methyl-1*H*-pyrazol-4-yl)-napthalene-2-carbonyl]-guanidine
BLM	**B**lack **l**ipid **m**embrane
BN-PAGE	**B**lue **n**ative **p**age
$\mathbf{C_8E_5}$	*n*-octyl-hydroxypenta(oxyethylene)
Cov	**Co**rona**v**irus
CLEANEX	**CLEAN** chemical **EX**change
CNBr	**C**yanoge**n** **br**omide
CMC	**C**ritical **mi**celle **c**oncentration
CSP	**C**hemical **s**hift **p**erturpation
CHCA	α-**C**yano-4-**h**ydroxy**c**innamic **a**cid
Cyclofos-6	6-cyclohexyl-1-hexylphosphocholine
1D	One-dimensional
DAA	**D**irect **a**cting **a**ntiviral
DGAT1	**D**ia**g**lycerol O-**a**cyl**t**ransferase 1
DHB	**D**roplet **h**ydrogel **b**ilayer
DH6PC	1,2-**di****h**exanoyl-*sn*-glycero-3-**p**hospho**c**holine
DH7PC	1,2-**di****h**eptanoyl-*sn*-glycero-3-**p**hospho**c**holine
DIB	**D**roplet **i**nterface **b**ilayer
DMPG	1,2-**di****m**yristoyl-*sn*-glycero-3-**p**hosp**g**lycerol

DMSO	**D**i**m**ethyl **s**ulf**o**xide
DOPC	1,2-**d**ioleoyl-*sn*-glycero-3-**p**hospho**c**holine
DPhPC	1,2-**di**-**ph**ytanoyl-*sn*-glycero-3 **p**hospho**c**holine
DSS	4,4-**d**imethyl-4-**s**ilapentane-1-**s**ulfonic acid
DTT	**D**i**t**hrio**t**hreitol
DPC	*n*-**d**odecyl**p**hospho**c**holine
E. coli	*Escherichia coli*
EDTA	**E**thylene**d**iamine**t**etraacetic **a**cid
EM	**E**lectron **m**icroscopy
ER	**E**ndoplasmic **r**eticulum
ESI-MS	**E**lectro**s**pray **i**onisation-**m**ass **s**pectrometry
FA	**F**ormic **a**cid
FID	**F**ree **i**nduction **d**ecay
Fmoc	**F**luorenyl**m**ethyl**o**xy**c**arbonyl
Fos-choline-10	*n*-decylphosphocholine
Fos-choline-12	*n*-dodecylphosphocholine
Fos-choline-14	*n*-tetradecylphosphocholine
FPLC	**F**ast **p**rotein **l**iquid **c**hromatography
FT	**F**ourier **t**ransformation
gt	**G**eno**t**ype
HA	**H**em**a**glutinin
HCl	**H**ydrogen **c**h**l**oride
HCV	**H**epatitis **C** **v**irus
HCVcc	**HCV** **c**ell **c**ulture system
hetNOE	**het**eronuclear **NOE**
HIV-1 Vpu	**H**uman **i**mmunodeficiency **v**irus-**1** **v**iral **p**rotein **u**
HLA	**H**uman **l**eukocyte **a**ntigen
HMA	5-(*N,N*-**H**exa**m**ethylene) **a**miloride
HMQC	**H**eteronuclear **m**ultiple **q**uantum **c**oherence
HPLC	**H**igh-**p**erformance **l**iquid **c**hromatography
HFIP	1,1,1,3,3,3-hexaflouro-2-propanol
HRSV	**H**uman **r**espiratory **s**yncytial **v**irus
HSQC	**H**eteronuclear **s**ingle **q**uantum **c**oherence

HTA	**H**ost-**t**argeted **a**gent
IAV	**I**nfluenza **A** **v**irus
IFNL3	**I**nterferon, **l**ambda 3
IMAC	**I**mmobilised **m**etal ion **a**ffinity **c**hromatography
INEPT	**I**nsensitive **n**ucleus **e**nhanced by **p**olarisation **t**ransfer
IPTG	**I**so**p**ropyl β-D-1-**t**hio**g**alactopyranoside
IRES	**I**nternal **r**ibosomal **e**ntry **s**ite
ITC	**I**so**t**hermal **c**alorimetry
JCV	**J**ohn **c**unningham polyoma**v**irus
JFH-1	**J**apanese **f**ulminant **h**epatitis-1
Kan	**Kan**amycin
LB	**L**ysogeny **b**roth
LCS	**L**ow **c**omplexity **s**equence
LD	**L**ipid **d**roplet
LMPG	1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-rac-glycerol)
LVP	**L**ipo**v**iro**p**article
M1	**M**atrix protein **1**
M2	**M**atrix protein **2**
MALDI-TOF	**M**atrix-**a**ssisted **l**aser **d**esorption/**i**onization - **t**ime **o**f **f**light
MAPS	**M**icrowave-**a**ssisted **p**eptide **s**ynthesis
MES	2-(*N*-**m**orpholino)**e**thane**s**ulfonic acid
miR-122	**mi**cro**R**NA-**122**
MS	**M**ass **s**pectrometry
MW	**M**embranous **w**eb
MWCO	**M**olecular **w**eight **c**ut **o**ff
NA	**N**euro**a**minidase
NaN_3	Sodium azide
NaOH	Sodium hydroxide
NH_4CL	Ammonium chloride
NMR	**N**uclear **m**agnetic **r**esonance
***N*N-DGJ**	***N*-n**onyl-**d**eoxy**g**alacto**j**irimycin
***N*N-DNJ**	***N*-n**onyl-**d**eoxy**n**o**j**irimycin
NOE	**N**uclear **o**verhauser **e**ffect

NOESY	**NOE** spectroscop**y**
NPHV	**N**on-**p**rimate **h**epaci**v**irus
ORF	**O**pen **r**eading **f**rame
O/N	**O**ver**n**ight
14-*O*-PC/6-*O*-PC	1,2-di-*o*-hexyl-*sn*-glycero-3-phosphocholine
PA	**P**hosphatidic **a**cid
PC	**P**hosphatidyl **c**holine
PCR	**P**olymerase **c**hain **r**eaction
PDB	**P**rotein **D**ata **B**ank
PE	**P**hosphatidyl**e**thanolamine
PMMA	**P**oly(**m**ethyl 2-**m**ethyl**p**ropenoate)
POPC	1-**P**almitoyl-2-**o**leoyl-3-*sn*-glycero-3-**p**hospho**c**holine
POPS	1-**P**almitoyl-2-**o**leoyl-3-*sn*-glycero-3-**p**hospho**s**erine
ppm	**p**arts **p**er **m**illion
PVDF	**P**oly**v**inyli**d**ene **f**luoride
RDC	**R**esidual **d**ipolar **c**oupling
RdRp	**R**NA-**d**ependent **r**NA **p**olymerase
rpm	**R**evolutions **p**er **m**inute
SARS	**S**evere **A**cute **R**espiratory **S**yndrome
SDS-PAGE	**S**odium **d**odecyl **s**ulphate **p**oly**a**crylamide **G**el **e**lectrophoresis
SEC	**S**ize-**e**xclusion **c**hromatography
SEC-MALS	**S**ize **e**xclusion **c**hromatography - **m**ulti-**a**ngle static **l**ight **s**cattering
SOFAST-HMQC	**S**elective **o**ptimized-**f**lip-**a**ngle **s**hort-**t**ransient-**HMQC**
SP	**S**ignal **p**eptidase
SPP	**S**ignal **p**eptide **p**eptidase
SPPS	**S**olid-**p**hase **p**eptide **s**ynthesis
SPR	**S**urface **p**lasmon **r**esonance
SRB1	**S**cavenger **R**eceptor class **B** member **1**
ssNMR	**S**olid **s**tate **NMR**
ssRNA	**S**ingle-**s**tranded **RNA**
SVR	**S**ustained **v**irological **r**esponse
SV	**S**indbis **V**irus
TCEP	**T**ris (2-**c**arboxy**e**thyl) **p**hosphine hydrochloride

TCL	**T**otal **c**ell **l**ysate
TFE	2,2,2-**T**ri**f**luoro**e**thanol
TM	**T**rans**m**embrane
TRACT	**T**ROSY for **r**ot**a**tional **c**orrelation **t**imes
TRL	**T**riglycerine-**r**ich **l**ipoproteins
TROSY	**T**ransverse **r**elaxation **o**optimised **s**pectroscopy
TOCSY	**To**tal **c**orrelation **s**pectroscop**y**
UTR	**U**n**t**ranslated **r**egion
VLDL	**V**ery **l**ow **d**ensity **l**ipoprotein
Wang resin	[4-hydroxymethyl-(phenoxymethyl)]-polystyrene

Chapter I

Introduction

I.1 Hepatitis C Virus

I.1.1 HCV Pathology

Infection with hepatitis C virus (HCV) has a big impact on global health. It is estimated that approximately 3 % of the world's population carry HCV, putting more than 200 million people at risk of developing severe liver disease, including chronic hepatitis, liver cirrhosis and hepatocellular carcinoma (reviewed in[1, 2]).

I.1.2 HCV virology

HCV displays high genetic heterogeneity and is currently classified into seven genotypes (gt 1-7) and more than a hundred subtypes[3]. It is a small, enveloped virus with a particle size ranging between 40 - 100 nm in diameter and it harbours a 9.6 kb positive-sense, single stranded RNA genome ((+) ssRNA). HCV was initially discovered and described as non-A, non-B hepatitis in 1975[4]. It was identified as HCV in 1989 in a screen of a random-primed cDNA library, constructed from HCV-containing plasma[5]. Within the *Flaviviridae* family[6], HCV is the type species of the *Hepacivirus* genus, with an assignment to a viral order still pending. The hypothesis that non-human primates are the source of HCV infection in humans is in discussion but controversial[7]. Other non-primate mammals have recently been found to be susceptible to hepaciviruses. One study found that non-primate hepacivirus (NPHV) causes pulmonary infections in dogs,

which was later also identified in horses. The latter are thought to be the true host of NPHV with sporadic cross-species infections to dogs ([7, 8], reviewed in[9]). RNA viruses in general and HCV in particular have a high *in vivo* mutation frequency, with HCV showing about 2.5 x 10^{-5} mutations per nucleotide per genome replication[10]. It is suggested that HCV like other RNA viruses exists as quasispecies, which are diverse, genetically-close variants linked to a master sequence, i.e. a genotype sequence[11].

I.1.3 The HCV genome

The HCV genome is composed of a 5'-UTR (untranslated region), which includes an IRES (internal ribosome entry site), an ORF (open reading frame) that encodes structural and non-structural proteins, and a 3'-UTR[12, 13]. It is translated via the 5' IRES as a polyprotein, which is co- and post-translationally cleaved by cellular and viral proteases to produce ten viral proteins (excluding an alternative translation product, see below)[12–14]. The amino-terminal part of the polyprotein consists of the structural proteins that make up the virion and include the capsid-forming core (C) protein and the two envelope glycoproteins E1 and E2. They are followed by the p7 viroporin and the non-structural proteins NS2 cysteine protease, NS3 serine protease and RNA helicase, the NS3 cofactor NS4A, the membranous web-inducing factor NS4B, the phosphoprotein NS5A and the NS5B RNA-dependent RNA polymerase[12–14] (Figure I.1). An alternative translation product referred to as F (frameshifted) protein or ARFP (alternative ribosomal frameshift protein) can be expressed through a ribosomal frameshift within the capsid-encoding sequence[12](Figure I.1). Its function is largely unknown but it is suggested to play a role in hepatocellular transformation in HCV patients ([15], reviewed in[16]).

I.1.4 HCV protein function

I.1.4.1 Core

Core is a dimeric, predominantly α-helical protein with a molecular weight of 21 kDa and a part of the HCV virus particle, with the paramount function to assemble into the viral capsid[17, 18]. Core is likely implicated in hepatic steatosis and carcinoma as shown in mouse studies, both of which are suggested to be caused by perturbation of cell

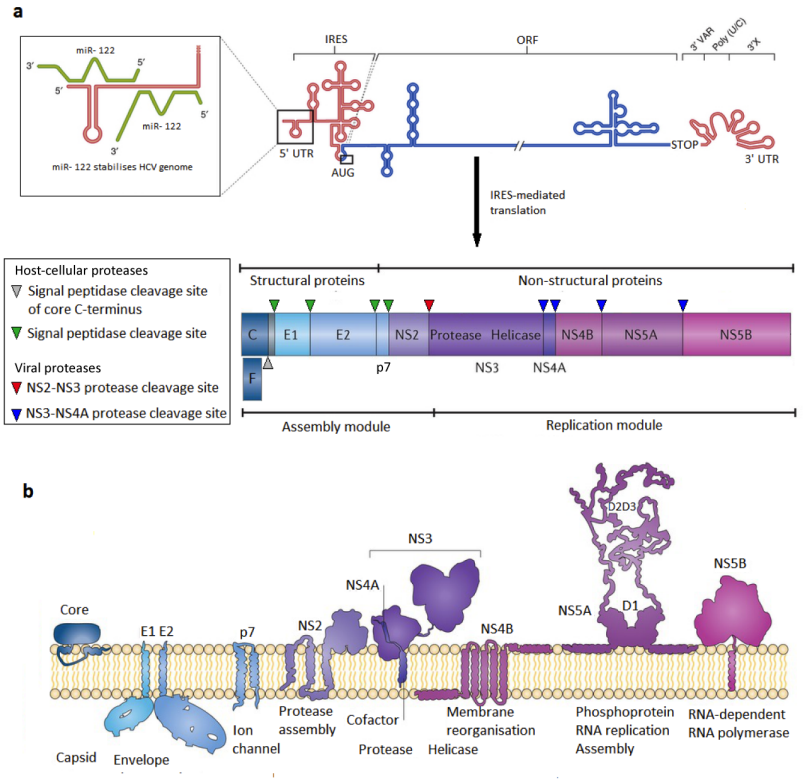

Figure I.1 – The HCV genome. a) The HCV genome contains an ORF (open reading frame, blue), with a 5' and 3' UTR (untranslated region), a start and stop codon flanking its ends (AUG and STOP, respectively). Binding of hepatic microRNA miR-122 (green) to the 5' UTR is shown in the left insert. The 5' UTR contains an IRES (internal ribosome entry site), via which translation of the polyprotein is initiated, that is co- and post-translationally processed into ten viral proteins (bottom). The structural proteins core (C) and glycoproteins E1 and E2 together with p7 and NS2 make up a module required for virus assembly. The F protein, whose function is largely unknown, is depicted below the core protein. The replication module for RNA replication consists of the remainder of the non-structural proteins, i.e. NS3, NS4A, NS4B, NS5A and NS5B. Cleavage of the polyprotein occurs via cellular signal peptidases and viral proteases. C-terminal cleavage of core signal peptide occurs via SPP (signal peptide peptidase, grey triangle) and SP (signal peptidase) cleaves core-E1, E1-E2 and E2-p7 (green triangles). The autocleavage of NS2-NS3 is highlighted by a red triangle. NS3-NS4A cleavage sites in NS3, NS4A, NS4B, NS5A and B are indicated by blue triangles. b) Membrane topologies and major functions of the HCV polyprotein cleavage products are shown. Single or multipass transmembrane helices connect most HCV proteins to intracellular compartments. The core protein and NS5A are tethered to the membrane via amphipathic helices, which do not span the membrane but rather dock on the lipid headgroups. NS5A is shown as dimer. All other proteins are depicted as monomers, while most if not all HCV proteins form homo- (core) or heterodimers (E1-E2), or oligomers (p7) (Figure adapted from [13, 14]).

proliferation regulation[19, 20]. Core consists of two distinct domains (D1 and D2) and an E1 signal sequence at its C-terminus[21, 22], which together make up an immature core protein of 191 aa (amino acids)[17, 18, 23]. The E1 signal sequence is cleaved off the E1 N-terminal end by SP (signal peptidase) and then gets further processed by SPP (signal peptide peptidase), via intramembrane proteolysis, which leads to the mature core protein of about 173 - 179 amino acids[12, 24, 25]. The highly basic D1 (aa 1-117) contains an RNA binding domain[26, 27], which also shows RNA chaperone functions[28, 29], e.g. enabling HCV 3'-UTR dimerisation[28, 30]. A solution NMR (nuclear magnetic resonance) structure of an N-terminal fragment (aa 2-45) was solved and shows a helix-turn-helix motif involving amino acids 15-41, several α-helices[22] and the immunodominant antigenic region of core (unpublished, see PDB (Protein Data Bank) entry 1CWX [M. Jolivet, M. F. Penin, P. Dalbon, L. Ladaviere, and X. Lacoux, patent application EP 1015481, WO 98/39360, 1997]). In isolation, the N-terminal core protein (aa 1-82) is suggested to be intrinsically unstructured, explaining its ability to interact with several host proteins[22, 31]. The C-terminal D2 domain (aa 118 to about 177) is hydrophobic and consists of two amphipathic α-helices (comprising aa 119-136 and 148-165) that are suggested to interact with the bilayer and thus tethering core to it[13, 22, 32]. The two helices are connected via a short and flexible loop (aa 137-147), which are suggested to interact with LDs (lipid droplets)[21, 23, 32, 33]. D2 is also crucial for the correct folding of full length core following SPP cleavage[32, 34].

I.1.4.2 E1 & E2

HCV envelope type I transmembrane (TM) glycoproteins E1 and E2 were first described in 1990[35], and contain an N-terminal ectodomain (between 160 and 260 aa) and a C-terminal transmembrane region of approximately 30 aa[22, 36]. E1 and E2 play important roles in the HCV life cycle, including particle assembly, HCV entry, and fusion with the endosomal membrane (reviewed in[37]). During maturation in the ER (endoplasmic reticulum), the C-terminus of the core protein acts as a signal sequence, which is responsible for the translocation of the E1 ectodomain into the ER lumen. There, both E1 and E2 are modified post-translation via N-linked glycosylation[38], with the glycosylated residues being crucial for correct protein folding and HCV entry[39]. After cleavage by SP, the TM domains change their orientation with the C-terminus facing the cytosol[40]. E1 and E2 form heterodimers that are noncovalently linked and suggested to

be the main components of the viral envelope[18, 33]. Dimerisation is dependent on the transmembrane components of E1 and E2[41]. Virion-associated E1 and E2 form large complexes via disulphide bridges between highly conserved cysteine residues[42]. HCV envelope protein maturation and folding thus involves ER chaperone folding, glycosylation and disulphide bond formation[22]. E1/E2 complexes are suggested to be displayed on HCV particles, which was shown by an elegant EM study that employs affinity-tagged E2 to purify and characterise HCV particles[43]. E2 shows an overall globular structure consisting mostly of β-strands and two short α-helices. It also includes regions lacking proper secondary structures, such as random coils and disordered loops (PDB entries 4MWF[44] and 2GSI[45]). A structural feature in the stem region of E2 (aa 322-332) comprises an α-helix and unstructured areas (PDB entry: 2KZQ), which are suggested to be essential for HCV cell entry[22, 46]. The observed immunoglobulin fold of E2, its compact structure and inability to oligomerise[44, 45] suggest a different mechanism to canonical class II fusion proteins[47]. The HCV fusion mechanism might be either based on E1 alone or on the E1/E2 heterodimer[45].

I.1.4.3 NS2

HCV NS2 was first described in 1993[48] and contains a cysteine-protease domain[49, 50], which cleaves the HCV polyprotein at the NS2/NS3 junction[22, 51]. It consists of a transmembrane domain, which is made up of three transmembrane segments (aa 4-23, 27-49, and 72-94) and an α-helix (aa 61-70), for all of which structures have been solved by solution-state NMR (PDB entries 2JY0, 2KWT and 2KWZ)[52, 53]. In its active form, NS2 is a homodimer, with its catalytic triad, i.e. histidine/glutamic acid/cysteine, being split between two monomers[49]. The crystal structure of the NS2 catalytic domain (PDB entry: 2HD0) shows a dimeric protease with two composite active sites[49]. In each site, the histidine and glutamic acid residues are situated on the same monomer, with the cysteine being contributed by another monomer[49, 50]. As a cofactor, NS2 requires the first two amino acids of NS3 for basal proteolytic activity and the NS3 Zn^{2+}-binding domain (NS3 aa 81-213) for enhanced proteolytic activity[50]. NS2 plays an important role in HCV particle assembly and interacts with structural and non-structural proteins, including E1, E2, p7, NS3 and NS5A[53–55]. Both, the NS2-E2 interaction and NS2 subcellular localisation rely on a part of the E2 transmembrane domain, which shows that NS2 functions can be independent of its protease activity[56].

NS2 also interacts with gene promoters to inhibit cytokine gene expression and to down regulate host cell antiviral responses[57].

I.1.4.4 NS3 & NS4A

NS3 is a polydomain and multifunctional 70 kDa protein, with a chymotrypsin-like serine protease at its N-terminus (aa 1-180) and an NTPase/superfamily 2 DExH/D-box RNA helicase[58, 59] at its C-terminal part (aa 181-631). The protease region is composed of two β-barrel subdomains separated by a cleft containing the active site, in which the catalytic triad is formed by His57, Asp81, and Ser139 and an α-helix (α0, aa 12-23), the latter tethering NS3 to intracellular membranes. It is also suggested to be responsible for subcellular localisation of the whole NS3/NS4A complex[13, 60–63] (Figure I.2). A Zn^{2+} stabilises the C-terminal NS3 protease subdomain (NS3 Zn^{2+} binding domain, aa 81-213). It is coordinated by Cys97, Cys99, Cys145, and His149[22] and aides in processing of the NS2/NS3 cleavage site[50]. The N-terminal protease part of NS3 is non-covalently complexed to a 54 amino acid NS4A polypeptide, which acts as a cofactor[22, 64–68](Figure I.2). Functions attributed to it include stabilisation of the NS3 protease and completion of the substrate recognition site by incorporation of a β-strand (aa 21-32) into the N-terminal NS3 β-barrel. It also interacts with the NS3 helicase and other proteins involved in HCV RNA replication, such as NS4B, and in assembly, such as NS3[13, 22, 69]. Like other helicases in its family[58, 59], the NS3 helicase hydrolyses ATP (adenosine triphosphate) and unwinds double-stranded or single-stranded RNA with extensive secondary structure content. The helicase is essential for HCV RNA replication[70] and particle assembly[71]. The exact function of NS4A, its role in the viral life cycle and why it is physically linked to the protease domain, are still being discussed[22].

I.1.4.5 NS4B

NS4B is a hydrophobic, polytopic and integral transmembrane protein of 27 kDa comprising 261 amino acids. Its transmembrane region (aa 70-190) is predicted to be composed of four transmembrane segments, an N-terminal part (aa 1-69) consisting of two amphipathic α-helices, designated AH1 and AH2 (aa 3-45 and 42-66, respectively) and

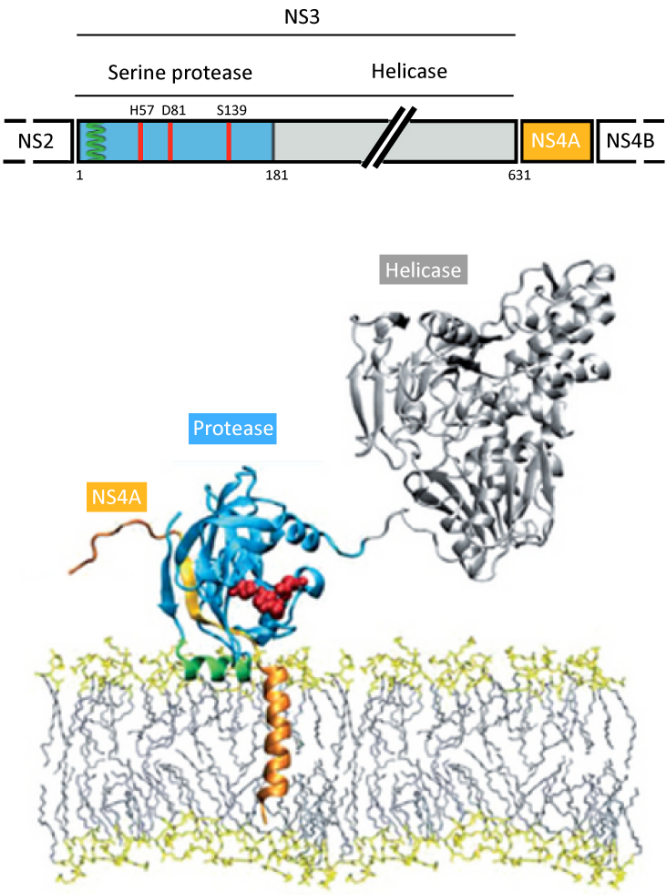

Figure I.2 – Structure and membrane topology of HCV NS3-4A. A schematic depiction of the HCV NS3A-NS4 genomic region of the HCV polyprotein is shown (top). The NS3 serine protease is depicted in blue and the helicase in grey. The cofactor NS4A is shown in orange. Cleavage sites are indicated by arrows (the white arrow indicates the NS2 cleavage site and the black arrows NS3-4A cleavage sites, respectively). The three catalytic residues of NS3A are shown in red (i.e. H57, D81 and S139). The amphipathic NS3 anchor helix (α0) is shown in green. The stabilising zinc ion is depicted as a white circle. The NS3-4A complex is shown in a modelled POPC (1-palmitoyl-2-oleoyl-3-sn-glycero-3-phospholcholine) bilayer (bottom). Bilayer headgroups are shown in yellow with the tails shown in light grey (Figure adapted from[13, 22]).

a C-terminal portion (aa ∼91-261)[22, 72–74]. NS4B is involved in HCV RNA replication and suggested to be the main inducer of membranous web formation[75–78]. AH2, as well as post-translational modifications in the C-terminal domain (i.e. palmitoylation) were identified as major determinants for NS4B oligomerisation, which is suggested to form a scaffold for the membranous web[74, 79, 80]. The membranous web is a membrane-associated multiprotein complex, i.e. the sum of vesicular membrane alterations harbouring viral replication sites, including all non-structural proteins which

are associated with ER-derived membranes[79, 81, 82]. An NMR structure of an amphipathic α-helix extending from amino acid 42 to 66 was solved using solution state NMR (PDB entry: 2JXF)[78], but no structure of the full length protein is available to date[22].

I.1.4.6 NS5A & NS5B

NS5A is a 447 amino acid phosphoprotein, with important roles in HCV RNA replication and particle assembly[13, 22, 83, 84]. It consists of an N-terminal amphipathic α-helix (aa 1-27) and three domains (D1 - D3; aa 28-213, 250-342 and 356-447, respectively), separated by two low-complexity sequences (LCS1 and LCS2)[85]. The N-terminal helix acts as a membrane anchor[86], whose structure was solved by solution state NMR (PDB entry: 1R7E)[87]. It suggests that the helix is embedded in-plane in the cytosolic leaflet of the ER membrane, burying a hydrophobic side in the membrane and presenting a polar side to the cytosol[87]. The anchor enables NS5A to interact with phospholipid monolayers, i.e. LDs or the LD-ER interface [22]. D1 and D2 play a role in HCV RNA replication and D1 also in LD binding[88–90], with D3 being implicated in core interaction[89, 91]. A crystal structure of D1 (PDB entry: 1ZH1)[92] reveals a dimeric structure with a groove facing away from the membrane that is suggested to accommodate either single- or double-stranded RNA[92]. It was suggested that NS5A binds zinc via a fully conserved quadruple cysteine motif[85], which was also seen in the D1 crystal structure, revealing a novel zinc binding motif at Cys39, Cys57, Cys59 and Cys80, respectively[22, 92]. A different crystal structure of the same domain was solved, revealing virtually the same monomer structure but a different dimer configuration, suggesting possibly different functions in RNA binding[13, 93]. HCV NS5A can be found in a basal phosphorylated (p56; 56 kDa) and hyperphosphorylated (p58; 58 kDa) form, with phosphorylation being mediated by several cellular kinases[84, 94–96]. It is suggested that the two different phosphorylation states of NS5A are modulating HCV RNA replication via regulation of replication and assembly host factors[13, 97, 98].

NS5B is a 591 amino acid RdRp (RNA-dependent RNA polymerase) of 65 kDa and the key enzyme in HCV RNA synthesis[13, 22, 99, 100]. It contains a C-terminal anchor (aa 570-591), a large N-terminal domain and a 40 amino acid linker in between[22], with the C-terminal anchor being disposable for enzyme function *in vitro*[79]. Several crystal structures of NS5B have been solved (PDB entries 1O55, 1CSJ, 1C2P and

1QUV)[93, 101–103]. All show the catalytic domain to be located in the N-terminal domain[100] and all share RdRp characteristic right hand structural organisation divided in fingers, palm, and thumb subdomains[13, 22, 79, 99, 100, 104].

I.1.5 The HCV life cycle

The HCV life cycle is a complex process and is currently not fully understood. The current accepted consensus is that it comprises HCV particles entering the cell[37], translation and polyprotein processing[105], viral RNA replication[79], virus particle assembly and release[106].

I.1.5.1 The HCV particle

The HCV particle is 40 - 100 nm in diameter. It consists of envelope glycoprotein E1 and E2 heterodimers that are embedded in a lipid bilayer. This surrounds a nucleocapsid composed of core protein and the single-stranded RNA genome[43, 107]. Patient sera and supernatant of cultured virus show that HCV RNA is found in low and high density fractions, i.e. at around 1.06 g/ml and 1.25 g/ml, respectively[108, 109]. HCV found in low density fractions shows a greater specific infectivity than HCV in high density fractions[110–112]. Virus derived from the infectious cell-culture system (HCVcc), which uses a genotype 2a genome (isolate JFH-1; Japanese fulminant hepatitis-1) and hepatoma Huh7-based cell lines[111], shows higher densities than sera-derived virus (1.15 - 1.20 to 1.03 - 1.16 g/ml, respectively)[112, 113]. This can be explained by the impaired VLDL (very-low-density lipoproteins) assembly pathways in Huh7- cell lines[114]. HCV particles are referred to as LVPs (lipoviroparticles), which was coined on the their appearance as large, spherical particles containing apoB (apolipoprotein B), HCV RNA and core[114, 115]. Viral particles also associate with other apolipoproteins, including apoE (apolipoprotein E)[116]. In general, HCV particles lack an icosahedral shape, which is caused by their association with host cell lipoproteins, i.e. LDL (low-density lipoprotein) and VLDL, such as apoB and apoE[14, 115, 116]. HCV particles also vary in their properties, depending on the host cell in which they are produced (reviewed in [18]), causing an innate heterogeneity in their appearance.

I.1.5.2 HCV cell entry

HCV entry and infection requires a minimal set of four host cell proteins, i.e. SRB1 (scavenger receptor class B member 1)[117], the tetraspanin CD81[118], claudin 1[119] and occludin[120], the latter two being tight junction proteins (Figure I.3). Other proteins and receptors have been identified which aid entry and infection (reviewed in [37, 121]), including the Niemann-Pick C1-like 1 cholesterol absorption receptor[122]. It is suggested that HCV spreads via cell-free and cell-cell routes *in vitro*[123], with the latter being proposed to be the possible predominant route *in vivo*[37]. After successful attachment, HCV employs a host cell mechanism and enters cells via clathrin-mediated endocytosis[124], followed by genome release into the cytoplasm where translation occurs[14].

I.1.5.3 Translation of the HCV polyprotein

HCV protein translation occurs in the rough ER in a cap-independent manner through the use of an IRES in the 5' UTR, which directly recruits the 40S ribosomal subunit to the viral RNA without the need of other cofactors[125] (FigureI.3). The translated HCV polyprotein is co- and post-translationally cleaved by cellular proteases (signal peptidase and signal peptide peptidase) and the viral NS2-NS3 and NS3-NS4A proteases to release ten HCV proteins[9, 12, 14, 22] (FigureI.1). The hepatic microRNA miR-122, which normally regulates genes in hepatic host cholesterol and lipid metabolism (reviewed in [126]) binds to the HCV 5' UTR[127] (Figure I.1 a) and has a stabilising and stimulatory effect on HCV genome translation[128, 129]. Furthermore, miR-122 protects the HCV genome from degradation by host exonucleases[130] and shields the genome from innate immune recognition[131].

I.1.5.4 HCV RNA replication

HCV viral RNA replication is catalysed by NS5B, an RNA-dependent RNA polymerase[79, 100] and takes place in ER-derived membrane vesicles, referred to as the membranous web (Figure I.3). It is composed of ER-derived single-, double- and multimembrane vesicles and LDs that assemble with HCV replication proteins as a complex, suggested to harbour the HCV replication sites[79, 81]. RNA replication is supported by the host

protein cyclophilin A, which is a petidyl prolyl isomerase[132]. Its function in the replication process is still unclear but it has been shown to interact with NS5A domains D2 and D3[133–135] and is suggested to support proper NS5A folding[13].

I.1.5.5 HCV assembly and release

HCV particle assembly is tightly linked with host cell lipid synthesis. It can be distinguished between an early nucleocapsid formation phase and a late envelopment and lipid incorporation phase[18, 106]. HCV particle assembly is located at LDs[21, 136]. LDs are lipid metabolism storage organelles that are composed of a monolayer of phospholipids harbouring a triacylglycerol and cholesterol ester core[18, 88, 108]. DGAT1 (diaglycerol O-acyltransferase) triggers a transfer of ER membrane bound core to the surface of LDs, with nucleocapsid formation involving core and NS5A[137]. NS5A associates to LDs via an amphipatic helix and interacts with core via D3 [22, 87]. Delivery of HCV genomes to the assembly site is suggested to be facilitated by the close proximity of RNA replication with assembly and budding sites[14, 18]. NS2 is crucial in HCV virion assembly by orchestrating protein-protein interactions to recruit other important components needed at the assembly site, i.e. E1/ E2[55], p7[138–141], NS3[71] and NS5A[53]. It was also shown that apoE interaction with NS5A is crucial for virus assembly[142]. It is suggested that HCV hijacks the VLDL pathway of its host cells to release mature HCV particles, which resemble TRLs (triglyceride-rich lipoproteins), which consist of apolipoproteins, a phospholipid monolayer and a neutral lipid core[18, 106, 143–145] (Figure I.3).

I.1.6 Antiviral strategies

I.1.6.1 Ribavirin and interferon

There is currently neither a prophylactic nor therapeutic HCV vaccine available (reviewed in[146]). In cases where HCV infection does not spontaneously clear, treatment with antiviral drugs is necessary. The previous standard treatment of care was a combination of pegylated interferon-α and ribavirin, with greatly varying response rates ranging between 36 % and 76 %, heavily depending on disease progression, the viral genotype and type of pegylated interferon used[147]. Also several host factors affect therapy outcome, including HLA (human leukocyte antigen) type, ethnicity, gender, age

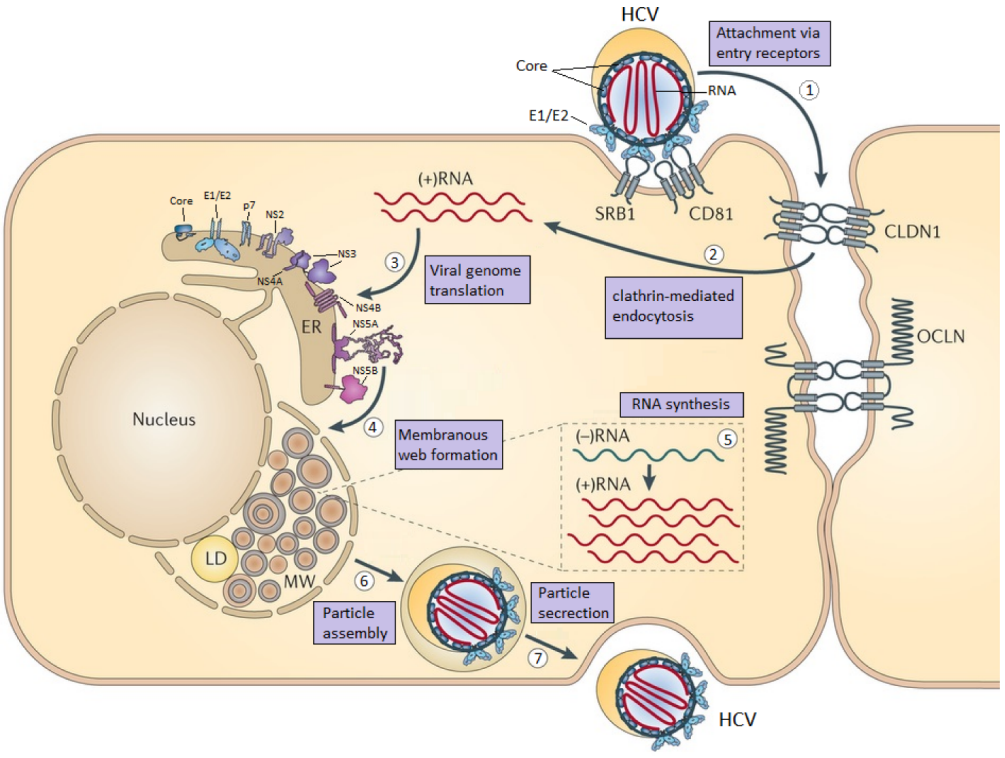

Figure I.3 – The HCV life cycle. (1) Binding of an HCV virus particle to SRB1 (scavenger receptor class B member 1) and CD81 triggers downstream interaction with the other two entry receptors CLDN1 (claudin 1) and OCLN (occludin), which leads to (2) uptake via clathrin-mediated endocytosis. (3) Following release of the HCV RNA genome into the cytoplasm and translation in the ER, the resulting HCV polyprotein is cleaved by viral and host proteases. (4) Viral and host cell proteins induce the membranous web (MW), consisting of lipid droplets (LD) and several types of membrane vesicles (single-, double- and multi-layered). (5) After RNA synthesis in the replication complexes (see insert) (6) assembly occurs at LD/ER membranes and viral particles bud into the ER-lumen (7). Mature HCV is suggested to be released via the VLDL secretory pathway (Figure adapted from[13]).

and obesity[14]. Recently, treatment success of interferon therapy has been shown to be dependent and predictable on a single-nucleotide polymorphism upstream of the IFNL3 (interferon, lambda 3) locus, which is linked to the innate immune response (reviewed in[148]). Genotype 1 patients only respond in about 45 % of cases successfully to interferon and ribavirin treatment[13, 149]. Interferons are cytokines that form part of the human cell-mediated immune response especially directed against but not acting directly on viruses[150]. Virally infected cells release interferon in order to alarm neighbouring cells, which then reduce protein synthesis to halt viral replication and spreading, creating an antiviral state within the cell and its neighbours[151, 152]. Ribavirin is a synthetic nucleoside analogue that interferes with viral RNA metabolism by resembling

purine RNA nucleotides. It is capable of incorporation into RNA, where it creates ambiguous basepairing, mimicking either guanosine or adenosine, successfully pairing with both cytidine and thymidine[153]. Ribavirin is suggested to act primarily by inducing mutations into the viral RNA genome, referred to as mutagenesis and error catastrophe or lethal mutagenesis, in which too many genome errors accumulate and start to be deleterious to the virus population[151, 153], which is supported and predicted by the quasispecies theory[11]. Its 5' mono-, di- and tri-phosphates, moreover, are inhibitors of viral RNA-dependent RNA polymerases. Other routes of action being discussed include direct inhibition of HCV replication, inosine-monophosphate-dehydrogenase inhibition and immunomodulation[151, 153] (reviewed in[149]).

I.1.6.2 Direct acting antivirals

The low efficacy against the common genotype 1, long treatment duration and undesirable side effects of interferon treatment make the development of interferon-free regimes desirable and necessary[13]. The goal of current HCV drug development is to establish an easy administrable, highly pan-genotype effective, interferon free therapy[149, 154]. The recently introduced HCV NS3/4A serine protease inhibitors, boceprevir[155] and teleprevir[156] (Figure I.4 b and c) are first generation, direct acting antivirals (DAAs) approved for treatment of genotype 1 patients. They increase sustained virological responses (SVRs) in combination therapy with pegylated interferon-α and ribavirin in previously unresponsive patients to up to 79 %[150], however, monotherapy with either compound leads to HCV escape mutants[157, 158], with implicated resistance residues of HCV NS3/4A being Val36, His57, Thr54, Val55, Asp81, Gln80, Ser139, Arg155, Ala156, Asp168 and Val170 (Figure I.4).

Besides different classes and second generation NS3/4A inhibitors, nearly all other HCV proteins are heavily implicated in drug development efforts, most prominently NS5A and NS5B, but also NS3[159], NS4B and p7[13, 154]. Second-generation NS3/4A inhibitors are required to have a high-resistance threshold, activity against HCV variants that are resistant to first generation compounds and activity against different HCV genotypes[154]. This definition can also be applied more broadly to inhibitors of other HCV proteins[160, 161]. There are three classes of NS3/NS4A protease inhibitors. i) Linear peptidomimetics that covalently but reversibly bind the enzyme active site, to which class both boceprevir and telaprevir belong, ii) linear noncovalent peptidomimetic,

and iii) macrocyclic noncovalent peptidomimetic inhibitors, both of which do not form covalent adducts (reviewed in[154]). There are two classes of NS5A inhibitors, which are non-symmetric or palindromic inhibitors. An example for a palindromic NS5A inhibitor is daclatasvir (BMS-790052), which counts as a first generation inhibitor[160] (Table I.1). NS5B inhibitors have being categorised in nucleoside/nucleotide and non-nucleoside inhibitors[154]. Sofosbuvir (GS-7977) is the most advanced of NS5B nucleotide inhibitors, currently in a phase III trial[149]. There are a multitude of other inhibitors in development and in clinical trials including host-targeted agents (HTAs) (Table I.1)[149].

Table I.1 – DAAs and HTAs in clinical development (adapted from[149]).

Compound	Target	Manufacturer	Trial Phase in 2013
Telaprevir	NS3/4A protease	Vertex & Janssen	Approved
Boceprevir	NS3/4A protease	Merck	Approved
Simeprevir	NS3/4A protease	Janssen	III
Faldaprevir	NS3/4A protease	Boehringer-Ingelheim	III
Danoprevir/r	NS3/4A protease	Roche/Genentech	II
Vaniprevir	NS3/4A protease	Merck	II
Narlaprevir/r	NS3/4A protease	Merck	II
Asunaprevir	NS3/4A protease	Bristol-Myers Squibb	II
Sovaprevir	NS3/4A protease	Achillion	II
GS-9256	NS3/4A protease	Gilead	II
GS-9451	NS3/4A protease	Gilead	II
ABT-450/r	NS3/4A protease	Abbott	II
MK-5172	NS3/4A protease	Merck	II
Sofosbuvir	NS5B RdRp	Gilead	III
Mericitabine	NS5B RdRp	Roche/Genentech	II
Tegobuvir	NS5B RdRp	Gilead	II
Filibuvir	NS5B RdRp	Pfizer	II
Setrobuvir	NS5B RdRp	Roche/Genentech	II
BI207127	NS5B RdRp	Boehringer-Ingelheim	II
ABT-333	NS5B RdRp	Abbott	II
ABT-072	NS5B RdRp	Abbott	II
VX-222	NS5B RdRp	Vertex	II
Daclatasvir	NS5A	Bristol-Myers Squibb	II
PPI-461	NS5A	Presidio	II
GS-5885	NS5A	Gilead	II
BMS-824293	NS5A	Bristol-Myers Squibb	II
ACH-2928	NS5A	Achillion	II
SCY-465 (HTA)	Cyclophilin	Scynexis	II
Miravirsen (HTA)	miR-122	Santaris	Ib

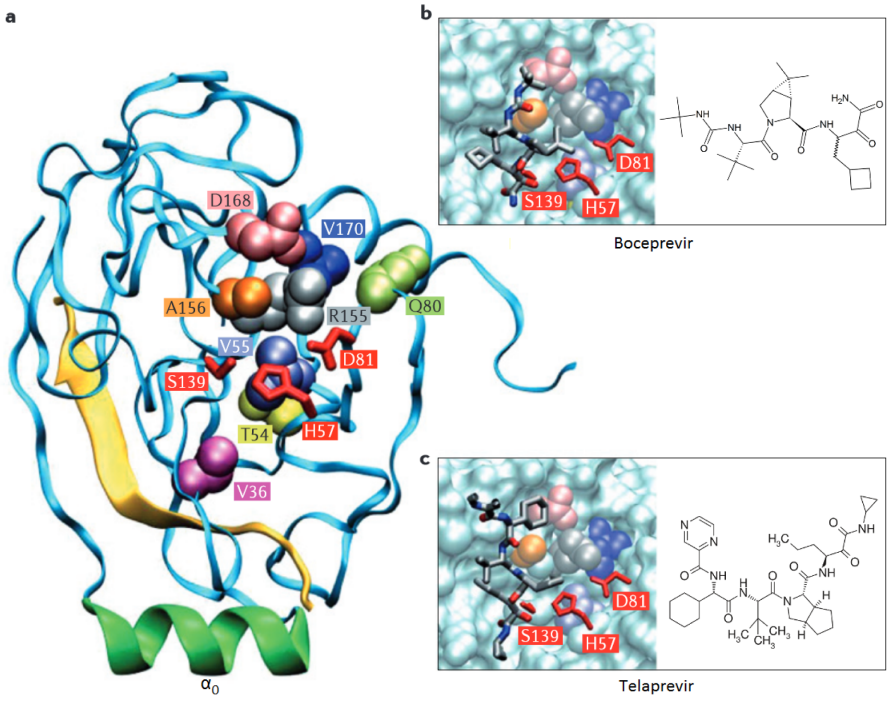

Figure I.4 – The NS3 protease/NS4A activation domain and binding of boceprevir and telaprevir. a) A ribbon diagram of the NS3 protease domain (PDB entry: 1CU1) with the central NS4A activation domain (yellow) is shown[60]. The NS3 catalytic triad (His57, Asp81 and Ser139) is indicated as red stick representations. Residues giving rise to resistance mutations against NS3-4A inhibitors (i.e. boceprevir and telaprevir) are highlighted and represented as coloured van der Waals spheres (Val36, His57, Thr54, Val55, Asp81, Gln80, Ser139, Arg155, Ala156, Asp168 and Val170). b & c) Boceprevir (top) and telaprevir (bottom) are shown in complex to the NS3 active site (surface representation) and their chemical structures are depicted in the inserts, respectively. The residues of the catalytic triad are highlighted in red and residues Arg155, Ala156, Asp168 and Val170 are shown as van der Waals spheres (PDB entries 2OC8 and 3SV6, respectively) (Figures adapted from[13]).

I.1.7 Viroporins

I.1.7.1 Viroporin topology and function

Viroporins are small, hydrophobic proteins able to oligomerise in membranes to form ion-conducting pores. They are involved in viral genome replication, assembly, as well as virus particle entry into and release from infected cells (reviewed in[162–164]). Examples of viroporins include influenza M2 (matrix protein 2)[165–167], HIV-1 Vpu (human immunodeficiency virus-1 viral protein U)[168–170] and HCV p7. Nieva *et al.* proposed a classification system based on viroporin structural characteristics, i.e. classes IA, IB, IIA and IIB[163]. Class IA viroporins have a single transmembrane domain with lumenal

N-termini, examples include M2[171, 172] and Vpu[170, 173]. Class IB viroporins have a cytosolic N-terminus and Class IIB cytosolic N- and C-termini[163]. Class IIA viroporins are characterised by helix-turn-helix hairpin motifs that span the membrane and have ER lumenal N- and C-termini, with HCV p7 being an example[163] (Figure I.5).

Viroporin functions are numerous[162, 163]. Functions carried out at the cell plasma membrane include alteration of host cell homeostasis and membrane depolarisation, but also roles in viral particle assembly and release. Intracellular effects include alterations of Ca^{2+}-homeostasis at the ER, dissipation of H^{+} gradients in the Golgi membrane remodelling to house replication sites (Figure I.6) (reviewed in[163]).

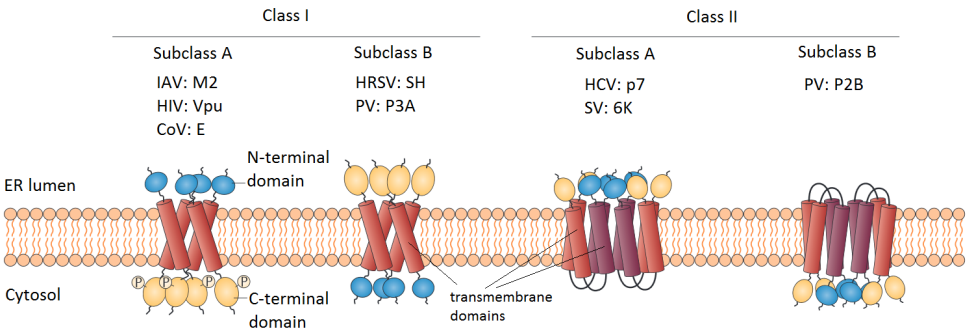

Figure I.5 – Viroporin classification based on membrane topology. Class I viroporins have a single transmembrane domain, with A and B subclasses having the lumenal and cytosolic termini on the opposite sites. Class A has a lumenal N-terminus, whereas class B has a cytosolic N-terminus and vice versa for the C-termini. In addition, class IA members are usually phosphorylated at the C-terminus. Examples for class IA viroporins include influenza A virus (IFA) M2, human immunodeficiency virus (HIV) Vpu, and coronavirus (Cov) E. Class IB members are human respiratory syncytial virus (HRSV) SH, and poliovirus (PV) P3A. Class II viroporins have two transmembrane domains which are characterised by helix-turn-helix hairpin motifs. Class IIA members have lumenal N- and C-termini, whereas members of subclass B have cytosolic N- and C-termini. Examples of Class IIA are HCV p7 and Sindbis virus (SV) 6K. PV P2B is an example for class IIB (Figure adapted from[163]).

I.1.7.2 Influenza M2

Influenza is an enveloped virus and has a negative -sense, single-stranded RNA genome ((-) ssRNA), which is made up of eight segments. Influenza belongs to the *Orthomyxoviridae* family and there are three serotypes (A, B and C). Influenza A viruses are further characterised by the subtype of their surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). The influenza A particle shows a round to oval shape and is approximately 90 nm in diameter[174]. The viral membrane contains M2. M2 is a pH-regulated, highly selective proton channel, which is encoded on the seventh RNA segment

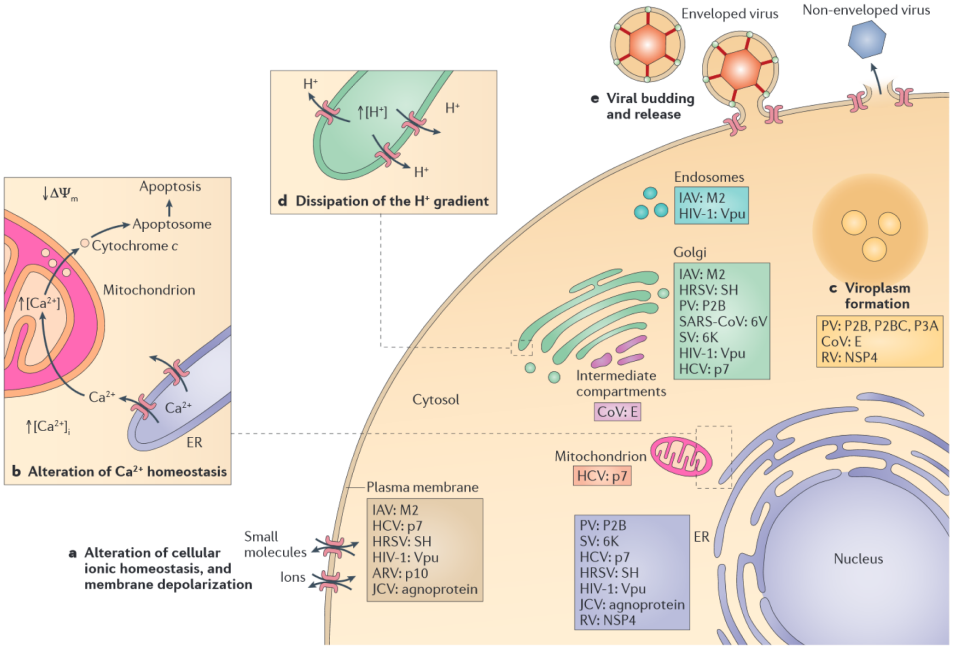

Figure I.6 – Viroporin functions in host cells during the viral life cycle. Cellular organelles that are targeted by viroporins are depicted, including the nucleus and surrounding ER in blue, mitochondrion in magenta, intermediate compartments in purple and the Golgi apparatus in green. a) At the plasma membrane, several viroporins have been shown to alter homeostasis of the cell and to cause membrane depolarisation, i.e. IAV (influenza A virus) M2, HCV p7, HRSV (human respiratory syncytial virus) SH, HIV-1 Vpu, ARV (avian reovirus) p10 and JCV (JC polyomavirus) agnoprotein. b) In the ER, viroporins cause leakage of Ca^{2+}, thereby altering Ca^{2+} homeostasis that can decrease the inner mitochondrial membrane potential ($\downarrow\Delta\Psi$ m) and lead to cytochrome C release, which in turn can trigger apoptosis. Implicated viroporins are listed in the blue box, such as RV NSP4 (rotavirus non-structural protein 4) and SV 6K (sindbis virus 6K). c) Other viroporins induce changes in intracellular membranes to create an environment suitable for viral replication (viroplasm). In case of HCV, this is triggerd by NS4B, which is not a viroporin (Figure I.1). d) Proton gradient dissipation in the Golgi which leads to decreased pH in the Golgi and other acidic vesicular compartments aiding inhibiting glycoprotein trafficking, is induced by IAV M2 and also suggested for HCV p7 and SARS-CoV 6V (severe acute respiratory syndrome corona virus 6V). e) Several viroporins are suggested to play a role in viral assembly and budding (Figure adapted from[163]).

together with the matrix protein 1 (M1) and generated by alternative splicing[163, 174–176]. M2 is a 97 amino acid, class IA viroporin and consists of three domains, i.e. an N-terminal ectodomain (aa 1-23), a transmembrane domain (aa 24-42) and a cytosolic tail (aa 43-97), which contains an amphipathic helix (aa 47-62). The cytosolic tail harbours a phosphorylation site[177, 178] and is needed for infectious virus production, genome packing[179] and interaction with M1[180]. The transmembrane domain is responsible for forming an M2 homotetramer and channel formation[181] and also harbours interaction sites for inhibitors such as amantadine and rimantadine[176]. Activation of M2 channel activity occurs via acidification of the endosome, which leads to opening of

M2 channels, allowing protons to pass from the endosome to the virus interior. This leads to release of viral RNA into the host cell cytoplasm enabling replication of the viral genomes[182]. It also avoids acidification of the trans-Golgi to enable adequate maturation of HA proteins, and it is implicated in membrane bussing and scission by generating negative membrane curvature[183–185]. Several structures have been solved employing various techniques, including solution[171] and solid state NMR (ssNMR)[186] and x-ray crystallography[172] (reviewed in[187–189]). All structures of the M2 transmembrane domains show a tetrameric pore, with side chains of highly conserved residues, His37 and Trp41 facing the pore lumen. The histidine can be protonated and is involved in proton flow, whereas the tryptophan can block the channel at neutral pH[163]. Several structures obtained at different pH suggest that the aromatic residues function as a gate, being open at low and closed at neutral pH[171, 172, 190]. The interaction site with adamantane inhibitors has also been identified and mapped to a cluster around Ser31[171, 172, 191]. It is unclear which of the first solved structures is most similar to the native M2 structure. The solution state NMR structure harbours a C-terminal part of M2 (aa 47-62), which the other two structures are lacking. The amphipathic segment (aa 47-62) together with the transmembrane domain is referred to as the conductance domain. An MD analysis of all three initial structures suggested that the X-ray structure is the most stable during MD simulations, mainly in its tri-protonated form, with the solution NMR structure being the least stable, which however can be partly attributed to the added C-terminal region[189]. Newly solved structures show a greater agreement for the backbone of the TM region and a similar position of the amphipathic helix with near consensus achieved in the closed state[192] (Figure I.7). Recent efforts are put into solving structures of the N-terminal ectodomain and the C-terminal cytoplasmic tail[187].

