**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**NONSTRUCTURAL 5A (NS5A) PROTEIN of
HEPATITIS C VIRUS: SELF-ASSOCIATION, MEMBRANE
ANCHORING and INTERACTIONS with ANTIVIRALS**

Serap BELDAR

School of Biological Sciences

2017

**Nonstructural 5A (NS5A) Protein of
Hepatitis C Virus: Self-Association, Membrane
Anchoring and Interactions with Antivirals**

Serap BELDAR

School of Biological Sciences

**A Thesis Submitted to the Nanyang Technological University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy**

2017

To my beloved family for their constant support.

ACKNOWLEDGEMENTS

Firstly, I would like to express the gratitude to my supervisor, Prof. Ho Sup Yoon. He has introduced me to the field of biochemistry and structural biology with patience. I am thankful to him for his constant support, guidance, and kindness. It was a unique experience to be a member of his group. This thesis would not be possible without his help considering how challenging the project was. I feel more than happy to have him as a supervisor.

I would like to thank Prof. Gerhard Grüber sincerely for his valuable insights and helps. I wish to thank Dr. Malathy Sony Subramanian Manimekalai for performing SAXS measurements and data analysis. We very much enjoyed collaborating with both Prof. Grüber and Dr. Sony. Their contribution paved the way for exploring new directions in my project.

I wish to thank Prof. Nam-Joon Cho for introducing me the Prof. Yoon's lab. I would like to thank him also for his guidance for the QCM-D experiments. He always warmly welcomed me and allowed me to use the equipment in his laboratory. I would like to thank my friend and Prof. Cho's Ph.D. student Şaziye Yorulmaz for teaching me the QCM-D technique.

I am always thankful to the members of Prof. Yoon's Lab for being kind and creating a very peaceful working environment. I would like to thank Dr. Shin Joon and Dr. Ye Hong for helping me to learn the fundamental aspects of NMR technique and conducting NMR experiments. I would like to thank Dr. Sreekanth Rajan, for teaching me the technical part of protein crystallization and helping me to perform crystallization experiments. I wish to express my appreciation to Dr. Toan Quoc Nguyen, Dr. Ajit Prakash, Hui Ting Toh, MinJoo Choi, Yeen Shian Ngow and Jonathan Gan for their friendship and helpfulness whenever I need. Especially, the jokes of Dr. Toan and Dr. Ajit made our lab life very joyful.

My friends not only in Singapore but also from other parts of the world have been priceless during the tough time of Ph.D. I am thankful to Duygu Kalkan, Cem Sonat, İsmail Can, Ufuk Borucu, and H.Enis Karahan for being caring, thoughtful and fun. I

would like to express my special appreciations to Ufuk and particularly Enis for their fruitful discussions and suggestions regarding my research as well.

Lastly, I would like to thank my family for their love and support. Especially, I am grateful to my mom, her strong standing and visions inspired me in every part of the life.

Table of Contents

ABSTRACT	i
ACKNOWLEDGEMENTS	i
LIST of TABLES	vii
LIST of FIGURES	viii
ABBREVIATIONS	xi
Chapter 1: BACKGROUND	1
1.1. Viral Hepatitis	1
1.1.1. History of the viral hepatitis	3
1.1.2. Discovery of hepatitis C virus	4
1.1.3. Taxonomy of hepatitis C virus	6
1.1.4. Epidemiology of the disease	8
1.2. Structural and Genomic Features of HCV	13
1.2.1. Structural organization of HCV viral particle	13
1.2.2. Genomic organization of hepatitis C virus	14
1.2.3. Hepatitis C virus proteins	16
1.2.4. The life cycle of hepatitis C virus	21
1.3. NS5A Protein	23
1.3.1. Structure and Function of NS5A	24
1.3.2. Interacting partners of NS5A	27
1.3.3. Phosphorylation sites of NS5A	28
1.3.4. NS5A is an RNA-binding protein	29
1.3.5. Emerging inhibitors targeting NS5A	30
1.4. The Aim of the Current Study	34
Chapter 2: MATERIALS and METHODS	35
2.1. Materials	35
2.1.1. Chemicals, enzymes, and kits	35
2.1.2. Vectors and cell strains	36
2.1.3. Instrumentation	36
2.1.4. Buffers and solutions	36
2.2. Methods	37
2.2.1. Polymerase chain reaction (PCR)	37

2.2.2. Plasmid and DNA product purification.....	39
2.2.3. Restriction of plasmids and inserts	40
2.2.4. DNA agarose gel electrophoresis	40
2.2.5. Ligation	41
2.2.6. Site-directed mutagenesis.....	41
2.2.7. Competent cell preparation by CaCl ₂ treatment method.....	41
2.2.8. Transformation of plasmid DNA into <i>E. coli</i>	42
2.2.9. Colony PCR	42
2.2.10. Media preparation	42
2.2.11. Expression of recombinant protein	42
2.2.12. Cell disruption by ultrasonication	43
2.2.13. Purification of soluble proteins by Ni ⁺² -NTA affinity column.....	43
2.2.14. SUMO-tag cleavage and second round purification by Ni ⁺² -NTA affinity column	43
2.2.15. Regeneration of Ni ⁺² -NTA beads.....	44
2.2.16. Purification of membrane proteins.....	44
2.2.17. SDS-PAGE.....	45
2.2.18. Western blot analysis	45
2.2.19. Fast-Performance Liquid Chromatography (FPLC)	46
2.2.20. Sample processing for ¹ H NMR experiment (1D NMR)	46
2.2.21. ¹⁵ N labeling of recombinant proteins.....	46
2.2.22. Sample processing for ¹⁵ N-labeled NMR experiment (2D NMR)	47
2.2.23. Protein concentration determination	47
2.2.24. Chemical crosslinking.....	47
2.2.25. Small-angle X-ray Scattering (SAXS).....	47
2.2.26. Thermal denaturation assay.....	48
2.2.27. Dynamic light scattering (DLS).....	51
2.2.28. Circular dichroism (CD)	51
2.2.29. Quartz crystal microbalance with dissipation (QCM-D) monitoring.....	51
2.2.30. Supported lipid bilayer (SLB) formation	52
2.3. Principles of SAXS and QCM-D	53
2.3.1. Structural biology techniques and small angle X-ray scattering.....	53
2.3.2. Quartz crystal microbalance with dissipation monitoring.....	61

Chapter 3: CLONING, EXPRESSION, and PURIFICATION of DOMAIN I of NS5A PROTEIN from HEPATITIS C VIRUS	63
3.1. Motivation.....	63
3.2. Bioinformatics.....	63
3.3. Cloning, Expression, and Purification of NS5A-DI.....	63
3.3.1. His-SUMO-NS5A(29-213) construct	65
3.3.2. His-SUMO-GSGSG-NS5A(29-213).....	72
3.3.3. UBI-NS5A(29-213)-His construct	74
3.3.4. NS5A(36-198)-His construct	75
3.3.5. His-SUMO-NS5A(36-198)-DDDDK construct.....	77
3.4. 1D and 2D NMR of NS5A(36-198)-DDDDK	79
3.5. Optimizing the buffer conditions for protein stability by using thermal denaturation assay	83
3.6. Overview of chapter.....	87
Chapter 4: SELF-ASSOCIATION of NS5A DOMAIN I of HEPATITIS C VIRUS: BIOPHYSICAL INVESTIGATIONS in SOLUTION	89
4.1. Motivation.....	89
4.2. NS5A-DI of genotype 1a and genotype 1b of HCV share high sequence similarity	89
4.3. NS5A-DI from genotype 1b of HCV self-associates in solution and forms reversible higher-order oligomers.....	90
4.4. Low-resolution structure: The oligomers of NS5A-DI from genotype 1b is an elongated polymer of dimer subunits	97
4.5. Self-association behavior of NS5A-DI of genotype 1a differs from genotype 1b	103
4.6. Low-resolution solution structure of NS5A-DI from genotype 1a	109
4.7. The effect of daclatasvir and RNA on NS5A-DI	113
4.8. Overview of chapter.....	117
Chapter 5: MOLECULAR CHARACTERIZATION of NS5A-DI with AMPHIPATHIC HELIX in SUPPORTED LIPID BILAYERS.....	119
5.1. Motivation.....	119
5.2. Cloning, expression, solubilisation, and purification of NS5A-DI including amphipathic helix (AH)	119
5.2.1. Cloning strategy of NS5A-DI with amphipathic helix.....	119
5.2.2. Expression screening of HCV NS5A-DI with Amphipathic Helix (AH)	120
5.2.3. Detergent screening and solubility check of HCV 1a NS5A(2-198)-His-tag	121

5.2.4. Purification of NS5A(2-198)-His-tag by imidazole gradient IMAC	124
5.2.5. Size-exclusion chromatography and secondary structure analysis of NS5A(2-198)-His-tag.....	125
5.3. The interactions of simulated membranes and NS5A-DI with AH: Towards understanding the mechanism of action of a direct-acting antiviral, daclatasvir.....	127
5.3.1. The formation of charged supported bilayer platform by solvent-assisted lipid bilayer (SALB) method.....	127
5.3.2. The effects of buffers on the association of membranes and NS5A(2-198)	129
5.3.3. The effect of daclatasvir dihydrochloride on supported bilayer platforms	133
5.3.4. The effect of DCV-HCl on the association of NS5A(2-198)-His-tag to the membranes	134
5.3.5. Daclatasvir-resistance mutations on NS5A and its associations with membranes with the drug: Towards an understanding of the mechanism of resistance mutations.....	138
5.4. Overview of chapter.....	141
Chapter 6: General Discussion & Perspectives	143
6.1. General Discussion	143
6.2. Perspectives.....	147
APPENDICES	149
BIBLIOGRAPHY	156

LIST of TABLES

Table 1-1. Current hepatitis C drugs and their mechanism of action.....	13
Table 2-1. The components of PCR reaction mixtures.	38
Table 2-2. Primers used in this study for the vector constructs	39
Table 2-3. Plasmids and their properties.....	40
Table 2-4. A typical restriction digestion protocol.	40
Table 2-5. Agarose gel electrophoresis percentage calculation.	41
Table 2-6. Buffers screened for thermal denaturation assay	50
Table 2-7. Comparison of structural biology techniques	54
Table 4-1: Data collection and X-ray scattering derived parameters for HCV NS5A-DI from genotypes 1b and 1a.....	102
Table 4-2. The statistics of the SAXS analysis of the NS5A-DI protein from genotype 1a of HCV.	113
Table 5-1. Properties of the detergents screened for solubilization of NS5A(2-198) membrane protein	123
Table 5-2. Statistics of QCM-D measurements on charged supported bilayers	133
Table 5-3. Protein stock concentration for SBL studies.....	141

LIST of FIGURES

Figure 1-1. Leading causes of deaths worldwide in 1990 and 2013	2
Figure 1-2. Targets of reduction in the global number of new infection cases and deaths	3
Figure 1-3. Timeline of the milestones in hepatitis C virus research.....	5
Figure 1-4. The taxonomy of hepatitis C virus	7
Figure 1-5. The estimated prevalence of hepatitis C virus infections worldwide.....	9
Figure 1-6. Outcomes associated with hepatitis C virus infection	11
Figure 1-7. The advances of HCV treatments from 1989 till now	12
Figure 1-8. The schematic structure of hepatitis C virus particle	14
Figure 1-9. Genomic organization of hepatitis C virus.....	15
Figure 1-10. The hepatitis C virus proteins.....	17
Figure 1-11. Life cycle of hepatitis C virus	21
Figure 1-12. The schematic organization of NS5A protein	24
Figure 1-13. The three different crystal structures.....	26
Figure 1-14. Electrostatic potential of the three different crystal structures for the domain I of NS5A protein.	27
Figure 1-15. The NS5A inhibitor resistance mutations.....	31
Figure 1-16. 2D chemical structures of direct-acting antivirals.....	32
Figure 1-17. Possible effects of NS5A inhibitors on the protein	33
Figure 2-1. Workflow/strategies for the purification of membrane proteins	45
Figure 2-2. Schematic representation of a SAXS experiment	55
Figure 2-3. Data analysis on a simulated SAXS data	58
Figure 2-4. Schematic for SAXS data solution structure modelling.....	59
Figure 2-5. Principles of QCM-D	62
Figure 3-1. PCR amplicons of the constructs and their schematic map.....	65
Figure 3-2. Protein expression control for His-SUMO-NS5A(29-213) construct	67
Figure 3-3. The purification of His-SUMO-NS5A(29-213) protein by affinity column	69
Figure 3-4. SUMO-tag cleavage SDS-PAGE analysis of His-SUMO-NS5A(29-213)	71
Figure 3-5. Expression and purification of His-SUMO-GSGSG-NS5A(29-213) protein	73
Figure 3-6. Protein purification of UBI-NS5A(29-213)-His	75
Figure 3-7. Expression and purification of NS5A(36-198)-His protein	77
Figure 3-8. Purification of the monomeric state NS5A(36-198)-DDDDK protein	79
Figure 3-9. NMR experiments showing a well-folded state of purified NS5A(36-198)-DDDDK	81

Figure 3-10. 2D ^1H , ^{15}N HSCQ spectra of NS5A(36-198)-DDDDK with 15U RNA	83
Figure 3-11. A representative fluorescence measurement of thermal denaturation assay of NS5A(36-198).....	84
Figure 3-12. pH 7.4 and HEPES buffer improves the thermal stability of NS5A(36-198)	85
Figure 3-13. The effect of salts on the thermal stability of NS5A(36-198)-DDDDK	86
Figure 3-14. The effect of reducing agents on the thermal stability of NS5A(36-198).....	86
Figure 4-1. Pairwise sequence alignment of NS5A(36-198)-DDDDK from genotypes 1a and 1b.....	90
Figure 4-2. Monomer state of NS5A-DI from genotype 1b of HCV	91
Figure 4-3. The size distribution analysis of purified NS5A-DI from genotype 1b by DLS	92
Figure 4-4. Elution profiles of the purified NS5A-DI from genotype 1b	94
Figure 4-5 Higher-order oligomers of NS5A-DI from genotype 1b are reversible	95
Figure 4-6. (a) Same amount of NS5A-DI from genotype 1b (500 μg) at indicated concentrations were loaded. (b) Normalized SEC elution profiles of the NS5A-DI from genotype 1b at different concentrations are shown.....	96
Figure 4-7. BMH crosslinking of NS5A-DI of genotype 1b.....	97
Figure 4-8. SAXS studies of NS5A-DI from genotype 1b of HCV	100
Figure 4-9. Size-exclusion chromatography of NS5A-DI from genotype 1a of HCV.....	103
Figure 4-10.DLS analysis of the proteins	104
Figure 4-11. SEC analysis the proteins	106
Figure 4-12. The SEC analysis of the proteins NS5A-DI from genotype 1a and genotype 1b	107
Figure 4-13. BMH crosslinking of NS5A-DI from genotype 1a	108
Figure 4-14. SAXS analysis of NS5A-DI from genotype 1a of HCV	110
Figure 4-15. The tetramer crystal structure of NS5A-DI from genotype 1a of HCV	112
Figure 4-16. NS5A-DI from genotype 1a of HCV exist as a dimer in solution.....	112
Figure 4-17. The effect of RNA and DCV-HCl on self-association of NS5A-DI of genotype 1b	115
Figure 4-18. SAXS studies of NS5A-DI of genotype 1b of HCV with 15U RNA and DCV-HCl.....	116
Figure 5-1. Cloning of NS5A(2-198)-His-tag.....	120
Figure 5-2. The protein expression control of NS5A(2-198)-His-tag.....	121
Figure 5-3. Dot blot analysis for detergent screening of the solubilisation of NS5A(2-198)-His-tag.....	124
Figure 5-4. Imidazole gradient purification of NS5A(2-198)-His-tag	125

Figure 5-5. The homogeneity and folding check of NS5A(2-198)-His-tag	126
Figure 5-6. The design of supported lipid bilayers membranes using charged and uncharged lipid molecules	129
Figure 5-7. The effect of buffers on supported lipid bilayer platforms.....	131
Figure 5-8. Associations of NS5A(2-198)-His-tag with supported bilayer platform.....	132
Figure 5-9. Investigations of the interactions between DCV-HCl and charged lipid bilayer platforms	134
Figure 5-10. The effect of DCV-HCl on the membrane association of NS5A(2-198) protein.....	136
Figure 5-11. The effect daclatasvir on the protein associated with the lipid bilayer.....	138
Figure 5-12. The interactions of daclatasvir-resistant NS5A on supported bilayers.....	140
Figure 6-1. The schematic illustrations of in solution characterization of NS5A-DI from genotype 1b and 1a	145

ABBREVIATIONS

1D:	one-dimension
Å:	Angstrom
Abs:	absorbance
AH:	amphipathic helix
APS:	ammonium persulfate
BCA:	bicinchoninic acid
BME:	β-mercaptoethanol
Bp:	base pair
BSA:	Bovine serum albumin
CD:	circular Dichroism
DCV:	daclatasvir
DCV-HCl:	daclatasvir dihydrochloride
DLS:	dynamic light scattering
DOEPC:	1,2-Dioleoyl-sn-glycero-3-ethylphosphocholine
DOPC:	1,2-dioleoyl-sn-glycero-3-phosphocholine
DTT:	DL-dithiothreitol
<i>E. coli:</i>	<i>Escherichia coli</i>
EC50:	half maximal effective concentration
Ek:	enterokinase, light chain
EM:	electron microscopy
ER:	endoplasmic reticulum
FPLC:	Fast-Performance Liquid Chromatography
HBV:	Hepatitis B virus
HCC:	hepatocellular carcinoma
HCV:	Hepatitis C virus
Hz:	hertz

IFN:	interferon
IFN- α :	interferon-alpha
IPTG:	isopropyl β -D-1-thiogalactopyranoside
JFH-1:	Japanese fulminant hepatitis
Kb:	kilo base
kDa:	kilodalton
LB:	Luria-Bertani
MoA:	mode of action
MW:	molecular weight
NANBH:	non-A non-B hepatitis
NMR:	nuclear magnetic resonance
NS:	nonstructural proteins
PAGE:	polyacrylamide gel electrophoresis
PCR:	polymerase chain reaction
PDB:	Protein Data Bank
PMSF:	phenylmethylsulfonyl fluoride
POPG:	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
PPI:	protein-protein interaction
RdRp:	RNA-dependent RNA polymerase
SALB:	solvent-assisted lipid bilayer
SAXS:	small angle X-ray scattering
SBL:	supported bilayer
SDS:	sodium dodecyl sulfate
SEC:	size exclusion buffer
SVR:	sustained viral response
UBI:	ubiquitin
UV-vis:	ultraviolet-visible
WHO:	World Health Organization

ABSTRACT

Direct-acting antivirals (DAAs) targeting the nonstructural 5A (NS5A) protein of Hepatitis C virus (HCV) are essential drugs with their exceptional clinical success in actual patients. However, surprisingly, their modes of action (MoA) are still unknown, and even more importantly, resistant types of the virus to DAAs are also quickly emerging. Therefore, it is vital to better understand the behavior of NS5A in every aspect, which may facilitate the development of new classes of inhibitors or methods that improve the efficacy of currently available DAAs. To this end, we have employed a facile method based on an expression vector with an N-terminal SUMO tag and have successfully purified NS5A-domain I (DI), where most resistance mutations are mapped on so far. Our system enabled us to achieve soluble proteins in high yields, which allowed to investigate the structural dynamics of NS5A and self-association behavior in vitro. By employing dynamic light scattering (DLS), size-exclusion chromatography (SEC), and chemical crosslinking experiments, we have shown that NS5A-DI from genotype 1b of HCV intrinsically self-associates in solution at physiological conditions. Notably, they remain in an equilibrium state, where monomers, dimers (transient), and higher-order oligomers coexist as a heterogeneous mixture. Small angle X-ray scattering (SAXS) studies suggest that the dimer units of NS5A-DI assemble as a helical polymer to form higher order structures. Interestingly, NS5A-DI of genotype 1a exhibits an alternative form of dynamic equilibrium of the proteins in solutions. To understand its MoA, we have tested the effect of daclatasvir (DCV), the most prominent DAA, on self-association and membrane-anchoring behavior of the protein. We have analyzed the effect of DCV on NS5A using DLS and SAXS. Our results showed that DCV does not shift the dynamic self-association behavior of the protein. However, DCV triggered the aggregation of protein, which is consistent with the features observed earlier. The multimers of NS5A in the cell culture were detected by fluorescent tags covalently attached to NS5A, and the formation of larger foci of NS5A in the presence of NS5A inhibitors was reported. As the mechanism of HCV infection is membrane-associated, we also explored the possible effect of DCV on the membrane anchoring behavior of NS5A protein by utilizing quartz crystal microbalance with dissipation monitoring (QCM-D) on supported bilayer platforms. These studies demonstrated that DCV does not completely block the membrane-anchoring of domain I of NS5A. We have tested wild-type and DCV-resistant mutant NS5A-DI proteins and observed a slight difference

in their membrane-anchoring behavior when incubated with DCV together. This study highlights the conformational variability of the NS5A domain 1 of HCV, which might be playing a significant role in the modulating of the multiple functions of the protein.

Chapter 1: BACKGROUND

1.1. Viral Hepatitis

In simple terms, hepatitis refers to the inflammation of the liver: “hepat” stands for “liver” and the suffix “-itis” indicates “inflammation”. Severe alcohol consumption, exposure to certain chemicals, a variety of autoimmune diseases and metabolic disorders, and several viruses can lead to hepatitis. And, among these, viral hepatitis is the most common reason behind hepatitis disease worldwide (Lavanchy, 2009).

The primary viral agents of the disease are named as hepatitis A virus (HAV), hepatitis B virus (HBV), hepatitis C virus (HCV), hepatitis D virus (HDV), and hepatitis E virus (HEV). However, viruses such as yellow fever virus, Epstein-Barr virus, cytomegalovirus, measles virus, and herpes simplex virus are also known to cause the inflammation of liver cells (Hadziyannis, 1998).

The hepatitis viruses vary in many ways from each other, like the family they belong, the way they are transmitted, and the viral genome type. For example, HAV is in the family of *Picornaviridae* and transmitted through fecal/oral ways with a single-stranded RNA (ssRNA) genomic content. HCV, HDV, and HEV are also +ssRNA viruses. HBV, on the other hand, is in the *Hepadnaviridae* family of a double-stranded DNA (dsDNA) with a transmission route of permucosal exposures (Zuckerman, 1996). Again importantly, although HAV and HEV do not carry envelope structure and do not normally cause chronic infections, HBV, HCV, and HDV are enveloped viruses and commonly cause chronic infections (Hall, 2007; Lindenbach, Thiel, & Rice, 2007; Trepo, 2014).

Viral hepatitis is estimated to cause more than one million deaths per year worldwide according to the data collected in between 1990 and 2013. And, hepatitis B and hepatitis C diseases account for more than 90% of viral hepatitis associated deaths and disabilities, in equal contribution (Shlomai, de Jong, & Rice, 2014; Stanaway et al., 2016). Moreover, viral hepatitis associated deaths jumped from tenth to seventh place as the leading cause of mortality with a significant increase in its relative importance

(Figure 1.1). At least partially relying on the emergence of new treatments, World Health Organization (WHO) targets to reach a dramatic reduction in the total new infection cases and deaths of hepatitis B and C by 2020 and beyond (Figure 1.2) (WHO, 2016). At this point, the major motivation of hepatitis researchers is to maintain such targets by better understanding the disease and the mechanism of current treatments with the hope of total eradication of the disease in the future.

Figure 1-1. Leading causes of deaths worldwide in 1990 and 2013. The figure is reproduced based on (Stanaway et al., 2016). Viral hepatitis made a significant jump from tenth place to seventh as a leading cause of global deaths.

Figure 1-2. Targets of reduction in the global number of new infection cases and deaths caused by hepatitis B and C in total in the near future (WHO, 2016).

1.1.1. History of the viral hepatitis

Hepatitis has been one of the major infectious diseases of humankind for centuries, and the infection was characterized by jaundice, which is due to the mild-zone necrosis of the liver. The first description was found on Sumerian clay tablets from 3000 B.C. as a report of jaundice. The Sumerian believed that a devil named Ahhazu that attacks the liver causes this symptom. “Hippocratic corpus”, a collection of medical work (400 B.C.) gave the first description of the clinical features of the disease. During the eighth century, the disease was recognized better, and the people with jaundice were identified to be infectious and suggested to be isolated. In 1885, in a Bremen shipyard, Lürhman observed that the long-term jaundice was transmitted to people who were vaccinated against smallpox, but not to other employees and the infection was postulated to be blood-borne (Trepo, 2014).

At the beginning of the twentieth century, the infecting agent carrying the disease was considered to be viral (McDonald, 1918). Due to the poor sanitary conditions of World War II (WWII), a large outbreak of hepatitis took place and right after WWII, viral hepatitis classified into two types, the Viral Hepatitis A and B, the infectious virus and serum virus, respectively (MacCallum, 1947). Immediately, after the discovery of Australia antigen (Ag) and serum hepatitis antigen (Prince, 1968), the virion structure of HBV was detected in Ag-positive patients serum (Dane, Cameron, & Briggs, 1970).

Soon after, the HAV viral particles were also identified by electron microscopy (EM) as well (Feinstone, Kapikian, & Purceli, 1973). Just in less than ten years after the identification of virus particles, the first significant attempts were made for the generation of HBV vaccines (Blumberg & London, 1982).

1.1.2. Discovery of hepatitis C virus

Hepatitis C was first discovered in the 1970s by two independent groups. Both groups almost simultaneously realized there was another virus causing liver damage that did not belong to HAV and HBV, leading to the emergence of terminology non-A and non-B hepatitis (NANBH) (Feinstone, Kapikian, Purcell, Alter, & Holland, 1975; Prince et al., 1974). Chimpanzees were the first model organism of NANBH and one of the main features of the virus discovered by EM, was the formation of double-membrane and tubular structures of the endoplasmic reticulum (ER). Next, the chloroform-inactivated NANBH was proved to be an enveloped virus (Bradley et al., 1985).

Filtration studies revealed that it had a size of 45-60 nm in diameter. The next fifteen years of efforts devoted to the cultivation, usual serological assays, and EM studies were all unsuccessful to further characterize the virus properties till the emergence of revolutionary molecular techniques (Trepo, 2014). In 1989, the genomic content of NANBH virus was identified using a novel molecular cloning approach, and it was shown to be derived from a new flavi-like virus, named “hepatitis C virus” (Choo et al., 1989). Since then, our understanding of HCV expanded enormously. Figure 1-3 summarizes the major breakthroughs in HCV research. (We will further explain the majority of these developments in upcoming sections.)

Figure 1-3. Timeline of the milestones in hepatitis C virus research. History of treatments for hepatitis C, significant developments in structural biology of hepatitis C virus and antiviral development. The figure is prepared based on the data obtained from the literature (R. F. Bartenschlager, Rice, & Sofia, 2016; Hajarizadeh, Grebely, & Dore, 2013; Scheel & Rice, 2013).

1.1.3. Taxonomy of hepatitis C virus

HCV is classified within the *Flaviviridae* family, which are positive-sense single-stranded RNA viruses (Figure 1-4a). The *Flaviviridae* family includes the flaviviruses such as dengue and yellow fever viruses, the animal pestiviruses such as bovine viral diarrhea virus (BVDV), and the pegiviruses (*e.g.* GBV-A, GBV-B, and GBV-C) as seen in Figure 1-4b (Robertson et al., 1998; Smith et al., 2014).

The main characteristic of the HCV is its high genetic diversity. Particularly, the envelope proteins are highly variable, and 5' and 3' non-coding regions have the highest level of conservation (Simmonds, 2004). There are seven major genotypes of HCV (named as 1, 2, 3, etc.) that vary in their nucleotide sequence by approximately 30-35%. Genotypes of HCV is divided into 67 subtypes those differ around 20-25% in their nucleotide sequences. Lastly, subtypes consist of several viral quasispecies (isolates) based on their genetic variations (Smith et al., 2014). Figure 1-4c shows the phylogenetic tree of HCV genotypes of 129 different complete coding sequences.

The genetic variations of HCV alter the pathogenesis, natural course, and treatment of the disease. Therefore, the genotype of an HCV infection is clinically important as it correlates to the patient response to antiviral therapy. So far, the patients infected with HCV genotype 1 were reported to show the lowest sustained viral response (SVR) by interferon- α (IFN- α) therapy, which was the most common HCV treatment to date (Schlutter, 2011).

Figure 1-4. The taxonomy of hepatitis C virus. (a) Baltimore classification of the pathogenic human viruses. HCV is in the *Flaviviridae* family of ssRNA viruses. (b) The members of *Flaviviridae* family are shown on the phylogenetic tree. The conserved motifs on RNA-dependent RNA polymerase are compared, and the distance scale represents amino acids substitutions (Romero-Brey & Bartenschlager, 2014). (c) Phylogenetic tree of HCV genotypes, 129 diverse and complete coding regions are exhibited (Smith et al., 2014).

1.1.4. Epidemiology of the disease

HCV is considered to have coevolved with humans for over 100,000 years and originated from Africa. HCV infects only humans and chimpanzees. The slave trade from Africa starting from the earlier 1700s likely spread the virus in a global scale (Oxford, 2016). And currently, HCV infection worldwide is estimated to be nearly 2-3% of the whole population, accounting for 130-150 million people (Gravitz, 2011). HCV mainly transmits through blood to blood associated with parenteral routes involving medical treatment, intravenous drug abuse, and sexual contact (Lavanchy, 2009).

1.1.4.1. Geographical distribution of hepatitis C virus

HCV is genetically highly diverse, and the geographic distribution of its genotypes varies significantly as shown in Figure 1-5. Accurate data for global distribution and prevalence of HCV is not available mainly for two reasons. First, most of the chronic infections are asymptomatic. And secondly, routine HCV screenings are not accessible in lower-income countries (Lavanchy, 2009). Although the prevalence of the genotypes varies significantly from country to country, the most prevalent types of HCV are genotype 1 subtypes (*i.e.* genotype 1a and genotype 1b) and genotype 3 (Hajarizadeh et al., 2013). It is estimated that genotype 1 subtypes and genotype 3 infect around 83.4 and 54.3 million people worldwide, respectively (Messina et al., 2015). The HCV infections may reach even above 10% of the local populations. In some countries, such as Egypt, the number of people infected with HCV is as high as >10% of the population. In contrast, some other (developed) countries managed to control HCV prevalence very remarkably. In the United States of America, HCV prevalence is estimated to be around 1%. In UK, Germany, and Canada, where routine tests are available, the prevalence is between 0.6 to 0.9% (Hajarizadeh et al., 2013).

Figure 1-5. The estimated prevalence of hepatitis C virus infections worldwide (Hajarizadeh et al., 2013).

1.1.4.2. HCV model systems

HCV can infect only human and chimpanzee. Therefore, the chimpanzee has been used as a model organism to study adaptive immunity and vaccine response (Bukh, 2012). However, current regulations limited the usage of chimpanzees as a model organism mainly due to ethical reasons. In order to compensate the gap, chimeric mice models and immunodeficient mice engrafted with human hepatocytes were used to explore the HCV entry and replication processes of HCV and innate immunity mechanisms developed in response to HCV infections (Mercer et al., 2001).

The discovery of subgenomic replicons (SGR) of HCV allowed studying of HCV molecular biology *in vitro* (Lohmann et al., 1999). SGR is a nucleic acid sequence that can self-replicate. However, as it lacks the sequences of structural proteins, and hence cannot form infectious virion. It also has a selection marker, *i.e.* neomycin resistance, that allows the replication of the SGR in Huh7-derived hepatoma cells under the selection marker. SGR of HCV gives an insight regarding the molecular biology of HCV proteins and their impact on HCV replication. However, the use of SGR has limitations

such as understanding the assembly and complete life cycle of HCV (Blight & Norgard, 2006).

The genotype 2a isolate from a Japanese patient, JFH1, was identified in 2005. JFH1 was very special as it did not need any cell culture adaptive mutations and could replicate in Huh7-derived cell culture by itself (Kato et al., 2003). Later, chimeric and adapted variants of JFH1 isolate resulted in *in vitro* model to study the complete life cycle of HCV from virus entry to assembly (Wakita et al., 2005). JFH1 HCV models allowed the monitoring of whole life cycle of HCV in *in vitro* conditions, which made an impact in facilitating further research activities (Ross-Thriepland & Harris, 2015).

1.1.4.3. Clinical features of HCV infection and long-term consequences

The outcomes of the HCV infection changes depending on the host and viral factors. The host factors, such as ethnicity, gender, alcohol, and drug abuse, and the viral factors such as the genotype of the HCV and co-infection with other viruses differ the progression of the infections (Chen & Morgan, 2006).

Once the HCV infection occurs, the virus might be eliminated by a strong immune response of the patients in 10-40% of the acute hepatitis cases. However, the majority of the infections, which accounts 60-90% of infected individuals, progress as a persistent and slowly progressing asymptomatic disease. As chronic hepatitis establishes, it may either stay as an asymptomatic persistent infection for several years, or it may provoke severe liver diseases like hepatocellular carcinoma (HCC) and cirrhosis as summarized in Figure 1-6 with estimated numbers (Scheel & Rice, 2013). HCC is a prevalent cancer type with 610,000 deaths worldwide, and around 75% of the cases are HCV related. The total number of HCV-associated cirrhosis cases is not clear. According to a report, there were around 738,000 deaths due to cirrhosis in 2002, and ~27% of them were HCV-related (McGlynn & London, 2011).

Figure 1-6. Outcomes associated with hepatitis C virus infection. The schematic representation was produced based on the data from (Scheel & Rice, 2013).

1.1.4.4. Treatment of hepatitis C disease

The success of HCV treatments is measured by sustained virologic response (SVR), which is defined as the absence of detectable HCV RNA viruses in the serum by an assay after six months following the completion of treatment (Pearlman & Traub, 2011). Combination therapies partially achieved treatment of hepatitis over last fifteen years (Rice, 2011). Until recently, the most common regimen was the combination of PEGylated IFN- α 2 and ribavirin (drug), which gave an SVR rate between 48% (genotypes 1, 4, 5, and 6) to 88% (genotypes 2 and 3) (Heim, 2013; Poynard, Yuen, Ratziu, & Lai, 2003). Currently, this type of combinational therapy is losing popularity as a new class of drugs emerged that directly targets HCV proteins, hence called as direct-acting antivirals (DAA) (Bidell, McLaughlin, Faragon, Morse, & Patel, 2016). Figure 1-7 summarizes the treatment of HCV infections and their SVR rates between 1989 till today.