I.1.7.3 HIV Vpu

Human immunodeficiency virus (HIV) is a Lentivirus of the family of *Retroviridae*. There are two HIV viruses characterised, HIV-1 and HIV-2. HIV-1 has a single-stranded, positive sense RNA genome ((+) ssRNA), which encodes nine proteins. Vpu (viral protein U) is translated from a bicistronic RNA and only found in HIV-1[195]. Vpu is a class IA viroporin (aa 77-86), with an N-terminal domain facing the ER lumen[163]. It further consists of a single transmembrane and a cytosolic domain made up of two short amphiphatic

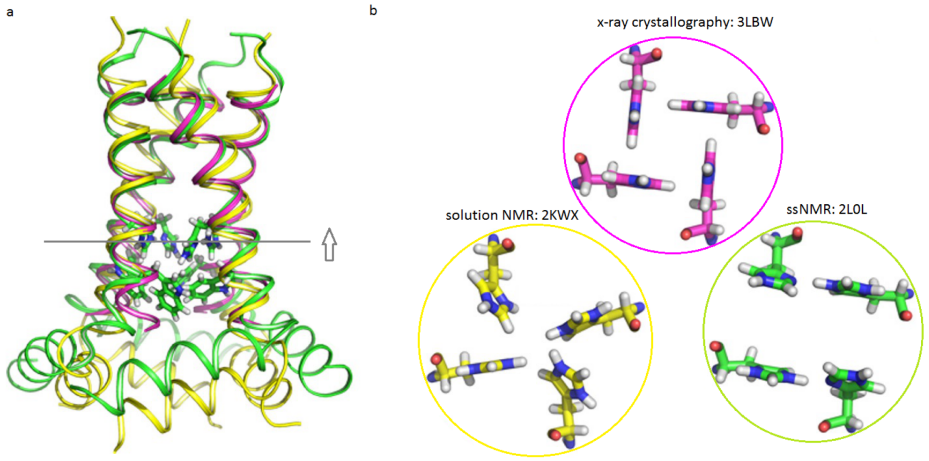

Figure I.7 – Influenza M2 structures solved by different techniques and in different conditions. a) An overlay of three recent M2 structures is shown with the backbone as a helical ribbon and the His37 and Trp41 sidechains as stick representations. The yellow structure (PDB entry: 2KWX) depicts the conductance domain and was obtained by solution NMR spectroscopy (aa 18-60; TM plus C-terminal part) in detergent micelles at pH 7.5[193]. The green structure (PDB entry: 2L0J) was obtained by ssNMR (solid state NMR spectroscopy) on the M2 conductance domain (aa 22-62) in liquid crystalline lipid bilayers at pH 7.5[194]. The pink structure (PDB entry: 3LBW) was obtained by X-ray crystallography on the TM domain (aa 22-46) in an octylglucoside environment at pH 6.5[190]. The superimposed TM helices of the ssNMR and X-ray structures are very similar, with a slight deviation at the C-terminus. The tilt of the solution NMR TM helix is significantly less than in the other two structures. Relative to the ssNMR structure, the amphipathic helices in the solution NMR structure have a different location with regard to both depth in the 'membrane' environment and lateral position. The His37 and Trp41 sidechains are shown for 2L0J. b) shows a comparison of the positioning of the histidine sidechains in the three M2 structures displayed in the overlay as viewed from the amphipathic helix side (C-terminal end). The view direction is indicated by a grey arrow in a) as is the height in M2 by a grey line. The His37 sidechain torsion angles for 2KWX are similar to those for 2L0J (green circle), but the imidazole-imidazolium hydrogen bonds are not formed. As such the His37 sidechain conformations in 2KWX appear to be unstable due to charge repulsion (yellow circle). The His37 sidechain conformations in 3LBW provide an alternative, but more limited mechanism through cation-π interactions for charge dispersion and for structural stabilisation (pink circle) (Figure adapted from[192]).

α-helices (aa 33-49 and 57-70) connected by a short loop[195, 196]. Vpu has two primary functions, i.e. CD4 receptor downregulation and tetherin antagonisation[195, 197]. Vpu is suggested to form an oligomeric channel via its transmembrane domain[198, 199]. Although structures of the TM helix have been solved by solution and ssNMR[173, 200], no full length structure besides a molecular model is available to date[201].

I.1.8 Viroporin inhibitors

Viroporins are important for the viral life cycle. As it is possible to screen for inhibitory compounds that block their membrane-permeabilising ability, viroporins are attractive targets for antiviral drug development[163], with a number of compounds in development to prevent channel activity or formation (Table I.2).

I.1.8.1 Adamantane compounds

The first drug studied that inhibited a viroporin was amantadine. It was used to block the activity of influenza M2[202]. Early reports show the rapid emergence of viral escape mutations[202], also against the very similar drug rimantadine[203], highlighting that both compounds are unsuitable for influenza monotherapy (Figure I.8). Amantadine was also suggested to be useful in treating HCV infections, as it was found to block p7 ion channel activity[204, 205], albeit with conflicting efficacies reported[141, 206]. Amantadine also showed no benefits in HCV combination therapy with interferon and ribavirin[207].

I.1.8.2 Acylguanidine compounds

HMA (5-(*N,N*-hexamethylene) amiloride) was shown to inhibit HIV Vpu[208] and p7[209]. BIT225 (*N*-[5-(1-methyl-1*H*-pyrazol-4-yl)-napthalene-2-carbonyl]-guanidine (Figure I.8) was identified as an inhibitor against p7 ion channel function[210] and has since been shown to inhibit Vpu[211], in a manner which is not based on interference with tethering antagonism[212]. It also shows an inhibitory effect on E protein of coronavirus[213]. There is also a report indicating that BIT225 inhibits the RdRp of coxsackievirus B3 (reviewed in[214]), which suggests that the antiviral mechanism of action of BIT225 is still not fully elucidated and might be multifactorial.

I.1.8.3 Long-alkyl chain iminosugars

*N*N-DNJ and *N*N-DGJ (*N*-nonyl-deoxynojirimycin and *N*-nonyl-deoxygalactojirimycin, respectively) are long alkyl iminosugar derivatives. Iminosugars are sugar analogues with the cyclic oxygen being replaced by nitrogen. Derivatives include the glucose analogue

Table I.2 – Viroporin inhibitors with respective test systems (Table adapted from[163]).

Viroporin(virus)	*Inhibitor(class)*	*Testsystem*
M2 (IAV)	Amantadine and rimantadine (adamantanes)	Artificial membranes, Cell culture[172, 191, 221]
Vpu (HIV-1)	HMA (acylguanidines)	Cell culture[208]
	BIT225 (acylguanidines)	Cell culture[211]
p7 (HCV)	Amantadine	Artificial membranes, Cell culture, Clinical studies[204–207]
	HMA	Artificial membranes, Cell culture[206, 209]
	*N*N-DNJ and *N*N-DGJ (long/alkyl chain iminosugars)	Artificial membranes, Cell culture, Clinical study[215, 222–224]
	BIT225	Artificial membrane, Clinical study[210]
E (SARS-CoV)	HMA	Cell culture[213]

HMA (5-(*N,N*-hexamethylene) amiloride), BIT225 (*N*-[5-(1-methyl-1*H*-pyrazol-4-yl)-napthalene-2-carbonyl]-guanidine, *N*N-DNJ (*N*-nonyl-deoxynojirimycin), *N*N-DGJ (and *N*-nonyl-deoxygalactonojirimycin), E SARS-CoV (E protein severe acute respiratory syndrome coronavirus)

deoxynojirimycin (DNJ) and the galactose analogue deoxygalactonojirimycin (DGJ), with alkyl chains of varying length added to their ring nitrogen (Figure I.8). *N*N-DGJ and *N*N-DNJ both inhibit HCV p7 *in vitro*[141, 215]. The glucose analogues (DNJs) also inhibit host ER-resident α-glucosidases, which prevents proper maturation of envelope glycoproteins, wielding antiviral activity against a wide range of viruses[216–220].

I.2 HCV p7

I.2.1 HCV p7 topology and function

HCV p7 is a small hydrophobic, endoplasmic reticulum (ER)-resident protein[225], encompassing 63 amino acids, which form two transmembrane α-helices[226] connected by a short basic loop, with the N- and C-termini facing towards the ER lumen[227]. It is assumed that HCV virions assemble at the ER membranes (i.e LD-ER membrane interface) and bud into the ER lumen, however, it is not yet clear if p7 is incorporated into the viral envelope[206, 228]. HCV p7 is critical for the assembly and secretion of

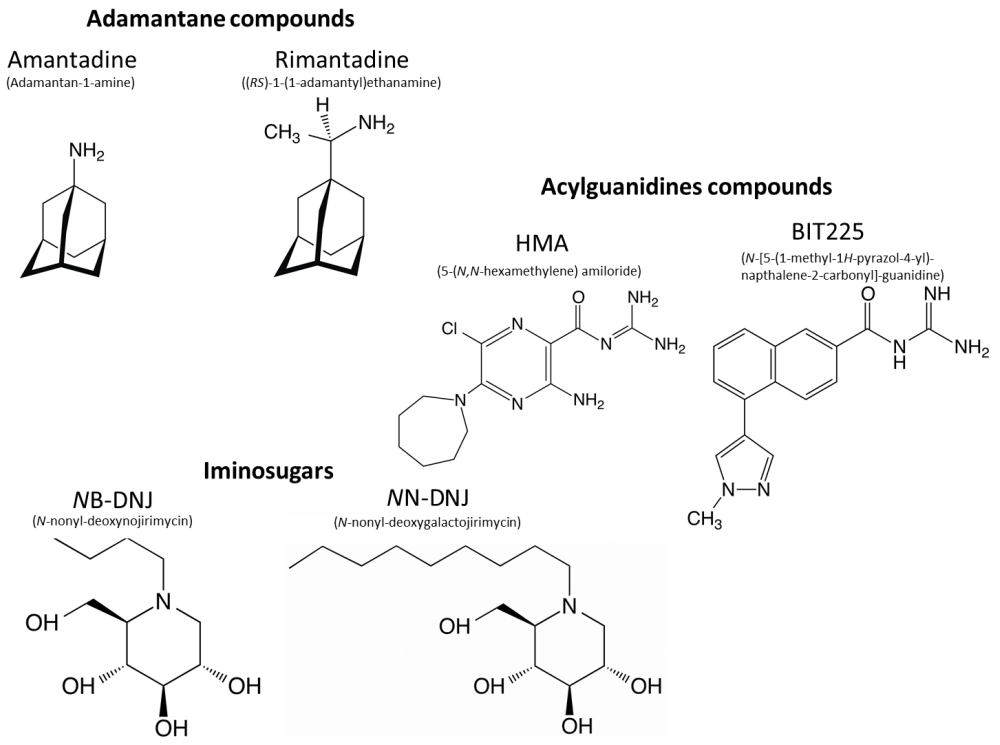

Figure I.8 – Inhibitors of HCV p7. The two adamantane compounds, amantadine and rimantadine are shown at the top. The middle row shows the acylguanidine compounds HMA and BIT225. The bottom row shows two glucose based iminosugars. The short-chained *N*B-DNJ, which shows no p7 inhibitory effects, and the long-alkyl chained *N*N-DGJ, which inhibits p7. The IUPAC names of each compound are listed in parentheses.

infectious viral particles *in vitro* and *in vivo*[139–141, 229–231], but it is neither crucial for viral genomic replication[232, 233] nor for entry[141, 231]. HCV p7 is a class IIA viroporin (Figure I.5), which are small, hydrophobic proteins able to oligomerise in membranes to form ion conducting pores (reviewed in[162, 163]).

I.2.2 HCV p7 structure

The structure of p7 has been characterised using most available structural methods, including EM (electron microscopy) and NMR. EM studies using GST-tagged HCV J4 p7 (gt 1b) show that p7 oligomers can exist as hexamers in artificial unilamellar vesicles of composition not described by the authors[204] and as heptamers in PA and PC (1:1, phosphatidic acid and phosphatidylcholine) liposomes[234]. In the absence of high-resolution structural information, an early modelling study used only secondary structure predictions and a computational hierarchical methodology to generate a hexameric p7 assembly[226]. Electrophysiological measurements using Cu^{2+} showed that

the divalent cation is chelated by the single histidine residue located in the N-terminal transmembrane domain of HCV strain H77 p7 (gt 1a) - thereby blocking the activity of the channel - indicating that the N-terminal helix is pore lining[235]. An EM study by Luik *et al.* determined the first three-dimensional structure of full-length p7, using single-particle electron microscopy, random conical tilting and negative staining[236]. The chemically synthesised, DH7PC (1,2-diheptanoyl-*sn*-glycero-3-phosphocholine) solubilised, HCV JFH-1 p7 (gt 2a) monomers assembled into a complex of 42 kDa. The structure with a resolution of 16 Å reveals a flower-shaped, hexameric protein architecture with protruding petals oriented toward the ER lumen, the latter being determined with p7 N- and C-terminal specific antibodies. A model was fitted into the EM volume and showed that 97.3 % of the atoms are located within the EM volume[236] (Figure I.9 a).

Several studies address the secondary and tertiary structure of monomeric p7. Using solution state NMR, Montserret *et al.* solved structures of TM1 and TM2 of HCV p7 J (C27S; gt 1b) in 50 % TFE (2,2,2-trifluoroethanol) and used molecular dynamic simulations in POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) bilayers to assemble a high-resolution structure of monomeric p7. Shown secondary structure elements include an N-terminal α-helix, an unstructured cytosolic loop and a C-terminal helical region[237] (Figure I.10). In a later study, using this monomeric p7 model, oligomeric complexes with four to seven subunits were simulated in hydrated POPC bilayers, which match the thickness of ER membranes. HCV p7 oligomers are shown to have a bias towards six or seven subunits in the assembled complexes, with tetrameric and pentameric stages suggested as intermediates[238]. These results support unpublished data showing the coexistence of hexa- and heptameric HCV JFH-1 p7 (gt 2a) in PC:PE (phosphatidylcholine:phosphatidylethanolamine; 4:1) membranes (Zitzmann *et al.*, unpublished). The same study showed the adaptability, fluidity and ability of p7 to align itself from the flower shaped, bent EM structure in the shorter chain DH7PC detergent micelles[236] into the cylindrical, upright complex observed in longer chain POPC bilayers[238] (Figure I.9 b), which more closely resembles the membrane thickness encountered in the ER. A study employing solution state NMR showed a p7 secondary structure consisting of two transmembrane segments each containing two helical segments, separated by a loop. Using magnetically aligned 14-*O*-PC/6-*O*-PC (1,2-di-*O*-hexyl-*sn*-glycero-3-phosphocholine) phospholipid membranes it was possible to determine helix tilt angles

by solid state NMR[239]. The secondary structure boundaries are overall similar to the previously mentioned NMR study, except kinks in TM 1 and TM 2[237] (Figure I.10, bottom). The same group published the first full-length three-dimensional structure of monomeric p7 using (1,2-dihexanoyl-*sn*-glycero-3-phosphocholine) solubilised HCV J4 p7 (C27S; gt 1b) with a final backbone root square deviation of 2.18 Å, at a pH of 4.0 and 50 °C. They found a structure that consists of two transmembrane domains, which are further split into two dynamic and two stable helical segments[240] (Figure I.10). The secondary structure is in close agreement with previous secondary structural studies, besides details in helix boundaries, tilt angles and position of kinks[237, 239]. A recently published monomeric structure of FLAG-tagged HCV J4 p7 (C27S, gt 1b) in methanol shows similar secondary elements as previously reported, with a small helix in the loop region and a TM 1 which shows no kink in His17 (PDB entry: 3ZDO[241]) (Figure I.10, bottom).

OuYang *et al.* reported the first three-dimensional structure of oligomeric p7 using 200 mM Fos-Choline-12 (DPC) solubilised HCV EUH1480 p7 (T1G, C2A, A12S, C27T, C44S; gt 5a) with a final backbone mean square root of 0.79 Å[242] (Figure I.10). The structure was supported by Fos-choline-12 solubilised and negative stained p7 EM particles of the same genotype and with the same mutations, although using a low (i.e. 3 mM) detergent concentration[242]. The particles looked similar to the ones seen in the previously mentioned single particle reconstruction EM study[236]. The secondary structure observed varies greatly from previous reports[237, 240] and the oligomeric subunit assembly is more complex than previously found in oligomeric p7 simulations in DH7PC micelles[236] and in POPC bilayers[238]. The structure reported by OuYang *et al.* is characterised by three helical segments in which the loop region that in other studies clearly separates the transmembrane helices, becomes part of a helical segment (Figure I.10, H2, yellow). There are also no intra-monomer helical contacts but rather inter-monomer contacts between helices of different monomers[242]. The final NMR structure was fitted with an inverted orientation of the termini into the Luik *et al.* EM envelope. The C-terminus is membrane buried in the hydrophobic interface of the Fos-choline-12 micelle[242].

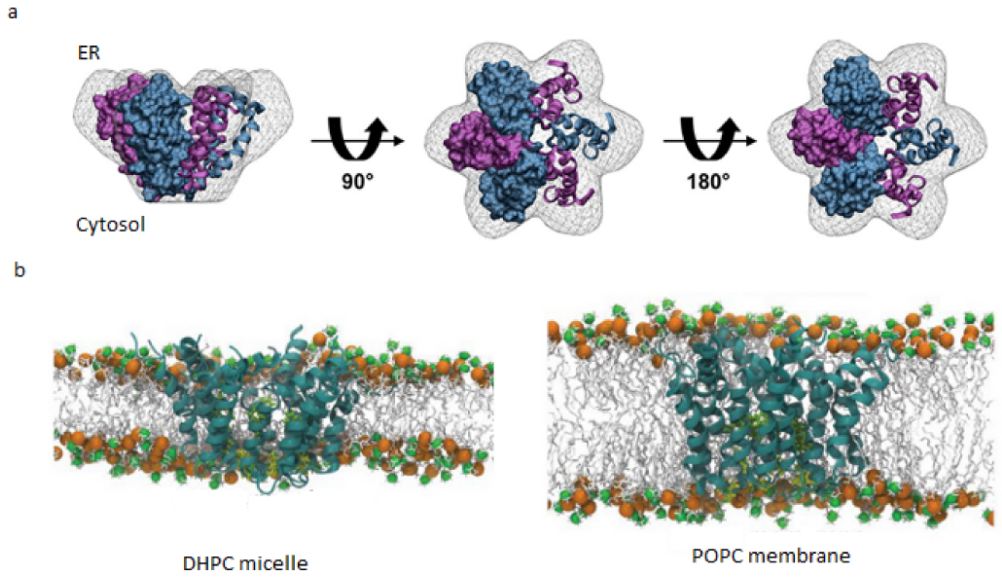

Figure I.9 – EM single particle reconstruction of HCV JFH-1 p7 (gt 2a). a) Fitting of simulated HCV JFH-1 p7 monomers into the hexameric EM volume is shown. Atomistic models of HCV JFH-1 p7 monomers were fitted into the EM density with their C- and N-termini oriented toward the petal tips. The monomers are depicted in alternating colours (blue and purple). In each model, three neighbouring monomers are surface represented; for the other three, the peptide backbones are shown. b) EM-fitted hexameric model derived from the solution NMR structure of HCV J p7 (C27S) in a DH7PC bilayer with a thickness of a DH7PC micelle, and POPC lipid bilayer are shown in side views (ribbon representation). The p7 monomers are depicted in green. The model DH7PC and POPC bilayer (labeled DH7PC micelle and POPC membrane, respectively) are described by means of a stick representation (light grey), with phosphorus and nitrogen atoms displayed as orange and green van der Waals spheres, respectively (Figures adapted from[236, 238]).

I.2.3 HCV p7 as a drug target

HCV p7 is a viroporin and forms cation-selective ion channels when reconstituted into planar lipid bilayers, which can be inhibited by several compounds (reviewed in[138]), including amantadine[204, 223, 237], rimantadine[206], HMA[209, 237], BIT225[210] and long-alkyl chain iminosugars (*N*N-DNJ and *N*N-DGJ)[215, 218, 223], whereas short-chain iminosugars (i.e. *N*B-DNJ and *N*B-DGJ) show no inhibitory effect against p7[223] (Figure I.11). HCV p7 was first shown to be able to conduct ions across membranes in BLM (black lipid membrane) experiments, either as a tagged fusion protein[204, 234, 243] or untagged[209, 215, 235, 244]. Liposome based assays also showed p7 ion flux[205, 237] and experiments with extracted intracellular membranes from cells expressing a FLAG-tagged p7 suggest that p7 is able to conduct protons[229, 245] (Figure I.11). Biotin-tagged HCV H77 p7 (gt 1a) ion channel function could be completely blocked *in vitro* using long-alky chain iminosugars at concentrations of 105-180 μM[243]. IC_{50} values,

	Montserret *et al.*	Cook *et al.*	Foster *et al.*	Ouyang *et al.*
p7 isolate	J	J4	J4	EUH1480
Genotype	1b	1b	1b	5a
Mutations	C27S	C27S	C27S	T1G, C2A, A12S, C27T, C44S
Technique	Solution NMR			
Detergent	50% TFE-d2	DHPC	Methanol	DPC (Fos12)
Temperature	25°C	50°C	25°C	30°C
pH	7.0	4.0	7.0	6.5
PDB entry	2K8J	n/a	3ZD0	2M6X

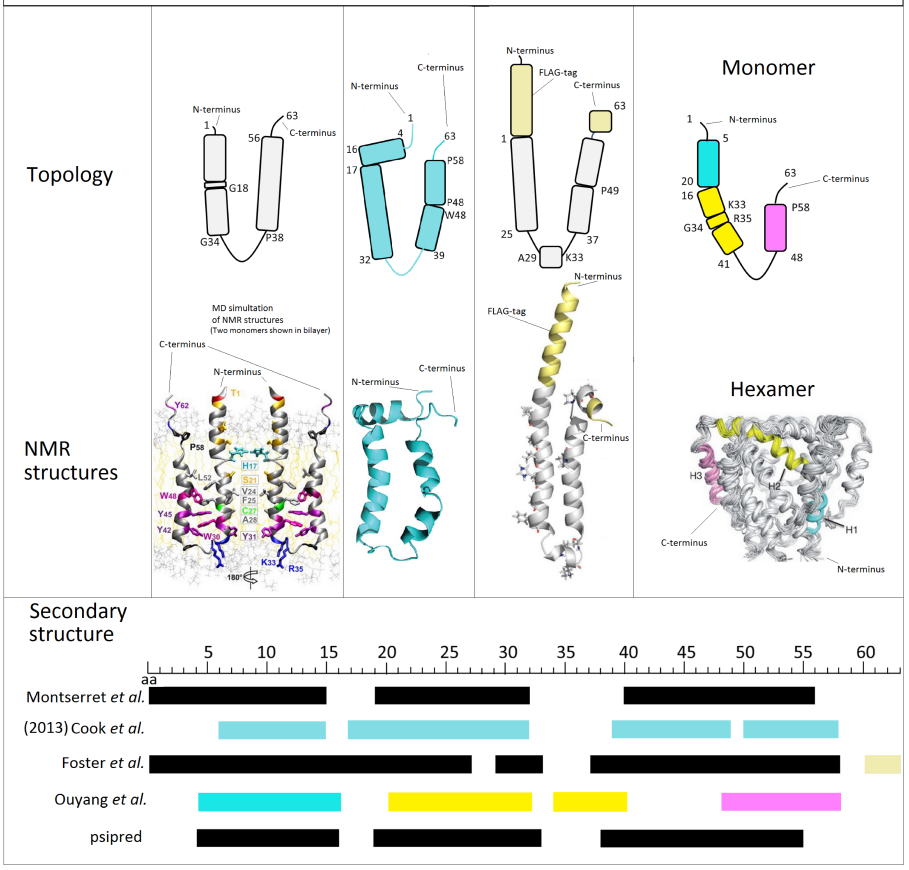

Figure I.10 – Comparison of solution NMR structures of HCV p7. The top table shows first author names of p7 NMR structural data publications (current in 2013)[237, 239–242]. The p7 isolate, genotype, construct mutations and sample conditions used, are listed. All structures were solved using solution NMR. The Montserret *et al.* structure is an MD simulation of three separate NMR structures of p7 peptides spanning full length p7[237]. The topology shows the similarities between the monomeric structures but also highlights the differences in some of the boundaries. See for example the helical segment in the Foster *et al.* loop region, or the C-terminal, small helix in the same structure[241]. The hexameric structure is upside down with regards to the other structures depicted[242]. A comparison of all secondary structure elements is shown in the bottom panel. Where feasible the colour code from the structures has been mirrored. For comparison, a HCV J4 p7 (C27S) psipred prediction is added at the bottom. The cytoplasmic loop between aa 32-38 in the monomeric structures becomes a structural feature in the hexameric structure of helix H2 (yellow).

which are the half maximal inhibitory concentration, of *N*N-DNJ and *N*N-DGJ on HCV H77 p7 (gt 1a) were 31 and 36 μM, respectively. Complete blockage of p7 was reached at concentrations ranging between 80-100 μM. Cell-culture experiments also showed a 60 % - 80 % reduction of HCV infected cells when treated with 50 μM *N*N-DGJ and *N*N-DNJ, respectively. Further, HCV RNA copy numbers dropped 50-fold when passaged in the presence of 5 μM *N*N-DGJ[223]. Limited genotype dependence could be shown using p7 genotypes 1-3, with a less efficient inhibition of HCV 452 p7 (gt 3a)[206]. This was later suggested to be caused by a single mutation in the HCV 452 p7 sequence, i.e. F25A[224], which is found in genotype 3 virus sequences only. HMA was found to be cytotoxic in cell-culture assays[206], whereas BIT225 is currently in phase II clinical trial against HCV and HIV co-infections.

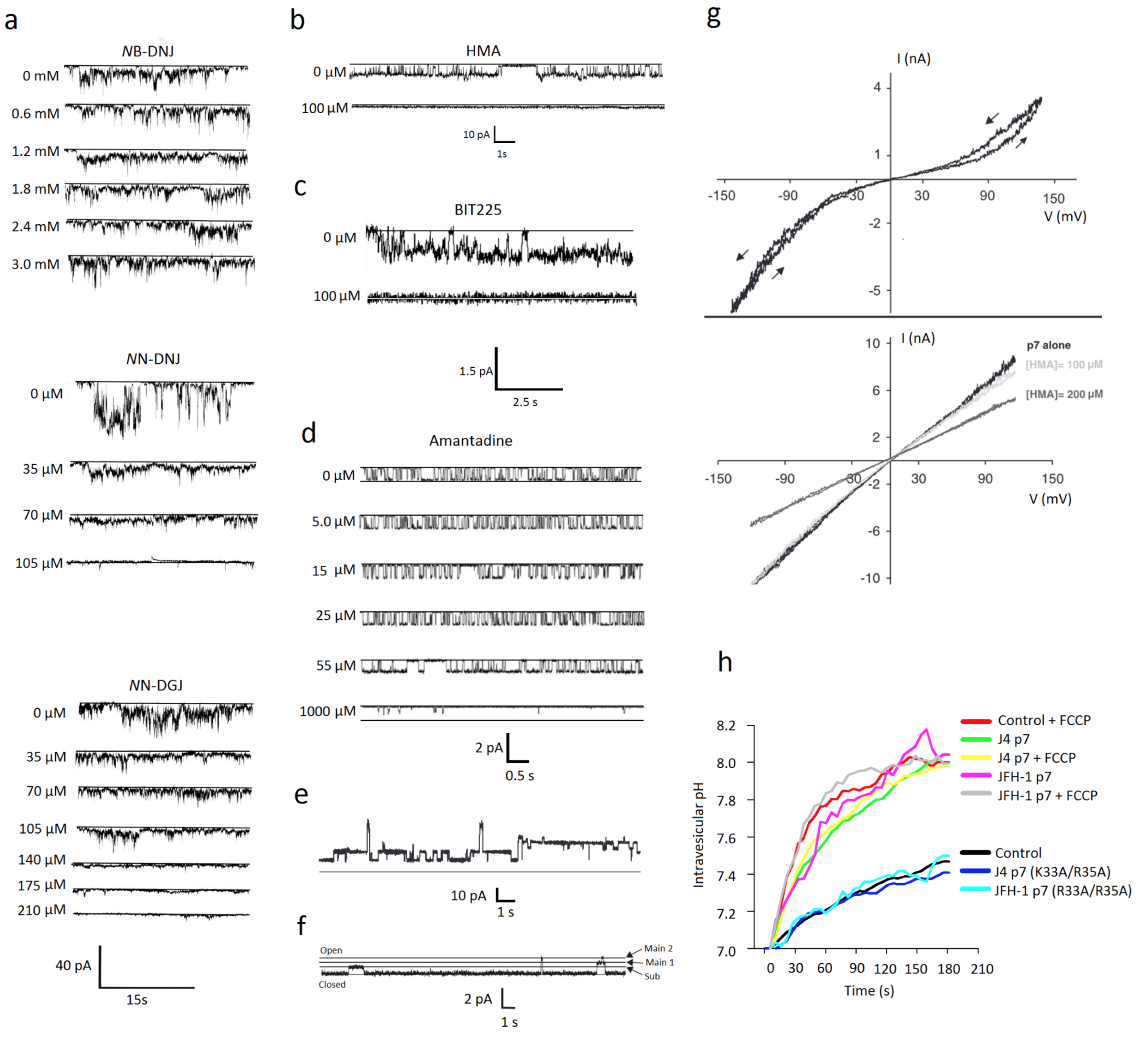

Figure I.11 (previous page) – Electrophysiology of HCV p7. a) Three BLM experiments of HCV H77 p7 (gt 1b) in POPE:POPC (4:1, w:w) membranes are shown. Short (*N*B-DNJ) and long-alkyl chain iminosugars (*N*N-DNJ and *N*N-DGJ) show different inhibitory capabilities. *N*B-DNJ shows no effect up to an inhibitor concentration of 3 mM (top). Both *N*N-DNJ and *N*N-DGJ show p7 inhibition at much lower concentrations, i.e. 105 and 210 μM, respectively. The closed state is shown as a solid, black line. Scale bars are depicted at the bottom. A positive voltage of 100 mV was used as the holding potential[215]. b) A BLM experiment of HCV H77 p7 (gt 1a) in a POPE:POPS:POPC (5:3:2) bilayer is shown with the addition of HMA leading to complete inhibition at 100 μM (bottom). The closed state is shown at a solid, black line. Scale bars are depicted below[209]. c) A BLM experiment of HCV H77 p7 (gt 1a) in POPE:POPS:POPC (6:3:1) bilayers is shown with BIT225. Complete block of conductance is reached at 100 μM BIT225[210]. d) Amantadine inhibition of HCV JFH-1 p7 (gt 2a) is shown in a BLM experiment using DOPC:POPS:POPE (4:1:1) bilayers. Full inhibition was reached at a concentration of 1 mM and is showed at the bottom, indicating that amantadine is not an efficient p7 inhibitor. The closed state is shown at a solid, black line. Scale bars are depicted at the bottom[223]. e) Single-channel recordings of HCV J p7 (gt 1b) in asolectin liposomes showing different conduction states are shown. Several conductions states can be distinguished, which are suggested to represent multiple copies of assembled p7 channels[237]. f) Shows a representative trace of HCV JFH-1 p7 (gt 2a) in PC:PE (1:4) membranes, highlighting two main and one sub-conductance state[246]. g) An I/V curve showing the current-voltage relationship of HCV J p7 (gt 1b) in asolectin liposomes is depicted. The arrows indicate the application direction of the applied voltage ramp. The I/V curve below shows HCV J p7 (gt 1a) after addition of HMA, showing a positive correlation of decreasing slope with increasing HMA concentration[237]. h) Shows an assay depicting the pH change in vesicle pellets isolated from HEK-293FT cells, which were transfected with either FLAG-tagged HCV J4 p7 (gt 1b), HCV JFH-1 p7 (gt 2a), the mutants stated or a control. All vesicles are loaded *ex-vivo* with a pH-dependent fluorophore. FCCP (carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone), a protonophore inducing proton leakage, is used as a further control. Experiments highlight that both FLAG-tagged HCV p7 isolates lead to proton leakage, suggesting that p7 is a proton channel. The two double mutants (FLAG HCV J4 p7 (K33A/R35A) and FLAG HCV JFH-1 (R33A/R35A) show no activity, both being at level with the control[245]. (POPE - 1-palmitoyl-2-oleoyl-3-*sn*-glycero-3-phospholethanolamine, POPC - 1-palmitoyl -2-oleoyl-3-*sn*-glycero-3-phospholcholine), POPS - 1-palmitoyl -2-oleoyl-3-*sn*-glycero-3-phosphoserine , DOPC - 1,2-dioleoyl-*sn*-glycero-3-phosphocholine) (Figures adapted from[209, 210, 223, 237, 245, 246]).

Amantadine inhibition of p7 is a field of on-going controversial discussion. His- tagged HCV J4 p7 (gt 1b) was shown to be completely blocked by 1 μM amantadine, whereas GST-His-p7 of the same genotype was not affected by concentrations up to 1 mM[204]. GST-FLAG HCV J4 p7 (gt 1b) could be inhibited in liposome release assays with concentrations as low as 10 μM, showing an IC_{50} of 2 μM[205]. The same study shows an even more potent inhibitory effect of rimantadine in the liposome assay[205], which was confirmed by cell culture studies[206]. Another study using HCV H77 p7 (gt 1a) in BLM experiments shows a bi-phasic inhibition curve for amantadine, with IC_{50} values of 69 and 253 μM, with near complete p7 blockage at physiologically irrelevant 1 mM. The same study observed a monophasic behaviour of long-alkyl chain iminosugars *N*N-DNJ and *N*N-DGJ[223]. HCVcc cell culture experiments also show that amantadine has no

effect on HCV H77 p7 (gt 1a), HCV J4 p7 (gt 1b) and HCV JFH-1 p7 (gt 2a) at concentrations of 50 μM[206, 223]. A study employing a liposome assay looking at residue determinants of HCV p7 ion channel function found that neither the two helices nor loop regions are determinants of amantadine and rimantadine sensitivity[247]. However, a more recent study suggested an adamantine binding site at L20 in a HCV JFH-1 p7 (gt 2a) context, showing adamantane resistance conferred by mutation L20F[224]. Mutating the strictly conserved dibasic motif (residues 33 and 35) in the p7 loop region shows that it is necessary for ion channel function of FLAG-tagged HCV J4 p7 (gt 1b)[247]. The same residues are also implicated in infectious virus production. Using HCV JFH-1 p7 (gt 2a) mutants K33Q and R35Q, a 100 and 1,000 fold decrease of total infectivity could be observed, respectively[141, 228]. The double alanine mutant R33A/R35A in JFH-1 (gt 2a) was also used in a cell culture assay to show that influenza M2 or J4 p7 (gt 1b) could restore infectious virus production, suggesting the proton channel function of p7 being solely responsible for it[229]. It was previously suggested that aberrant processing of E2-p7-NS2 could also be the cause for the effects observed linked to this mutation[141, 248]. M2 substitution is controversially discussed, with some studies reporting its ability to substitute for p7[229, 249], whereas others using the HCVcc system show the opposite[230]. The solution NMR structure of FLAG-tagged HCV J4 p7 (gt 1b) confirmed the previously reported adamantane binding site clustered around leucine 20 using rimantadine titration[241]. The NMR study by Opella and co-workers, looking at HCV J4 p7 (C27S, gt 1b) in DH6PC micelles, suggests that *N*N-DNJ and amantadine bind to the terminal ends of p7, whereas HMA binds preferentially to the loop region[240] (Figure I.12).

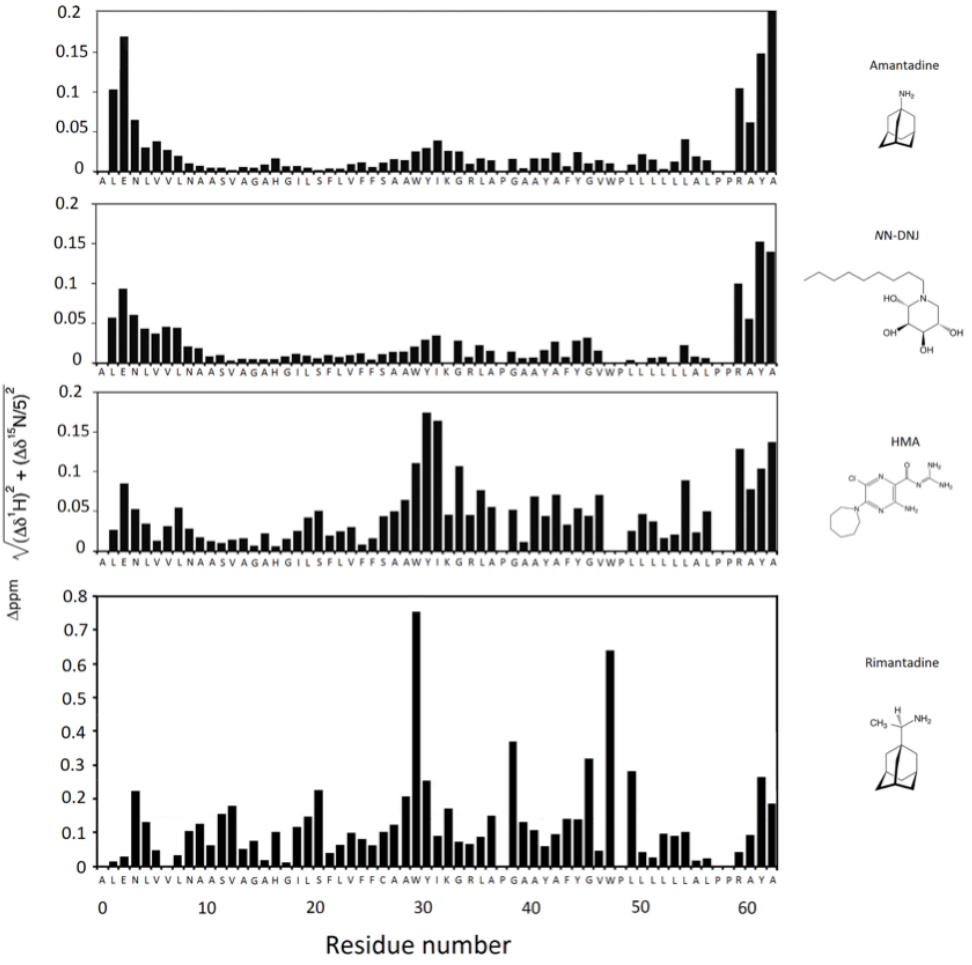

Figure I.12 – Chemical shift perturbations of HCV p7 inhibitors using solution NMR. All plots show chemical shift differences plotted as a function of residue number. Changes in chemical shifts are observed when different p7 inhibitors are added and interact with p7 and change the chemical environment of the interacting residue. The top graph shows the addition of 10 mM amantadine in H_2O to HCV J4 p7 (C27S, gt 1b) at a concentration of 0.5 mM. Perturbations are mainly visible in the N- and C-terminal part depicted by the large black bars. The graph below shows the addition of 10 mM *N*N-DNJ solubilised in DMSO (dimethyl sulfoxide) to 0.5 mM HCV J4 p7 (C27S, gt 1b). Again, most chemical shift perturbations are visible at the termini. HMA at a concentration of 10 mM in H_2O was added to HCV J4 p7 (C27S, gt 1b) in the experiment below. Perturbations are visible at the C-terminus but mainly in the loop region, the latter being defined by the monomeric HCV J4 p7 (C27S, gt 1b) structure in DH6PC micelles[240]. The graph at the bottom depicts FLAG-tagged p7 of the same genotype, lacking the C27S mutation. It shows a two molar excess addition of rimantadine at a p7 concentration between 0.3 and 0.6 mM. The biggest perturbations are visible at W30 and W48. Using a L20F mutant, interaction with rimantadine could be prevented, i.e. no shifts were observed (data not shown). Chemical shift differences were calculated using the formula depicted next to the y-axis (Figures adapted from[240, 241]).

Chapter II

Nuclear Magnetic Resonance

II.1 Basic Resonance Theory

NMR is a form of spectroscopy that can detect certain nuclei absorbing and reemitting electromagnetic radiation in an applied magnetic field (B_0). NMR spectra appear at low frequencies in the range of about but not exclusively 10 to 1000 MHz, which correspond to radiofrequency wavelengths of 30 m to 30 cm, respectively[250]. Nuclear spin is an intrinsic form of angular momentum in atomic nuclei, the latter being composed of neutrons and protons. The spin quantum number I of NMR-active nuclei is defined as I≠0. A non-zero spin is always associated with a non-zero magnetic moment (μ) via the relation, $\mu = \gamma$I, where γ is the gyromagnetic ratio of the nucleus, which is the ratio of the nucleus' magnetic dipole moment to its angular momentum. It is specific to each isotope and arises from the structure of the nucleus[251–253]. The non-zero magnetic moment allows the observation of NMR absorption spectra, caused by transitions between nuclear spin levels. When an external magnetic field, usually denoted as B_0, is applied to these nuclei, their magnetic dipoles align in spin states $I = 2I + 1$ relative to B_0 and they precess at the Larmor frequency ω, given by $\omega_0 = \gamma_0$ with γ being the gyromagnetic ratio of the respective nucleus and B_0 the applied magnetic field in Tesla. The gyromagnetic ratio dictates the sensitivity of the nuclei. A higher gyromagnetic ratio results in larger energy level differences for a given B_0 that is caused by larger population differences between ground and excited state (α and β). In protein NMR, the most studied nuclei are ^{1}H, ^{13}C and ^{15}N, which have spin of ½, which results in only two energy levels and the observed spectra are thus relatively simple[252–254] (Table II.1). The Boltzmann

distribution function explains that in B_0, spin systems exist at equilibrium of differing energy levels. The two possible states are the upper and ground state (Figure II.1 a). At equilibrium, there is excess in the ground state and a fraction of the spins will align with the external field B_0 along the z-axis. The spins give rise to the bulk magnetisation referred to as M_z (z magnetisation, longitudinal) (Figure II.1). The spins precess around the magnetic field with the Larmor frequency ($\omega = \gamma B_0$,), which is proportional to the magnetic field. This equilibrium can be perturbed by application of a radiofrequency field (B_1), which leads to a redistribution of populations and transitional magnetisation along the xy plane (M_{xy}, xy-magnetisation, transverse), which is followed by the return of the spins to their equilibrium state, referred to as relaxation (Figure II.1 c to b).

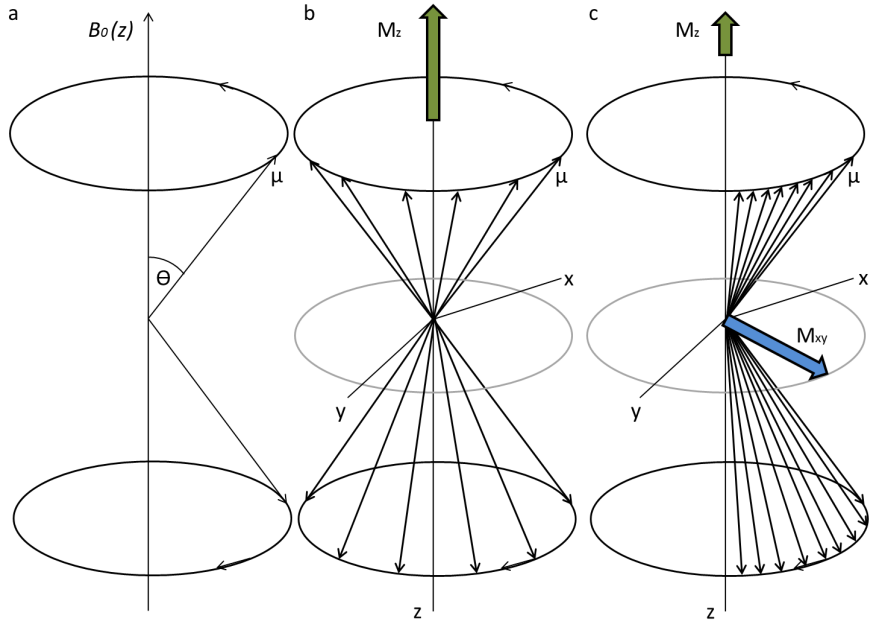

Figure II.1 – Representation of precessing magnetic moments μ. a) In a magnetic field B_0, aligned along z, the magnetic moment μ of a spin ½ system can take up two possible energy states, the low energy state aligned with B_0 and the high energy state aligned against it. Angular momentum and fixed Θ make μ precess around B_0. The population of the two cones along +z and -z is nearly equal, with the upper cone (i.e the low energy sate) containing slight excess. b) The magnetisation along B_0, is M_z, referred to as longitudinal magnetisiation. All xy spin components cancel leaving a net magnetisation along +z. c) Magnetisation is induced along xy (M_{xy}), by application of a radiofrequency field B_1, referred to as transverse magnetisation, which allows the passage of the low energy state to the high energy state (i.e. the absorption of energy). The return of the magnetisation back to M_z, is called relaxation by returning of the population to the low energy state (Figure adapted from[251, 252, 255, 256]).

A metal coil placed close to the sample can detect spins precessing in the xy-plane.

Table II.1 – Relevant nuclei characteristics in protein NMR (adapted from[250, 252, 253, 257]).

Nucleus	Spin	γ [MHz T^{-1}]	Frequency at 11.74 T [MHz]	Frequency at 22.3 T [MHz]	Natural abundance[%]
^{1}H	1/2	42.576	~500.0	~950.0	99.98
^{2}H	1	6.536	76.75	145.75	1.5 x 10^{-2}
^{13}C	1/2	10.705	125.72	238.72	1.1
^{14}N	1	3.077	36.12	68.62	99.63
^{15}N	1/2	-4.316	50.66	96.25	0.37

It is standard to refer to magnets in the protein NMR field according to their proton frequency at the given magnetic field, i.e. a 11.74 T magnet is a 500 MHz magnet, and a 22.3 T is a 950 MHz magnet.

The signal induced by the precessing spins decays over time and is called the free induction decay (FID). Since samples typically contain more than one nucleus having different resonance frequencies, the FID is typically Fourier transformed (FT) to generate a frequency-dependent spectrum[252]. There are two relaxation types that lead to signal decay: T_1- or longitudinal relaxation (spin-lattice) and T_2 or transverse relaxation (spin-spin), and their respective rates are referred to as R_1 and R_2. T_1 is the time constant that characterises the recovery of M_z magnetisation along B_0. Application of B_1, and T_2 describes the decay of M_{xy} back to equilibrium. Nuclei with a spin of 1, e.g. ^{14}N, are generally not used in protein NMR as their T_1 is too short, resulting in very broad or no visible peaks[257, 258]. Protein tumbling on the nanosecond timescale influences the rate of R_2 relaxation. Large molecules have high R_2 values, caused by their slow tumbling time. This results in signal broadening after FT and imposes a limit on the size of proteins that can be studied by solution NMR since for very large complexes such as membrane proteins in lipid bilayers the broadening results in signal intensities that fall under the detection threshold [250–253, 255, 256, 259].

Movement of electrons induced by the B_0 field causes a local magnetic field that opposes B_0. This shielding effect results in a shift in the nuclear resonance frequency. The amount of shielding depends on the chemistry and conformation of the molecule, and thus creates unique resonance frequencies referred to as chemical shifts. Methyl groups have well-shielded protons, whereas amide protons are relatively deshielded due to the presence of the electronegative nitrogen [250, 252, 254, 260, 261]. The chemical shift is measured by reference to the protons in a standard compound, i.e. DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), which has highly shielded methyl protons. The chemical

shifts of other nuclei are given by the difference in parts per million (ppm) to the reference frequency of the DSS proton: δ(ppm) = (γ - γ_{ref}/ γ_{ref}), where γ is given in Hz. This chemical shift scale has the advantage of being independent of B_0[250–253, 255, 256]. Other shielding effects observed in NMR are caused by interactions between nuclear spins that are mediated by the shared electrons in chemical bonds. Since spin ½ nuclei can exist in two orientations, differences in the effective B_0 fields of a nearby nucleus results in the signals for that nucleus being split in two. For each additional nearby spin ½ nucleus the nuclear signals are split again. This phenomenon is referred to as J-coupling, with a spin-spin coupling constant J measured in Hz (through-bond)[251–254]. Nuclear spins can also interact through space. Interaction between local dipoles in spatially close nuclei through-space is referred to as dipole-dipole interactions, which is the dominant cause of relaxation for spin ½ nuclei. If two nuclear-spin dipoles are spatially close (up to 5 Å), then application of a strong B_1 to one spin can cause intensity changes in the spectra of other nearby spins through cross-relaxation. This phenomenon is called the nuclear Overhauser effect (NOE) and is one of the most useful tools for the study of protein structure[250–253, 255, 256, 259].

II.1.1 Multidimensional NMR experiments

In protein NMR in general but especially in the study of membrane proteins with detergent and/or lipids present, the relatively large size of the molecules and the number of nuclei make one dimensional (1D) NMR spectra very crowded and overlapped. 1D experiments are valuable to look at signal to noise ratio of a sample and optimisation of sample conditions before moving on to multidimensional (2D and 3D) experiments. By correlating the protons with second or third spins and spreading the signals into multiple dimensions the overlaps can be mostly resolved. Furthermore, new information can be revealed by the correlations to other nuclei either in homonuclear (e.g. ^{1}H and ^{1}H) or heteronuclear (e.g. ^{1}H and ^{15}N) experiments[252, 257–259]. Most modern experiments apply H_2O suppression techniques, as H_2O protons are much more numerous than the protons of the protein sample. Experiments are usually carried out in H_2O buffers if exchangeable protons such as backbone amides are being observed[257, 258, 262].

II.1.2 Two-dimensional NMR experiments

II.1.2.1 Heteronuclear single quantum coherence

Heteronuclear single quantum coherence (HSQC) spectroscopy can be used to observe correlations between the chemical shifts of nitrogen (or carbon) atoms and covalently attached protons[250–253, 255, 256, 259, 263]. The ^{1}H,^{15}N-HSQC is particularly useful for looking at backbone amide groups in proteins[250, 252, 254, 256, 260, 261]. It is used as a 'fingerprint' experiment in protein NMR firstly because the sensitivity of the chemical shifts to chemistry, structure, and environment and secondly because all nonproline residues should, in theory, generate a crosspeak in a ^{1}H,^{15}N-HSQC spectrum (Figure II.2). It is also very useful for assessing the influence of detergent, temperature and/or pH conditions. In addition, since many NMR experiments are based on this experiment, without a high quality HSQC (or TROSY for larger proteins[264]), further NMR analysis is difficult.

II.1.2.2 Heteronuclear multiple quantum coherence

A heteronuclear multiple quantum coherence (HMQC) spectrum provides essentially the same spectral results as HSQC experiments[265, 266]. It employs fewer radio frequency pulses, which can also be implemented over a shorter duration, resulting in lower sensitivity losses due to pulse imperfections and relaxation. Selective Optimized-Flip-Angle Short-Transient-HMQC (SOFAST-HMQC) is an HMQC variant that enables very fast acquisition of 2D heteronuclear correlation spectra. It employs two (band-selective) ^{1}H pulses, ensuring minimal perturbation of the undetected proton spins[254, 260], thereby decreasing the inter-scan relaxation delay and decreasing experiment time[254]. In this way, analysis of low concentration protein samples is possible within a reasonable amount of time (Figure II.2).

II.1.2.3 Total correlation Spectroscopy

A 2D ^{1}H,^{1}H-TOCSY (Total correlation Spectroscopy) experiment correlates ^{1}H chemical shifts within one amino acid or part of one amino acid, i.e. correlations are seen between distant protons as long as there are couplings between every intervening proton.

Spin systems are isolated from one another by the peptide bond, which lacks a proton on its carbonyl group. TOCSY experiments are helpful in assignment of ^{1}H chemical shifts. It consists of a 90°pulse, a τ_1 period and a spin lock period. The latter transfers magnetisation throughout the ^{1}H spin system via ^{1}H,^{1}H J-coupling and its length (e.g. 90-110 ms) controls the numbers of magnetisation transfer steps[256, 267].

II.1.2.4 Nuclear Overhouser Spectroscopy

As discussed above, NOEs provide information on through space spin-spin interactions via dipole-dipole cross-relaxation[257]. In NOE spectroscopy (NOESY), crosspeaks reflecting the interaction are observed and the crosspeak intensities can be related to the distance between the spins[250, 252]. The homonuclear 2D ^{1}H,^{1}H-NOESY experiment consists of three sequential 90°pulses[257] (Figure II.3). The disruption in spin population close to another spin can be correlated to the chemical shift of the initially excited proton[250, 256]. The pulse sequence of a homonuclear 2D ^{1}H,^{1}H-NOESY experiment is shown in Figure II.3[257].

II.1.3 Three-dimensional NMR experiments

II.1.3.1 3D-NOESY-HSQC

The heteronuclear 3D ^{15}N-edited-NOESY HSQC, essentially employs a 2D ^{1}H,^{15}N-HSQC, in which the first 90°pulse is replaced by a 2D ^{1}H,^{1}H-NOESY sequence. It detects NOEs to the protein amide protons from any other proton in the molecule, providing the chemical shift of the two protons and the attached ^{15}N of the amide proton[250, 261]

II.1.3.2 3D-TOCSY HSQC

The heteronuclear 3D ^{15}N-edited-TOCSY HSQC employs a 2D ^{1}H,^{1}H-TOCSY followed by a 2D ^{1}H,^{15}N-HSQC. It resolves the TOCSY peaks along the ^{15}N plane and correlates each HSQC peak with the protons in the same spin system, i.e. its side chain protons[256].

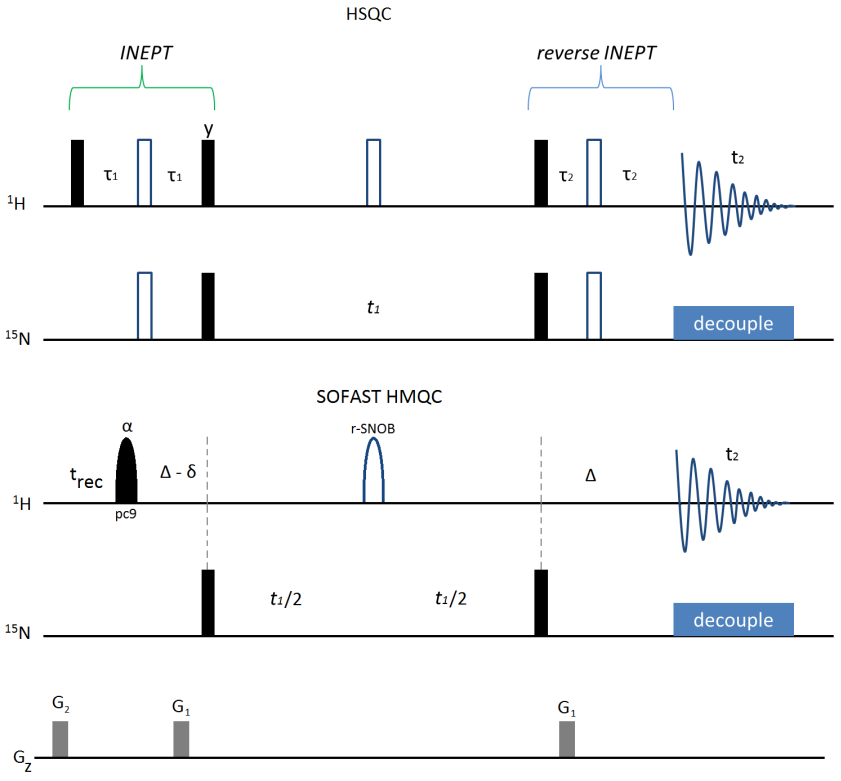

Figure II.2 – 2D ^{1}H,^{15}N-HSQC and SOFAST-HMQC pulse sequences. The top shows the pulse sequence for a decoupled ^{1}H,^{15}N-HSQC. Black and white rectangles represent 90°and 180°pulses, respectively. The HSQC transfers magnetisation from an amide ^{1}H to its attached ^{15}N and back for detection. It can be divided into four parts, the INEPT sequence (see below)[255], ^{15}N-chemical shift encoding, reverse INEPT, and ^{1}H-chemical shift acquisition. The INEPT (insensitive nucleus enhanced by polarisation transfer) sequence uses ^{15}N-^{1}H J-coupling to create ^{15}N-excitation, with ^{15}N being the insensitive nucleus due to its lower γ. The optimal value for τ_1 and τ is $1/(4J_{NH})$. The first 90°pulse transfers ^{1}H magnetisation to the -y axis, and during the two τ_1 periods magnetisation is transferred from ^{1}H to ^{1}N. The 180°pulses on both spins, prevents ^{1}H chemical shift evolution and the generation of multiple quantum magnetisation. The two τ_1 periods are followed by 90°pulses on both spins, resulting in transverse ^{15}N and longitudinal ^{1}H magnetisation. With ^{15}N magnetisation in the transverse plane, the τ_1 period is then used to encode the ^{15}N chemical shift. The 180°pulse during τ_1 refocuses the NH J-coupling resulting in a singlet peak in the ^{15}N-dimension. Magnetisation is then transferred back to ^{1}H during a reverse INEPT, and the ^{1}H chemical shift is recorded in the FID, during which a rapid sequence of inversion pulses ensures decoupling of the ^{1}H magnetisation from ^{15}N spins. The bottom shows a pulse sequence for a decoupled ^{1}H, ^{15}N-SOFAST-HMQC. The first ^{1}H excitation pulse is applied with a flip angle α and has a polychromatic PC9 shape. Band-selective ^{1}H refocusing is realised using an r-SNOB profile and the transfer delay Δ is set to $1/(2J_{NH})$. Delay δ accounts for spin evolution during the PC9 pulse, and t_{rec} is the recycle delay between scans (Figures adapted from[250, 254, 256, 260]).

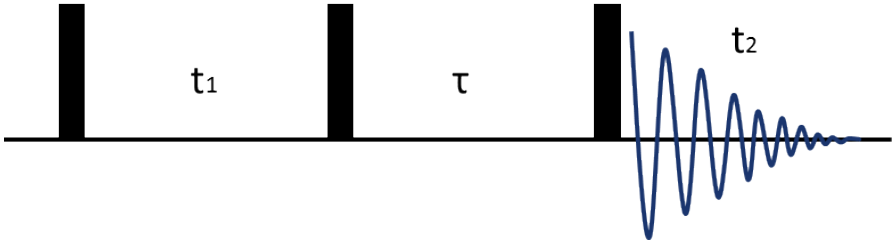

Figure II.3 – The 2D ^{1}H,^{1}H-NOESY pulse sequence. It consists of three sequential 90 °pulses (black rectangles). Magnetisation is transferred via the NOE. The first pulse excites protons and a delay (t_1) encodes their chemical shift, followed by a second pulse that transfers some of the magnetisation back to the z direction. During the mixing time τ, the altered populations of the first spin perturb nearby spins via cross-relaxation. They carry a label identifying their source, which is solely dependent on the chemical shift of the first proton. Magnetisation in -z can transfer the equilibrium disturbance to nearby, spatially close spins. The third pulse moves the magnetisation back to the xy plane for detection (t_2) (Figure adapted from[250, 256]).

II.1.4 Protein assignment

Resonance assignment is the determination of the chemical shift for particular nuclei within the protein, which can be done for relatively small proteins (<10 kDa) by a combination of 3D-HSQC NOESY and 3D-TOCSY NOESY[252, 253, 268, 269]. Such an approach to backbone assignment only requires ^{15}N-labelled protein[251, 252, 256, 259, 268, 270]. The ^{15}N-NOESY-HSQC provides crosspeaks to NH groups from all ^{1}H resonances, which are within about 5 Å. Critical to the assignment process is the ^{15}N-TOCSY-HSQC, which provides crosspeaks to NH groups from only intra-residue protons, therefore enabling unambiguous assignment of the intra- and inter-residue crosspeaks in the NOESY[253, 259, 271].

II.1.5 Protein dynamics

Proteins generally undergo motions, some of which are crucial for function. Flexible proteins accommodate rapid changes in their direct environment and binding of interaction partners[258, 259]. Some biological processes rely on information transduction, in which proteins play a role through conformational changes, such as ligand binding[252, 272]. Solution-state NMR is a unique technique that enables measurement of protein dynamics over a timescale of picoseconds to hours, spanning 12 orders of magnitude[252, 259, 272–275]. Protein dynamics on the fast timescale, which includes local motions such as side-chain bond rotations or movements within unstructured loops, can be quantified

in NMR by measuring nuclear relaxation rates. Commonly, the ^{15}N T_1 and T_2 relaxation times (described above), and the ^{1}H,^{15}N heteronuclear NOE (hetNOE) are measured[250, 252, 259], which can provide information on backbone flexibility for every nonproline residue. T_1 and T_2 are often reported as their reciprocal relaxation rates R_1 and R_2, the latter also being linked to the molecular size of the protein, and in case of membrane proteins, the whole complex including the detergent micelle or lipid bicelles[259, 263, 272]. In addition, the ratio of R_2 and R_1 and their product, can reveal indications for chemical exchange. R_2/R_1 identifies residues that undergo chemical exchange, whereas the product of R_1 and R_2 distinguishes further between motional anisotropy and chemical exchange[250, 252, 256, 276]. The ^{1}H,^{15}N hetNOE is similar to the NOE, but rather than using homonuclei the magnetisation transfer occurs from ^{1}H to ^{15}N. It can be used to determine flexible residues in a protein, which is based on the NOEs transfer rate being influenced by protein dynamics. Since the NOE depends both on distance and dynamics, and the distance between bonded amide ^{1}H and ^{15}N is essentially fixed, the hetNOE provides a sensitive measure of the motion of the amide group, i.e. the constant chemical bond length defines the intermolecular distance[264, 277]. ^{1}H,^{15}N hetNOEs provide information for NH-bond vectors in proteins, i.e. their motion. If they tumble faster than the whole protein, the hetNOE intensity decreases, which is typically observed at the N- and C-termini. Negative values can be observed for random coil-like behaviour[265, 266, 278], but for residues in folded domains, values up to 0.8 can be found and indicate inflexible structures[259] The ^{15}N T_1 and T_2 and hetNOE can be interpreted in terms of motional frequencies and amplitudes. The motional amplitude is given by the generalised order parameter S_2, based on model-free analysis of Lipari and Szabo[251, 253, 259, 279]. S_2 values are dimensionless and can be between 0 for freely moving residues and up to 1 for a completely rigid residue[259, 272]. Membrane buried residues are often found to have values close to the maximum value ($S_2 = 1$)[259, 262, 280]. The exchange between two conformations in a protein, or the exchange of binding of an interaction partner, can be either in fast-exchange, leading to one average signal, in slow exchange, giving rise to two separate signals, or broadening of the two in intermediate exchange[259]. In the slow exchange regime, bound and free form exchange slower than the differences in chemical shifts, which leads to two separate signals of the free and bound form. With the addition of increasing concentration of ligand, the free form signal becomes weaker until it disappears at saturation, while at the same time the bound form signal intensity increases. In the fast exchange regime,

bound and free form exchange faster than the difference in chemical shifts. This will lead to only one visible signal that changes position according to the ratio between free and bound form. The final position is reached once the protein has been completely saturated with the binding ligand[259, 262]. In between the fast and slow exchange regimes is the intermediate regime in which the exchange rate of bound and free state is comparable to the difference in chemical shifts, resulting in a more complex situation, consisting of a mixture of shifting signals as seen as in fast-exchange, and appearance/disappearance of signals as seen as in slow-exchange[259, 262].

Chapter III

Expression and purification of HCV p7

III.1 Expression and purification of HCV p7

In order to carry out biophysical analyses of HCV p7 using NMR, an expression system needed to be established. NMR experiments, such as resonance assignment and interaction studies, require milligram quantities of pure, isotopically labelled protein (i.e. ^{15}N, ^{13}C and ^{2}H). The latter can also be achieved using chemical synthesis, however, it is not feasible as the purchase of isotopically labelled amino acids is cost-prohibitive.

The plasmid pMM-LR6 (gift from S.C. Blacklow, Harvard Medical School, Boston; provided by J. Schnell, Department of Biochemistry, Oxford University), which carries a trpΔLE sequence and employs *E. coli (Escherichia coli)* as an expression host, is generally regarded to be useful for production of small membrane proteins[281] and has been successfully used previously to express other viroporins, such as influenza M2[171, 282] (Figure III.1). TrpΔLE induces inclusion body formation in *E. coli*, which reduces toxicity and increases resistance to proteolysis. It is based on a pET-17b (Kan$_r$) vector (Novagen) and carries a 9-His-tag. TrpΔLE consists of the leader sequence of the *E. coli trp* operon fused to a sequence of 97 residues found near the C-terminus of the anthranilate synthase gene[283, 284]. The gene of interest is inserted via HindIII (5') and BamHI (3') restriction sites using custom synthesised DNA fragments (GenScript or GeneArt). The fusion protein contains a C-terminal, unique methionine. The target

sequence has to either naturally lack any methionine or any occurring methionine has to be mutated to a different amino acid (i.e. leucines or isoleucine), as cleavage of the heterologous fusion protein is catalaysed by CNBr (cyanogen bromide), which cleaves specifically after methionine. Cleavage results in release of the N-terminal His_9-tagged trpΔLE, and the C-terminal p7, and depending on cleavage efficiency, uncut fusion protein. Purification is carried out via IMAC (immobilised metal affinity chromatography), followed by HPLC (high-performance liquid chromatography) or FPLC (fast protein liquid chromatography).

III.1.1 HCV p7 sequence analysis

There are seven HCV genotypes and hundreds of subtypes described[3]. An alignment of p7 sequences in figure III.2 shows similarities and conserved residues of the different isolates with at least one example sequence of each respective genotype. Several residues such as Gly18 and Trp30, are 100 % conserved, whereas conservation of other residues can be quite varied and different, and in the example for threonine at position 1 in HCV EUH1480 p7 (gt 5a) (UniProt: O39928) quite low (Figure III.2). For the chosen expression system, isolates containing no naturally occurring methionines in their sequence were favoured, and three isolates, HCV J4 p7 (UniProt: O92972), HCV 452 p7 (UniProt: Q1W5D5) and HCV SA13 p7 (UniProt: O91936), spanning different genotype families (gt 1b, 3a and 5a, respectively), were selected for small scale test expressions.

III.1.2 HCV p7 test expression using the pMM-LR6 vector

The three p7 isolates HCV J4, HCV 452 and HCV SA13 were ordered as synthetic genes (GeneSript) in a carrier plasmid (pUC57), containing the protein sequence plus the required restriction sites at their 5' and 3'-ends. *E. coli* codon optimisation was carried out according to the manufacturers' instructions. Following transformation in a suitable *E. coli* strain (e.g. DH5α) for plasmid multiplication and minipreparation (Qiagen), restriction digest was carried out using restriction endonucleases BamHI and HindIII (NEB). The same restriction digest was carried out with the plasmid pMM-LR6 (pMM1) (Figure III.3). Following agarose gel purification (Qiagen), the digested inserts and plasmid were ligated to give rise to three constructs, pMM1 J4 p7, pMM1 452 p7 and pMM1 SA13 p7. The sequences were verified by DNA-sequencing (Source

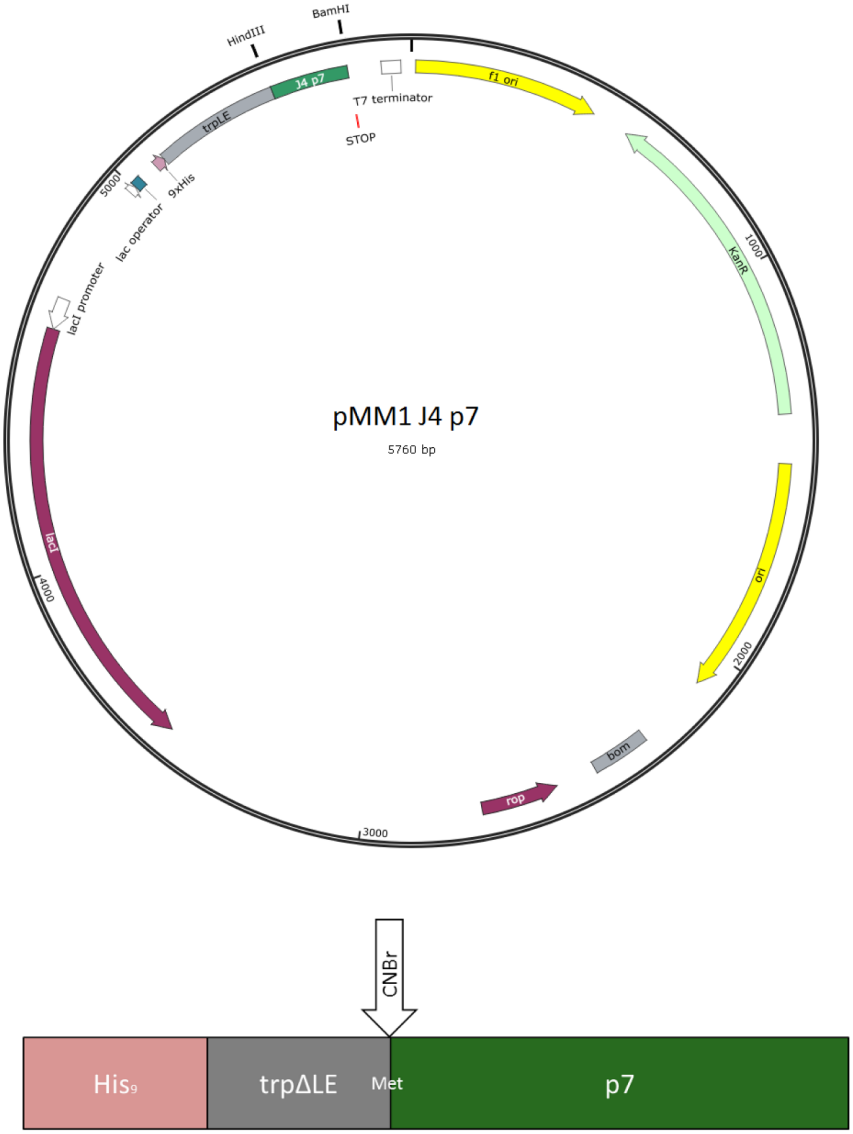

Figure III.1 – Vector map of pMM1 J4 p7. Vector map of pMM1 J4 p7 (gt 1b). The plasmid map shows the pMM-LR6 vector with inserted HCV J4 p7 (wt; wildtype) gene, referred to as pMM1 J4 p7. It contains a trpΔLE sequence (grey) that directs overexpressed protein into inclusion bodies. The fusion protein consists of a 9xHis-tag (pink) followed by the trpΔLE sequence and p7 (green) (top). Inserted between trpΔLE and p7 is a sole methionine, which is used for CNBr (cyanogen bromide) cleavage (bottom). SnapGene 2.1 was used to create the vector map.

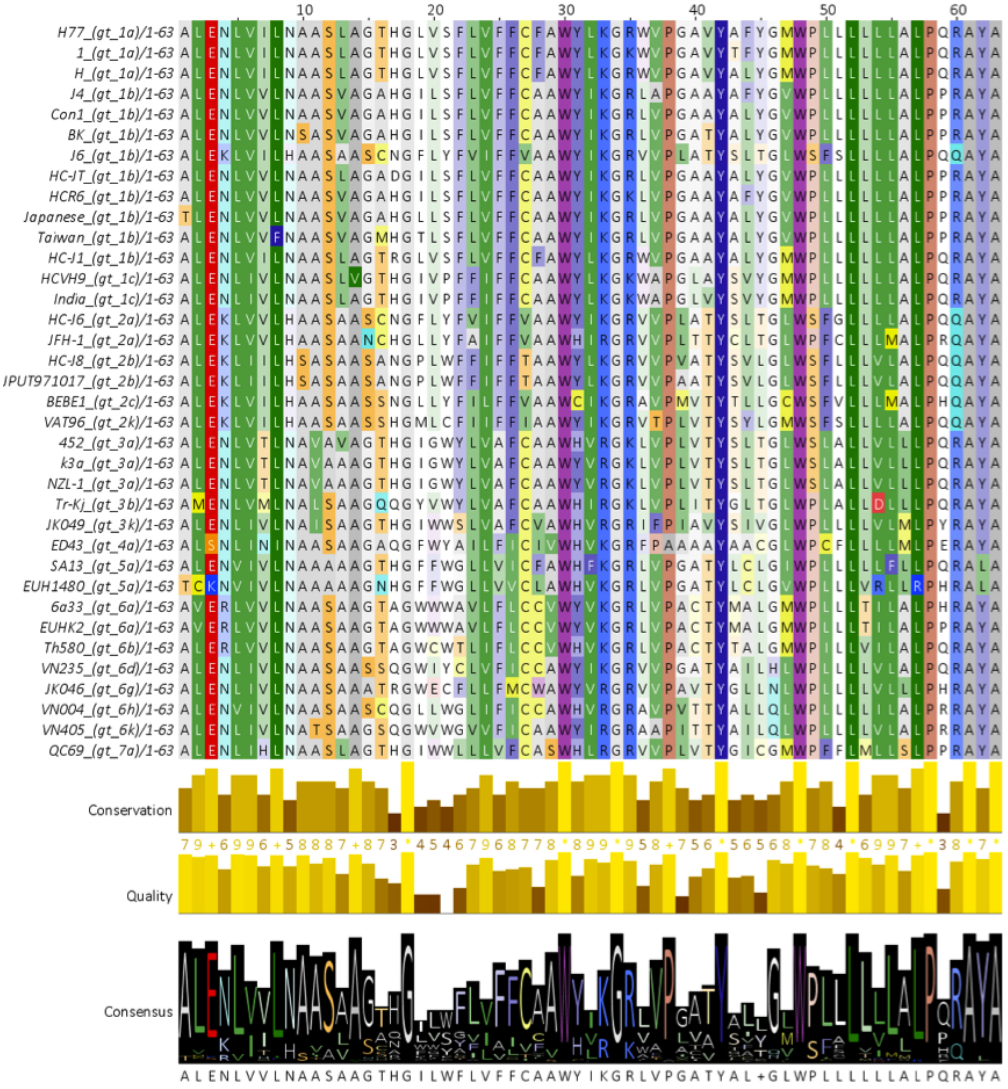

Figure III.2 – Sequence alignment of selected HCV p7 sequences. UniProt reviewed HCV p7 sequences are listed as single-letter amino acid sequences, which are depicted in RasMol colours, i.e. N and Q in cyan, K and R in blue, F and Y in mid blue, A in grey, H in pale blue, C and M in yellow, S and T in orange, E and D in red, L, V and I in green, W in purple and P in brown. Isolate names are shown next to the sequences with the corresponding genotype in parentheses. At least one example of the seven recognised genotypes is listed. The isolates of particular interest in this study are HCV J4 p7 (gt 1b, UniProt: O92972), HCV 452 p7 (gt 3a, UniProt: Q1W5D5) and HCV SA13 p7 (gt 5a, UniProt: O91936). The overall conservation is depicted as a bar graph below the alignment followed by the consensus of the isolate sequences listed (Jalview)[285].

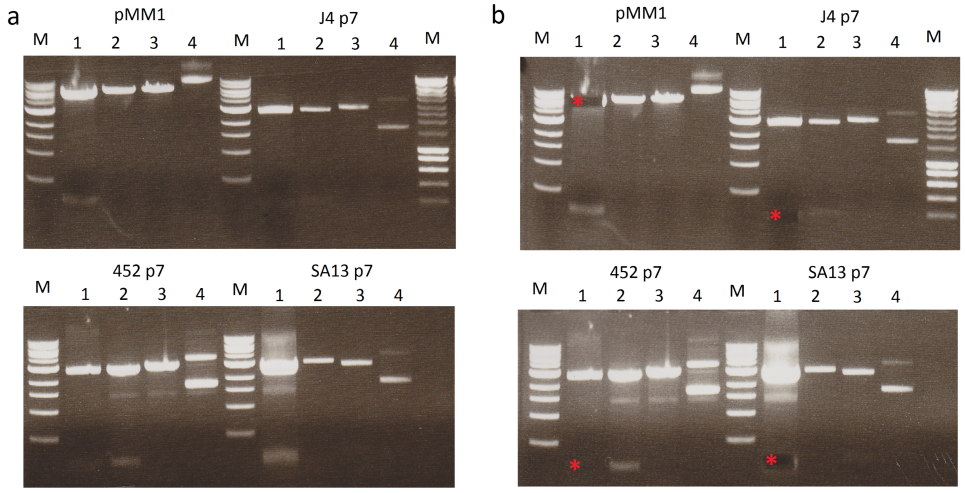

Figure III.3 – Restriction digest of the pMM1 vector and inserts HCV J4 p7, HCV 452 p7 and HCV SA13 p7. a) In order to clone the synthetic genes (genescript) into the expression vector, restriction enzymes BamHI and HindIII (NEB) were used to digest previously amplified copies of pMM1, pUC57 J4, pUC57 452 and pUC57 SA13. A volume of 50 μL per digest miniprep (Qiagen) was used. Lane 1 shows the double digest of both, BamHI and HindIII added to the plasmid DNA. Two bands are expected in this digest, the plasmid and the cut out insert, visible towards the bottom of each lane 1. Lane two to four are controls, i.e. lane two is the single BamHI and lane three the single HindIII digest, in both of which no insert but only supercoiled DNA should be visible. In lane four no restriction enzymes are present, which leads to vector specific, supercoiled running behaviour, with pMM1 running higher than pUC57, the latter being also split in multiple bands. b) The same gels are depicted with bands highlighted by red asterisk which were excised for agarose gel extraction (Qiagen) and ligation. In all gels 1 kb ladder (NEB) and SmartLadder (Eurogentech) were used, on the left and right, respectively.

Bioscience, Oxford). In order to carry out expression in *E. coli*, especially suited protein expression strains have to be used. BL21(DE3)pLysS was initially chosen for trial expressions, as it is a standard cell line in heterologous protein expression (Chapter VII). BL21(DE3)pLysS cells contain a T7-promoter, which is inducible with allolactose and its analogue IPTG (isopropyl β-D-1-thiogalactopyranoside). After heat-shock transformation, the p7-containing constructs were grown in 5 mL of LB (lysogeny broth) media up to an OD_{600} (optical density measured at 600 nm) of around 0.6 and induced with 1 mM IPTG. Fractions were analysed using SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis). The fusion protein of p7, trpΔLE and His-tag has a molecular weight of about 20 kDa. The observed overexpression band at 12 h runs a bit lower as expected on the SDS-PAGE. This size anomaly is not an uncommon phenomenon in SDS-PAGE and previously described[286]. Trial expression shows that all constructs express under these standard conditions, and there is no significant difference observed in band intensity for the fusion protein (trpΔLE-p7) at 12 h and 24 h

post-induction (Figure III.4).

III.1.3 HCV p7 M9 minimal media test expression

After the successful growth of HCV J4 p7 in rich (i.e. full) LB media, a trial scale expression in M9 minimal media was carried out next. M9 media is a basic media, which makes isotopic labelling of *E. coli*-expressed proteins possible (Chapter VII), a prerequisite for NMR studies. An LB starter culture was prepared to inoculate a 5 mL M9 culture. At an OD_{600} of 0.6, the cultures were induced with 1 mM IPTG. The SDS-PAGE analysis shows that all three isolates were expressed (black arrow, Figure III.5). Expression of HCV p7 was shown to be possible in LB and M9 minimal media.

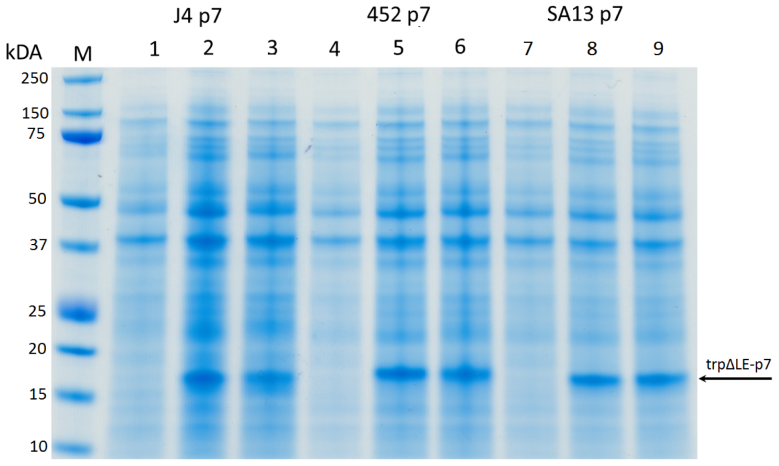

Figure III.4 – SDS-PAGE analysis of HCV p7 test expressions in LB. Whole-cell lysates of three p7 constructs (HCVJ4, HCV 452 and HCV SA13) are shown uninduced (1, 4 and 7), 12 h post-induction (2, 5 and 8) and 24 h post-induction (3, 6 and 9) on a 4 - 12 % gradient Bis-Tris gels in MES (2-*N*-morpholino)ethanesulfonic acid) buffer (NuPAGE system, Life). Running voltage applied was 200 V and the gels typically ran for 35 minutes if not otherwise stated. The marker and the respective weights are indicated at the left (Biorad Dual Xtra). The fusion protein is visible as a strong, blue band between 15 and 20 kDa (black arrow).

III.1.4 HCV p7 expression and inclusion body preparation

A large scale expression consisted typically of 5-6 L of M9 media. When tested, similar amounts of autoinduction media have been used. In order to purify inclusion bodies, an

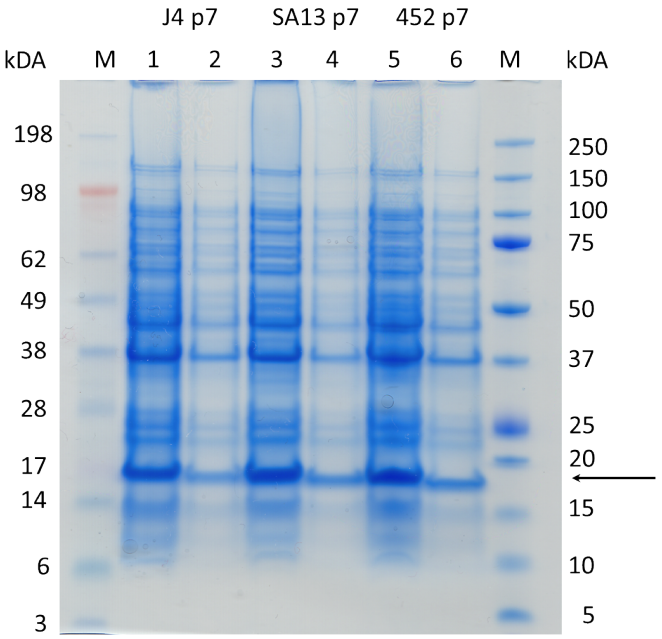

Figure III.5 – SDS-PAGE analysis of HCV p7 test expressions in M9. 12 h post induction fractions are shown. Two different sets of MW (molecular weight) markers are shown with their respective MW indicated next to the gel (left: Biorad Dual Xtra, right: Invitrogen See Blue Plus2). The p7 fusion is seen as a strong, blue band between 15 and 20 kDa, (black arrow). The difference between lane one and two is the amount of cells prepared for each lane (i.e. 250 μL vs 50 μL), with lane 1, 3 and 5 containing the higher concentration.

M2-based method was applied to p7[171, 282]. In brief, cells were harvested 12 h post-induction and resuspended in lysis buffer using a Dounce homogeniser. After sonication and centrifugation, the pellet was resuspended in lysis buffer and centrifuged again. The inclusion body containing pellet was resuspended in a 6 M guanidine buffer and loaded onto a Ni-NTA gravity column. After binding O/N (over night), the resin was washed and eluted. Several types of affinity resin can be used to purify His-tagged proteins. In general, either Nickel or Cobalt are able to coordinate and bind His-tags. Ni-NTA is nickel chelated to agarose beads via nitroloacetic acid. An initial purification using HCV J4 p7 with nickel-based HisPur Ni-NTA (Pierce) lead to unsatisfactory results, i.e. degradation and/or *E. coli* protein contamination in the eluted fractions (Figure III.6). The lack of binding and poor purification performance of HisPur could be caused by low binding capacity, which is limited by the resin volume used and/or insufficient washing pre-elution, respectively. In order to evaluate if a cobalt resin would be more suitable, TALON (Clontech) and His-Pur, were tested simultaneously with equal amounts of purified HCV 452 p7. The experiments show that the TALON resin does not bind efficiently

to the fusion protein, whereas HisPur binds larger amounts of it. Degradation products or host cell proteins are still visible (Figure III.7).

These results suggest that HisPur is more promising and suitable for optimisation, although it cannot be ruled out that TALON would have performed equally well after further optimisation. Introduced adaptions to the inclusion body purification include more washing steps (Table VII.12) and addition of imidazole in sub-elution concentrations (i.e. 50 to 150 mM) to the washing buffers, which lead to clean elution bands of trpΔLE-p7 from the Ni-NTA resin (Figure III.8). The binding capacity per purification was also increased from suggested standard values, i.e. from 5 mL per purification to up to 5 mL of resin per 1 L culture, which together with the improved washing of the resin lead to pure fusion protein bands visible in SDS-PAGE gels (Figure III.8). Fractions containing the protein of interest were dialysed against H_2O and followed by CNBr digest, which is found not to be 100 % efficient (Figure III.9).

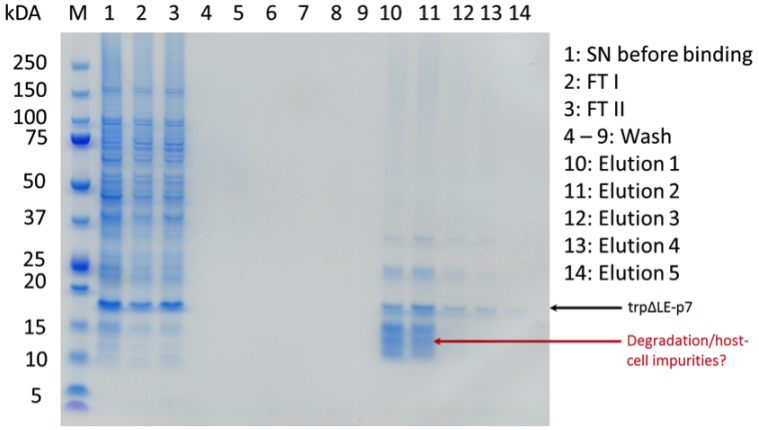

Figure III.6 – SDS-PAGE analysis of HisPur Ni-NTA IMAC purified HCV ^{15}-J4 p7. HCV J4 p7 was grown in M9 (1 L) and purified as previously described[171, 282]. Lane one shows the SN (supernatant) before application to the Ni-NTA column. FT (Flow through) lanes one and two show sequentially how much protein gets bound after each gravity flow through. The resin volume is 4 mL of HisPur. After elution with 400 mM imidazole (elution 1-5), fractions were chloroform/methanol precipitated (Chapter VII) and analysed. The eluted fractions show that besides HCV J4 p7, cellular protein and/or degradation products are eluted (red arrow). No protein bands are visible in the washing buffers lanes.

III.1.5 HPLC purification of HCV p7

In order to separate the cleaved trpΔLE-tag from p7 and uncleaved fusion protein HPLC was employed. A C4 column (Jupiter, Phenomenex), with an acetonitrile gradient in water (Buffer A: 0.1 % FA (Formic acid) in H_2O, Buffer B: 0.1 % FA in acetonitrile)

was used. HCV SA13 p7 was grown in standard M9 media (3 L) using the p7-adapted inclusion body purification (Table VIII.12). The HPLC chromatogram shows good separation of two main fractions (Figure III.10 a). MS (mass spectrometry) analysis shows that HCV SA13 p7 is contained in the late smaller peaks, and the trpΔLE tag in the earlier peak (Figure III.10 b).

Using an autoinduction protocol[287], which uses allolactose as an inducer of the T7 promoter, HCV SA13 p7 shows again a separation in two peaks in the HPLC chromatogram (Figure III.10 c). SDS-PAGE analysis of its fractions reveals a bleeding of the uncleaved fusion protein into the p7 fractions (Figure III.10 c). Autoinduction did not give higher amounts of protein in comparison to IPTG induction, with a yield of about 0.5 to 1 mg / 5 L culture.

In order to improve separation, a column with an 18 carbon chain length resin (C18) was used, which is the most hydrophobic resin applicable in HPLC with IMAC-purified and CNBr-cleaved HCV J4 p7. The HPLC chromatogram shows limited separation and the previously observed impurities in SDS-PAGES are still present in purified p7 fractions (Figure III.11).

Wildtype p7 sequences purification using HPLC was achieved for HCV J4 p7 and HCV SA13 p7 and partially for HCV 452 p7. MS analysis shows that it is possible to obtain purified p7 of the correct mass, but the impurities visible in the SDS-PAGE needed to be addressed. All isolates contain at least one cysteine. Free cysteine residues are often the cause for problems in protein expression and purification, as they can form intermolecular disulphide bonds and aggregation.

III.1.6 FPLC purification of HCV J4 p7 (C27S)

To improve the purity of p7, a previously published study employing a p7 mutant of HCV J4 p7 (C27S) and FPLC was adapted[288]. The C27S point mutation was inserted into the pMM1 J4 p7 construct using site-directed mutagenesis (QuikChange, Agilent) and the sequence confirmed by DNA-sequencing (SourceBiosciene, Oxford). The inclusion body purification and CNBr cleavage slightly differed to the previously used protocol (Table VII.12) (Figure III.12 a & b left), but the main difference is the use of FPLC, which enables the use of detergent in the running buffer. The Ni-NTA eluted and CNBr cleaved HCV ^{15}N-J4 p7 (C27S) is taken up in 10 % SDS, diluted with SDS-containing running buffer, and injected into a Sephacryl column on an Äkta pure system (GE

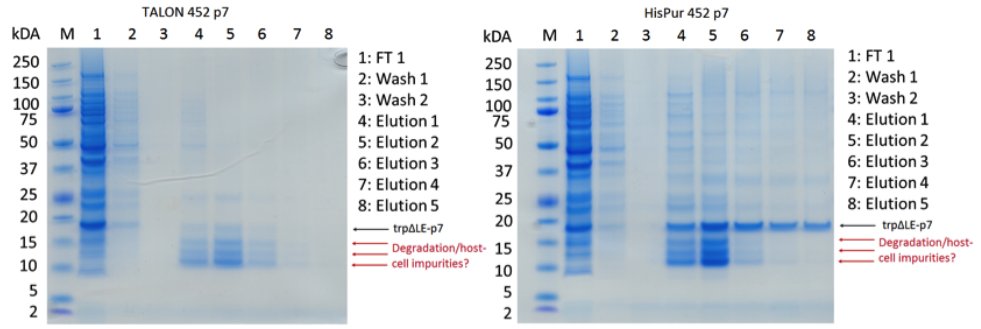

Figure III.7 – SDS-PAGE analysis of TALON and HisPur Ni-NTA IMAC purified HCV ^{15}N-452 p7. SDS-PAGE analysis of TALON and HisPur Ni-NTA IMAC purified HCV 452 p7. HCV 452 p7 was grown in M9 (3 L each) and purified as described (Chapter VII). On the left, the SDS-PAGE of the TALON purified fractions is shown. The FT shows a significant unbound amount of HCV 452 p7 and all elutions lanes show a lack of binding for HCV 452 p7 (black arrow). Red arrows indicate degradation products or contaminant host cell proteins. HisPur bound more p7 and there is a clear band visible at around 20 kDa, corresponding to the fusion protein of His_9-trpΔLE-452 p7 (black arrow). There are impurities visible in the elution fractions but not in the washes (lanes two and three in both gels, respectively).

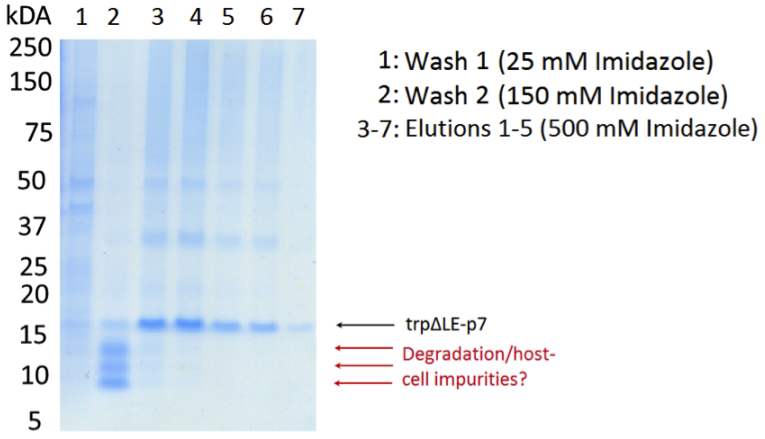

Figure III.8 – SDS-PAGE analysis of HCV p7 optimised purification conditions. An HCV SA13 p7 purification is shown with washing buffers containing increasing imidazole concentrations (25 and 150 mM, respectively), which elutes previously observed degradation and/or host-cell protein impurities. HCV 452 p7 requires a higher amount of imidazole (150 mM) in the washing buffers in comparison to HCV SA13 and HCV J4 p7 (50 mM), in order to elute the impurities before elution.

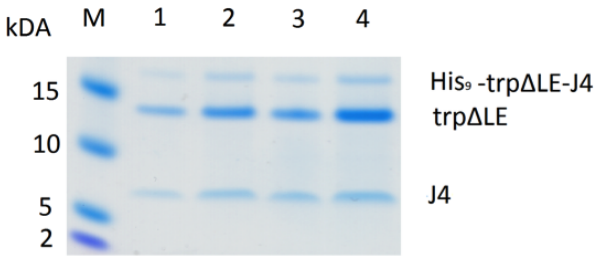

Figure III.9 – SDS-PAGE analysis of a CNBr cleavage efficiency. Lanes 1 to 4 are loaded with increasing amounts of a CNBr cleavage product (2 μL to 10 μL, respectively). Cleavage is not 100 % efficient, with the fusion protein band still visible at the top.

Healthcare). An isocratic gradient is used to separate cleaved p7 from the His_9-trpΔLE-tag and any uncleaved fusion protein (Figure III.12 b). The mass was confirmed using MS (Figure III.12 c).

In order to increase the final yield and make more economical usage of isotopic labels, a condensation protocol was established. In brief, HCV J4 p7 (C27S) was grown in 4-5 L of LB medium until it reached an OD_{600} of 0.6. Cells were harvested, washed with M9 salts and resuspended in 1 L M9-minimal media, which is the condensation step. After incubation for 20-30 minutes at 37 °C, cells were induced with 1.5 mM IPTG and grown for 12 h. Cells were harvested and purified at 24 h. The same method was also employed to grow triple labelled, deuterated (^{2}H, ^{13}C and ^{15}N) J4 (C27S). H_2O was replaced by D_2O after the cells were grown in LB and washed the first time in 1x M9 salts. ^{13}C-glucose and ^{15}N-ammonium chloride were the only carbon and nitrogen source supplied to the M9 minimal media. The final growth was typically 4 L of LB condensed into 1 L M9 medium. Induction with 1.5 mM IPTG was repeated after 12 h growth. The deuteration level of triple labelled HCV J4 p7 (C27S) was difficult to determine as MS analysis showed two separate species (data not shown) (Figure III.13).

III.1.7 HPLC purification of HCV EUH1480 p7 (T1G, C2A, A12S, C27T and C44S)

The recently published hexameric, solution NMR structure of p7 uses HCV EUH1480 p7, the second known genotype 5 sequence other than SA13 p7[242]. EUH1480 p7 is one of the few isolate sequences harbouring a threonine at its N-terminus, which is followed by a cysteine (Figure III.2). Five mutations were introduced in comparison to wildtype EUH1480, i.e. T1G, C2A, A12S, C27T and C44S, referred to as HCV EUH1480 p7

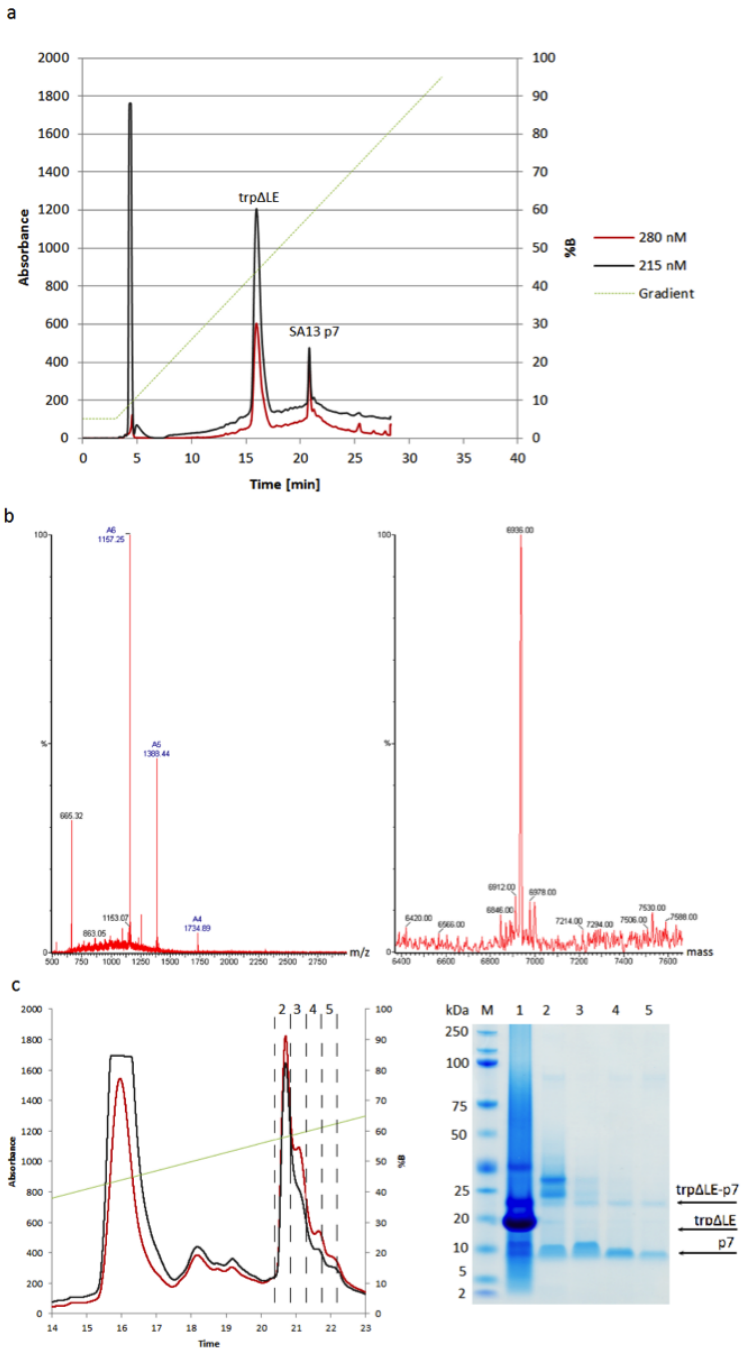

Figure III.10 – HPLC and MS analysis of HCV ^{15}N-SA13 p7. a) An HPLC chromatogram of HCV SA13 p7 is shown. A C4 column (Jupiter 300 Å, 10 μ, Phenomenex) was used with a Gilson PrepLC system and an acetonitrile gradient in water (Buffer A: 0.1 % FA in H_2O, Buffer B: 0.1 % FA in acetonitrile). Absorbance at the two wavelengths, 215 nm (black trace) and 280 nm (red trace), is shown. The trpΔLE tag and p7 get separated in the HPLC trace. b) The MS analysis shows HCV SA13 p7 of the expected mass. c) A different HCV SA13 p7 purification (HPLC trace detail left) and SDS-PAGE analysis of fractions shows that uncleaved trpΔLE-p7 fusion proteins leaks into the pure p7 fractions.

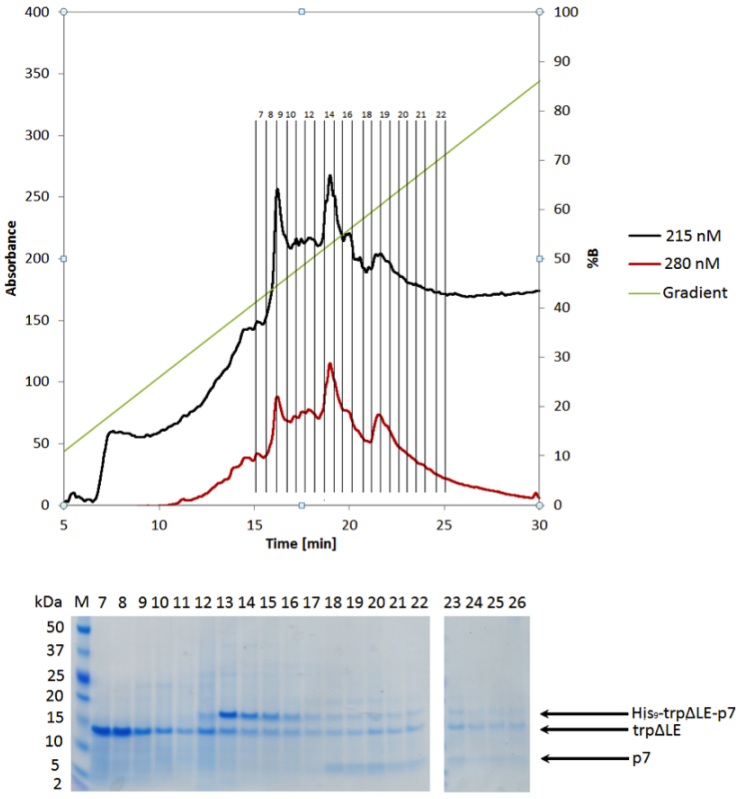

Figure III.11 – HPLC purification of HCV ^{15}N-J4 p7. The HPLC chromatogram of HCV J4 p7 (wildtype) shows the bleeding of the full length fusion protein into the trpΔLE tag and p7 fractions. A C18 column (Jupiter 100 Å, 5 μ, Phenomenex) was used with a Gilson PrepLC system and an acetonitrile gradient in water (Buffer A: 0.1 % FA in H_2O, Buffer B: 0.1 % FA in acetonitrile).

(mt5) (Figure III.14). The inclusion body preparation differs in a few details from the one employed for HCV J4 p7 wildtype and the C27S mutant, but it is overall very similar to the published M2 protocol[282]. The latter was initially used for HCV J4 p7 wildtype and uses no detergent in the inclusion body purification[282]. The protocol optimised for HCV J4 p7 (C27S) includes a detergent step (1 % deoxycholate and 1 % IGEPAL) which in the purification of HCV EUH1480 p7 (mt5) are replaced by a 1 % Triton-X100 step in the guanidine buffer[242]. The wildtype and five mutations harbouring HCV EUH1480 construct were ordered as DNA-strings (GenArt), digested and ligated into the pMM1 vector as previously described and trialled for expression in LB media. HCV EUH1480 wildtype and mt5 both show high expression levels, especially in comparison to previously studied isolates (Figure III.15). A condensation expression using 8 L LB cultures for a 2 L M9 media final volume was used to express ^{15}N-labelled HCV EUH1480 p7 (mt5). Expression levels were high and similar to intensities

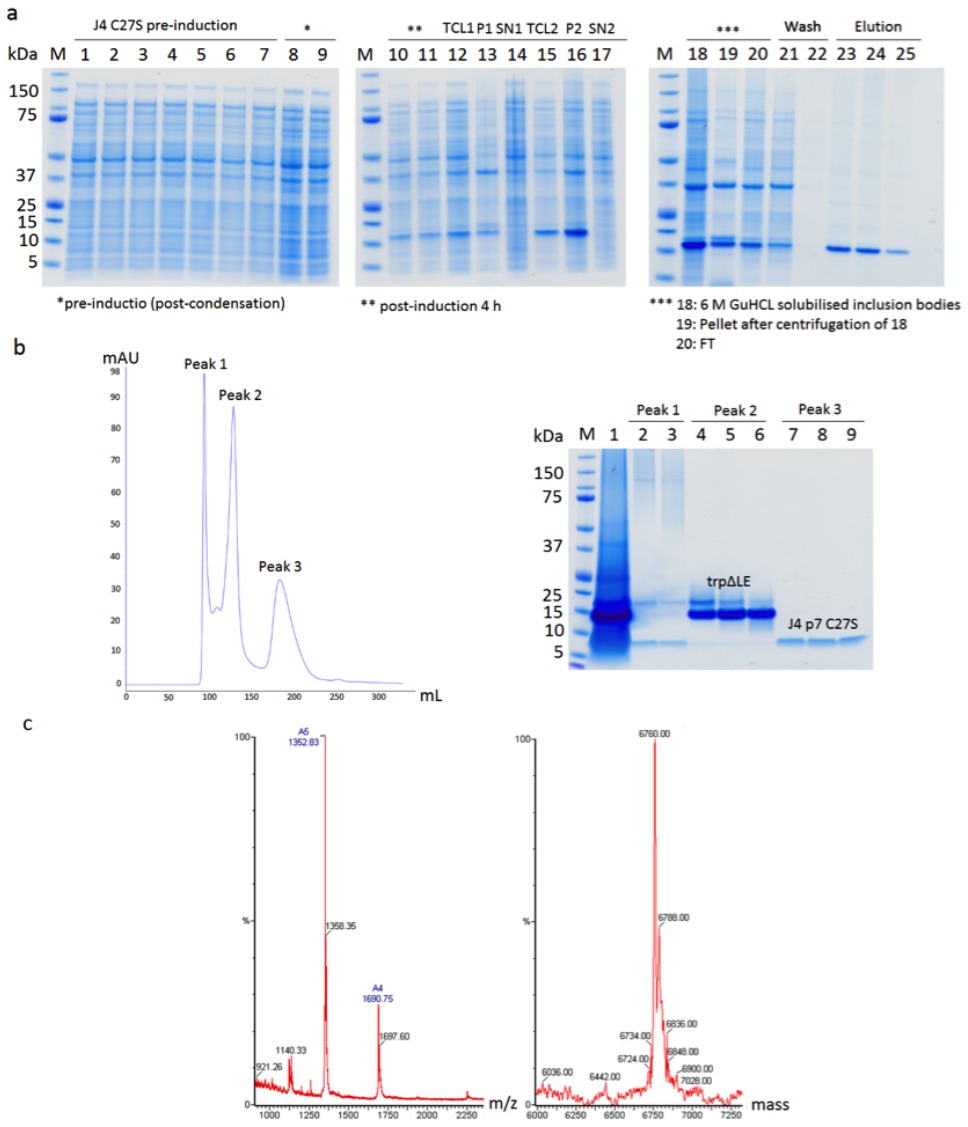

Figure III.12 – FPLC purification of HCV ^{15}N-J4 p7 (C27S). a) An SDS-PAGE analysis of purification steps performed to obtain pure fusion protein (His_9-trpΔLE-p7) is shown. The left gel shows the *E. coli* in LB media until they reach an OD_{600} of around 0.6. Following the harvest and a wash in M9 salts, cells were resuspended in M9 minimal media. In this case the media contained ^{15}N-ammonium chloride to isotopically label the protein. The middle gel shows the inclusion body purification. TCLs (total cell lysates), Ps (pellets) and SNs (supernatants) are all part of the resuspension-centrifugation cycle to separate the inclusion body from the rest of lysed *E. coli* cells. After each centrifugation step, a solubilisation in a Dounce homogeniser and sonication leads to the TCLs. The HCV J4 p7 (C27S) containing inclusion bodies are found in the insoluble pellet fractions. The resin is washed and fusion protein is eluted with 500 mM imidazole buffer. b) The fusion protein is shown post CNBr cleavage on the SDS-PAGE on the left, with full-length fusion protein, trpΔLE, and p7 visible (indicated by the black, red and green asterisk, respectively). The sample is then solubilised in 10 % SDS, mixed with SDS-containing FPLC running buffer (Chapter VII) and run on a Sephacryl 200H column, using an Äkta pure FPLC system (GE Healthcare). Three peaks are visible. Peak 1 is the aggregation peak and peak 2 contains the His_9-trpΔLE tag. Peak 3 contains J4 p7 (C27S), c) whose mass is confirmed by MS (Biochemistry Department, Oxford University).

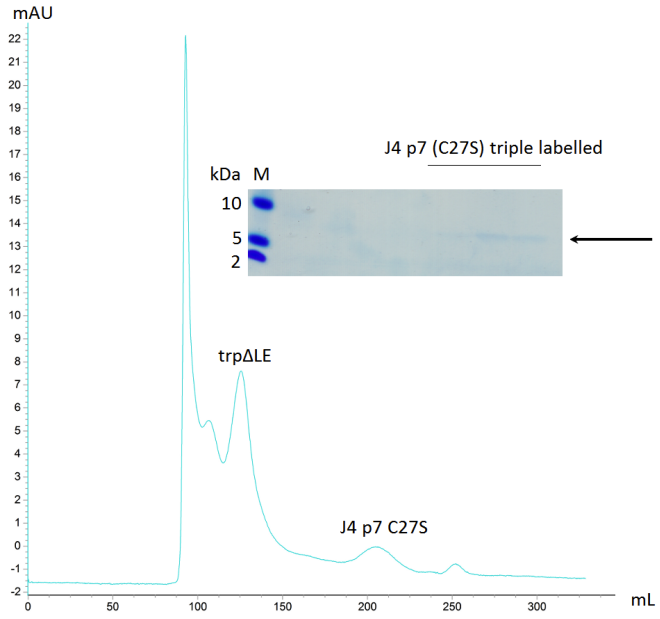

Figure III.13 – FPLC purification of HCV ^{2}H, ^{13}C and ^{15}N-J4 p7 (C27S). As described for HCV ^{15}N-J4 p7 (C27S), FPLC was used to purify triple labelled J4 using the condensation protocol, with D_2O replacing H_2O, and ^{13}C-glucose and ^{15}N-ammonium chloride as the sole isotope sources. Because of these growing restrictions, the resulting yield is lower. Typically an OD_{280} of 7-9 is injected per run, whereas here a 8 L condensation growth only gave an OD_{280} of 4. SDS-PAGE analysis (insert) confirms the low protein content. Fractions were collected and lyophylised O/N and analysed using MALDI-TOF (data not shown).