Figure 1-7. The advances of HCV treatments from 1989 till now. The picture is reproduced and updated from (Heim, 2013). Recombinant IFN- α was the first therapy that had a very low SVR *i.e.* less than 25%. The combination of IFN- α with an analog of ribonucleic (ribavirin) improved the therapy up to 38% SVR. The PEGylated IFN- α increased the treatment success 10% more and became a standard treatment for HCV infections for ten years. The first DAAs (telaprevir and boceprevir) were approved in 2011 and provided an SVR up to 75%. In 2014, the NS5A inhibitors were approved, and the patients who treated with them resulted in an SVR rate up to 98% (Pol, Corouge, & Vallet-Pichard, 2016).

There are four classes of DAAs classified by their therapeutic targets, NS3/4A protease inhibitors (PIs), NS5B nucleoside polymerase inhibitors (NPIs), NS5B non-nucleoside polymerase inhibitors (NNPIs) and NS5A inhibitors. These newly emerging drugs dramatically changed the landscape of treatment regimens over the last three years, which can be easily assessed by comparing the success rates of conventional treatment and DAAs. Using DAAs, combined therapeutic regimes are accessible based on the genotype of the virus, and all have very similar success rates up to 98% SVR (Afdhal & Serfaty, 2016). Table 1-1 shows the DAAs currently available in the market with their combinatorial regimens.

Table 1-1. Current hepatitis C drugs and their mechanism of action (Pockros, 2016).

Therapeutic target	Drugs	Mechanism
NS5B	Sofosbuvir (Nucleoside inhibitor)	Binds to the catalytic site of RNA-dependent RNA polymerase
	Dasabuvir (Non- Nucleoside inhibitor)	Binds to near the active site of the polymerase
NS5A	Daclatasvir Elbasvir Ledipasvir Ombitasvir Velpatasvir	Not completely known, inhibits the viral RNA replication and the assembly
NS3/4A	Grazoprevir Paritaprevir Simeprevir	Inhibits the processing of the HCV polyprotein

1.2. Structural and Genomic Features of HCV

1.2.1. Structural organization of HCV viral particle

HCV is an enveloped virus, thus covered by a lipid bilayer and envelope glycoproteins. Like all members of *Flaviviridae* family, HCV particles have three structural proteins involved in their formation: E1, E2, and core protein. E1 and E2 are glycosylated envelope proteins that have multiple functions and can anchor to the membrane of the host with an ER retention signal sequences. Inside the lipid envelope, the icosahedral nucleocapsid layer (formed by core proteins) lies. Besides proteins, the last component of the HCV particle is single-stranded (+) strand RNA genome that is located inside the nucleocapsid structure (Op De Beeck & Dubuisson, 2003). Figure 1-8 illustrates the organization of an HCV particle.

The shape of HCV particles does not seem icosahedral, and they are very

pleomorphic as they are associated with low-density and very-low-density lipoproteins (Andre et al., 2002). The neutralization of HCV by antibodies is a major objective for vaccination approaches for both the prevention HCV infections and as a strategy to prevent re-infection cases. Studies of the HCV neutralizing response improved our understanding regarding the epidemiology of the chronic viral infection. However, effective antibody-based vaccines and treatments for the disease are still not available (Ball, Tarr, & McKeating, 2014). One mechanism suggested so far is that the close-proximity of the HCV particles with lipids may protect the virus from neutralization by antibodies (R. Bartenschlager, Penin, Lohmann, & Andre, 2011). And very recently, researchers showed that highly conserved region of the E2 protein is very flexible making the development of a vaccine against HCV very challenging (Kong et al., 2016).

Figure 1-8. The schematic structure of hepatitis C virus particle. HCV particle contains E1 and E2 envelope proteins outside the lipid envelope.

1.2.2. Genomic organization of hepatitis C virus

The HCV genome is a (+) sense 9.6 kb ssRNA that containing only one open reading frame terminated by 5' and 3' noncoding regions (NCRs). This single open reading frame is translated into one polyprotein product of around 3100 amino acids, which are further cleaved by viral and host proteases. Subsequently, the precursor polyprotein is processed to generate structural proteins (*i.e.* E1, E2, and core) and nonstructural (NS) proteins (*i.e.* NS2, NS3, NS4A, NS4B, NS5A, NS5B) (Penin, Dubuisson, Rey, Moradpour, & Pawlotsky, 2004). Cleavage also results in the formation of a small

protein, p7 that forms ion-channel and acts as a viroporin (S. D. Griffin et al., 2003; Pavlovic et al., 2003) and, it is proposed that p7 is a part of the virion (S. Griffin et al., 2008). Figure 1-9 shows the ss(+)RNA of HCV, its corresponding polyprotein, and products of the post-processed polyprotein.

Figure 1-9. Genomic organization of hepatitis C virus. A schematic representation of HCV genome *i.e.* positive-sense 9.6 kb of single-stranded RNA is shown. A single polyprotein precursor generated by IRES-dependent RNA translation. The polyprotein is processed by proteases and generates structural proteins (orange colored): core (C), envelope 1 (E1), envelope 2 (E2) and nonstructural proteins (NS), NS2, NS3, NS4A, NS4B, NS5A, and NS5B (blue colored) and p7.

1.2.2.1. 5' and 3' noncoding regions (NCRs) of genomic RNA

NCRs are the nucleotide sequences, located at the 3' and 5' ends of the genomic RNA, which are not translated into protein but are crucial for the replication and

translation of viral proteins (Song et al., 2006). The 5' NCR is a highly-conserved sequence comprising 341 nucleotides. It has four domains, named as I, II, III, and IV based on the analysis of their secondary structure (Brown, Zhang, Ping, & Lemon, 1992). The 5' NCR has an internal ribosome entry site (IRES), and the structure of IRES facilitates the cap-independent translation of viral RNA by binding 40S ribosomal subunit of the host in eukaryotic organisms (Ji, Fraser, Yu, Leary, & Doudna, 2004). Upstream sequences of IRES are necessary for virus replication, and first 125 nucleotides are sufficient for viral replication. However, such replicons having only 125 nucleotides exhibit an impaired phenotype for multiplication. Therefore, complete 5' NCR is more efficient for viral RNA synthesis (Friebe, Lohmann, Krieger, & Bartenschlager, 2001).

The 3' NCR contains a variable region, poly(U/UC) tract and a highly conserved 80 nucleotides, forming a tripartite structure (Tanaka, Kato, Cho, Sugiyama, & Shimotohno, 1996). HCV replicon systems were employed for the characterization of the regions that are necessary for RNA replication. It was revealed that the complete deletion of variable region is tolerated. However, a significant reduction in replication was observed in cell culture studies. Replicons having a poly(U/UC) region replaced by homouridine of 26 nucleotides were found to capable of replication, whereas when this region was replaced by homopolymers of adenine, guanine, cytosine, and uridine sequences, the system was unable to replicate (Friebe & Bartenschlager, 2002; Yi & Lemon, 2003). The highly-conserved region comprises three stable stem-loop structures, called X-tail or 3' X. The deletion of individual or all stem-loop structures on 3' X resulted in inactive replicons (Friebe & Bartenschlager, 2002). Moreover, the 3' NCR has a role in enhancing IRES-mediated translation of viral RNA by interacting with 5' end of viral RNA (Song et al., 2006). Notably, 3' NCRs increase the efficiency of translation by terminating translation process and likely involving in ribosome recycling (Bradrick, Walters, & Gromeier, 2006).

1.2.3. Hepatitis C virus proteins

The polyprotein of HCV produces three structural (*i.e.* core, E1 and E2), six nonstructural proteins (*i.e.* NS2, NS3, NS4A, NS4B, NS5A, and NS5B), and p7 protein. We will introduce the function of these proteins in HCV life cycle below. (The structures

and their membrane associations are illustrated Figure 1-10.)

Figure 1-10. The hepatitis C virus proteins on endoplasmic reticulum membrane are represented. Cellular ER signal peptidases cleave the regions between core, E1, E2, p7, and NS2, shown with scissors. Noted that known structures are shown with ribbon representations, and the others are given as generic elliptic shapes. The figure is taken from (Moradpour, Penin, & Rice, 2007).

1.2.3.1. Structural proteins of HCV

1.2.3.1.1. Core protein

Core protein is one of the structural proteins of HCV, placed at the N-terminus of the viral polyprotein, which forms the viral capsid. Once the polypeptide is translated, it is targeted to the ER membrane. Immature protein has 191 amino acids (23 kDa), with an E1 signal peptide at C-terminus. Then, it undergoes cleavage processes and transforms into the mature (21 kDa) form of core protein (Hope, Murphy, & McLauchlan, 2002).

Core proteins are located on lipid droplets (LD), the intracellular storage sites for triacylglycerols and cholesterol esters. These lipid droplets are assumed to serve as virus assembly sites (Hope & McLauchlan, 2000). The core protein is a dimeric α -helical protein that has two domains based on their hydrophobicity. Domain I of core protein has a positive charge and interacts with RNA (Boulant, Vanbelle, Ebel, Penin, & Lavergne, 2005). And, domain II core protein is crucial for targeting of core proteins to LDs. The deletion of this domain II from core sequence blocks the protein-lipid interaction (Boulant et al., 2006). Alanine scanning mutation analysis of the core of

HCV showed the importance of core protein for virus production (Murray, Jones, Tassello, & Rice, 2007). The core protein interacts with a variety of cellular proteins that might have a role in the natural course and pathogenesis of the hepatitis C. Those host proteins have roles in apoptosis (TNF- α , Fas), cell cycle, and regulation of gene expressions (c-myc) (Tellinghuisen & Rice, 2002).

1.2.3.1.2. Envelope 1 and 2 proteins

The HCV envelope proteins (E1 and E2) are essential proteins due to their roles in viral attachment and cell fusion. Therefore, their potential uses in vaccine and antiviral strategies became very critical in HCV research. E1 (30-35kDa) and E2 (70-72 kDa) are glycoproteins anchored on the envelope of virus particles, and they form heterodimers that function for the entry of virus (Dubuisson, 2007). Both proteins have C-terminal transmembrane domains that anchor the protein on ER membrane, whereas the N-terminal of E1 and E2 stay in the ER lumen. E1 and E2 have glycosylation motifs, and they undergo N-linked glycosylation that helps glycoprotein folding and facilitate HCV entry as well (Goffard et al., 2005).

E2 seems to associate with the cellular receptor(s) while E1 is suggested to process fusion activity (Op De Beeck, Cocquerel, & Dubuisson, 2001). As mentioned before, E1 and E2 are highly variable. Hypervariable region of E2 can differ as much as around 80% of within HCV isolates. However interestingly, although the sequence changes a lot, the three-dimensional structure of the protein is found to be highly conserved (Penin et al., 2001).

The crystal structures of the core of E2 (Khan et al., 2014; Kong et al., 2013), have revealed significant findings to depict the molecular interactions with CD81 and broadly neutralizing antibodies. Three available structures showed that E2 of HCV is primarily globular fold with a basic front face and a more hydrophobic back layer. The basic face contains a central β -sandwich motif, and back face contains another β -sheet that is almost perpendicular to the central β -sandwich. CD81 interaction site is shown to be at the interface of these two layers (Khan et al., 2014; Kong et al., 2013).

The crystal structure of the N-terminal of E1 protein (flexible regions excluded) was solved in 2014 (El Omari et al., 2014). The structure of N-terminal E1 consists a β -hairpin and 16 amino acids long α -helix at the N-terminal, followed by three strands

antiparallel β -sheet. Their oligomeric arrangement showed that they exist as a dimer of dimers.

The crystallization studies of the major envelope proteins in the *Flaviviridae* family showed elongated structures (Vieyres, Dubuisson, & Pietschmann, 2014). The class II fusion proteins insert their hydrophobic region into the cell membrane and refold to mediate fusion. However, unexpectedly, the crystal structures of the N-terminal ectodomains of E1 and E2 showed more compact and globular structures compared to class II fusion proteins. Therefore, HCV seems to have an alternative fusion mechanism from other viruses (Freedman, Logan, Law, & Houghton, 2016). Therefore, the mechanism of the fusion of HCV and the formation of the heterodimers of E1/E2 are still incomplete and growing area that may facilitate the emerging of new antivirals against HCV infection.

1.2.3.2. Nonstructural proteins

1.2.3.2.1. Nonstructural 2 and 3 proteases

NS2 is a 23 kDa protein that has multiple N-terminal transmembrane domains localizing on ER membrane and a C-terminal protease domain. Functionally, it is a protease responsible for the autoproteolytic cleavage of the junction between NS2 and NS3. It has been shown that this cleavage requires NS3 serine protease. NS3 enhances the activity of NS2 by the cofactor activity of its domain containing a zinc atom (Pieroni et al., 1997).

The structure of C-terminal catalytic part of NS2 was shown to have a globular structure placed in the cytosolic part of ER (Lorenz, Marcotrigiano, Dentzer, & Rice, 2006). NS2 is dispensable for viral replication. However, its activity is shown to be crucial for the formation of infectious particle production (Lohmann et al., 1999; Pietschmann et al., 2006).

1.2.3.2.2. Nonstructural 3-4A protein complex

NS3 (68 kDa) has multiple functions having an N-terminal serine-type protease activity and C-terminal RNA helicase/NTPase activity. NS4A is 6 kDa in size with two parts: an ER-anchoring N-terminal hydrophobic sequence and a C-terminal acidic

component that has an effect on NS5A phosphorylation and RNA replication (Lindenbach, Pragai, et al., 2007). NS3 interacts with NS4A, and together they localize to ER membrane through the N-terminal part of NS4A transmembrane domain (Wolk et al., 2000). The serine-type protease domain of NS3 associates with NS2 to enhance the cleavage of the NS2/NS3 junction. Further, NS3 cleaves the links between NS3/NS4A, NS4A/NS4B, NS4B/NS5A, and NS5A/NS5B (R. Bartenschlager, Ahlborn-Laake, Mous, & Jacobsen, 1993). The C-terminal RNA helicase/NTPase binds to RNA as a monomer, and the dimerization of NS3 facilitates unwinding of RNA (Serebrov & Pyle, 2004). Therefore, it has been proposed to be involved in unwinding RNA secondary structures at the termini of viral RNA (Dubuisson, Helle, & Cocquerel, 2008).

1.2.3.2.3. Nonstructural 4B protein

NS4B is 27 kDa in size, and it is tightly bound to ER membrane with four transmembrane domains. The current knowledge regarding the structure of NS4B is very limited as its very hydrophobic nature makes its investigation challenging. Consequently, the exact orientation of NS4B on ER membrane is uncertain, although some parts are thought to reside on either cytosolic or luminal sides of the ER membrane (Lundin, Monne, Widell, Von Heijne, & Persson, 2003). The expression of NS4B involves in the rearrangement of cellular ER membranes resulting in the formation of membrane-associated foci that are critical for viral replication (Egger et al., 2002). Moreover, the mutagenic analysis of NS4B demonstrated that NS4B influences the replication of HCV genome (Lindstrom, Lundin, Haggstrom, & Persson, 2006).

1.2.3.2.4. Nonstructural 5B protein

NS5B is an RNA-dependent RNA polymerase (RdRp) that performs the HCV replication. It synthesizes a complementary (–) strand RNA from the genome of HCV and then it subsequently produces (+) strand RNAs by using (–) strand RNAs as templates. The enzymatic functions of the protein were characterized at biochemical and structural level (Bressanelli, Tomei, Rey, & De Francesco, 2002; Lesburg et al., 1999). Like all (+) strand RNA viruses, nonstructural proteins of HCV are thought to form a replication complex on the membrane (Shirota et al., 2002). It has been shown that NS5A and NS5B directly interact with each other (Shirota et al., 2002).

1.2.4. The life cycle of hepatitis C virus

The life cycle of HCV contains six steps: (1) viral attachment through several cell receptors, (2) cytoplasmic release and uncoating of the viral genome, (3) HCV RNA translation and post-translational modifications of the polyprotein, (4) formation of RNA replication sites and RNA replication, (5) virus packaging and assembly, and (6) virion maturation and release. Figure 1-11 illustrates the viral life cycle.

Figure 1-11. Life cycle of hepatitis C virus

1.2.4.1. Virus attachment, entry, and uncoating

Viral attachment and entry of HCV are likely multi-step processes that require several receptors on the host cell and unfortunately, it is not fully understood yet. The receptors proposed are CD81, a tetraspanin molecule found on hepatocytes (Pileri et al., 1998), the low density lipoprotein receptor (LDLR), glycosaminoglycans (Agnello, Abel, Elfahal, Knight, & Zhang, 1999), scavenger receptor class B type I (SR- BI) (Scarselli et al., 2002), claudin-1 (CLDN1), and occludin (OCLN) (S. Liu et al., 2009).

LDLR and glycosaminoglycans regulate the low-affinity cell association, and the virus can bind to the specific receptor-like CD81 and SR-BI receptors through E2 of

HCV (Bartosch & Dubuisson, 2010). Then, CLDN1 and OCLN are also necessary for the entry of the virus. They form tight junction complexes, which are suggested to take a role in the late-stage fusion of the virus (S. Liu et al., 2009). Clathrin-mediated endocytosis processes the viral uptake that occurs in low-pH compartments, probably in endosomes (Koutsoudakis et al., 2006; Tscherne et al., 2006). The late stages of the HCV uptake and viral uncoating are poorly understood. The structural basis for low-pH-induced membrane fusion has been studied for other viruses, like dengue and TBE. It is shown that an internal fusion peptide of enveloped proteins is exposed in low-pH environments of the membrane and undergoes conformational changes (Bressanelli et al., 2004). As the scaffolds of these proteins are similar within *Flaviviridae* family, including HCV, it is suggested to behave the same way in the entry process of HCV (Lindenbach & Rice, 2013; Moradpour et al., 2007). The association of E2 with CD81 is thought to activate the HCV glycoproteins to react to low pH in the endocytic compartment and to induce the uncoating of the viral envelope (Sharma et al., 2011)

1.2.4.2. Translation and polyprotein posttranslational modification

HCV genome has an ordered structure, 5'NCRs, 3'NCRs and the first 30nt of the core -coding region. ER-associated translation is initiated by binding of the 40S ribosomal subunit to the IRES located in the HCV 5' NCRs (Hoffman & Liu, 2011). The precursor polyprotein is co- and post-translationally processed by cellular and viral proteases to form mature HCV proteins. The structural proteins (E1, E2, core) and p7 are cleaved by ER signal peptidase, and nonstructural proteins are cleaved by NS2-3 and NS3-4A complex proteases of HCV (Grakoui, Wychowski, Lin, Feinstone, & Rice, 1993).

1.2.4.3. The replication of hepatitis C virus genome

HCV, similar to other (+) strand RNA viruses forms a membrane-associated replication complex that includes viral proteins, RNA, and altered membrane structures (Egger et al., 2002). HCV replication occurs at DMVs, derived from ER membrane and named as “membranous web” (Moradpour et al., 2003). Although the DMV formation is induced by the action of several viral proteins such as NS3-4A, NS4B, NS5A and NS5B, only the domain I of NS5A (NS5A-DI) and the helicase domain of NS3 are found to be required for DMV biogenesis (Romero-Brey et al., 2015).

The NS3 protein has a helicase domain that might have a role in the unwinding of RNA structures during replication (Raney, Sharma, Moustafa, & Cameron, 2010). The RNA-dependent RNA polymerase, *i.e.* NS5B, catalyzes the replication process. The (+) strand RNA is used both as the genomic information and as messenger RNA as well for the synthesis of (–) strand RNA. Next, the (–) strand RNA is utilized to produce (+) strand RNA and genomes of HCV particles (R. Bartenschlager, Frese, & Pietschmann, 2004).

1.2.4.4. The assembly and release of hepatitis C virus

The late stages of HCV life cycle are unclear, but it is thought to be regulated by host cell lipid synthesis (R. Bartenschlager et al., 2011). The mature core protein is transferred from ER- membranes to near cytoplasmic lipid droplets, the places for virus assembly (McLauchlan, Lemberg, Hope, & Martoglio, 2002). The current nucleocapsid formation model includes the association of the core protein with NS5A protein (Masaki et al., 2008).

1.3. NS5A Protein

NS5A is a membrane-associated multifunctional phosphoprotein that contains three domains separated by two low-complexity sequences (LCI and LCII). Due to its potential function(s) in IFN signaling pathway, NS5A once became popular in the field (Enomoto et al., 1995). Recently, it attracted broad research interest as the several HCV inhibitors that are in current use were found to be targeting NS5A, albeit the mechanism of their inhibition is not known yet (Berger et al., 2014; Kwon et al., 2015; Nakamoto, Kanda, Wu, Shirasawa, & Yokosuka, 2014; O'Boyle et al., 2013; Sun et al., 2015). It is thought that NS5A mediates its multiple functions through its interactions with viral and host cell proteins. NS5A is an active player of HCV replication machinery with no known enzymatic function (de Chassey et al., 2008; Tellinghuisen, Foss, Treadaway, & Rice, 2008). And, it has other roles in cellular processes from innate immunity to dysregulated cell growth and assembly of viral particles (Macdonald, Crowder, Street, McCormick, & Harris, 2004). (Figure 1-12 illustrates the organization of NS5A protein indicating important residues/regions, all of which will be further explained in next sections.)

Figure 1-12. The schematic organization of NS5A protein. AH is an amphipathic helix, from 1 to 33 aa, domain I (DI) is from 1 to 213, low complexity region I (LCI) is from 213-250, domain II (DII) is from 250 to 342, low complexity region II (LCII) is from 342 to 356, domain III (DIII) is from 356 to 447. LCI, LCII, and DIII are the main phosphorylations sites. Cartoon representations of the structures are produced using Protein Data Bank (PDB) entries: membrane anchor domain 2AJN (Sapay et al., 2006) and Zn-binding domain 1zh1 (Tellinghuisen, Marcotrigiano, & Rice, 2005).

1.3.1. Structure and Function of NS5A

The structure of complete NS5A is not available. However, the structures of different domains are solved. We will briefly introduce these studies below by correlating with their identified functions.

The very N-terminal region of NS5A is a highly conserved alpha helix and the NMR structure of this peptide is available (Brass et al., 2002). It is an amphipathic helix (AH) necessary to mediate the association of NS5A with ER membrane. This region is proposed to lie parallel to the ER membrane such that hydrophobic residues are buried within the bilayer while hydrophilic residues stay on top of the membrane exposed to the surface (Penin, Brass, et al., 2004).

DI of NS5A (NS5A-DI) has a tetracysteine motif coordinating a Zn atom, which is crucial for RNA replication (Tellinghuisen, Marcotrigiano, Gorbalenya, & Rice, 2004).

DI is the well-characterized part of NS5A where X-ray crystallographic structures of DI reveal that it can form four different dimeric conformations, shown in Figure 1-13. The monomer structures are virtually the same in all dimer forms of DI and contain two subdomains. The pioneering work, the first crystal structure of NS5A-DI from genotype 1b, was published more than ten years ago (Tellinghuisen et al., 2005). It reports an asymmetric dimer structure with an intermolecular interface of ~ 830 Å, seen in Figure 1-13a. This work showed that DI harbors disulfide bonds that maintain the overall fold. (See the yellow sticks given in Figure 1-13a.) This interface accommodates a large groove, which is proposed as the RNA-binding area. Second crystal structure of NS5A-DI from genotype 1b is an alternative dimer structure of two parallel monomers with an interface of 910 Å as seen in Figure 1-13b (Love, Brodsky, Hickey, Wells, & Cronin, 2009). The recent study revealed two alternative dimers of NS5A-DI from genotype 1a of HCV that can form a tetrameric structure, seen in Figure 1-13c (Lambert et al., 2014). Very importantly, although all these crystal structures of NS5A-DI showed the existence of dimers, there is no satisfactory biochemical evidence regarding the formation of dimers under *in vitro* conditions (Hwang et al., 2010; Lim et al., 2012). (We will further discuss this point in the upcoming sections of the thesis in the light of our findings.) It is also essential to emphasize that it is still unknown how, and in which conformation(s), the NS5A protein carries out its function(s). In the current status, the biological relevance of the dimerization of NS5A is a matter of debate (Ross-Thriepland & Harris, 2015). The Figure 1-14 shows the electrostatic potential of the crystal structures for the domain I of NS5A.

Figure 1-13. The three different crystal structures for the domain I of NS5A protein. The orange spheres in structures show the primary resistance mutation of NS5A inhibitors, tyrosine 93. (a) The “clam-like” dimer from genotype 1b with a large groove forming a potential RNA binding pocket (from PDB entry of 1zh1). (b) “Back-to-back” dimer from genotype 1b with a parallel orientation of monomers (from PDB entry of 3fqg). (c) The tetramer from genotype 1a from PDB entry of 4cl1 (Lambert et al., 2014).

DII and DIII of the protein are reported to have a flexible and disordered structure in nature, shown by NMR and biophysical methods (Hanouille et al., 2009; Liang, Kang, & Yoon, 2006). The structural flexibility of these regions is thought to be important to mediate its rather complex interactions of NS5A (de Chassey et al., 2008). Recently, an NMR study of DII with the LCI and LCII regions showed the existence of three α -helical folds (Feuerstein et al., 2012).

Figure 1-14. Electrostatic potential of the three different crystal structures for the domain I of NS5A protein. . (a) and (b) The “clam-like” dimer from genotype 1b with a large groove forming a potential RNA binding pocket (from PDB entry of 1zh1). (c) and (d) “Back-to-back” dimer from genotype 1b with a parallel orientation of monomers (from PDB entry of 3fqg). (e) and (f) The tetramer from genotype 1a from PDB entry of 4cl1 (Lambert et al., 2014).

1.3.2. Interacting partners of NS5A

Over one hundred interactions of NS5A with different proteins are documented so far (Han, Niu, Wang, & Li, 2016). It is suggested that NS5A might present a variety of isoforms that differ in phosphorylation patterns to mediate such an expanded protein-protein interaction (PPI) network (Ross-Thriepland & Harris, 2015). (One may also

consider different possible conformations of NS5A and perhaps their varying spatial localizations within the cell in connection with the possible different interactions of the protein.)

NS5A binds to the cellular SH3-domain-containing proteins through a conserved motif, PxxPxR, present on the LCII of protein (Hughes et al., 2009). The mutagenesis of this motif had no effect on the replication and assembly of the virus. However, this motif was critical for the inhibition of apoptotic response against oxidative stress (Mankouri et al., 2009). An ER-located form of lipid kinases, *i.e.* type III phosphatidylinositol 4-kinase alpha (PI4KIII α), binds to NS5A and this binding event activates the kinase (Reiss et al., 2011). The activation of PI4KIII α results in the increase of phosphatidylinositol 4-phosphate (PI4P) lipids that are required for the alteration of cytoplasmic membranes. These membranous structures serve as the replication site of the virus. So far, seven residues on the C-terminus of the DI were identified as the binding sites for PI4KIII α (Reiss et al., 2013).

Cyclophilin A is a protein with a peptidyl-prolyl isomerase activity that is essential for HCV replication as cyclophilin inhibitors suppress viral RNA binding and assembly of infectious HCV (Liu, Yang, Robotham, & Tang, 2009). The association of DII of NS5A with cyclophilin A stimulates RNA-binding in an isomerase dependent manner (Foster, Galloway, Stonehouse, & Harris, 2011). NS5A is involved in cellular signaling pathways, too (Macdonald & Harris, 2004). Such as Ras-ERK pathway, the interaction of NS5A with this pathway alters cell growth. It also interacts with protein kinase R (PKR), an IFN-induced protein. This interaction prevents the shutting down of protein translation, which is suggested to enhance the translation of virus protein (Gale et al., 1997). NS5A also interferes with MAPK signaling and inhibits the MAPK-ERK pathway (Macdonald et al., 2003). These findings and others suggest that NS5A could interfere with cellular pathways to improve the HCV life cycle. Therefore, NS5A is an emerging target for the inhibition of HCV replication as it has an essential role in RNA replication of HCV, although its roles are not fully clear yet (Buhler & Bartenschlager, 2012).

1.3.3. Phosphorylation sites of NS5A

NS5A is highly phosphorylated, and SDS-PAGE analysis detects two forms, *i.e.* basally phosphorylated (56 kDa) and hyperphosphorylated (58 kDa) form (Y. Huang,

Staschke, De Francesco, & Tan, 2007). The phosphorylation sites on the protein are complex and still under investigation. The majority of the phosphorylations occur on serine residues that lie on LC1, DII, and DIII (Y. Huang et al., 2007). The phosphorylation status of NS5A is proposed to be a control mechanism for viral replication as reducing hyperphosphorylation results in enhanced RNA synthesis (Appel, Pietschmann, & Bartenschlager, 2005).

The phosphorylation sites may vary in different genotypes of the virus. For example, the phosphorylation site on LC1 region of the protein was reported as only one residue (pS222) and suggested to regulate the replication (Lemay, Treadaway, Angulo, & Tellinghuisen, 2013). In other studies (Masaki et al., 2014; Ross-Thriepland & Harris, 2014), more than one sites were identified on the LC1 region of the protein. LCII also contains several phosphorylation sites on the protein, and those are correlated with HCV replication (Ross-Thriepland & Harris, 2014). DIII of the protein has phosphorylation sites, and their alanine mutagenesis caused impaired virus assembly (Masaki et al., 2008; Tellinghuisen et al., 2008).

1.3.4. NS5A is an RNA-binding protein

The interaction of NS5A with RNA has been shown *in vitro* by different groups (Foster, Belyaeva, Stonehouse, Pearson, & Harris, 2010; L. Huang et al., 2005). NS5A interacts with 3'-ends of HCV (+) and (–) strand RNAs. The polypyrimidine tract of 3'-end of the (+) strand, which is crucial for viral replication and was identified as a binding site for NS5A (L. Huang et al., 2005).

NS5A has a high binding affinity to single-stranded and G/U-rich RNA sequences. Unphosphorylated NS5A, casein kinase II phosphorylation sites deleted protein, showed an evidence of a more stable binding to RNA (L. Huang et al., 2005). In another report, the biochemical characterization of the RNA-NS5A protein was further studied (Hwang et al., 2010). They reported that 5 mM MgCl₂, 100 mM NaCl, less than 1 mM EDTA and pH 7.5 are optimal conditions for RNA-NS5A interactions. It was also found that a 5 to 6 nucleotide RNA size is enough for NS5A binding and the C-4 carbonyl of uridine base exhibits specific binding to NS5A. Lastly, they showed that DI and LCI are the mandatory binding sites on NS5A (Hwang et al., 2010). Whereas another study revealed that all three domains of NS5A can bind to RNA (Foster et al., 2010).

Another recent study aimed to understand the effect of potent NS5A inhibitors (BMS-790052 and AZD7291) on the protein and its RNA binding behavior (Ascher et al., 2014). According to this study, inhibitors directly bind to DI of NS5A. Based on their microscale thermophoresis based *in vitro* results, inhibitor binding decreases the RNA binding, and at the same time, RNA binding diminishes the inhibitor binding as well. However, this model seems limited to explain the seemingly nonstoichiometric interactions of drugs and NS5A, although some reports give approximate protein-to-drug values. (We will touch this point later on in length.)

Moreover, it was shown that monomer-dimer equilibrium shifts toward dimer side when U-rich RNA is present, which suggests that the dimerization of NS5A is RNA-binding mediated (Hwang et al., 2010). In one of the crystal structure studies of NS5A-DI (Lambert et al., 2014), the RNA-binding site of NS5A was estimated based on the sulfate groups bound to the crystal structure, which originally came from the crystallization buffer. The authors considered both sulfates and RNAs are negatively charged so that the sites where sulfates groups bind should mimic for the binding site of RNAs, as well. As a result, the authors suggested Arg41 and Arg81 sites as possible RNA-interacting sites on NS5A.

Taken together, despite great efforts to understand the association of NS5A with RNA molecules *in vitro*, the exact binding motifs and the detailed binding mode of NS5A and RNA molecules is still not known. In accordance, the functional value of this relationship is still vague (Ross-Thriepland & Harris, 2015).

1.3.5. Emerging inhibitors targeting NS5A

NS5A is a validated drug target for treating HCV infections. The cell-based HCV replicon systems have been used to screen a large database of small molecules. And as a result, several inhibitors emerged against HCV infection (Lemm et al., 2011; Lemm et al., 2010; Macdonald & Harris, 2004; O'Boyle et al., 2005). Very importantly, the sequence analysis of drug-resistant HCV mutants reveals amino acid substitutions in NS5A protein. The primary inhibitor resistance mutations mapped to NS5A-DI, which are all in the first one hundred amino acid sequence of N-terminus (Fridell, Qiu, Wang, Valera, & Gao, 2010; Lemm et al., 2010). The first of these inhibitors was discovered by the company Bristol-Myers Squibb and named as BMS-790052, which was named daclatasvir (DCV) later. DCV has an impressive potency and specificity with half

Figure 1-16. 2D chemical structures of direct-acting antivirals in use targeting NS5A. (a) DCV-HCl/BMS-790052 (Bristol-Myers Squibb), (b) LDV/GS-5885 (Gilead) (c) Ombitasvir/ABT (Abbott). Note that all these three drugs have dimeric symmetry in their chemical structures. And, the dimeric symmetry of DAAs is widely thought to be significant on their drug activity presumably through targeting the dimeric form of NS5A (Gao et al., 2010; Lambert et al., 2014).

Although NS5A-targeting drugs are in current clinical use, our understanding towards their function is quite limited. A significant effort has been applied on NS5A and DAAs resulting in a dramatic increase in the number of publications in the last five years. However, the mode of actions (MoA) of DAAs remains a mystery. We will represent some of the crucial findings regarding the MoAs of NS5A inhibitors below. (See Figure 1-17 for the summary of these findings.)

Figure 1-17. Possible effects of NS5A inhibitors on the protein

First, the incubation of biotin-tagged derivative of DCV resulted in the immunoprecipitation of NS5A. Therefore, DCV seems directly binding to NS5A protein. Again in the same study, considering the dimeric nature of drug and dimeric structures of the protein, the researchers proposed DCV might be stabilizing NS5A dimers (Gao et al., 2010). By using UV-activated crosslinking and affinity isolation techniques, DCV and some other DAAs were found to interact with NS5A with high affinity and specificity (Kwon et al., 2015; O'Boyle et al., 2013). There is an ongoing controversy about the stoichiometry of NS5A-DAA interactions. Based on an *in vitro* radioligand binding assay, the stoichiometry of NS5A-LDV interaction is 1 LDV molecule per 15 NS5A monomers / 7.5 dimers (Kwon et al., 2015). In another study, by employing an *in vitro* assay based on microscale thermophoresis method, it was shown that DCV binds tightly to NS5A-DI and reduce RNA binding. The same study revealed that the compounds did not affect NS5A dimerization (Ascher et al., 2014). Furthermore, it was demonstrated that DCV might change the subcellular localization and fractionation patterns of NS5A (Lee et al., 2011; Targett-Adams et al., 2011). Then, in a cell-based assay, fluorescent tags covalently attached to NS5A measured the intracellular proximity of NS5A proteins. DCV treatment resulted in the relocation of both N and C termini of NS5A, including the disruption of the close positioning of the N-termini of two different NS5A molecules in a multimolecular complex.