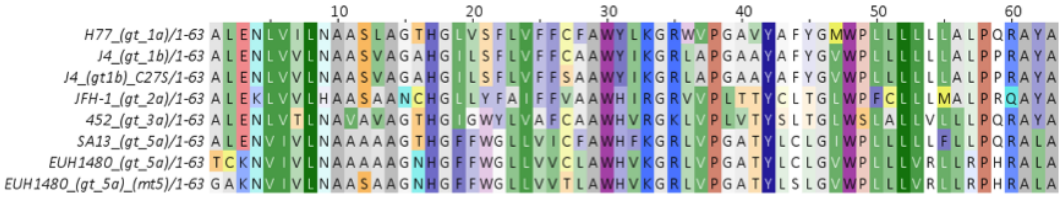

Figure III.14 – Sequence alignment of HCV EUH1480 p7 (gt 5a). Selected isolates spanning relevant genotype families in comparison to wildtype and mutated isolate EUH1480 p7 (gt 5a, UniProt: O39928) are shown. Note the threonine and cysteine at position one and two, respectively. The five mutations introduced are T1G, C2A, A12S, C27T and C44S (Jalview)[285].

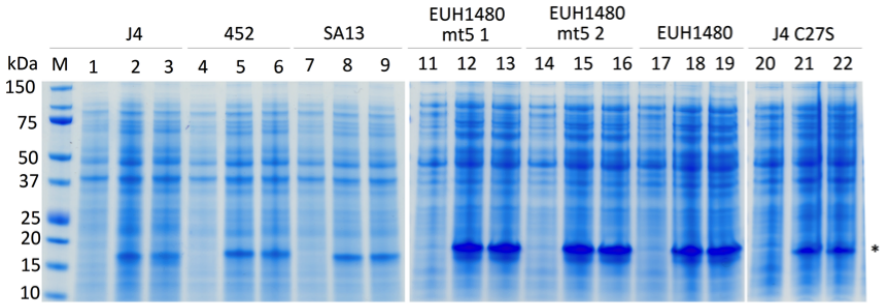

Figure III.15 – SDS-PAGE analysis of HCV EUH1480 and HCV EUH1480 p7 (mt5) LB test expressions. Expression of native and the five mutation harbouring HCV EUH1480 p7 (mt5) is compared to constructs used previously. All samples were treated equally, i.e. an overnight culture was used to inoculate a LB culture (5 mL) the next day. At OD_{600} = 0.6 all cultures were induced with 1 mM IPTG, and samples harvested at 12 h and 24 h post induction. Fractions of three initially used p7 constructs (J4, 452 and SA13) are shown on the left gel with uninduced (lanes 1, 4 and 7), 12 h post induction (lanes 2, 5 and 8) and 24 h post-induction fractions (lanes 3, 6 and 9). The fusion protein is visible as a strong, blue band between 15 and 20 kDA (black asterisk). Next are two identical clones of the HCV EUH1480 p7 (mt5) construct, picked from different single *E. coli* colonies off an agar plate (lanes 11-16) followed by wildtype EUH1480 p7. An example for a HCV J4 p7 (C27S) expression is shown on the left (20-21). The HCV EUH1480 p7 constructs express stronger under the same conditions.

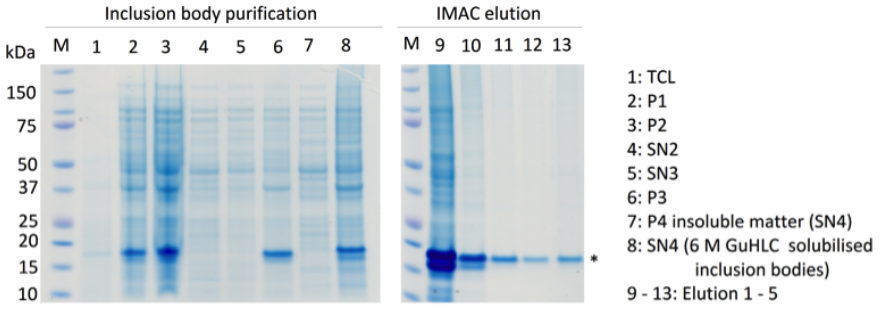

Figure III.16 – SDS-PAGE analysis of HCV ^{15}N-EUH1480 p7 (mt5) condensation expression. The inclusion body purification was performed as described[242], using Triton-X100 in the guanidine buffer instead of previously used detergents in the washing steps. HCV EUH1480 p7 (mt5) is found in the insoluble pellet fraction. After solubilisation in 6 M guanidine HCl (lane 8) and centrifugation to separate insoluble parts (lane 7), IMAC was performed using HisPur. After washing steps, protein was eluted with a 500 mM imidazole buffer (lanes 9-13). There is unspecifically bound protein visible in elutions one and two, which were handled separately from elutions 11-13. The fusion protein is indicated by a black asterisk.

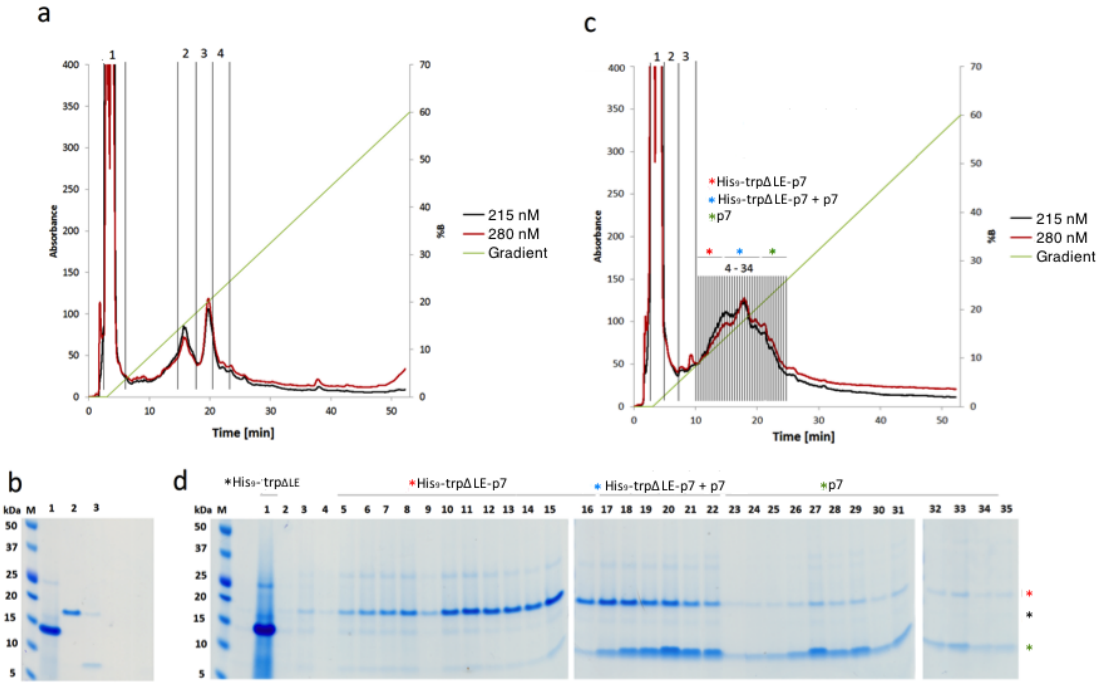

Figure III.17 – Purification of HCV ^{15}N-EUH1480 p7 (mt5). a) Shows an HPLC chromatogram with labelled fractions corresponding to the SDS-PAGE in b). Fraction 1 contains the trpΔLE-tag, fraction 2 the full-length fusion protein (His$_9$-trpΔLE-p7) and fractions 3 and 4 (the latter not shown on SDS-PAGE), contain HCV EUH1480 p7 (mt5). c) Shows the same HPLC running conditions but with a higher loading concentration (10 mg/ml vs. 18 mg/ml). The separation is no longer achieved and components are not clearly separated based on the HPLC chromatogram. d) A SDS-PAGE analysis of the fraction shows that peak 1 corresponds to the His$_9$-trpΔLE tag (black asterisk), with peaks 5-16 containing mainly the uncleaved fusion protein His$_9$-trpΔLE-p7 (red asterisk) with pure p7 components also present. Peaks 17-22 contain a mixture (blue asterisk) and peaks 23-34 contain pure p7 (green asterisk) with smaller components of uncleaved fusion protein.

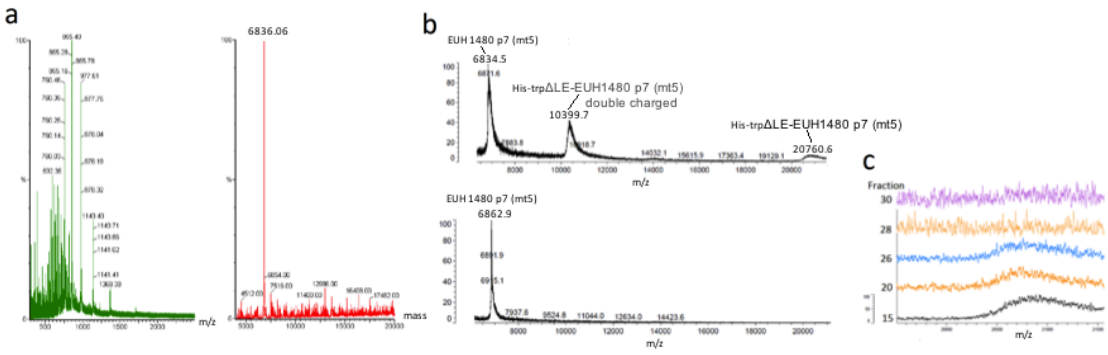

Figure III.18 – MS of HCV ^{15}N-EUH1480 p7 (mt5). a) MS analysis of fraction 3 shows the correct mass of ^{15}N-labelled EUH1480 p7 (mt5). b) MALDI analysis of fractions 15 (top) and 28 (bottom). The chromatogram shows that the fusion protein is present in fraction 15, but not in fraction 28. c) Shows fractions 15, 20, 26 and 30 in a MALDI chromatogram in the region of the fusion protein.

seen in the test expressions, with a blue to brown discolouration of the IMAC resin occurring, which was more intense than the usually observed colour change due to the reducing agent TCEP (Tris (2-carboxyethyl) phosphine hydrochloride). Inclusion bodies were purified from host-cell membrane and debris as described in detail in chapter VII. Solubilised inclusion bodies were bound to HisPur and eluted with an imidazole buffer. Fractions one and two contained a lot of impurities and were handled separately from fractions three to five. All fractions were lyophilised and purified using HPLC on a C18 column (Proto, 300 μ, Higgins Analytical), with buffers according to the published NMR study[242] (Figure III.16). The HPLC chromatogram shows a different behaviour to that previously reported, with HCV EUH1480 p7 (mt5) eluting earlier (Figure III.17 a and b). MS analysis confirmed the mass of ^{15}N-labelled HCV EUH 1480 p7 (mt5). The expression and purification of HCV EUH1480 p7 (mt5) was successful, with yields of about 20 mg fusion protein pre-CNBr cleavage, and only using elutions 3 to 5, a final yield of about 1 mg. The efficiency of the CNBr cleavage was not high, as there is still uncleaved fusion protein visible (Figure III.17 b and d). For the second HPLC run, a higher loading concentration than before was used (18 mg/ml vs. 10 mg/ml), resulting in merging of peaks. SDS-PAGE analysis reveals the bleeding of fusion protein into the p7 only peaks and vice-versa (Figure III.17 d). MALDI analysis confirms the presence of fusion protein in fraction 15 and, although visible as a faint band on the gel, not in fraction 28 (Figure III.18 f). A comparison of the MALDI chromatogram showing the fusion protein section of fractions 15, 20, 26, 28 and 30, confirms its presence in fraction 15, 20 and 26 and its absence in fraction 28 and 30 (Figure III.18 g).

III.2 Discussion

The aim of the work described in this chapter was to develop and establish a bacterial expression system and purification method for HCV p7. Up until the start of this project, no system was established that made the expression of native p7 sequences without addition of tags and mutations possible. Also the low protein yields previously achieved were prohibitive for thorough analysis using NMR, which requires large quantities of isotopically labeled protein.

The expression system based on the pMM-LR6 vector, employing a trpΔLE-tag that directs expressed heterologous protein into inclusion bodies, was previously shown to

enable expression and purification of influenza M2[171], another member of the viroporin family. This prompted the decision to adapt this system for usage with HCV p7. The inherent limitation to the system is that the target sequences must not contain a methionine residue. CNBr, which is used to separate the tag from the target protein, specifically cleaves after methionine residues. As the Hepatitis C virus is an ensemble of related genotypes and subtypes, encompassing hundreds of sequences, identifying wild-type sequences that do not contain methionine is straightforward Three isolates were initially chosen, spanning genotype families 1, 3 and 5. The genotype 1 isolate J4 was previously used in experiments and was a prime candidate, as infection with genotype 1 is currently difficult to treat. Genotype 3 isolate 452 and genotype 5 isolate SA13, show also interesting characteristics. The former is implicated in a reported controversial genotype specific iminosugar resistance[206] and the latter by being only one of two recognised sequences of its genotype family, respectively. Trial and large scale expression worked satisfactorily after several optimisation steps. It was found that the choice of bacterial cell is not critical. BL21 Rosetta 2, which carry a codon for rare eukaryotic genes was tested, given that HCV infects and propagates in humans, but did not yield more protein. Standard BL21 were found to be the easiest to handle and yielding good results in LB expression and were chosen for minimal media expression. Minimal media are needed to express isotopically labeled protein, but the yields obtained are generally lower than the yields obtained in full media. Autoinduction, which is reported to increase yields in other studies[287], did not convincingly give more than a standard IPTG expression. It also requires more ^{15}N-ammonium chloride and ^{13}C-glucose for labeling, which taken together were reasons not to pursue it further.

A previously published protocol for the viroporin M2 was initially adapted but lead to impurities in the elution from IMAC columns. This was overcome by the addition of imidazole to the IMAC washing buffers and detergents in the inclusion body preparation combined with high salt washes. It is also noteworthy that the stated binding capacities of tested resins were not sufficient to bind all the p7 fusion protein. Instead of 5 mL resin, typically up to 30 mL had to be used to obtain sufficient amounts of protein from one protein preparation, which included up to 8 L M9 media. The purification using HPLC and a gradient using acetonitrile and water with added 0.1 % FA worked and was shown to separate the compounds on the chromatogram. SDS-PAGE however showed that the uncleaved fusion protein was also found in the p7 fractions.

In the course of this work, an expression and purification protocol was published by

Cook *et al.*[288], which also uses the p7 isolate J4, but with a mutation at position 27 to change the sole cysteine to a serine residue, thus inhibiting the formation of undesired disulphide interactions. Up until then, the reducing agent TCEP was used in all our buffers in order to prevent the formation of disulphide bonds. A first test with the HCV J4 p7 (C27S) construct showed similar expression in LB and M9 media to wildtype HCV J4 p7. The protocol also suggested changes in the CNBr cleavage protocol, i.e. addition of 1N NaOH in order to stop the cleavage reaction that requires an acidic environment[282, 289]. It was observed that omitting this step lead to non-separation in the next purification step of HCV J4 p7 (C27S), the FPLC. Using an isocratic gradient with SDS-containing running buffer made it possible to repeatedly purify HCV J4 p7 (C27S), with yields between 0.5 to 1.0 mg per 5 L culture.

Preparing large amounts of M9 media is laborious and expensive, especially if ^{13}C-labelling and deuteration of samples are needed. The adoption of a condensation protocol exploited the preference of *E. coli* to grow in a fully supplemented media (e.g. LB) and small M9 media volumes, which are more readily prepared. For isotopically labeled HCV J4 p7 (C27S), 4-5 L of LB were prepared and condensed in 1 L of M9 media, which is easier to handle and more cost-effective. Similar amounts of protein were obtained from a 5 to 1 L condensation in comparison to a slightly larger M9 expression (1.2 to 1.3 mg per purification, respectively). The difference in reported yields might be explained by the usage of bioreactors to grow the cells, which optimises *E. coli* nutrition supply at the later stages of a growth[288]. A triple labeled sample for NMR studies was made, but yields need to be much improved, such as by scaling up the condensation steps performed at the same time, e.g. 15 L LB media condensed into 3 L M9 media or double colony selection[290]

The recently published first hexameric solution NMR structure of a HCV isolate, also using a trpΔLE vector system for expression and purification, triggered the adaptation of this isolate to our system for further study. Interestingly, it is besides HCV SA13 p7 one of the other two reported genotype 5 isolates. The sequence of the EUH1480 isolate is one of the least hydrophobic ones found in the p7 sequence family. Other mutations, such as an exchange of a cysteine at position 44 with a serine residue, made the sequence even more hydrophilic. Trial expressions show that both the wildtype and the mutated HCV EUH1480 p7 isolate express with visible higher yields than all constructs previously used in this work, which is obviously an advantage in structural biology studies. The washing steps suggested were found to lead to unsatisfactory results as a lot of

unspecific protein was eluted in the p7 fractions. This made a final estimate of protein yields difficult, as the repurification of the first two elutions was necessary. Increasing amounts of imidazole in the washing buffers, as used for other isolates in this work, should prevent this in the future.

The HPLC purification also showed a different behaviour to the reported study[242]. Rather than eluting late at the end of an acetonitrile gradient, the p7 protein was found to elute approximately 25 to 30 minutes earlier. A first run showed a satisfactory separation and MS confirmed the correct mass of 6837 Da of the ^{15}N-labelled HCV EUH1480 p7 harbouring the five described mutations. A run with a higher loading concentration, 18 mg/ml instead of 10 mg/ml, lead to an unresolved peak. Its fractions needed to be analysis via SDS-PAGE to identify their content. Using MALDI-TOF, fusion protein, the trpΔLE-tag and p7 were successfully identified. It is worth pointing out that the cleavage efficiency was low, with significant amounts of uncleaved fusion visible in the SDS-PAGE analysis. An increase in cleavage time or a similar protocol to the one employed for HCV J4 p7 (C27S) might be necessary to improve the CNBr cleavage reaction. The discrepancy found in the elution time during HPLC purification, with an early eluting p7 in our hands, might be explained by the use of two running buffers with very similar and high organic solvent content (Buffer A: 40 % acetonitrile and buffer B: 60 % acetonitrile)[242]. Instead, a less organic buffer A should be tried, i.e. 5 % acetonitrile.

In conclusion, a system to successfully express and purify HCV J4 p7 (C27S) and others has been implemented, including the isolate that led to the first published hexameric p7 structure. Initial trials employing using the cost-effective DTT (dithiothreitol) as a reducing agent in the FPLC SDS-running buffer to purify wildtype J4 p7 wildtype should be followed up. Expression and study of native sequences is worth pursuing, as possible changes in behaviour or structural changes by mutations can be ruled out. The study of naturally methionine and cysteine-free isolates might also prove interesting. The p7 isolates 452 and SA13 have been adapted for the FPLC expression system, by introducing suitable cysteine to serine mutations and should be tested. The genotype 3 p7 isolate 452 is implicated in a suggested iminosugar resistance via a genotype-specific point mutation at position 25 (F25A)[206], which will be worthwhile to follow up using solution or solid-state NMR.

Chapter IV

Solution-state nuclear magnetic resonance of HCV p7

IV.1 Solution-state nuclear magnetic resonance of HCV p7

HCV p7 is a viroporin[163], which is crucial for infectious HCV production[139, 141] and a potential target for therapeutic small molecule inhibitors. The binding of inhibitors to their targets, like other physiologically relevant protein-ligand interactions, such as antibody-antigen recognition or enzyme-substrate binding, requires specific recognition of the ligand. In most cases binding occurs over a defined area on the target protein surface[262]. Understanding the molecular details of an interaction can help to elucidate the underlying inhibitory mechanism of a drug, possibly paving the way for the rational design of more potent inhibitors and the ability to predict their binding energetics. Several techniques are available to study protein-ligand interactions that allow determination of binding characteristics such as stoichiometry or thermodynamic properties and binding constants. These techniques include isothermal titration calorimetry (ITC), MS or surface plasmon resonance (SPR)[262].

Solution-state NMR is uniquely able to determine affinities and structural information on an atomistic level simultaneously, with low false positive rates. As the experiments are in solution, studies under physiologically relevant pH and temperature conditions are possible. Another advantage is that it can determine low-affinity, weak binding events,

which are difficult to study using alternative methods[262]. Its main limitations are low sensitivity (collecting high-quality data is usually only possible on samples reaching concentrations in the order of 0.1 mM), low throughput, and the need for isotopically labelled protein. Isotopically labelled protein can be produced by chemical synthesis and recombinant expression. The former is usually prohibitively expensive. The latter can be costly albeit much less than fully labelled, chemically synthesised proteins and also time consuming. This is especially the case for membrane proteins, where expression is often hampered by low yields (Chapter III).

The key experiment for the study of protein-ligand interactions by solution NMR is the 2D ^{1}H,^{15}N-HSQC. ^{15}N- and ^{13}C-labelling is readily achieved once expression and purification yields are sufficiently high, by employing the ability of *E. coli* to grow in minimal media, in which the sole nitrogen and/or carbon source can be controlled and modified. Uniformly ^{15}N-labelled protein gives at least one signal per nonproline amino acid in a 2D ^{1}H,^{15}N-HSQC, which transfers magnetisation from the amide proton to the attached nitrogen and back (Chapter II), giving a 'fingerprint' of the protein backbone conformation and environment. Interactions in HSQCs are observed by recording a spectrum without interaction partner present, followed by recording of successive spectra after addition/titration of the interaction partner to the sample. Either direct-binding interactions or binding-induced changes in structure will alter the environment of the amino acids and cause a change in their chemical shift. This perturbation is the basis for detection and characterisation of the binding event, which is generally referred to as chemical shift mapping or chemical shift perturbation (CSP)[291].

Assignment of each of the chemical shifts in the 2D ^{1}H,^{15}N-HSQC to their respective amino acid in the protein sequence makes it possible to identify amino acids, which are involved in the interaction at atomistic detail[252, 259, 270]. In many cases, a chemical shift is indicative of a direct interaction (short-distance effect), but allosteric changes causing structural rearrangements (long-distance effect) can also induce chemical shift perturbation[259]. Discrimination between direct and indirect effects of binding can often be delineated by mutagenesis or observation of NOEs between protein and ligand.

In order to carry out interaction studies using previously reported HCV p7 inhibitors, such as *N*N-DNJ[223], isotopically labelled HCV p7 was produced (Chapter III) and used for initial condition screens. Other studies report the use of organic solvents in NMR structure determination or solvent use for drug solubilisation in titration experiments, as well as tagged p7 of non-native sequence, high temperatures and/or low pH

in several interaction studies. This taken together with the general lack of consensus in secondary and tertiary structures triggered the search for more physiologically relevant sample conditions at near neutral pH and 37 °C, that would also permit recording of high quality NMR data.

IV.1.1 Detergent screen

Initial detergent screens were performed using wild-type HCV J4 p7 and HCV SA13 p7. DH7PC (1,2-diheptanoyl-*sn*-glycero-3-phosphocholine, Avanti Polar Lipids) is a short-chain phospholipid, which is suitable for studying membrane proteins[292, 293]. The reported hexameric EM structure of HCV JFH-1 p7 in DH7PC micelles[236] further implied its possible use for studying p7 in solution-state NMR. HCV J4 p7 and HCV SA13 p7 readily dissolve in DH7PC, which has a critical micelle concentration (CMC) of 1.4 mM. For NMR experiments, Shigemi microtubes (matched to D_2O) and a 40 mM sodium phosphate buffer were used and all samples contained 5 mM TCEP, and 5 % D_2O and DSS as lock and reference standards, respectively, if not otherwise stated. Lyophilised HCV SA13 p7 was solubilised in 60 mM DH7PC by adding the detergent in 40 mM sodium phosphate buffer at pH 7.0 to the protein, followed by incubation at 37 °C on a thermoshaker (Eppendorf).

The final HCV SA13 p7 sample concentration was about 40 μM. A 2D ^{1}H,^{15}N-HMQC was chosen instead of a 2D ^{1}H,^{15}N-HSQC because of its superior sensitivity (Chapter II). Visible peaks in the HMQC spectrum are glycine residues in the top area of the spectrum (∼104 ppm in ^{15}N) and doublet NH_2 peaks from asparagine and glutamine residues at ∼112 ppm (^{15}N). The indole NH side chain peaks of tryptophan are typically shifted downfield in ^{1}H and ^{15}N and are observed in the region of ∼9-10 ppm (^{1}H) at the bottom left corner. However, the SOFAST-HMQC was recorded with a narrow, shaped pulse at 8.2 ppm, and so the tryptophan sidechain resonances are not often detected in these experiments. A crosspeak that likely arises from the C-terminal residue is visible in the bottom right of the spectrum (∼128 ppm in ^{15}N). Some peaks are well dispersed and individual peaks can be distinguished.

The 2D ^{1}H,^{15}N-HMQC of HCV SA13 p7 in DH7PC shows that there are a small number of peaks resolved, such as the C-terminal alanine residue and one faint glycine out of a

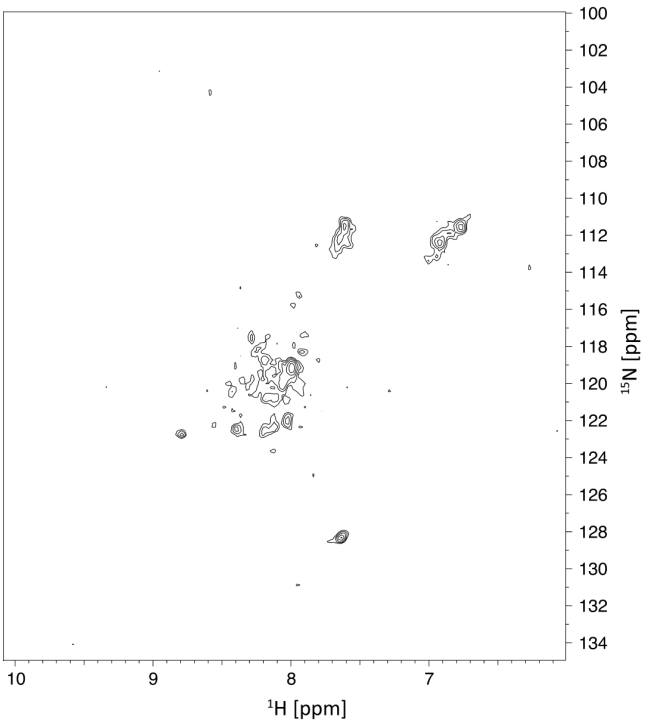

Figure IV.1 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-SA13 p7 in DH7PC. The conditions used are based on the published EM p7 structure. 60 mM DH7PC, 40 mM sodium phosphate and 5 % D_2O, with DSS used as a chemical shift reference. The pH was adjusted to 7.0. The proton dimension is shown as the x-axis (^{1}H) and the nitrogen dimension is depicted on the y-axis to the right (^{15}N). The sample concentration was 42 μM. The C-terminal residue likely corresponds to the peak visible at ∼128 ppm (^{15}N), and a single weak glycine (out of five total) is visible near the top. Asparagine side chain resonances are apparent at ∼112 ppm (^{15}N). Overall, the observable number of peaks is low, and the dispersion and peak homogeneity is poor. The spectrum was recorded at 37 °C on a 600 MHz ^{1}H-frequency Omega spectrometer.

possible five. The overall number of peaks is low, also likely due to the low protein concentration. Only a few peaks are visible and are clustered in the central area indicating poor dispersion (Figure IV.1).

The next detergent analysed was Fos-choline-12 (*n*-dodecylphosphocholine, Affymetrix), which contains a phospholipid-like headgroup attached to a single acyl chain (CMC of 1.5 mM). The same buffer was used as in the sample with DH7PC. Lyophilised HCV SA13 p7 was solubilised in 250 mM Fos-choline-12 and the final HCV SA13 p7 sample concentration was 200 μM. A greater overall number of peaks is visible in this spectrum, which is likely partly due to the increased protein concentration. More glycines were visible and other individual peaks throughout the spectrum. However, the middle of the spectrum remained poorly resolved, indicating aggregation or formation of variable

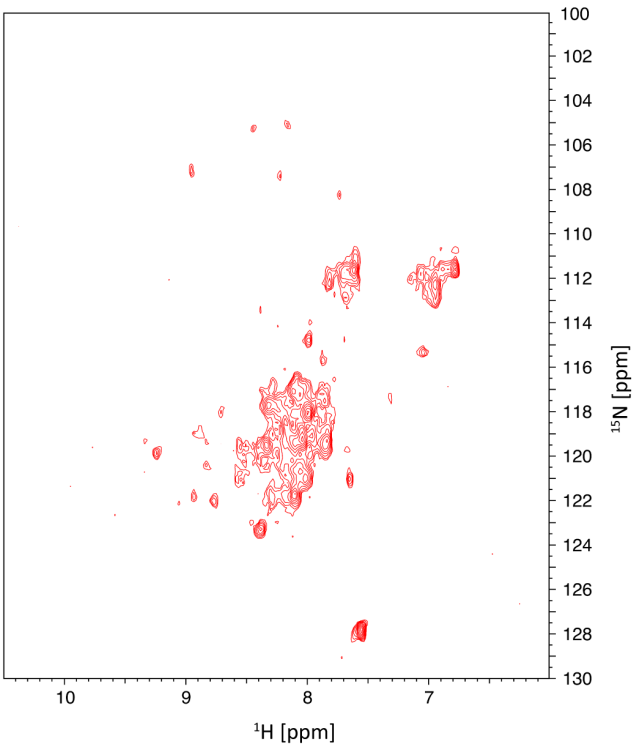

Figure IV.2 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-SA13 p7 in Fos-choline-12. Fos-choline-12, which contains a lipid-like headgroup and a single acyl chain, was tested at a concentration of 250 mM. The pH was adjusted to 7.0 and the temperature was 37 °C. The concentration of the sample was 200 μM. The C-terminus is the peak visible at ∼128 ppm (^{15}N), and several weak glycine peaks can be seen in the region ∼104-108 ppm (^{15}N). The two putative asparagine side chains are visible at 112 ^{15}N ppm and 7 and 7.6 ^{1}H ppm, respectively. Overall, the number of peaks increased, but the backbone region in the centre is not well resolved. The spectrum was recorded on a 600 MHz ^{1}H-frequency Omega spectrometer.

oligomers, sample heterogeneity and possibly interfering dynamic effects. The various NH_2 doublets visible for two asparagine and the single glutamine residue indicate multiple conformers (Figure IV.2).

Since Fos-choline-12 was more promising than DH7PC, we sought to improve the spectra by adding Fos-choline-10 (*n*-decylphosphocholine, Affymetrix), which is very similar to Fos-choline-12, only differing in two methylene groups in the acyl tail with a CMC of 11 mM. Fos-choline-10 was added to make a final detergent concentration of 125 mM Fos-choline-10 and 250 mM Fos-choline-12. Some more resolved peaks became visible but other peaks disappeared. Addition of 250 mM Fos-choline-14 (*n*-tetradecylphosphocholine, Affymetrix), which has the same headgroup as Fos-choline-10 and Fos-choline-12, but with 14 carbons in the acyl chain, (CMC of 0.12 mM) showed no overall improvement (Figure IV.3).

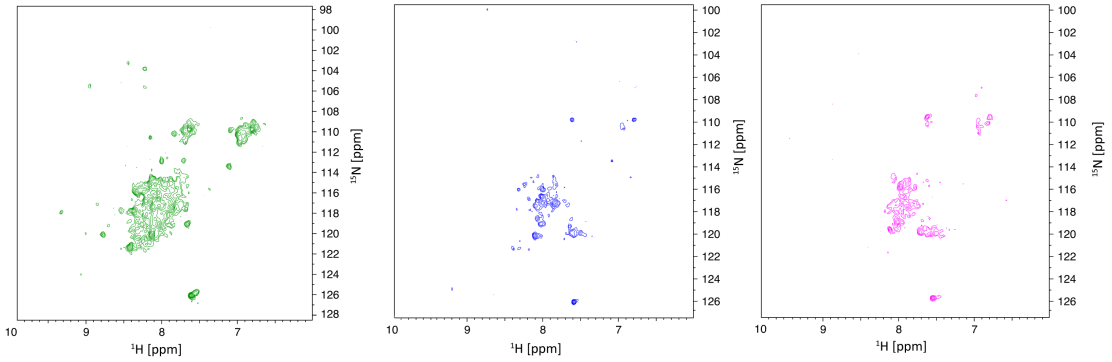

Figure IV.3 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-SA13 p7 in Fos-cholines. Fos-choline-10 was added at 125 mM to the sample from Figure IV.2, resulting in the spectrum in green (left). Another spectrum of the same sample after the pH was lowered to pH 6.0 is shown in blue (middle). Addition of 250 mM Fos-choline-14 resulted in the pink spectrum (right). Experiments were recorded at 37 °C on a 600 MHz ^{1}H-frequency Omega spectrometer.

In order to test whether the spectra were dependent on the refolding conditions, which have been shown to be important for other membrane proteins before[259], a urea-thin-film refold protocol was employed, which used Fos-choline-14 and SDS (sodium dodecyl sulfate)[294]. Protein and Fos-choline-14 were co-dissolved and dried to a thin film under a stream of N_2 gas. SDS was added to urea, which then was used to dissolve the thin film. Following dialysis into an NMR suitable buffer (20 mM sodium phosphate, pH 6.8) and addition of 5 % (v/v) D_2O, NMR experiments were carried out. The final sample had a HCV SA13 p7 concentration of 52.3 μM, and contained 250 mM Fos-choline-14 and 25 mM SDS (Figure IV.4 a). The spectrum represents a minor improvement over previous experiments, with a more clearly resolved central area, which is indicative of reduced aggregation. To see if the spectrum was temperature sensitive, another spectrum was recorded at 50 °C. A high temperature has been previously reported to yield high quality spectra of p7 in other detergent systems[239, 240].

The spectrum at the higher temperature deteriorated in the central area of the spectrum, and exhibited severe peak overlap. Although the doublet NH_2 peaks from asparagine became stronger and some additional glycine peaks were visible, a titration of 500 μM *N*N-DNJ induced no major peak shifts (Figure IV.4 b), implying that these conditions would not yield useful drug interaction data.

The next detergent tested was the lyso-lipid LMPG (1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-*rac*-glycerol), Affymetrix), which contains a lipid headgroup but only a single acyl chain (CMC of 0.05 mM)[295]. The sample contained 202 μM HCV SA13 p7,

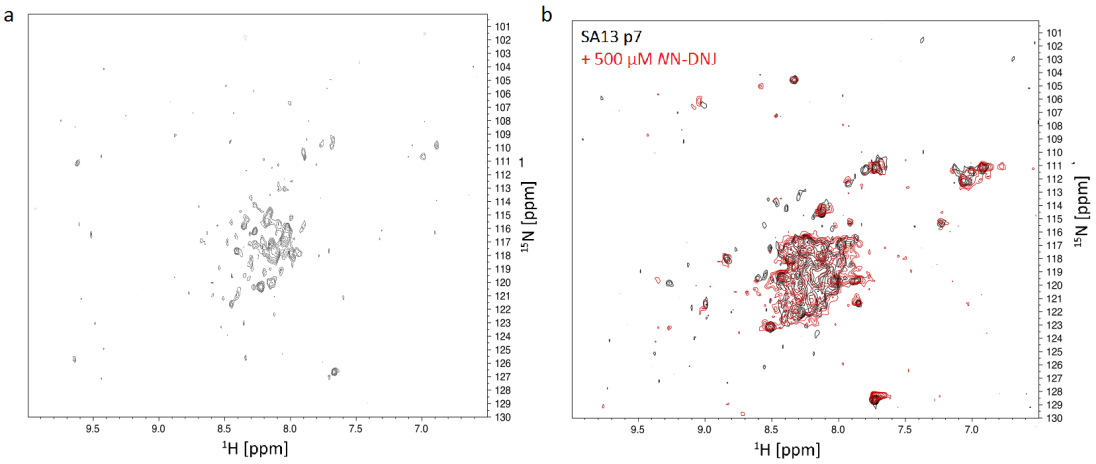

Figure IV.4 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-SA13 p7 in Fos-choline-14 and SDS. a) Urea refolded HCV SA13 p7 was thin filmed, and solubilised in 250 mM Fos-choline-14 and 25 mM SDS. The SA13 concentration was 52 μM, with 40 mM sodium phosphate and 5 % D_2O. Some individual peaks are visible and are better resolved than in the higher concentrated Fos-choline-12 sample (Figure IV.2). The spectrum in a) was recorded at 37 °C and the spectrum in b) (black) is the same sample at 50 °C, showing aggregation or oligomer formation. The red spectrum in overlay shows the addition of 500 μM *N*N-DNJ. The spectra were recorded on a 950 (left) or 750 MHz (right) ^{1}H-frequency Omega spectrometer.

40 mM LMPG, and 40 mM sodium phosphate at a pH of 6.5. In comparison to previous detergents, the LMPG spectrum of HCV SA13 p7 showed few promising characteristics. No glycines were visible, and the resolution of the central part of the spectrum was poor, with no major peaks visible besides one of the asparagine or glutamine sidechains, which is indicative of multiple protein conformations present and general sample heterogeneity (Figure IV.5).

IV.1.2 HCV J4 p7 and HCV SA13 p7 in Cyclofos-6

As the detergent systems tested so far yielded unsatisfactory results, alternatives were investigated. Cyclofos-6 (6-cyclohexyl-1-hexylphosphocholine, Affymetrix) contains a lipid-like headgroup, identical to the Fos-cholines, but contains a cyclohexane group at the end of the alkyl tail rather than a methyl group. It has a CMC of 2.68 mM and can be used to produce bicelles with the lipid DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine)[296]. A 10x CMC concentration was chosen as a starting point. HCV SA13 p7 was taken up as lyophilised powder in Cyclofos-6 prepared in sodium phosphate buffer containing TCEP, incubated at 37 °C for 10 minutes until the solution cleared, and transferred into a Shigemi microtube. The spectrum showed an improved overall peak

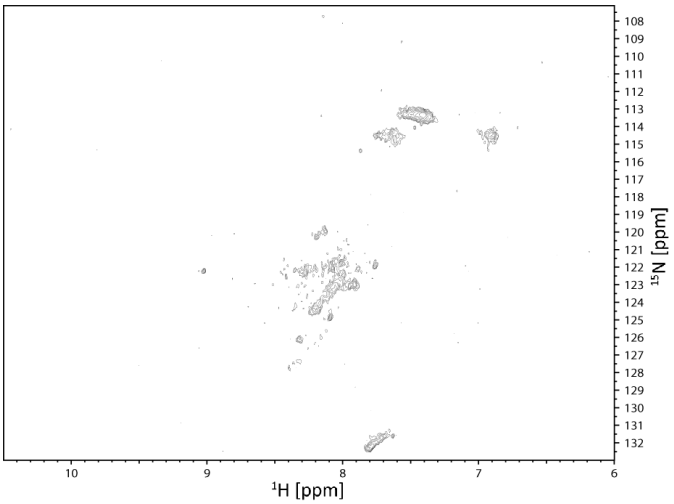

Figure IV.5 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-SA13 p7 in LMPG. LMPG is an ionic detergent and was tested at a concentration of 40 mM. The NMR sample was buffered at pH 6.5 and the spectrum collected at 37 °C. The concentration of the sample was 202 μM. The C-terminus is the sole peak visible at the bottom. Other peaks were resolved, but the spectrum shows a low number of peaks in total, which are also not well dispersed. The spectrum was recorded on a 950 MHz ^{1}H-frequency Omega spectrometer.

number, with at least two glycines, the expected asparagine and glutamine sidechains, and several other strong peaks in the central part visible (Figure IV.6). HCV SA13 p7 in Cyclofos-6 was screened for its dependence on temperature and pH in order to improve the previously obtained spectrum. It has been reported for other systems that an increase in temperature or change of pH[297] can have a positive effect on spectral quality. Three temperatures between 25 °C and 42 °C were tested. The most promising spectrum was obtained at 37 °C (Figure IV.7, left). Since any exchange of the backbone amide protons with water can be reduced by lowering pH, spectra were recorded with pHs in the range between 6.0 and 7.0. However, no extensive differences were visible between pH 6.0, pH 6.5 and pH 7.0 (Figure IV.7, right). In order to maximise solubilisation, the samples were incubated at 37 °C for 30 - 60 minutes on an Eppendorf shaker prior to NMR analysis. The resulting spectra in Figure IV.8 show the resulting improvement of overall spectral qualities.

IV.2 Drug titrations and protein dynamics

The optimised HCV SA13 p7 sample was used in a drug titration with the long-alkyl chain iminosugar *N*N-DNJ. *N*N-DNJ has been previously shown to inhibit HCV p7[223,

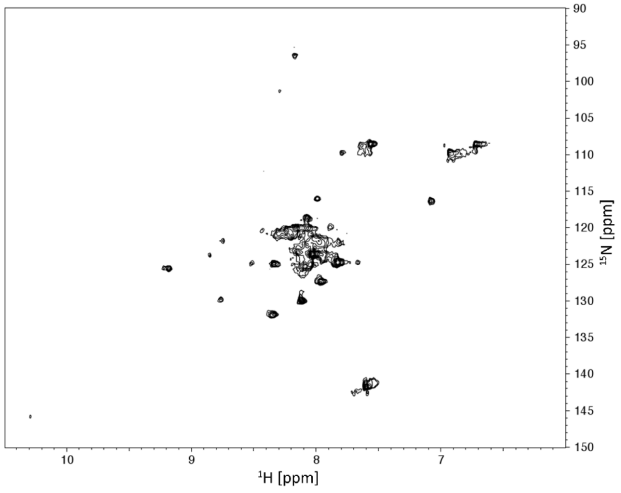

Figure IV.6 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-SA13 p7 p7 in Cyclofos-6. Cyclofos-6 was used at a concentration of 26.8 mM. The pH was adjusted to 6.5 and the spectrum was collected at 37 °C. The protein concentration was 210 μM. The spectrum is well dispersed compared with that of previously tested detergents. Besides the C-terminus, two glycine peaks are visible at the top. The two asparagine and glutamine NH_2 side chains are visible at 110 ^{15}N ppm and 7 and 7.7 ^{1}H ppm, respectively. Around the centre of the spectrum several single resonances are visible, with the highly overlapped cluster of peaks usually found in the centre of the spectrum lessened in its intensity. The spectrum was recorded on a 750 MHz ^{1}H-frequency Omega spectrometer.

243]. A concentrated *N*N-DNJ stock solution was prepared using the NMR buffer in Cyclofos-6 (26.8 mM Cyclofos-6, 40 mM sodium phosphate, pH 6.0). Titration steps ranged from 0.85 mM to 20.8 mM *N*N-DNJ. Lower concentration points were omitted based on the experience gained from the previous titration, which revealed no peak shifts at 500 μM. After each titration step, a 2D ^{1}H,^{15}N HMQC was recorded and analysed. A small number of peaks shifted their position and the majority of the peaks stayed in place, which is consistent with and points towards a specific interaction (Figure IV.9).

IV.2.1 HCV J4 p7 (C27S) in Cyclofos-6

In order to investigate if the spectra could be further improved in the established detergent system, a HCV J4 p7 cysteine mutant was used, in which cysteine 27 of HCV J4 p7 was substituted with a serine residue[288]. After expression was optimised (Chapter III), a 2D ^{1}H,^{15}N HMQC in the conditions previously determined to be optimal showed excellent spectral resolution and close to all expected peaks were visible (Figure IV.10).

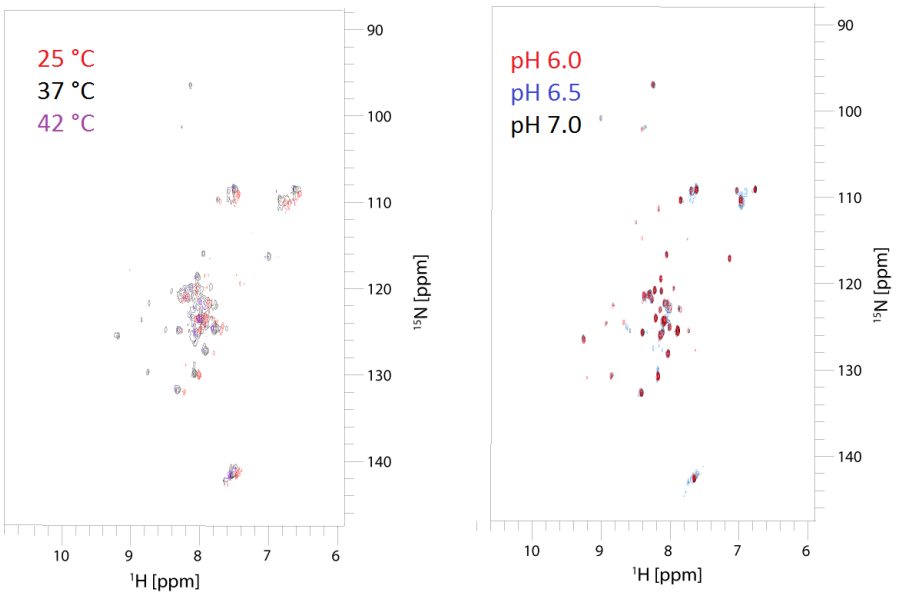

Figure IV.7 – Temperature and pH dependence of HCV ^{15}N-SA13 p7 in Cyclofos-6. The temperature and pH dependence of HCV SA13 p7 in Cyclofos-6 was analysed. On the left three temperatures were tested, 25 °C, 37 °C and 42 °C. HCV SA13 p7 concentration was 210 μM. The spectral quality at 25 °C was low, with fewer peaks visible than at 37 °C. Increasing the temperature to 42 °C decreased the number of visible peaks. On the right are shown spectral overlays of a 260 μM HCV SA13 p7 sample in Cyclofos-6 at different pH values, i.e. 6.0, 6.5 and 7.0. Increased incubation times after detergent addition to the lyophilised powder appeared to result in spectral improvements. However, no major differences are visible between the three pH values tested. The spectra were recorded on a 950 and 750 MHz ^{1}H-frequency Omega spectrometer, respectively.

A drug titration using *N*N-DNJ on the same sample was consistent with a specific interaction, i.e. a subset of peaks shifted their position upon addition of the iminosugar, whereas others were unperturbed and stayed in place (Figure IV.11).

IV.2.2 Assignment of HCV J4 p7 (C27S) in Cyclofos-6

In order to identify the binding site, HCV J4 p7 (C27S) was assigned using a 3D ^{15}N-edited NOESY-HSQC and a 3D ^{15}N-edited TOCSY-HSQC[259, 270]. A higher detergent concentration of 268 mM was used for the experiments as it aided HCV J4 p7 (C27S) solubilisation. A comparison of 2D ^{1}H,^{15}N HMQCs of HCV J4 p7 (C27S) in low (26.8 mM) and high (268 mM) detergent concentration indicated only small changes as a function of detergent concentration over that range (Figure IV.12). The through-space NOE crosspeaks in the 3D ^{15}N-edited NOESY-HSQC were used in conjunction with the through-bond crosspeaks in the 3D ^{15}N-edited TOCSY-HSQC to completely assign the

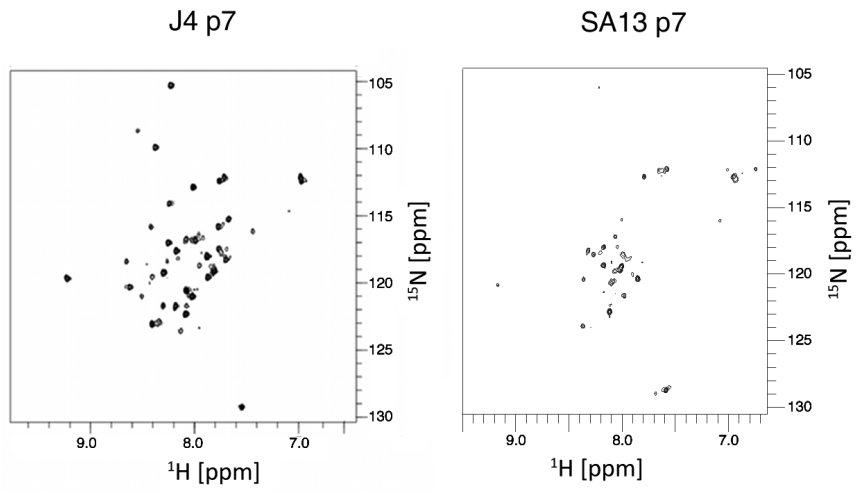

Figure IV.8 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-J4 and HCV ^{15}N-SA13 p7 p7 in Cyclofos-6. Spectra of the p7 constructs from HCV isolates SA13 (left) and J4 (right) in Cyclofos-6 at pH 6.5 and 37 °C have better spectral qualities compared to all previously tested conditions. The spectra were recorded on a 950 MHz ^{1}H-frequency Omega spectrometer.

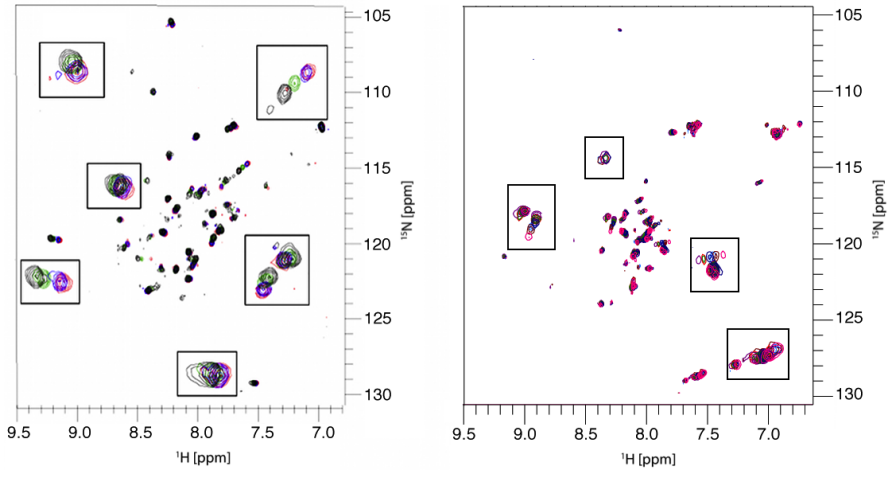

Figure IV.9 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-J4 and HCV ^{15}N-SA13 p7 in Cyclofos-6 with NN-DNJ. Left: The titration steps of *N*N-DNJ were 0.85 mM (black), 5.1 mM (green), 15.3 mM (blue), and 19.6 mM (red). Right: In the titration with HCV SA13 p7 the steps included 2.6 mM (purple), 5.2 mM (red), 7.8 mM (blue), 10.4 mM (light blue), 13 mM (maroon) and 20 mM (pink) *N*N-DNJ. Some local peak shifts are visible (see inserts) but the majority of peaks do not shift. Spectra were recorded on a 750 or 950 MHz ^{1}H-frequency Omega spectrometer.

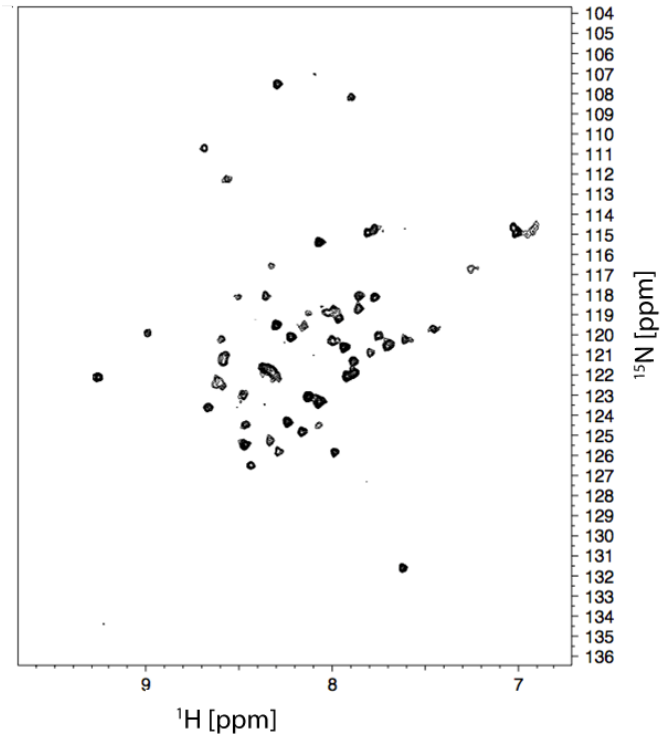

Figure IV.10 – 2D ^{1}H,^{15}N-HMQC of HCV ^{15}N-J4 (C27S) in Cyclofos-6. The protein concentration was 125 μM, using the same NMR buffer as before but without the addition of a reducing agent. The spectrum was collected at 37 °C and shows nearly all of the expected peaks, indicating that essentially the complete protein structural information could be obtained under these conditions. The spectrum was recorded on a 950 MHz ^{1}H-frequency Omega spectrometer.

backbone amides of HCV J4 p7 (C27S), as well as a large number of sidechain proton resonances (Figure IV.13). The spectra were analysed using CARA[298]. The 3D ^{15}N-edited TOCSY-HSQC was first used to identify the probable amino acid types, since it contained crosspeaks for most hydrogens in the residue spin system, which could be compared with characteristic chemical shift values[299]. The C-terminal alanine was identified first, followed by the preceding residues up to Pro59. Identified next were glycine residues, which have distinct amide chemical shifts; they were assigned based on the characteristic chemical shifts of the inter-residue NOEs. The rest of the assignments were made from inter-residue connectivities identified in the NOESY spectrum. Particularly useful were the large number of alanine residues, which have a distinctive, intense TOCSY crosspeak at the H$_\beta$ resonance at ∼1.4 ppm. The assignment using the NOESY and TOCSY was very successful with only a small number of tentative assignments, which will be confirmed using triple-resonance approaches (Figure IV.14).

Chemical shift mapping of the *N*N-NDJ titration shown in Figure IV.15 indicated that

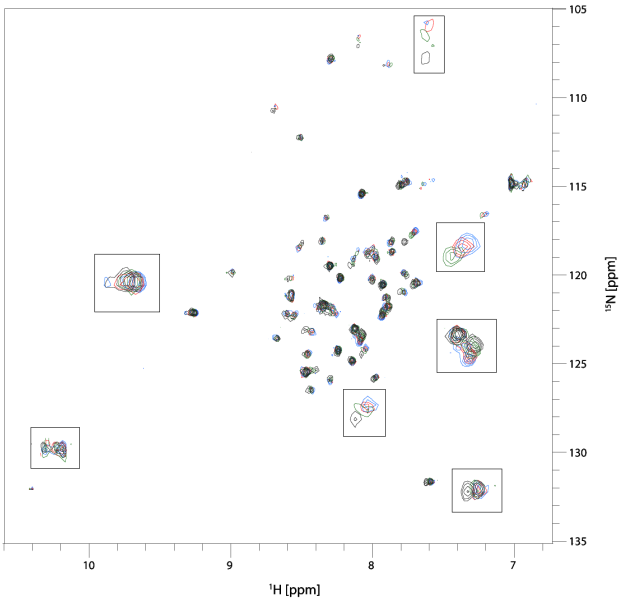

Figure IV.11 – Titration of HCV ^{15}N-J4 p7 (C27S) and NN-DNJ. The titration steps of *N*N-DNJ were 5 mM (green), 10 mM (red), and 20 mM (blue). The Cyclofos-6 concentration was 26.8 mM and the pH was 6.5. A subset of peaks shifted (see inserts) but the majority of peaks were unperturbed. The bottom left insert shows the tryptophan side chains, which do not appear to be affected. The spectra were recorded on a 750 MHz ^{1}H-frequency Omega spectrometer..

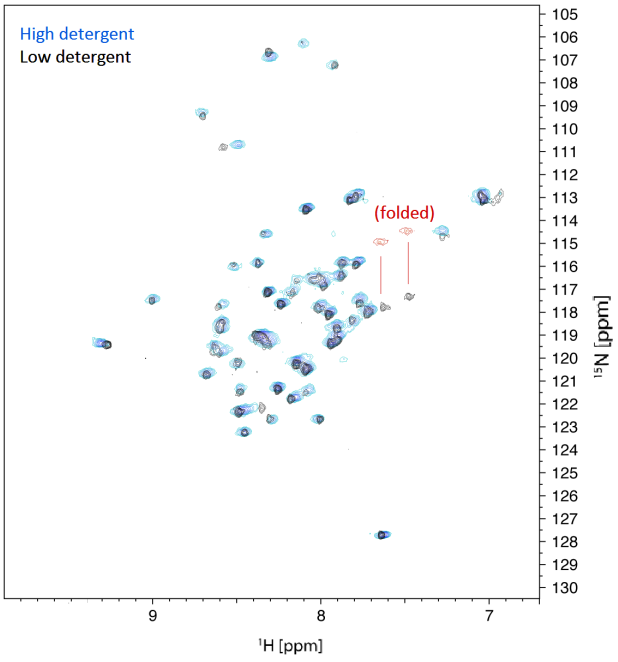

Figure IV.12 – Detergent concentration dependency of HCV ^{15}N-J4 p7 (C27S). Spectral overlay of 2D ^{1}H,^{15}N-HMQCs of HCV J4 p7 (C27S) at two concentrations of Cyclofos-6, i.e 26.8 mM and 268 mM. The 2D ^{1}H,^{15}N-HMQC overlay shows that the two spectra are very similar. The spectra were recorded on a 950 MHz ^{1}H-frequency Omega spectrometer.

the largest changes occurred at the N-terminal region in residues 7 and 8, in the region around His17, Lys33, between residues 43 - 45, and at the C-terminal end (Figure IV.16 c). A titration that was carried out with low drug concentrations in the range of 10 μM to 1 mM induced no shifts. Titration steps shown were between 0.75 mM, 2.5 mM, 20 mM and 125 mM *N*N-DNJ (Figure IV.16). Peak shifts are plotted against the sequence, which indicates three interaction sites (Figure IV.16 b). The higher detergent and inhibitor concentration used lead to basically same results as in lower detergent and inhibitor concentrations (Figure IV.16 c). The changes upon *N*N-DNJ titration were saturable, consistent with a specific interaction.

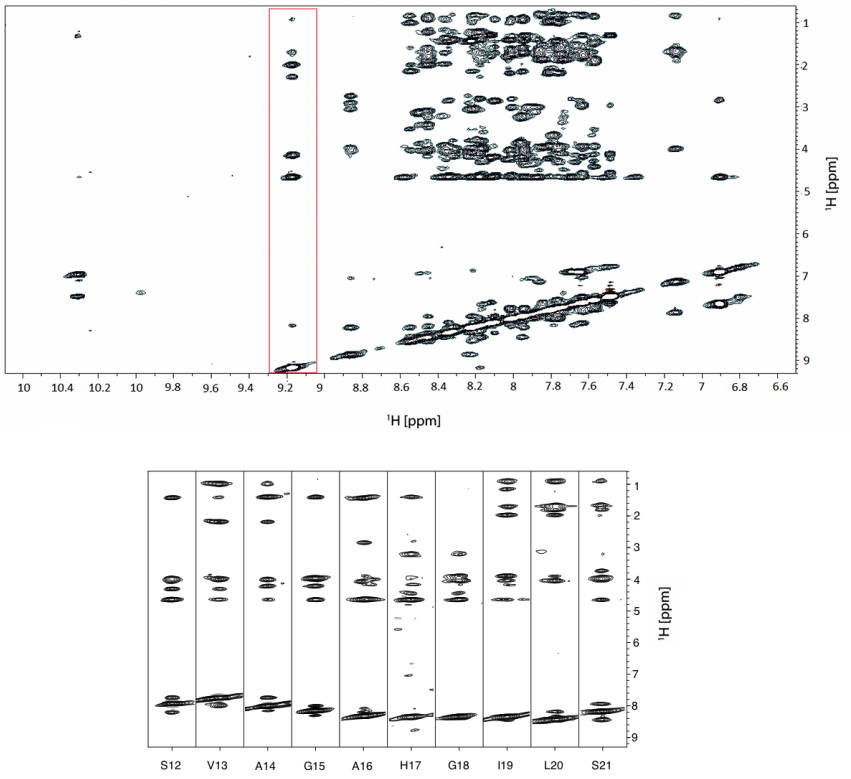

Figure IV.13 – 3D ^{15}N-edited NOESY-HSQC of HCV ^{15}N-J4 p7 (C27S). Top: Shown is the ^{1}H-^{1}H plane projection of the NOESY experiment, showing crosspeaks used for assignment. At the top of the spectrum the NOEs to aliphatic hydrogens are seen. Amide to amide and amide to aromatic protons can be observed near the diagonal at the bottom of the spectrum. The protein concentration was 384 μM. In each strip, (e.g. red strip), NOEs from one NH group to all other hydrogen atoms are spatially close. Bottom: Strip plots of residues 12 to 21 are shown as an example how the intraresidue and sequential NOEs were used for backbone assignment. The spectra were recorded on a 950 MHz ^{1}H-frequency Omega spectrometer.

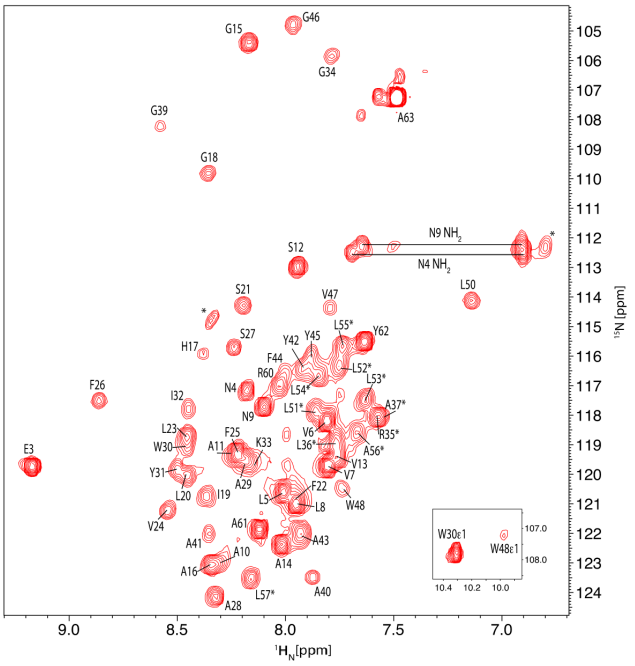

Figure IV.14 – Assignment of HCV J4 p7 (C27S) in Cyclofos-6. 2D ^{1}H,^{15}N-HMQC showing crosspeak assignments of HCV J4 p7 (C27S). Amino acids are listed in single letter code with their residue position. The C-terminal alanine residue is folded. The insert shows the tryptophan side chain residues. Asterisks indicate tentative peak assignments, which will be confirmed by triple resonance experiments. The spectrum was recorded on a 600 MHz ^{1}H-frequency Bruker spectrometer.

IV.2.3 Protein dynamics and water accessibility of HCV J4 p7 (C27S) in Cyclofos-6

To characterise the dynamic properties of HCV J4 p7 (C27S), ^{15}N-R_2 and R_2, and ^{1}H-^{15}N hetNOEs were determined. In addition, the relative magnitudes of proton exchange between the protein amides and water were measured using a CLEANEX (CLEAN chemical exchange) experiment[300]. The amide-water proton exchange rate provides insight into flexibility and hydrogen bonding of the backbone amides. The calculated values for ^{15}N-R_1, and R_2, the hetNOEs and amide-water proton exchange are shown in Figure IV.17. R_1 values were mostly uniform with a decrease between residues 20 to 37 and 45 to 56 (Figure IV.17 top). The R_2 values were also similar, with a decreased region around the N-terminus (residues up to 16) and again at the C- terminal region. Gly34 had a very large R_2 value in comparison to the rest of the sequence, possibly indicating chemical exchange (Figure IV.17 middle top). HetNOE data indicated flexibility in the C- and N-terminal region, but also around His17 (residues 10 to 20) and in the region of residues 35 to 45 (Figure IV.17 middle bottom). CLEANEX experiments indicated

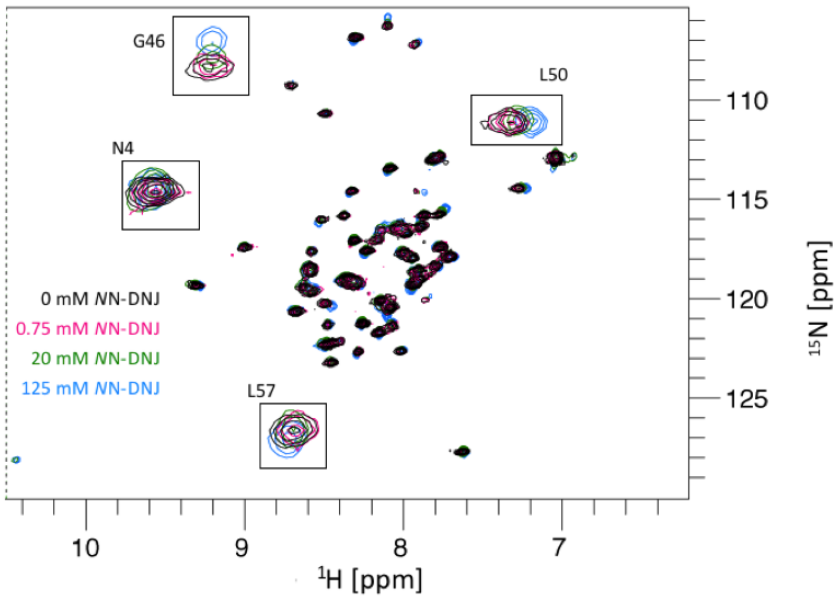

Figure IV.15 – Titration HCV J4 p7 (C27S) and NN-DNJ in Cyclofos-6. The titration steps of *N*N-DNJ were 0.75 mM (pink), 20 mM (green) and 125 mM (blue). The Cyclofos-6 concentration was 268 mM and the pH was 6.5. A subset of peaks shift (see inserts) but the majority of peaks are unperturbed. The higher concentration needed to induce shifts in comparison to previous titrations is due to the higher detergent concentration present, with *N*N-DNJ likely to be sequestered in detergent micelles that do not contain p7. The spectra were recorded on 950 and 750 MHz ^{1}H-frequency Omega spectrometers.

that the regions with strongest amide-water proton exchange were at around His17 and Gly39 (Figure IV.17 bottom).

In summary a region of increased flexibility is observed from residues 38 to 44, with a significant H_2O exchange for amides in position 16 to 18, and 39 to 41. This suggests a consensus loop region of HCV J4 p7 (C27S) around residues 33 to 40. A secondary structure prediction using TALOS+[301], which uses chemical shifts (i.e. HN, N, HA) for empirical prediction of phi and psi backbone torsion angles[302] was obtained for HCV ^{15}N-J4 p7 (C27S) and showed generalised order parameter values (S^2) that were lower in the region of His17 and around Gly49 (Figure IV.18 a). Modelfree analysis of R_1, R_2 and hetNOE values were not used to calculate S^2 as the former were only measured in one field, which would lead to unreliable values. The artificial neural network (ANN)-predicted[303] secondary structure is shown below, indicating flexibility, i.e. less likelihood of a helical fold, again around His17, from Leu36 to Ala40, Val47 to Pro49 and from Ala56 to Pro59 (Figure IV.18 b). This suggests that HCV J4 p7 (C27S) in Cyclofos-6 has three to four helices (Figure IV.18 c).

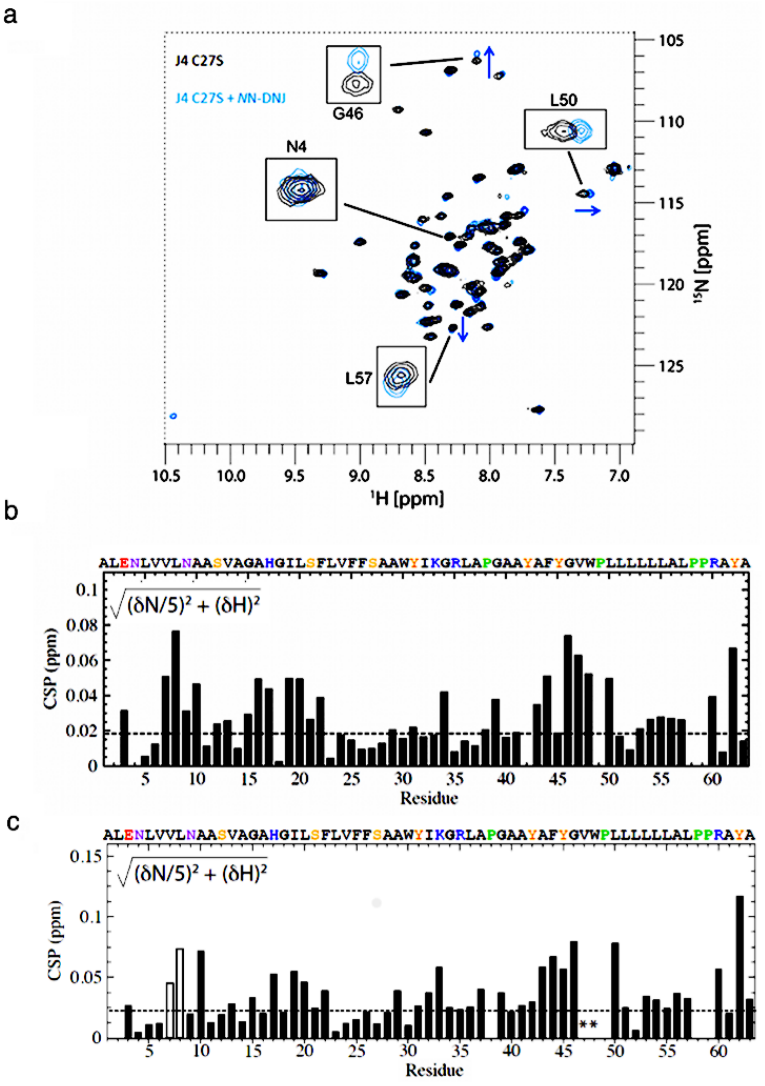

Figure IV.16 – Chemical shift mapping (CSP) of HCV J4 p7 (C27S) and NN-DNJ. a) 2D ^{1}H,^{15}N-HMQC in black shows HCV J4 p7 (C27S) without *N*N-DNJ added. The spectrum in blue shows the addition of 125 mM *N*N-DNJ. Inserts show peaks shift of selected peaks. Note that some peaks do not shift (e.g. N4). b) For chemical shift mapping the maximal titration point of *N*N-DNJ was used. The dashed horizontal line is the standard deviation for all changes. Three interaction areas were identified, i.e. around L8, H17, and Y45. c) The data used here are HCV J4 p7 (C27S) without inhibitor present and 20 mM *N*N-DNJ from the titration in Figure IV.11. The detergent concentration in this experiment was 10x less compared to the titration seen in a). The dashed horizontal line is the standard deviation for all changes. Asterisks and open bars indicate weak peak intensities or overlap, respectively. Similar interactions albeit on a weaker level can be observed. The amino acids at each position are shown above the plot. Polar residues and prolines are shown in different colours.

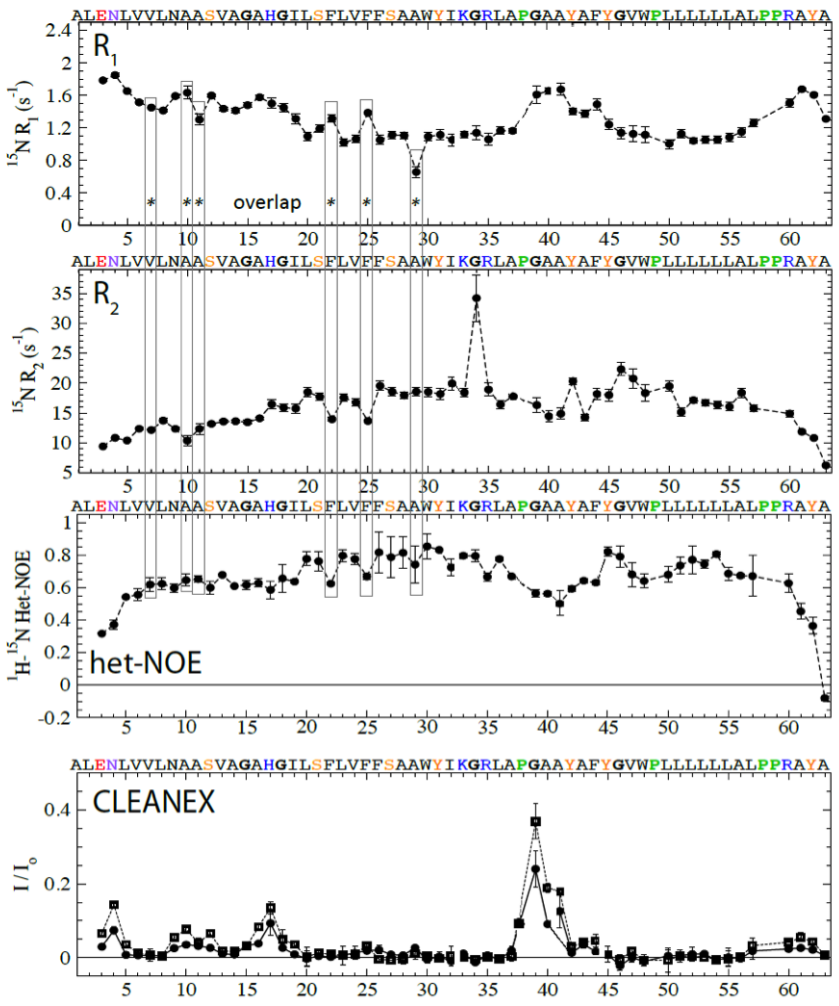

Figure IV.17 – Dynamic properties and water accessibility of HCV J4 p7 (C27S) in Cyclofos-6. An ^{15}N-labelled HCV J4 p7 (C27S) sample was used to measure ^{15}N-R_1 and R_2, the ^{1}H- ^{15}N heteronuclear NOEs, and amide proton exchange (CLEANEX). Residues for which the crosspeaks overlapped - with their data therefore potentially skewed - are indicated with a box and asterisk. Mixing times in the CLEANEX were 20 ms and 50 ms. The experiments were recorded on a 600 MHz ^{1}H-frequency Bruker spectrometer.

IV.2.4 HCV J4 p7 (C27S) structure calculation using Rosetta

Chemical shifts were used to generate a *de novo* Rosetta model[304] of HCV J4 p7 (C27S) (Figure IV.19 left). The Rosetta results confirm the previously predicted tri-helical segmentation. The dibasic motif K33/R35 is part of helix 1 and no longer in the suggested consensus loop. Plotting of *N*N-DNJ interaction data and CLEANEX accessibility reveals where residues are positioned on the predicted structure (Figure IV.19 middle and right). The Rosetta HCV J4 p7 (C27S) structure indicates that its C-terminal part would protrude from the detergent micelle, although an unspecific interaction with detergent cannot be ruled out.

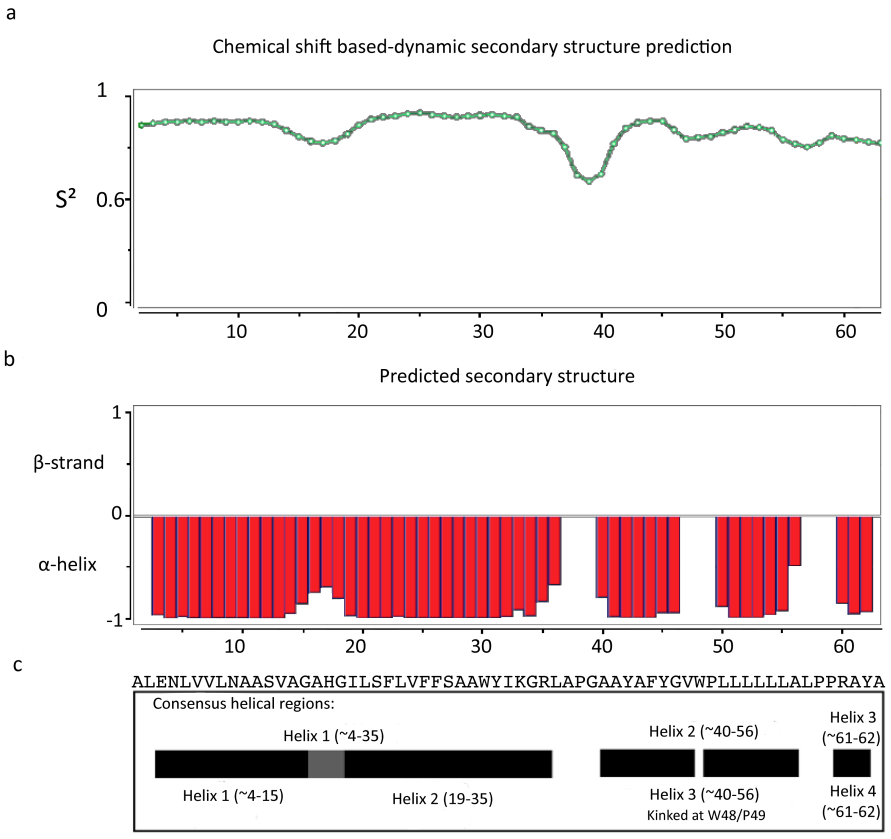

Figure IV.18 – Determination of consensus helical regions of HCV J4 p7 (C27S) in Cyclofos-6. a) Using TALOS+, ^{15}N-chemical shift based predictions were performed (i.e. HN, N, HA). An S^2 value of 1 predicts complete rigid structures and values of 0 indicate complete flexibility. Areas of flexibility in HCV J4 p7 (C27S) were predicted to be around H17 and G39. b) Shows the ANN-prediction of secondary structures, which indicate that HCV J4 p7 (C27S) has three to four helices in Cyclofos-6, with unstructured loop areas centred on H17, P38, W48 and P59. c) Consensus secondary structure organisation for HCV J4 p7 (C27S) in Cyclofos-6 is shown on top (Helix 1 black and grey rectangles), which is based on chemical shift information and the amide-water proton exchange (Figure IV.17).

IV.2.5 HCV EUH1480 p7 (mt5) in Fos-Choline-12

The p7 genotype 5 isolate EUH14580 (mt5) was investigated next since it was used to solve the first experimental high resolution structure of an oligomeric p7 channel[242]. The expression and purification was described in Chapter III, and the sample was prepared using the previously described guanidine refolding protocol[242].

A comparison of a collected 2D ^{1}H,^{15}N-HMQC with a previously published 2D ^{1}H,^{15}N-TROSY of HCV EUH1480 p7 (mt5) [242] shows that the two spectra are highly similar, indicating that the expressed and purified HCV HCV EUH1480 p7 (mt5) was successfully refolded and that it adopted a comparable if not identical conformation(Figure IV.20).

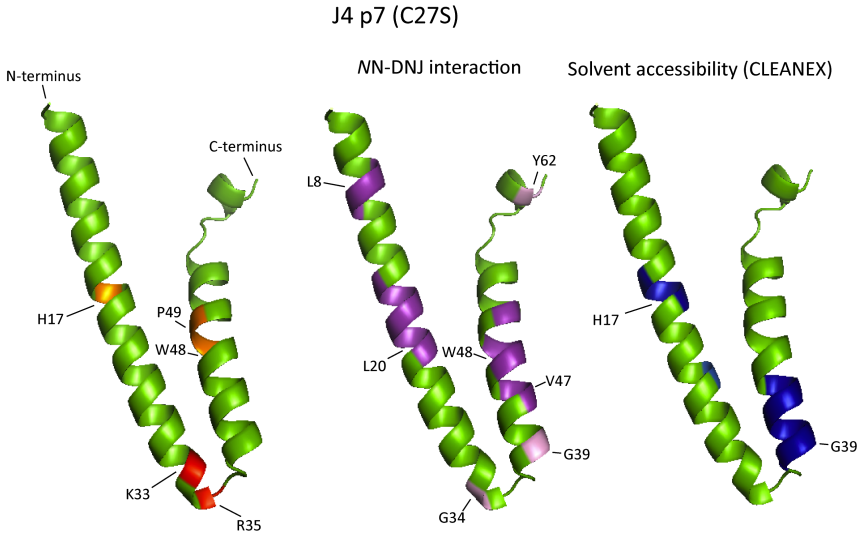

Figure IV.19 – Rosetta model of HCV J4 p7 (C27S) in Cyclofos-6. Left: The best fit using chemical shifts for *de novo* modelling is shown. Three helices are predicted. Residues highlighted are H17, K33, R35, W48 and P49. H17 is in an area of increased flexibility, K33 and R35 are the dibasic motif and W48 and P49 are an area of helix kinks. Middle: *N*N-DNJ binding sites clustering around L8, G17 and W48 are shown in magenta. Single residues perturbed outside of these clusters are highlighted in orange (i.e. G34, G39 and Y62). Right: Solvent accessible residues are shown in blue, clustering around H17 and G39. F25 is slightly elevated in comparison to baseline. The Rosetta calculations were performed by J.Brady, Membrane Protein Interactions Lab, Biochemistry Department, Oxford University. The figures were prepared in MacPyMOL (Delano Scientific, Palo Alto, CA).

To further characterise HCV EUH1480 p7 (mt5) and explore whether it exhibited similar conformational properties to HCV J4 p7 (C27S), it was explored whether HCV EUH1480 p7 (mt5) could be studied at higher temperatures closer to the 37 °C at which the HCV J4 p7 (C27S) was studied. 2D ^{1}H,^{15}N-HMQC spectra of HCV EUH1480 p7 (mt5) were collected at three temperatures: 30 °C, 34 °C and 37 °C. An overlay of all three spectra indicates that the sample is largely insensitive to temperature changes in the range tested (Figure IV.21).

IV.2.6 Protein dynamics and water accessibility of HCV EUH1480 p7 (mt5) in Fos-choline-12

To characterise the protein dynamics of HCV EUH1480 p7 (mt5) in Fos-choline-12, hetNOEs and amide-water proton exchange using CLEANEX experiments were measured at 37 °C. The hetNOE data was overall very similar to that of HCV J4 p7 (C27S), showing a flexibility in the N- and C-terminal regions, around His17 (residues 9 to 20)

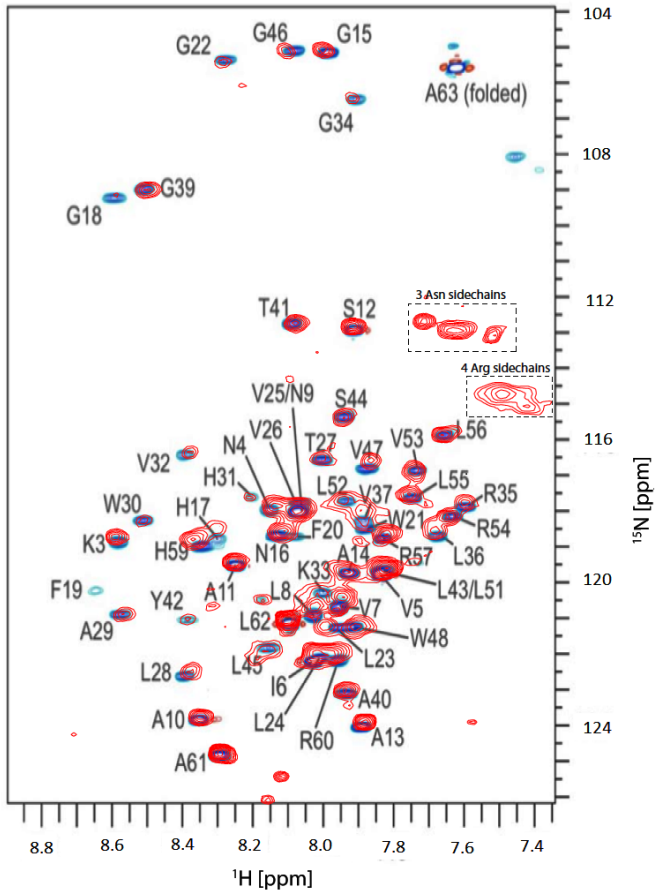

Figure IV.20 – Overlay of 2D ^{1}H,^{15}-HMQC with a 2D ^{1}H,^{15}N-TROSY of HCV ^{15}N-EUH1480 p7 (mt5)[242]. The expressed and refolded HCV EUH1480 p7 (mt5) in Fos-Choline-12 (red) shows a high degree of similarity with the published 2D ^{1}H,^{15}N-TROSY (blue) from Ouyang *et al.*. The sample concentration was 170 μM. The red spectrum was recorded on a 600 ^{1}H-frequency Bruker spectrometer.

and in the region of residues 35 to 45 (Figure IV.22 top). The CLEANEX data indicated regions of amide-water proton exchange at His17, Phe20, Gly39 and Thr41 (Figure IV.22 bottom). The exchange at Phe20 is likely due to overlap with Ala16. A region of increased flexibility was identified between residues 35 to 44, with a significant amide-water proton exchange for positions 17 and 39 to 42. This suggested a consensus loop region of HCV EUH1480 p7 (mt5) around residues ∼39 to 43.

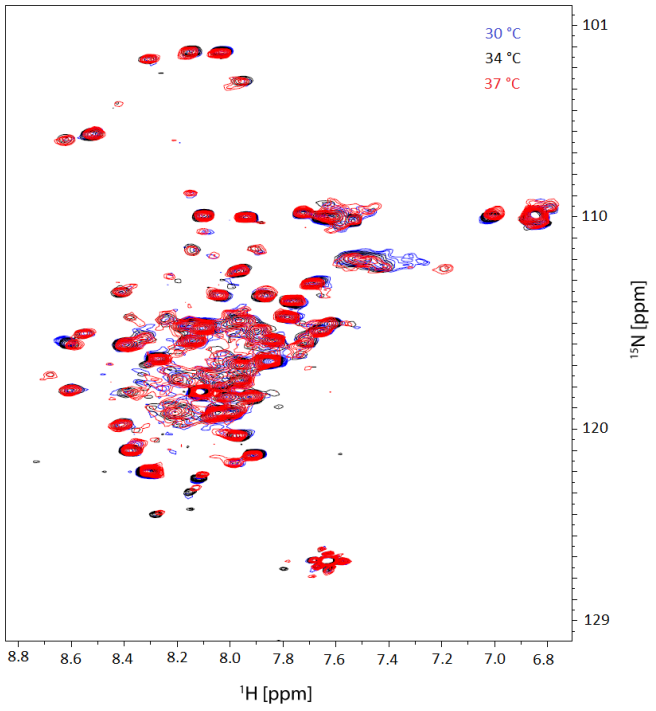

Figure IV.21 – Temperature sensitivity of HCV ^{15}N-EUH1480 p7 (mt5) in Fos-choline-12. Three temperatures were tested, 30 °C (blue), 34 °C (black) and 37 °C. The 2D ^{1}H,^{15}N-HMQC overlay shows that there were no major changes between 30 °C and 37 °C. The spectra were recorded on a 600 ^{1}H-frequency Bruker spectrometer, and referencing was adjusted to minimise the non-specific temperature dependent amide chemical shifts in order to more clearly observe any conformational changes with temperature.

IV.2.7 Comparison of HCV J4 p7 (C27S) in Cyclofos-6 and HCV EUH1480 p7 (mt5) in Fos-choline-12

The protein dynamics data obtained was used to compare both studied p7 isolates in their respective detergent systems. The hetNOE data showed that both isolates look very similar and have the same overall hetNOE profile (Figure IV.23). The CLEANEX experiments both showed similarities for the water accessibility, although HCV EUH1480 p7 (mt5) showed higher values in the His17 region and also at the respective N- and C-termini (Figure IV.23). This suggested that both isolates have similar dynamic properties and water accessibility.

The intrinsic hydrogen exchange rates were predicted for both sequences [305]. For HCV J4 p7 (C27S), the predicted and observed non-transmembrane amide protons were closely correlated. In contrast, several regions within the transmembrane domains were predicted to have intrinsic exchange but show little or no exchange in the CLEANEX experiments. The results for HCV EUH1480 p7 (mt5) were similar to that of HCV J4

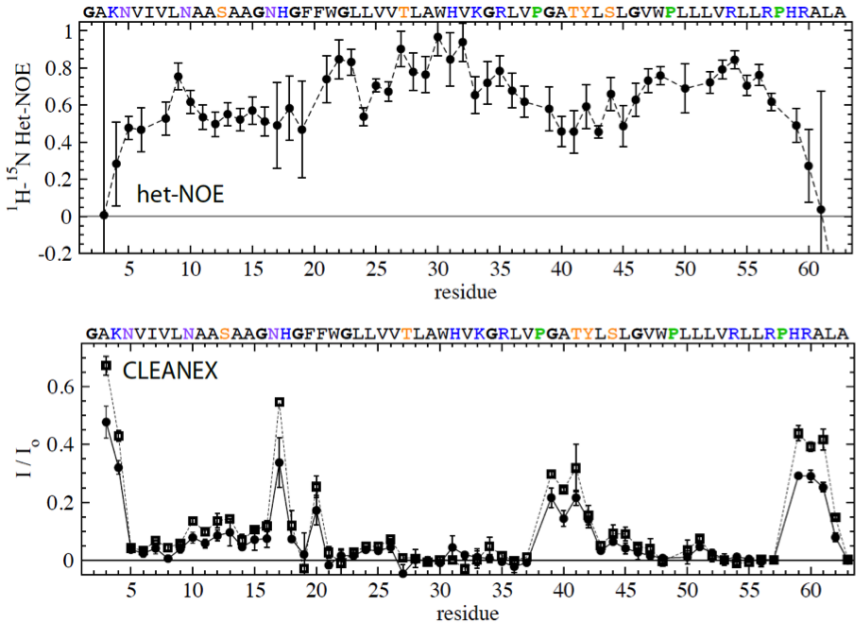

Figure IV.22 – Dynamic properties and water accessibility of EUH1480 p7 (mt5) in Fos-choline-12. An ^{15}N-labelled HCV EUH1480 p7 (mt5) sample was used to determine the hetNOEs (top) and amide proton exchange (CLEANEX) (bottom), revealing the flexible regions. Mixing time in the CLEANEX was 20 ms and 50 ms and sample temperature was 37 °C. The experiments were recorded on a 600 MHz ^{1}H-frequency Bruker spectrometer.

p7 (C27S). Notably, the C-terminal region of HCV EUH1480 p7 (mt5) that was determined to be within the buried region of the micelle in the hexameric structure[242], was predicted to have intrinsic exchange with water(Figure IV.24), which was confirmed in the CLEANEX experiments (Figure IV.22).

IV.2.8 Molecular weight estimation of HCV J4 p7 (C27S) in Cyclofos-6 and HCV HCV EUH1480 p7 (mt5) in Fos-choline-12

Preliminary calculations of the molecular weight of the two studied p7 isolates in their detergent micelles were performed[306]. The T_1 and T_2 values were calculated from ratios of intensities in 1D experiments, assuming a change in H_2O viscosity as a function of temperature. In the case of HCV J4 p7 (C27S) in Cyclofos-6, T_1 was found to be 685 ms and T_2 was 72 ms. For HCV EUH1480 p7 (mt5) in Fos-choline-12, the values determined for T_1 and T_2 were 741 ms and 69 ms, respectively. The estimated rotational correlation time τ_c was 9.4 ns for HCV J4 p7 (C27S) and 10 ns for EUH1480 p7 (mt5). This gave an apparent molecular weight for HCV J4 p7 (C27S) in Cyclofos-6 micelles of

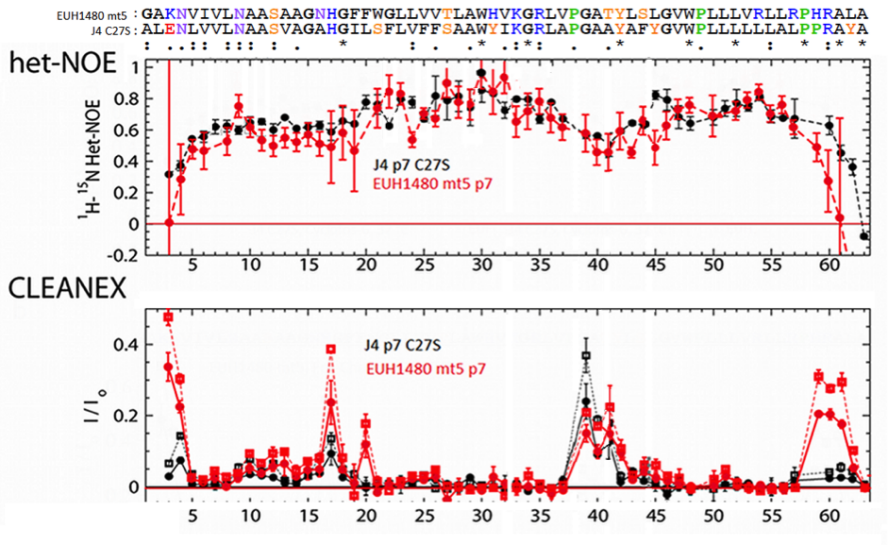

Figure IV.23 – Comparison of HCV ^{15}N-EUH1480 p7 (mt5) in Fos-choline-12 and HCV ^{15}N-J4 p7 (C27S) in Cyclofos-6. a) Previously discussed hetNOE data is shown in an overlay for both isolates, HCV J4 p7 (C27S) in black and HCV EUH1480 p7 (mt5) in red. b) The CLEANEX data for both isolates is shown, HCV EUH1480 p7 (mt5) (red) and HCV J4 p7 (C27S) (black), highlighting in both cases their similarity.

22.7 kDa and for HCV EUH1480 p7 (mt5) in Fos-choline-12 micelles of 20.8 kDa.

1D ^{15}N,^{1}H-TRACT (TROSY for rotational correlation times) experiments were also used for the measurement of the rotational correlation times (Figure IV.25). The resulting approximate molecular weight of HCV J4 p7 (C27S) in Cyclofos-6 micelles is around 20 kDa and 4 kDa for HCV EUH1480 p7 (mt5) in Fos-Choline-12 micelles. The molecular weight of 4 kDa for HCV EUH1480 p7 (mt5) is below the expected minimum value of a monomer in micelles. In cases of flexible peptides, the 1D TRACT approach can yield apparent molecular weights that are significantly lower than the overall rotational correlation time for an equivalent sphere representing the protein would indicate[307]. The inserted mutations in HCV EUH1480 p7 (mt5) make it one of the most hydrophilic p7 sequence studied and it must contain extensive flexible polypeptide segments (Figure III.14).

To further characterise the molecular weights, SEC (size-exclusion chromatography) was used to determine if the high-detergent NMR conditions differ from the low-detergent EM conditions previously described to preform single particle EM (i.e. 200 mM vs. 3 mM Fos-Choline-12)[242]. A high-detergent containing sample of HCV EUH1480 p7 (mt5) was prepared and run into the low-detergent buffer using SEC. The resulting

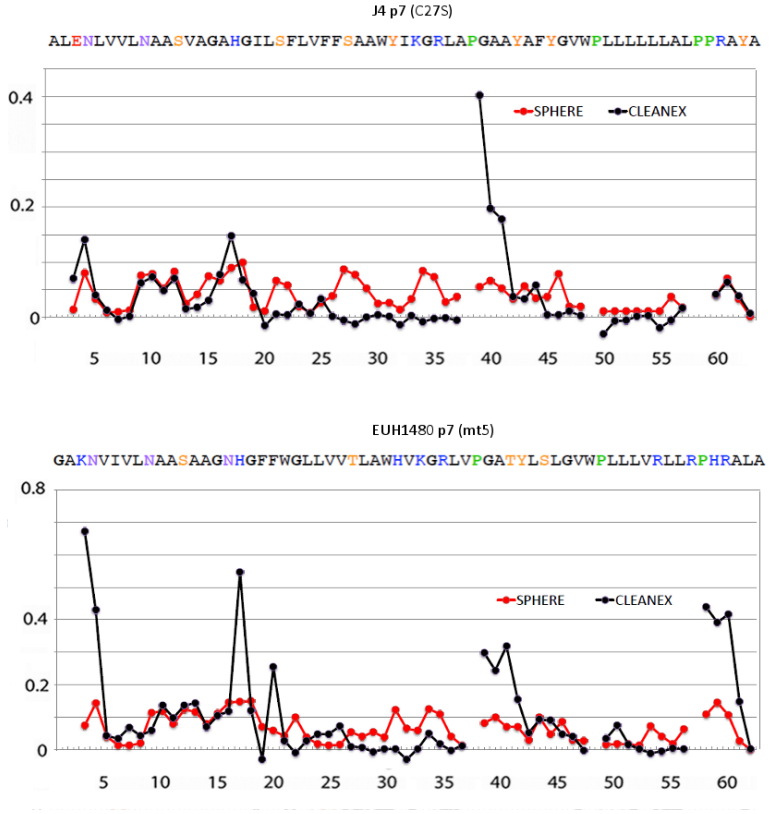

Figure IV.24 – Intrinsic hydrogen exchange rates of HCV ^{15}N-EUH1480 p7 (mt5) in Fos-choline-12 and HCV ^{15}N-J4 p7 (C27S) in Cyclofos-6. Top: The intrinsic hydrogen exchange rate for HCV J4 p7 (C27S) was calculated (SPHERE, red) and plotted versus the measured CLEANEX results (black). Bottom: The CLEANEX results for HCV EUH1480 p7 (mt5) (black) are plotted versus the sequence intrinsic hydrogen exchange rate as calculated by SPHERE (red)[305].

chromatogram showed a main peak corresponding to a monomer (Figure IV.26 a). Lowering to even less detergent in the same experiments (i.e. 1 mM Fos-Choline-12) showed no change in oligomerisation state and also a monomer (data not shown). SEC-MALS (SEC - Multi-Angle Static Light Scattering) analysis of HCV EUH1480 p7 (mt5) in 3 mM Fos-Choline-12 suggests a protein size of around 7 kDa - corresponding to a p7 monomer - with a micelle of about 39 kDa (Figure IV.26 b).

IV.3 Discussion

The work described in this chapter established sample conditions at physiologically relevant pH and temperature to analyse HCV p7 using solution-state NMR. Previous NMR

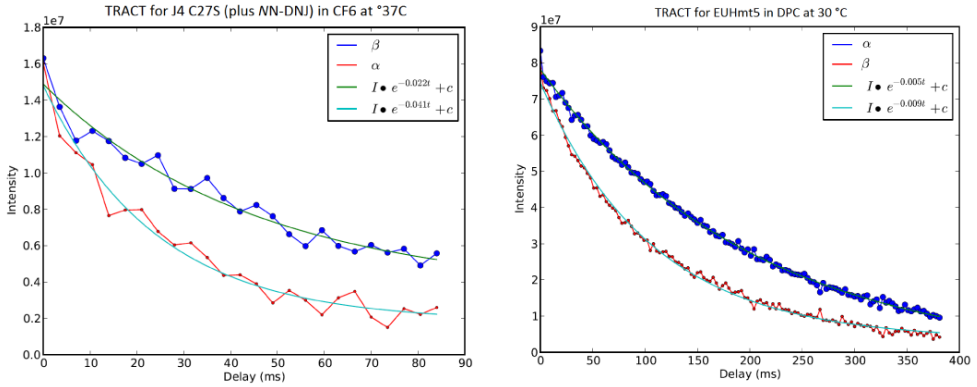

Figure IV.25 – 1D ^{1}H,^{15}N-TRACT of HCV ^{15}N-EUH1480 p7 (mt5) in Fos-choline-12 and HCV ^{15}N-J4 p7 (C27S) in Cyclofos-6. 1D ^{1}H ^{15}N-TRACT showing the decay of the relative intensity of the ^{1}H NMR signal, due to ^{15}N relaxation of HCV J4 p7 (C27S) in Cyclofos-6 and HCV EUH1480 p7 (mt5) in Fos-Choline-12. The upper and lower curves correspond to the slowly relaxing α-spin state of ^{15}N and the more rapidly relaxing β-spin state, respectively. Exponential fits (solid blue and green lines) yielded R_α and R_β values as indicated. For HCV J4 p7 (C27S) relaxation rates are $R_\alpha = 22$ Hz and $R_\beta = 41$ Hz. For EUH1480 p7 (mt5) relaxation rates are $R_\alpha = 5$ Hz and $R_\beta = 9$ Hz, respectively. For HCV EUH1480 p7 (mt5) in Fos-choline-12, the effective rotational correlation time (τc) was calculated using equations as previously described[307] and was 2.7 ns for the protein-detergent complex. Such a low (τc) value implies a high number of flexible residues and is in favour of a monomeric state of HCV EUH1480 p7 (mt5) under these conditions. The approximate molecular weight for the protein-detergent complex is 4 kDa. An increase in flexibility can skew the result below the minimal expected molecular weight. The complex of HCV J4 p7 (C27S) in Cyclofos-6 has an approximate molecular weight of 20 kDa.The experiments were recorded on a 500 MHz ^{1}H-frequency Bruker spectrometer.

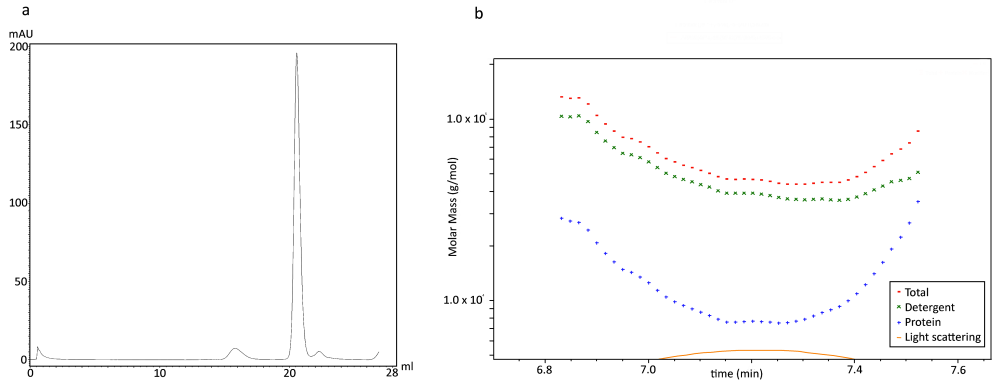

Figure IV.26 – Molecular weight determination of HCV EUH1480 p7 (mt5) in Fos-Choline-12 micelles. Left: SEC-MALS of HCV EUH1480 p7 (mt5) from high detergent NMR conditions (i.e. 200 mM) into low detergent EM conditions (i.e. 3 mM). The sample was solubilised as described[242]. The SEC running buffer contained 3 mM Fos-Choline-12, which corresponds to the buffer used to solve the EM structures byOuyang *et al.*. The chromatogram at 280 nM shows that the protein in the low detergent buffer is monomeric. Right: SEC-MALS shows the protein conjugate data vs. time of the 3 mM Fos-Choline-12 sample. The apparent size of the protein is approximately 7 kDa (i.e. 7745.89 Da, red crosses) and the micelle size is about 39 kDa (i.e. 39900 Da, green crosses). The combined size is approximately 47 kDa (red points).

studies looking at p7 structure employed either non-physiological pH and/or temperatures. To reach a physiological pH was paramount in our study, in order to acquire relevant inhibitor interaction data.

The pH values encountered in the human body vary, i.e. the lysosomal lumen has a low pH of 5.5, whereas the ER lumen has a near neutral pH of 7.2 and human blood pH 7.4[308]. Inhibitor pKa values define the point at which 50 % of a drug is ionised. Values can vary greatly, e.g. *N*N-DNJ has an approximate pKa of 7.2, amantadine 10.1 and rimantadine 9.0. At neutral pH, it can be predicted that ∼50 % of *N*N-DNJ is protonated, and ∼50 % is in the form of free base. At a pH of 4.0 practically all of it will be protonated. The pKa is not a predictor if the protonated or free-base of the compound are the pharmacologically active form[309], which makes it important to match the pH in the experiment to the pH where the compound is active. The exact pKa in the detergent surroundings used in the experiments is not known but might shift the pKa. The constraint of staying in a limited pH and temperature range made condition screening difficult, as pH and temperature can have a profound effect on NMR spectral quality.

The study of membrane proteins also requires membrane mimetics. Detergents are not ideal as they lack several important properties of the lipid-bilayer, however, even for detergents that are considered strong, the structural properties of proteins in detergents are often similar to those in well-folded tertiary structures[310]. For smaller proteins like p7, the published structures and this study show that the influence of the chosen detergent can be more pronounced by influencing oligomerisation state and secondary structure. Thus, despite their limitations, the comparatively small size of detergent micelles provide an opportunity to explore structural propensities that can lead to physiologically relevant functional insights.

The lipid-like detergent DH7PC did not yield 2D ^{1}H,^{15}N-HMQC data of sufficient quality, although it has been used to determine the first low resolution oligomeric structure of p7 using EM[236]. Although in this study only the hexamer was analysed in detail, gel filtration data suggests the presence of non-hexameric species; thus it is possible that oligomer dynamics of p7 in DH7PC were responsible for the problems observed in NMR, the latter being more sensitive to heterogeneous samples than EM, in which classes are picked before data fitting[236]. It is reported that p7 can be hexameric or heptameric in lipids[204, 234] and both oligomeric states can also be found simultaneously in the same sample preparation (Luik *et al.*, unpublished). We also observed a

multidisperse behaviour of HCV JFH-1 p7 (gt 2a) in DH7PC in gel filtration experiments (data not shown). A monomer structure of HCV J4 p7 (C27S) was solved in the very similar detergent DH6PC (1,2-dihexanoyl-*sn*-glycero-3-phosphocholine), which has one less methylene in each acyl chain than DH7PC, although under relatively acidic conditions and at high temperature, which may have prevented oligomerisation and improved sample conditions for spectral analysis[240]. Neither the anionic detergent LMPG nor combinations of Fos-cholines were found to provide suitable conditions to study p7. The solubilisation of membrane proteins can be aided by organic or chaotropic reagents, such as HFIP (1,1,1,3,3,3-hexaflouro-2-propanol), urea or guanidine hydrochloride. Thin-film protocols, which use these reagents to ensure that all the protein and detergent is solubilised and thereby minimising aggregation and aiding proper refolding[259], resulted in only minor improvements in the case of HCV p7.

The initial spectral quality of p7 in the detergent Cyclofos-6 was promising. After optimisation of sample solubilisation, it was determined that the optimal temperature to study p7 in Cyclofos-6 is 37 °C. The spectral quality was largely independent of pH in a range between pH 6 and pH 7. Thus, a set of sample conditions were identified that fulfilled the initial goal of physiological temperature and pH. The study of the wild type sequences of HCV SA13 p7 and HCV J4 p7, however, proved to be challenging. Although well resolved spectra could be obtained using Cyclofos-6, only a fraction of the expected number of peaks was observed. Substitution of the sole cysteine at position 27 in the HCV J4 p7 sequence with a serine resulted in a large improvement in spectral quality, with essentially all expected peaks now being observed. The concentration dependency of HCV J4 p7 (C27S) in Cyclofos-6 was also investigated as it was found to solubilise more efficiently in higher Cyclofos-6 concentrations. The experiments showed that HCV J4 p7 (C27S) spectra were essentially similar between 28.6 and 286 mM Cyclofos-6, indicating an insensitivity towards the detergent concentration in this range.

The assignment of HCV J4 p7 (C27S) was successfully completed using the combination of a 3D ^{15}N-edited NOESY-HSQC and a 3D ^{15}N-edited TOCSY-HSQC, although a small number of overlapping and ambiguous residues remain to be confirmed by sequential assignment using a triple labelled sample. The expression and purification of a triple labelled sample of HCV J4 p7 (C27S) was successful, but the sample concentration was too low to record the necessary experiments.

Interaction studies using *N*N-DNJ confirmed a specific interaction and showed binding

saturation. Three regions of HCV J4 p7 (C27S) show the greatest chemical shift changes upon *N*N-DNJ addition, i.e. around Val9, His17, and Trp48. In the high detergent titration, there were no differences in peak shift visible between addition of 75 and 125 mM *N*N-DNJ. The titration in the lower amount of detergent showed the same chemical shift changes, albeit some residues could not be observed, as the protein concentration was too low.

Protein dynamics, based on ^{15}N-R_1 and R_2, hetNOE and amide-water exchange experiments of HCV J4 p7 (C27S) in Cyclofos-6 were indicative of different loop region boundaries than previously reported[237, 239–242]. Notably, the reported consensus region containing the dibasic motif, which is highly conserved among p7 isolates, becomes a part of what is predicted to be the first transmembrane helix (helix 1). The HCV J4 p7 (C27S) secondary structure suggested by our experiments is unique in its helix 1 boundaries, and the loop region most closely resembles that of Cook *et al.* and Montserret *et al.* structures. HCV EUH1480 p7 (mt5), which was used to determine a hexameric solution NMR structure in Fos-Choline-12[242] also shows the dibasic motif as part of a helix, but has a loop region moved towards the C-terminus with different boundaries. The similarity in an overlay of the 2D ^{1}H,^{15}N-HMQC and 2D ^{1}H,^{15}N-TROSY of HCV EUH1480 p7 (mt5) confirms that the sample reconstitution is working and that the conformational state of the protein must be similar.

The performed protein dynamics experiments of HCV EUH1480 p7 (mt5), consisting of hetNOE and amide-water proton exchange measurements, suggested a similar degree of flexibility between the two isolates in their respective detergents. In addition, 1D-based measurements of ^{15}N-R_1 and R_2 for HCV J4 p7 (C27S) in Cyclofos-6 and HCV EUH1480 p7 (mt5) in Fos-Choline-12 indicated very comparable overall tumbling times and also similar molecular weights. ^{15}N,^{1}H-TRACT experiments, measuring the rotational correlation times also result in monomeric molecular weights for both isolates. SEC and SEC-MALS further showed that HCV EUH1480 p7 (mt5) is monomeric in Fos-Choline-12, with a lack of any observable multimeric species. In addition, chemical cross-linking experiments and AUC could be used to better characterise the molecular weights of the p7 samples under NMR conditions.

In summary, a comparison to previously published structures shows that the secondary structure determined for HCV J4 p7 (C27S) in Cyclofos-6 has a different loop region, with the dibasic motif, which is implicated in viral fitness[230, 231], being part of a helical segment (Figure IV.18). This is supported by a Rosetta model of HCV J4 p7

(C27S). The N-terminal helix 1 very likely protrudes out of the detergent micelle, in a conformation that is similar to that observed by Foster *et al.*[241]. No function can be assigned to the protruding helix, but it does not appear to have significant amphipathic character[311] and a similar extended helical conformation has been reported elsewhere[312]. The three-dimensional structure of HCV J4 p7 (C27S) in Cyclofos-6 will confirm the preliminary dynamics and model data reported here.