Interestingly, the existence of large aggregates of NS5A also reported in the same study (Bhattacharya et al., 2014). In another important recent study, by inserting a fluorescent protein to the C-terminal domain of NS5A, it is shown that DCV decreases the motility of the protein within the cell and also causes hyper-accumulation of PI4P (Chukkapalli et al., 2015). And even perhaps more critically, they also clearly observed that DCV leads to the aggregation of not only NS5A but also NS5B. (The involvement of DCV in the reduced motility and increased aggregation behavior of NS5A indicate the critical role of DCV in NS5A dynamics. We will later discuss this observation in connection with the *in vitro* results presented in upcoming sections of the thesis.) On the other hand, NS5A inhibitors were shown to block HCV replication by preventing the formation of the membranous web, which was not correlated with the inhibition of PIKIII α (Berger et al., 2014). And, they also interestingly reported that the subcellular distribution of NS5A was not altered by DCV treatment, too.

1.4. The Aim of the Current Study

The NS5A research expanded dramatically in the last ten years and provided us invaluable information. However, there are many open questions in the field. It is not clear how NS5A-DI exist in solution environment. Crystal structures showed the dimeric forms of NS5A-DI, but no study so far clearly showed the existence of dimeric forms in solution. By combining a variety of biophysical characterization tools such as size-exclusion chromatography (SEC), dynamic light scattering (DLS), small-angle X-ray scattering (SAXS) in combination with crosslinking assays, we have carried out a systematic investigation of the in-solution behavior of NS5A. The knowledge we gained from the studies of free NS5A-DI allowed us to further investigate the interactions of NS5A with RNAs and the major DAA, DCV, *in vitro*. By doing so, we mainly aimed to reach an insight regarding the MoA of DAA in the inhibition of HCV through NS5A-DI. Moreover, considering the membrane-anchoring nature of NS5A-DI, we also furthered extended our studies by shifting from in-solution to surface-bound state using supported bilayer platforms (SBL). Our goal was to check whether NS5A behaves differently with DCV treatment on bilayers. Thus, the overall aim of this study is to widen our understanding towards the in solution behavior of NS5A and its interactions with DCV and RNA.

Chapter 2: MATERIALS and METHODS

2.1. Materials

2.1.1. Chemicals, enzymes, and kits

Antibodies against His-tag and UBI-tag were ordered from Santa Cruz (CA, USA). The chemicals, phenylmethylsulfonyl fluoride (PMSF), DL-dithiothreitol (DTT), 2-mercaptoethanol (BME), the primers, and lactose agarose were purchased from Sigma (Saint Louis, MO, USA). The DNA and protein ladders, bovine serum albumin (BSA) and Western Blotting Detection kit were purchased from Bio-Rad (Hercules, CA, USA). Protease inhibitor cocktail tablets (EDTA-free) were purchased from Roche (BASEL, Switzerland). Luria-Bertani broth was obtained from BD (Sparks, MD, USA). SYPRO orange dye was purchased from Life Technologies. T4 DNA Ligase, restriction enzymes, and their buffers were from New England Biolabs (NEB). *Pfu* and taq DNA Polymerase were purchased from Fermentas (Glen Burnie, MD, USA). Miniprep Plasmid Extraction kit was purchased from Axygen (Union City, CA, USA). QIAquick gel extraction kit was purchased from QIAGEN (Venlo, Limburg, Netherlands). Ni²⁺-NTA High-performance, His-Trap FF Crude, PD-10 Desalting, Superdex 75 Hi-Load 16/60, Superdex 200 Increase 10/300 GL and Superdex 75 5/150 GL columns were purchased from GE Healthcare (Uppsala, Sweden). Bio-Rad Protein assay (Bradford) was from Bio-Rad (Hercules, CA, USA) and BCA Protein Assay was from The Thermo Scientific Pierce (USA). The protein concentrators, Amicon concentrator (3 kDa MCO) were purchased Millipore (Co-cork, Ireland). 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (sodium salt) (POPG) and 1,2-Dioleoyl-sn-glycero-3-ethylphosphocholine (chloride salt) (DOEPC) were purchased either in powder form from Avanti Polar Lipids (Alabaster, AL, USA). Daclatasvir dihydrochloride was purchased from MedChemExpress (USA). RNA molecules (15U RNA) were ordered from Integrated DNA Technologies (IDT), (IOWA, USA).

2.1.2. Vectors and cell strains

pETSUMO (Kan)	LifeSensors (Malvern, PA, USA)
pETUBI (Kan)	Addgene (Cambridge, MA)
PCG1(Chloramphenicol)	A gift from Prof. Craig Cameron (of Penn State University)
pET29b (Kan)	Addgene (Cambridge, MA)
<i>Escherichia coli</i> (<i>E. coli</i>)	BL21 (DE3), Rosetta, DH5 α (Novagen), C41

2.1.3. Instrumentation

ÄKTA FPLC UPC-900 system was from GE Healthcare (Uppsala, Sweden). Quartz crystal microbalance with dissipation (QCM-D) measurements were performed using a Q-Sense E4 instrument equipped with flow modules (Q-Sense AB, Gothenburg, Sweden). Small angle X-scattering (SAXS) data were collected on a BRUKER NANOSTAR SAXS instrument equipped with a Metal-Jet X-ray source (Excillum, Germany) and VÅN-TEC 2000-detector system. DLS analysis was performed using a Zetasizer NanoZS (Malvern Instruments).

2.1.4. Buffers and solutions

2.1.4.1. DNA agarose gel electrophoresis running buffers and the loading dye

TAE (Tris-acetate-EDTA) was prepared as 50 $\times$. 242 g of Tris base, 57.1 mL glacial acetic acid, 100 mL of 0.5 M EDTA were mixed, and the solution was topped up to 1 L with distilled water. The pH of the buffer was adjusted to 8.0 by HCl. TE (Tris-EDTA) buffer preparation was done accordingly. 10 ml of 1 M Tris-HCl pH 8.0 and 2 ml of 0.5 M EDTA were mixed and topped up to 1L with distilled water.

DNA loading dye was prepared as 6 $\times$. 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 40% (w/v) sucrose were dissolved in distilled water. The loading dye was stored at -20 °C for the long term.

2.1.4.2. Stocks of the IPTG and antibiotics

IPTG stock was prepared as 1M, approximately 1000 $\times$. 7.1 g IPTG was dissolved in 30 ml of distilled water and sterile filtered. The aliquots of the IPTG stocks were kept at -20 °C. Stocks of kanamycin and chloramphenicol antibiotics were prepared as

1000×. 30 mg of kanamycin and chloramphenicol were dissolved in distilled water and ethanol, accordingly and sterile filtered. Aliquots were kept at -20 °C for long term storage.

2.1.4.3. The SDS-PAGE running buffer and the loading dye

Running buffer of SDS-PAGE electrophoresis was prepared as 5×. 15.1 g of Tris-base, 72 g of glycine and 5 g of SDS were dissolved in 500 mL of distilled water on a magnetic stirrer for 2 hours, once all the chemicals were dissolved, the solution was topped up to 1L with distilled water. The buffer was kept at room temperature.

SDS-PAGE loading dye was prepared as 2×. 100 mM Tris-HCl pH 6.8, 4% SDS, 0.2% Bromophenol blue, 10 mM BME and 20% (v/v) glycerol were mixed in distilled water, and the pH was checked again and adjusted to 6.8 with HCl. The mixture was stored at -20 °C.

2.1.4.4. Buffers and reagents used for western blot

Transfer buffer was prepared as 1×. 200 mL of 5X SDS-PAGE running buffer, 200 mL methanol, 600 mL of distilled water were mixed. It was kept at room temperature for long time and it was cooled to 4 °C before the usage.

10×TBS (Tris-buffered saline) was prepared by mixing 24 g of Tris and 88 g of NaCl were dissolved in distilled water and the pH was adjusted to 7.5 using HCl. Membrane washing buffer (TBST) was the 1× TBS containing 0.05% tween-20 as detergent. Blocking buffer was prepared by dissolving 2.5 g of non-fat dry milk in 50 mL of TBST.

2.2. Methods

2.2.1. Polymerase chain reaction (PCR)

The DNA taq polymerase or pfu polymerase were used in this study; pfu was used for the amplification of gene constructs and taq polymerase specifically used for colony PCR to check successful transformations. The PCR reaction conditions are shown in Table 2-1 for both polymerases. And, the primers used in this study are available in Table 2-2. The reactions were performed as the following cycle: 95 °C / 5 minutes, 95 °C / 30 second (denaturation), 60 °C / 30 seconds (primer annealing), 72 °C / 90 seconds (primer extension), 72 °C / 5 minutes. Cycle number was set as twenty five for each

PCR reaction.

Table 2-1. The components of PCR reaction mixtures.

Reagents	taq polymerase PCR (μL)	<i>pfu</i> polymerase PCR (μL)
10× PCR buffer	5	5
dNTPs (10 mM)	1	1
MgCl ₂ (25 mM)	3	-
Forward primer (10 μM)	1	1
Reverse primer (10 μM)	1	1
Template	1	1
DNA polymerase	0.5	0.5
Autoclaved water	37.5	40.5

Table 2-2. Primers used in this study for the vector constructs

Primer code	Name	Sequence
HYD 1618	1a HCV NS5A f29	5'-ATGGTCTCTAGGTCCACAACCTGCCTGGG-3' (BsaI site at 5')
HYD 1619	1a HCV NS5A r213	5'-ATATTCTCGAGCTACGCCTCTGCTGTTAT-3' (XhoI site and STOP codon at 5' end)
HYD 1661	1a HCV NS5A f29-GSGSG	5'-ATGGTCTCTAGGTGGCAGCGGCAGCGGCCTGCCTGGGATTCCC-3' (BsaI site and GSGSG at 5' end)
HYD 1680	1a HCV NS5A f36	5'-ATATTCCGCGG C TTTGTGTCCTGCCAG-3' (SacII site at 5' end) with shift special for pETUBI
HYD 1685	1a HCV NS5A f36	5'-ATATTCATATGTTTGTGTCCTGCCAG-3' (NdeI site at 5' end)
HYD 1686	1a HCV NS5A r198	5'-ATATTCTCGAGCACGGCTACGTCCGG -3' (XhoI site at 5' end)
HYD 1817	1a HCV NS5A f 2	5'-ATATTCATATGGGTTTCCTGGCTAAGGGAC-3' NdeI
HYD 1676	1a HCV NS5A f36	5'-ATGGTCTCTAGGTTTTGTGTCCTGCCAG-3' (BsaI site at 5')
HYD 1786	1aHCV NS5A r198	5'-ATATTCTCGAGCTATTTATCATCATCATCCA CGGCTACGTCCGG-3' (XhoI site+stopcodon+DDDDK)
HYD 1912	1a HCV NS5A fY93H	5' - CCC ATT AAC GCC CAC ACC ACG GGC CCC -3'
HYD 1913	1a HCV NS5A r y93h	5'- GGG GCC CGT GGT GTG GGC GTT AAT GGG- 3'

2.2.2. Plasmid and DNA product purification

The bacteria were grown in up to 10 mL LB with antibiotics overnight and then harvested by centrifugation at 10.000 g for two minutes, see Table 2-3 for the antibiotic selection markers of the plasmids. Plasmids were purified from the bacterial culture by Axygen plasmid miniprep kit following the manual manufacturer.

Table 2-3. Plasmids and their properties

Vector	Resistance	Expression host	Remarks
pETSUMO	Kanamycin	BL21 or Rosetta	Protein expression with SUMO-tag
pETUBI	Kanamycin	BL21(DE3)	Protein expression with UBI-tag
PCG1	Chloramphenicol	BL21(DE3)	Cleavage of UBI-tag <i>in vivo</i>
pET29	Kanamycin	BL21(DE3)	C-terminal His-tag

2.2.3. Restriction of plasmids and inserts

All enzymes were ordered from New England Biolabs. Purified plasmid or DNA product were digested with restriction enzymes. In total 50 μL of reactions were prepared. The reactions were optimized based on the manufacture's suggestion; typical reaction components are shown in Table 2-4. The generally restriction reactions were incubated for 4 hours at 37 °C.

Table 2-4. A typical restriction digestion protocol.

Reagent	Plasmid digestion
H ₂ O	17.5 μL
Buffer	5 μL
BSA	0.5 μL
Plasmid/DNA product	25 μL (1 μg)
Restriction Enzyme 1	1-2 μL (~10 units)
Restriction Enzyme 2	1-2 μL (~10 units)
Incubation temperature	Enzyme dependent

2.2.4. DNA agarose gel electrophoresis

The percentage of agarose gel was calculated based on Table 2-5 below. The agarose was dissolved in TAE buffer by heating; then it was left for cooling. Gel red dye (1000 $\times$) was added to the solution and mixed homogeneously, used to visualize DNA bands. After solidification of the gel, an estimated of 500 ng DNA samples were added to the wells, and the samples were electrophoresed at 90 V for 30-50 minutes. UV light was utilized to visualize the bands. DNA markers (1 kb or 100 bp DNA marker) bands were

also run simultaneously to check the size of the DNA molecules

Table 2-5. Agarose gel electrophoresis percentage calculation.

Optimum Agarose (%)	Size of the Linear DNA (bp)
0.7	800–12.000
1.0	500–10.000
1.2	400–7.000
1.5	200–3.000
2.0	50–2.000

2.2.5. Ligation

PCR products or plasmids were purified by using PCR clean-up kits (The QIAquick PCR Purification Kit) to remove buffers/enzymes left from restriction reactions. 10 μ L of reaction mix, containing 1 μ L of ligation buffer, 3 μ L of digested plasmid, 5 μ L of digested PCR product and 1 μ L of *T4 ligase* enzyme was incubated overnight at 16 °C.

2.2.6. Site-directed mutagenesis

We performed site-directed mutagenesis for single amino acid substitutions. We designed the primers in a way that it introduces the mutation (changed base) in the middle of the primer. A plasmid containing wild-type gene of the target was utilized as a template for PCR reactions. After PCR reactions, the product was purified using PCR clean up kit. Then, it was treated with DpnI enzyme for 2 hours at 37 °C. The new mutated plasmid product was transformed into XL-blue competent cells. DNA sequencing confirmed the mutations.

2.2.7. Competent cell preparation by CaCl₂ treatment method

We employed CaCl₂ chemical treatment method to make competent cells of *E. coli*. A single colony was picked up from the LB agar plate and grown in 5 mL culture tube for overnight. Next day it was transferred to 100 mL of fresh LB, and it was grown till the OD of the culture reaches 0.5-0.6. The cells were collected by centrifugation and resuspended in ice cold and 100 mM CaCl₂ treatment was applied approximately half an hour to 1 hour. Importantly, all the steps were conducted on ice. As cells were very fragile, all the steps were performed very gently after the chemical treatment step. The

cells were harvested and resuspended in 8 mL of 100 mM CaCl₂ and 2 ml of glycerol. The suspension was then aliquoted in 50 µL of volumes and stored at –80 °C.

2.2.8. Transformation of plasmid DNA into *E. coli*

50 µL of the DH5α competent cell was placed on ice and thawed. 2 µL of plasmid DNA or 10 µL of ligation product was added to the tube and incubated on ice for 30 minutes. Later the tube was heated for 45 seconds at 42 °C and immediately put in the ice box for 3 to 5 minutes to create a thermal shock. 250 µL of LB with no antibiotic was added to the tube and incubated at 37 °C for 45-60 minutes. Lastly, the culture was placed on LB agar plate with antibiotic and incubated overnight at 37 °C.

2.2.9. Colony PCR

Colony PCR screening was performed to know whether plasmid carrying the gene of interest was transformed. Each colony, generally ten colonies, were inoculated into 2 mL of LB with antibiotic and grown at 37 °C for nearly 4 hours. 2 µL of the culture was used as a template for PCR with Taq polymerase. The positive colonies that had amplified PCR products were identified. The plasmid isolation was done from positive colonies, and they were sent for DNA sequencing. Once sequence was confirmed, the plasmids were transformed into competent cell lines of *E. coli* for expression, BL21 (DE3) or Rosetta.

2.2.10. Media preparation

We used two different media to grow in *E. coli*. Luria-Bertani (LB) media was used for expression of recombinant proteins. NaCl (5 g), yeast extract (5 g) and back to tryptone (10 g) were mixed in 950 mL of distilled water and autoclaved at 121 °C for 15 minutes. For the preparation of LB-agar plate, 15 g/L of agar was added to the medium before autoclaving. Different antibiotics were added to the mediums as required before bacterial culture.

2.2.11. Expression of recombinant protein

A single colony was inoculated and grown overnight in 50 mL of LB medium supplemented with antibiotic/s based on different plasmids, see Table 2-3 for antibiotic selection markers. 10 mL of overnight culture was used to inoculate 1 liter of LB

medium. The 1-liter culture was grown at 37 °C to an OD₆₀₀ of ~0.6-0.7. The culture was cooled at 4 °C for half an hour before IPTG induction. 1 mM of IPTG added to induce protein expression and then 1-liter culture were grown at 25 °C for an additional 6 hours. The cells were harvested by centrifugation in a Beckman JLA-10.500 rotor at 8.000 rpm for 20 minutes.

2.2.12. Cell disruption by ultrasonication

The cell pellet from 1L of culture was suspended in 35 mL lysis buffer. Phenylmethylsulfonyl fluoride (PMSF) was included as a final concentration of 1 mM. The resuspended cells were lysed by sonication ~30 minutes (3 seconds on, 1 second off). The cells disrupted by high-frequency sound and 50 µl of lysed sample was separated for SDS-PAGE analysis.

2.2.13. Purification of soluble proteins by Ni⁺²-NTA affinity column

The extract obtained from sonication step was centrifuged in a Beckman JA-20.50 rotor for 30 minutes at 18,000 rpm at 4 °C. 50 µL of supernatant (clarified lysate) was kept for SDS-PAGE analysis. 5 mL of beads (2.5 mL of exact beads) were topped up 30 mL with water and washed six times with water. Then beads were equilibrated with lysis buffer ~10 to 15 minutes at room temperature on a shaker. 35 mL of clarified cell lysate was added to beads and incubated at 4 °C on shaker 1 to 2 hours. Clarified lysate with beads was loaded on the column. The column was washed with washing buffer (20 mM potassium phosphate pH 8.0, 300 mM NaCl, 25 mM imidazole and later 50 mM imidazole). The protein was eluted with elution buffer (20 mM potassium phosphate pH 8.0, 250 mM NaCl, 10% glycerol, 500 mM imidazole). The collected fractions were concentrated to 2.5 mL, and its buffer was exchanged to buffer B (25 mM HEPES of pH 7.4, 150 mM NaCl, 10% glycerol) by using PD 10 column.

2.2.14. SUMO-tag cleavage and second round purification by Ni⁺²-NTA affinity column

The recombinant proteins with the SUMO-tag were incubated with 40 µL of SUMO protease (Invitrogen) for overnight at 4 °C. Next day, the mixture that had tag-free recombinant protein, SUMO-tag and SUMO protease was loaded to Ni⁺²-NTA column. SUMO-tag and SUMO protease had His-tag, so they were bound to the column. Tag-free recombinant protein collected in the flow-through. The purified proteins were run

on a 12% SDS-PAGE to analyze the purity of protein.

2.2.15. Regeneration of Ni²⁺-NTA beads

Ni²⁺-NTA agarose beads were either washed with 0.5 M NaOH or regenerated after each purification step. The following steps are usually applied for regeneration of Ni²⁺-NTA agarose beads. The column is washed with 2 volumes of 6 M guanidine hydrochloride in 0.2 M acetic acid solution.; then it is washed with 2% SDS one volume. The ethanol washing steps are performed, one volume of 25, 50, 75, and 100% of ethanol. Followed by five-volume of distilled water and 100 mM EDTA (pH 8.0). Beads are regenerated with two volumes of 100 mM NiSO₄. Finally, they are washed with one five-volume of distilled water and three volumes of 20% ethanol. The beads are stored in 20% ethanol at 4 °C.

2.2.16. Purification of membrane proteins

Cell disruption was performed in cell resuspension buffer (50 mM potassium phosphate, pH 7.0, 300 mM NaCl, 2 mM MgCl₂, supplemented with 1 mM PMSF), 50 mL of buffer per 1 L of culture was used for sonication. The cell lysate was centrifuged at 40.000 g for 60 minutes, the supernatant containing soluble proteins are discarded. Next, the pellet the pellet was resuspended in the membrane resuspension buffer (50 mM potassium phosphate, pH 7.0, 300 mM NaCl, 2 mM MgCl₂, 5 glycerol%, %2 *Triton X-100* supplemented with 1 mM PMSF and Roche EDTA-free protease inhibitor). The cell pellet was mixed with a glass homogenizer on the ice, and the bacterial membrane proteins were solubilized at 4 °C for overnight. Next day the solubilized membranes were centrifuged at 45.000 g for 1 hour, the supernatant containing the solubilized membrane proteins were loaded on the His-Trap FF column for gradient imidazole protein purification. Buffer A (25 mM potassium phosphate, pH 7.0, 300 mM NaCl, 2 mM MgCl₂, 0.003% *Triton X-100*, 25 mM imidazole) and buffer B (25 mM potassium phosphate, pH 7.0, 300 mM NaCl, 2 mM MgCl₂, 0.003% *Triton X-100*, 500 mM imidazole). Proteins eluted after 330 mM imidazole concentrations were collected.

Figure 2-1. Workflow/strategies for the purification of membrane proteins

SDS-PAGE was used to check purification steps of protein based on the separation of proteins according to their weight. The protein sample was mixed with 6× SDS-loading buffer and heated at 95 °C for 5 minutes. Once samples were loaded into SDS-PAGE gel, the gel was run at 90 volts for 110 minutes.

2.2.18. Western blot analysis

Western blotting was performed to detect proteins using an antibody. In this study, the validation of protein purification was done by using His-tag specific antibody as proteins designed with N/C-terminal His-tags. (See Figure 3-1 in the next section for the constructs designed in this study.) SDS-PAGE was run according to the protocol above with dual-color protein marker. Later the PDVF membrane, SDS gel, filter papers were soaked into transfer buffer. The cassette was assembled according to manufacturer's manual. The sample was run at 100 volts for 1 hour, and so the transfer of proteins from SDS gel to PDVF membrane was performed. After the transfer of proteins to the membrane, the membrane was placed into blocking solution, 5% skim milk for 1 hour at room temperature. Then the membrane was incubated with primary antibody

overnight at 4 °C. Next day membrane was washed three times for 15 minutes with 1× TBST. The membrane was treated with secondary antibody for 1 hour at room temperature, and the later membrane was washed three times with 1× TBST again. Finally, enzyme kit (Bio-Rad) mixture was added to the membrane, and it was incubated for three minutes and detection was performed on a gel imaging system.

2.2.19. Fast-Performance Liquid Chromatography (FPLC)

The proteins that were purified by Ni⁺²-NTA affinity column were checked on SDS-PAGE gel. In order to increase the purity of the recombinant proteins, the sample was purified with FPLC based on size-exclusion gel filtration. In this study, three columns were used. For Superdex75 HiLoad 16/60, 0.5 mL/minute flow rate was used, and the sample fractions were concentrated to 2 mL and injected to Superdex75 column. Superdex 200 increase 10/300 GL, 0.3 mL/minute flow rate was set, and 0.5 mL of sample was injected. Superdex 75 5/150 GL column is used for analytical purposes. 0.2 mL/minute flow rate was used, and 150 µL of sample was injected into the column. SEC buffer for NS5A-DI lacking AH was 25 mM HEPES of pH 7.4, 150 mM NaCl, 10% glycerol, 0.01 NaN₃. SEC buffer for NS5A-DI with AH was 25 mM HEPES, pH 6.3, 300 mM NaCl, 2 mM MgCl₂, 0.002% Triton X-100, 0.01 NaN₃.

2.2.20. Sample processing for ¹H NMR experiment (1D NMR)

After the purification of the protein as described above, it was concentrated to at least 0.1 mM in 20 mM potassium phosphate, 25 mM potassium phosphate, 150 mM NaCl, 5% glycerol, pH 7.5, 10% D₂O solution and 1D NMR analysis was done. (The NMR instrument used is given above.)

2.2.21. ¹⁵N labeling of recombinant proteins

NS5A(36-198) protein, labeled with ¹⁵N, were purified for 2D NMR experiments. Rosetta cells (DE3) cells were inoculated into 10 mL M9 media containing ¹⁵N labeled ammonium chloride. M9 media components are shown below. For ¹⁵N labeling the following recipe was used. 52.7 mM Na₂HPO₄, 26.5 mM KH₂PO₄, 10 mM NaCl, 0.1 mM CaCl₂, 1.2 mM MgSO₄, 4.5 g/L Glucose, 5 ng/L Thiamine, and 1 g/L ¹⁵NH₄Cl were prepared. All the chemicals were filtered/autoclaved as well. Kanamycin and 0.2 mM ZnCl₂ were added into the stated M9 media too.

2.2.22. Sample processing for ^{15}N -labeled NMR experiment (2D NMR)

^{15}N -labeled ZnCl_2 added M9 medium was used to grow cultures. Protein purification was performed as described in Section 2.2.13. Then, protein was concentrated to 0.1 mM in the same buffer.

2.2.23. Protein concentration determination

We used three different methods for the protein concentration determination. Bio-Rad Protein assay (Bradford) and Nanodrop were employed for the water soluble and detergent-free proteins. We used the bicinchoninic acid assay (BCA assay) for the detergent-solubilized proteins. We prepared a standard curve with a known protein (BSA) to obtain a standard plot. Different concentrations of the protein (BSA) was mixed with 1 mL of Bradford/BCA solution and absorbance recorded at 595/562 nm, respectively.

2.2.24. Chemical crosslinking

NS5A(36-198) were then crosslinked at cysteine residues by incubation with 1 mM 1,6-Bismaleimido-hexane (BMH, Pierce) for 1 hour at room temperature. The crosslinking reaction was quenched by 4 mM DTT. Later, the crosslinked samples were analyzed using 10% SDS-PAGE.

2.2.25. Small-angle X-ray Scattering (SAXS)

SAXS-data for the NS5A domain 1 of HCV from genotypes 1b and 1a were collected with the BRUKER NANOSTAR SAXS instrument equipped with a Metal-Jet X-ray source (Excillum, Germany) and VÅNTEC 2000-detector system as described in (Balakrishna, Basak, Manimekalai, & Gruber, 2015; Dip et al., 2014; Saw et al., 2015). The scattering patterns were measured using a sample to detector distance of 0.67 m and a wavelength of $\lambda=1.3414 \text{ \AA}$, which covers a range of momentum transfer of $0.016 < q < 0.4 \text{ \AA}^{-1}$ ($q = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle). SAXS experiments were carried out at 15 °C in two different concentrations in the buffer in 25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol with a sample volume of 40 μl in a vacuum tight quartz capillary. To monitor for radiation damage, ten 1 min exposures were collected for each protein sample and no radiation effect was observed. The data were normalized to the intensity of the transmitted beam, and the scattering of the buffer was subtracted

and the difference curves were scaled for concentration.

All the data processing steps were performed using the program package PRIMUS (Petr V. Konarev, Petoukhov, Volkov, & Svergun, 2006) from ATSAS package version 2.7.1. The experimental data obtained for all protein samples were analyzed for aggregation using the Guinier region and the forward scattering $I(0)$ and the radius of gyration, R_g , were computed using the Guinier approximation. These parameters were also computed from the extended scattering patterns using the indirect transform package GNOM (D. I. Svergun, 1992), which provides the distance distribution function $P(r)$ hence the maximum particle dimension, D_{max} , as well as the radius of gyration, R_g . Qualitative particle motion was inferred by plotting the scattering patterns in the normalized Kratky plot $((qR_g)^2(I(q)/I(0))$ vs qR_g) (Durand et al., 2010). *Ab initio* low-resolution models of the proteins were built by the program DAMMIF (Franke & Svergun, 2009) considering low angle data ($q < 2 \text{ nm}^{-1}$). Twenty independent *ab initio* reconstructions were performed for each protein and then averaged using DAMAVER (Volkov & Svergun, 2003). The averaged and filtered *ab initio* model was superimposed with the atomic model using SUPCOMB (Kozin & Svergun, 2001). The theoretical scattering curves from atomic structures were generated and evaluated against experimental scattering curves using CRY SOL (D. Svergun, Barberato, & Koch, 1995). Utilizing the intensities from each component, the volume fraction of each component can be determined by OLIGOMER (P.V. Konarev, Volkov, Sokolova, Kochb, & Svergun, 2003). The oligomeric state of the protein was confirmed from the molecular mass calculation based on $I(0)$, Porod volume (V_p), excluded volume (V_{ex}) and volume of correlation (V_c) (Mylonas & Svergun, 2007; Rambo & Tainer, 2013)

2.2.26. Thermal denaturation assay

Protein samples were mixed with SYPRO orange dye and buffer to a final volume of 25 μL per well into 96- well QPCR plate. A final 10 μM concentration of NS5A(36-198)-DDDDK protein mixed with 10X SYPRO orange dye and the plates were centrifuged at 1000 g for 1 minute to mix each sample before assay. The plates were heated incrementally by 0.4 $^{\circ}\text{C}$ steps from 10 to 100 $^{\circ}\text{C}$ approximately in 2 hours. The SYPRO orange dye fluorescence readings were obtained by an Applied Biosystems real-time PCR machine. The raw data was analyzed by using Protein Thermal Shift Software Version 1.2. Protein Thermal Shift Software calculates the melting

temperature (T_m) from each fluorescence profile by using Boltzmann method. Buffers and agents used in thermal stability screening are presented on the Table 2-6.

Table 2-6. Buffers screened for thermal denaturation assay. Final pH, buffering systems, salt types and concentrations, and reducing agent types and concentrations are varied. Note that all these experiments are grouped by color coding in the table, and, within each experimental group, varying conditions are underlined. (Conditions indicated by asterisk (*) are repeated. For example, Buffers #11 and #22 are repeated while studying the effect of buffer system only using the conditions given by #17-to-20 in the table.)

Buffer No	pH	Buffering system (50 mM)	NaCl	ZnCl ₂	DTT	BME
			(mM)			
1	<u>2.7</u>	Potassium phosphate	100	0	0	0
2	<u>3.6</u>	Potassium phosphate	100	0	0	0
3	<u>3.9</u>	Potassium phosphate	100	0	0	0
4	<u>5.0</u>	Potassium phosphate	100	0	0	0
5	<u>5.2</u>	Potassium phosphate	100	0	0	0
6	<u>5.4</u>	Potassium phosphate	100	0	0	0
7	<u>5.7</u>	Potassium phosphate	100	0	0	0
8	<u>6.1</u>	Potassium phosphate	100	0	0	0
9	<u>6.6</u>	Potassium phosphate	100	0	0	0
10	<u>7.0</u>	Potassium phosphate	100	0	0	0
11*	<u>7.4</u>	Potassium phosphate	100	0	0	0
12	<u>7.8</u>	Potassium phosphate	100	0	0	0
13	<u>8.2</u>	Potassium phosphate	100	0	0	0
14	<u>8.4</u>	Potassium phosphate	100	0	0	0
15	<u>9.0</u>	Potassium phosphate	100	0	0	0
16	<u>9.4</u>	Potassium phosphate	100	0	0	0
17	7.4	<u>Tris-HCl</u>	100	0	0	0
18	7.4	<u>Sodium phosphate</u>	100	0	0	0
19	7.4	<u>BICES</u>	100	0	0	0
20	7.4	<u>MOPS</u>	100	0	0	0
21	7.4	HEPES	<u>0</u>	0	0	0
22*	7.4	HEPES	<u>100</u>	0	0	0
23	7.4	HEPES	<u>200</u>	0	0	0
24	7.4	HEPES	<u>400</u>	0	0	0
25	7.4	HEPES	<u>600</u>	0	0	0
26	7.4	HEPES	<u>800</u>	0	0	0
27*	7.4	HEPES	100	<u>0</u>	0	0
28	7.4	HEPES	100	<u>2</u>	0	0
29	7.4	HEPES	100	<u>5</u>	0	0
30	7.4	HEPES	100	<u>10</u>	0	0
31	7.4	HEPES	100	<u>15</u>	0	0
32	7.4	HEPES	100	0	<u>1</u>	0
33	7.4	HEPES	100	0	<u>2</u>	0
34	7.4	HEPES	100	0	<u>3</u>	0
35	7.4	HEPES	100	0	<u>4</u>	0
36	7.4	HEPES	100	0	0	<u>1</u>
37	7.4	HEPES	100	0	0	<u>2</u>
38	7.4	HEPES	100	0	0	<u>3</u>
39	7.4	HEPES	100	0	0	<u>4</u>

2.2.27. Dynamic light scattering (DLS)

The average size and polydispersity of the proteins were analyzed using a Zetasizer Nano ZS instrument (Malvern Instruments, UK). Each measurement was performed modifying the protocol given in the literature (Karahan, Karakus, & Birer, 2015). All measurements were performed at 25 °C. Approximately 50 µL of protein samples were collected in a low-volume quartz batch cuvette (ZEN2112). Equilibration time was adjusted as 120 seconds in all measurements. At least twenty measurements with 15 seconds accumulations were recorded. Intensity-averaged hydrodynamic diameter results of minimum fifteen data points were combined using the software of the instrument, DTS Version 6.7. And, the results are presented as diameter distributions.

2.2.28. Circular dichroism (CD)

CD spectra of the proteins were collected in the far UV light (180-260 nm) using a CHIRASCAN spectropolarimeter (Applied Photophysics). Data were collected in a 60 µL quartz cell (Hellma) at 20 °C at a step resolution of 1 nm. CD spectrum of proteins was recorded in SEC buffers, and three measurements were collected for each sample. The CD spectrum of the buffer alone was subtracted from the spectra of the protein using the software of the instrument. The baseline corrected and averaged spectra was analyzed in BestSel software (Micsonai et al., 2015). This software is specifically designed for the secondary structure analysis of proteins containing beta-sheets. As NS5A is a protein rich in beta-sheets, we have employed BestSel for data analysis. Importantly, BestSel is also able to classify beta-sheet structures like anti-parallel and parallel beta-sheets.

2.2.29. Quartz crystal microbalance with dissipation (QCM-D) monitoring

A Q-sense E4 instrument (Q-sense Ab, Gothenburg, Sweden) was employed to characterize the interaction between model lipid bilayers and proteins by monitoring the changes in resonance frequency and energy dissipation of acoustic transducers, *i.e.* piezoelectric crystal as a function of time. Piezoelectric materials accumulate electric charge in response to applied mechanical stress, in our system silicon oxide-coated quartz (model no. QSX303) was used as an acoustic transducer. The silicon-coated sensor substrates were washed with %2 SDS, water, and %70 ethanol sequentially and dried by nitrogen gas. Oxygen plasma treatment (Harrick Plasma, Ithaca, NY) was

applied to the sensors at the maximum radio frequency for further cleaning of the surfaces. The QCM-D measurements' parameters were set as follows. Continuous flow rate of 100 μ L/minute was used for all steps in the bilayer formation and 41.8 μ L for the protein injections. Temperature was fixed to 24 °C. And, the third, fifth, seventh, ninth, and eleventh overtones were collected during the measurements. The Sauerbrey equation (Equation 2-1) was utilized to calculate the deposited mass of the biomolecules.

2.2.30. Supported lipid bilayer (SLB) formation

Supported lipid bilayers were constructed by using solvent-assisted lipid bilayer (SALB) method. Powder forms of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and 1,2-Dioleoyl-*sn*-glycero-3-ethylphosphocholine (chloride salt) (DOEPC) dissolved in isopropanol at the final concentration of 10 mg/mL. The 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (POPG) lipids were dissolved in ethanol to 5 mg/mL in water bath 42 °C to improve solubility. The targeted lipid compositions were mixed and diluted to the final experimental lipid concentration of 0.5 mg/mL before experiments. SALB method comprises four steps: 1) baseline measurement with physiological buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl), 2) buffer exchange with isopropanol, 3) injection of lipid mixtures dissolved in isopropanol (0.5 mg/mL) and incubation for 10 minutes, 4) driving bilayer formation by buffer exchange with 10 mM Tris-HCl pH 7.5, 150 mM NaCl.

2.3. Principles of SAXS and QCM-D

2.3.1. Structural biology techniques and small angle X-ray scattering

Biological macromolecules regulate the cell through the reactions they catalyze, conformations they adopt and interactions that they are involved in. The knowledge of the structures, their three-dimensional atomic configurations, is critical to elucidate the mechanistic understanding of these interactions. Structural biology aims to achieve high resolution models of macromolecules/virions that range from 1 to 100 nm in diameter. The wavelength of the visible light is around 500 nm, and for that reason, biological macromolecules are far too small for the visualization by light microscopy. Since we require atomic resolution structures, we utilize radiation with a wavelength as small as dimensions of the atoms (~0.2 nm) (Campbell, 2002).

Methods in structural biology involve several techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), small-angle X-ray scattering (SAXS) and cryo-electron microscopy (cryo-EM). Each method has its advantages and disadvantages which are summarized in Table 2-7. The main challenges of structural studies are conformational changes, protein dynamics and polydispersity. Moreover, it is reported that most of the significant problems would involve proteins/complexes whose structures would not be easily revealed by any single biophysical technique (Putnam, Hammel, Hura, & Tainer, 2007). It is suggested that the combination of several techniques, like X-ray crystallography and SAXS, will be valuable in the future for such challenging studies.

SAXS is a biophysical method to study the structural features of ordered and disordered macromolecules in solution. It gives information about the molecular weight and shape of proteins and their complexes under a variety of conditions (Skou, Gillilan, & Ando, 2014). This part of Chapter 2 provides a brief overview of the SAXS method and how to interpret SAXS data.

Table 2-7. Comparison of structural biology techniques

	Advantages	Disadvantages	Sample requirements
X-ray Crystallography	Various molecular weights High resolution Easy for modelling and established protocol	Difficult for crystallization Time consuming Inability to study behavior of the molecules in solution Unnatural and non-physiological environment	Crystallizable samples Water/ detergent Solublized proteins, membrane proteins, ribosomes, DNA/RNA and protein complexes
NMR	High resolution 3D structure in solution Give information regarding dynamic properties	Sample preparation is difficult and costly. Requires high purity at high concentrations Size limitation.	Less than 25kDa is preferred Water soluble samples Stability at higher concentrations for long time (~ 3 days)
SAXS	Structure in native state Easy sample preparation Low concentrations of sample can be studied. No size limitation	Low-resolution (10-50 Å)	
Cryo-EM	Easy sample preparation Structure in native state	Relatively low resolution (compared to NMR and X-ray). Applicable to samples of high molecular weights only Highly dependent on EM techniques. Costly EM equipment.	Bigger than 150 kDa, such as virions, large proteins, and complex macromolecules

2.3.1.1. Experimental set-up of SAXS

SAXS and X-ray crystallography are basically similar methods in that they both detect interactions between X-ray with matter. In most of the experiment set-ups, a collimated monochromatic beam of X-rays irradiates a sample and the X-ray detectors measure the intensities of the diffracted and scattered X-ray for crystallography and SAXS respectively. The major difference between solution scattering and X-ray crystallography is the relative organization of the target molecules. In X-ray crystallography, the molecules are highly ordered within a crystal lattice. In contrast, in solution scattering, the intensities from all orientations of the target molecules, relative to one another, is collected and averaged. Thus, the solution scattering data is continuous and radially symmetric as shown in Figure 2-2b (Putnam et al., 2007).

Figure 2-2. Schematic representation of a SAXS experiment. (a) The object is placed in front of a highly collimated X-ray beam and the scattered X-rays are measured. 2θ is the angle of any scattered position with the direct beam. (b) Scattered X-ray from a solution of yeast PCNA with a maximum resolution of 23.9 Å, the figure is taken from (Putnam et al., 2007).

2.3.1.2. Data collection and interpretation

This section gives brief information regarding the measurement and processing of SAXS data that was summarized from the following resources (Kikhney & Svergun, 2015; Putnam et al., 2007). SAXS measures the scattering signal derived from the difference in the average of electron density, ($\Delta p(r)$), of the sample solution ($p(r)$) and solvent (p_s). Therefore, the resulting signal ($\Delta p(r)$) is directly correlated with the overall shape and the size of the particles and highly contrast dependent between the solution and the particles (Mertens & Svergun, 2010)

$$\Delta p(r) = p(r) - p_s$$

The orientations of the dispersed particles in solution are random and, which makes scattering pattern isotropic. Therefore, the scattering pattern of the SAXS experiments collected by two-dimensional X-ray detector can be radially averaged. Scattering intensity (I) is shown as a function of the momentum transfer of $q = (4\pi \sin(\theta))/\lambda$, where 2θ is the scattering angle and λ is the wavelength of the X-ray beam.

$$\langle I(q) \rangle = \langle I(q) \rangle_{\Omega} = \langle A(q) A^*(q) \rangle_{\Omega}$$

$A(q)$ and $\langle \rangle_{\Omega}$ refer to scattering amplitude and the spherical average respectively where $A(q)$ is a Fourier transform of the excess electron density.

The radially averaged one-dimensional curves of the isotropic scattering patterns are plotted against q as shown in Figure 2-3a, referred as SAXS pattern or profile. The experimental scattering pattern of a sample provides a significant amount of information regarding to the particle. like Guinier plot (protein aggregation analysis), radius of gyration (R_g), molecular weight, particle volume and the maximum dimension (D_{max})

Firstly, one of the analyses directly extracted from SAXS pattern is the Guinier plot (the scattering intensity $\ln(I)$ vs. q^2). This plot should be a linear function for a particle of any shape based on Guinier approximation (Kikhney & Svergun, 2015), and the protein exposing a linear Guinier plot means that there is no aggregation in the system and that data quality is good. Similarly, non-linear behavior indicates aggregation or inappropriate q -range. Another critical analysis of SAXS profile of a protein is the distance distribution function ($P(r)$). A particle can be conveniently described by distance distribution function which is a histogram of distances between all possible

pairs of atoms within a particle and it can be calculated from the intensity using a Fourier transform [6] as shown below.

$$P(r) = \frac{r}{2\pi^2} \int_0^\infty I(q) q \sin(qr) dq$$

Figure 2-3b shows the plot of $P(r)$ vs. r for the following samples (dark blue: globular, light blue: partially unfolded, gray: natively unfolded). $P(r)$ analysis serves as a very critical parameter for the shape determination. Moreover, the maximum dimension of the particle (D_{max}) can be calculated where D_{max} is the value of r at $P(r)=0$ for larger r . High-frequency oscillations or discontinuities in $P(r)$ may refer to problematic Fourier transform processes.

Next, R_g is one of the most important parameter directly extracted from the small-angle scattering pattern of a sample and provides an estimated measurement for the overall size of the macromolecule. R_g mathematically means the root mean square distance of the object's parts from its center of mass. R_g value can be calculated by using several methods (by using Guinier plot or $P(r)$).

R_g from Guinier approximation is shown below

$$\ln[I(q)] = \ln[I(0)] - \frac{q^2 R_g^2}{3}$$

and R_g value defined by $P(r)$ as follows

$$R_g^2 = \frac{\int_0^{D_{max}} r^2 P(r) dr}{\int_0^{D_{max}} P(r) dr}$$

Lastly, Kratky plot ($q^2 I(q)$ vs q) provides the information about the potential flexible behavior and/or extended shape of the examined particle (Durand et al., 2010). For a well-folded globular protein, the plot will show a defined bell curve profile as seen in Figure 2-3c. Extended molecules, unfolded peptides don't show a peak but they have a plateau in the larger q range.

Figure 2-3. Data analysis on a simulated SAXS data, the figure was adopted from (Kikhney & Svergun, 2015). (a) Logarithmic plot of the scattering intensity $I(q)$ vs q . (b) Distance distribution function $p(r)$. (c) Kratky plot $q^2 I(q)$ vs. q . (d) Normalized Kratky plot. Three proteins (60 kDa) are represented, dark blue: globular, light blue: 50% unfolded, gray: disordered protein

2.3.1.3. Combination of crystallography and solution structure modelling

After extracting several parameters from measured scattering intensity profile (as explained above), solution structure modelling is performed. The overall schematic of the solution structure modelling steps is illustrated in Figure 2-4. This part will briefly describe methodology of structure determination from the following the sources (Kikhney & Svergun, 2015; Putnam et al., 2007; Skou et al., 2014). If the particle in solution indicates a monodisperse sample, the measured scattering profile can be directly used for the particle shape determination. Ab-initio structure modelling can be utilized for low-resolution structures by using the programs like DAMMIN, SAXS3D, GA_STRUCT or GASBOR. Generation of multiple models and comparison of several reconstructions are very critical to improve model quality and preventing erroneous shapes. The DAMMAVER program compares, averages and checks the uniformity of the generated structures. Then, if the atomic structure of the macromolecule is known, the theoretical SAXS profile of the known structure should be compared with the

experimental data to evaluate structure modelling. CRY SOL program calculates the theoretical SAXS profile for known atomic structures. A theoretical scattering curve computed using CRY SOL having a fit to the experimental scattering curve of $\chi^2 < 3.0$ agreement is common. Alternatively, the *ab-initio* model and the atomic resolution models can be aligned by utilizing programs like SUPCOMB and SITUS as a model validation step.

Figure 2-4. Schematic for SAXS data solution structure modelling, the workflow is taken from (Putnam et al., 2007).

If the atomic resolution models and the SAXS result are not in agreement, there might be several reasons for that. One of them is that the sample might exist in an heterogeneous mixture (multimerization). In that case, OLIGOMER program can be utilized which calculates the scattering profiles of several states of the macromolecules and gives a fit score for the experimental data. Another reason for bad fit score between atomic models and SAXS data might arise due to flexible and truncated regions. To overcome this

problem, a variety of several states for the flexible loops are generated to improve the overall fit.

2.3.2. Quartz crystal microbalance with dissipation monitoring

Quartz crystal microbalance with dissipation (QCM-D) is a type of quartz crystal microbalance with nanogram sensitivity. QCM-D monitors changes in the resonance frequency (Δf) and energy dissipations (Δd) of the surface, which are correlated to the mass and viscoelastic properties of materials/macromolecules adsorbed to the surface (Reviakine, Rossetti, Morozov, & Textor, 2005)

Quartz crystal microbalance (QCM) utilizes the inherent property of the quartz crystals, piezoelectricity, that is the electric charge that accumulates in quartz in response to applied mechanical stress (Bunde, Jarvi, & Rosentreter, 1998). Therefore, the major component of QCM is a thin quartz disk sandwiched between a pair of electrodes, as shown Figure 2-5a (Johannsmann, 2008). A quartz material with alternating current (AC) applied to the electrodes causes the crystal to oscillate at its acoustic resonance frequency (Figure 2-5b). The frequency of the oscillating crystal changes as the mass (mechanical stress) on the sensor changes. Therefore, the changes in the resonance frequency, Δf gives us information regarding the mass of the materials/macromolecules adsorbed to the sensors. The Sauerbrey equation is utilized to estimate the deposited mass of the biomolecules and correlates with the difference in frequency (Δf) as follow:

Equation 2-1. Sauerbrey equation.

$$\Delta m = -C \frac{\Delta f}{n}$$

where Δm is the deposited mass (ng /cm²), C is the proportionality constant (17.7 ng/Hz cm²), Δf is the frequency change (Hz) and n is the odd overtone number of the resonance frequency (Sauerbrey 1959).

The QCM-D measures another real-time parameter, the energy loss or dissipation factor (Δd), compared to the conventional QCM. If AC is turned off (Figure 2-5c), the oscillation decays as shown in Figure 2-5d and the decay rate or dissipation factor is also recorded in addition to resonance frequency and it is related to the viscoelastic properties of the mass on the sensor (Dixon, 2008).

Figure 2-5. Principles of QCM-D, the figure is taken from (Dixon, 2008). (a) A QCM-D sensor with Au surface. (b) The application of alternating current (AC) voltage across Au surfaces. (c) Short circuiting the AC. (d) The oscillation decay of the quartz material.

The monitoring the frequency and dissipation on a surface provide us an outstanding platform to analyze the state of molecular layers bound to the sensor surface, their mass, thickness and structural (viscoelastic) properties.

Chapter 3: CLONING, EXPRESSION, and PURIFICATION of DOMAIN I of NS5A PROTEIN from HEPATITIS C VIRUS

3.1. Motivation

In the present section of this study, we focused on the development of the efficient bacterial expression systems to produce large amounts of soluble NS5A-DI for structural studies. This section involves molecular cloning approaches to construct recombinant NS5A-DI for the expression and purification of various fusion proteins. Furthermore, we introduce strategies to obtain tag free native proteins and buffer screening methods to improve the solubility and stability of the protein.

3.2. Bioinformatics

We performed a bioinformatics analysis of NS5A protein sequence to understand variations within the protein from different genotypes of HCV. NS5A is a unique protein that has no human or viral homologs other than NS5A of GB virus B (GBV-B) with 21% identity (Mankouri et al., 2008). Herein, UniProt database was used to retrieve NS5A protein sequences from the different genotypes of HCV (*i.e.* 1a, 1b, 2a, 3a, 4a, 5a, 6a) (Magrane & Consortium, 2011). Clustal Omega tool aligned the corresponding sequences using NS5A from genotype 1a as the reference sequence for comparison. The sequence similarity scores, the source of sequences and UniprotIDs are listed in the Appendix A. The NS5A sequence similarity between different genotypes of HCV ranges from 63% to 80%, suggests that NS5A sequence is heterogeneous within different genotypes of HCV. The result of multiple sequence alignment is also available in Appendix A. Significant sequence variations appear within the carboxyl terminus (C-terminus) of NS5A, while the N-terminus where NS5A-DI is located is mostly conserved. It is quite possible that the N-terminal region of the protein could have functional importance as it is highly conserved.

3.3. Cloning, Expression, and Purification of NS5A-DI

To date, the constructs of the recombinant forms of the NS5A-DI that were used for

the crystallization studies are NS5A(25-215)-X-DDDDK (genotype 1b, PDB: 1zh1), AL-NS5A(33-202)-DDDDK (genotype 1b, PDB: 3fqg) and GS-NS5A(33-202)-DDDDK (genotype 1a, PDB: 4cl1). The reported construct did not define the X specifically (Love et al., 2009; Tellinghuisen et al., 2005). The alignment of sequences of the proteins is shown in Appendix B. We tested five different constructs of NS5A-DI for the efficient purification of the protein. The DNA sequence of NS5A-DI from genotype 1a of HCV without AH was cloned with N-terminal SUMO-tag, N-terminal ubiquitin tag, and C-terminal His-tag. Figure 3-1a and b show the PCR amplicons and the schematic representation of the constructs, respectively. This section will discuss the cloning, expression, and protein purification results of these constructs.

a)

b)

Figure 3-1. PCR amplicons of the constructs and their schematic map. (a) 0.8% agarose gel analysis of the following PCR products. M: 100 bp marker, lane 1: NS5A(29-213), lane 2: GSGSG-NS5A(29-213), lane 3: NS5A(29-213), lane 4: NS5A(36-198), lane 5: NS5A(36-198)-DDDDK with expected sizes of 576, 591, 569, 506, and 521 bp, respectively. **(b)** The names and designs of the constructs are illustrated here.

3.3.1. His-SUMO-NS5A(29-213) construct

pETSUMO vector utilizes a small ubiquitin-like modifier to enhance the solubility of proteins. The tertiary structure of SUMO fusion protein is recognized and cleaved by SUMO protease enzyme, and that ideally results in the generation of high-yield and

soluble native recombinant protein. Therefore, pETSUMO protein expression system was our first choice to obtain native NS5A-DI. We amplified the DNA sequence of NS5A(29-213) from genotype 1a of HCV by PCR with the primers that had *BsaI* and *XhoI* restriction enzyme cutting site. Table 2-2 shows the sequences of the primers that were used, *i.e.* HCV NS5A f29 and 1a HCV NS5A r213. The amplified gene product (seen in Figure 3-1a) and the pETSUMO vector were digested with *BsaI* and *XhoI*. Then, digested PCR product and pETSUMO vector were ligated, and the construct was transformed into *E. coli* DH5 α . After transformation, DH5 α cells were plated on LB agar containing supplemented with kanamycin. Two colonies, confirmed as positive clones by colony PCR, were sent for sequencing. Thus, their sequences were found to contain the correct reading frame and no mutation.

Once His-SUMO-NS5A(29-213) vector was amplified and purified from the DH5 α strain, it was transformed into the BL21 cell line of *E. coli* for the expression of recombinant protein. The induction and expression of the recombinant protein was performed with 1 mM IPTG at different temperatures (*i.e.* 15 and 25 °C). The theoretical molecular weight of the His-SUMO-NS5A(29-213) is 32.342 kDa. The recombinant protein was expressed better and at a higher yield at 25 °C than at 15 °C, as seen in Figure 3-2.

Figure 3-2. Protein expression control for His-SUMO-NS5A(29-213) construct. 12% SDS-PAGE analysis of the cell lysates for the following conditions. Lane 1: marker, lane 2: uninduced at 15 °C, lane 3: 3 hours after induction with 1 mM IPTG at 15 °C, lane 4: 3: overnight (O/N) induction with 1 mM IPTG at 15 °C, lane 5: uninduced at 25 °C, lane 6: 3 hours after induction with 1 mM IPTG at 25 °C, lane 7: overnight (O/N) induction with 1 mM IPTG at 25 °C.

After the protein expression had been verified, cells were lysed by sonication and purified by Ni-NTA affinity column. Fractions from purification steps (*i.e.* flow-through, wash and elute fractions) were collected and analyzed by 12% SDS-PAGE. Our initial trials of the protein purification resulted in a very small yield of the protein, as expression levels were generally low and most of the protein appeared in the cell pellet fractions (Figure 3-3a). Therefore, we needed to optimize our routine further by improving the solubility of the protein and expression conditions, as well. First, we changed the host strain to Rosetta, a derivative of BL21, which enhances the expression of proteins that contain codons rarely used in *E. coli* such as AGG/AGA (arginine), CGG (arginine), AUA (isoleucine), CUA (leucine), CCC (proline), and GGA (Glycine). Secondly, 10% of glycerol was added to lysis and elution buffers to increase solubility and stability of the protein. The purification results after optimizations are available in Figure 3-3b. Notably, the addition of 10% glycerol to the lysis buffer significantly increased the amount of soluble protein, as evidenced by the appearance of a pronounced

and specific protein band in the clarified lysate fraction. Therefore, to keep NS5A-DI soluble and stable, we included 10% glycerol in our purification buffers (lysis and elution buffers) for further experiments.

Figure 3-3. The purification of His-SUMO-NS5A(29-213) protein by affinity column. The samples, uninduced, induced, cell lysate, clarified cell lysate, the cell pellet, flow through, wash and elution fractions were analyzed on Coomassie-stained 12% SDS-PAGE. (a) Purification of His-SUMO-NS5A(29-213) from BL21 host cell line, lacking glycerol in lysis and elution buffer. (b) Purification of His-SUMO-NS5A(29-213) from Rosetta host cell line, 10% glycerol added into purification buffers (lysis and elution). Note that increased an amount of soluble protein was achieved, seen in clarified lysate and elution fractions of the gel on the right.

Next, we aimed to obtain native NS5A(29-213) without SUMO- and His-tags by using SUMO protease. It is important to note that SUMO protease cannot work efficiently at high concentrations of imidazole (more than 150 mM), making it necessary to perform buffer exchange of Ni-NTA elution fractions. Once buffer exchange was carried out using PD-10 desalting column (GE Healthcare Life Sciences), the fusion protein was mixed with SUMO protease and incubated at 4 °C overnight. As SUMO protease was unable to cleave SUMO protein from NS5A(29-213), we tested different temperatures (4 °C, room temperature around 25 °C, 30 °C) and different buffer exchange methods (PD-10 column and dialysis) to cleave the SUMO-tag. The sizes of the fusion protein and SUMO-tag were 32.342 kDa and 12.036 kDa, respectively. In a successful cleavage, we expected to see the appearance of a band around 20 kDa, corresponding to NS5A(29-213). However, we observed that the fusion protein remained uncut in all conditions and migrated around 32 kDa, as seen in the 12% SDS-PAGE analysis of the samples in Figure 3-4. Interestingly, we observed the appearance of another band around the size of 20 kDa, even though SUMO protease was not added (see the control sample, named as “uncut”). This suggested to us that the protein might be broken down into smaller polypeptides due to protease contamination.

Figure 3-4. SUMO-tag cleavage SDS-PAGE analysis of His-SUMO-NS5A(29-213). Following the purification of recombinant protein, its buffer was exchanged. SUMO protease was added, and they were incubated at different temperatures. Lane 1: marker, lane 2: only His-SUMO-NS5A(29-213) fusion protein (32 kDa), without SUMO protease, lane 3: buffer exchange with PD-10 column and SUMO protease added (26 kDa), mixture was incubated at 4 °C for overnight, lane 4: same conditions with lane 3, at 25 °C, lane 5: same conditions with lane 3 at 30 °C, lane 6: buffer exchange was performed by dialysis, SUMO protease added and incubated at 4 °C for overnight, lane 7: same conditions with lane 6 at 25 °C, lane 8: same conditions with lane 6 at 30 °C. Note that the 32-kDa fusion protein is seen in all SUMO protease added lanes which mean SUMO-tag cleavage failed in all conditions.

In this construct, we could not get the native protein. One reason for this could be that the N-terminal region of NS5A(29-213) is quite close to the membrane anchoring region and harbors hydrophobic residues, especially the first nine residues “PQLPGIPFV”. Such hydrophobic residues might in tight association with the SUMO-tag, thereby making it inaccessible for cleavage by the SUMO protease. Hence, SUMO protease may not recognize the shape of the tag, which is necessary for the successful cleavage of the SUMO-tag.

3.3.2. His-SUMO-GSGSG-NS5A(29-213)

As we could not cleave the His-SUMO-tag, we designed another pETSUMO construct to circumvent SUMO-tag cleavage problem. The newly constructed His-SUMO-GSGSG-NS5A(29-213) has additional five residues Gly-Ser-Gly-Ser-Gly (GSGSG) between the SUMO-tag and NS5A(29-213). The aim was to interrupt the putative association between SUMO and the N-terminal region of NS5A protein and to achieve SUMO-tag cleavage. The cloning and purification of His-SUMO-GSGSG-NS5A(29-213) were performed as described in the previous section. The gene fragment, shown in Figure 3-1a, was amplified by using the primers 1a HCV NS5A f29-GSGSG and 1a HCV NS5A r213 from the template- of NS5A from genotype 1a, the primer sequences are available in Table 2-2 . The positive clones were identified by colony PCR and their sequences were confirmed by DNA sequencing.

We induced protein expression with 1 mM IPTG at 25 °C. The clarified cell lysate samples (uninduced, 4 hours and 12 hours after IPTG induction) were analyzed on 12% SDS-PAGE. As can be seen in Figure 3-5a, the fusion protein expression was successful, since a band corresponding to His-SUMO-GSGSG-NS5A(29-213) appeared around 33 kDa. Protein purification was performed as described in Section 2.2.13 and the fractions collected from each purification step were run on 12% SDS-PAGE (Figure 3-5b). The elution fractions revealed that the yield of desired protein was very low. The expected size of the fusion protein was around 32.915 kDa. Notably, however, another band running at a lower molecular weight than the desired protein (around 20 kDa) appeared, suggesting a breakdown of the protein. In comparison with the previous construct, the amount of this smaller size protein was increased significantly. Furthermore, since the small protein band appears in Ni-NTA elution fractions but not in the cell lysate or clarified cell lysate, it is indicative of the protein degrading during the course of purification. Presumably, the addition of five extra amino acid residues between the N-terminus of NS5A(29-213) and the SUMO-tag may have disrupted the association of the SUMO-tag with the NS5A(29-213). Thereby, the N-terminal region of the protein may have been rendered more accessible to proteolytic degradation, which in turn may provide an explanation for the observed increased levels of the shorter protein fragment.

Figure 3-5. Expression and purification of His-SUMO-GSGSG-NS5A(29-213) protein. 12% SDS-PAGE analysis of the protein at following conditions. (a) Induction of protein expression was tested at 25 °C. Cells, from Rosetta strain of *E. coli*, were induced with 1mM IPTG for 4 and 12 hours as indicated. His-SUMO-GSGSG-NS5A(29-213) was expressed, as seen around 33 kDa. (b) Fractions from each step of purification of His-SUMO-GSGSG-NS5A(29-213) protein are shown. The size of protein was 32.915 kDa, which is clearly seen in both cell lysate and clarified cell lysate fractions. Note that the elution fractions revealed another band smaller than expected size around 20 kDa, suggesting the breakdown of the recombinant protein at a significant level.

3.3.3. UBI-NS5A(29-213)-His construct

As previous purifications of NS5A(29-213) using an N-terminal SUMO-tag were unsatisfactory, we decided to use an expression strategy that involved co-transformation of pETUBI and PCG1 plasmids for the efficient purification of native NS5A-DI (Marcotrigiano & Tellinghuisen, 2009). The plasmid pETUBI carries a ubiquitin-tag that is fused to the N-terminus of the recombinant protein. Due to its highly stable structure, the ubiquitin-tag is thought to serve as a scaffold to facilitate the folding of the fusion protein. The PCG1 plasmid, on the other hand, encodes for a yeast ubiquitinase enzyme that removes the ubiquitin-tag during expression, resulting in the generation of untagged native proteins.

First, PCG1 plasmid was transformed into *E. coli* BL21 using the selection marker of the PCG1 plasmid, namely chloramphenicol resistance. Subsequently, the BL21 cell line carrying PCG1 was made chemo-competent by CaCl₂ treatment. Resulting competent cells were called as BL21(PCG1). NS5A(29-213) was amplified and cloned into pETUBI plasmid, named as UBI-NS5A(29-213)-His. The UBI-NS5A(29-213)-His construct was then transformed into BL21 (PCG1), and cells were plated on LB agar plates containing 30 µg/µL of kanamycin and 25 µg/µL chloramphenicol. Protein expression was induced with 1 mM IPTG at 25 °C for 6 hours. The protein was well-expressed, since a band migrating slightly above 25 kDa appeared (“induced” lane of Figure 3-6). Next, we purified the protein and performed 12% SDS-PAGE analysis of the purification process as shown in Figure 3-6. The size of UBI-NS5A(29-213)-His was 28.374 kDa, and the size of the protein without UBI-tag, NS5A(29-213)-His, was 20.441 kDa. The size at which the eluted protein band migrated was bigger than the theoretical size of desired protein, as seen in the elution fractions of Figure 3-6. The band was verified as UBI-NS5A-DI by mass spectrometry. In this section, we showed that UBI-tag could also not be removed from NS5A(29-213). Moreover, the expression level of the fusion protein was significantly lower than observed in the pETSUMO expression systems, as well. Therefore, we stopped using pETUBI expression plasmids for further studies.

Figure 3-6. Protein purification of UBI-NS5A(29-213)-His. The samples from each purification step were analyzed on a Coomassie-stained SDS gel (12%). Lane 1: marker, lane 2: cell lysate before IPTG, lane 3: cell lysate after IPTG induction, lane 4: cell lysate, lane 5: clarified cell lysate, lane 6: pass through, lane 7: wash, lane 8-9-10: subsequent elution fractions. UBI-tag could not be removed in the cell, and the fusion protein (with UBI-tag) was purified.

3.3.4. NS5A(36-198)-His construct

We added SUMO- and-ubiquitin tags at the N-terminal of the target protein to increase the solubility of the recombinant protein. However, we had problems with removing tags. This was undesired for structural studies, particularly to perform NMR, which is a suitable method for proteins smaller in size. Therefore, we decided to clone the protein in a plasmid that does not carry any tag. We selected pET29b plasmid, which had only His-tag (~0.8 kDa) at the C- terminal of the multiple cloning sites. Moreover, to overcome the protein breakdown problem, we shortened the length of the protein by seven amino acids from N-terminal, the sequence of the first crystal structure was taken as reference (Tellinghuisen et al., 2005). The gene fragment was amplified by using primers 1a HCV NS5A f36 with a NdeI restriction enzyme recognition site and 1a HCV NS5A r198 with XhoI restriction site, the PCR amplicon is shown in Figure 3-1.

We induced the protein expression with 1 mM IPTG at 25 °C and purified NS5A(36-198)-His-tag protein by Ni-NTA column. The fractions collected from protein purification steps were run on 12% SDS-PAGE, Figure 3-7a. The molecular weight of the protein is 19.15 kDa, and it is seen that the expression and purification of the protein

were successful. We wanted to improve the purity of the protein further by using size-exclusion chromatography (SEC). The fractions eluted from Ni-NTA column were combined, concentrated, and then loaded to FPLC. The SEC profile of the protein is seen in Figure 3-7b. We observed two peaks around 40 to 50 mL. The protein was supposed to elute around 85-100 mL from the Superdex 75 16/60 column, based on the size of the protein. These results suggested that NS5A(36-198)-His forms aggregates. The theoretical isoelectric point (pI) of this protein was 8.15, calculated by ProtParam tool. To prevent the formation of aggregates, the pH of the buffers should be around minimum one unit below or above of the pI of the protein, which makes the overall charge of the protein positive or negative. Therefore, we tested buffers (potassium phosphate at pH 6.0, 7.0, and 9.0) for the purification of the protein. However interestingly, it did not help much in our case, and we could not concentrate the protein more than 2 mg/mL in our preliminary attempts.

Figure 3-7. Expression and purification of NS5A(36-198)-His protein. (a) 12% SDS-PAGE analysis of the protein expression for uninduced and induced with 1 mM IPTG at 25 °C and protein purification steps. Note that the molecular weight of the protein is 19.15 kDa and both expression and purification are successful. (b) The elution fractions from a were combined and concentrated and were loaded to SEC, Superdex 75 16/60. Irreversible protein aggregates were eluted around 40-50 mL.

3.3.5. His-SUMO-NS5A(36-198)-DDDDK construct

The protein with only His-tag was quite prone to form aggregates, particularly when it was concentrated to >2 mg/mL. As the protein alone without any tag was unstable,

we sought other strategies to obtain stable and soluble protein at higher concentrations for structural studies. The addition of enterokinase sequence (Asp-Asp-Asp-Asp-Lys) was reported to increase the solubility of NS5A-DI (Lambert et al., 2014; Love et al., 2009). For this reason, we produced another construct spanning the NS5A(36-198) with an N-terminal SUMO-tag and C-terminal enterokinase tag (DDDDK). The PCR product was amplified by using primers 1a HCV NS5A f36 (BsaI cut site) and 1a HCV NS5A r198 (XhoI cut site + DDDDK sequence). The PCR product is shown in Figure 3-1a above.

The His-SUMO-NS5A(36-198)-DDDDK plasmid construct was transformed into Rosetta strain of *E. coli*. The protein expression was induced with 1mM IPTG at 25 °C for 6 hours. We purified the fusion protein by Ni-NTA column and collected the fractions from each purification step. The 12% SDS-PAGE analysis of the fractions are shown in Figure 3-8a. The elution fractions revealed that the yield of the protein was improved significantly. In Figure 3-8b, the first lane (control) shows the elution fractions from Figure 3-8a without any SUMO protease. The second lane (S.P. cut) shows the elution fractions incubated with SUMO protease at 4 °C for overnight. And, the last lane shows the second-round Ni-NTA purification after SUMO protease cleavage. The SUMO protease cleavage was successful and second round of Ni-NTA column purification was done. The proteins after the second round of Ni-NTA column was concentrated and loaded to Superdex 75 16/60 column. The native protein without any tag, NS5A(36-198)-DDDDK eluted at around 90 mL, which corresponds to the monomer size of the protein, shown in Figure 3-8c. The inset figure in Figure 3-12c shows the 12% SDS-PAGE analysis of the monomer size peak, shaded area, showing protein was highly pure.