Chapter V

Electrophysiology and oligomerisation of HCV p7

V.1 Electrophysiology and oligomerisation of HCV p7

The oligomeric state of HCV p7 is linked to its function as a viroporin. An EM study showed that HCV JFH-1 p7 (gt 2a) in the detergent DH7PC assembled into a hexamer[236]. Other suggested oligomerisation states comprise heptameric assemblies observed in another EM study[234], tetrameric to hexameric complexes simulated in an MD study[238], and a hexamer in the recently published first high-resolution structure of HCV EUH1480 p7 (mt5) in the detergent Fos-choline-12[236, 242]. In order to carry out biophysical analyses of HCV p7, large amounts of pure protein are required. We used peptide synthesis to chemically synthesise HCV p7. Isolates representing HCV genotypes 1 to 6 had been previously synthesised in our group and purified using HPLC. The goal of this work was to complete the synthesis of missing relevant p7 genotypes and pertinent mutants, and to analyse their oligomeric state and ion channel function. Mutations of interest included the dibasic motif (in HCV JFH-1 p7 (K33/R35)), which is implicated in viral fitness. Mutations of either or both of these residues to either glutamine or alanine impair ion channel function and severely reduce the amount of infectious virus produced[229–231, 234]. They also function as a signal sequence and are important to obtain the correct topology of HCV p7. Mutation in both residues to alanine abolishes the sorting signal, whereas mutation to glutamine leaves it intact[141,

227, 230, 238, 313]. It was not known whether the double mutants can assemble into a functioning oligomeric viroporin, or if the secretion defects associated with it are caused by other mechanisms directly involving the presence of the two basic residues, perhaps as part of a binding interface for viral or cellular proteins.

It has been previously suggested by our lab (T.Whitfield, unpublished) and others[224] that iminosugars disassemble or prevent assembly of p7 oligomers. A F25A mutant discovered in a genotype 3 background is suggested to confer iminosugar resistance and has been shown to be resistant to iminosugar-induced disassembly in native gel experiments[224]. In order to investigate these suggested mechanisms, we employed AUC (analytical ultra centrifugation), as the first choice to analyse channel oligomerisation. Several conditions and p7 genotypes, including HCV JFH-1 p7, HCV JFH-1 p7 (K33Q/R35Q), HCV J4 p7 and HCV 452 p7, have been trialled using the neutrally buoyant detergent C_8E_5 (*n*-octyl-hydroxypenta(oxyethylene), Sigma) which is most suitable for AUC. C_8E_5 solubilisation of all tested p7 was low and required extensive optimisation. A solubilisation protocol using several steps of sonication, vortexing, and heating at 37 °C was employed and reached protein concentrations acceptable for AUC analysis. However, unlike in previous experiments conducted in our group (unpublished), only the monomeric state could be observed, indicating that under the conditions tested oligomerisation had not occurred, and therefore this approach was unsuitable to investigate the influence of inhibitors on p7 assembly. In order to analyse the oligomeric state, a native gel protocol was developed instead, using the NativePAGE system (Life technologies), which is a technique for the analysis of native proteins and protein complexes. The previously reported native gel experiments using FLAG-tagged p7[224, 236] could not be repeated using chemically synthesised p7 of native length. NativePAGE is suggested to be able to estimate molecular mass and oligomeric states, and is based on blue native polyacrylamide gel electrophoresis (BN-PAGE)[234, 314]. It was used to determine the oligomeric states of HCV JFH-1 p7 (gt 2a) and other isolates, genotypes and mutants, including HCV JFH-1 p7, HCV 452 p7 (gt 3a), HCV H77 p7 (gt 1a) and HCV JFH-1 p7 (R33Q/R35Q).

In order to characterise HCV p7 electrophysiology and link possible mutations to p7 ion channel function, droplet interface bilayers (DIBs) developed by Dr Mark Wallace (Chemistry Department, Oxford) were used. DIBs are a way to generate artificial bilayers suitable for bulk and single-channel electrophysiological measurements. An aqueous

droplet in an oil solution containing phospholipids, is covered by a phospholipid monolayer. Another monolayer is formed on an agarose hydrogel coated cover slip, which upon contact with the aqueous droplet forms the bilayer at the oil-water interface. They are suitable for electrical recording and imaging of pores and ion channels[315]. The advantages over other systems, such as Black Lipid Membranes (BLM), include stability, possible application of asymmetric lipid bilayers and optical imaging (reviewed in[316]). It is also possible to manipulate the bilayer size of the bilayer after it is formed[317]. They are also referred to as droplet hydrogel bilayers (DHBs) when a solid hydrogel (i.e. agarose) is used as a base[316].

The protocol in brief entails (i) droplet generation, (ii) monolayer assembly and (iii) bilayer formation and (iv) measurement of activity (Figure V.1)[316].

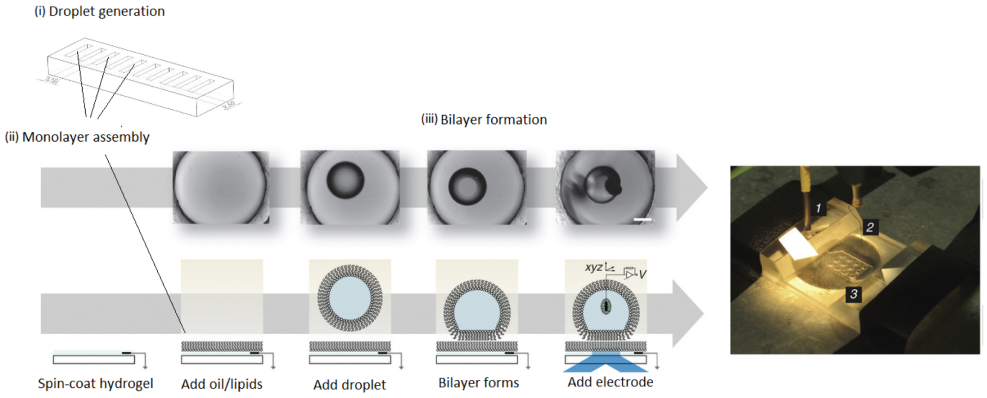

Figure V.1 – Formation of droplet-hydrogel bilayers. Left: A glass cover slip is covered in an agarose hydrogel, which is covered by an oil lipid solution. Droplets are prepared in a PMMA (Poly(methyl 2-methylpropenoate)). An aqueous droplet is also covered in a monolayer of lipid and added to a device. Once the bilayer is formed, two Ag/AgCl electrodes are inserted into the PMMA device. Scale bars are 200 μM. Right: A PMMA device is shown: 1) the ground Ag/AgCl electrode connecting to the agarose hydrogel. 2) The 100 μm Ag/AgCl electrode used for electrical measurements inside droplets. 3) The DHB device attached to the bottom of the Faraday cage. Tape is used to seal the agarose channel and to mount the device on the microscope table (Figures adapted from[316]).

V.1.1 Chemical synthesis of HCV p7

In order to perform biophysical experiments, HCV p7 needs to be either expressed in *E. coli* (Chapter III) or chemically synthesised. Peptide synthesis is the chemical production of polypeptides and proteins. SPPS (solid-phase peptide synthesis) involves the stepwise addition of protected amino acids to a growing peptide chain, which is bound by a covalent bond to a solid resin particle[318]. HCV p7 used in this thesis was

synthesised using two different synthesis methods, MAPS (microwave-assisted peptide synthesis) and SPPS. Both approaches employ the same chemistry but MAPS uses a microwave to heat up the reaction vessel, whereas SPPS is carried out at room temperature, which makes MAPS a special form of SPPS. MAPS in general shows higher yields and faster throughput than expression, and a faster throughput than SPPS. HCV p7 isolates previously synthesised in our laboratory by SPPS include H77 (gt 1a), JFH-1 (gt 2a), JFH-1 (gt 2a, R33A/R35A), 452 (gt 3a), ED43 (gt 4a), SA13 (gt 5a), Th580 (gt 6b) thus spanning all genotype families with the exception of genotype 7 (isolate QC69). HCV J4 p7 (gt 1b) was synthesised first for use in NMR experiments. Synthesis was carried out using fluroenylmethyloxycarbonyl (Fmoc SPPS) on a CS-Bio Peptide Synthesizer (CS536), as previously described for the other p7 isolates listed above. A 4.0 mmol synthesis using Fmoc as the protective group and Wang resin was performed. The synthesis steps in brief include repeated cycles of deprotection-wash-coupling-wash. HPLC was performed using acetonitrile as the organic phase (Buffer A: H_2O with 0.1 % FA and buffer B: acetonitrile 0.1 % FA), with an increasing gradient B from five to 95 % in 15 to 20 minutes. Two peak fractions were collected (Figure V.2) and the purity and mass of the protein were determined by mass spectrometry (ESI-MS) (Figure V.2). The synthetic peptide was pure, although a discrepancy in mass was observed. The expected mass of HCV J4 p7 is 6706.0 Da and the observed mass is 6790.23 Da. Using Edman degradation and N-terminal protein sequencing of the first 40 residues, the difference of 84 Da was found to be due to an exchange of Ala29 and Ala30 with two leucine residues. Despite the observed mass difference in HCV J4 p7, the applied HPLC protocol was used to purify other p7 genotypes.

SPPS using the CS-Bio synthesiser is time and labour consuming, with a maximum of ten amino acid couplings per day in an optimal system set up. In order to speed up synthesis, a CEM Liberty microwave synthesiser (CEM Microwave Technology) was used. MAPS of HCV J4 p7 was carried out at a smaller scale of 0.1 mmol and was completed after three days. HPLC was performed with a longer, 20 minute gradient. MS analysis of crude HCV J4 p7 confirmed the correct mass (Figure V.3). The HCV p7 genotype 7 isolate QC69 was synthesised using the same synthesiser and scale. MS analysis of crude HCV QC69 p7 showed a reduced amount of full length HCV QC69 p7 (Figure V.4 a, peak at the far left). SDS-PAGE analysis revealed a correct sized band (Figure V.4 a), with possible aggregates visible in the bands that have been loaded with higher concentrations of crude HCV QC69 p7 (Figure V.4 b).

The synthesis of HCV J4 p7 was successful using the CS Bio synthesiser and also MAPS. The former yielded very pure, crude protein on a higher scale than MAPS. The latter was faster, enabling higher throughput but gave rise to lower yields. MS analysis confirmed the correct mass of MAPS HCV J4 p7, whereas MAPS of HCV QC69 p7 showed a decreased amount of final product. HCV QC69 p7 is pending HPLC purification and MS mass confirmation.

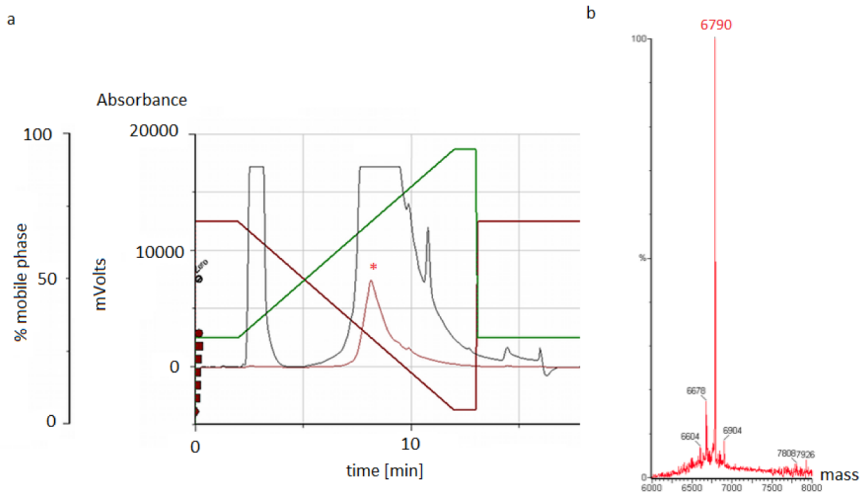

Figure V.2 – Purification of chemically synthesised HCV J4 p7.a) The HPLC chromatogram of synthesised, crude HCv J4 p7 is shown. HCV J4 p7 is found in the big central peak (red asterisk). A C4 column (Jupiter, 300 Å, 10 μ, Phenomenex) was used with a Gilson PrepLC system (321 Pump, UV/VIS-155 detector, UniPoint software). Two wavelengths of the UV spectrum, 215 nm (black) and 280 nm (red) were recorded. b) The MS spectrum of HPLC purified J4 p7 is shown. Purity and mass of HCV J4 p7 were assessed using electrospray ionisation-mass spectrometry (ESI-MS). The analysis shows one main mass species of about 6790 Da. The calculated mass of HCV J4 p7 is 6706.0 Da; the discrepancy of 84 Da is due to an exchange of two alanine residues at position 29 and 30 with leucine residues, as determined by N-terminal sequencing.

V.1.2 NativePAGE of HCV p7

NativePAGE analysis was used to study the oligomeric state of HCV p7, using chemically synthesised p7. A previously published BN-PAGE and the EM-structure reconstruction using HCV JFH-1 p7 and DH7PC was used as a starting point[236]. Sample optimisation was carried out (Figure V.5 a), during which it was found that the addition of surplus DH7PC is needed to induce p7 band formation, whereas the sample additive G250 had no visible effect (Figure V.5 b). Putative oligomeric HCV JFH-1 p7 bands were visualised using other isolates with representatives of all HCV p7 genotype families excluding

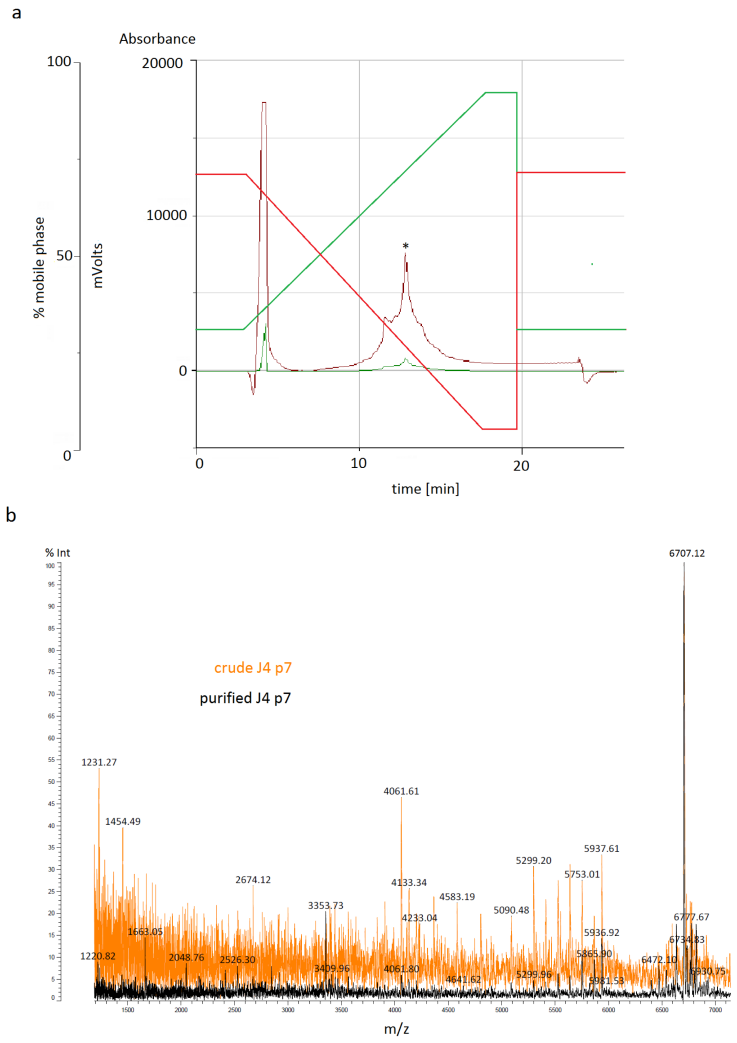

Figure V.3 – HPLC purification and MS analysis of crude microwave-synthesised HCV J4 p7. a) The HPLC chromatogram of synthesised, crude, microwave synthesised HCV J4 p7 is shown. Two wavelengths of the UV spectrum, 215 nm (red) and 280 nm (green) were recorded. Pure HCV J4 p7 is found in the centre of big central peak (black asterisk). A C4 column (Jupiter, 300 Å, 10 μ, Phenomenex) was used. b) The MS analysis shows the purity of the central fraction of synthetic, crude HCV J4 p7 (orange) in overlay with the purified HCV J4 p7 (black) analysed with MALDI.

genotype 7. All samples ran at approximately the height of a theoretical hexamer or heptamer. This suggests that isolates which span genotype families 1 to 6 and the double mutant HCV JFH-1 p7 (R33A/K35A) all form oligomeric complexes (Figure V.5 c). Cyclofos-6 was also tested with HCV J4 p7 and ran lower than other isolates in DH7PC, indicating that p7 is present in a monomeric form under these conditions (Figure V.5 d)

In order to confirm these results and to identify the bands as p7 multimers, which

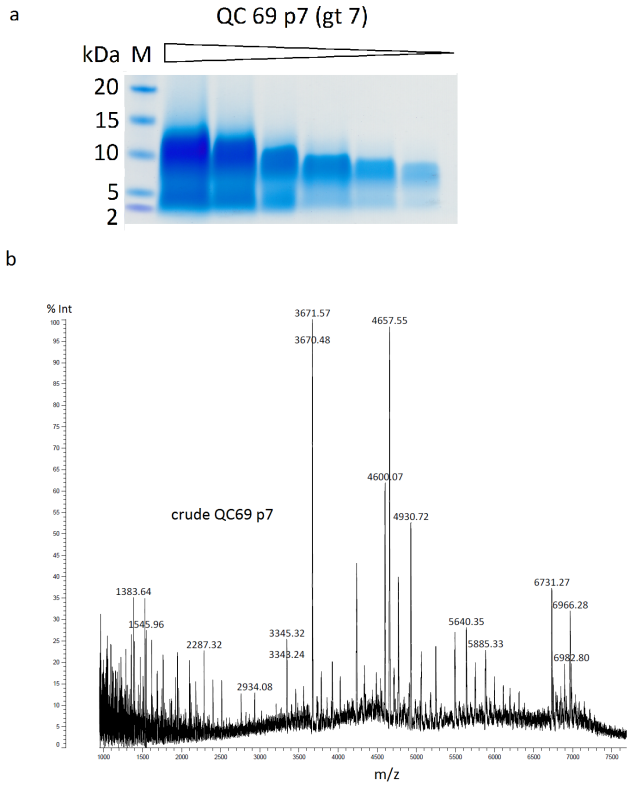

Figure V.4 – SDS-PAGE and MS analysis of crude microwave-synthesised HCV p7 QC69. a) SDS-PAGE analysis of crude QC69 p7 shows different concentrations loaded onto the lanes (indicated by triangle), with higher to lower concentration loaded from left to right, respectively. b) Shows the MALDI analysis of crude HCV QC69 p7. The peaks on the far left are the full-length HCV QC69 p7 (MW 7078.5 Da), with intermediate length pieces being visible in the middle and on the right of the spectrum. The discrepancy of sizes of the peaks on the right to the excepted mass are due to the calibration of the MALDI being not adjusted at the time of experiments performed.

run very low on the gel close to the Coomassie front, a HCV JFH-1 p7 N-terminal specific antibody (developed by phage display, Zitzmann group, unpublished) was used for western blotting. This antibody specifically bound to HCV JFH-1 p7 containing lanes only, when analysing a similar pan-genotype p7 collection as in the gel described before. Although the western blot was overloaded, it nevertheless indicated that the bands seen on the NativePAGE are oligomeric p7 channels. This also suggested that the dibasic mutant HCV JFH-1 p7 (R33Q/R35Q) and its associated defects are not linked to an inability to form oligomers (Figure V.5).

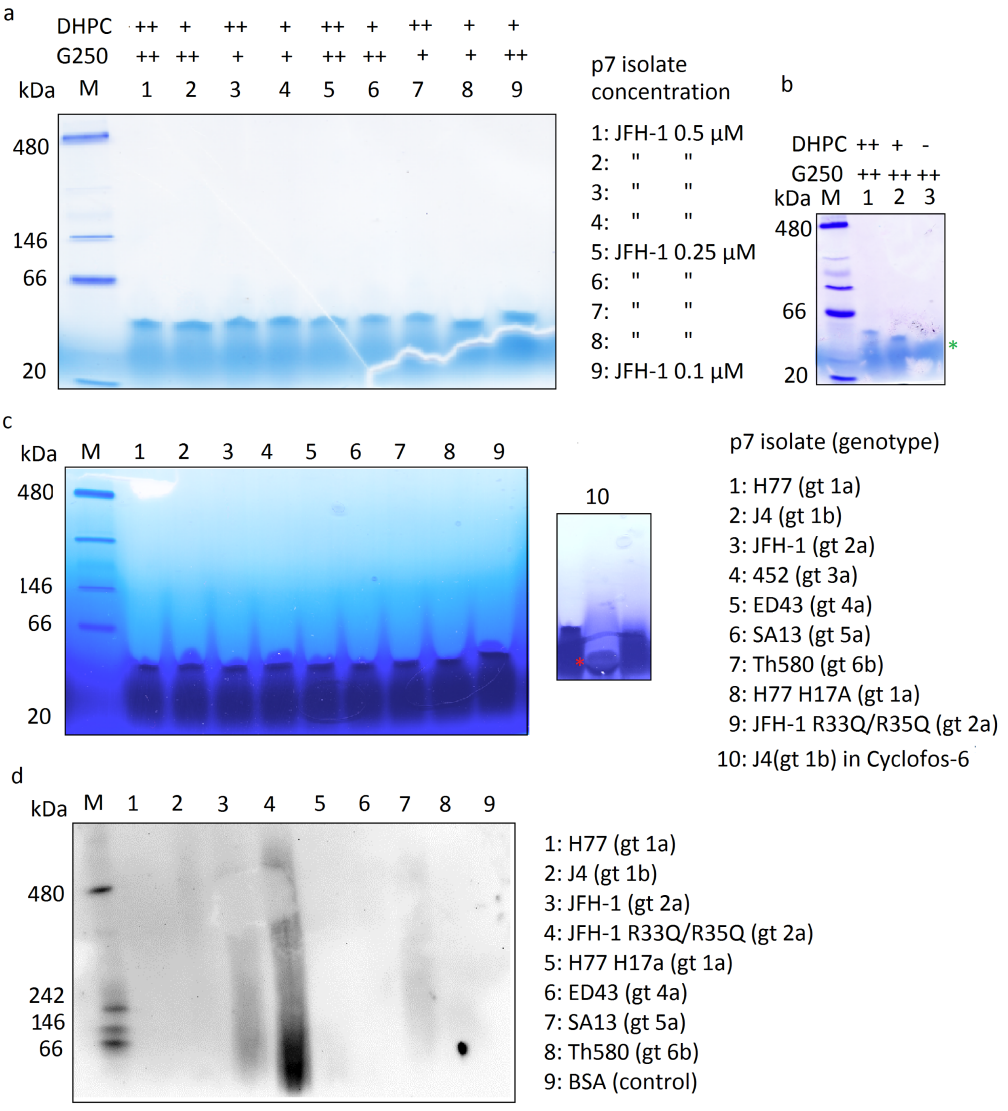

Figure V.5 – NativePAGE analysis of HCV isolates. a) The top 4-16 % Bis-Tris gel shows the detergent and G250 addition optimisation process. Protein concentrations are listed next to the gel. In brief, JFH-1 was solubilised in 60 mM DHPC, and the final sample was prepared according to the manufacturers' instructions, which suggested the addition of G250. b) Without addition of further DH7PC, no bands were visible (green asterisk). c) The optimised conditions were used to analyse a panel of HCV p7 genotypes. Cyclofos-6 solubilised p7 ran at the size expected for a p7 monomer (red asterisk) d) Western blot analysis of a 3-12 % Bis-Tris gel, using a HCV JFH-1 p7 specific N-terminal antibody, shows that the antibody only reacts with HCV JFH-1 p7 sequences. NativeMark was used as a molecular weight marker in all gels (Life technologies).

V.2 Droplet Hydrogel Bilayers of HCV JFH-1 p7

Chemically synthesised HCV JFH-1 p7 was used to investigate the possibility to apply the DHB system to study the electrophysiological characteristics of p7. The goals were to investigate and characterise selected isolated from all seven genotypes in their electrophysiological behaviour in different lipids (e.g. ER-like lipid compositions) and their respective responsiveness to iminosugar and other inhibitors, including BIT225 and amantadine. Especially the reported inability of *N*N-DNJ to disassemble a genotype 3 isolate[224] and the possible single point mutation F25A conferring resistance to iminosugar treatment, possibly also in other isolates, made further characterisation using a thorough electrophysiological analysis necessary. The overall reported discrepancies in amantadine efficacies in blocking p7 could also be investigated in a more ER-like membrane. It has been shown that p7 ion channel characteristics are dependent on lipid compositions[246], which might explain previously reported differences besides the varying sequence differences of p7 used[204, 223].

In order to evaluate the DIB system for compatibility with HCV p7, the lipid DPhPC (1,2-diphytanoyl-*sn*-glycero-3 phosphocholine) was used as a model system in hexadecane with HCV JFH-1 p7. A stock solution of HCV JFH-1 p7 was prepared in DMSO (dimethyl sulfoxide) and diluted in DHB buffer (0.5 M KCl/5 mM Hepes/1 mM $CaCl_2$, pH 7.4). The procedure describing how to set up the device is depicted in Figure V.1 and reviewed in[316]. HCV JFH-1 p7 was added to the aqueous buffer that was used to prepare droplets, giving a final volume of 69 nL. Upon bilayer formation the droplets were stabbed using an Ag/AgCl microelelectrode under a microscope. In order to characterise ion channel activity of synthesised HCV JFH-1 p7, the current-voltage (I-V) relationship was determined. A voltage ramp was applied and the resulting I-V curve shows linear, ideal ohmic behaviour. This indicates firing of a non voltage-gated, leaky ion channel, which allows ions to pass in both directions under these experimental conditions (Figure V.6).

V.3 Discussion

The establishment of a quick chemical synthesis makes the study of a large variety of p7 mutants and isolates possible. Genotype and sequence specific difficulties in purification

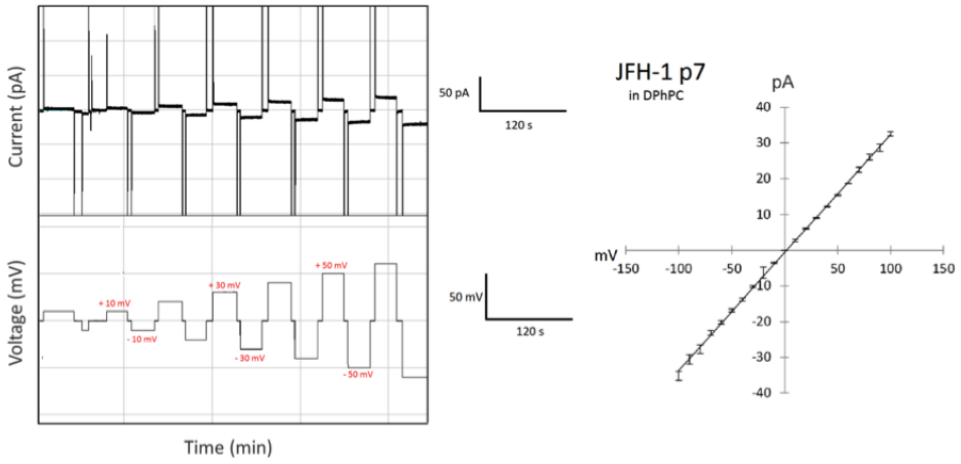

Figure V.6 – DHB electrophysiological characterisation of HCV JFH-1 p7 in DPhPC. Shown is the Current-Voltage (I-V) relationship as determined by a square voltage ramp. Averages of values hold for the time stated were used to calculate the position in the I-V curve (top left). Data points represent mean value for three technical replicates +/-S.E.M. The I-V curve shows ideal ohmic behaviour, which is indicative of a non voltage-gated ion channel.

cannot be ruled out, as observed in the synthesis of HCV QC69 p7.

Using native gel analysis it was possible to show, albeit in a crude way, that all tested isolates form oligomers. Different percentage gels could be used to shift the band further up and increase resolution. The antibody specific against the N-terminus of HCV JFH-1 p7 (C. Fotinou, Zitzmann group, unpublished) enabled the identification of the bands on the native gels as oligomeric p7 complexes. The interaction with both HCV JFH-1 p7 wildtype and the dibasic mutant R33Q/R35Q, confirms the specificity of the antibody and its ability to recognise HCV JFH-1 p7 as native protein, and it strongly suggests that the negative effects of the dibasic mutant of JFH-1 viral assembly and secretion is not linked to an inability of HCV JFH-1 p7 (R33Q/R35Q) ion channel to assemble, as complexes can be seen on the Western blot. The latter needs to be repeated with decreased loading concentrations on the NativePAGE in order to decrease smearing and hopefully resulting in a clear band separation. The dibasic JFH-1 mutant R33A/R35A, which has a more severe negative influence on virus production[229–231] should also be analysed. In previous work of the Zitzmann group the double arginine mutant was non-functional when its ion conductance was tested using a BLM system (Whitfield, unpublished). The confirmation of this result is currently awaiting DHB analysis. If confirmed, this would indicate that the JFH-1 R33Q/R35Q mutant p7 ion channel can assemble but not conduct ions.

The AUC protocol should be adapted and a previously published AUC method employing the co-lyophilisation of detergent and p7 prior rehydration should be tested[237]. We established for the first time the DHB system for HCV JFH-1 p7 analysis using DPhPC as a model lipid. The direction of current is defined as the flow of positive charge. With regard to a cell inward current, this is cation flow into the cell or anions out; conversely, outward current is positively charged particles flowing out of the cell, or negatively charged particles flowing into the cell. The observation that HCV JFH-1 p7 shows linear behaviour in the I-V relationship indicates that p7 is a non voltage-gated, leaky channel. The process of bilayer assembly used in this experiment will allow HCV JFH-1 p7 to insert in both directions into the membrane, as p7 is already present in the droplet before bilayer formation.

Previous experiments report HCV J4 p7 conductance as a function of the membrane potential, showing asymmetric behaviour in planar lipid bilayers and liposomes, suggesting that HCV J4 p7 inserts in a preferred orientation[237]. The discrepancies could be due to varying insertion orientations, different lipids and p7 concentration used or genotype specific differences. I-V behaviour of synthetic HCV J4 p7 using the DHB system in DPhPC bilayers should be used to compare the findings. It is also possible to add p7 via a microinjector to the droplet with an already formed bilayer. The eventual goal is to analyse all seven genotype in a ER-derived membrane, which would be more physiological relevant than previous studies using liposomes and BLM[204, 205, 223, 237, 246]. The DHB system is also well suited to look at inhibitor interactions, by either addition of the inhibitor to the aqueous solution making the droplets or via injections directly into the droplet with an already formed bilayer with a microinjector. Both will be used to look at iminosugar interactions against all seven genotypes. As a lipid environment dependency is implicated in p7 oligomerisation[246], the application of ER-like membrane or asymmetric bilayers will be interesting to investigate, and see how and if it influences inhibitor efficacies.

Chapter VI

Conclusion

HCV p7 is important for infectious virus production[139, 141, 231]. It is a cation-selective viroporin[163, 236], which has also been suggested to conduct protons[245]. Its ion channel function is the target of extensive inhibitor development efforts[139, 141, 210, 218, 236, 241, 288]. The need for alternative HCV drug targets other than the NS3/4A protease, even after the introduction of the protease inhibitors boceprevir and telaprevir is still high, as protease inhibitors and next generation DAA are likely not able to deal with viral escape mutants in highly mutating viruses like HCV[10, 13]. The ambition for future HCV treatment is an interferon free, affordable and pan-genotype effective drug, if possible all combined in one compound. However, a combination of a potent fast acting DAA with another pangenotypic compound with a high genetic resistance barrier and a favourable toxicity profile and no adverse drug interactions seems currently more realistic, such as the combination of anti-NS3 and anti-NS5A inhibitors with the novel NS5B polymerase inhibitor Sofosbuvir, which reportedly allows up to 100 % efficiency in HCV infected patients[319].

HCV p7 is interesting as a drug target as once inhibited, HCV can no longer produce infectious viral particles. The conservative approach of inhibiting assembled channels, e.g. with amantadine, may not be as promising, as mutations in only a small subset of amino acid side chains may render the drug non-effective. This was observed in clinical trials with amantadine in a recent study suggesting different targets other than p7 are conferring drug resistance[320].

Amantadine and rimantadine, which were initially discovered as Influenza A M2 ion channel inhibitors, are generally regarded as not suitable for HCV treatment. BIT225,

a putative p7 inhibitor currently in clinical trials shows promising results in combination therapy with interferon and ribavirin, although statistical relevance has yet to be reported. Another class of p7 inhibitors are iminosugars, which are suggested to disassemble and/or prevent assembly of the p7 ion channel (Whitfield *et al.* unpublished and[224]), showing high efficacies in BLM and cell-culture experiments[223]. We speculated that this might be due to interaction of these iminosugars with the aromatic amino acids located at the phospholipid headgroup/lipophilic side chain interface of the lipid membrane, which are crucial to keep the viroporin anchored in the membrane. It was hypothesised that if HCV were to introduce mutations within p7 to escape 'disassembling' iminosugar interactions, it would thereby incapacitate its own channel functionality, which is important for viral fitness and survival. Therefore we set out to determine the specific p7 amino acids that iminosugars interact with. For this, solution-state NMR is well suited.

Current literature does not report to which specific p7 amino acid residues either BIT225 or iminosugars bind. The recently published interaction data obtained for *N*N-DNJ from a study using monomeric HCV J4 p7 (C27S) in DH6PC (1,2-dihexanoyl-*sn*-glycero-3-phosphocholine) micelles shows interaction especially in water exposed areas, which may not be very informative to determine the binding site[240]. Knowledge of the specific interaction sites would contribute to our understanding of how p7 functions as an ion channel, and why so far no viral escape mutant to long alkylchain iminosugar treatment could be observed in one laboratory that specifically tried to create and describe such mutants[321]. Such information would also aide in understanding the reported iminosugar resistant mutant of a HCV 452 p7[206]. The determination of drug binding sites on the p7 monomer will be a first step towards determining the specificity and perhaps mechanism of action of these drugs. Indeed, a recent study which used a monomeric NMR structure of a flag-tagged HCV J4 p7 (C27S) for structure-guided compound design, based on the suggested p7 adamantanes binding site, reports an increase in compound efficiency of 10,000 fold[241]. However, more relevant information would be gained from interaction site mapping on the assembled oligomer, representing the active form of p7.

As p7 in its monomeric form cannot function as an ion channel, the hunt for a high-resolution structure of a functional oligomer was fuelled by a 16 Å EM-structure of HCV JFH-1 p7 in DH7PC (1,2-diheptanoyl-*sn*-glycero-3-phosphocholine) micelles[236]. Recently, the first structure at atomistic resolution of a p7 channel was reported, using a

sequence of HCV isolate EUH1480 with five introduced mutations, which is one of only two genotype 5 sequences. The authors report a hexameric channel, with the opposite orientation compared to that reported in the hexameric EM study[236], and an unusual helix packing[242]. Several techniques were employed by several groups to study drug interactions with p7 and to gain an understanding on how the inhibitors work, all of them sharing the need for functional p7 protein. Chapter III of this thesis addressed this by the development and adaption of an *E. coli* expression system, in order to enable an economically viable production of isotopically labeled protein. At the same time this work was being carried out, a similar expression protocol was reported[288]. The establishment of a working expression system was a step forward, as until then only chemical synthesis of p7 was established to produce untagged p7 of native length and sequence in the large quantities required.

In order to avoid as many non-physiological parameters as possible in the experimental setup - as a membrane protein, the use of membrane mimics was unavoidable - wildtype p7 sequences at sample conditions of physiological pH and temperature were initially studied. Expression and purification of wildtype p7 isolates was partly achieved, however, the presence of cysteine residues in the sequences interfered with their purification. The introduction of a serine mutation at the sole cysteine in the J4 p7 sequence at position 27 overcame this. Because yields were initially low compared to other reports[288], a condensation protocol was successfully adapted and subsequently yielded sufficient quantities for NMR analysis.

Chapter IV deals with the screening for detergent conditions suitable for the envisaged interaction studies at neutral pH and at a physiological relevant temperature. The latter two are normally key determinants in solution NMR, which can have an immense influence on NMR spectral quality. We successfully identified Cyclofos-6 as a suitable membrane mimic, combining sufficient solubilisation capabilities with very promising NMR spectra at neutral pH and 37 °C. Promising results were obtained not only for HCV J4 p7, but also for HCV SA13 p7. Initially, both of them were used as wildtype sequences. The adaption of the HCV J4 p7 (C27S) sequence improved spectral qualities. It is worth noting that the same will be likely the case for HCV SA13 p7 and HCV 452 p7. The latter is a genotype 3 isolate of particular interest, as it is implicated in a proposed iminosugar resistance[224]. The purification and sample condition system established should be used to assign a cysteine-free HCV 452 p7 and determine if iminosugars interact with the proposed alanine residue in position 25, as no elevated interaction was

detected for the Phe25 residue in HCV J4 p7 (C27S). The successful assignment of HCV J4 p7 (C27S) made it possible to determine interaction sites on p7, which should be now dissected to identify the minimum residue requirement for *N*N-DNJ binding. Our solvent access data shows that *N*N-DNJ binds to some regions, which are, and others which are not, H_2O accessible, indicating a specific interaction. It is interesting to note that all three putative binding sites contain highly conserved residues, i.e. Leu8, His17 and Trp48. The working hypothesis how iminosugars assert their antiviral effect was that they bind to conserved aromatic residues (i.e. Trp30 and Trp48) and cause disassembly and/or prevent assembly. Aromatic residues found in membrane proteins are suggested to localise towards the membrane interface where they form hydrogen bonds and anchor the channel to the membrane[322, 323]. Aromatic residues found in the membrane axis are sometimes suggested to indicate a closed state of a channel. In the open conformation of an inward rectifying potassium channel (KirBac1.1), aromatic residues are clustered at the membrane interface, whilst in the closed conformation the aromatic residues are found evenly distributed along the membrane axis[323]. The observation that *N*N-DNJ did not interact with the channel anchoring aromatic amino acids that - if the channel was inserted into a membrane - would be located at the membrane interface between hydrophilic phospholipid headgroups and lipophilic side chains, suggests that our working hypothesis must be revised and that the combinatory effect of the conserved newly identified residues is a hurdle the virus cannot overcome without loosing its viral fitness. Triple labeled experiments will further confirm the assignment and collect structural information restraints, which can be used to support the determined Rosetta model of HCV J4 (C27S) p7 in Cyclofos-6. It highlights a unique conformation in the secondary structure, containing the dibasic motif in helix 1, which has not been reported previously. Pending confirmation by structure determination, helix 1 might protrude from the detergent micelle, which also represents a possible novel conformation of p7. Its function remains unclear. It is not amphipathic and no functional data is available, but one can envisage how it could possibly aide oligomerisation or interact with cytosolic/host cell partners if it was also found in a native membrane environment. The hydrophobic core size of detergents in micelles compared to that of native membranes is different. The calculated size for the hydrophobic core of a Cyclofos-6 micelle is between 20 - 24 Å, assuming a C-C bond length of 120 to 150 pm. A native ER membrane of rat hepatocytes has an approximate width of 37.5 $\pm$0.4 Å[324], which might lead to different conformations encountered in a cellular lipid membrane environment[238]. The determination

of a three-dimensional structure of HCV J4 p7 (C27S) in Cyclofos-6 would answer the question if a protruding helix exists. As we are looking at a monomeric form of p7, it is worth speculating that the encountered flexibility at His17 might be a place of possible rearrangements that happen upon oligomerisation. RDC (Residual dipolar coupling) would allow us to determine whether helix 1 and helix 2 are co-linear, similar to the RDC determination in a study looking at HCV J4 (C27S) p7 in DH6PC micelles[240]. The conditions determined here could also be used as a starting point to explore bicelles with the lipid DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine)[296]. Other possibilities to determine a structure of more physiological value include the application of solid-state NMR, for which some preliminary data have been published[239, 240], and the possibility of x-ray crystallography in the lipid cubic phase to determine a high-resolution structure of oligomeric p7[325].

The dynamic properties and solvent accessibility of EUH1480 p7 (mt5) in DPC show that both are very similar to HCV J4 p7 (C27S) in Cyclofos-6. The solvent accessibility is slightly elevated for the former, which could be due to the EUH1480 sequence being much less hydrophobic than the J4 sequence. In fact, HCV EUH1480 p7 was chosen for its relative polarity and was mutated further to make it even more polar[242]. This might affect secondary structure and micelle interactions, which could explain the higher exchange rates.

The analysis of molecular tumbling times indicates a monomeric rather than a hexameric protein-bicelle complex, which needs to be confirmed using cross-linking experiments. AUC would be an alternative approach to address this issue. The effect of detergent concentration on the oligomerisation state of HCV EUH1480 p7 (mt5) needs to be elaborated. The observed EM particles in the same study clearly show a hexameric assembly but are solubilised in a lower amount of detergent[242]. The influence of detergent concentration on oligomerisation was tested by gel filtration analysis and SEC-MALS and again no mutlimeric HCV EUH1480 p7 (mt5) was observed.

The oligomeric state can be analysed with several techniques and was addressed in chapter V. It also attempts a preliminary characterisation of electrophysiological properties of p7. The goal of using AUC as a medium-throughput readout of oligomer and inhibitor mechanism analysis was hampered first by low protein solubilisation and second by non-oligmerisation in the detergent used. The use of NativePAGE was optimised and a recently developed anti p7 antibody showed that the bands observed with nativePAGE represent oligomeric channels, confirming that several isolates form oligomers in DH7PC,

including the dibasic mutant of HCV JFH-1 p7. In conjunction with unpublished data from the Zitzmann group showing that the dibasic mutant is non-functional in the BLM electrophysiological set-up, this indicates that p7 ion channel function is needed for virus production and not the hexameric protein assembly itself. We are planning to repeat and confirm the electrophysiological measurements with the dibasic mutant in the DIB experimental set up.

The electrophysiological analysis of p7 is extensive, with several model membrane and systems used[204, 210, 223, 235, 237, 244]. The analysis of p7 using the novel DIB system was explored. Preliminary data using a model lipid are encouraging and the establishment of an ER-like membrane should be addressed next to analyse all genotypes in the first pan-genotype analysis in a lipid composition resembling native ER membranes as closely as possible. The mutant HCV p7 isolate EUH1480 used for the hexameric NMR structure should be analysed by DIB, as it was not possible to show electrophysiological activity in an oocyte assay[242]. The DIB set up would also lend itself to the detailed interrogation of a series of interesting mutants that should help delineating details of p7 ion channel function, using single ion channel recordings. Setting up the DIB system for a protein other than the α-hemolysin channel, which was extensively used for its development and description[326, 327], was a challenging and time-consuming undertaking. However, the proof of principle has been established here, and opens the way for future testing of relevant p7 channels and pertinent mutants.

The merging of electrophysiology and structural biology is a powerful tool to interrogate the structure function relationships of p7 wildtype and mutants of all genotypes, and to analyse anti-p7 compounds for efficacy, and will continue to further our understanding of HCV p7.

Chapter VII

Material & Methods

VII.1 Cloning of HCV p7 isolates

HCV p7 J4 (gt 1b) was cloned into the pMM-LR6 expression vector, referred to as pMM1 (Figure VII.1). The procedure was carried out according to the manufacturers' instructions, which is in brief as follows: Site-directed mutagenesis on a pGEX-6p2 J4 construct was performed using the QuickChange Site directed mutagenesis kit (Stratagene). Insert and vector were digested with restriction endonucleases BamHI and HindIII (NEB). Fragments were purified using the QIAquick Gel Extraction kit (Qiagen) and ligated with T4 DNA ligase at room temperature for 1 h into pMM1 vector. Preparations of plasmid DNA were carried out using the QIAprep Spin Miniprep kit (Qiagen). Bacterial clones were pre-selected by colony PCR, checked by restriction digest and DNA sequencing (Source Bioscience, Oxford). Synthetic and codon-optimised p7 genes for HCV J4 (gt 1b, UniProt: O92972), HCV 452 (gt 3a, UniProt: Q1W5D5) and HCV SA13 (gt 5a, UniProt: O91936) were obtained from Genscript and HCV p7 genotype 5a isolates EUH1480 wildtype (UniProt: O39928) and EUH1480 harbouring mutations T1G, C2A, A12S, C27T and C44S (EUH1480 p7 mt5) were obtained from GeneArt (Table VII.1). To be consistent, the initially used HCV J4 p7 construct was replaced by its codon optimised version.

Synthetic genes were cloned into pMM1 according to the manufacturer's instructions. In brief, the genes of HCV J4 p7, HCV 452 p7 and HCV SA13 p7 were transformed into DH5α or XL10 gold *E. coli*, amplified from a carrier pUC57 plasmid and purified using the QIAprep Spin Miniprep kit (Qiagen). Restriction endonucleases BamHI and

Table VII.1 – DNA expression constructs

Construct	HCV p7 isolate (genotype)	Gene	Supplier
pMM1 J4	J4 (1b)	Synthetic gene (pUC57 vector)	Genscript
pMM1 J4 (C27S)	J4 (1b)	Synthetic gene (pUC57 vector)	Genscript
pMM1 J4 (C27S)	J4 (1b)	Synthetic gene (DNA String)	GeneArt
pMM1 452	452 (3a)	Synthetic gene (pUC57 vector)	Genscript
pMM1 SA13	SA13 (5a)	Synthetic gene (pUC57 vector)	Genscript
pMM1 EUH1480	EUH1480 (5a)	Synthetic gene (DNA String)	GeneArt
pMM1 EUH1480 (mt5)	EUH1480 (5a)	Synthetic gene (DNA String)	GeneArt

HindIII were used to digest the plasmids. The digest was separated using agarose gel electrophoresis and purified with the QIAquick Gel Extraction Kit (Qiagen). Using T4 DNA ligase (NEB), the digested genes were ligated in pre-digested pMM1 plasmids. The presence of an insert was checked by restriction digests and the positive clones by DNA sequencing. HCV EUH1480 p7 wildtype, HCV EUH1480 p7 (mt5) were ordered as double stranded DNA Strings (GeneArt), which were directly digested with BamHI and HindIII, subcloned into pMM1 and transformed into NEB5α (NEB) cells (Table VII.2). Positive clones were checked by DNA-sequencing.

VII.1.1 Enzymes

All restriction endonucleases, as well as T4 DNA ligase, VentR and Phusion DNA Polymerase were purchased from NEB, whereas Pfu DNA polymerase and DpnI were part of the QuickChange site-directed mutagenesis kit (Stratagene).

VII.1.2 Vectors

Two different vector systems were used during the course of this work. The vector pGEX-6p-2 contained the initially used HCV J4 p7 sequence, which was excised and cloned into pMM1. The expression vector pMM1 has a kanamycin resistance (Kanr), an N-terminal His_9-tag and a trpΔLE tag (Figure VII.1). The trpΔLE tag allows protein expression in inclusion bodies and under denaturing conditions.

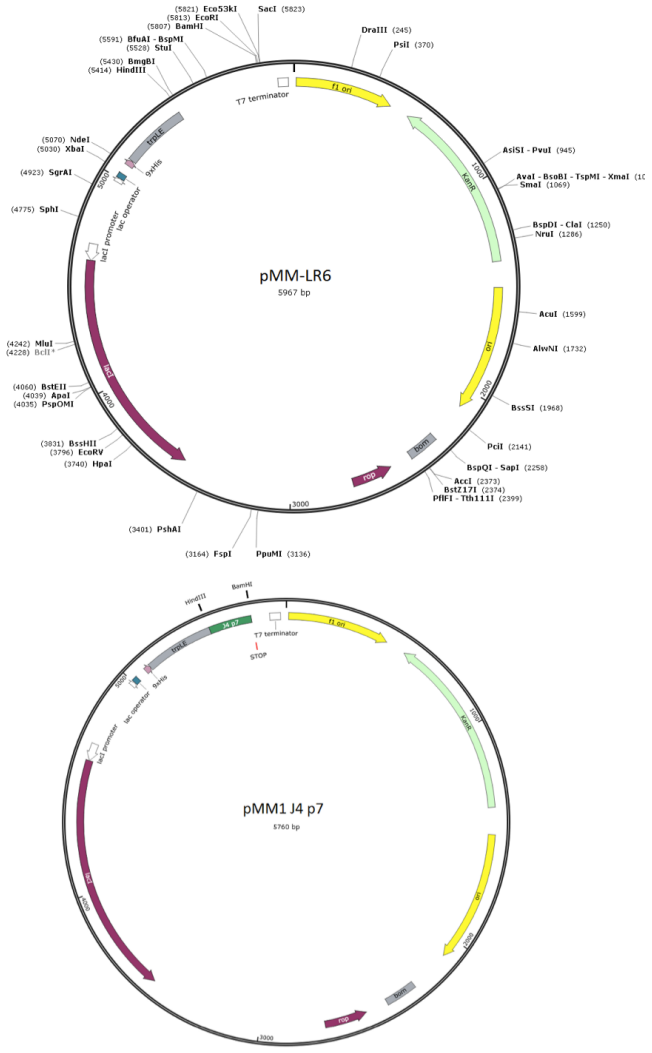

Figure VII.1 – Map of pMM-LR6 and pMM1 J4 p7 plasmids. The vector pMM-LR6 (gift from S.C. Blacklow, Harvard Medical School, Boston, provided by J. Schnell, Biochemistry Department, Oxford University) is based on a pET-17b (Kan^r) vector (Novagen), and contains a 9xHis-tag (pink) and a trpΔLE sequence (grey), latter containing the leader sequence of the *trp* operon of *E. coli*, which is fused to the 97 amino acid residues of the C-terminus of the anthranilate synthase gene. HCV J4 p7 (green) and all other isolates used were cloned into pMM-LR6 via two restriction sites (HindIII and BamHI, respectively). The sequences were confirmed by DNA sequencing (Source Bioscience, Oxford). Expression with a trpΔLE tag directs protein into inclusion bodies. Methionines in the target protein sequence have to be mutated (e.g. to leucine) to allow for cleavage with CNBr, which cleaves specifically after a single methionine at the fusion site. SnapGene (GSL Biotech) was used to generate the vector maps.

Table VII.2 – Bacterial strains

Strain	Genotype	Application
XL10-Gold	Tet^r textitΔ(mcrA)183 Δ (mcrCB-*hsdSMR-mrr)173 endA1 cloning, supE44thi-1 recA1 gyrA96 r elA1 lac Hte [F' proAB lacIqZ* Δ*M15* Tn*10* (Tet^r Amy Cam^r]	subcloning, plasmid amplification
DH5α	F- Φ80*lacZ*ΔM15 Δ*(lac*ZYA *-arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK-, mK+) *pho*A *sup*E44 λ- *thi*-1 *gyr*A96*rel*A1	subcloning, plasmid amplification
NEB5α	*fhuA2* Δ*(argF-lacZ)U169 phoA glnV44* Φ*80*Δ *(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	subcloning, plasmid amplification, high transformation efficiency
BL21 (DE3) pLysS	F$^-$ *ompT gal dcm Ion hsdS*$_B$(r$_B$$^-$ mB$^-$) λ(DE3) pLysS (Cam^r)	protein expression (T7-promotor)
BL21 Rosetta 2 (DE3)	F$^-$ *ompT hsdS*$_B$(r$_B$$^-$ m$_B$$^-$) *gal dcm* (DE3) pLysSpRARE2 (Cam^r)	protein expression (rare eukaryotic codons)

VII.1.3 Competent cells

The $CaCl_2$ method was used to make competent *E. coli* cells. 10 ml LB (Lysogeny broth) with appropriate antibiotic (e.g. 10 μL chloramphenicol for BL21 cells), were inoculated with a single colony or with small volume of a glycerol stock and incubated overnight at 37 °C and 250 rpm. The next morning cells were subcultured into 100 ml LB. The culture was incubated until the OD_{600} reached 0.5 - 0.7, then put on ice and further incubated for 20 minutes. Cells were harvested by centrifugation at 2,500 g for 10 minutes at 4 °C. The cell pellet was resuspended in 10 ml cold 0.1 M $CaCl_2$ per 50 ml culture and incubated on ice for 30-60 minutes. Cells were again harvested by centrifugation at 2,500 g for 10 minutes at 4 °C. The cell pellet was resuspended in 4 ml cold CaC_2 (2 ml per 50 ml initial culture) and 1 ml of 80 % glycerol was added. Competent cells were stored at -80 °C.

VII.1.4 Transformation

XL10 Gold (Stratagene), DH5α, NEB5α and BL21 cells (Life technologies) were transformed by heat shock according to the manufacturers' instructions (Table VII.1). In

Table VII.3 – PCR primer

Construct	Primer sequence
pGEX-6P-2 p7 J4 Met	for: 5'-AGGATGACGACGATAAGATGGCCTTAGAGAACTTGGT rev: 5'-ACCAAGTTCTCTAAGGCCATCTTATCGTCGTCATCCT
pMM1 J4 p7	for: 5'-TTCGTAGCAAGCTTTGGATGGCCTTAGAGAACTTGGTGG rev: 5'-CTAGGGATCCCATGGCGTAAGCTCGTGGTGGTA
pMM1 J4 p7 (C27S)	for: 5'- TCTGTCCTTTCTGGTCTTCTTCGCTGCGGCCTGGTATATTAA rev: 5'- TTAATATACCAGGCCGCAGCGAAGAAGACCAGAAAGGACAGA

brief, cells were thawed on ice and 50-100 μl were aliquoted into pre-chilled 14 ml round-bottom tubes. 1-2 μl of a ligation mix or 5 to 10 ng of plasmid DNA were added, mixed and incubated on ice for 30 min. BL21 cells were pre-incubated for 10 min with 1 μl 1:10 diluted β-mercaptoethanol. Cells were put into a water bath at 42 °C for 45 seconds and incubated on ice for 2 minutes. 450 μl pre-warmed LB or SOC media (NEB) was added and cells were incubated at 37 °C at 250 rpm for 60 minutes For a ligation reaction of a digested PCR product with vector, all cells were transferred to a 1.5 ml Eppendorf test tube and centrifuged at 5,000 rpm for 5 minutes to harvest the cells. The cell pellet was resuspended in 50-100 μl supernatant and streaked out on a suitable LB agar plate. In case of plasmid re-transformations, 50-100 μl were plated out without prior centrifugation. Plates were incubated at 37 °C overnight.

VII.1.5 Site-directed mutagenesis

Site-directed mutations in pGEX-6p-2 J4 p7 to prepare cloning from the initial construct into the pMM1 vectors, i.e insertion of the methionine pivotal for cleaving prior to the protein sequence and restriction cleavage sites, were generated using pGEX-6P-2 J4 p7 as a template, and specific primers (Table VII.3). The C27S mutant of HCV J4 p7 was created using specific primers (Table VII.3) and pMM1 J4 as a template following the manufacturers' instruction, both using the QuickChange kit (Stratagene).

VII.1.6 DNA sequencing

Sequencing of DNA constructs was conducted by Geneservice Oxford (Source Bioscience) in their DNA sequencing facility in the Department of Biochemistry, University of Oxford.

VII.1.7 Commercial Kits and reagents

The QIAprep Spin Miniprep kit (Qiagen) was used to prepare plasmid DNA from *E. coli* cultures. The QIAquick Gel Extraction kit (Qiagen) was used to extract DNA from agarose gels and the QuickChange Site-directed mutagenesis kit (Stratagene) was used to insert point mutations on plasmid DNA. If not otherwise stated, all materials and chemicals have been purchased from Sigma-Aldrich, Fisher or VWR.

VII.1.8 Polymerase Chain Reaction

Primer oligonucleotides were ordered from Sigma-Aldrich (Table VII.3). PCR (polymerase chain reaction) was used to amplify DNA fragments required for cloning of the HCV J4 p7 construct and to introduce the required restriction endonuclease recognition sites. VentR and Phusion DNA Polymerase (NEB) and the corresponding ThermoPol and Phusion HF buffers were used according to the manufacturers' instructions.