Figure 3-8. Purification of the monomeric state NS5A(36-198)-DDDDK protein. (a) 12% SDS-PAGE analysis of Ni-NTA column purification of His-SUMO-NS5A(36-198)-DDDDK protein. (b) SDS-PAGE analysis of the following samples, lane 1: elution fractions from (a), lane 2: Sumo protease was added and incubated at °C 4 for overnight, lane 3: after second round Ni-NTA column purification. (c) The protein after second round of purification was loaded to Superdex 75 16/60 column. The native protein without any tag, NS5A(36-198)-DDDDK, eluted at around 90 mL, corresponds to monomer size. The inset shows the 12% SDS-PAGE analysis of the shaded area, showing protein was highly pure.

3.4. 1D and 2D NMR of NS5A(36-198)-DDDDK

We successfully purified NS5A(36-198)-DDDDK protein and then we performed 1D and 2D NMR studies to get information about its folding state and tertiary structure. 0.1 mM of purified protein was used for 1D NMR. Peaks in 1D NMR spectrum were

observed to be well-dispersed methyl peaks of methyl proton (−0.5-1.5 ppm), α -proton (3.5-6 ppm), and amide proton (6-10 ppm), shown in Figure 3-9a. Next, we performed a 2D NMR (^1H , ^{15}N HSQC spectrum) that gives unique peaks for all amino acids except proline. The well-dispersed peaks of backbone amide provide information about the tertiary structure of the protein. Figure 3-9b represents the 2D NMR analysis of NS5A(36-198)-DDDDK indicated that protein folded well.

Figure 3-9. NMR experiments showing a well-folded state of purified NS5A(36-198)-DDDDK. (a) 1D ^1H NMR spectrum of NS5A(36-198)-DDDDK Methyl H appeared between 0 to -1 ppm which further indicates the protein is in a folded state. (b) 2D ^1H , ^{15}N -HSQC spectrum of the NS5A(36-198)-DDDDK. Spectrum was collected on 600 MHz spectrometer at 298 K, protein was at 100 μM in 25 mM potassium phosphate, 150 mM NaCl, 5% glycerol, pH 7.5. Most peaks are resolved individually and the backbone amides are dispersed well.

After we had obtained 2D ^1H , ^{15}N -HSQC of the protein, we performed NMR titration experiments to explore the interaction between the protein and 15U RNA molecules. We have selected U-rich RNA molecules as it was reported that NS5A has the highest affinity to such RNAs (Hwang et al., 2010). We collected the 2D spectrum of the free proteins at 100 μM in a buffer composed of 25 mM potassium phosphate, 150 mM NaCl and 5% glycerol (pH 7.5) as shown with red peaks in Figure 3-10. Then, 2D spectrum of the proteins with 50 μM of 15U RNA (1 to 0.5 ratio) and 100 μM of 15U RNA (1 to 1 ratio) were collected (shown as green and blue peaks accordingly in Figure 3-10). We observed that several peaks shifted upon the addition of 15U RNA compared to the apo protein and some peaks disappeared. Such a behavior in NMR spectra may suggest that the RNA molecules do not specifically interact with a certain motif of protein, whereas it affects the overall protein. It would be necessary to gather complementary information, such 3D NMR based backbone assignments of the peaks, to draw a solid conclusion regarding this picture. Therefore, we attempted to make backbone and side-chain assignments of the protein using 3D NMR. These experiments require the protein to be stable at room temperature for around three days. However, as the protein was not stable enough at room temperature, we first had to conduct buffer optimization studies to improve its stability. We introduce these efforts below.

Figure 3-10. 2D ^1H , ^{15}N HSCQ spectra of NS5A(36-198)-DDDDK with 15U RNA. The experiments conducted in 25 mM potassium phosphate, 150 mM NaCl, 5% glycerol, pH 7. 5. Red peaks, the protein alone at 100 μM , green peaks the protein (100 μM) with 15U RNA (50 μM), blue peaks the protein (100 μM) with 15U RNA (100 μM).

3.5. Optimizing the buffer conditions for protein stability by using thermal denaturation assay

There are several factors to be considered when designing optimal buffers like pH, buffering systems, salts, reducing agents, and some additives/stabilizers. These factors play important roles in stabilizing proteins. And, they may alter unfolding/melting temperature of the proteins. The thermal denaturation assay measures the thermal stability of the proteins using a fluorescent protein-binding dye. Fluorescence dye (Sypro orange dye) is mixed with the protein and heated, and as protein unfolds, the dye binds to the hydrophobic regions of the proteins and that results in altered fluorescence measurement in thermal denaturation assay in different buffers. A representative fluorescence measurement is shown in Figure 3-11.

Figure 3-11. A representative fluorescence measurement of thermal denaturation assay of NS5A(36-198). The assay used to screen the effect of various pH conditions here. The data shows the melting profiles of protein at pH 2.7, 5.0, 6.6, 7.4, 8.4, and 10. Data were collected at 1 °C intervals ranging from 25 to 100 °C.

Protein thermal denaturation assay was utilized to perform high-throughput screening of buffer conditions to determine the optimal conditions for stable NS5A(36-198). First, we screened the effect of different pH values on HCV NS5A(36-198)-DDDDK protein. 50 mM potassium phosphate buffers, varying from pH 2.6 to 9.4, were screened. In this experiment, we kept buffering system (potassium phosphate) and salt (150 mM NaCl) the same in all conditions to observe the changes only coming from the effect of pH, eliminating the effect of ionic strength of buffering systems on the melting temperature of proteins. We observed that the pH around 7.4 stabilized the protein, NS5A(36-198)-DDDDK, Figure 3-12a. The buffering systems, Tris-HCl, potassium phosphate, HEPES, MOPS, sodium phosphate was tested as well. HEPES buffer was appeared to increase the stability of the protein very slightly. HEPES is an organic chemical buffering agent zwitterionic in nature, carrying multiple positive and negative charges. These multiple ionic effects of HEPES might be playing a role in stabilizing the protein.

Figure 3-12. pH 7.4 and HEPES buffer improves the thermal stability of NS5A(36-198). (a) Melting temperature (T_m) at varying pH values in of 50 mM potassium phosphate buffer NS5A(36-198)-DDDDK and approximately pH 7.4 has the highest stabilizing effect. (b) NS5A(36-198)-DDDDK in different buffering systems at pH 7.4 were tested, and HEPES buffer appeared to increase T_m slightly.

Next, the effect of certain salts (ZnCl₂ and NaCl) were tested to optimize further the conditions for enhancing the stability of the NS5A(36-198). As the N-terminal region of NS5A is a Zn-binding domain, the effect of ZnCl₂ (0-15 mM) were screened. The melting temperature curves did not suggest any stabilization effect of ZnCl₂ as shown in Figure 3-13a. Many buffers contain NaCl to keep protein soluble and to mimic the cellular environments. In this part, we tested the NaCl concentrations from ranging from 0 to 800 mM. The data given in Figure 3-13b shows the increase in melting temperature with the increasing concentrations of NaCl, indicating higher concentrations of NaCl, *i.e.* 400 to 600 mM, have a role in the stability of the protein.

Figure 3-13. The effect of salts on the thermal stability of NS5A(36-198)-DDDDK. The effect of (a) ZnCl₂, (b) NaCl on the thermal stability of the protein. ZnCl₂ did not improve, however NaCl has a stabilizing effect at around 400 to 600 mM.

After checking the effect of salts, we also tested the impact of the reducing agents on the stability of NS5A-DI using two reducing agents, BME and DTT. Interestingly, BME significantly destabilized the protein as it can be seen in Figure 3-14a. DTT, on the other hand, made rather a negligible impact on the stability of the protein (Figure 3-14b).

Figure 3-14. The effect of reducing agents on the thermal stability of NS5A(36-198). (a) BME and (b) DTT on the thermal stability of the protein. BME destabilizes the protein, however DTT had no significant stabilizing or destabilizing effect.

Although we have optimized the pH, buffering systems, salt, and reducing agent conditions, we still observed that the purified NS5A(36-198) protein exists as a heterogeneous mixture, different in size based on DLS analysis. (We present these

results in the next section). In our preliminary studies, we also tested various stabilizers like glycerol, arginine, LiCl, Triton X-100 detergent and PEG varieties to obtain homogenous protein sample at higher concentrations suitable for crystallization and NMR studies. Again, we could obtain no clear benefit of additives. Thus, all these observations showed the inherent and strong self-association tendency of the protein as it is not easy to keep it homogeneous in monomer form.

3.6. Overview of chapter

We explored five different NS5A-DI constructs by using three different plasmids pETSUMO, pETUBI, and pET-29b (+). In longer constructs of the protein residues from 29-213, we had problems to remove SUMO and UBI-tags. And, another construct that has extra amino acids between SUMO-tag and the NS5A(29-213) resulted in an instant breakdown of proteins into shorter fragments. These results showed us that very N-terminal part of the protein is highly difficult to handle as it was sticky most probably due to the hydrophobic nature of the amino acids. And, the protein also looks like quite prone to protein proteolysis based on the observation of fragmented proteins frequently. A shortened version of the NS5A(36-198) with a C-terminal His-tag formed rapid, irreversible aggregates. For this reason, we produced another construct spanning the NS5A(36-198) with an N-terminal SUMO-tag and C-terminal enterokinase tag (DDDDK). In this construct, we could cleavage SUMO-tag, removing 7 residues “PQLPGIPFV” from N-terminal of NS5A circumvented both cleavage and degradation problems. Moreover, the solubility of the protein increased strikingly, perhaps the charged residues of enterokinase tag helped. The enterokinase tag altered the pI of the protein as well. Important to note, pI of the protein alone was around 8, and it dropped to 5.7 with the addition of tag.

We performed 1D and 2D NMR, ^1H and ^{15}N HSQC experiments, for the NS5A(36-198)-DDDDK and they indicated that the protein is well-folded. 15U RNA titration experiments of the protein showed the shifts in several peaks and some of them were disappeared. The protein was not stable for long term to perform protein backbone assignment, which is crucial to see peaks (amino acids) altered upon the addition of RNA molecule. Therefore, we screened several buffer conditions, which helped us improve the stability of the protein. However, we still observed nonhomogeneous mixtures of purified protein in solution, which stands as a bottleneck for performing

NMR measurements. Therefore, we tended towards biophysical tools for in solution characterization of the protein. These efforts are presented in the next section.

Chapter 4: SELF-ASSOCIATION of NS5A DOMAIN I of HEPATITIS C VIRUS: BIOPHYSICAL INVESTIGATIONS in SOLUTION

4.1. Motivation

In the previous chapter, we described our strategies to obtain homogeneous monomer form of the NS5A-DI of genotype 1a that we wanted to use for NMR studies to identify RNA and drug binding interfaces on the protein. However, neither the other groups nor we achieved that yet. And, we still do not have NMR backbone assignments of the DI, although several crystal structures were previously obtained by other groups (see Figure 1-13 in the Introduction Section).

The HCV genotypes are very diverse and classified into seven different genotypes and genotype 1a, and 1b are the most prevalent species of the virus with 83.4 million cases around the world (46.2% of the total infections) (Messina et al., 2015). Experiments to identify similarities and differences in the profile of resistance NS5A from genotype 1a and genotype 1b revealed that DCV was around 5 to 10 times less efficient against genotype 1a (Fridell et al., 2010). Moreover, these variations, the amino acid substitutions in the N-terminal region of NS5A, caused the resistant type of the virus against the drug (see Figure 1-15 in the Introduction Section). They also reported that the drug response and effect of mutations were greater in genotype 1a rather than 1b. Taken all these together, we decided to perform a more comprehensive study of NS5A-DI to investigate its behavior in solution. We primarily aimed to compare the differences in between NS5A-DI from genotypes 1b and 1a under the same buffer conditions in solution.

4.2. NS5A-DI of genotype 1a and genotype 1b of HCV share high sequence similarity

To investigate the differences between the NS5A-DI from genotypes 1a and 1b in solution, we have designed two recombinant forms of NS5A-DI proteins that lack AH

and comprises residues from 36 to 198. The first form is from recombinant H77 (genotype 1a) and the second form is from isolate Con1 (genotype 1b). The 3' primers were designed in a way to encode (Asp)4-Lys residues which are the enterokinase recognition sequences that enhanced the solubility of NS5A-DI at higher concentrations. The Figure 4-1 shows the protein sequence of NS5A(36-198)-DDDDK of genotypes 1a and 1b. They are aligned by Clustal Omega tool, and the similarity score is 91.9%. Please note that in this chapter, we only used NS5A(36-198)-DDDDK from genotype 1a and genotype 1b. NS5A(36-198)-DDDDK will be called as NS5A-DI during this chapter.

```

1a  FVSCQRGYRGVWRGDGIMHTRCHCGAEITGHVKNGTMRIVGPRTCRNMWSGTFPINAYTT
1b  FFSCQRGYKGVWRGDGIMQTTCPGQAQITGHVKNKSMRIVGPRTCSNTWHGTFPINAYTT
    *.*****:*****:* * *:*****:***** * * *****

1a  GPCTPLPAPNYKFALWRVSAEEYVEIRRVGDFHYVSGMTTDNLKCPCQIPSPFEFFTELDG
1b  GPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVPAPFEFFTEVDG
    ***** . *****:*****: *****:*****:*****:*:*****:**

1a  VRLHRFAPPCKPLLREEVSFRVGLHEYVGSQLPCEPEPDVAVLTSMDDDDK
1b  VRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPCEPEPDVAVLTSMDDDDK
    *****:* *****:* *:*: *****:*****

```

Figure 4-1. Pairwise sequence alignment of NS5A(36-198)-DDDDK from genotypes 1a and 1b using Clustal Omega tool. The total number of the amino acids are 172 in both constructs. 144 of 172 (83.7%) are identical, represented as (*). 17 of them (9.9 %) are conserved substitutions, and semi-conserved substitutions marked with (:), and (.), respectively. 11 of them (6.4%) are different, indicated as ( ).

4.3. NS5A-DI from genotype 1b of HCV self-associates in solution and forms reversible higher-order oligomers

The His-SUMO-tagged recombinant form of NS5A-DI (genotype 1b, con1), lacking N-terminal amphipathic helix was expressed in *E. coli*, Rosetta (DE3) strain, using pETSUMO vector. Western blot analysis confirmed the His-SUMO-tagged protein by His-tag antibody and then His-SUMO-tag was successfully cleaved. The protein was further purified by SEC (Superdex 200- 10/300GL) resulted in highly pure protein as shown SDS-PAGE in Figure 4-2. A mixture of protein standards from Bio-Rad, containing thyroglobulin (670 kDa), γ -globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa), and vitamin B12 (1.35 kDa) was run on Superdex 200-10/300 column. (Elution volumes of gel filtration standards are shown in Appendix E). The

purified NS5A-DI using Superdex 200 column revealed that most of the protein appeared at monomer size (~19 kDa), around the same elution volume of myoglobin which was 17 kDa.

Figure 4-2. Monomer state of NS5A-DI from genotype 1b of HCV. The protein eluted at around 17 mL from Superdex 200 10/300, and the inset shows the purity of the eluted fraction (shaded area) checked by SDS-PAGE. Note that the calculated molecular weight of the protein is 19.011 kDa.

Crystallization studies of NS5A-DI (both genotype 1a and 1b) have appeared as dimers (Lambert et al., 2014; Love et al., 2009; Tellinghuisen et al., 2005). For that reason, we analyzed NS5A-DI by dynamic light scattering (DLS) to investigate the particle size in solution to assess whether they exist as monomeric, dimeric or multimeric form. The monomers (shaded area from Figure 4-2) were collected and concentrated to 2.5 mg/mL in SEC buffer (25 mM HEPES, 150 mM NaCl, 10% glycerol) before DLS analysis. DLS assay revealed a bimodal distribution of two peaks with peak maxima of ~15 and ~255 nm, as seen in Figure 4-3. The DLS data given in Figure 4-3 is based on the intensity-averaged size distribution and the intensity of scattered light is proportional to the sixth power of the particle diameter (Karahana et al., 2015). As the intensity of smaller particles (peak-1: ~15 nm) are more than two-fold higher than the intensity of larger particles (peak-2: ~255 nm), we could reach to a

conclusion that the smaller particles are much more populated than the larger particles. The molecular weight of NS5A-DI monomer is ~19 kDa. And, based on the reported crystal structures of protein dimer, the monomer size of NS5A can be estimated to be around 2.3 nm in width and 5 nm in height approximately (Appendix-D). Therefore, the smaller particles should correspond to monomers and/or dimers in the first place. However, it should be kept in mind that, this peak (peak-1) is quite wide covering a range up to 50 nm, which may include higher-order structures or small aggregates, as well. We will interpret this with the support of SEC, crosslinking assay, and SAXS analysis below.

Figure 4-3. The size distribution analysis of purified NS5A-DI from genotype 1b by DLS. (2.5 mg/mL at 25 °C in 25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol.)

DLS suggested that the purified NS5A-DI of genotype 1b was heterogeneous in solution and indicated that the protein self-associates in solution. Therefore, we further investigated the self-association behavior of the protein using SEC. We collected the monomer fraction of the protein (shaded area in Figure 4-2), concentrated at varying concentrations and reloaded to the same column. Broadly speaking, we observed a clear shift toward smaller elution volumes at different concentrations, ranging from 58 to 268 μ M, as shown in Figure 4-4. SEC column contains hydrated beads made of crosslinked polymers. The elution profile of the protein through the column depends on partitioning between two solvent spaces, the solvent spaces around the spherical beads and the

solvent within the composite dextran/crosslinked agarose matrix inside the beads. Smaller proteins require a greater volume of buffer to pass through column before they elute (Gell, Grant, & Mackay, 2012). As we observed a clear continuous shift towards to smaller elution volumes, data suggest that monomers formed higher-order structures. In self-associating systems appearance of single peaks would be observed if the dissociation constant $k_d < 1 \mu\text{M}$ (single oligomer peak) or $k_d > 1 \text{ mM}$ (single monomer peak) (Stevens, 1989). Yet, an intermediate k_d value would cause a continuous shift due to the diffusion and dispersion of the sample peak, similar to what we observed in Figure 4-4.

Figure 4-4. Elution profiles of the purified NS5A-DI from genotype 1b. The monomer size of the protein was concentrated and reloaded to SEC (Superdex 200 10/300 GL) at different concentrations, ranging from 58 to 268 μM , are shown. The elution volume of the NS5A-DI peak decreases with increasing loading concentrations of the protein (from 58 to 268 μM), is an evidence of the self-association behavior of the protein. M: monomer, D: dimer, H: higher order oligomers. The protein in standards (shown in the Appendix-E) were overlaid as a reference.

After demonstrating that NS5A-DI of genotype 1b forms oligomers at higher concentrations, as the next step, we examined whether these oligomers are reversible. To address this question, the size corresponding to the higher-order forms of the proteins eluted before 12.5 mL (indicated as a red box inset in Figure 4-5) were collected and injected to SEC again. The elution profile of the higher-order molecules (Figure 4-5), demonstrate that they dissociated to smaller size subunits, which confirms that higher-order structures are not irreversible aggregates.

Figure 4-5 Higher-order oligomers of NS5A-DI from genotype 1b are reversible. Higher-order oligomers fractions of NS5A-DI (shown in the red box) were collected and reloaded to S200 column, and their dispersion to smaller states of protein observed. This confirms that they were not aggregates of NS5A-DI.

Next, we also wanted to see the effect of concentration on the self-association behavior of the proteins. The proteins at varying concentration (51-164 μM) were reloaded to the same column, and we targeted to load the same amount of proteins (approximately 500 μg) as seen in Figure 4-6a. Elution volumes of gel filtration standards are shown in Appendix-E as a reference. We could see that the intensity of the absorbance spectra of the monomer are significantly decreasing while that of higher-order oligomers are slightly increasing as the protein concentration increases. To better analyze this data, we normalized all absorbance spectra by the peak maxima of monomer region for each concentration, as shown Figure 4-6b. The normalized SEC profiles of the proteins suggest that as the concentration of the protein increases, the equilibrium shifts toward higher-order states of the protein. Here, SEC study provided an additional insight regarding the self-association behavior of the protein. An interesting observation of the SEC analysis is that the dimeric forms were almost undetectable, despite all the available crystal structures are dimeric.

Figure 4-6. (a) Same amount of NS5A-DI from genotype 1b (500 μ g) at indicated concentrations were loaded. (b) Normalized SEC elution profiles of the NS5A-DI from genotype 1b at different concentrations are shown.

The protein mixture was further examined by chemical crosslinking experiments with bismaleimido-hexane (BMH). BMH is a maleimide crosslinker for covalent and irreversible conjugation of sulfhydryl groups (cysteines in proteins) with an arm length of 13.0 Å, 11-atom spacer arm. The protein has nine cysteine residues on each monomer form and the Appendix-C shows the cysteines on the dimer structure of NS5A-DI. The fractions, corresponding to monomer size, were collected and concentrated, then crosslinked through their cysteine residues by using BMH. The crosslinked, heated samples at different concentrations (ranging from 40 to 200 μ M) were run on 10% SDS-

PAGE, shown in Figure 4-7. The irreversible chemical crosslinking of protein resulted in monomers, dimers, and higher-order oligomers. Interestingly, dimers seem abundant on the chemically-crosslinked proteins whereas the monomer and higher-order bands were very faint.

Figure 4-7. BMH crosslinking of NS5A-DI of genotype 1b. The purified NS5A-DI from genotype 1b was crosslinked by BMH, and the reaction was terminated by DTT. Irreversible crosslinked samples were analyzed on 10% SDS-PAGE. The gel image on the left shows the same amount (10 μg) of crosslinked samples at 40, 75, 105, 150, and 200 μM concentrations. On the right gel image, 10 μL of samples were loaded at the indicated concentrations.

4.4. Low-resolution structure: The oligomers of NS5A-DI from genotype 1b is an elongated polymer of dimer subunits

Small angle X-ray scattering (SAXS) is a biophysical method to study the overall shape and the size of the particles in solution, and it applies to a broad variety of experimental conditions. Like X-ray crystallography, SAXS deduct structural information from the interaction of X-rays with the three-dimensional distribution of electrons in the sample. Several information is obtained from one single experimental scattering profile like molecular weight, foldedness state, flexibility, particle volume,

and maximum dimension (D_{max}) of the particles. Furthermore, with the recent advances in the data analysis algorithms, it provides quantitative structural information for the systems that are dynamic and polydisperse (Kikhney & Svergun, 2015; Skou et al., 2014). We thus decided use SAXS as a tool to further characterize the self-association behavior of protein mixture.

In homogenous samples, the scattering data of a single particle is directly correlated with the overall parameters of the particles in the system as they are identical. However in heterogeneous samples, the experimentally determined values represent the averaged characteristics of the systems, and they need to be further analyzed to get information of individual components in the mixture (Vestergaard, 2016). One of the major advantages of SAXS analysis is that the modern SAXS devices may resolve complex data using advanced data analysis tools.

To investigate the structural features of NS5A-DI from genotype 1b in solution, we performed SAXS measurements of the protein at two concentrations, 2.5 and 3.7 mg/mL (Figure 4-8a). The scattering patterns were measured using a sample to detector distance of 0.67 m and a wavelength of $\lambda=1.3414 \text{ \AA}$, which covers a range of momentum transfer of $0.016 < q < 0.4 \text{ \AA}^{-1}$ ($q = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle). The Guinier plots at low angles appeared linear and confirmed good data quality with no indication of protein aggregation (Figure 4-8a, *inset*). The R_g -values derived from the Guinier region was determined to be 37.55 ± 1.35 and $46.10 \pm 1.41 \text{ \AA}$ at 2.5 and 3.7 mg/mL, respectively (Table 4-1). The distance distribution function ($P(r)$) showed a single peak and had a right-skewed distribution (Figure 4-8) with the maximum particle dimension (D_{max}) of 127 ± 10 and $145 \pm 10 \text{ \AA}$ at 2.5 and 3.7 mg/mL, respectively, indicating that NS5A-DI from genotype 1b has an elongated structure in solution. The R_g -values extracted from the $P(r)$ function were 39.38 ± 0.48 and $46.80 \pm 0.34 \text{ \AA}$ at 2.5 and 3.7 mg/mL, respectively, are in agreement with the R_g -values extracted from the Guinier region within the error range (Table 4-1). The normalized Kratky plot provides the information about the potential flexible behavior and/or extended shape of the examined particle (Durand et al., 2010). For a well-folded globular protein like lysozyme, the plot will show a defined bell curve profile as seen in Figure 4-8c. In comparison, the normalized Kratky plot of NS5A-DI from genotype 1b exhibited a very broad bell-shaped profile and its peak is well shifted toward the right with respect to standard globular proteins, indicating the presence of motion and/or being extended. Based on

the Porod and excluded volumes and volume of correlation the molecular mass of NS5A-DI from genotype 1b was calculated to be around 65 kDa at 2.5 mg/mL concentration and around 100 kDa at 3.7 mg/mL concentration in solution. The increase in R_g and D_{max} values, as well as the molecular mass, indicates that as the solute concentration increases the particle size increases in solution suggesting the formation of higher order oligomers. The two different crystal structure dimers from genotype 1b, “clam-like” dimer (PDB ID: 1zh1) and “back-to-back” dimer (PDB ID: 3fqj) (Figure 1-13), had a poor fitting to the experimental scattering data with χ^2 value of 4 for the 2.5 mg/mL concentration data. Neither the monomer nor the tetramer (generated based on symmetry-related molecules in the crystal structure dimer) fit the data ($\chi^2 = 8.7$ for monomer and 1.5 for tetramer for the 2.5 mg/mL concentration data). Since no single oligomer population was fitting the experimental data, a mixture of oligomers population were considered for further analysis of the scattering data for NS5A-DI from genotype 1b as evidenced also by SEC studies exhibiting a polydisperse population in solution. The OLIGOMER program was employed, which fits the experimental scattering intensities from a multi-component mixture of proteins and estimates the volume fractions of each independent scattering components in the solution. The analysis revealed that NS5A-DI from genotype 1b has a mixture of population composed of monomer, hexamer and decamer with a volume fraction of 61:18:21 at 2.5 mg/mL and 41:17:42 at 3.7 mg/mL of protein, and reveal a good agreement with the experimental scattering data with a χ^2 value of 0.5 (Figure 4-8d). The analysis suggested that the higher-order oligomer ratio is increasing as the protein concentration increases which correlates well with the SEC data above.

Figure 4-8. SAXS studies of NS5A-DI from genotype 1b of HCV. (a) SAXS pattern ($\circ$) of NS5A-DI from genotype 1b in 25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol buffer at 2.5 mg/mL (green) and 3.7 mg/mL (forest green) concentration. (*Inset*) Guinier plots show linearity at all concentrations used, indicating no aggregation. The scattering profiles are offset for clarity by applying arbitrary scale factors. (b) Overlapping of pair-distance distribution function $P(r)$ at 2.5 mg/mL (—; green) and 3.7 mg/mL (—; forest green) concentration show differences in the profile. (c) Normalized Kratky plot of NS5A-DI from genotype 1b at 2.5 mg/mL ($\bullet$; green) and 3.7 mg/mL ($\bullet$; forest green) concentration compared to the compact globular lysozyme ($\bullet$; grey) with a peak (---; grey), representing the theoretical peak and assuming an ideal Guinier region of a globular particle. The scattering pattern of NS5A-DI from genotype 1b exhibits a broad bell-shaped profile shifted toward right with respect to standard globular proteins, indicating the presence of motion and/or extended nature of the protein. (d) Theoretical scattering curve for the mixture of monomer-M, hexamer-H, and decamer-D of NS5A-DI from genotype 1b (—; red) calculated using OLIGOMER program fits to the experimental scattering pattern at

2.5 mg/mL (○; green) and 3.7 mg/mL (○; forest green) concentration with a χ^2 -value of 0.5.

It is critical to emphasize two main outcomes of the SAXS analysis. First, the larger particles observed in DLS and SEC analysis are indeed higher-order structures rather than being irregular aggregates. And, higher-order structures form by dimer subunits of NS5A-DI. Secondly, the dimeric form is virtually absent in free form. This suggests that NS5A-DI monomers form dimers in solution and these dimers do not remain freely for a long time in solution and turn into higher-order structures. At this point, one may wonder the reason behind the abundance of dimers in BMH crosslinking assay. One possible explanation of this duality might be the looseness of higher-order structures. We suspect that the reaction conditions that we perform BMH crosslinking might be destabilizing higher-order structures giving rise to their subunits, dimers. We are currently trying to further optimize the chemical crosslinking experiments.

Table 4-1: Data collection and X-ray scattering derived parameters for HCV NS5A-DI from genotypes 1b and 1a.

Data collection parameters				
Instrument (source & detector)	Bruker NanoStar equipped with MetalJet eXcillum X-ray source and VÅNTEC-2000 detector			
Beam geometry	100 μm slit			
Wavelength ($\text{\AA}$)	0.134			
q range ($\text{\AA}^{-1}$)	0.016 – 0.4			
Exposure time (min)	10 (10 frames x 1 min)			
Temperature (K)	288.15			
Protein sample	NS5A-DI (genotype 1b)		NS5A-DI (genotype 1a)	
Concentration (mg mL^{-1})	2.5	3.7	2.5	3.0
Structural parameters				
$I(0)$ (arbitrary units) [from P(r)]	18.98 ± 0.26	27.20 ± 0.25	9.51 ± 0.20	9.35 ± 0.18
R_g ($\text{\AA}$) [from P(r)]	39.38 ± 0.48	46.80 ± 0.34	30.38 ± 0.77	30.57 ± 0.77
$I(0)$ (arbitrary units) (from Guinier)	18.96 ± 0.49	27.76 ± 0.66	9.56 ± 0.24	9.40 ± 0.21
R_g ($\text{\AA}$) (from Guinier)	37.55 ± 1.35	46.10 ± 1.41	29.51 ± 1.11	29.55 ± 0.98
D_{max} ($\text{\AA}$)	127 ± 10	145 ± 10	106 ± 6	107 ± 6
Porod volume estimate (V_p) ($\text{\AA}^3$)	~ 102931	~ 157557	~ 52067	~ 52772
DAMMIF excluded volume (V_{ex}) ($\text{\AA}^3$)	~ 137400	~ 208100	~ 74558	~ 73721
Dry volume from sequence ($\text{\AA}^3$) ‡	~ 23003	~ 23003	~ 23466	~ 23466
Molecular mass determination				
Calculated monomeric MM (kDa) [from sequence*]	~ 19	~ 19	~ 19	~ 19
MM from Porod invariant ($V_p/1.6$) (kDa)	64 ± 6	98 ± 10	33 ± 4	34 ± 4
MM from excluded volume ($V_{ex}/2$) (kDa)	68 ± 7	104 ± 10	37 ± 4	37 ± 4
MM from volume of correlation (V_c) (kDa)	60 ± 6	97 ± 10	35 ± 3	35 ± 3
Oligomer	Mixture† (61:18:21)	Mixture† (41:17:42)	Dimer	Dimer
Software employed				
Primary data reduction	BRUKER SAS			
Data processing	PRIMUS			
<i>Ab initio</i> analysis	DAMMIF			
Validation and averaging	DAMAVR			
Computation of model intensities	CRY SOL			
Computation of volume fractions of mixtures	OLIGOMER			
3D graphics representations	PyMOL			

* http://web.expasy.org/compute_pi/

‡ <http://www.basic.northwestern.edu/biotools/proteincalc.html>

† Monomer: Hexamer: Decamer

4.5. Self-association behavior of NS5A-DI of genotype 1a differs from genotype 1b

Analysis to identify similarities and differences in the profile of resistance NS5A from genotype 1a and genotype 1b revealed that DCV was around 5 to 10 times less efficient against genotype 1a (Fridell et al., 2010). Moreover, these variations, the amino acid substitutions in the N-terminal region of NS5A, caused the resistant type of the virus against the drug. All these findings motivated us to investigate the differences between NS5A domain 1 of genotype 1a and genotype 1b. We have characterized NS5A domain 1 from genotype 1a in the same way as we performed for NS5A-DI from genotype 1b in the previous part by using SEC, DLS, chemical crosslinking, and SAXS methods. The NS5A-DI from genotype 1a was purified by using the same method with as NS5A-DI from genotype 1b. The SEC elution profile and its corresponding SDS-PAGE analysis for NS5A-DI from genotype 1a are shown in Figure 4-9.

Figure 4-9. Size-exclusion chromatography of NS5A-DI from genotype 1a of HCV. Superdex 200 10/300. The protein was eluted at 17.9 mL, and the inset figure shows the SDS-PAGE analysis of the eluted fraction proteins. Note that the protein product is quite pure and has a molecular weight around 19 kDa, as seen as a single band.

Next, the shaded area seen in Figure 4-9, corresponding to monomer state of the protein was collected and concentrated to 2.5 mg/mL for DLS analysis in the SEC buffer (25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol). DLS measurements of purified NS5A-DI from genotype 1a revealed a bimodal size distribution with peak maxima at around 10 and 165 nm, as seen in Figure 4-10. Although the particle sizes of NS5A-DI from genotype 1b and NS5A-DI from genotype 1a are not same, the smaller particles again dominate NS5A-DI from genotype 1a samples.

Figure 4-10. DLS analysis of the proteins. (a) Hydrodynamic diameter distribution of NS5A-DI from genotype 1a of HCV by DLS at 25 °C using 2.5 mg/mL of protein. Purified NS5A-DI revealed a bimodal distribution in 25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol. (b) The DLS analysis of the NS5A-DI from genotype 1a and 1b were overlayed.

To further investigate the self-association behavior of NS5A-DI from genotype 1a,

we reloaded the monomeric protein to SEC after collecting and concentrating to different concentrations, *i.e.* 52, 109 and 176 μ M. In this experiment, we again targeted to load the same amount of protein, ~500 μ g (Figure 4-11a). Figure 4-11b shows the normalized SEC profiles of the protein at different concentrations. So, the SEC data suggests that the exchange rates between different conformations are significantly different in NS5A-DI from genotype 1a and NS5A-DI from genotype 1b. While the higher-order structures are clearly observed in SEC profile of NS5A-DI from genotype 1b, such forms are almost undetectable for NS5A-DI from genotype 1a. (The overlapped SEC profiles given in Figure 4-12 highlight these differences). This crucial finding was surprising as the protein sequences are highly similar with a similarity ratio of around 91% as shown in Figure 4-1.

Figure 4-11. SEC analysis of the proteins. (a) Same amount of NS5A-DI from genotype 1a at indicated concentrations reloaded. (b) Normalized SEC elution profiles of the NS5A-DI from genotype 1b at different concentrations are shown.

Figure 4-12. The SEC analysis of the proteins NS5A-DI from genotype 1a and genotype 1b were overlayed. It is clearly seen that the NS5A-DI from genotype 1b shows shifts towards higher-orders (bigger size) while NS5A-DI from genotype 1a showed an altered profile on SEC column.

Lastly, we performed crosslinking experiments for NS5A-DI from genotype 1a, the SDS-PAGE analysis of irreversible crosslinking results is shown in Figure 4-13. Here, it is seen that monomer/dimer ratio is much bigger than we observed for NS5A-DI of genotype 1b.

Figure 4-13. BMH crosslinking of NS5A-DI from genotype 1a. The purified NS5A-DI from genotype 1a was crosslinked by bismaleimido-hexane (BMH), and the reaction was terminated by DTT. Irreversibly crosslinked samples were analyzed by 12% SDS-PAGE analysis. 10 µg of samples were loaded on the left gel-image, 10 µL of samples were loaded on the right image.

4.6. Low-resolution solution structure of NS5A-DI from genotype 1a

SAXS studies of NS5A-DI from genotype 1a of HCV was done in accordance with the same buffer of NS5A-DI from genotype 1b of HCV, SEC buffer (25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol). The scattering intensity profile of the protein (buffer subtracted) and the Guinier plot are shown in Figure 4-14a and Figure 4-14b. The SAXS-patterns were of good quality and the Guinier plots at low angles appeared linear (Figure 4-14a, inset) with the derived R_g -values of 29.51 ± 1.11 and 29.55 ± 0.98 Å for 2.5 and 3.0 mg/mL concentration data (Table 4-1). The $P(r)$ function are similar for both concentrations and showed a single peak and right-skewed (Figure 4-14b) with D_{max} of 106 ± 6 Å (Table 4-1), reflecting that NS5A-DI from genotype 1a is elongated. The normalized Kratky plot of NS5A-DI from genotype 1a had a broad bell-shaped profile reflecting that it is extended and less rigid in solution (Figure 4-14c). Molecular mass determined based on the Porod and excluded volumes and volume of correlation confirmed that NS5A-DI from genotype 1a exists as dimer in solution at the concentration used, Table 4-1.

Figure 4-14. SAXS analysis of NS5A-DI from genotype 1a of HCV. (a) SAXS pattern ($\circ$) of NS5A-DI from genotype 1a in 25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol buffer at 2.5 mg/mL (orange) and 3.0 mg/mL (red) concentration. (Inset) Guinier plots show linearity at all concentrations used, indicating no aggregation. The scattering profiles are offset for clarity by applying arbitrary scale factors. (b) Overlapping of pair-distance distribution function $P(r)$ at 2.5 mg/mL (—; orange) and 3.0 mg/mL (—; red) concentrations show similar profile. (c) Normalized Kratky plot of NS5A-DI from genotype 1a at 3.0 mg/mL ($\bullet$; red) concentration compared to the compact globular lysozyme ($\bullet$; grey) with a peak (---; grey), representing the theoretical peak and assuming an ideal Guinier region of a globular particle. The scattering pattern of NS5A-DI from genotype 1a exhibits a broad bell-shaped profile shifted toward right with respect to standard globular proteins, indicating the presence of motion and/or extended nature of the protein. (d) Theoretical scattering curve fit (—) of dimer 1 (cyan) and dimer 2 (blue)

to the experimental scattering pattern (○) of 2.5 mg/mL (orange) and 3.0 mg/mL (red) concentration. Dimer 2 fits better to the experimental data with χ^2 -value of 0.46 and 0.55 respectively to 2.5 mg/mL (orange) and 3.0 mg/mL (red) concentration data.

The crystal structure of NS5A domain 1 from genotype 1a (PDB ID: 4cl1) contains four molecules in the asymmetric unit which forms two types of dimers. In “dimer 1” the monomers are arranged in antiparallel head to tail configuration and in “dimer 2” it forms an N-terminal head to head configuration (Figure 4-15). A theoretical scattering curve computed using CRY SOL for both the dimers had a fit to the experimental scattering curve of $\chi^2 = 0.82$ for dimer 1 and 0.46 for dimer 2 for 2.5 mg/mL concentration data ($\chi^2 = 1.13$ for dimer 1 and 0.55 for dimer 2 for 3.0 mg/mL concentration data) (Figure 4-14d). The experimental data are in good agreement to the dimer 2 curve, suggesting the presence of N-terminal head to head dimers in solution. The low-resolution shapes of NS5A-D1-g1a were reconstructed *ab initio* using DAMMIF, and the final averaged and filtered model with a normalized spatial discrepancy (NSD) of 0.56 ± 0.04 is shown in Figure 4-16. The NS5A-D1-g1a solution model showed an elongated conformation and superimposed nicely to the crystallographic dimer 2 structure (Figure 4-16). The unoccupied density in the bead model around one of the monomers (represented by the arrow in Figure 4-16) suggests that the monomers have a highly dynamic behavior and are in rapid motion in solution due to the lower interface area ($311 \text{ \AA}^2$) of the dimer. The SAXS analysis of NS5A-D1-g1a confirmed the presence of an N-terminal head to head dimer structure in solution.

a) Tetramer of NS5A domain I from genotype 1a of HCV (PDB ID: 4Cl1)

Figure 4-15. The tetramer crystal structure of NS5A-DI from genotype 1a of HCV., (PDB ID: 4cl1). The dimers of the tetramer named as “dimer 1” (blue and green monomers) and “dimer 2” (yellow and orange monomers)

Figure 4-16. NS5A-DI from genotype 1a of HCV exist as a dimer in solution The averaged and filtered envelope (grey) from twenty independent ab initio reconstructions using DAMMIF superimposed onto the cartoon representation of the crystallographic dimer 2 structure of NS5A-DI from genotype 1a (PDB ID: 4cl1). The two subunits of the dimer are colored green and cyan. The red arrow indicates unoccupied density in the solution shape.

The Table 4-2 summarized the overall statistics of the SAXS analysis of the protein. Rg values from Guinier approximation were $29.51 \pm 1.11 \text{ \AA}$ and $29.55 \pm 0.98 \text{ \AA}$ at 2.5

and 3.0 mg/mL, accordingly. Maximum particle dimensions were 106 and 107 Å at 2.5 and 3.0 mg/mL, respectively. The CRY SOL program confirmed the existence of dimer2 (100%) at both concentrations.

Table 4-2. The statistics of the SAXS analysis of the NS5A-DI protein from genotype 1a of HCV. The overall statistics: radius of gyration values from Guinier approximation and P(r) function, maximum particle dimensions and analysis of OLIGOMER program of the protein at 2.0 and 3.0 mg/mL are summarized.

Conc. (mg/ml)	Rg(Å)		Dmax (Å)	MW (kDa)	Oligomer
	Guinier	Gnom			
2.5	29.51 ± 1.11	30.38 ± 0.48	106	31	Dimer 2
3.0	29.55 ± 0.98	30.57 ± 0.34	107	31	Dimer 2

The comparison of DLS, SEC, crosslinking assay, and SAXS results of NS5A-DI of genotype 1a and 1b clearly showed that these two proteins behave differently in solution. The most striking difference is the presence of higher-order structures in genotype 1b whereas genotype 1a exists merely in a monomer-dimer equilibrium. Such a difference in the dynamic equilibrium of these two proteins might be the reason behind the observation of their different drug responses.

4.7. The effect of daclatasvir and RNA on NS5A-DI

Neither NS5A full-length protein nor NS5A-DI has any enzymatic function that can be tested *in vitro* (Mankouri et al., 2008). However, the DI is known to be interacting with the RNA and DCV (Gao et al., 2010; Hwang et al., 2010). To test whether the oligomeric state of NS5A-DI from genotype 1b is altered in the presence of RNA and DCV, we performed DLS, BMH crosslinking and SAXS studies in the presence of 15U RNA and daclatasvir dihydrochloride (DCV-HCl). DCV-HCl is the water-soluble version of the DCV. DLS analysis of the protein alone at 105 µM (green line), 105 µM of protein incubated with 210 µM of 15U RNA (orange line) and 105 µM of protein incubated with 105 µM of DCV-HCl (red line) are shown in Figure 4-17a.

DLS profile of the protein incubated with RNA slightly shifted towards to smaller particles compared to the apo protein. The DLS profile of the DCV-HCl-treated protein immediately formed quite large particles. We crosslinked the samples by BMH and run SDS-PAGE, as shown in Figure 4-17b. We analyzed the following five samples, protein alone, the protein with DCV-HCl (ratios tested 1:1, 1:0.5, 1:0.1), and protein with RNA (ratio of 1:2). Unfortunately, there were no significant differences between apo protein, the protein incubated with DCV-HCl, and the protein incubated with RNA from BMH crosslinking experiments that could explain DLS observations.

Figure 4-17. The effect of RNA and DCV-HCl on self-association of NS5A-DI of genotype 1b. (a) The particle size distributions in the presence and absence of RNA/DCV. (b) SDS-PAGE analysis of BMH crosslinked protein in the presence and absence of RNA/DCV-HCl.

Figure 4-18. SAXS studies of NS5A-DI of genotype 1b of HCV with 15U RNA and DCV-HCl. (a) SAXS profiles of NS5A-DI apo protein (green), protein plus DCV (red), and protein plus RNA (orange) in 25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol. The protein plus DCV condition shows aggregates. (b) The Guinier plots for apo protein and protein plus RNA at low angles are linear, showing no aggregation. (c) Distance distribution functions of apo protein at 2 mg/mL (green) and protein plus RNA (orange). (d) Analysis of apo protein and protein plus RNA by OLIGOMER program (M (monomer), H: hexamer, D: decamer), M: H: D ratios are indicated.

We proceeded with SAXS studies to explain the phenomenon that observed by the DLS analysis. The three SAXS patterns of the protein at 2 mg/mL apo, with DCV-HCl (1:1) and with RNA (1:2) were collected. The scattering profiles of those samples are represented in Figure 4-18a. Low-angle scattering patterns of the protein with DCV-HCl suggested the formation of the aggregates, which provides a crucial insight towards understanding our previous observations regarding the behavior of protein when treated with DCV-HCl. As seen from SAXS data, DCV-HCl triggers the formation of protein aggregates. We believe for this reason we previously observed a clear shift to larger sizes in the DLS profile where no difference seen in the BMH crosslinking assay. The Guinier plots of the apo protein, protein exposed to RNA are linear, indicating no aggregation for both samples (Figure 4-18b). The distance distribution function of the samples correlated with our previous studies skewed right distribution, and profile of the apo protein is slightly broadened compared to RNA added the sample (Figure 4-18c). Moreover, OLIGOMER analysis suggests that monomer: hexamer: decamer ratio is 42:27:31 when protein is alone, and it is 51:37:12 respectively when RNA is added to the protein sample as shown in Figure 4-18d. SAXS data confirm the observation of the DLS profile that was observed in the protein with RNA sample. Our findings would suggest that the apo protein exists in an expanded conformation and it might be more compact upon association of RNA with protein.

4.8. Overview of chapter

This study was carried out to characterize the NS5A-DI of HCV in solution mainly to compare NS5A-DI of genotypes 1a and 1b. They revealed different dimers by crystallization studies and show different drug responses in the literature. (Please refer to Introduction Section for detailed information.) In this part, we made efforts to depict the in-solution behavior of the proteins by combining several complementary techniques. To do so, we combined DLS, SEC, and SAXS studies.

Here, we show that NS5A-DI of genotype 1b has an intrinsic behavior to oligomerize in solution at physiological conditions and it exists in an equilibrium of monomer/dimer and higher-order oligomers. NS5A-DI of genotype 1a also exists as a heterogeneous mixture as well, but its equilibrium state is significantly different from that of NS5A-DI

of genotype 1b. DLS, BMH crosslinking and SEC analysis give us a strong body of evidence regarding the differences between these two proteins. The protein from genotype 1a tends to exist at smaller sizes more like as monomers and dimers in equilibrium. In the case of the protein from genotype 1b, it can form higher-order structures (multimers), as seen in the apparent shift in SEC results.

SAXS model suggests that the higher-order oligomers of NS5A-DI of genotype 1b are formed of dimers of the crystal structure, 1zh1, structure and have an elongated helical shape. The SAXS model of NS5A-DI of genotype 1a confirmed the existence of the N-terminal head-to-head dimer conformation of the crystal structure, 4cl1. We also tested the effect of DCV-HCl and the 15U RNA on the equilibrium of NS5A-DI from genotype 1b. The DCV-HCl triggered the aggregate formation of the protein, whereas RNA resulted in a more compact conformation of the protein.

Chapter 5: MOLECULAR CHARACTERIZATION of NS5A-DI with AMPHIPATHIC HELIX in SUPPORTED LIPID BILAYERS

5.1. Motivation

The replication of HCV, like all (+) strand RNA viruses, takes place in distinct membranous compartments of the infected cells (Chatel-Chaix & Bartenschlager, 2014). Mainly, for HCV-infected cells, double- and multi-membrane vesicles accumulate, and these meshwork structures are called as “membranous web” (MW), which serve as the replication factory of the virus (Romero-Brey et al., 2012). The first systematic study to investigate the individual proteins of HCV those contribute to the ER-derived membrane rearrangements reported that the NS5A-DI and the helicase domain of NS3 were the determinants for the biogenesis of MWs (Romero-Brey et al., 2015). More recent evidence highlights that HCV-induced membranous rearrangements are blocked in DCV-treated infected cells, whereas in the DCV-resistant mutant of NS5A, they are still observable (Berger et al., 2014). Taken all together, these findings motivated us to perform further investigations regarding the NS5A-DI on biomimicry supported bilayer (SBL) platforms. By doing so, we aimed to characterize the behavior of protein on membranous environments alone and in the presence of DCV to better understand the mode of action (MoA) of DCV.

5.2. Cloning, expression, solubilisation, and purification of NS5A-DI including amphipathic helix (AH)

NS5A-DI is a membrane-associating protein where it exists in a lipid environment of cytoplasmic membranous structures. Therefore, purification and handling of this protein require stringently optimized conditions, compared to water-soluble proteins. In this part of the study, we will present our strategies to obtain soluble NS5A-DI with AH.

5.2.1. Cloning strategy of NS5A-DI with amphipathic helix

We cloned NS5A(2-198) gene from the template of HCV from genotype 1a (H77) into pET29 b (Kan) plasmid by using primers; 1a HCV NS5A f 2 and 1a HCV NS5A

r198, the sequences of the primers are shown in Table 2-2 . His-tag was placed at the C-terminal end of the protein to specifically reduce the possibility of any unwanted artificial changes in the membrane anchoring behavior of the protein. The PCR amplicon and detailed the illustration of recombinant protein, *i.e.* NS5A(2-198)-His-tag construct are shown in Figure 5-1a and b. DNA sequencing confirmed the correct position of recombinant protein without any mutations.

Figure 5-1. Cloning of NS5A(2-198)-His-tag. (a) The PCR product corresponding NS5A(2-198) which is 621 bp is shown. B) The representation of C-terminal tagged protein, NS5A(2-198)-His-tag. (c) The sequence of the amphipathic helix is shown. Red color, acidic residues, blue color, basic residues, green color hydrophobic residues. The DCV-resistance amino acids on AH are shown, M28 and L31.

5.2.2. Expression screening of HCV NS5A-DI with Amphipathic Helix (AH)

There is no study of NS5A-DI purification protocol that includes AH region from a bacterial source of expression systems in the literature, and all the structural studies have been reported with the constructs lacking N-terminal AH (Lambert et al., 2014; Love et al., 2009; Tellinghuisen et al., 2005). However, very recently, the full-length NS5A was successfully purified from an insect cell expression system (Kwon et al., 2015).

We tested two different strains, BL21 (DE3) and C41 (DE3) of *E. coli* for the expression of NS5A(2-198)-His-tag. Once the construct was transformed into BL21 and C41, the protein expression was induced with 1 mM IPTG at 37 °C. 12% SDS-PAGE analysis of uninduced and induced cell lysate samples are shown in Figure 5-2.

The theoretical molecular weight of the NS5A(2-198)-His-tag protein is 23.6 kDa, and it appeared at around ~25 kDa, slightly higher than the expected size as shown in Figure 5-2. It is seen that both strains of *E. coli* showed modest expression of the recombinant protein and the longer period of induction (*i.e.* overnight) resulted in less amount of the protein expression for both strains. And in fact, it almost disappeared in BL21 strain. Overexpression of the recombinant integral membrane proteins is grueling due to their physical properties as their association with cellular membranes might cause toxic effects on expression systems that result in low-level expressions and inclusion bodies. As the C41 (DE3) cells are the strains that have a higher tolerance for toxic proteins, we choose the C41 for the overexpression of the NS5A(2-198)-His-tag.

Figure 5-2. The protein expression control of NS5A(2-198)-His-tag. 12% SDS-PAGE analysis of the cell lysate samples are shown. Lane 1: marker, lane 2: uninduced at 37 °C, C41 strain lane 3: 1 mM IPTG, at 37 °C for 2 hours, C41 strain, lane 4: 1 mM IPTG, at 37 °C for 5 hours, C41 strain, lane 5: 1 mM IPTG, at 37 °C for O/N, C41 strain, lane 6: Uninduced at 37 °C, BL21 strain, lane 7: 1 mM IPTG, at 37 °C for 2 hours, BL21 strain, lane 8: 1 mM IPTG, at 37 °C for 5 hours, BL21 strain, lane 9: 1 mM IPTG, at 37 °C for O/N, BL21 strain. The recombinant protein appeared at around 25 kDa, slightly higher than the expected size of the protein.

5.2.3. Detergent screening and solubility check of HCV 1a NS5A(2-198)-His-tag

Integral membrane proteins exist in cellular membranes as surrounded by lipid

molecules. However, they are required to be dispersed in aqueous solutions, which requires a special effort to find the suitable conditions. Insolubility problem is circumvented by adding a detergent into membrane suspension buffer that disrupts cellular membranes and solubilizes the membrane protein with\without lipids and they prevent the formation of the inclusion bodies. In contrast, peripheral membrane proteins can be dissociated from the membranes by weakening the ionic interactions between membrane and protein. For instance, high salt concentration like 1 M NaCl can be used for this purpose. So that, before we performed large-scale protein expression, we screened buffers with detergents and high salt to obtain the protein, NS5A(2-198)-His-tag, free from membranes.

We induced protein expression at 30 °C for 6 hours with the C41 strain of *E. coli*. Then, we lysed the cells with the lysis buffer (50 mM potassium phosphate, pH 7.0, 300 mM NaCl, 2 mM MgCl₂, supplemented with 1mM PMSF). Next, we extracted the cellular membrane fractions as a pellet by using ultracentrifugation method by following the work-flow seen Figure 2-1. Isolated cellular membranes were suspended by using eight different membrane resuspension buffers that include seven different detergents and one buffer alone (without any detergent but at high salt concentration, 1 M NaCl) as a negative control and also to check if it was a peripheral membrane protein. In total five nonionic (*Triton-100*, *Brij 35*, *Tween-20*, *n-octyl β-D-maltoside* and *NP-40*), one zwitterionic (*CHAPS*) and one ionic (*SDS*) detergents were screened, as listed in Table 5-1. *SDS* was used a positive control since it solubilizes many membrane proteins most likely by denaturing them.

Table 5-1. Properties of the detergents screened for solubilization of NS5A(2-198) membrane protein

Detergent name	CMC (mM)	Ave. Mw	Agg. #	Class of detergent
C12E8 Octaethylene glycol monododecyl ether	0.11	66.000	123	Non-ionic
Triton-100	0.2-0.9	625-80.000	100-155	Non-ionic
Brij 35	0.09	1125	40	Non-ionic
Tween-20	0.06	1228	n.a.	Non-ionic
CHAPS	6.5	615	n.a.	Zwitterionic
n-octyl β -D-maltoside	23.4	38.000	84	Non-ionic
NP-40	0.02	640		Non-ionic
SDS	7-10	18.000	62	Ionic

Next, membrane fractions were solubilized by stirring at 4 °C for 3 hours and overnight. Then, the solubilized proteins were recovered by ultracentrifugation (100.000 g for 1 hour) from membrane suspension. The supernatants were analyzed by dot (western) blot method to detect NS5A(2-198)-His-tag by using anti-His-tag antibody. The dot-blotting results are shown in Figure 5-3; numbers represent detergents of the membrane suspension buffers. The last sample (8) is the negative control, in which no detergent but high salt was added. We did not detect any protein in the supernatant when there was no detergent in the membrane suspension buffer, as shown in Figure 5-3. Moreover, high salt concentration also did not help to dissociate the protein from membranes, which suggests that NS5A(2-198)-His-tag is not a peripheral membrane protein and indeed behaves like an integral membrane protein, tightly associated with the membrane. This lends support to the previous findings (Brass et al., 2002). Among the detergents, *Brij-35*, labeled as “2” and *tween-20*, marked as “4”, failed to solubilize the target protein. However, *Triton X-100*, *CHAPS*, *n-octyl β -D-maltoside* and *NP-40* successfully solubilized the NS5A(2-198). We chose *Triton X-100* as the solubilizing agent for further steps as it is nonionic, which means that it does not denature the protein.

Moreover, it is also cheap compared to the *CHAPS*, *n-octyl β -D-maltoside*. Lastly, its critical micelle concentration (CMC) is low, so that the detergent amount will be less in the buffers, which is essential for the functional studies on the membrane-mimicking environments. Therefore, we picked up *Triton X-100* for further studies. Table 5-1 shows the properties of the detergents tested in this study.

Figure 5-3. Dot blot analysis for detergent screening of the solubilisation of NS5A(2-198)-His-tag. Membrane fractions were solubilized at 4 °C for 3 hours (a) and overnight (b) in different detergents 1) Triton X-100, 2) Brij-35, 3) CHAPS, 4) Tween-20, 5) *n-octyl β -D-maltoside*, 6) NP-40, 7) SDS (positive control), 8) buffer alone (no detergent) with 1 M of NaCl (negative control).

5.2.4. Purification of NS5A(2-198)-His-tag by imidazole gradient IMAC

After the overexpression and solubilisation of the NS5A(2-198)-His-tag protein was optimized, the solubilized protein was loaded into Ni-NTA column and purified in a similar way with previous proteins. After our several attempts (the column was washed with large volumes of 50 mM imidazole with 1 M of NaCl salt), the elution fraction was contaminated with unwanted proteins and the yield achieved was quite low.

We utilized HisTrap FF column as a neater and more practical solution to solve this problem. We performed the protein purification by HisTrap FF column on Bio-Rad system with an imidazole gradient washing, from 25 to 500 mM. The protein fractions during imidazole gradients washing steps were collected and analyzed on 12% SDS-PAGE (Figure 5-4). Most the target proteins coeluted with impurities until ~330 mM of imidazole, and we could collect the fractions eluted after 330 mM of imidazole. Purification of the protein with His-Trap FF column optimized our imidazole-washing step precisely so that we could get pure proteins for next steps.

Figure 5-4. Imidazole gradient purification of NS5A(2-198)-His-tag. (a) Chromatogram showing the elution profile of NS5A(2-198)-His-tag on HisTrap FF column. Imidazole gradient, *i.e.* 25 to 500 mM is shown with red coloring and the absorbance at 280 nm is shown as blue color. (b) 12% SDS-PAGE analysis of the elution fractions from imidazole gradient washing are shown. The fractions eluted after 330 mM imidazole were collected for further experiments.

5.2.5. Size-exclusion chromatography and secondary structure analysis of NS5A(2-198)-His-tag

Imidazole gradient purification resulted in highly pure protein, as shown by the SDS-PAGE analysis above. Although the protein product seems highly pure, the homogeneity of the proteins should be tested by using gel filtration chromatography for detergent-solubilized proteins as the detergents might cause the formation of heterogeneous mixtures. We loaded NS5A(2-198)-His-tag protein to the Superdex S200 column, and the protein eluted at around 12.6 mL, seen in Figure 5-5a. Please note that the expected elution volume of a protein with size of 23.6 kDa would be around 17 mL. As this shape of protein corresponds to a detergent-solubilized protein, it is surrounded

by detergent molecules. Therefore, it is very typical that we observed the elution volume lower than expected. Interestingly, there is another peak around 11 mL which suggests that this protein might also be self-associating similar to the behavior of NS5A-DI. The circular dichroism (CD) spectra of the protein, measured the ellipticity of the protein within the wavelength range of 190-250 nm is shown in Figure 5-5b. BestSel tool (Micsonai et al., 2015) was used to predict the secondary structure of the protein. The outcome of this analysis is available in Figure 5-5c. The protein contains 24.7% helix, 66.2% beta sheet, and 9.1% turns.

Figure 5-5. The homogeneity and folding check of NS5A(2-198)-His-tag. (a)The elution profile of NS5A(2-198)-His-tag on Superdex S200 (10/300GL). The protein eluted at around 12.5 mL. (b) The CD spectra of the protein. (c)The secondary structure content analysis of the protein by using BestSel tool (Micsonai et al., 2015).

5.3. The interactions of simulated membranes and NS5A-DI with AH: Towards understanding the mechanism of action of a direct-acting antiviral, daclatasvir

SBL platforms contain two-dimensional membrane mimicking surfaces that can be studied by surface-sensitive techniques such as acoustic or optical biosensors (Jackman & Cho, 2012; Sackmann, 1996). Solvent-assisted lipid bilayer (SALB) formation method is a simple and adaptable approach *i.e.* lipid depositions in the organic solvents and subsequently exchange aqueous buffers to fabricate reproducible bilayers for non-experts (Tabaei, Choi, Haw Zan, Zhdanov, & Cho, 2014). The chemical composition of the bilayers such as electrical surface charge, membrane fluidity, and lipid content, also its surrounding buffers can be optimized. By combining with a label-free biosensor, QCM-D, SBLs serve as an outstanding tool to track molecular interactions.

In order to investigate NS5A(2-198)-His-tag on the biomimicked membranes, we employed a surface sensitive technique, quartz microbalance with dissipation (QCM-D). QCM-D monitors the changes in the resonance frequency (Δf) and energy dissipations (Δd) of the surface, which are correlated to the mass and viscoelastic properties of the materials/macromolecules adsorbed to the surface (Reviakine et al., 2005). Utilizing this relatively rapid and very sensitive method, we explored NS5A(2-198)-His-tag protein on SBL platforms quantitatively, in real-time.

5.3.1. The formation of charged supported bilayer platform by solvent-assisted lipid bilayer (SALB) method

We designed planar bilayer membrane mimicking platforms with a variety of surface charges by altering the lipid compositions to study the effect of membrane electrostatic surface charge on the interactions of NS5A(2-198)-His-tag with membranes. To test this, three different synthetic lipid molecules, *i.e.* DOPC (zwitterionic), DOEPC (cationic), and POPG (anionic) mixed to obtain positively-, negatively-, and neutral-charged membranes. To form negatively- and positively-charged SBL platforms, 30% molar ratio of POPG or DOEPC lipids were mixed with 70% molar ratio of DOPC accordingly. For neutral membranes, 100% of DOPC lipid were used. The representative image of the charged bilayer platforms by SALB method on SiO₂ are shown in Figure 5-6a.

Figure 5-6b shows the typical QCM-D profile for the SBL formation on SiO₂. At

time zero, step 0 (S0), the aqueous buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl) injected to collect the baseline measurement for the frequency and energy dissipation signals of QCM-D. After the baseline signal gets stabilized, isopropanol was run (S1). The notable signal changes were observed in Δf and Δd as isopropanol, and the aqueous buffer has different density and viscosity properties. The freshly prepared lipid mixtures were injected at S2, and the isopropanol solution was replaced by aqueous solution at S3 to establish the formation of a complete planar lipid bilayer. The complete bilayer was reported to have a characteristic profile of ~ 26 Hz (Tabaei et al., 2014; Zan, Jackman, & Cho, 2014). The final changes in frequency and dissipation values checked and found to be correlated with the literature. Sauerbrey equation was used to correlate the frequency changes with the mass of deposited/dissolved material on the rigid films. The formation of charged lipid bilayers is summarized as accumulated mass in Figure 5-6.

Figure 5-6. The design of supported lipid bilayers membranes using charged and uncharged lipid molecules. (a) SBLs formed by SALB method. (The illustrations from left to right are as follows: DOPC: POPG (70:30), DOPC, DOPC: DOEPC (70:30)). (b) QCM-D measurement of the SBL. Changes in the frequency (Δf) and energy dissipation (Δd) of the overtone, $n=5$, were monitored as a function of time. Note that (S0) refers to the injection of Tris buffer; (S1) refers to the injection of isopropanol; (S2) refers to the loading of 0.5 mg/mL of lipid mixtures dissolved in isopropanol; (S3) refers to the exchange of isopropanol to Tris buffer. (c) The lipid mass surface density was calculated by the Sauerbrey equation, and formation of SBLs are summarized as deposited mass.

5.3.2. The effects of buffers on the association of membranes and NS5A(2-198)

Once the formation of SBLs are completed, we investigated the effect of the buffer that NS5A(2-198)-His-tag was stored, *i.e.* 20 mM HEPES pH 6.3, 300 mM NaCl, 5% glycerol, 0.015% Triton X-100. SBLs were formed as explained in Section 5.3.1. The steps after the formation of SBLs is shown in Figure 5-7. (Note that the data were normalized, which means Δf equalized to zero after completion of all charged bilayers, shown in S0.). Then the protein (2 μ M) in 20 mM HEPES pH 6.3, 300 mM NaCl, 5% glycerol, 0.015% Triton X-100 was added on the lipid bilayers to observe the changes

in the frequency and dissipations shown as S1 and washing with Tris buffer is shown as S2. The notable decrease in the frequency at S1 indicates the adsorption of the protein and perhaps the buffer components like detergent micelles by lipid bilayers. At S2, washing with Tris buffer resulted in the frequency changes around 25 Hz, which means that everything on the sensor was washed away, including the lipid molecules. This study showed us that the detergent, Triton X-100 at 0.015%, has a significant effect on the SBL as it destroys the bilayer formation. By using this knowledge, we put more efforts to reduce the amount of Triton X-100 and glycerol in the binding buffer for further experiments.

Figure 5-7. The effect of buffers on supported lipid bilayer platforms. QCM-D measurement of the SBL. The changes in the frequency (Δf) and energy dissipation (Δd) of the overtone, $n=5$, were monitored as a function of time. Note that (S0) refers to the loading of 0.5 mg/mL of particular lipid mixtures dissolved in isopropanol; and exchange of isopropanol to Tris buffer; (S1) refers to the injection of NS5A(2-198)-His-tag with the storage buffer of protein (*i.e.* 20 mM HEPES pH 6.3, 300 mM NaCl, 5% Glycerol, 0.01% Triton X-100); (S2) refers to the washing of the surfaces with Tris buffer.

As Triton X-100 and glycerol caused significant changes in the frequency and dissipation measurements, the binding buffer was replaced with the Tris buffer, which was 10 mM Tris-HCl pH 7.5, 150 mM NaCl. (Please note that this was the buffer used to form the bilayers.). The protein was diluted in binding buffer and injected on lipid bilayers. The representative QCM-D measurement is shown in Figure 5-8a. S0 shows the time where bilayers were completed and the frequency is equalized to zero. The protein is diluted to 2 μ M and injected onto the cationic (DOPC:DOEPC, 70:30), anionic (POPG:DOPC, 70:30), and zwitterionic (DOPC) bilayers at S1. And, the surface was washed with the binding buffer at S2.

Figure 5-8. Associations of NS5A(2-198)-His-tag with supported bilayer platform. Frequency and dissipation changes of the cationic (DOPC:DOEPC, 70:30), anionic (POPG:DOPC, 70:30) and zwitterionic (DOPC) bilayers were recorded by QCM-D. The bilayers were formed, and the frequency equalized to zero at S0. NS5A(2-198)-His-tag protein stock (30 μ M in 25 mM HEPES, 300 mM NaCl, 0.015% *TritonX-100*, 5% glycerol 6.3). was diluted to 1.5 μ M in binding buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl) and added to bilayers at S1. The surfaces were washed with 10 mM Tris-HCl pH 7.5, 150 mM NaCl at S2.

The changes in the frequency and dissipation on the negatively-charged bilayer were “ 21.45 ± 2.23 ” and “ 1.42 ± 0.32 ”, respectively after the injection of protein. They were “ 15.23 ± 2.07 ” and “ 1.32 ± 0.45 ” for zwitterionic lipid bilayer, suggesting that NS5A(2-198)-His-tag was adsorbed by mostly negatively-charged and neutral lipid bilayers. However, the frequency and dissipation changes of the positively-charged bilayer were “ 0.99 ± 0.34 ” and “ 0.99 ± 0.27 ”, suggesting almost no binding of the protein to the surface. The overall statistics of the changes in frequency, dissipation, and theoretical mass adsorbed by surfaces were shown in Table 5-2. Herein, we demonstrated that NS5A(2-198)-His-tag showed the highest affinity for negatively-charged bilayer, *i.e.* DOPC:POPG (70:30). Therefore, we conducted our further investigations on negatively-charged SBL platform.

Table 5-2. Statistics of QCM-D measurements on charged supported bilayers. The average and standard deviation values of the changes in the frequencies, dissipation, and theoretical adsorbed mass calculated by Sauerbrey equation values of the three independent experiments are shown.

SLB platform	ΔF (Hz)	ΔD (10^{-6})	Sauerbrey mass (ng cm^{-2})
DOPC	15.23 ± 2.07	1.32 ± 0.45	269.571 ± 36.639
DOPC:DOEPC (70:30)	0.99 ± 0.34	0.99 ± 0.27	17.523 ± 4.779
DOPC:POPG (70:30)	21.45 ± 2.23	1.40 ± 0.32	379.665 ± 39.471

5.3.3. The effect of daclatasvir dihydrochloride on supported bilayer platforms

Daclatasvir dihydrochloride (DCV-HCl) is the water-soluble form of the drug DCV. And, when it is solubilized in water, it keeps the pH around 5.0. The solubility of DCV-HCl highly depends on the pH of the buffer. If pH increases, the solubility of the DCV-HCl should decrease and a visible aggregate of the DCV should form, based on our observations and the reported technical sheets of the product.

To investigate the interactions of DCV-HCl with the lipid bilayer, we dissolved DCV-HCl in a binding buffer, 10 mM Tris-HCl pH 5.0, 50 mM NaCl. Herein, we did two alterations in the binding buffer. First, we changed the pH from 7.5 to 5.0 to improve the solubility of the DCV-HCl, and we reduced the NaCl concentration as the binding of the protein to the lipid bilayers decreases at higher salt concentration. Figure 5-9a is the illustration of the steps seen in the QCM-D measurements. And, Figure 5-9b shows the frequency and dissipation changes in response to the DCV-HCl injection on the negatively-charged SBL. S1 shows the addition time of variety of DCV concentrations (*i.e.* 0.25, 0.5, and 0.75 mg/mL) onto the lipid bilayer. It is seen that DCV-HCl did not adsorb on membrane, as only a quite negligible change took place.