VII.1.9 Agarose gel electrophoresis

After completion of the PCR reaction, agarose gel electrophoresis was used to separate and purify the DNA fragments. 0.6 to 1 % gels were used, for which 0.3 to 0.5 g of agarose were added to 50 ml 1x TBE buffer. The agarose was dissolved in a microwave. After the solution cooled down to approximately 50 °C, 0.5 g/ml ethidium bromide was added and the mixture poured into a pre-assembled gel cast system. Each PCR sample was mixed with 6x DNA loading buffer. Electrophoresis was carried out at 80 - 100 V for 45 minutes.

VII.1.10 DNA gel extraction

Bands containing the DNA fragments were excised under a UV light source and DNA was extracted using the QIAquick gel extraction kit according to the instructions provided by the manufacturer.

VII.1.11 Restriction endonuclease cleavage

To digest the purified DNA fragments, vectors and synthetic genes (i.e. HCV p7 isolates, pGEX-6p-2 and pMM1) restriction endonucleases BamHI and HindIII and the corresponding buffers were used according to the manufacturers' instruction. Both reaction mixtures were combined and briefly centrifuged. 1 μl of each restriction enzyme (BamHI and HindIII, respectively) were added to the PCR product reaction mix. The vector reaction mixture was split into four samples and enzyme was added.

VII.1.12 Ligation

Ligation was performed using a molar ratio between 1:5 and 1:10, using T4 Ligase and the corresponding buffer. Reactions were incubated at room temperature for 30 minutes and stored at -20 °C without heat inactivation.

VII.1.13 Colony PCR

Colony PCR was used to screen bacterial colonies for the presence of clones that contain the correct insert. Bacterial cells were lysed by adding reaction mix and heating during the first PCR cycle. PCR was performed alongside a positive control (e.g. pGEX HCV J4 p7). Resulting reaction mixes were analysed on an agarose DNA gel. PCR reactions were performed in a total volume of 25 μl (i.e. after template was added). A mastermix containing everything except the colonies and the template in case of the positive control was prepared. Selected colonies were transferred from the plate into the PCR tubes and mixed with the reaction mix. The positive control was the PCR template that has been originally used to clone or subclone the inserted gene from. 0.5 μL of the plasmid and 1.5 μL dH_2O (distilled H_2O) were added to the positive control reaction mix. Standard PCR was performed, with 10 added cycles to the cycle number used in the original PCR reaction. All samples were checked on a 1 % agarose DNA gel. Positive clones showed a band matching the band in the positive control and were further verified by restriction digest and DNA sequencing.

VII.1.14 Solid-Phase Synthesis of p7

HCV J4 p7 was synthesised using fluorenylmethoxycarbonyl solid-phase peptide synthesis (Fmoc SPPS) on a CS-Bio Peptide Synthesizer (CS536). Derived peptides were cleaved from Wang resin ([4-hydroxymethyl-(phenoxymethyl)]-polystyrene), using triufloroacetic acid/triisopropylsilane/1,2-ethanedithiol/H_2O (94:2:2:2 vol %) and extracted using diethylether. Crude peptides were resuspended in 1 - 5 % formic acid, 60 % acetonitrile and 35 - 39 H_2O. The peptide was purified by preparative reverse phase-HPLC, using a Phenomenex C4 column (Jupiter, 300 Å, 10 μ) and a Gilson PrepLC system (321 Pump, UV/VIS-155 detector, UniPoint software), in a gradient of increasing acetonitrile concentration. The gradient was established using two buffers (A: H_2O, 0.1 % formic acid; B: acetonitrile in H_2O, 0.1 % formic acid) ranging from 10-100 % buffer B over a 10 minute period at a flow rate of 30 ml/min. All collected fractions were lyophilised overnight. Purity and molecular mass of HCV J4 p7 were assessed using electrospray ionisation-mass spectrometry (ESI-MS) (Department of Biochemistry, University of Oxford). HCV J4 p7 and HCV QC69 p7 were synthesised on a CEM Liberty microwave synthesiser (CEM Microwave Technology Ltd) using Fmoc chemistry. Peptides were cleaved off HMPB NovaPEG resin (Novabiochem) using TFA (triufloroacetic acid) and precipitated using diethylether acid and extracted using diethylether.

VII.1.15 SDS-PAGE

SDS-PAGE (Sodium Dodecyl Sulphate-PolyAcrylamide Gel Electrophoresis) was used to estimate the molecular weight of proteins and to evaluate sample purity. Precast 4 - 12 % Bis-Tris NuPAGE gels in MES buffer were used according to the manufacturers' protocol (Life technologies). Precision Plus Protein Dual Xtra (Biorad) was used as the protein marker if not otherwise stated. Quick Coomassie stain (Generon) was employed to stain the gels. Voltage applied was 200 V unless otherwise stated and the gels typically ran for 35 minutes. SDS-running buffer typically contained 35 μL 4 X LDS (lithium dodecyl sulphate) sample buffer (Life technologies), 35 μL 8 M urea, 5 μL 10 % SDS and if disulphides are present 5 μL DTT (Dithiothreitol) (or 5 μL 8 M urea if no disulphides were present). NativePAGE (Life) with precast 4 - 16 % Bis-Tris Gels were used according to the manufacturers' instructions to estimate the size native protein complexes, using NativeMark (Life technologies) as a marker.

VII.1.16 Western blotting

Western blots of NativePAGE gels were performed using a XcellII blot module (Novex) according to the manufacturers' instructions. The dark blue NativePAGE running buffer was exchanged half way through the run with a light blue running buffer. The gel was transferred to a PVDF (polyvinylidene fluoride, Millipore) membrane, which was wetted in methanol beforehand. For blocking, cooled I-Block (life technologies) at 2 % in 0.1 % PBST (0.1 % Tween in Phosphate Buffered Saline) was used for 1-2 h at room temperature or at 4 °C overnight. Two wash steps followed with 2x PBST (10 minutes each) and then the primary antibody was added. The HCV p7 antibodies used were raised using phage display technology against an N-terminal peptide of JFH-1 p7 (gt 2a) (Zitzmann laboratory, unpublished). After 2 h incubation at room temperature, the antibody was collected and the membrane washed four times with 2x PBST. Secondary-antibody was added for 1 h, followed by 4x wash with 2x PBST. It was developed using a chemilumenescence kit (Amersham) and imaged with a Las-1000 CCD (charge-coupled device) camera in an intelligent dark box (FujiFilm).

VII.1.17 Methanol/Chloroform precipitation

Protein samples for SDS-PAGE analysis, which contain 6 M guanidine hydrochloride were precipitated by a methanol/chloroform method (Cold Spring Harbour Laboratory Manuals). One volume of protein solution (for purifications: elution, wash and flowthrough 100 μL), was mixed with four volumes of methanol (400 μL) and vortexed. One volume of chloroform was added (100 μL) and vortexed again. After addition of three volumes of H_2O (300 μL) and vortexing, the mixture was centrifuged for 2 minutes at maximum speed in an Eppendorf table top centrifuge. An interface forms with an aqueous top layer that contains the guanidine, and a bottom layer, which contains the protein. The top layer was removed with a pipette and four volumes of methanol were added, which precipitates the protein (N.B. flakes are not always visible). The mixture was centrifuged at the maximum setting to pellet the protein, supernatant was aspirated and the pellet was dried and prepared for SDS-PAGE, by adding SDS running buffer.

VII.2 Protein expression

The pMM vector system used in this study provides a variety of options for producing polyhistidine-tagged fusion proteins, which can be purified using IMAC. HisPur Ni-NTA Resin (Pierce Biotechnology) binds selectively to polyhistidine sequences using nickel ions and is used to purify the proteins. TALON resin (Takara), which uses cobalt instead of nickel was also used. The pMM vector (Figure VII.1) system is based on the ability of the trpΔLE-fusion protein to direct protein expression into inclusion bodies. CNBr specifically cleaves after methionine residues and is used to cleave off the trpΔLE and His_9-tag. In brief, p7 constructs were transformed into *E. coli* BL21 (DE3) pLys cells (Table VII.2). Expression was assessed by analysing overexpression bands of overnight cultures, which are inoculated from using single colonies picked from agar plates. All growths were carried out at 37 °C and between 250 and 350 rpm on orbital shakers with suitable antibiotics, i.e. kanamycin and chloramphenicol. The next morning upon saturation growth, cells were spun down and resuspended in LB and added to fresh growth medium. Protein expression was induced at OD_{600}= 0.6 using 1 mM IPTG. After overnight expression, cells were harvested, the pellet redissolved in 40 μL H_2O, 20 μL loading buffer, 5 μL DTT and 5 μL 10 % SDS and analysed by SDS-PAGE. For uniform isotope labelling, expression in isotope-enriched M9 minimal media was used (Table VII.4). For ^{15}N-labelled protein, ^{15}N-H_4Cl (ammonium chloride, Sigma-Aldrich) (1 g/L) was used as the sole nitrogen source. For uniform carbon labelling, ^{13}C-glucose (Sigma-Aldrich) (2-4 g/L) was used as the sole carbon source. To achieve partial deuteration, H_2O was substituted with D_2O (98-99 %, Sigma-Aldrich). For large scale expression, IPTG induction and autoinduction[287], which uses allolactose as an inducer, were used. All isolates were expressed in BL21 (DE3) pLys or BL21 Rosetta 2 cells (Table VII.2). In brief, starter cultures were grown in LB-media which was added to M9-minimal media or auto-induction media (NPS, 5052 plus ^{15}N-H_4Cl; Table VII.4). Cultures were induced with 1 mM IPTG at an optical density of 0.6.-0.8 and harvested after 24 h. To achieve a higher yield and optimise use of isotope-labels, a condensation protocol was used, in which cells were four-fold condensed. Cells were grown in LB media until they reached an OD_{600} of 0.6. Between three to four litres of LB media were used per growth. The cells were spun down at 3,350 g for 20 minutes at 4 °C and washed in 200 ml (per 1 L LB growth) of 1 x M9 salts and subsequently centrifuged again. The cell pellet was resuspended in 1 L of M9 media, containing the appropriate nitrogen,

carbon and/or deuteration source (i.e D_2O). The cells were incubated for a further 30 minutes on an orbital shaker at 37 °C. Protein expression was induced by addition of 1 mM IPTG, and cells were harvested after 24 h. For deuteration, cells were further induced 12 h into the growth, and grown for up to 36 h.

VII.3 Protein purification

After cell harvest, inclusion bodies were purified. Several methods were used. The first is a protocol previously applied for the purification of influenza M2[282]. In brief, cells were taken up in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl (sodium chloride) and solubilised using a Dounce homogeniser. One Complete Protease Inhibitor tablet (Roche) per 50 mL lysis buffer was added. In case of wildtype HCV J4 p7, HCV 452 p7 and HCV SA13 p7, all buffers were supplemented with 5 mM TCEP (Tris (2-carboxyethyl) phosphine hydrochloride) (Pierce). Resuspended cells were sonicated and centrifuged. Cells were harvested by centrifugation at 7,000 rpm for 30 minutes in an ultracentrifuge (Beckmann). Cell pellets containing the inclusion bodies were washed by resuspension in lysis buffer using a Dounce homogeniser, which was followed by three minutes sonication. The inclusion bodies and cell debris were spun down at 15,000 rpm for 25 minutes at 4°C. The inclusion body containing pellets were again solubilised in lysis buffer and spun down at 15,000 rpm for 25 minutes at 4°C in order to further remove water-soluble contaminants. The pellets were dissolved in 50 mL (per 1 L culture) of guanidine buffer (6 M guanidine HCl, 50 mM Tris (pH 8.0), 200 mM NaCl) using a Dounce homogeniser. Undissolved matter was pelleted by centrifugation at 18,000 rpm for 1.5 h (Table VII.4). The second protocol, used mainly for HCV J4 p7, HCV 452 p7 and HCV SA13 p7 (wildtype) contained added detergent washing steps in order to remove membrane fractions that complicate downstream HCV p7 purification. IGEPAL CA-630 (octylphenoxypolyethoxyethanol) (Sigma-Aldrich) and deoxycholic acid (sodium salt, Fisher), both at 1 % (v/v and w/v, respectively) were used. A higher than previously used salt wash was added as well (1 M NaCl instead of 200 mM). Inclusion bodies were taken up in guanidine buffer containing 6 M guandine HCl, 10 mM imidazole and 2 mM TCEP, and incubated with stirring overnight at 4 °C. The next day, the mixture was sonicated for 4 min and centrifuged to separate undissolved matter as described above (Table VII.4). The third protocol used to purify HCV J4 p7 (C27S)[288], was

adapted for the usage of the condensation protocol described above. The inclusion body purification differs in the use of 1 mM EDTA (ethylenediaminetetraacetic acid), 1 mM NaN_3 (sodium azide) and 15 % (v/v) glycerol in the lysis buffer. 30 mL of lysis buffer per one litre culture were used. The washing step contained as the protocol above the two detergents IGEPAL and deoxycholic acid and also 1 mM EDTA and 1 mM NaN_3. The inclusion bodies were solubilised in a guanidine buffer (6 M guanidine HCl, 0.5 M NaCl, 20 mM Tris (pH 8.0) and 5 mM imidazole), sonicated and centrifuged again to pellet undissolved matter (Table VII.4). The last protocol employed was used to purify HCV EUH1480 p7 (mt5)[242]. The lysis buffer contained 50 mM Tris HCl (pH 8.0), 200 mM NaCl and 1 mM EDTA. The guanidine buffer contained 6 M guanidine HCl, 50 mM Tris HCl (pH 8.0), 200 mM NaCl and 1 % (v/v) Triton X-100. 50 mL of lysis buffer per 1 L culture was used, and the resuspension and centrifugation was the same as for the methods described above. Purified inclusion bodies were resuspended in 50 mL per 1 L culture guanidine buffer (Table VII.4). In all cases, the fusion protein was purified the following day using IMAC and HisPur resin (Pierce). Washing with imidazole containing wash buffers is critical for purity of eluted p7 fractions (15-150 mM, respectively) (Table VII.4). Fractions containing protein were dialysed against H_2O until protein precipitated out of solution. The protein was then lyophilised in the HCV J4 p7 (C27S) protocol, before CNBr cleavage, whereas in the other protocols cleavage was performed on the wet pellet. In either case, the fusion protein was taken up in 70 % formic acid and cleaved using CNBr, (0.2 g/ ml), which releases the His_9-tagged trpΔLE from p7. Cleavage duration ranges between one to three hours (Table VII.4) and was performed under a nitrogen stream being stirred by a magnetic bar. In one hour protocols, the cleavage reaction was dialysed against water using a 3,500 MWCO (molecular weight cut off) dialysis cassette (Pierce) or a 1,000 MWCO membrane (Spectrum laboratories) and lyophilised overnight. In the HCV J4 p7 (C27S) protocol, the CNBr reaction was stopped by additions of 1N NaOH (sodium hydroxide), which was then diluted with H_2O and dialysed until neutral pH was achieved, followed by lyophilisation.

VII.4 High Performance Liquid Chromatography

Reverse-phase HPLC was used to purify crude synthesised and expressed HCV p7 (J4, 452, SA13 and EUH1480), on a Gilson ProStar instrument. Columns used are C4

Table VII.4 – Buffers for inclusion body purification protocols

	Purification protocol			
	M2-based	p7-adapted	J4 p7 (C27S)	EUH1480 p7 (mt5)
Lysis buffer	50 mM Tris HCl (pH 8.0), 200 mM NaCl, 5 mM TCEP, (plus protease inhibitor)	50 mM Tris HCl (pH 8.0), 5 mM TCEP, (plus protease inhibitor)	50 mM Tris HCl (pH 8.0), 1 mM NaN_3, 1 mM EDTA, 15 % (v/v) glycerol, (plus protease inhibitor)	50 mM Tris HCL (pH8.0), 200 mM NaCl, 1 mM EDTA, (plus protease inhibitor)
Wash buffer(s)	50 mM Tris HCl (pH 8.0), 200 mM NaCl, 5 mM TCEP	50 mM Tris HCl (pH 8.0), 5 mM TCEP, 1 % (w/v) deoxycholate + 1% (v/v) IGEPAL CA-630	50 mM Tris HCl (pH 8.0), 1 mM NaN_3, 1 mM EDTA, 1 % (w/v) deoxycholic acid, 1% IGEPAL CA-630	50 mM Tris HCl (pH 8.0), 200 mM NaCl, 1 mM EDTA
		50 mM Tris HCl (pH 8.0), 5 mM TCEP, 1 M NaCl		
Guanidine buffer	50 mM Tris HCl, 200 mM NaCl, 6 M guanidine HCl, 15 mM imidazole, pH 8.0, 5 mM TCEP	50 mM Tris HCl (pH 8.0), 6 M guanidine HCl, 10 mM Imidazole, 2 mM TCEP	20 mM Tris HCl (pH 8.0), 6 M guanidine HCl, 0.5 M NaCl, 5 mM imidazole	50 mM Tris HCl (pH 8.0),6 M guanidine HCl, 200 mM NaCl, 1 % Triton X-100
Resin Wash buffer(s)	50 mM Tris HCl, 200 mM NaCl, 6 M guanidine HCl,15 mM imidazole, pH 8.0, 5 mM TCEP	50 mM Tris HCl (pH 8.0), 6 M guanidine HCl, 25 mM imidazole, 2 mM TCEP	20 mM Tris HCl (pH 8.0), 6 M guanidine HCl, 0.5 M NaCl, 50 mM imidazole	50 mM Tris HCl (pH 8.0),6 M guanidine HCl, 200 mM NaCl, 1 % Triton X-100
		50 mM Tris HCl (pH 8.0), 6 M guanidine HCl, 100-150 mM imidazole, 2 mM TCEP		
Elution Buffer	50 mM Tris HCl, 200 mM NaCl, 6 M guanidine HCl, 400 mM imidazole, pH 7.0, 5 mM TCEP	50 mM Tris HCl (pH 8.0), 6 M guanidine HCl, 500 mM imidazole, 2 mM TCEP	20 mM Tris HCl (pH 8.0), 6 M guanidine HCl, 0.5 M NaCl, 500 mM imidazole	50 mM Tris HCl (pH 8.0), 200 mM NaCl, 6 M guanidine HCl, and 400 mM imidazole, pH 7.0
CNBr cleavage	5 ml of 70 % FA and adding 1 g of CNBr	5 ml of 70 % FA and adding 0.5-1.0 g of CNBr	10 mg/ml of the lyophilised fusion protein in 70 % FA, with 100 mg/ml CNBr	5 ml 70 % FA (1 hr, CNBr 0.2 g/ml)

HCl: Hydrochloride, NaN_3: sodium azide, EDTA: ethylenediaminetetraacetic acid, NaCl: sodium chloride, TCEP: Tris (2-carboxyethyl) phosphine hydrochloride, CNBr: cyanogen bromide, FA: Formic acid

(Jupiter 4 μm, 300 Å), C8 (Jupiter 4 μm, Proteo 90 Å, AXIA) and C18 (Luna 10 μm, 100 Å) preparative columns (250 mm x 21.2 mm) (Phenomenex). A C18 preparative column (Proto 5 μm, 300 Å) was used for work on one HCV p7 construct (EUH1480 p7 (mt5)) (Higgins Analytical). For synthetic HCV J4 p7 and HCV QC69 p7, an acetonitrile gradient (0.1 % formic acid) was used (Buffer A: H_2O, buffer B: acetonitrile). For EUH1480 p7 (mt5), buffer A was 40 % acetonitrile and buffer B was 60 % acetonitrile, both containing 0.1 % TFA. The sample was injected in 100 % FA. For the other buffer system, the sample was taken up in up to 1 % FA and then mixed with 60 % acetonitrile. All solvents and samples were filtered using a 0.22 μM membrane. Fractions of interest were lyophilised overnight. 1 ml of fractions to be analysed by SDS-PAGE were lyophilised separately. Lyophilised protein was checked by MS.

VII.5 Fast protein liquid chromatography

FPLC was used to purify HCV J4 p7 (C27S), on an Äkta Pure FPLC system, using a Sephacryl column (HiPrep 26/60 Sephacryl S-200 (GE Healthcare). An isocratic gradient was used to elute the protein (20 mM sodium phosphate, 1 mM EDTA, 20 mM NaN_3 and 40 mM SDS). The dried, cleaved protein was taken up in 10 % SDS, sonicated for 15 minutes, mixed with same volume of running buffer and filtered through a 0.22 μM membrane. OD_{280} was typically between 5 and 9 for one condensation type M9 growth. Sample volume was 1 or 2 ml per run. Elution samples were analyses by SDS-PAGE and fractions of interest dialysed (1,000 MWCO membrane) extensively against H_2O until protein precipitated. Following lyophilisation, protein mass was checked by MS.

VII.6 Size-exclusion chromatography-multi-angle laser light scattering (SEC-MALS)

SEC (=FPLC) and SEC-MALS have been used to analyse the molecular weight of HCV J4 p7 and HCV EUH1480 p7 in Cyclofos-6 and Fos-Choline-12, respectively. SEC was performed with a Superdex200 column on an Äkta Pure FPLC system (GE Healthcare). SEC-MALS was performed with a Superose6 column on a Nexera HPLC (Shimadzu)

and a MALS DAWN HELEOS II and a refractive index Optilab T-rEX detector (Wyatt technology).

VII.7 Mass spectrometry

MS (mass spectrometry) was carried out at the Department of Biochemistry, using LC-ESI-MS (Liquid chromatography-electrospray ionisation-mass spectrometry), or in the Glycobiology Institute using MALDI-TOF (Matrix-assisted laser desorption/ionisation). For ESI, the protein samples in acetonitrile and water (1:1 + 0.1 % formic acid) were introduced at a flow rate of 10 μl/min by electrospray ionisation into a Micromass LCT orthogonal acceleration reflecting TOF mass spectrometer in positive ion mode. The instrument has been adapted specifically for protein work and the mass spectrometer was calibrated using myoglobin. The resultant m/z spectra were converted to mass spectra by using the Maximum Entropy analysis MaxEnt in the MassLynx suite of programmes. For MALDI, protein was taken up in 1 % FA and 40 % H_2O. The protein sample was mixed with CHCA (α-Cyano-4-hydroxycinnamic acid) (0.5 μL each) and dried down. The instrument was calibrated using ribonuclease A and insulin (Sigma-Aldrich). Expected protein masses were calculated with the ProtParam online tool (https://web.expasy.org/protparam) or the Scripps protein calculator (v3.3) (http://www.scripps.edu cdputnam/protcal.html), the latter was also used to determine weights of isotopically labelled proteins.

VII.8 Protein concentration determination

Protein concentrations were determined by measuring the absorbance at 280 nm, using a Nanodrop 1000 (Thermo scientific). The extinction coefficient of each protein was calculated using the ProtParam online tool (http://web.expasy.org/protparam).

VII.9 Sample reconstitution for NMR

Detergent evaluation for NMR samples was performed by preparing detergent solutions and adding them to lyophilised HCV p7. All NMR samples contained 5 % D_2O and DSS

(4,4-dimethyl-4-silapentane-1-sulfonic acid). The standard buffer used is 40 mM sodium phosphate in a pH range between 6.5 and 7.5. Detergents screened include DH7PC (1,2-diheptanoyl-sn-glycero-3-phosphocholine), Fos-choline-10, Fos-choline-12 (DPC), Fos-choline-14, SDS and Cyclofos-6. All detergents were purchased from Avanti polar lipids or Affymetrix. NMR samples were added to 350 μl Shigemi microtubes (Sigma-Aldrich).

VII.9.1 Thin film reconstitution

The detergent of choice was dissolved in a glass vial in an organic solvent, HFIP (1,1,1,3,3,3-hexaflouro-2-propanol). A thin film of Fos-choline-12 was produced under a weak nitrogen flow. The detergent thin film was re-dissolved in 400 μl of HFIP, and lyophilised HCV J4 p7 was added. Freeze-thaw cycles were used when the protein was not soluble at first. A second thin film containing detergent and protein was formed at the bottom of the glass tube under weak nitrogen flow. This process was repeated until a clear thin-film was visible. Lyophilisation was used to remove all remaining HFIP. 8 M Urea was added and the sample dialysed into NMR buffer (40 mM sodium phosphate). D_20 (5 %) was added after the dialysis.

VII.9.2 Guanidine reconstitution

Fos-choline 12 was dissolved in 6 M guanidine HCl and 40 mM sodium phosphate buffer (pH 6.5) followed by addition of HCV EUH1480 p7 (mt5). After a brief incubation, dialysis against 2 L of 40 mM sodium phosphate buffer (pH 6.5) using a 1,000 MWCO membrane removed the guanidine. D_2O (5 %) was added after the dialysis.

VII.10 NMR

All NMR spectra were recorded on NMR spectrometers with Oxford Instrument magnets with ^{1}H frequencies between 500 to 950 MHz, equipped with home-built triple resonance probes with triple axis gradients5 or with Bruker TCI CryoProbes with single Z-axis gradients (500 and 600 MHz). The Oxford Instrument magnet triple resonance probes allow for pulsing on up to four nuclei in a single experiment. TCI CryoProbes are equipped with cooled preamplifiers giving more sensitivity for double and triple

resonance NMR experiments[328]. The very low temperature caused by the helium cooling circuit around the tuning and matching circuits and the NMR coils reduce the Johnson-Nyquist noise, which is caused by the random thermal motion of electrons in conductors, leading to an improvement of signal-to-noise ratio of up to a factor of five (www.bruker.com). Experiments were unless otherwise stated performed at 37 °C and pH 6.5. NMR spectra were referenced in the direct dimension against DSS at 0 ppm (parts per million). The H_2O proton resonance frequency at 4.65 ppm was used as an internal proton reference for all experiments. NMR data was processed using NMRPipe and analysed using NMRDraw[329] and analysis[330]. Chemical shifts were assigned using the program CARA[298]. Experimental parameters for the main experiments can be found in Table VII.5.

VII.11 Droplet interface bilayers

For electrophysiological analysis, we used the method described by Lepthin *et al.* [316]. In brief, to create Droplet Interface Bilayers (DIBs) 10 mM DPhPC (1,2- Diphytanoyl-*sn*-glycero-3-phosphocholine, Avanti Polar Lipids) in hexadecane (C16, Sigma-Aldrich) was used as the lipid/oil solution. DPhPC monolayers spontaneously self-assemble around aqueous volumes immersed in this solution. Upon contact of two monolayers of two components, a bilayer is spontaneously formed. DIBs are formed by contacting aqueous droplets on an agarose layer. A plasma treated cover slip is covered by 1 % agarose, taken up in the final buffer solution (0.5 M KCl/5 mM Hepes/1 mM $CaCl_2$, pH 7.4). The device itself is cast with 2 % agarose, made up with the buffer solution. A stabilisation period of at least 15 minutes is required before contacting monolayers to prevent fusion. DIB bilayers are observed on an inverted microscope. A 100 μm diameter Ag/AgCl electrode is inserted into the droplets using a M3301L micromanipulator (World Precision Instruments), with a corresponding Ag/AgCl ground electrode in the agarose support10. A 1 mM stock solution in DMSO was diluted to a 10 nM working solution, using a DHB buffer (0.5 M KCl/5 mM Hepes/1 mM $CaCl_2$, pH 7.4.). WinEDR (SIPBS - University of Strathclyde) and pCLAMP (Molecular Devices) were used to collect and analyse data.

Table VII.5 – NMR experimental parameters

Experiment	^{1}H frequency [MHz]	Blocksize	rd [s]	Direct ^{1}H	Indirect ^{1}H	^{15}N	t1dw [μs]	^{15}Nppm	mixtime [s]
2D ^{1}H,^{15}N-HMQC (*)	950	1024	0.25	32 - 512	0	256	200	117.5	
2D ^{1}H,^{15}N-HMQC (**)	950	1024	0.25	32 - 512	0	256	200	120	
2D ^{1}H,^{15}N-HMQC (*)	750	1024	0.25	32 - 512	0	256	200	117.5	
2D ^{1}H,^{15}N-HMQC (**)	750	1024	0.25	32 - 512	0	256	120	200	
3D ^{15}N-edited NOESY-HSQC (*)	950	512	1.2	8	64	192	80	150	0.9
3D ^{15}N-edited NOESY-HSQC (*)	950	512	1.2	8	80	224	80	200	1.0
3D ^{15}N-edited TOCSY-HSQC (*)	950	512	1.2	8	64	192	80	150	90

(*)J4 p7 (C27S) in Cyclofos-6, (**)EUH1480 p7 (mt5) in Fos-choline-12
rd (recycle delay), sw (spectral width), t1dw - (t1 dwell time)

Bibliography

[1] Behzad Hajarizadeh, Jason Grebely, and Gregory J Dore. Epidemiology and natural history of HCV infection. *Nature reviews. Gastroenterology & hepatology*, 10 (9):553–562, September 2013.

[2] Leonard B Seeff. Natural history of chronic hepatitis C. *Hepatology*, 36(5 Suppl 1):S35–46, November 2002.

[3] Tatsunori Nakano, Gillian M G Lau, Grace M L Lau, Masaya Sugiyama, and Masashi Mizokami. An updated analysis of hepatitis C virus genotypes and subtypes based on the complete coding region. *Liver international : official journal of the International Association for the Study of the Liver*, 32(2):339–345, February 2012.

[4] S M Feinstone, A Z Kapikian, R H Purcell, H J Alter, and P V Holland. Transfusion-associated hepatitis not due to viral hepatitis type A or B. *The New England journal of medicine*, 292(15):767–770, April 1975.

[5] Q L Choo, G Kuo, A J Weiner, L R Overby, D W Bradley, and M Houghton. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, 244(4902):359–362, April 1989.

[6] E G Westaway, M A Brinton, SYa Gaidamovich, M C Horzinek, A Igarashi, L Kääriäinen, D K Lvov, J S Porterfield, P K Russell, and D W Trent. Flaviviridae. *Intervirology*, 24(4):183–192, 1985.

[7] Sinéad Lyons, Amit Kapoor, Colin Sharp, Bradley S Schneider, Nathan D Wolfe, Geoff Culshaw, Brendan Corcoran, Bruce C McGorum, and Peter Simmonds. Nonprimate hepaciviruses in domestic horses, United kingdom. *Emerging infectious diseases*, 18(12):1976–1982, December 2012.

[8] Peter D Burbelo, Edward J Dubovi, Peter Simmonds, Jan L Medina, Jose A Henriquez, Nischay Mishra, Jason Wagner, Rafal Tokarz, John M Cullen, Michael J Iadarola, Charles M. Rice, W Ian Lipkin, and Amit Kapoor. Serology-enabled discovery of genetically diverse hepaciviruses in a new host. *Journal of virology*, 86(11):6171–6178, June 2012.

[9] Peter Simmonds. The origin of hepatitis C virus. *Current topics in microbiology and immunology*, 369:1–15, 2013.

[10] Ruy M Ribeiro, Hui Li, Shuyi Wang, Mark B Stoddard, Gerald H Learn, Bette T Korber, Tanmoy Bhattacharya, Jeremie Guedj, Erica H Parrish, Beatrice H Hahn, George M Shaw, and Alan S Perelson. Quantifying the diversification of hepatitis C virus (HCV) during primary infection: estimates of the in vivo mutation rate. *PLoS Pathogens*, 8(8):e1002881, 2012.

[11] Adam S Lauring and Raul Andino. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathogens*, 6(7):e1001005, 2010.

[12] Darius Moradpour, François Penin, and Charles M. Rice. Replication of hepatitis C virus. *Nature reviews. Microbiology*, 5(6):453–463, June 2007.

[13] Ralf Bartenschlager, Volker Lohmann, and François Penin. The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. *Nature reviews. Microbiology*, 11(7):482–496, July 2013.

[14] Troels K H Scheel and Charles M. Rice. Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nature medicine*, 19(7):837–849, July 2013.

[15] Hsin-Chieh Ma, Ta-Wei Lin, Huichun Li, Sanae M M Iguchi-Ariga, Hiroyoshi Ariga, Yu-Li Chuang, Jing-Hsiung Ou, and Shih-Yen Lo. Hepatitis C virus ARFP/F protein interacts with cellular MM-1 protein and enhances the gene transactivation activity of c-Myc. *Journal of biomedical science*, 15(4):417–425, July 2008.

[16] Niki Vassilaki and Penelope Mavromara. The HCV ARFP/F/core+1 protein: production and functional analysis of an unconventional viral product. *IUBMB life*, 61(7):739–752, July 2009.

[17] E Santolini, G Migliaccio, and N La Monica. Biosynthesis and biochemical properties of the hepatitis C virus core protein. *Journal of virology*, 68(6):3631–3641, June 1994.

[18] Ralf Bartenschlager, François Penin, Volker Lohmann, and Patrice André. Assembly of infectious hepatitis C virus particles. *Trends in microbiology*, 19(2):401–412, February 2011.

[19] Kyoji Moriya, Hajime Fujie, Yoshizumi Shintani, Hiroshi Yotsuyanagi, Takeya Tsutsumi, Kotaro Ishibashi, Yoshiharu Matsuura, Satoshi Kimura, Tatsuo Miyamura, and Kazuhiko Koike. The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. *Nature medicine*, 4(9):1065–1067, September 1998.

[20] K Moriya, H Yotsuyanagi, Y Shintani, H Fujie, K Ishibashi, Y Matsuura, T. Miyamura, and K Koike. Hepatitis C virus core protein induces hepatic steatosis in transgenic mice. *The Journal of general virology*, 78 (Pt 7):1527–1531, July 1997.

[21] Anna Shavinskaya, Steeve Boulant, François Penin, John McLauchlan, and Ralf Bartenschlager. The lipid droplet binding domain of hepatitis C virus core protein is a major determinant for efficient virus assembly. *The Journal of biological chemistry*, 282(51):37158–37169, December 2007.

[22] Darius Moradpour and François Penin. Hepatitis C virus proteins: from structure to function. *Current topics in microbiology and immunology*, 369:113–142, 2013.

[23] J. McLauchlan. Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. *Journal of viral hepatitis*, 7(1):2–14, January 2000.

[24] John McLauchlan, Marius K Lemberg, Graham Hope, and Bruno Martoglio. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *The EMBO journal*, 21(15):3980–3988, August 2002.

[25] Verena Oehler, Ana Filipe, Roland Montserret, Daniel Da Costa, Gaie Brown, François Penin, and John McLauchlan. Structural analysis of hepatitis C virus core-E1 signal peptide and requirements for cleavage of the genotype 3a signal sequence by signal peptide peptidase. *Journal of virology*, 86(15):7818–7828, August 2012.

[26] N Mamiya and H J Worman. Hepatitis C virus core protein binds to a DEAD box RNA helicase. *The Journal of biological chemistry*, 274(22):15751–15756, May 1999.

[27] L R You, C M Chen, T S Yeh, T Y Tsai, R T Mai, C H Lin, and Y H Lee. Hepatitis C virus core protein interacts with cellular putative RNA helicase. *Journal of virology*, 73(4):2841–2853, April 1999.

[28] Gaël Cristofari, Roland Ivanyi-Nagy, Caroline Gabus, Steeve Boulant, Jean-Pierre Lavergne, François Penin, and Jean-Luc Darlix. The hepatitis C virus Core protein is a potent nucleic acid chaperone that directs dimerization of the viral (+) strand RNA in vitro. *Nucleic Acids Research*, 32(8):2623–2631, 2004.

[29] Kamal kant Sharma, Pascal Didier, Jean-Luc Darlix, Hugues de Rocquigny, Hayet Bensikaddour, Jean-Pierre Lavergne, François Penin, Jean-Marc Lessinger, and Yves Mély. Kinetic analysis of the nucleic acid chaperone activity of the hepatitis C virus core protein. *Nucleic Acids Research*, 38(11):3632–3642, June 2010.

[30] Roland Ivanyi-Nagy, Igor Kanevsky, Caroline Gabus, Jean-Pierre Lavergne, Damien Ficheux, François Penin, Philippe Fossé, and Jean-Luc Darlix. Analysis of hepatitis C virus RNA dimerization and core-RNA interactions. *Nucleic Acids Research*, 34(9):2618–2633, 2006.

[31] Jean-Baptiste Duvignaud, Christian Savard, Rémi Fromentin, Nathalie Majeau, Denis Leclerc, and Stéphane M. Gagné. Structure and dynamics of the N-terminal half of hepatitis C virus core protein: an intrinsically unstructured protein. *Biochemical and Biophysical Research Communications*, 378(1):27–31, January 2009.

[32] Steeve Boulant, Roland Montserret, R Graham Hope, Maxime Ratinier, Paul Targett-Adams, Jean-Pierre Lavergne, François Penin, and John McLauchlan. Structural determinants that target the hepatitis C virus core protein to lipid droplets. *The Journal of biological chemistry*, 281(31):22236–22247, August 2006.

[33] Jean Dubuisson, François Penin, and Darius Moradpour. Interaction of hepatitis C virus proteins with host cell membranes and lipids. *Trends in cell biology*, 12 (11):517–523, November 2002.

[34] Steeve Boulant, Christophe Vanbelle, Christine Ebel, François Penin, and Jean-Pierre Lavergne. Hepatitis C virus core protein is a dimeric alpha-helical protein

exhibiting membrane protein features. *Journal of virology*, 79(17):11353–11365, September 2005.

[35] K. Takeuchi, Y. Kubo, S. Boonmar, Y. Watanabe, T. Katayama, Q L Choo, G Kuo, M Houghton, I. Saito, and T. Miyamura. The putative nucleocapsid and envelope protein genes of hepatitis C virus determined by comparison of the nucleotide sequences of two isolates derived from an experimentally infected chimpanzee and healthy human carriers. *The Journal of general virology*, 71 (Pt 12): 3027–3033, December 1990.

[36] Cécile Voisset and Jean Dubuisson. Functional hepatitis C virus envelope glycoproteins. *Biology of the cell / under the auspices of the European Cell Biology Organization*, 96(6):413–420, August 2004.

[37] Mirjam B Zeisel, Daniel J Felmlee, and Thomas F Baumert. Hepatitis C virus entry. *Current topics in microbiology and immunology*, 369:87–112, 2013.

[38] A Op De Beeck, L. Cocquerel, and J. Dubuisson. Biogenesis of hepatitis C virus envelope glycoproteins. *The Journal of general virology*, 82(Pt 11):2589–2595, November 2001.

[39] Anne Goffard, Nathalie Callens, Birke Bartosch, Czeslaw Wychowski, François-Loïc Cosset, Claire Montpellier, and Jean Dubuisson. Role of N-linked glycans in the functions of hepatitis C virus envelope glycoproteins. *Journal of virology*, 79 (13):8400–8409, July 2005.

[40] Laurence Cocquerel, Anne Op de Beeck, Michel Lambot, Juliette Roussel, David Delgrange, André Pillez, Czeslaw Wychowski, François Penin, and Jean Dubuisson. Topological changes in the transmembrane domains of hepatitis C virus envelope glycoproteins. *The EMBO journal*, 21(12):2893–2902, June 2002.

[41] A Op De Beeck, R Montserret, S Duvet, L. Cocquerel, R Cacan, B Barberot, M Le Maire, F Penin, and J. Dubuisson. The transmembrane domains of hepatitis C virus envelope glycoproteins E1 and E2 play a major role in heterodimerization. *The Journal of biological chemistry*, 275(40):31428–31437, October 2000.

[42] Gabrielle Vieyres, Xavier Thomas, Véronique Descamps, Gilles Duverlie, Arvind H Patel, and Jean Dubuisson. Characterization of the envelope glycoproteins associated with infectious hepatitis C virus. *Journal of virology*, 84(19):10159–10168, October 2010.

[43] Maria Teresa Catanese, Kunihiro Uryu, Martina Kopp, Thomas J Edwards, Linda Andrus, William J Rice, Mariena Silvestry, Richard J Kuhn, and Charles M. Rice. Ultrastructural analysis of hepatitis C virus particles. *Proceedings of the National Academy of Sciences of the United States of America*, 110(23):9505–9510, June 2013.

[44] Leopold Kong, Erick Giang, Travis Nieusma, Rameshwar U Kadam, Kristin E Cogburn, Yuanzi Hua, Xiaoping Dai, Robyn L Stanfield, Dennis R Burton, Andrew B Ward, Ian A Wilson, and Mansun Law. Hepatitis C virus E2 envelope glycoprotein core structure. *Science*, 342(6162):1090–1094, November 2013.

[45] Abdul Ghafoor Khan, Jillian Whidby, Matthew T Miller, Hannah Scarborough, Alexandra V Zatorski, Alicja Cygan, Aryn A Price, Samantha A Yost, Caitlin D Bohannon, Joshy Jacob, Arash Grakoui, and Joseph Marcotrigiano. Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. *Nature*, February 2014.

[46] Anna Albecka, Roland Montserret, Thomas Krey, Alexander W Tarr, Eric Diesis, Jonathan K Ball, Véronique Descamps, Gilles Duverlie, Félix Rey, François Penin, and Jean Dubuisson. Identification of new functional regions in hepatitis C virus envelope glycoprotein E2. *Journal of virology*, 85(4):1777–1792, February 2011.

[47] Thomas Krey, Jacques d'Alayer, Carlos Kikuti, Aure Saulnier, Laurence Damier-Piolle, Isabelle Petitpas, Daniel Johansson, Rajiv Tawar, Bruno Baron, Bruno Robert, Patrick England, Mats Persson, Annette Martin, and Félix Rey. The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule. *PLoS Pathogens*, 6(2):e1000762, February 2010.

[48] A Grakoui, D W McCourt, C Wychowski, S M Feinstone, and C M Rice. A second hepatitis C virus-encoded proteinase. *Proceedings of the National Academy of Sciences of the United States of America*, 90(22):10583–10587, November 1993.

[49] Ivo C Lorenz, Joseph Marcotrigiano, Thomas G. Dentzer, and Charles M. Rice. Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature*, 442(7104):831–835, August 2006.

[50] V. Schregel, S. Jacobi, F Penin, and N. Tautz. Hepatitis C virus NS2 is a protease stimulated by cofactor domains in NS3. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13):5342–5347, March 2009.

[51] K E Reed, A Grakoui, and C M Rice. Hepatitis C virus-encoded NS2-3 protease: cleavage-site mutagenesis and requirements for bimolecular cleavage. *Journal of virology*, 69(7):4127–4136, July 1995.

[52] Vlastimil Jirasko, Roland Montserret, Nicole Appel, Anne Janvier, Leah Eustachi, Christiane Brohm, Eike Steinmann, Thomas Pietschmann, François Penin, and Ralf Bartenschlager. Structural and functional characterization of nonstructural protein 2 for its role in hepatitis C virus assembly. *The Journal of biological chemistry*, 283(42):28546–28562, October 2008.

[53] Vlastimil Jirasko, Roland Montserret, Ji Young Lee, Jérôme Gouttenoire, Darius Moradpour, François Penin, and Ralf Bartenschlager. Structural and functional studies of nonstructural protein 2 of the hepatitis C virus reveal its key role as organizer of virion assembly. *PLoS Pathogens*, 6(12):e1001233, 2010.

[54] Bertrand Boson, Ophélia Granio, Ralf Bartenschlager, and François-Loïc Cosset. A concerted action of hepatitis C virus p7 and nonstructural protein 2 regulates core localization at the endoplasmic reticulum and virus assembly. *PLoS Pathogens*, 7 (7):e1002144, July 2011.

[55] Kenneth A Stapleford and Brett D Lindenbach. Hepatitis C virus NS2 coordinates virus particle assembly through physical interactions with the E1-E2 glycoprotein and NS3-NS4A enzyme complexes. *Journal of virology*, 85(4):1706–1717, February 2011.

[56] Costin-Ioan Popescu, Nathalie Callens, Dave Trinel, Philippe Roingeard, Darius Moradpour, Véronique Descamps, Gilles Duverlie, François Penin, Laurent Héliot, Yves Rouillé, and Jean Dubuisson. NS2 protein of hepatitis C virus interacts with structural and non-structural proteins towards virus assembly. *PLoS Pathogens*, 7 (2):e1001278, 2011.

[57] Pasi Kaukinen, Maarit Sillanpää, Laura Nousiainen, Krister Melén, and Ilkka Julkunen. Hepatitis C virus NS2 protease inhibits host cell antiviral response by inhibiting IKKε and TBK1 functions. *Journal of Medical Virology*, 85(1):71–82, January 2013.

[58] Frances V Fuller-Pace. DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Research*, 34 (15):4206–4215, 2006.

[59] Laxmi Yeruva and KevinD Raney. Protein Displacement by Helicases. *Methods in molecular biology*, 587:85–98, 2010.

[60] N Yao, P Reichert, S S Taremi, W W Prosise, and P C Weber. Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure*, 7(11):1353–1363, November 1999.

[61] Volker Brass, Jan Martin Berke, Roland Montserret, Hubert E Blum, François Penin, and Darius Moradpour. Structural determinants for membrane association and dynamic organization of the hepatitis C virus NS3-4A complex. *Proceedings of the National Academy of Sciences of the United States of America*, 105(38): 14545–14550, September 2008.

[62] Kevin D Raney, Suresh D Sharma, Ibrahim M Moustafa, and Craig E. Cameron. Hepatitis C virus non-structural protein 3 (HCV NS3): a multifunctional antiviral target. *Journal of Biological Chemistry*, 285(30):22725–22731, July 2010.

[63] Ying He, Leiyun Weng, Rui Li, Li Li, Tetsuya Toyoda, and Jin Zhong. The N-terminal helix $\alpha(0)$ of hepatitis C virus NS3 protein dictates the subcellular localization and stability of NS3/NS4A complex. *Virology*, 422(2):214–223, January 2012.

[64] A Grakoui, D W McCourt, C Wychowski, S M Feinstone, and C M Rice. Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites. *Journal of virology*, 67(5):2832–2843, May 1993.

[65] R Bartenschlager, L Ahlborn-Laake, J Mous, and H Jacobsen. Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for

cleavage at the NS3/4 and NS4/5 junctions. *Journal of virology*, 67(7):3835–3844, July 1993.

[66] M R Eckart, M Selby, F Masiarz, C Lee, K Berger, K Crawford, C Kuo, G Kuo, M Houghton, and Q L Choo. The hepatitis C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. *Biochemical and Biophysical Research Communications*, 192(2): 399–406, April 1993.

[67] L Tomei, C Failla, E Santolini, R de Francesco, and N La Monica. NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *Journal of virology*, 67(7):4017–4026, July 1993.

[68] Rudolf K F Beran, Brett D Lindenbach, and Anna Marie Pyle. The NS4A protein of hepatitis C virus promotes RNA-coupled ATP hydrolysis by the NS3 helicase. *Journal of virology*, 83(7):3268–3275, April 2009.

[69] Tung Phan, Andrew Kohlway, Peniel Dimberu, Anna Marie Pyle, and Brett D Lindenbach. The acidic domain of hepatitis C virus NS4A contributes to RNA replication and virus particle assembly. *Journal of virology*, 85(3):1193–1204, February 2011.

[70] Yasuo Ariumi, Misao Kuroki, Ken-ichi Abe, Hiromichi Dansako, Masanori Ikeda, Takaji Wakita, and Nobuyuki Kato. DDX3 DEAD-box RNA helicase is required for hepatitis C virus RNA replication. *Journal of virology*, 81(24):13922–13926, December 2007.

[71] Yinghong Ma, Jeremy Yates, Yuqiong Liang, Stanley M Lemon, and MinKyung Yi. NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. *Journal of virology*, 82(15):7624–7639, August 2008.

[72] Marika Lundin, Magnus Monné, Anders Widell, Gunnar von Heijne, and Mats A A Persson. Topology of the membrane-associated hepatitis C virus protein NS4B. *Journal of virology*, 77(9):5428–5438, May 2003.

[73] M Francisca Palomares-Jerez, Henrique Nemesio, and José Villalaín. Interaction with membranes of the full C-terminal domain of protein NS4B from hepatitis C virus. *Biochimica et biophysica acta*, 1818(11):2536–2549, November 2012.

[74] Jérôme Gouttenoire, Philippe Roingeard, François Penin, and Darius Moradpour. Amphipathic alpha-helix AH2 is a major determinant for the oligomerization of hepatitis C virus nonstructural protein 4B. *Journal of virology*, 84(24):12529–12537, December 2010.

[75] Menashe Elazar, Ping Liu, Charles Rice, and Jeffrey S Glenn. An N-terminal amphipathic helix in hepatitis C virus (HCV) NS4B mediates membrane association, correct localization of replication complex proteins, and HCV RNA replication. *Journal of virology*, 78(20):11393–400, October 2004.

[76] Jason Aligo, Shuaizheng Jia, David Manna, and Kouacou V Konan. Formation and function of hepatitis C virus replication complexes require residues in the carboxy-terminal domain of NS4B protein. *Virology*, 393(1):68–83, October 2009.

[77] David Paul, Inés Romero-Brey, Jérôme Gouttenoire, Savina Stoitsova, Jacomine Krijnse-Locker, Darius Moradpour, and Ralf Bartenschlager. NS4B self-interaction through conserved C-terminal elements is required for the establishment of functional hepatitis C virus replication complexes. *Journal of virology*, 85(14):6963–6976, July 2011.

[78] Jérôme Gouttenoire, Valérie Castet, Roland Montserret, Naveen Arora, Vincent Raussens, Jean-Marie Ruysschaert, Eric Diesis, Hubert E Blum, François Penin, and Darius Moradpour. Identification of a novel determinant for membrane association in hepatitis C virus nonstructural protein 4B. *Journal of virology*, 83(12): 6257–6268, June 2009.

[79] Volker Lohmann. Hepatitis C virus RNA replication. *Current topics in microbiology and immunology*, 369:167–198, 2013.

[80] Guann-Yi Yu, Ki-Jeong Lee, Lu Gao, and Michael M C Lai. Palmitoylation and polymerization of hepatitis C virus NS4B protein. *Journal of virology*, 80(12): 6013–6023, June 2006.

[81] Denise Egger, Benno Wölk, Rainer Gosert, Leonardo Bianchi, Hubert E Blum, Darius Moradpour, and Kurt Bienz. Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *Journal of virology*, 76(12):5974–5984, June 2002.

[82] Inés Romero-Brey, Andreas Merz, Abhilash Chiramel, Ji Young Lee, Petr Chlanda, Uta Haselman, Rachel Santarella-Mellwig, Anja Habermann, Simone Hoppe, Stephanie Kallis, Paul Walther, Claude Antony, Jacomine Krijnse-Locker, and Ralf Bartenschlager. Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. *PLoS Pathogens*, 8(12): e1003056, 2012.

[83] Y Tanji, T Kaneko, S Satoh, and K Shimotohno. Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. *Journal of virology*, 69(7):3980–3986, July 1995.

[84] S I Asabe, Y Tanji, S Satoh, T Kaneko, K Kimura, and K Shimotohno. The N-terminal region of hepatitis C virus-encoded NS5A is important for NS4A-dependent phosphorylation. *Journal of virology*, 71(1):790–796, January 1997.

[85] Timothy L Tellinghuisen, Joseph Marcotrigiano, Alexander E Gorbalenya, and Charles M. Rice. The NS5A protein of hepatitis C virus is a zinc metalloprotein. *The Journal of biological chemistry*, 279(47):48576–48587, November 2004.

[86] Volker Brass, Elke Bieck, Roland Montserret, Benno Wölk, Jan Albert Hellings, Hubert E Blum, François Penin, and Darius Moradpour. An amino-terminal amphipathic alpha-helix mediates membrane association of the hepatitis C virus nonstructural protein 5A. *The Journal of biological chemistry*, 277(10):8130–8139, March 2002.

[87] François Penin, Volker Brass, Nicole Appel, Stephanie Ramboarina, Roland Montserret, Damien Ficheux, Hubert Blum, Ralf Bartenschlager, and Darius Moradpour. Structure and function of the membrane anchor domain of hepatitis C virus nonstructural protein 5A. *The Journal of biological chemistry*, 279(39): 40835–43, September 2004.

[88] Yusuke Miyanari, Kimie Atsuzawa, Nobuteru Usuda, Koichi Watashi, Takayuki Hishiki, Margarita Zayas, Ralf Bartenschlager, Takaji Wakita, Makoto Hijikata, and Kunitada Shimotohno. The lipid droplet is an important organelle for hepatitis C virus production. *Nature cell biology*, 9(9):1089–1097, September 2007.

[89] Stephanie T. Shi, Stephen J. Polyak, Hong Tu, Deborah R. Taylor, David R. Gretch, and Michael M C Lai. Hepatitis C virus NS5A colocalizes with the core

protein on lipid droplets and interacts with apolipoproteins. *Virology*, 292(2): 198–210, January 2002.

[90] Timothy L Tellinghuisen, Katie L Foss, Jason C Treadaway, and Charles M. Rice. Identification of residues required for RNA replication in domains II and III of the hepatitis C virus NS5A protein. *Journal of virology*, 82(3):1073–1083, February 2008.

[91] Takahiro Masaki, Ryosuke Suzuki, Kyoko Murakami, Hideki Aizaki, Koji Ishii, Asako Murayama, Tomoko Date, Yoshiharu Matsuura, Tatsuo Miyamura, Takaji Wakita, and Tetsuro Suzuki. Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. *Journal of virology*, 82(16):7964–7976, August 2008.

[92] Timothy L Tellinghuisen, Joseph Marcotrigiano, and Charles M. Rice. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. *Nature*, 435(7040):374–379, May 2005.

[93] Robert A Love, Oleg Brodsky, Michael J Hickey, Peter A Wells, and Ciarán N Cronin. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. *Journal of virology*, 83(9):4395–4403, May 2009.

[94] Timothy L Tellinghuisen, Katie L Foss, and Jason Treadaway. Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. *PLoS Pathogens*, 4(3):e1000032, March 2008.

[95] Manuela Quintavalle, Sonia Sambucini, Chiara Di Pietro, Raffaele de Francesco, and Petra Neddermann. The alpha isoform of protein kinase CKI is responsible for hepatitis C virus NS5A hyperphosphorylation. *Journal of virology*, 80(22): 11305–11312, November 2006.

[96] Manuela Quintavalle, Sonia Sambucini, Vincenzo Summa, Laura Orsatti, Fabio Talamo, Raffaele de Francesco, and Petra Neddermann. 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. *The Journal of biological chemistry*, 282(8):5536–5544, February 2007.

[97] Matthew J Evans, Charles M. Rice, and Stephen P Goff. Phosphorylation of hepatitis C virus nonstructural protein 5A modulates its protein interactions and

viral RNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, 101(35):13038–13043, August 2004.

[98] Nicole Appel, Thomas Pietschmann, and Ralf Bartenschlager. Mutational analysis of hepatitis C virus nonstructural protein 5A: potential role of differential phosphorylation in RNA replication and identification of a genetically flexible domain. *Journal of virology*, 79(5):3187–3194, March 2005.

[99] S E Behrens, L Tomei, and R de Francesco. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *The EMBO journal*, 15(1): 12–22, January 1996.

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

[101] Robert A Love, Hans E Parge, Xiu Yu, Michael J Hickey, Wade Diehl, Jingjin Gao, Hilary Wriggers, Anne Ekker, Liann Wang, James A Thomson, Peter S Dragovich, and Shella A Fuhrman. Crystallographic identification of a noncompetitive inhibitor binding site on the hepatitis C virus NS5B RNA polymerase enzyme. *Journal of virology*, 77(13):7575–7581, July 2003.

[102] S. Bressanelli, L Tomei, A. Roussel, I. Incitti, R. L. Vitale, M. Mathieu, R de Francesco, and F. A. Rey. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, 96(23):13034–13039, November 1999.

[103] H Ago, T Adachi, A Yoshida, M Yamamoto, N Habuka, K Yatsunami, and M Miyano. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure*, 7(11):1417–1426, November 1999.

[104] Kyung H Choi and Michael G Rossmann. RNA-dependent RNA polymerases from Flaviviridae. *Current opinion in structural biology*, 19(6):746–751, December 2009.