Figure 5-9. Investigations of the interactions between DCV-HCl and charged lipid bilayer platforms. (a) The schematic representation of the steps in the measurement. (b) Frequency and dissipation changes of the anionic (POPG: DOPC, 70:30) bilayer were recorded, at overtone 5. The bilayers (all negatively-charged) were formed and the frequency equalized to zero at S0. (b) The DCV-HCl at different concentrations, *i.e.* 0.25, 0.5, and 0.75 mg/mL were added at S1. The surfaces were washed with 10 mM Tris-HCl pH 5.0, 50 mM NaCl at S2. Note that the changes in the frequency and dissipation were negligible.

5.3.4. The effect of DCV-HCl on the association of NS5A(2-198)-His-tag to the membranes

After showing the interaction of NS5A(2-198)-His-tag with negatively-charged bilayer, as seen in Figure 5-8, we explored the MoA of DCV-HCl on SBL platforms by focusing on the impact of the DCV-HCl on the association of NS5A with membranes. For this purpose, we employed two strategies to determine the effect of DCV-HCl. Our first strategy is illustrated in Figure 5-10a. First, we mixed NS5A(2-198)-His-tag (3 μ M) with different concentrations of DCV-HCl, and then DCV-HCl incubated protein was added to SBL. We designed this approach to understand the effect of DCV-HCl on the binding of the protein to membranes that is crucial for the biogenesis of the replication sites. In our second strategy, illustrated in Figure 5-11a, we added the protein on SBL first, and then we injected the DCV-HCl on the membrane-associated proteins. This

strategy aimed to see the effect of DCV-HCl on the replication sites already formed.

Figure 5-10b shows an exemplary QCM-D measurement of our first approach. The frequency was equalized to zero after the formation negatively charged bilayers completed in 10 mM Tris-HCl pH 7.5, 150 mM NaCl at S0. At S1, the buffer was replaced with binding buffer 10 mM Tris-HCl pH 5.0, 50 mM NaCl. Then, the proteins incubated with DCV-HCl were added on the membrane. The protein stock (in 12.7 μ M in 20 mM HEPES pH 6.3, 300 mM NaCl, 5% glycerol, 0.015% Triton X-100) was diluted to 3 μ M in the binding buffer and incubated with different concentrations of DCV-HCl at room temperature for 30 minutes on the shaker. Accordingly, 0 μ M, 3 μ M, 15 μ M, and 30 μ M of DCV-HCl concentrations were tested, and the samples were labeled as NS5A, NS5A + DCV-HCl (1:1), NS5A + DCV-HCl (1:5), NS5A + DCV-HCl (1:10) in Figure 5-10. The frequency change at t=60 minute was used to calculate the mass adsorbed on membrane and Sauerbrey mass was calculated. The average and standard deviations of Sauerbrey mass from three independent measurements are shown in Figure 5-10c. The binding of protein alone at t= 60 minutes was normalized as “100” , and the normalized Sauerbrey mass and concentrations of DCV-HCl were plotted in Figure 5-10d. We observed 15% less binding of proteins when they were incubated with DCV-HCL compared to apo version. This analysis suggested that DCV-HCl might be inhibiting the binding of NS5A(2-198)-His-tag to the membrane. However, it can only inhibit 15% of the total protein. More importantly, increasing the concentrations of the DCV-HCl (1:1, 1:5, 1:10 protein: DCV-HCl ratio) did not affect binding of the protein in a drug concentration dependent manner. These experiments gave us a clue that maybe DCV-HCL is associating with a subpopulation NS5A protein. Thus, for future experiments, we would like to test our hypothesis again and we would like to change the DCV-HCl concentrations, as well. Maybe if we lower the drug concentrations such as - 1: 0.1, 1:0.5 and 1:1 (protein/DCV-HCl ratio), we could observe an inhibition in the binding to the membranes in a DCV-HCl dependent manner.

Figure 5-10. The effect of DCV-HCl on the membrane association of NS5A(2-198) protein (a) The schematic representation of the measurement. (b) Frequency and dissipation changes of the negatively charged bilayers recorded at overtone 5. The bilayers were formed, and the frequency equalized to zero at S0. The buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl) was exchanged to the binding buffer (10 mM Tris-HCl pH 5.0, 50 mM NaCl) at S1. NS5A(2-198)-His-tag protein stock (30 μ M in 25 mM HEPES, 300 mM NaCl pH 6.3). was diluted to 3 μ M in binding buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl). The diluted protein was incubated with 0, 3, 15 or 30 μ M of DCV. (The samples represented NS5A, NS5A + DCV-HCl (1:1), NS5A + DCV-HCl (1:5), NS5A + DCV-HCl (1:10).) The samples were added to bilayers at S2. The surfaces were washed with 10 mM Tris-HCl, pH 5.0, 50 mM NaCl at S3. (c) The average and standard deviation values of the changes in the frequencies, dissipation, and theoretical adsorbed mass calculated by Sauerbrey equation from three independent experiments are shown. (d) The normalized Sauerbrey mass versus DCV-HCl concentrations is plotted.

Our second strategy, to investigate the impact of DCV-HCl on membrane-associated protein is shown in Figure 5-11a. The QCM-D measurement of this strategy is shown in Figure 5-11b, as the previous experiment, S0 shows the normalized frequency after formation of negatively-charged bilayers were completed, S1 is the buffer exchange step. The stock protein was diluted to 3 μ M in the binding buffer; they were injected to lipid bilayers at S2. The surface was washed with the binding buffer to remove any unbound proteins at S3. Buffer alone and DCV-HCl at 0.2, 0.4, and 0.6 mg/ mL were injected to lipid bilayers and finally the surface was again washed with binding buffer to remove unbound molecules at S5. This experiment provided us critical clues for the association of NS5A with DCV-HCl. We can see that DCV-HCl cannot disturb the associations of NS5A with the membrane. Please note that the concentration of DCV-HCl was high here. As the concentrations of DCV-HCl is very high and this observation cannot explain the mode of action of the drug. We performed this experiments with the lower concentrations of DCV-HCl however we have not seen any change. Therefore, we increase the concentration of the drug so high to get some clue for this investigation.

Figure 5-11. The effect daclatasvir on the protein associated with the lipid bilayer. (a) The schematic representation of the measurement. (b) The frequency and dissipation changes of the negatively charged bilayers recorded at over tone 5. The bilayers were formed, and the frequency equalized to zero at S0. The buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl) was exchanged to the binding buffer (10 mM Tris-HCl pH 5.0, 50 mM NaCl) at S1. NS5A(2-198)-His-tag protein stock (30 μ M in 25 mM HEPES, 300 mM NaCl pH 6.3) was diluted to 3 μ M in binding buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl). The diluted proteins were (, and the surface was washed with binding buffer at S3. The DCV-HCl at different concentrations, (0.2 mg/mL, 0.4 mg/mL, and 0.6 mg/mL) were added at S4. And, all the surfaces were washed with 10 mM Tris-HCl, pH 5.0, 50 mM NaCl at S5.

5.3.5. Daclatasvir-resistance mutations on NS5A and its associations with membranes with the drug: Towards an understanding of the mechanism of resistance mutations

DCV has already been approved for the treatment of hepatitis C disease, and its success is proven in real patients (Pol et al., 2016). However, naturally occurring DCV-resistant viral variants arise quickly and they may give rise to the obstruction of the treatments of the disease shortly. The selection of DCV resistant virus variants revealed several mutations as listed in Figure 1-15. NS5A Y93H mutation was identified in all screened genotypes *in vitro* and (Fridell et al., 2011; Scheel, Gottwein, Mikkelsen, Jensen, & Bukh, 2011). Therefore, we decided to characterize the resistance type of NS5A, NS5A Y93H on SBLs. We performed site-directed mutagenesis on the wild-type

NS5A(2-198)-His-tag construct, by using primers 1a HCV NS5A f Y93H and 1a HCV NS5A r Y93H, available in Table 2-2. DNA sequencing confirmed the sequence and the amino acid substitution, and the protein was purified like wild-type protein by using HisTrap FF column.

In Figure 5-12a, the steps of the experiment were illustrated. A representative QCM-D measurement of the three independent QCM-D experiments is shown in Figure 5-12b. At S0, negatively-charged lipid bilayers were formed, and the frequency was normalized to zero. At S1, the buffer was exchanged to with binding buffer. The stock NS5A(2-198) Y93H-His-tag (at 12.7 μ M in 20 mM HEPES pH 6.3, 300 mM NaCl, 5% glycerol, 0.015% Triton X-100) was diluted to 3 μ M in binding buffer, and they were incubated with different concentrations of DCV at room temperature for 30 minutes on the shaker. Similar to the wild-type protein, 0, 3, 15, and 30 μ M of DCV concentrations were tested and samples were labelled as Y93H, Y93H + DCV (1:1), Y93H + DCV (1:5), Y93H + DCV (1:10) on the Figure 5-12 accordingly. The frequency shift at t=60 minutes was used to correlate mass adsorbed by membrane and Sauerbrey mass calculated. The average and standard deviations of Sauerbrey mass from three independent measurements are shown in Figure 5-12c. The binding of the protein alone at t=60 minutes was normalized as “100” and normalized Sauerbrey mass and concentrations of DCV were plotted in Figure 5-12d. It is seen that DCV does not affect the binding of the protein on the membrane. Table 5-3 shows the protein stock and working concentrations were used in SBL studies.

Figure 5-12. The interactions of daclatasvir-resistant NS5A on supported bilayers. The schematic representation of the measurement. (b) Frequency and dissipation changes of the negatively charged bilayers recorded at overtone 5. The bilayers were formed, and the frequency equalized to zero at S0. The buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl) was exchanged to the binding buffer (10 mM Tris-HCl pH 5.0, 50 mM NaCl) at S1. NS5A (2-198) Y93H-His-tag protein stock (30 μ M in 25 mM HEPES, 300 mM NaCl pH 6.3) was diluted to 3 μ M in binding buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl). The diluted protein was incubated with 0, 3, 15 and 30 μ M of DCV. And, the samples represented Y93H, Y93H + DCV (1:1), Y93H + DCV (1:5), Y93H + DCV (1:10) accordingly. The samples were added to bilayers at S2. The surfaces were washed with 10 mM Tris-HCl pH 5.0 50 mM NaCl at S3. (c) The average and standard deviation values of the changes in the frequencies, dissipation, and theoretical adsorbed mass calculated by Sauerbrey equation from three independent experiments are shown. (d) The normalized Sauerbrey mass versus DCV concentrations is plotted.

Table 5-3. Protein stock concentration for SBL studies.

Protein	NS5A with AH WT	NS5A with AH Y93H
Molecular weight, kDa	23.59518	23.56915
Number of amino acids	206	206
Stock concentration, mg/mL	0.3	0.3
Stock concentration, μM	12.71	12.73
Working concentration, μg/mL	75	75
Molarity, μM	3.18	3.18

5.4. Overview of chapter

From our constructs (in Chapter 3), we learned that N-terminal region of NS5A, especially beyond the location 36, is very hydrophobic. And, therefore adding a tag to those areas is quite challenging. By utilizing this information, we cloned NS5A(2-198) with only C-terminal His-tag, and we solubilized the protein using detergents. Based on our membrane extraction method, we confirmed that NS5A(2-198)-His-tag is not a peripheral membrane protein as high salt concentration was not enough to remove the protein from membranes. However, we could extract the soluble protein from the membrane by using the detergents: *Triton X-100*, *CHAPS*, *n-octyl β -D-maltoside* and *NP-40*. This suggests that NS5A(2-198) behaves like an integral membrane protein. The folding of the protein was checked by CD.

Figure 5-6 shows a typical bilayer formation QCM-D measurement, for each measurement that we performed, we checked the complete bilayer formation, which should give a frequency difference of ~ 26 Hz (Tabaei et al., 2014; Zan et al., 2014). Our SEC buffer of NS5A(2-198) was 20 mM HEPES pH 6.3, 300 mM NaCl, 5% glycerol, 0.015% Triton X-100 and we checked the effect of the buffer to SBLs where the buffer disrupted all bilayers formed, as seen in Figure 5-7. Therefore, we tried to reduce the detergent and glycerol from our binding buffers for further experiments. Next, by utilizing QCM-D combined-SBLs platforms, we first investigated the effect of surface charge on the association of the NS5A(2-198)-His-tag to the membranes. We showed that the protein primarily interacts with negatively-charged membrane surfaces and it almost did not show any affinity to positively-charged membranes as seen in Figure 5-8.

Then, the association of DCV-HCl with membranes were tested where it almost did not interact with SBLs, shown in Figure 5-9. The protein incubated with DCV-HCl showed a reduced amount of binding around (15% of less binding), Figure 5-10. Interestingly, we have not seen a change as we increase DCV-HCl concentrations. Different concentrations of the drug should be tested maybe to see DCV-HCl dependent changes (inhibition of binding to the membranes). The DCV-HCl did not affect the NS5A(2-198)-membrane complex those already formed, shown in Figure 5-11. The association of resistance type of the NS5A with SBLs did not change in the presence of DCV-HCl, shown in Figure 5-12.

Chapter 6: General Discussion & Perspectives

The treatment of hepatitis C is undergoing a revolution with the emerging DAAs and their combinatorial regimens (Bidell et al., 2016), especially in last three years with the approval of DCV and DCV-like drugs (Pol et al., 2016). These dimeric molecules are targeting NS5A protein of HCV with an unknown mechanism (Stanaway et al., 2016) and its resistance mutations are arising quickly. NS5A plays critical roles in life cycle of HCV, but we still do not know the exact function of this protein. DCV resistance mutations are mapped in the N-terminal region of NS5A that comprises amphipathic helix (AH) and domain I (DI). Therefore, we aimed to better understand the behavior of NS5A in many aspects, which may facilitate the development of new classes of inhibitors or methods that improve the efficacy of currently available DAAs.

Our cloning and purification efforts resulted in two proteins. Water soluble NS5A(36-198)-DDDDK and detergent-solubilized NS5A(2-198)-His-tag. NS5A(36-198)-DDDDK protein was very pure and stable at higher concentrations. Therefore, we could use it for structural studies and in solution characterizations as well. NS5A(2-198)-His-tag was also very pure. However, maximum concentrations we could achieve was 0.3 mg/mL. Therefore, we utilized this protein for functional studies like the characterization on membrane mimicking platforms with/without drugs. In next part, we will discuss our results with the findings that are available in the literature.

6.1. General Discussion

Crystallization studies of NS5A-DI resulted in different conformations of homodimers and revealed identical monomer fold as shown in Figure 1-13. As second structure resulted in an alternative conformation of NS5A-DI and as they observed a heterogeneous mixture of protein in solution, they proposed a model of the multimeric state of NS5A-DI (Love et al., 2009). Their theoretically assembled model contains both dimer conformations (1zh1 and 3fqj based on simultaneous recruitment) and forms a superhelical array. It is suggested that this oligomeric state of NS5A-DI might form the membranous network of HCV replication machinery. With the discovery of two more different conformations of NS5A-DI from genotype 1a, the authors of the third crystal structure study proposed that these new dimers and their interfaces could be utilized for the formation of a more expanded network of multimeric NS5A (Lambert et al., 2014).

However, there is no biochemical evidence for the existence of dimers *in vitro* (Lim et al., 2012). Observation of different dimer conformations could be an artifact of crystallization studies, as they were conducted in various conditions (protein purification and crystallization buffers, and additives). Therefore, we wanted to characterize both proteins by purifying in the same manner and in same buffers to understand how they behave and exist in solution. Our results suggest that NS5A-DI remains in a dynamic equilibrium in the solution. The dynamic and self-association behavior of the proteins from genotype 1a and genotype 1b are quite different from each other, although we observe sign of dimers in both genotypes. This changes in dynamic equilibrium between genotype 1a and genotype 1b might be the reason behind their altered drug responses. Figure 6-1 illustrates our findings regarding in solution characterization of the NS5A-DI from genotype 1a and genotype 1b. Note that, different conformations of the proteins exist in solution for genotype 1a and genotype 1b.

Figure 6-1. The schematic illustrations of in solution characterization of NS5A-DI from genotype 1b and 1a. (a) NS5A-DI from genotype 1b in solution dynamic behavior. (b) NS5A-DI from genotype 1a in solution dynamic behavior.

After the multimer model of NS5A was proposed (Love et al., 2009), the model became very popular to explain the extraordinary MoA of NS5A inhibitors. A very recent study estimated that there was one DCV molecule for 47,000 of NS5A proteins in cells and they suggested that a small molecule should have an impact on a large number of NS5A molecules (Sun et al., 2015). By utilizing the multimeric model of NS5A, they proposed an allosteric regulation model for the MoA of DCV where a single DCV bound to NS5A could disturb the communication between NS5A molecules and disrupting the formation of the replication complex. There is no *in vivo* or *in vitro* evidence in the literature to support this model. As we observed the self-association of NS5A-DI, we wanted to test the effect of DCV on the self-association behavior of NS5A-DI. However, our analyses both DLS and SAXS studies resulted in the formation of NS5A aggregates upon DCV addition. The BMH crosslinking revealed that there was

not any change with the addition of DCV. Initially, we found that our results were disappointing, as we could not prove any allosteric regulation of the DCV as it caused aggregates of NS5A *in vitro*. However interestingly, we learned that large aggregates of NS5A in the cells upon DCV treatments were observed in several studies in the literature too. (Bhattacharya et al., 2014; Chukkapalli et al., 2015).

Our results demonstrated that NS5A domain 1 has intrinsic dynamics and represents the varied conformations in solution. Moreover, small changes in the sequence of the protein can shift the equilibrium toward the alternative conformational states as we observed while comparing NS5A-DI from genotype 1a and genotype 1b. In the broad sense, the dynamic behaviors and conformational variability are considered to be key features of proteins in maintaining evolutionary adaptability (Tokuriki & Tawfik, 2009). Adaptability of proteins structures/conformational motions may enable functional interactions (Haliloglu & Bahar, 2015). Therefore, we believe that the alternative conformations of NS5A might be modulating its functional versatility.

We used NS5A(36-198)-DDDDK protein to perform 2D NMR titration, DLS and SAXS analysis in the presence of 15U RNA molecules. In RNA titration experiments, we observed that several peaks shifted upon addition of 15U RNA and even some residues disappeared. Our DLS and SAXS analysis of the protein with 15U RNA resulted in a more compact form of the protein. Taken all together, our preliminary analysis suggested that the interaction of the NS5A-DI with RNA molecules might change the conformation of NS5A, from a flexible structure to compact form. Perhaps that is why we observed shifts in several peaks in NMR titration experiment. However, it should be noted that these results give us only limited information, and require further investigations regarding NS5A-D1 and RNA molecules.

Formation of altered membranous structures, called as membranous web (MW) are the key features of HCV-infected cells. Proper MW formation needs the concerted roles of all replication machinery factors, but NS5A was the only protein inducing DMV-like structures (Romero-Brey et al., 2012). The major and most critical effect of DCV has been just recently published, *i.e.* inhibition of biogenesis of HCV-induced membranous alterations (Berger et al., 2014). By using correlative light electron microscopy, they expose that the formation of the membranous web (MW) was blocked with DCV treatment, even though cells expressed a high amount of viral proteins. Strikingly, DMV

number was greatly decreased in wild type NS5A, on the other hand, in the NS5A Y93H mutant, the DMV diameter was slightly increased, and DMV number was not altered. It should be noted that the DCV-resistance mutations of NS5A protein, are located on the membrane association-proximal site of the dimer structures, as seen in Figure 1-13. These findings motivated us to characterize the N-terminal of NS5A protein on supported bilayer (SBL) platforms with/without drugs. We also introduced a DCV-resistance type of the protein *i.e.* NS5A Y93H to understand the mechanism of resistance to the drug. The wild type NS5A-DI incubated with DCV showed 15% less binding compared to untreated protein. Importantly, there was no change in the membrane anchoring behavior of DCV-resistance variant of the protein between DCV treated and untreated conditions. This shows that DCV does not completely inhibit the association of NS5A-DI with membranes. A slight reduction in the binding of proteins to the SBLs in DCV treatment condition (~15%) might suggest that DCV only targets some conformations of the protein, not all of them.

6.2. Perspectives

Our findings suggest that the further investigations can be undertaken in the following areas:

Mutation analysis of the N-terminal part of NS5A: In this thesis, we showed that NS5A-DI self-associates in solution. However, the biological significance of this self-association behavior is not fully clear. By introducing several mutations in the N-terminal region of NS5A-DI, the important residues those are critical for the multimerization of NS5A can be identified. Later, those mutations can be further tested in HCV-models as well, like JFH-1 cell culture model, to understand the importance of self- association and dynamic behavior of the protein in HCV life cycle.

Understanding the aggregation mechanism of the DCV- treated NS5A-DI proteins: In this study, DCV-treated NS5A-DI resulted in the large aggregates of NS5A-DI in solution. Surprisingly, the aggregates of the proteins with DCV treatment (not only NS5A protein but also the other HCV proteins) were reported in other studies. Therefore, the understanding of the particular mechanism that drives the formation of large aggregates of the NS5A could be key in the future to explain the extraordinary mode of action of NS5A inhibitors.

Characterization of NS5A on SBLs: In this study, we have presented our preliminary results to test the effect of DCV on the membrane binding activity of NS5A-DI. Importantly, we have observed some differences between wild type and DCV-resistance type of the NS5A-DI which is very promising. However, it should be noted that we started our investigations by using a very basic lipid model. In the future, the model membrane can be tuned by changing its physicochemical properties that can better mimic the natural features of the membranous web of HCV-infected cells. This alteration might improve the simulations of the associations between NS5A protein with membranes with/without DCV.

APPENDICES

Appendix-A: Bioinformatics analysis of NS5A of different genotypes

Sequence comparison of NS5A protein from seven genotypes of HCV are tabulated below. NS5A protein sequences from the genotypes of HCV (1a, 1b, 2a, 3a, 4a, 5a, 6a) are retrieved from Uniprot database. Clustal Omega alignment scores are shown, the sequences similarity score varies from 62.95% to 79.64% compared to the reference sequence, *i.e.* NS5A from genotype 1a.

Genotype	Isolate	Uniprot ID	% identity
1a	H77C	P26664	100
1b	Japanese	P26662	80
2a	HC-J6	P26660	63
3a	k3a	Q81495	69
4a	ED43	O39929	69
5a	EUH1480	O39928	70
6a	6a33	Q5I2N3	71

Clustal Omega multiple sequence alignment of NS5A from different genotypes of HCV are given in the figure below. In this analysis, the alignment of the NS5A protein of various genotypes (1b, 1a, 5a, 6a, 3a, 4a, 2a) are represented. Identical amino acids, conserved substitutions, and semi-conserved substitutions are marked with (*), (:), and (.), respectively. Note that sequence variations seem to appear at the C-terminal region of the protein.

1b SGSWLKDVWDWICTVLSDFKTLWLSKLLPRLPGLPFLSCQRYKGVWRGDMQTTCPG 60
1a SGSWLRDINDWICEVLSDFKTLWLSKLLPRLPGLPFLSCQRYKGVWRGDMQTTCPG 60
5a DGTWLRADWDWCTALTDFKAWLQAKLLPQLPGVPPFSCQRYKGVWRGDMNSTKCPG 60
6a ATSWLRDWDWCTVLSDFKWLQAKLLPRLPGLPFLSCQRYKGVWRGDMNSTKCPG 60
3a SGDWLRDIWDWVCSVSDFKTLWLSKLLPRLPGLPFLSCQRYKGVWRGDMNSTKCPG 60
4a AESWLWEVWDWVHLVLSDFKTLWLSKLLPRLPGLPFLSCQRYKGVWRGDMNSTKCPG 60
2a SGSWLRDWDWCTILTDFKAWLQAKLLPRLPGLPFLSCQRYKGVWRGDMNSTKCPG 60
* * : * * : : * * * : * : * : * : * : * : * : * : * : * : *

1b AQITGHVKNGSMRIVGPKTCSNTWHGTFPINAYTTGPTSPAPNYSRALWRVAEEYVE 120
1a AEITGHVKNGSMRIVGPKTCSNTWHGTFPINAYTTGPTSPAPNYSRALWRVAEEYVE 120
5a ATISGHVKNGSMRIVGPKTCSNTWHGTFPINAYTTGPTSPAPNYSRALWRVAEEYVE 120
6a AVIAGHVKNKMTKITGPKTCSNTWHGTFPINAYTTGPTSPAPNYSRALWRVAEEYVE 120
3a ASIAGHVKNKMTKITGPKTCSNTWHGTFPINAYTTGPTSPAPNYSRALWRVAEEYVE 120
4a ADLAGHVKNKMTKITGPKTCSNTWHGTFPINAYTTGPTSPAPNYSRALWRVAEEYVE 120
2a ANISGNVRLGSMRITGPKTCSNTWHGTFPINAYTTGPTSPAPNYSRALWRVAEEYVE 120
* : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

1b VTRVGDFHYVTGMTTQNVKCPQVPAPEFFTEVDGVRHRYAPVCKPLLREEVVFQVGLN 180
1a IRQVGDFHYVTGMTTQNVKCPQVPAPEFFTEVDGVRHRYAPVCKPLLREEVVFQVGLN 180
5a VRRVGDFHYVTGMTTQNVKCPQVPAPEFFTEVDGVRHRYAPVCKPLLREEVVFQVGLN 180
6a VRRVGDFHYVTGMTTQNVKCPQVPAPEFFTEVDGVRHRYAPVCKPLLREEVVFQVGLN 180
3a VRRVGDFHYVTGMTTQNVKCPQVPAPEFFTEVDGVRHRYAPVCKPLLREEVVFQVGLN 180
4a VRRVGDFHYVTGMTTQNVKCPQVPAPEFFTEVDGVRHRYAPVCKPLLREEVVFQVGLN 180
2a VTQHGSDHYITGLTQNVKCPQVPAPEFFTEVDGVRHRYAPVCKPLLREEVVFQVGLN 180
: : * : * : * : * : * : * : * : * : * : * : * : * : * : *

1b ATCTTHHDSPADLIENALLWRQEMGGNITRVESENKVVILDSFDPIRAVE-DEREISVP 299
1a ATCTANHDSPADLIENALLWRQEMGGNITRVESENKVVILDSFDPLVAEE-DEREISVP 299
5a ATCTIQGHHPADLIENALLWRQEMGGNITRVESENKVVILDSFDPLVAEE-DEREISVP 299
6a ATCTTSKDHHPADLIENALLWRQEMGGNITRVESENKVVILDSFDPLVAEE-DEREISVP 299
3a ATCTQTHRHPADLIENALLWRQEMGGNITRVESENKVVILDSFDPLVAEE-DEREISVP 299
4a ATCTAPHDSPADLIENALLWRQEMGGNITRVESENKVVILDSFDPLVAEE-DEREISVP 299
2a ATCTTHGKAYDVMVDANLFM----GGDVTRIESENKVVILDSFDPLVAEE-DEREISVP 299
* * : * : * : * : * : * : * : * : * : * : * : * : * : *

1b AEILRK-PRKFPPALPIWARPDYNPPLLESWKDPDYPPVHGCPLPSTKAPPIPPRRK 358
1a AEILRK-SRRFAQALPVWARPDYNPPLVETWKKPDYPPVHGCPLPSTKAPPIPPRRK 358
5a ADCFKK-GPAFPPALPVWARPDYNPPLLETWKKPDYPPVHGCPLPSTKAPPIPPRRK 358
6a AECHRPPIPKFPPALPIWARPDYNPPLQAWQMPGYEPPVHGCPLPSTKAPPIPPRRK 358
3a AECFKK-PPKYPPALPIWARPDYNPPLDRWKSPDYPPVHGCPLPSTKAPPIPPRRK 358
4a AEILRK-TKKFPPALPIWARPDYNPPLTETWKKQDYQAPTGHGALPPAKQPPVSPRRK 355
2a SEYMLP-KKRFPFALPAWARPDYNPPLVESWKRPDYQATVAGCALPPPKKTPTPPRRR 354
: : : * : * : * : * : * : * : * : * : * : * : * : *

1b R-TVVLTESTVSSALAEATKTFGS-----SGSSAVDS-GTATGPPDQASDGDGKGS 409
1a R-TVVLTESTLSTALAEATKTFGS-----SSTSGITG-DNTTSSSEAPSGCPSDSD 409
5a RKPMELESDTVSQVMADLADARFKVDTPSIE-GQDSALGT-SSQHDGSGPEEKRD--DNSD 414
6a R-LVHLDDESTVSRALAEATKTFGS-----SGSSAVDS-GTATGPPDQASDGDGKGS 414
3a R-TIQLDGSMVSAALAEATKTFGS-----SGSSAVDS-GTATGPPDQASDGDGKGS 414
4a R-TVQLTESVSTALAEATKTFGS-----SSDRDQDL-TPTETDQSGPIVDDASD 409
2a R-TVGLSESSIADALQQLAKSFGQPPP-----SGDGLSTGADAADSGSRTPPDELALSE 409
* : * : * : * : * : * : * : * : * : * : * : *

1b VESYSSMPPLEGEPGDPDL-----DGSWSTVSGEA-GEDVWCC 447
1a AESYSSMPPLEGEPGDPDL-----DGSWSTVSGEANAEDVWCC 448
5a AASYSSMPPLEGEPGDPDL-----SGSWSTVSG--EDNVWCC 450
6a AESYSSMPPLEGEPGDPDL-----DGSWSTVSD--QDDVWCC 451
3a SESCSSMPPLEGEPGDPDL-----CDSWSTVSDSE-EQSVWCC 452
4a DGSYSSMPPLEGEPGDPDL-----SDSWSTVSG--SEDVWCC 445
2a TGSISSMPPLEGEPGDPDL-----DGSWSTVSG--EDNVWCC 446
* * * * * : * * * * * : * * * * *

Appendix-B: Constructs of the recombinant NS5A-DI

a) The sequence alignment of the proteins from crystallization of NS5A-DI.

```

4c1l GSGIPFVSCQRGYRGVWRGDGIMHTRCHCGAEITGHVKNGTMRIVGPRTCRNMWSGTFPI
1zh1 -----FFSCQRGYKGVWRGDGIMQTTCPGQAQITGHVKNNGSMRIVGPRTCSNTWHGTFPI
3fqq ALGVPPFFSCQRGYKGVWRGDGIMQTTCPGQAQITGHVKNNGSMRIVGPRTCSNTWHGTFPI
      *.*****:*****:* * ***:*****:***** * * *****

4c1l NAYTTGPCTPLPAPNYKFALWRVSAEEYVEIRRVGDFHYVSGMTTDNLKCPCQIPSPPEFF
1zh1 NAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVPAPPEFF
3fqq NAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVPAPPEFF
      ***** *****. *****:*****: *****:*****:*****:*:*****

4c1l TELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPDVAVLTSMDDDDK
1zh1 TEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPCEPEPDVAV-----
3fqq TEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPCEPEPDVAVLTSMDDDDK
      **:*****:* *****:* ***: * *****

```

b) The properties of the proteins used in this study.

Component	1a NS5A(36-198)-DDDDK	1b NS5A(36-198)-DDDDK
Number of a.a.	172	172
Molecular weight (Da)	19394.15	19011.57
Theoretical pI	6.64	5.78
The instability index (II)	39.27	31.82
Grand average of hydropathicity	-0.383	-0.288
Ext. coefficient	225940	27430

Appendix-C: Cysteines on the dimer of NS5A-DI

The carton representation of the clam-like dimer structure of NS5A-DI (Tellinghuisen et al., 2005). The figure below shows the cysteine residues (CYS) on the dimer and the distances less than 13.0 Å between two cysteines are shown with yellow dots. The table below shows the number of the cysteines residues and those having distance less than 13.0 Å are indicated.

CYS position	A CYS 39	A CYS 57	A CYS 59	A CYS 80	A CYS 98	A CYS 140	A CYS 142	A CYS 165	A CYS 190	B CYS 39	B CYS 57	B CYS 59	B CYS 80	B CYS 98	B CYS 140	B CYS 142	B CYS 165	B CYS 190
A CYS 39	0																	
A CYS 57	6.1	0																
A CYS 59	5.2	5.3	0															
A CYS 80	6.8	7.8	7	0														
A CYS 98					0													
A CYS 140						0												
A CYS 142						6.5	0											
A CYS 165								0										
A CYS 190							5.3		0									
B CYS 39	11.5							8.7		0								
B CYS 57	12.6		12.1						12.6	0								
B CYS 59	12.8	12.2	12.2					11.8	6.4	5.3	0							
B CYS 80								8.4	6.6	8.2	7	0						
B CYS 98													0					
B CYS 140														0				
B CYS 142															0			
B CYS 165	8.7	3.8	11.8	8.3												0		
B CYS 190																	0	0

Appendix-D: Particle size estimation of NS5A-DI monomer based on crystal structure

The crystal structure (PDB: 1zh1) is shown below on the left. The distances between two atoms are calculated by using PYMOL visualization program and the average expected size of the NS5A-DI monomer size is shown on the right. A monomer should be around 2.3 nm in width and 5 nm in height approximately. (Please note that it is very same for both genotype 1a and 1b.)

Appendix-E: Gel filtration Standards for S200 (25 mL) column

The buffer (25 mM HEPES of pH 7.4, 150 mM NaCl, 10% glycerol, 0.01 NaN₃) was used during the run.

MW (kDa)	Elution volume (mL)
1.350	20.46
17	17.40
44	15.38
158	12.65
670	9.60

Appendix-F: Gel filtration Standards of Superdex 75 5/150 GL

The buffer (25 mM HEPES of pH 7.4, 150 mM NaCl, 10% glycerol, 0.01 NaN₃) was used during the run.