[105] Michael Niepmann. Hepatitis C virus RNA translation. *Current topics in microbiology and immunology*, 369:143–166, 2013.

[106] Brett D Lindenbach. Virion assembly and release. *Current topics in microbiology and immunology*, 369:199–218, 2013.

[107] Pablo Gastaminza, Kelly A Dryden, Bryan Boyd, Malcolm R Wood, Mansun Law, Mark Yeager, and Francis V Chisari. Ultrastructural and biophysical characterization of hepatitis C virus particles produced in cell culture. *Journal of virology*, 84(21):10999–11009, November 2010.

[108] Tetsuro Suzuki. Assembly of hepatitis C virus particles. *Microbiology and immunology*, 55(1):12–18, January 2011.

[109] R Thomssen, S Bonk, C Propfe, K H Heermann, H G Köchel, and A Uy. Association of hepatitis C virus in human sera with beta-lipoprotein. *Medical microbiology and immunology*, 181(5):293–300, 1992.

[110] Søren U Nielsen, Margaret F Bassendine, Alastair D Burt, Caroline Martin, Wanna Pumeechockchai, and Geoffrey L Toms. Association between hepatitis C virus and very-low-density lipoprotein (VLDL)/LDL analyzed in iodixanol density gradients. *Journal of virology*, 80(5):2418–2428, March 2006.

[111] Brett D Lindenbach, Matthew J Evans, Andrew J Syder, Benno Wölk, Timothy L Tellinghuisen, Christopher C Liu, Toshiaki Maruyama, Richard O Hynes, Dennis R Burton, Jane A McKeating, and Charles M. Rice. Complete replication of hepatitis C virus in cell culture. *Science*, 309(5734):623–626, July 2005.

[112] Brett D Lindenbach, Philip Meuleman, Alexander Ploss, Thomas Vanwolleghem, Andrew J Syder, Jane A McKeating, Robert E Lanford, Stephen M Feinstone, Marian E Major, Geert Leroux-Roels, and Charles M. Rice. Cell culture-grown hepatitis C virus is infectious in vivo and can be recultured in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 103(10): 3805–3809, March 2006.

[113] Pablo Gastaminza, Sharookh B Kapadia, and Francis V Chisari. Differential biophysical properties of infectious intracellular and secreted hepatitis C virus particles. *Journal of virology*, 80(22):11074–11081, November 2006.

[114] Eva Herker and Melanie Ott. Unique ties between hepatitis C virus replication and intracellular lipids. *Trends in endocrinology and metabolism: TEM*, 22(6): 241–248, June 2011.

[115] P André, F Komurian-Pradel, S Deforges, M Perret, J L Berland, M Sodoyer, S Pol, C Bréchot, G Paranhos-Baccalà, and V Lotteau. Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. *Journal of virology*, 76(14):6919–6928, July 2002.

[116] Andreas Merz, Gang Long, Marie-Sophie Hiet, Britta Brügger, Petr Chlanda, Patrice André, Felix Wieland, Jacomine Krijnse-Locker, and Ralf Bartenschlager. Biochemical and morphological properties of hepatitis C virus particles and determination of their lipidome. *The Journal of biological chemistry*, 286(4):3018–3032, January 2011.

[117] Birke Bartosch, Alessandra Vitelli, Christelle Granier, Caroline Goujon, Jean Dubuisson, Simona Pascale, Elisa Scarselli, Riccardo Cortese, Alfredo Nicosia, and François-Loïc Cosset. Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *The Journal of biological chemistry*, 278(43):41624–41630, October 2003.

[118] Emmanuel G Cormier, Fay Tsamis, Francis Kajumo, Robert J Durso, Jason P Gardner, and Tatjana Dragic. CD81 is an entry coreceptor for hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, 101(19):7270–7274, May 2004.

[119] Matthew J Evans, Thomas von Hahn, Donna M Tscherne, Andrew J Syder, Maryline Panis, Benno Wölk, Theodora Hatziioannou, Jane A McKeating, Paul D Bieniasz, and Charles M. Rice. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature*, 446(7137):801–805, April 2007.

[120] Alexander Ploss, Matthew J Evans, Valeriya A Gaysinskaya, Maryline Panis, Hana You, Ype P de Jong, and Charles M. Rice. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature*, 457(7231):882–886, February 2009.

[121] Luke W Meredith, Garrick K Wilson, Nicola F Fletcher, and Jane A McKeating. Hepatitis C virus entry: beyond receptors. *Reviews in medical*, 22(3):182–193, January 2012.

[122] Bruno Sainz, Naina Barretto, Danyelle N Martin, Nobuhiko Hiraga, Michio Imamura, Snawar Hussain, Katherine A Marsh, Xuemei Yu, Kazuaki Chayama, Waddah A Alrefai, and Susan L Uprichard. Identification of the Niemann-Pick C1-like 1 cholesterol absorption receptor as a new hepatitis C virus entry factor. *Nature medicine*, 18(2):281–285, February 2012.

[123] Jennifer M Timpe, Zania Stamataki, Adam Jennings, Ke Hu, Michelle J Farquhar, Helen J Harris, Anne Schwarz, Isabelle Desombere, Geert Leroux Roels, Peter Balfe, and Jane A McKeating. Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. *Hepatology*, 47(1):17–24, January 2008.

[124] Emmanuelle Blanchard, Sandrine Belouzard, Lucie Goueslain, Takaji Wakita, Jean Dubuisson, Czeslaw Wychowski, and Yves Rouillé. Hepatitis C virus entry depends on clathrin-mediated endocytosis. *Journal of virology*, 80(14):6964–6972, July 2006.

[125] Brett Hoffman and Qiang Liu. Hepatitis C viral protein translation: mechanisms and implications in developing antivirals. *Liver international : official journal of the International Association for the Study of the Liver*, 31(10):1449–1467, November 2011.

[126] Gyongyi Szabo and Shashi Bala. MicroRNAs in liver disease. *Nature reviews. Gastroenterology & hepatology*, 10(9):542–552, September 2013.

[127] Catherine L Jopling, MinKyung Yi, Alissa M Lancaster, Stanley M Lemon, and Peter Sarnow. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science*, 309(5740):1577–1581, September 2005.

[128] Jura Inga Henke, Dagmar Goergen, Junfeng Zheng, Yutong Song, Christian G Schüttler, Carmen Fehr, Christiane Jünemann, and Michael Niepmann. microRNA-122 stimulates translation of hepatitis C virus RNA. *The EMBO journal*, 27(24):3300–3310, December 2008.

[129] Rohit K Jangra, MinKyung Yi, and Stanley M Lemon. Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122. *Journal of virology*, 84(13):6615–6625, July 2010.

[130] You Li, Takahiro Masaki, Daisuke Yamane, David R McGivern, and Stanley M Lemon. Competing and noncompeting activities of miR-122 and the 5' exonuclease Xrn1 in regulation of hepatitis C virus replication. *Proceedings of the National Academy of Sciences of the United States of America*, 110(5):1881–1886, January 2013.

[131] Tetsuro Shimakami, Daisuke Yamane, Rohit K Jangra, Brian J Kempf, Carolyn Spaniel, David J Barton, and Stanley M Lemon. Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex. *Proceedings of the National Academy of Sciences of the United States of America*, 109(3):941–946, January 2012.

[132] G Fischer, B Wittmann-Liebold, K Lang, T Kiefhaber, and F X Schmid. Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins. *Nature*, 337(6206):476–478, February 1989.

[133] Xavier Hanoulle, Aurélie Badillo, Jean-Michel Wieruszeski, Dries Verdegem, Isabelle Landrieu, Ralf Bartenschlager, François Penin, and Guy Lippens. Hepatitis C virus NS5A protein is a substrate for the peptidyl-prolyl cis/trans isomerase activity of cyclophilins A and B. *The Journal of biological chemistry*, 284(20): 13589–13601, May 2009.

[134] Dries Verdegem, Aurélie Badillo, Jean-Michel Wieruszeski, Isabelle Landrieu, Arnaud Leroy, Ralf Bartenschlager, François Penin, Guy Lippens, and Xavier Hanoulle. Domain 3 of NS5A protein from the hepatitis C virus has intrinsic alpha-helical propensity and is a substrate of cyclophilin A. *Journal of Biological Chemistry*, 286(23):20441–20454, June 2011.

[135] Claire Rosnoblet, Bernd Fritzinger, Dominique Legrand, Hélène Launay, Jean-Michel Wieruszeski, Guy Lippens, and Xavier Hanoulle. Hepatitis C virus NS5B and host cyclophilin A share a common binding site on NS5A. *The Journal of biological chemistry*, 287(53):44249–44260, December 2012.

[136] Kazuya Ogawa, Takayuki Hishiki, Yuko Shimizu, Kenji Funami, Kazuo Sugiyama, Yusuke Miyanari, and Kunitada Shimotohno. Hepatitis C virus utilizes lipid droplet for production of infectious virus. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, 85(7):217–228, 2009.

[137] Nicole Appel, Margarita Zayas, Sven Miller, Jacomine Krijnse-Locker, Torsten Schaller, Peter Friebe, Stephanie Kallis, Ulrike Engel, and Ralf Bartenschlager. Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. *PLoS Pathogens*, 4(3):e1000035, March 2008.

[138] Eike Steinmann and Thomas Pietschmann. Hepatitis C virus p7-a viroporin crucial for virus assembly and an emerging target for antiviral therapy. *Viruses*, 2(9): 2078–2095, September 2010.

[139] Christopher T Jones, Catherine L Murray, Dawnnica K Eastman, Jodie Tassello, and Charles M. Rice. Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *Journal of virology*, 81(16):8374–8383, August 2007.

[140] Akito Sakai, Marisa St Claire, Kristina Faulk, Sugantha Govindarajan, Suzanne U Emerson, Robert H Purcell, and Jens Bukh. The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 100(20):11646–11651, September 2003.

[141] Eike Steinmann, François Penin, Stephanie Kallis, Arvind H Patel, Ralf Bartenschlager, and Thomas Pietschmann. Hepatitis C virus p7 protein is crucial for assembly and release of infectious virions. *PLoS Pathogens*, 3(7):e103, July 2007.

[142] Wagane J A Benga, Sophie E Krieger, Maria Dimitrova, Mirjam B Zeisel, Marie Parnot, Joachim Lupberger, Eberhard Hildt, Guangxiang Luo, John McLauchlan, Thomas F Baumert, and Catherine Schuster. Apolipoprotein E interacts with hepatitis C virus nonstructural protein 5A and determines assembly of infectious particles. *Hepatology*, 51(1):43–53, January 2010.

[143] Hua Huang, Fang Sun, David M Owen, Weiping Li, Yan Chen, Michael Gale, and Jin Ye. Hepatitis C virus production by human hepatocytes dependent on assembly and secretion of very low-density lipoproteins. *Proceedings of the National Academy of Sciences of the United States of America*, 104(14):5848–5853, April 2007.

[144] Pablo Gastaminza, Guofeng Cheng, Stefan Wieland, Jin Zhong, Wei Liao, and Francis V Chisari. Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *Journal of virology*, 82(5):2120–2129, March 2008.

[145] Daniel M Jones and John McLauchlan. Hepatitis C virus: assembly and release of virus particles. *Journal of Biological Chemistry*, 285(30):22733–22739, July 2010.

[146] Stephen M Feinstone, Dale J Hu, and Marian E Major. Prospects for prophylactic and therapeutic vaccines against hepatitis C virus. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 55 Suppl 1: S25–32, July 2012.

[147] T Poynard, V Leroy, M Cohard, T Thevenot, P Mathurin, P Opolon, and J P Zarski. Meta-analysis of interferon randomized trials in the treatment of viral hepatitis C: effects of dose and duration. *Hepatology*, 24(4):778–789, October 1996.

[148] Stacy M Horner and Michael Gale. Regulation of hepatic innate immunity by hepatitis C virus. *Nature medicine*, 19(7):879–888, July 2013.

[149] Jean-Michel Pawlotsky. Treatment of Chronic Hepatitis C: Current and Future. *Current topics in microbiology and immunology*, 369:321–342, 2013.

[150] Christoph Welsch, Arun Jesudian, Stefan Zeuzem, and Ira Jacobson. New direct-acting antiviral agents for the treatment of hepatitis C virus infection and perspectives. *Gut*, 61 Suppl 1:i36–46, May 2012.

[151] Jordan J Feld and Jay H Hoofnagle. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature*, 436(7053):967–972, August 2005.

[152] Leonidas C Platanias. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nature reviews. Immunology*, 5(5):375–386, May 2005.

[153] Jason D. Graci and Craig E. Cameron. Mechanisms of action of ribavirin against distinct viruses. *Reviews in medical virology*, 16(1):37–48, January 2006.

[154] Leen Delang, Johan Neyts, Inge Vliegen, Sergio Abrignani, Petra Neddermann, and Raffaele de Francesco. Hepatitis C virus-specific directly acting antiviral drugs. *Current topics in microbiology and immunology*, 369:289–320, 2013.

[155] Srikanth Venkatraman, Stéphane L. Bogen, Ashok Arasappan, Frank Bennett, Kevin Chen, Edwin Jao, Yi-Tsung Liu, Raymond Lovey, Siska Hendrata, Yuhua Huang, Weidong Pan, Tejal Parekh, Patrick Pinto, Veljko Popov,

Russel Pike, Sumei Ruan, Bama Santhanam, Bancha Vibulbhan, Wanli Wu, Weiying Yang, Jianshe Kong, Xiang Liang, Jesse Wong, Rong Liu, Nancy Butkiewicz, Robert Chase, Andrea Hart, Sony Agrawal, Paul Ingravallo, John Pichardo, Rong Kong, Bahige Baroudy, Bruce Malcolm, Zhuyan Guo, Andrew Prongay, Vincent Madison, Lisa Broske, Xiaoming Cui, Kuo-Chi Cheng, Yunsheng Hsieh, Jean-Marc Brisson, Danial Prelusky, Walter Korfmacher, Ronald White, Susan Bogdanowich-Knipp, Anastasia Pavlovsky, Prudence Bradley, Anil K. Saksena, Ashit Ganguly, John Piwinski, Viyyoor Girijavallabhan, and F George Njoroge. Discovery of (1R,5S)-N-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]- 3-[2(S)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]- 6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(S)-carboxamide (SCH 503034), a selective, potent, orally bioavailable hepatitis C virus NS3 protease inhibitor: a potential therapeutic agent for the treatment of hepatitis C infection. *Journal of medicinal chemistry*, 49(20):6074–6086, October 2006.

[156] Kai Lin, Robert B Perni, Ann D Kwong, and Chao Lin. VX-950, a novel hepatitis C virus (HCV) NS3-4A protease inhibitor, exhibits potent antiviral activities in HCv replicon cells. *Antimicrobial agents and chemotherapy*, 50(5):1813–1822, May 2006.

[157] Nobuhiko Hiraga, Michio Imamura, Hiromi Abe, C Nelson Hayes, Tomohiko Kono, Mayu Onishi, Masataka Tsuge, Shoichi Takahashi, Hidenori Ochi, Eiji Iwao, Naohiro Kamiya, Ichimaro Yamada, Chise Tateno, Katsutoshi Yoshizato, Hirotaka Matsui, Akinori Kanai, Toshiya Inaba, Shinji Tanaka, and Kazuaki Chayama. Rapid emergence of telaprevir resistant hepatitis C virus strain from wildtype clone in vivo. *Hepatology*, 54(3):781–788, September 2011.

[158] Simone Susser, Christoph Welsch, Yalan Wang, Markus Zettler, Francisco S Domingues, Ursula Karey, Eric Hughes, Robert Ralston, Xiao Tong, Eva Herrmann, Stefan Zeuzem, and Christoph Sarrazin. Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus-infected patients. *Hepatology*, 50(6):1709–1718, December 2009.

[159] Dimitrios Vlachakis, Vassiliki Lila Koumandou, and Sophia Kossida. A holistic evolutionary and structural study of flaviviridae provides insights into the function and inhibition of HCV helicase. *PeerJ*, 1:e74, 2013.

[160] Oscar Belda and Paul Targett-Adams. Small molecule inhibitors of the hepatitis C virus-encoded NS5A protein. *Virus Research*, 170(1-2):1–14, December 2012.

[161] Paul Targett-Adams, Emily J S Graham, Jenny Middleton, Amy Palmer, Stephen M Shaw, Helen Lavender, Philip Brain, Thien Duc Tran, Lyn H Jones, Florian Wakenhut, Blanda Stammen, David Pryde, Chris Pickford, and Mike Westby. Small molecules targeting hepatitis C virus-encoded NS5A cause subcellular redistribution of their target: insights into compound modes of action. *Journal of virology*, 85(13):6353–6368, July 2011.

[162] Maria Eugenia Gonzalez and Luis Carrasco. Viroporins. *FEBS letters*, 552(1): 28–34, September 2003.

[163] José Luis Nieva, Vanesa Madan, and Luis Carrasco. Viroporins: structure and biological functions. *Nature reviews. Microbiology*, 10(8):563–574, August 2012.

[164] Bo OuYang and James J Chou. The minimalist architectures of viroporins and their therapeutic implications. *Biochimica et biophysica acta*, September 2013.

[165] R A Lamb, S L Zebedee, and C D Richardson. Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. *Cell*, 40(3): 627–633, March 1985.

[166] I V Chizhmakov, F M Geraghty, D C Ogden, A Hayhurst, M Antoniou, and A J Hay. Selective proton permeability and pH regulation of the influenza virus M2 channel expressed in mouse erythroleukaemia cells. *The Journal of physiology*, 494 (Pt 2):329–336, July 1996.

[167] J A Mould, H C Li, C S Dudlak, J D Lear, A Pekosz, R A Lamb, and L H Pinto. Mechanism for proton conduction of the M(2) ion channel of influenza A virus. *The Journal of biological chemistry*, 275(12):8592–8599, March 2000.

[168] K Strebel, T Klimkait, and M A Martin. A novel gene of HIV-1, vpu, and its 16-kilodalton product. *Science*, 241(4870):1221–1223, September 1988.

[169] E A Cohen, E F Terwilliger, J G Sodroski, and W A Haseltine. Identification of a protein encoded by the vpu gene of HIV-1. *Nature*, 334(6182):532–534, August 1988.

[170] U Schubert, A V Ferrer-Montiel, M Oblatt-Montal, P Henklein, K Strebel, and M Montal. Identification of an ion channel activity of the Vpu transmembrane domain and its involvement in the regulation of virus release from HIV-1-infected cells. *FEBS letters*, 398(1):12–18, November 1996.

[171] Jason R Schnell and James J Chou. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature*, 451(7178):591–595, January 2008.

[172] Amanda L Stouffer, Rudresh Acharya, David Salom, Anna S Levine, Luigi Di Costanzo, Cinque S Soto, Valentina Tereshko, Vikas Nanda, Steven Stayrook, and William F Degrado. Structural basis for the function and inhibition of an influenza virus proton channel. *Nature*, 451(7178):596–599, January 2008.

[173] Sang Ho Park, Anthony A Mrse, Alexander A Nevzorov, Michael F Mesleh, Myrta Oblatt-Montal, Mauricio Montal, and Stanley J Opella. Three-dimensional structure of the channel-forming trans-membrane domain of virus protein "u" (Vpu) from HIV-1. *Journal of molecular biology*, 333(2):409–424, October 2003.

[174] Nicole M Bouvier and Peter Palese. The biology of influenza viruses. *Vaccine*, 26 Suppl 4:D49–53, September 2008.

[175] L H Pinto, L J Holsinger, and R A Lamb. Influenza virus M2 protein has ion channel activity. *Cell*, 69(3):517–528, May 1992.

[176] Ruo-Xu Gu, Limin Angela Liu, and Dong-Qing Wei. Structural and energetic analysis of drug inhibition of the influenza A M2 proton channel. *Trends in pharmacological sciences*, 34(10):571–580, October 2013.

[177] J M Thomas, M P Stevens, N Percy, and W S Barclay. Phosphorylation of the M2 protein of influenza A virus is not essential for virus viability. *Virology*, 252 (1):54–64, December 1998.

[178] L J Holsinger, M A Shaughnessy, A Micko, L H Pinto, and R A Lamb. Analysis of the posttranslational modifications of the influenza virus M2 protein. *Journal of virology*, 69(2):1219–1225, February 1995.

[179] Matthew F McCown and Andrew Pekosz. The influenza A virus M2 cytoplasmic tail is required for infectious virus production and efficient genome packaging. *Journal of virology*, 79(6):3595–3605, March 2005.

[180] Matthew F McCown and Andrew Pekosz. Distinct domains of the influenza a virus M2 protein cytoplasmic tail mediate binding to the M1 protein and facilitate infectious virus production. *Journal of virology*, 80(16):8178–8189, August 2006.

[181] T Sakaguchi, Q Tu, L H Pinto, and R A Lamb. The active oligomeric state of the minimalistic influenza virus M2 ion channel is a tetramer. *Proceedings of the National Academy of Sciences of the United States of America*, 94(10):5000–5005, May 1997.

[182] Rafal M Pielak and James J Chou. Influenza M2 proton channels. *Biochimica et biophysica acta*, 1808(2):522–529, February 2011.

[183] Nathan W Schmidt, Abhijit Mishra, Jun Wang, William F Degrado, and Gerard C L Wong. Influenza virus A M2 protein generates negative Gaussian membrane curvature necessary for budding and scission. *Journal of the American Chemical Society*, 135(37):13710–13719, September 2013.

[184] Thom Leiding, Jun Wang, Jonas Martinsson, William F Degrado, and Sindra Peterson Arsköld. Proton and cation transport activity of the M2 proton channel from influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(35):15409–15414, August 2010.

[185] Tijana Ivanovic, Rutger Rozendaal, Daniel L Floyd, Milos Popovic, Antoine M van Oijen, and Stephen C Harrison. Kinetics of proton transport into influenza virions by the viral M2 channel. *PLOS ONE*, 7(3):e31566, 2012.

[186] J Wang, S Kim, F Kovacs, and T A Cross. Structure of the transmembrane region of the M2 protein H(+) channel. *Protein science : a publication of the Protein Society*, 10(11):2241–2250, November 2001.

[187] Shu Yu Liao, Keith J Fritzsching, and Mei Hong. Conformational analysis of the full-length M2 protein of the influenza A virus using solid-state NMR. *Protein science : a publication of the Protein Society*, 22(11):1623–1638, November 2013.

[188] Mei Hong and William F Degrado. Structural basis for proton conduction and inhibition by the influenza M2 protein. *Protein science : a publication of the Protein Society*, 21(11):1620–1633, November 2012.

[189] Hadas Leonov and Isaiah T Arkin. Structure and dynamics of the influenza A M2 channel: a comparison of three structures. *Journal of molecular modeling*, 15(11): 1317–1328, November 2009.

[190] Rudresh Acharya, Vincenzo Carnevale, Giacomo Fiorin, Benjamin G Levine, Alexei L Polishchuk, Victoria Balannik, Ilan Samish, Robert A Lamb, Lawrence H Pinto, William F Degrado, and Michael L Klein. Structure and mechanism of proton transport through the transmembrane tetrameric M2 protein bundle of the influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(34):15075–15080, August 2010.

[191] Sarah D Cady, Klaus Schmidt-Rohr, Jun Wang, Cinque S Soto, William F Degrado, and Mei Hong. Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature*, 463(7281):689–692, February 2010.

[192] Timothy A Cross, Hao Dong, Mukesh Sharma, David D Busath, and Huan-Xiang Zhou. M2 protein from influenza A: from multiple structures to biophysical and functional insights. *Current opinion in virology*, 2(2):128–133, April 2012.

[193] Rafal M Pielak and James J Chou. Solution NMR structure of the V27A drug resistant mutant of influenza A M2 channel. *Biochemical and Biophysical Research Communications*, 401(1):58–63, October 2010.

[194] Mukesh Sharma, Myunggi Yi, Hao Dong, Huajun Qin, Emily Peterson, David D Busath, Huan-Xiang Zhou, and Timothy A Cross. Insight into the mechanism of the influenza A proton channel from a structure in a lipid bilayer. *Science*, 330 (6003):509–512, October 2010.

[195] Mathieu Dubé, Mariana G Bego, Catherine Paquay, and Éric A Cohen. Modulation of HIV-1-host interaction: role of the Vpu accessory protein. *Retrovirology*, 7:114, 2010.

[196] Autumn Ruiz, John C Guatelli, and Edward B Stephens. The Vpu protein: new concepts in virus release and CD4 down-modulation. *Current HIV research*, 8(3): 240–52, April 2010.

[197] Mark Skasko, Yan Wang, Ye Tian, Andrey Tokarev, Jason Munguia, Autumn Ruiz, Edward B Stephens, Stanley J Opella, and John Guatelli. HIV-1 Vpu

protein antagonizes innate restriction factor BST-2 via lipid-embedded helix-helix interactions. *Journal of Biological Chemistry*, 287(1):58–67, January 2012.

[198] F S Cordes, A D Tustian, M S P Sansom, A Watts, and W B Fischer. Bundles consisting of extended transmembrane segments of Vpu from HIV-1: computer simulations and conductance measurements. *Biochemistry*, 41(23):7359–7365, June 2002.

[199] T Mehnert, A Routh, P J Judge, Y H Lam, D Fischer, A Watts, and W B Fischer. Biophysical characterization of Vpu from HIV-1 suggests a channel-pore dualism. *Proteins*, 70(4):1488–1497, March 2008.

[200] Sang Ho Park, Anna A de Angelis, Alexander A Nevzorov, Chin H Wu, and Stanley J Opella. Three-dimensional structure of the transmembrane domain of Vpu from HIV-1 in aligned phospholipid bicelles. *Biophysical journal*, 91(8):3032–3042, October 2006.

[201] V Lemaitre, D Willbold, A Watts, and W B Fischer. Full length Vpu from HIV-1: combining molecular dynamics simulations with NMR spectroscopy. *Journal of biomolecular structure & dynamics*, 23(5):485–496, April 2006.

[202] A J Hay, A J Wolstenholme, J J Skehel, and M H Smith. The molecular basis of the specific anti-influenza action of amantadine. *The EMBO journal*, 4(11): 3021–3024, November 1985.

[203] R B Belshe, B Burk, F Newman, R L Cerruti, and I S Sim. Resistance of influenza A virus to amantadine and rimantadine: results of one decade of surveillance. *The Journal of infectious diseases*, 159(3):430–435, March 1989.

[204] Stephen D C Griffin, Lucy P Beales, Dean S Clarke, Oliver Worsfold, Stephen D Evans, Joachim Jaeger, Mark P G Harris, and David J Rowlands. The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS letters*, 535(1-3):34–38, January 2003.

[205] Corine Stgelais, Tobias J Tuthill, Dean S Clarke, David J Rowlands, Mark Harris, and Stephen Griffin. Inhibition of hepatitis C virus p7 membrane channels in a liposome-based assay system. *Antiviral research*, 76(1):48–58, October 2007.

[206] Stephen Griffin, Corine Stgelais, Ania M Owsianka, Arvind H Patel, David Rowlands, and Mark Harris. Genotype-dependent sensitivity of hepatitis C virus to inhibitors of the p7 ion channel. *Hepatology*, 48(6):1779–1790, December 2008.

[207] Hanneke van Soest, Peter J van der Schaar, Ger H Koek, Richard A de Vries, Nancy A van Ooteghem, Bart van Hoek, Joost P H Drenth, Jan M Vrolijk, Rob J Lieverse, Peter Houben, Annet van der Sluys Veer, Peter D Siersema, Marguerite E I Schipper, Karel J van Erpecum, and Greet J Boland. No beneficial effects of amantadine in treatment of chronic hepatitis C patients. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, 42(7):496–502, July 2010.

[208] Gary D. Ewart, Kerry Mills, Graeme B Cox, and Peter W Gage. Amiloride derivatives block ion channel activity and enhancement of virus-like particle budding caused by HIV-1 protein Vpu. *European biophysics journal : EBJ*, 31(1):26–35, March 2002.

[209] A Premkumar, L Wilson, G D Ewart, and P W Gage. Cation-selective ion channels formed by p7 of hepatitis C virus are blocked by hexamethylene amiloride. *FEBS letters*, 557(1-3):99–103, January 2004.

[210] Carolyn A. Luscombe, Zhuhui Huang, Michael G. Murray, Michelle Miller, John Wilkinson, and Gary D. Ewart. A novel Hepatitis C virus p7 ion channel inhibitor, BIT225, inhibits bovine viral diarrhea virus in vitro and shows synergism with recombinant interferon-alpha-2b and nucleoside analogues. *Antiviral research*, 86 (2):144–153, May 2010.

[211] Gabriela Khoury, Gary Ewart, Carolyn Luscombe, Michelle Miller, and John Wilkinson. Antiviral efficacy of the novel compound BIT225 against HIV-1 release from human macrophages. *Antimicrobial agents and chemotherapy*, 54(2): 835–845, February 2010.

[212] Björn D Kuhl, Vicky Cheng, Daniel A Donahue, Richard D Sloan, Chen Liang, John Wilkinson, and Mark A Wainberg. The HIV-1 Vpu viroporin inhibitor BIT225 does not affect Vpu-mediated tetherin antagonism. *PLOS ONE*, 6(11): e27660, 2011.

[213] Lauren Wilson, Peter Gage, and Gary Ewart. Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication. *Virology*, 353(2):294–306, September 2006.

[214] Elena V Gazina and Steven Petrou. Viral targets of acylguanidines. *Drug discovery today*, 17(17-18):1039–1043, September 2012.

[215] D Pavlovic, W Fischer, M Hussey, D Durantel, S Durantel, N Branza-Nichita, S Woodhouse, R A Dwek, and N Zitzmann. Long Alkylchain Iminosugars Block the HCV p7 Ion Channel. In JohnS Axford, editor, *Advances in Experimental Medicine and Biology*, pages 3–4–4. Springer US, 2005.

[216] Shu-Fen Wu, Chyan-Jang Lee, Ching-Len Liao, Raymond A Dwek, Nicole Zitzmann, and Yi-Ling Lin. Antiviral effects of an iminosugar derivative on flavivirus infections. *Journal of virology*, 76(8):3596–3604, April 2002.

[217] N Zitzmann, A S Mehta, S Carrouée, T D Butters, F M Platt, J McCauley, B S Blumberg, R A Dwek, and T M Block. Imino sugars inhibit the formation and secretion of bovine viral diarrhea virus, a pestivirus model of hepatitis C virus: implications for the development of broad spectrum anti-hepatitis virus agents. *Proceedings of the National Academy of Sciences of the United States of America*, 96(21):11878–11882, October 1999.

[218] D Durantel, N Branza-Nichita, S Carrouée-Durantel, T D Butters, R A Dwek, and N Zitzmann. Study of the mechanism of antiviral action of iminosugar derivatives against bovine viral diarrhea virus. *Journal of virology*, 75(19):8987–8998, October 2001.

[219] T D Butters, L A G M van den Broek, George WJ Fleet, T M Krulle, M Wormald, R A Dwek, and F M Platt. Molecular requirements of imino sugars for the selective control. *Tetrahedron: Asymmetry*, 11(1):113–124, 2000.

[220] Robert J Nash, Atsushi Kato, Chu-Yi Yu, and George WJ Fleet. Iminosugars as therapeutic agents: recent advances and promising trends. *Future Medicinal Chemistry*, 3(12):1513–1521, September 2011.

[221] A P Kendal and H D Klenk. Amantadine inhibits an early, M2 protein-dependent event in the replication cycle of avian influenza (H7) viruses. *Archives of virology*, 119(3-4):265–73, January 1991.

[222] Jean-Michel Pawlotsky, Stéphane Chevaliez, and John G McHutchison. The hepatitis C virus life cycle as a target for new antiviral therapies. *Gastroenterology*, 132(5):1979–1998, May 2007.

[223] Eike Steinmann, Thomas Whitfield, Stephanie Kallis, Raymond A Dwek, Nicole Zitzmann, Thomas Pietschmann, and Ralf Bartenschlager. Antiviral effects of amantadine and iminosugar derivatives against hepatitis C virus. *Hepatology*, 46 (2):330–338, August 2007.

[224] Toshana L Foster, Mark Verow, Ann L Wozniak, Matthew J Bentham, Joseph Thompson, Elizabeth Atkins, Steven A Weinman, Colin Fishwick, Richard Foster, Mark Harris, and Stephen Griffin. Resistance mutations define specific antiviral effects for inhibitors of the hepatitis C virus p7 ion channel. *Hepatology*, 54(1): 79–90, July 2011.

[225] G Haqshenas, J M Mackenzie, X Dong, and E J Gowans. Hepatitis C virus p7 protein is localized in the endoplasmic reticulum when it is encoded by a replication-competent genome. *The Journal of general virology*, 88(Pt 1):134–142, January 2007.

[226] George Patargias, Nicole Zitzmann, Raymond Dwek, and Wolfgang B Fischer. Protein-protein interactions: modeling the hepatitis C virus ion channel p7. *Journal of medicinal chemistry*, 49(2):648–655, January 2006.

[227] Séverine Carrère-Kremer, Claire Montpellier-Pala, Laurence Cocquerel, Czeslaw Wychowski, François Penin, and Jean Dubuisson. Subcellular localization and topology of the p7 polypeptide of hepatitis C virus. *Journal of virology*, 76(8): 3720–3730, April 2002.

[228] Gabrielle Vieyres, Christiane Brohm, Martina Friesland, Juliane Gentzsch, Benno Wölk, Philippe Roingeard, Eike Steinmann, and Thomas Pietschmann. Subcellular localization and function of an epitope-tagged p7 viroporin in hepatitis C virus-producing cells. *Journal of virology*, 87(3):1664–1678, February 2013.

[229] Matthew J Bentham, Toshana L Foster, Christopher McCormick, and Stephen Griffin. Mutations in hepatitis C virus p7 reduce both the egress and infectivity of assembled particles via impaired proton channel function. *The Journal of general virology*, 94(Pt 10):2236–2248, October 2013.

[230] Christiane Brohm, Eike Steinmann, Martina Friesland, Ivo C Lorenz, Arvind Patel, François Penin, Ralf Bartenschlager, and Thomas Pietschmann. Characterization of determinants important for hepatitis C virus p7 function in morphogenesis by using trans-complementation. *Journal of virology*, 83(22):11682–11693, November 2009.

[231] Juliane Gentzsch, Christiane Brohm, Eike Steinmann, Martina Friesland, Nicolas Menzel, Gabrielle Vieyres, Paula Monteiro Perin, Anne Frentzen, Lars Kaderali, and Thomas Pietschmann. Hepatitis c Virus p7 is critical for capsid assembly and envelopment. *PLoS Pathogens*, 9(5):e1003355, 2013.

[232] V Lohmann, F Körner, J Koch, U Herian, L Theilmann, and R Bartenschlager. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science*, 285(5424):110–113, July 1999.

[233] Keril J Blight, Jane A McKeating, and Charles M. Rice. Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *Journal of virology*, 76(24):13001–13014, December 2002.

[234] Dean Clarke, Stephen Griffin, Lucy Beales, Corine St Gelais, Stan Burgess, Mark Harris, and David Rowlands. Evidence for the formation of a heptameric ion channel complex by the hepatitis C virus p7 protein in vitro. *The Journal of biological chemistry*, 281(48):37057–68, December 2006.

[235] Chee Foong Chew, Ranjit Vijayan, Jason Chang, Nicole Zitzmann, and Philip C Biggin. Determination of pore-lining residues in the hepatitis C virus p7 protein. *Biophysical journal*, 96(2):L10–2, January 2009.

[236] Philipp Luik, Chee Chew, Jussi Aittoniemi, Jason Chang, Paul Wentworth, Raymond A Dwek, Philip C Biggin, Catherine Vénien-Bryan, and Nicole Zitzmann. The 3-dimensional structure of a hepatitis C virus p7 ion channel by electron microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31):12712–12716, August 2009.

[237] Roland Montserret, Nathalie Saint, Christophe Vanbelle, Andrés Gerardo Salvay, Jean-Pierre Simorre, Christine Ebel, Nicolas Sapay, Jean-Guillaume Renisio, Anja Böckmann, Eike Steinmann, Thomas Pietschmann, Jean Dubuisson, Christophe Chipot, and François Penin. NMR structure and ion channel activity of the p7

protein from hepatitis C virus. *Journal of Biological Chemistry*, 285(41):31446–31461, October 2010.

[238] Danielle E Chandler, François Penin, Klaus Schulten, and Christophe Chipot. The p7 protein of hepatitis C virus forms structurally plastic, minimalist ion channels. *PLoS computational biology*, 8(9):e1002702, 2012.

[239] Gabriel A Cook and Stanley J Opella. Secondary structure, dynamics, and architecture of the p7 membrane protein from hepatitis C virus by NMR spectroscopy. *Biochimica et biophysica acta*, 1808(6):1448–1453, June 2011.

[240] Gabriel A Cook, Lindsay A Dawson, Ye Tian, and Stanley J Opella. Three-dimensional structure and interaction studies of hepatitis C virus p7 in 1,2-dihexanoyl-sn-glycero-3-phosphocholine by solution nuclear magnetic resonance. *Biochemistry*, 52(31):5295–5303, August 2013.

[241] Toshana L Foster, Gary S Thompson, Arnout P Kalverda, Jayakanth Kankanala, Matthew Bentham, Laura F Wetherill, Joseph Thompson, Amy M Barker, Dean Clarke, Marko Noerenberg, Arwen R Pearson, David J Rowlands, Steven W Homans, Mark Harris, Richard Foster, and Stephen Griffin. Structure-guided design affirms inhibitors of hepatitis C virus p7 as a viable class of antivirals targeting virion release. *Hepatology*, 59(2):408–422, February 2014.

[242] Bo OuYang, Shiqi Xie, Marcelo J Berardi, Xinhao Zhao, Jyoti Dev, Wenjing Yu, Bing Sun, and James J Chou. Unusual architecture of the p7 channel from hepatitis C virus. *Nature*, 498(7455):521–525, June 2013.

[243] Davor Pavlović, David C A Neville, Olivier Argaud, Baruch Blumberg, Raymond A Dwek, Wolfgang B Fischer, and Nicole Zitzmann. The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proceedings of the National Academy of Sciences of the United States of America*, 100(10):6104–6108, May 2003.

[244] Anita Premkumar, Xuebin Dong, Gholamreza Haqshenas, Peter W Gage, and Eric J Gowans. Amantadine inhibits the function of an ion channel encoded by GB virus B, but fails to inhibit virus replication. *Antiviral therapy*, 11(3):289–295, 2006.

[245] Ann L Wozniak, Stephen Griffin, David Rowlands, Mark Harris, MinKyung Yi, Stanley M Lemon, and Steven A Weinman. Intracellular proton conductance of the hepatitis C virus p7 protein and its contribution to infectious virus production. *PLoS Pathogens*, 6(9):e1001087, 2010.

[246] Thomas Whitfield, Andrew J Miles, Johanna C Scheinost, John Offer, Paul Wentworth, Raymond A Dwek, B A Wallace, Philip C Biggin, and Nicole Zitzmann. The influence of different lipid environments on the structure and function of the hepatitis C virus p7 ion channel protein. *Molecular membrane biology*, 28(5):254–264, August 2011.

[247] Corine Stgelais, Toshana L Foster, Mark Verow, Elizabeth Atkins, Colin W G Fishwick, David Rowlands, Mark Harris, and Stephen Griffin. Determinants of hepatitis C virus p7 ion channel function and drug sensitivity identified in vitro. *Journal of virology*, 83(16):7970–7981, August 2009.

[248] Yinghong Ma, Manu Anantpadma, Jennifer M Timpe, Saravanabalaji Shanmugam, Sher M Singh, Stanley M Lemon, and MinKyung Yi. Hepatitis C virus NS2 protein serves as a scaffold for virus assembly by interacting with both structural and nonstructural proteins. *Journal of virology*, 85(1):86–97, January 2011.

[249] Stephen D C Griffin, Ruth Harvey, Dean S Clarke, Wendy S Barclay, Mark Harris, and David J Rowlands. A conserved basic loop in hepatitis C virus p7 protein is required for amantadine-sensitive ion channel activity in mammalian cells but is dispensable for localization to mitochondria. *The Journal of general virology*, 85 (Pt 2):451–461, February 2004.

[250] J Keeler. *Understanding NMR Spectroscopy*. John Wiley & Sons. Ltd, September 2010.

[251] I D Campbell and R A Dwek. *Biological Spectroscopy*. The Benjamin/Cummings Publishing Company, University of Oxford, September 1984.

[252] I D Campbell. *Biophysical Techniques*. Oxford University Press, University of Oxford, September 2012.

[253] M H Levitt. *Spin Dynamics*. Wiley-Blackwell, September 2008.

[254] Paul Schanda, Ēriks Kupče, and Bernhard Brutscher. SOFAST-HMQC experiments for recording two-dimensional deteronuclear correlation spectra of proteins within a few seconds. *Journal of biomolecular NMR*, 33(4):199–211, 2005.

[255] Gareth A Morris and Ray Freeman. Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society*, 101 (3):760–762, January 1979.

[256] S M Pascal. *NMR PRIMER: An HSQC-Based Approach*. IM Publications LLP, September 2008.

[257] JMBH Jeener, B H Meier, P Bachmann, and R R Ernst. Investigation of exchange processes by twodimensional NMR spectroscopy. *The Journal of chemical physics*, 71(11):4546–4553, 2008.

[258] James A Jarvis, Ibraheem M Haies, Philip T F Williamson, and Marina Carravetta. An efficient NMR method for the characterisation of 14N sites through indirect 13C detection. *Physical chemistry chemical physics : PCCP*, 15(20):7613–7620, May 2013.

[259] Heike I Rösner and Birthe B Kragelund. Structure and dynamic properties of membrane proteins using NMR. *Comprehensive Physiology*, 2(2):1491–1539, April 2012.

[260] Paul Schanda and Bernhard Brutscher. Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *Journal of the American Chemical Society*, 127(22):8014–8015, June 2005.

[261] Walther Gerlach and Otto Stern. Das magnetische moment des silberatoms. *Zeitschrift für Physik A Hadrons and Nuclei*, 9(1):353–355, 1922.

[262] Michael Goldflam, Teresa Tarragó, Margarida Gairí, and Ernest Giralt. NMR studies of protein-ligand interactions. *Methods in molecular biology (Clifton, N.J.)*, 831:233–259, 2012.

[263] Geoffrey Bodenhausen and David J Ruben. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters*, 69(1):185–189, 1980.

[264] K Pervushin, R Riek, G Wider, and K Wüthrich. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America*, 94(23):12366–12371, November 1997.

[265] A D Bax, Richard H Griffey, and Bruce L Hawkins. Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *Journal of Magnetic Resonance (1969)*, 55(2):301–315, 1983.

[266] Luciano Mueller. Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *Journal of the American Chemical Society*, 101(16): 4481–4484, 1979.

[267] Katelijne Gheysen, Camelia Mihai, Karel Conrath, and José C Martins. Rapid identification of common hexapyranose monosaccharide units by a simple TOCSY matching approach. *Chemistry (Weinheim an der Bergstrasse, Germany)*, 14(29): 8869–8878, 2008.

[268] B Whitehead, C J Craven, and J P Waltho. Double and triple resonance NMR methods for protein assignment. *Methods in molecular biology (Clifton, N.J.)*, 60: 29–52, 1997.

[269] Dominique Marion, Lewis E Kay, Steven W Sparks, Dennis A Torchia, and Ad Bax. Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *Journal of the American Chemical Society*, 111(4):1515–1517, 1989.

[270] K Wüthrich, G Wider, G Wagner, and W Braun. Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance. *Journal of molecular biology*, 155(3):311–319, March 1982.

[271] C Bartels and K Wüthrich. A spectral correlation function for efficient sequential NMR assignments of uniformly (15)N-labeled proteins. *Journal of biomolecular NMR*, 4(6):775–785, November 1994.

[272] Arthur G Palmer. NMR characterization of the dynamics of biomacromolecules. *Chemical reviews*, 104(8):3623–3640, August 2004.

[273] T Szyperski, P Luginbühl, G Otting, P Güntert, and K Wüthrich. Protein dynamics studied by rotating frame 15N spin relaxation times. *Journal of biomolecular NMR*, 3(2):151–164, March 1993.

[274] Virginia A Jarymowycz and Martin J Stone. Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chemical reviews*, 106(5):1624–1671, May 2006.

[275] Tyler Reddy and Jan K Rainey. Interpretation of biomolecular NMR spin relaxation parameters. *Biochemistry and cell biology*, 88(2):131–142, April 2010.

[276] Julie M Kneller, Min Lu, and Clay Bracken. An effective method for the discrimination of motional anisotropy and chemical exchange. *Journal of the American Chemical Society*, 124(9):1852–1853, March 2002.

[277] Stanley J Opella and Francesca M Marassi. Structure determination of membrane proteins by NMR spectroscopy. *Chemical reviews*, 104(8):3587–3606, August 2004.

[278] H Schwalbe, K M Fiebig, M Buck, J A Jones, S B Grimshaw, A Spencer, S J Glaser, L J Smith, and C M Dobson. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry*, 36(29):8977–8991, July 1997.

[279] Giovanni Lipari and Attila Szabo. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *Journal of the American Chemical Society*, 104(17):4559–4570, 1982.

[280] Binyong Liang, Ashish Arora, and Lukas K Tamm. Fast-time scale dynamics of outer membrane protein A by extended model-free analysis of NMR relaxation data. *Biochimica et biophysica acta*, 1798(2):68–76, February 2010.

[281] J P Staley and P S Kim. Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor. *Protein science : a publication of the Protein Society*, 3(10):1822–1832, October 1994.

[282] Jolyon K Claridge and Jason Schnell. Bacterial production and solution NMR studies of a viral membrane ion channel. *Methods in molecular biology (Clifton, N.J.)*, 831:165–79, January 2012.

[283] Kevin Bertrand, Craig Squires, and Charles Yanofsky. Transcription termination in vivo in the leader region of the tryptophan operon of Escherichia coli. *Journal of molecular biology*, 103(2):319–337, May 1976.

[284] G F Miozzari and C Yanofsky. Translation of the leader region of the Escherichia coli tryptophan operon. *Journal of bacteriology*, 133(3):1457–1466, March 1978.

[285] Andrew M Waterhouse, James B Procter, David M A Martin, Michèle Clamp, and Geoffrey J Barton. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)*, 25(9):1189–1191, May 2009.

[286] Arianna Rath, Mira Glibowicka, Vincent G Nadeau, Gong Chen, and Charles M Deber. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 106(6):1760–1765, February 2009.

[287] F William Studier. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification*, 41(1):207–234, May 2005.

[288] Gabriel A Cook, Susanne Stefer, and Stanley J Opella. Expression and purification of the membrane protein p7 from hepatitis C virus. *Biopolymers*, 96(1):32–40, 2011.

[289] E Gross and B Witkop. Selective cleavage of the methionyl peptide bonds in ribonuclease with cyanogen bromide1. *Journal of the American Chemical Society*, 1961.

[290] Arun Sivashanmugam, Victoria Murray, Chunxian Cui, Yonghong Zhang, Jianjun Wang, and Qianqian Li. Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein science : a publication of the Protein Society*, 18(5):936–948, May 2009.

[291] Mike P Williamson. Using chemical shift perturbation to characterise ligand binding. *Progress in nuclear magnetic resonance spectroscopy*, 73:1–16, August 2013.

[292] H Hauser. Short-chain phospholipids as detergents. *Biochimica et biophysica acta*, 1508(1-2):164–181, November 2000.

[293] J Kessi, J C Poiree, E Wehrli, R Bachofen, G Semenza, and H Hauser. Short-chain phosphatidylcholines as superior detergents in solubilizing membrane proteins and preserving biological activity. *Biochemistry*, 33(35):10825–10836, 1994.

[294] Matthew E Call, Kai W Wucherpfennig, and James J Chou. The structural basis for intramembrane assembly of an activating immunoreceptor complex. *Nature immunology*, 11(11):1023–1029, November 2010.

[295] Daniel Schwarz, Friederike Junge, Florian Durst, Nadine Frölich, Birgit Schneider, Sina Reckel, Solmaz Sobhanifar, Volker Dötsch, and Frank Bernhard. Preparative scale expression of membrane proteins in Escherichia coli-based continuous exchange cell-free systems. *Nature protocols*, 2(11):2945–2957, 2007.

[296] Zhenwei Lu, Wade D Van Horn, Jiang Chen, Sijo Mathew, Roy Zent, and Charles R Sanders. Bicelles at low concentrations. *Molecular pharmaceutics*, 9 (4):752–761, April 2012.

[297] M A Eriksson, T Härd, and L Nilsson. On the pH dependence of amide proton exchange rates in proteins. *Biophysical journal*, 69(2):329–339, August 1995.

[298] RLJ Keller. *Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment.* PhD thesis, ETH Zurich Thesis, Zurich, 2004.

[299] K Wüthrich. *NMR of proteins and nucleic acids*. Wiley, New York (USA), October 1986.

[300] T L Hwang, P C van Zijl, and S Mori. Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *Journal of biomolecular NMR*, 11(2):221–226, February 1998.

[301] Yang Shen, Frank Delaglio, Gabriel Cornilescu, and Ad Bax. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR*, 44(4):213–223, August 2009.

[302] G Cornilescu, F Delaglio, and A Bax. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of biomolecular NMR*, 13(3):289–302, March 1999.

[303] Yu-Dong Cai, Xiao-Jun Liu, Xue-Biao Xu, and Kuo-Chen Chou. Artificial neural network method for predicting protein secondary structure content. *Computers & chemistry*, 26(4):347–350, June 2002.

[304] Carol A Rohl, Charlie E M Strauss, Kira M S Misura, and David Baker. Protein Structure Prediction Using Rosetta. pages 66–93. Elsevier, 2004.

[305] Yawen Bai, John S Milne, Leland Mayne, and S Walter Englander. Primary structure effects on peptide group hydrogen exchange. *Proteins*, 17(1):75–86, September 1993.

[306] L E Kay, D A Torchia, and A Bax. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, 1989.

[307] Donghan Lee, Christian Hilty, Gerhard Wider, and Kurt Wüthrich. Effective rotational correlation times of proteins from NMR relaxation interference. *Journal of magnetic resonance (San Diego, Calif. : 1997)*, 178(1):72–76, January 2006.

[308] Joseph R Casey, Sergio Grinstein, and John Orlowski. Sensors and regulators of intracellular pH. *Nature reviews. Molecular cell biology*, 11(1):50–61, January 2010.

[309] J H Lin and A Y Lu. Role of pharmacokinetics and metabolism in drug discovery and development. *Pharmacological reviews*, 49(4):403–449, December 1997.

[310] David V Tulumello and Charles M Deber. Efficiency of detergents at maintaining membrane protein structures in their biologically relevant forms. *Biochimica et biophysica acta*, 1818(5):1351–1358, May 2012.

[311] Romain Gautier, Dominique Douguet, Bruno Antonny, and Guillaume Drin. HELIQUEST: a web server to screen sequences with specific alpha-helical properties. *Bioinformatics (Oxford, England)*, 24(18):2101–2102, September 2008.

[312] Dianfan Li, Joseph A Lyons, Valerie E Pye, Lutz Vogeley, David Aragão, Colin P Kenyon, Syed T A Shah, Christine Doherty, Margaret Aherne, and Martin Caffrey. Crystal structure of the integral membrane diacylglycerol kinase. *Nature*, 497 (7450):521–524, May 2013.

[313] Gunnar von Heijne. Membrane-protein topology. *Nature reviews. Molecular cell biology*, 7(12):909–918, December 2006.

[314] Ilka Wittig, Hans-Peter Braun, and Hermann Schägger. Blue native PAGE. *Nature protocols*, 1(1):418–428, 2006.

[315] Lydia M Harriss, Brid Cronin, James R Thompson, and Mark I Wallace. Imaging multiple conductance states in an alamethicin pore. *Journal of the American Chemical Society*, 133(37):14507–14509, September 2011.

[316] Sebastian Leptihn, Oliver K Castell, Brid Cronin, En-Hsin Lee, Linda C M Gross, David P Marshall, James R Thompson, Matthew Holden, and Mark I Wallace. Constructing droplet interface bilayers from the contact of aqueous droplets in oil. *Nature protocols*, 8(6):1048–1057, June 2013.

[317] Linda C M Gross, Oliver K Castell, and Mark I Wallace. Dynamic and reversible control of 2D membrane protein concentration in a droplet interface bilayer. *Nano letters*, 11(8):3324–3328, August 2011.

[318] Robert B Merrifield. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *Journal of the American Chemical Society*, 85(14):2149–2154, 1963.

[319] Jean-Michel Pawlotsky. Chapter Five - Hepatitis C Virus: Standard-of-Care Treatment. In Erik De Clercq, editor, *Antiviral Agents*, pages 169–215. Academic Press, 2013.

[320] Annika Vermehren, Christoph Welsch, Ulrike Elsler, Johannes Vermehren, Eva Herrmann, Christoph Sarrazin, Michael von Wagner, Simone Susser, Wolf Peter Hofmann, Bernd Kronenberger, Stefan Zeuzem, and Ulrike Mihm. Investigation of viral escape mutations within HCV p7 during treatment with amantadine in patients with chronic hepatitis C. *Antiviral therapy*, July 2013.

[321] E. Steinmann, M Friesland, S Ciesek, N Zitzmann, and T Pietschmann. Culture long-term treatment with iminosugar-based HCV assembly inhibitors does not lead to viral breakthrough in tissue culture. *Journal of hepatology*, 54:S489, January 2011.

[322] Manuela Gebhardt, Franziska Hoffgaard, Kay Hamacher, Stefan M Kast, Anna Moroni, and Gerhard Thiel. Membrane anchoring and interaction between transmembrane domains are crucial for K+ channel function. *The Journal of biological chemistry*, 286(13):11299–11306, April 2011.

[323] Devaki A Kelkar and Amitabha Chattopadhyay. Membrane interfacial localization of aromatic amino acids and membrane protein function. *Journal of biosciences*, 31(3):297–302, September 2006.

[324] Kakoli Mitra, Iban Ubarretxena-Belandia, Tomohiko Taguchi, Graham Warren, and Donald M Engelman. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proceedings of the National Academy of Sciences of the United States of America*, 101(12):4083–4088, March 2004.

[325] J Clogston and M Caffrey. Controlling release from the lipidic cubic phase. Amino acids, peptides, proteins and nucleic acids. *Journal of controlled release : official journal of the Controlled Release Society*, 107(1):97–111, September 2005.

[326] Andrew J Heron, James R Thompson, Brid Cronin, Hagan Bayley, and Mark I Wallace. Simultaneous Measurement of Ionic Current and Fluorescence from Single Protein Pores. *Journal of the American Chemical Society*, 131(5):1652–1653, February 2009.

[327] Andrew J Heron, James R Thompson, Amy E Mason, and Mark I Wallace. Direct detection of membrane channels from gels using water-in-oil droplet bilayers. *Journal of the American Chemical Society*, 129(51):16042–16047, December 2007.

[328] Nick Soffe, Jonathan Boyd, and Mark Leonard. The construction of a high-resolution 750 MHz probehead. *Journal of Magnetic Resonance, Series A*, 116 (1):117–121, 1995.

[329] F Delaglio, S Grzesiek, G W Vuister, G Zhu, J Pfeifer, and A Bax. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR*, 6(3):277–293, November 1995.

[330] Wim F Vranken, Wayne Boucher, Tim J Stevens, Rasmus H Fogh, Anne Pajon, Miguel Llinas, Eldon L Ulrich, John L Markley, John Ionides, and Ernest D Laue.

The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, 59(4):687–696, June 2005.