MW (kDa)	Elution volume (mL)
1.350	2.752
17	2.001
44	1.7072
158	1.46
670	< 1.46

BIBLIOGRAPHY¹

- Afdhal, N. H., & Serfaty, L. (2016). Effect of Registries and Cohort Studies on HCV Treatment. *Gastroenterology*, 151(3), 387-390. doi:10.1053/j.gastro.2016.07.024
- Agnello, V., Abel, G., Elfahal, M., Knight, G. B., & Zhang, Q. X. (1999). Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proc Natl Acad Sci U S A*, 96(22), 12766-12771.
- Andre, P., Komurian-Pradel, F., Deforges, S., Perret, M., Berland, J. L., Sodoyer, M., . . . Lotteau, V. (2002). Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. *J Virol*, 76(14), 6919-6928.
- Appel, N., Pietschmann, T., & Bartenschlager, R. (2005). Mutational analysis of hepatitis C virus nonstructural protein 5A: potential role of differential phosphorylation in RNA replication and identification of a genetically flexible domain. *J Virol*, 79(5), 3187-3194. doi:10.1128/JVI.79.5.3187-3194.2005
- Ascher, D. B., Wielens, J., Nero, T. L., Doughty, L., Morton, C. J., & Parker, M. W. (2014). Potent hepatitis C inhibitors bind directly to NS5A and reduce its affinity for RNA. *Sci Rep*, 4, 4765. doi:10.1038/srep04765
- Balakrishna, A. M., Basak, S., Manimekalai, M. S., & Gruber, G. (2015). Crystal structure of subunits D and F in complex gives insight into energy transmission of the eukaryotic V-ATPase from *Saccharomyces cerevisiae*. *J Biol Chem*, 290(6), 3183-3196. doi:10.1074/jbc.M114.622688
- Ball, J. K., Tarr, A. W., & McKeating, J. A. (2014). The past, present and future of neutralizing antibodies for hepatitis C virus. *Antiviral Res*, 105, 100-111. doi:10.1016/j.antiviral.2014.02.013
- Bartenschlager, R., Ahlborn-Laake, L., Mous, J., & Jacobsen, H. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *J Virol*, 67(7), 3835-3844.
- Bartenschlager, R. F., Rice, C. M., & Sofia, M. J. (2016). Hepatitis C Virus-From Discovery to Cure: The 2016 Lasker-DeBaakey Clinical Medical Research Award. *JAMA*, 316(12), 1254-1255. doi:10.1001/jama.2016.13713
- Bartenschlager, R., Frese, M., & Pietschmann, T. (2004). Novel insights into hepatitis C virus replication and persistence. *Adv Virus Res*, 63, 71-180. doi:10.1016/S0065-3527(04)63002-8
- Bartenschlager, R., Penin, F., Lohmann, V., & Andre, P. (2011). Assembly of infectious hepatitis C virus particles. *Trends Microbiol*, 19(2), 95-103.

---

¹ Prepared following the instructions of APA style (6th Edition).

- Bartosch, B., & Dubuisson, J. (2010). Recent advances in hepatitis C virus cell entry. *Viruses*, 2(3), 692-709. doi:10.3390/v2030692
- Berger, C., Romero-Brey, I., Radujkovic, D., Terreux, R., Zayas, M., Paul, D., . . . Bartenschlager, R. (2014). Daclatasvir-like inhibitors of NS5A block early biogenesis of hepatitis C virus-induced membranous replication factories, independent of RNA replication. *Gastroenterology*, 147(5), 1094-1105 e1025. doi:10.1053/j.gastro.2014.07.019
- Bhattacharya, D., Ansari, I. H., Hamatake, R., Walker, J., Kazmierski, W. M., & Striker, R. (2014). Pharmacological disruption of hepatitis C NS5A protein intra- and intermolecular conformations. *J Gen Virol*, 95(Pt 2), 363-372. doi:10.1099/vir.0.054569-0
- Bidell, M. R., McLaughlin, M., Faragon, J., Morse, C., & Patel, N. (2016). Desirable Characteristics of Hepatitis C Treatment Regimens: A Review of What We Have and What We Need. *Infect Dis Ther*, 5(3), 299-312. doi:10.1007/s40121-016-0118-x
- Blight, K. J., & Norgard, E. A. (2006). HCV Replicon Systems. In S. L. Tan (Ed.), *Hepatitis C Viruses: Genomes and Molecular Biology*. Norfolk (UK).
- Blumberg, B. S., & London, W. T. (1982). Hepatitis B virus: pathogenesis and prevention of primary cancer of the liver. *Cancer*, 50(11 Suppl), 2657-2665.
- Boulant, S., Montserret, R., Hope, R. G., Ratniner, M., Targett-Adams, P., Laverne, J. P., . . . McLauchlan, J. (2006). Structural determinants that target the hepatitis C virus core protein to lipid droplets. *J Biol Chem*, 281(31), 22236-22247. doi:10.1074/jbc.M601031200
- Boulant, S., Vanbelle, C., Ebel, C., Penin, F., & Laverne, J. P. (2005). Hepatitis C virus core protein is a dimeric alpha-helical protein exhibiting membrane protein features. *J Virol*, 79(17), 11353-11365. doi:10.1128/JVI.79.17.11353-11365.2005
- Bradley, D. W., McCaustland, K. A., Cook, E. H., Schable, C. A., Ebert, J. W., & Maynard, J. E. (1985). Posttransfusion non-A, non-B hepatitis in chimpanzees. Physicochemical evidence that the tubule-forming agent is a small, enveloped virus. *Gastroenterology*, 88(3), 773-779.
- Bradrick, S. S., Walters, R. W., & Gromeier, M. (2006). The hepatitis C virus 3'-untranslated region or a poly(A) tract promote efficient translation subsequent to the initiation phase. *Nucleic Acids Res*, 34(4), 1293-1303. doi:10.1093/nar/gkl019
- Brass, V., Bieck, E., Montserret, R., Wolk, B., Hellings, J. A., Blum, H. E., . . . Moradpour, D. (2002). An amino-terminal amphipathic alpha-helix mediates membrane association of the hepatitis C virus nonstructural protein 5A. *J Biol Chem*, 277(10), 8130-8139. doi:10.1074/jbc.M111289200

- Bressanelli, S., Stiasny, K., Allison, S. L., Stura, E. A., Duquerroy, S., Lescar, J., . . . Rey, F. A. (2004). Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. *EMBO J*, 23(4), 728-738. doi:10.1038/sj.emboj.7600064
- Bressanelli, S., Tomei, L., Rey, F. A., & De Francesco, R. (2002). Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. *J Virol*, 76(7), 3482-3492.
- Brown, E. A., Zhang, H., Ping, L. H., & Lemon, S. M. (1992). Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Res*, 20(19), 5041-5045.
- Buhler, S., & Bartenschlager, R. (2012). New targets for antiviral therapy of chronic hepatitis C. *Liver Int*, 32 Suppl 1, 9-16. doi:10.1111/j.1478-3231.2011.02701.x
- Bukh, J. (2012). Animal models for the study of hepatitis C virus infection and related liver disease. *Gastroenterology*, 142(6), 1279-1287 e1273. doi:10.1053/j.gastro.2012.02.016
- Bunde, R. L., Jarvi, E. J., & Rosentreter, J. J. (1998). Piezoelectric quartz crystal biosensors. *Talanta*, 46(6), 1223-1236.
- Campbell, I. D. (2002). Timeline: the march of structural biology. *Nat Rev Mol Cell Biol*, 3(5), 377-381. doi:10.1038/nrm800
- Chatel-Chaix, L., & Bartenschlager, R. (2014). Dengue virus- and hepatitis C virus-induced replication and assembly compartments: the enemy inside--caught in the web. *J Virol*, 88(11), 5907-5911. doi:10.1128/JVI.03404-13
- Chen, S. L., & Morgan, T. R. (2006). The natural history of hepatitis C virus (HCV) infection. *Int J Med Sci*, 3(2), 47-52.
- Choo, Q. L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W., & Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, 244(4902), 359-362.
- Chukkapalli, V., Berger, K. L., Kelly, S. M., Thomas, M., Deiters, A., & Randall, G. (2015). Daclatasvir inhibits hepatitis C virus NS5A motility and hyper-accumulation of phosphoinositides. *Virology*, 476, 168-179. doi:10.1016/j.virol.2014.12.018
- Dane, D. S., Cameron, C. H., & Briggs, M. (1970). Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet*, 1(7649), 695-698.
- de Chassey, B., Navratil, V., Tafforeau, L., Hiet, M. S., Aublin-Gex, A., Agaoglu, S., . . . Lotteau, V. (2008). Hepatitis C virus infection protein network. *Mol Syst Biol*, 4, 230. doi:10.1038/msb.2008.66
- Dip, P. V., Kamariah, N., Subramanian Manimekalai, M. S., Nartey, W., Balakrishna, A. M., Eisenhaber, F., . . . Gruber, G. (2014). Structure, mechanism and ensemble formation of the alkylhydroperoxide reductase subunits AhpC and

- AhpF from *Escherichia coli*. *Acta Crystallogr D Biol Crystallogr*, 70(Pt 11), 2848-2862. doi:10.1107/S1399004714019233
- Dixon, M. C. (2008). Quartz crystal microbalance with dissipation monitoring: enabling real-time characterization of biological materials and their interactions. *J Biomol Tech*, 19(3), 151-158.
- Dubuisson, J. (2007). Hepatitis C virus proteins. *World J Gastroenterol*, 13(17), 2406-2415.
- Dubuisson, J., Helle, F., & Cocquerel, L. (2008). Early steps of the hepatitis C virus life cycle. *Cell Microbiol*, 10(4), 821-827. doi:10.1111/j.1462-5822.2007.01107.x
- Durand, D., Vives, C., Cannella, D., Perez, J., Pebay-Peyroula, E., Vachette, P., & Fieschi, F. (2010). NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. *J Struct Biol*, 169(1), 45-53. doi:10.1016/j.jsb.2009.08.009
- Egger, D., Wolk, B., Gosert, R., Bianchi, L., Blum, H. E., Moradpour, D., & Bienz, K. (2002). Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J Virol*, 76(12), 5974-5984.
- El Omari, K., Iourin, O., Kadlec, J., Sutton, G., Harlos, K., Grimes, J. M., & Stuart, D. I. (2014). Unexpected structure for the N-terminal domain of hepatitis C virus envelope glycoprotein E1. *Nat Commun*, 5, 4874. doi:10.1038/ncomms5874
- Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., . . . Sato, C. (1995). Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *J Clin Invest*, 96(1), 224-230. doi:10.1172/JCI118025
- Feinstone, S. M., Kapikian, A. Z., & Purceli, R. H. (1973). Hepatitis A: detection by immune electron microscopy of a viruslike antigen associated with acute illness. *Science*, 182(4116), 1026-1028.
- Feinstone, S. M., Kapikian, A. Z., Purcell, R. H., Alter, H. J., & Holland, P. V. (1975). Transfusion-associated hepatitis not due to viral hepatitis type A or B. *N Engl J Med*, 292(15), 767-770. doi:10.1056/NEJM197504102921502
- Feuerstein, S., Solyom, Z., Aladag, A., Favier, A., Schwarten, M., Hoffmann, S., . . . Brutscher, B. (2012). Transient structure and SH3 interaction sites in an intrinsically disordered fragment of the hepatitis C virus protein NS5A. *J Mol Biol*, 420(4-5), 310-323. doi:10.1016/j.jmb.2012.04.023
- Foster, T. L., Belyaeva, T., Stonehouse, N. J., Pearson, A. R., & Harris, M. (2010). All three domains of the hepatitis C virus nonstructural NS5A protein contribute to RNA binding. *J Virol*, 84(18), 9267-9277. doi:10.1128/JVI.00616-10
- Foster, T. L., Gallay, P., Stonehouse, N. J., & Harris, M. (2011). Cyclophilin A

- interacts with domain II of hepatitis C virus NS5A and stimulates RNA binding in an isomerase-dependent manner. *J Virol*, 85(14), 7460-7464. doi:10.1128/JVI.00393-11
- Franke, Daniel, & Svergun, Dmitri I. (2009). DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. *Journal of Applied Crystallography*, 42(2), 342-346. doi:10.1107/S0021889809000338
- Freedman, H., Logan, M. R., Law, J. L., & Houghton, M. (2016). Structure and Function of the Hepatitis C Virus Envelope Glycoproteins E1 and E2: Antiviral and Vaccine Targets. *ACS Infect Dis*, 2(11), 749-762. doi:10.1021/acsinfecdis.6b00110
- Fridell, R. A., Qiu, D., Wang, C., Valera, L., & Gao, M. (2010). Resistance analysis of the hepatitis C virus NS5A inhibitor BMS-790052 in an in vitro replicon system. *Antimicrob Agents Chemother*, 54(9), 3641-3650. doi:10.1128/AAC.00556-10
- Fridell, R. A., Wang, C., Sun, J. H., O'Boyle, D. R., 2nd, Nower, P., Valera, L., . . . Gao, M. (2011). Genotypic and phenotypic analysis of variants resistant to hepatitis C virus nonstructural protein 5A replication complex inhibitor BMS-790052 in humans: in vitro and in vivo correlations. *Hepatology*, 54(6), 1924-1935. doi:10.1002/hep.24594
- Friebe, P., & Bartenschlager, R. (2002). Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. *J Virol*, 76(11), 5326-5338.
- Friebe, P., Lohmann, V., Krieger, N., & Bartenschlager, R. (2001). Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication. *J Virol*, 75(24), 12047-12057. doi:10.1128/JVI.75.24.12047-12057.2001
- Gale, M. J., Jr., Korth, M. J., Tang, N. M., Tan, S. L., Hopkins, D. A., Dever, T. E., . . . Katze, M. G. (1997). Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. *Virology*, 230(2), 217-227. doi:10.1006/viro.1997.8493
- Gao, M., Nettles, R. E., Belema, M., Snyder, L. B., Nguyen, V. N., Fridell, R. A., . . . Hamann, L. G. (2010). Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nature*, 465(7294), 96-100. doi:10.1038/nature08960
- Gell, D. A., Grant, R. P., & Mackay, J. P. (2012). The detection and quantitation of protein oligomerization. *Adv Exp Med Biol*, 747, 19-41. doi:10.1007/978-1-4614-3229-6_2
- Goffard, A., Callens, N., Bartosch, B., Wychowski, C., Cosset, F. L., Montpellier, C., & Dubuisson, J. (2005). Role of N-linked glycans in the functions of hepatitis C virus envelope glycoproteins. *J Virol*, 79(13), 8400-8409. doi:10.1128/JVI.79.13.8400-8409.2005
- Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M., & Rice, C. M. (1993).

- Expression and identification of hepatitis C virus polyprotein cleavage products. *J Virol*, 67(3), 1385-1395.
- Gravitz, L. (2011). Introduction: a smouldering public-health crisis. *Nature*, 474(7350), S2-4. doi:10.1038/474S2a
- Griffin, S. D., Beales, L. P., Clarke, D. S., Worsfold, O., Evans, S. D., Jaeger, J., . . . Rowlands, D. J. (2003). The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett*, 535(1-3), 34-38.
- Griffin, S., Stgelais, C., Owsianka, A. M., Patel, A. H., Rowlands, D., & Harris, M. (2008). Genotype-dependent sensitivity of hepatitis C virus to inhibitors of the p7 ion channel. *Hepatology*, 48(6), 1779-1790. doi:10.1002/hep.22555
- Hadziyannis, S. J. (1998). Fulminant hepatitis and the new G/GBV-C flavivirus. *J Viral Hepat*, 5(1), 15-19.
- Hajarizadeh, B., Grebely, J., & Dore, G. J. (2013). Epidemiology and natural history of HCV infection. *Nat Rev Gastroenterol Hepatol*, 10(9), 553-562. doi:10.1038/nrgastro.2013.107
- Haliloglu, T., & Bahar, I. (2015). Adaptability of protein structures to enable functional interactions and evolutionary implications. *Curr Opin Struct Biol*, 35, 17-23. doi:10.1016/j.sbi.2015.07.007
- Hall, G. F. (2007). Hepatitis A, B, C, D, E, G: an update. *Ethn Dis*, 17(2 Suppl 2), S2-40-45.
- Han, Y., Niu, J., Wang, D., & Li, Y. (2016). Hepatitis C Virus Protein Interaction Network Analysis Based on Hepatocellular Carcinoma. *PLoS One*, 11(4), e0153882. doi:10.1371/journal.pone.0153882
- Hanoulle, X., Verdegem, D., Badillo, A., Wieruszeski, J. M., Penin, F., & Lippens, G. (2009). Domain 3 of non-structural protein 5A from hepatitis C virus is natively unfolded. *Biochem Biophys Res Commun*, 381(4), 634-638. doi:10.1016/j.bbrc.2009.02.108
- Heim, M. H. (2013). 25 years of interferon-based treatment of chronic hepatitis C: an epoch coming to an end. *Nat Rev Immunol*, 13(7), 535-542. doi:10.1038/nri3463
- Hoffman, B., & Liu, Q. (2011). Hepatitis C viral protein translation: mechanisms and implications in developing antivirals. *Liver Int*, 31(10), 1449-1467. doi:10.1111/j.1478-3231.2011.02543.x
- Hope, R. G., & McLauchlan, J. (2000). Sequence motifs required for lipid droplet association and protein stability are unique to the hepatitis C virus core protein. *J Gen Virol*, 81(Pt 8), 1913-1925.
- Hope, R. G., Murphy, D. J., & McLauchlan, J. (2002). The domains required to direct core proteins of hepatitis C virus and GB virus-B to lipid droplets share

- common features with plant oleosin proteins. *J Biol Chem*, 277(6), 4261-4270. doi:10.1074/jbc.M108798200
- Huang, L., Hwang, J., Sharma, S. D., Hargittai, M. R., Chen, Y., Arnold, J. J., . . . Cameron, C. E. (2005). Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. *J Biol Chem*, 280(43), 36417-36428. doi:10.1074/jbc.M508175200
- Huang, Y., Staschke, K., De Francesco, R., & Tan, S. L. (2007). Phosphorylation of hepatitis C virus NS5A nonstructural protein: a new paradigm for phosphorylation-dependent viral RNA replication? *Virology*, 364(1), 1-9. doi:10.1016/j.virol.2007.01.042
- Hughes, M., Gretton, S., Shelton, H., Brown, D. D., McCormick, C. J., Angus, A. G., . . . Harris, M. (2009). A conserved proline between domains II and III of hepatitis C virus NS5A influences both RNA replication and virus assembly. *J Virol*, 83(20), 10788-10796. doi:10.1128/JVI.02406-08
- Hwang, J., Huang, L., Cordek, D. G., Vaughan, R., Reynolds, S. L., Kihara, G., . . . Cameron, C. E. (2010). Hepatitis C virus nonstructural protein 5A: biochemical characterization of a novel structural class of RNA-binding proteins. *J Virol*, 84(24), 12480-12491. doi:10.1128/JVI.01319-10
- Jackman, J. A., & Cho, N. J. (2012). Model membrane platforms for biomedicine: case study on antiviral drug development. *Biointerphases*, 7(1-4), 18. doi:10.1007/s13758-011-0018-2
- Ji, H., Fraser, C. S., Yu, Y., Leary, J., & Doudna, J. A. (2004). Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc Natl Acad Sci U S A*, 101(49), 16990-16995. doi:10.1073/pnas.0407402101
- Johannsmann, D. (2008). Viscoelastic, mechanical, and dielectric measurements on complex samples with the quartz crystal microbalance. *Phys Chem Chem Phys*, 10(31), 4516-4534. doi:10.1039/b803960g
- Karahan, H. E., Karakus, K., & Birer, O. (2015). Simultaneous DLS–SLS study of titanium and titanium/silicon oxide sol growth. *J Sol-Gel Sci. and Technol*, 76, 251-259.
- Kato, T., Date, T., Miyamoto, M., Furusaka, A., Tokushige, K., Mizokami, M., & Wakita, T. (2003). Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon. *Gastroenterology*, 125(6), 1808-1817.
- Khan, A. G., Whidby, J., Miller, M. T., Scarborough, H., Zatorski, A. V., Cygan, A., . . . Marcotrigiano, J. (2014). Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. *Nature*, 509(7500), 381-384. doi:10.1038/nature13117
- Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett*, 589(19 Pt A), 2570-2577. doi:10.1016/j.febslet.2015.08.027

- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Kochb, M.H.J., & Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *JOURNAL OF APPLIED CRYSTALLOGRAPHY*, 36, 1277-1282.
- Konarev, Petr V., Petoukhov, Maxim V., Volkov, Vladimir V., & Svergun, Dmitri I. (2006). ATSAS 2.1, a program package for small-angle scattering data analysis. *JOURNAL OF APPLIED CRYSTALLOGRAPHY*, 39(2), 277-286. doi:doi:10.1107/S0021889806004699
- Kong, L., Giang, E., Nieuwsma, T., Kadam, R. U., Cogburn, K. E., Hua, Y., . . . Law, M. (2013). Hepatitis C virus E2 envelope glycoprotein core structure. *Science*, 342(6162), 1090-1094. doi:10.1126/science.1243876
- Kong, L., Lee, D. E., Kadam, R. U., Liu, T., Giang, E., Nieuwsma, T., . . . Law, M. (2016). Structural flexibility at a major conserved antibody target on hepatitis C virus E2 antigen. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1609780113
- Koutsoudakis, G., Kaul, A., Steinmann, E., Kallis, S., Lohmann, V., Pietschmann, T., & Bartenschlager, R. (2006). Characterization of the early steps of hepatitis C virus infection by using luciferase reporter viruses. *J Virol*, 80(11), 5308-5320. doi:10.1128/JVI.02460-05
- Kozin, M. B., & Svergun, D. I. (2001). Automated matching of high- and low-resolution structural models. *Journal of Applied Crystallography*, 34(1), 33-41. doi:10.1107/s0021889800014126
- Kwon, H. J., Xing, W., Chan, K., Niedziela-Majka, A., Brendza, K. M., Kirschberg, T., . . . Sakowicz, R. (2015). Direct binding of ledipasvir to HCV NS5A: mechanism of resistance to an HCV antiviral agent. *PLoS One*, 10(4), e0122844. doi:10.1371/journal.pone.0122844
- Lambert, S. M., Langley, D. R., Garnett, J. A., Angell, R., Hedgethorpe, K., Meanwell, N. A., & Matthews, S. J. (2014). The crystal structure of NS5A domain 1 from genotype 1a reveals new clues to the mechanism of action for dimeric HCV inhibitors. *Protein Sci*, 23(6), 723-734. doi:10.1002/pro.2456
- Lavanchy, D. (2009). The global burden of hepatitis C. *Liver Int*, 29 Suppl 1, 74-81. doi:10.1111/j.1478-3231.2008.01934.x
- Lee, C., Ma, H., Hang, J. Q., Leveque, V., Sklan, E. H., Elazar, M., . . . Glenn, J. S. (2011). The hepatitis C virus NS5A inhibitor (BMS-790052) alters the subcellular localization of the NS5A non-structural viral protein. *Virology*, 414(1), 10-18. doi:10.1016/j.virol.2011.03.026
- Lemay, K. L., Treadaway, J., Angulo, I., & Tellinghuisen, T. L. (2013). A hepatitis C virus NS5A phosphorylation site that regulates RNA replication. *J Virol*, 87(2), 1255-1260. doi:10.1128/JVI.02154-12
- Lemm, J. A., Leet, J. E., O'Boyle, D. R., 2nd, Romine, J. L., Huang, X. S., Schroeder, D. R., . . . Gao, M. (2011). Discovery of potent hepatitis C virus NS5A inhibitors with dimeric structures. *Antimicrob Agents Chemother*, 55(8), 3795-3802. doi:10.1128/AAC.00146-11

- Lemm, J. A., O'Boyle, D., 2nd, Liu, M., Nower, P. T., Colonno, R., Deshpande, M. S., . . . Gao, M. (2010). Identification of hepatitis C virus NS5A inhibitors. *J Virol*, 84(1), 482-491. doi:10.1128/JVI.01360-09
- Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. F., & Weber, P. C. (1999). Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat Struct Biol*, 6(10), 937-943. doi:10.1038/13305
- Liang, Y., Kang, C. B., & Yoon, H. S. (2006). Molecular and structural characterization of the domain 2 of hepatitis C virus non-structural protein 5A. *Mol Cells*, 22(1), 13-20.
- Lim, P. J., Chatterji, U., Cordek, D., Sharma, S. D., Garcia-Rivera, J. A., Cameron, C. E., . . . Gallay, P. A. (2012). Correlation between NS5A dimerization and hepatitis C virus replication. *J Biol Chem*, 287(36), 30861-30873. doi:10.1074/jbc.M112.376822
- Lindenbach, B. D., Pragai, B. M., Montserret, R., Beran, R. K., Pyle, A. M., Penin, F., & Rice, C. M. (2007). The C terminus of hepatitis C virus NS4A encodes an electrostatic switch that regulates NS5A hyperphosphorylation and viral replication. *J Virol*, 81(17), 8905-8918. doi:10.1128/JVI.00937-07
- Lindenbach, B. D., & Rice, C. M. (2013). The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol*, 11(10), 688-700. doi:10.1038/nrmicro3098
- Lindenbach, B. D., Thiel, H. J., & Rice, C. M. (2007). Flaviviridae: The viruses and their replication. *Fields Virology, Vol. 1*.
- Lindstrom, H., Lundin, M., Haggstrom, S., & Persson, M. A. (2006). Mutations of the Hepatitis C virus protein NS4B on either side of the ER membrane affect the efficiency of subgenomic replicons. *Virus Res*, 121(2), 169-178. doi:10.1016/j.virusres.2006.05.008
- Liu, S., Yang, W., Shen, L., Turner, J. R., Coyne, C. B., & Wang, T. (2009). Tight junction proteins claudin-1 and occludin control hepatitis C virus entry and are downregulated during infection to prevent superinfection. *J Virol*, 83(4), 2011-2014. doi:10.1128/JVI.01888-08
- Liu, Z., Yang, F., Robotham, J. M., & Tang, H. (2009). Critical role of cyclophilin A and its prolyl-peptidyl isomerase activity in the structure and function of the hepatitis C virus replication complex. *J Virol*, 83(13), 6554-6565. doi:10.1128/JVI.02550-08
- Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L., & Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science*, 285(5424), 110-113.
- Lorenz, I. C., Marcotrigiano, J., Dentzer, T. G., & Rice, C. M. (2006). Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature*, 442(7104), 831-835. doi:10.1038/nature04975

- Love, R. A., Brodsky, O., Hickey, M. J., Wells, P. A., & Cronin, C. N. (2009). Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. *J Virol*, 83(9), 4395-4403. doi:10.1128/JVI.02352-08
- Lundin, M., Monne, M., Widell, A., Von Heijne, G., & Persson, M. A. (2003). Topology of the membrane-associated hepatitis C virus protein NS4B. *J Virol*, 77(9), 5428-5438.
- MacCallum, F. (1947). Homologous serum hepatitis. *Lancet*, 2(691), 2.
- Macdonald, A., Crowder, K., Street, A., McCormick, C., & Harris, M. (2004). The hepatitis C virus NS5A protein binds to members of the Src family of tyrosine kinases and regulates kinase activity. *J Gen Virol*, 85(Pt 3), 721-729. doi:10.1099/vir.0.19691-0
- Macdonald, A., Crowder, K., Street, A., McCormick, C., Saksela, K., & Harris, M. (2003). The hepatitis C virus non-structural NS5A protein inhibits activating protein-1 function by perturbing ras-ERK pathway signaling. *J Biol Chem*, 278(20), 17775-17784. doi:10.1074/jbc.M210900200
- Macdonald, A., & Harris, M. (2004). Hepatitis C virus NS5A: tales of a promiscuous protein. *J Gen Virol*, 85(Pt 9), 2485-2502. doi:10.1099/vir.0.80204-0
- Magrane, M., & Consortium, U. (2011). UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)*, 2011, bar009. doi:10.1093/database/bar009
- Mankouri, J., Dallas, M. L., Hughes, M. E., Griffin, S. D., Macdonald, A., Peers, C., & Harris, M. (2009). Suppression of a pro-apoptotic K⁺ channel as a mechanism for hepatitis C virus persistence. *Proc Natl Acad Sci U S A*, 106(37), 15903-15908. doi:10.1073/pnas.0906798106
- Mankouri, J., Milward, A., Pryde, K. R., Warter, L., Martin, A., & Harris, M. (2008). A comparative cell biological analysis reveals only limited functional homology between the NS5A proteins of hepatitis C virus and GB virus B. *J Gen Virol*, 89(Pt 8), 1911-1920. doi:10.1099/vir.0.2008/001131-0
- Marcotrigiano, J., & Tellinghuisen, T. (2009). Purification and crystallization of NS5A domain I of hepatitis C virus. *Methods Mol Biol*, 510, 85-94. doi:10.1007/978-1-59745-394-3_7
- Masaki, T., Matsunaga, S., Takahashi, H., Nakashima, K., Kimura, Y., Ito, M., . . . Suzuki, T. (2014). Involvement of hepatitis C virus NS5A hyperphosphorylation mediated by casein kinase I-alpha in infectious virus production. *J Virol*, 88(13), 7541-7555. doi:10.1128/JVI.03170-13
- Masaki, T., Suzuki, R., Murakami, K., Aizaki, H., Ishii, K., Murayama, A., . . . Suzuki, T. (2008). Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. *J Virol*, 82(16), 7964-7976. doi:10.1128/JVI.00826-08
- McDonald, S. (1918). ACUTE YELLOW ATROPHY IN SYPHILIS: (A Preliminary Note.). *Br Med J*, 1(2977), 76-78.

- McGlynn, K. A., & London, W. T. (2011). The global epidemiology of hepatocellular carcinoma: present and future. *Clin Liver Dis*, 15(2), 223-243, vii-x. doi:10.1016/j.cld.2011.03.006
- McLauchlan, J., Lemberg, M. K., Hope, G., & Martoglio, B. (2002). Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *EMBO J*, 21(15), 3980-3988. doi:10.1093/emboj/cdf414
- Mercer, D. F., Schiller, D. E., Elliott, J. F., Douglas, D. N., Hao, C., Rinfret, A., . . . Kneteman, N. M. (2001). Hepatitis C virus replication in mice with chimeric human livers. *Nat Med*, 7(8), 927-933. doi:10.1038/90968
- Mertens, H. D., & Svergun, D. I. (2010). Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J Struct Biol*, 172(1), 128-141. doi:10.1016/j.jsb.2010.06.012
- Messina, J. P., Humphreys, I., Flaxman, A., Brown, A., Cooke, G. S., Pybus, O. G., & Barnes, E. (2015). Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology*, 61(1), 77-87. doi:10.1002/hep.27259
- Micsonai, A., Wien, F., Kernya, L., Lee, Y. H., Goto, Y., Refregiers, M., & Kardos, J. (2015). Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc Natl Acad Sci U S A*, 112(24), E3095-3103. doi:10.1073/pnas.1500851112
- Moradpour, D., Gosert, R., Egger, D., Penin, F., Blum, H. E., & Bienz, K. (2003). Membrane association of hepatitis C virus nonstructural proteins and identification of the membrane alteration that harbors the viral replication complex. *Antiviral Res*, 60(2), 103-109.
- Moradpour, D., Penin, F., & Rice, C. M. (2007). Replication of hepatitis C virus. *Nat Rev Microbiol*, 5(6), 453-463. doi:10.1038/nrmicro1645
- Murray, C. L., Jones, C. T., Tassello, J., & Rice, C. M. (2007). Alanine scanning of the hepatitis C virus core protein reveals numerous residues essential for production of infectious virus. *J Virol*, 81(19), 10220-10231. doi:10.1128/JVI.00793-07
- Mylonas, Efstratios, & Svergun, Dmitri I. (2007). Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *JOURNAL OF APPLIED CRYSTALLOGRAPHY*, 40(s1), s245-s249. doi:doi:10.1107/S002188980700252X
- Nakamoto, S., Kanda, T., Wu, S., Shirasawa, H., & Yokosuka, O. (2014). Hepatitis C virus NS5A inhibitors and drug resistance mutations. *World J Gastroenterol*, 20(11), 2902-2912. doi:10.3748/wjg.v20.i11.2902
- O'Boyle, D. R., 2nd, Nower, P. T., Lemm, J. A., Valera, L., Sun, J. H., Rigat, K., . . . Gao, M. (2005). Development of a cell-based high-throughput specificity screen using a hepatitis C virus-bovine viral diarrhea virus dual replicon assay. *Antimicrob Agents Chemother*, 49(4), 1346-1353. doi:10.1128/AAC.49.4.1346-1353.2005

- O'Boyle Li, D. R., Sun, J. H., Nower, P. T., Lemm, J. A., Fridell, R. A., Wang, C., . . . Gao, M. (2013). Characterizations of HCV NS5A replication complex inhibitors. *Virology*, 444(1-2), 343-354. doi:10.1016/j.virol.2013.06.032
- Op De Beeck, A., Cocquerel, L., & Dubuisson, J. (2001). Biogenesis of hepatitis C virus envelope glycoproteins. *J Gen Virol*, 82(Pt 11), 2589-2595. doi:10.1099/0022-1317-82-11-2589
- Op De Beeck, A., & Dubuisson, J. (2003). Topology of hepatitis C virus envelope glycoproteins. *Rev Med Virol*, 13(4), 233-241. doi:10.1002/rmv.391
- Oxford, J. , Kellann, P. Leslie, Collier. (2016). *Human Virology* (Fifth edition ed.): OXFORD.
- Pavlovic, D., Neville, D. C., Argaud, O., Blumberg, B., Dwek, R. A., Fischer, W. B., & Zitzmann, N. (2003). The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proc Natl Acad Sci U S A*, 100(10), 6104-6108. doi:10.1073/pnas.1031527100
- Pearlman, B. L., & Traub, N. (2011). Sustained virologic response to antiviral therapy for chronic hepatitis C virus infection: a cure and so much more. *Clin Infect Dis*, 52(7), 889-900. doi:10.1093/cid/cir076
- Penin, F., Brass, V., Appel, N., Ramboarina, S., Montserret, R., Ficheux, D., . . . Moradpour, D. (2004). Structure and function of the membrane anchor domain of hepatitis C virus nonstructural protein 5A. *J Biol Chem*, 279(39), 40835-40843. doi:10.1074/jbc.M404761200
- Penin, F., Combet, C., Germanidis, G., Frainais, P. O., Deleage, G., & Pawlotsky, J. M. (2001). Conservation of the conformation and positive charges of hepatitis C virus E2 envelope glycoprotein hypervariable region 1 points to a role in cell attachment. *J Virol*, 75(12), 5703-5710. doi:10.1128/JVI.75.12.5703-5710.2001
- Penin, F., Dubuisson, J., Rey, F. A., Moradpour, D., & Pawlotsky, J. M. (2004). Structural biology of hepatitis C virus. *Hepatology*, 39(1), 5-19. doi:10.1002/hep.20032
- Pieroni, L., Santolini, E., Fipaldini, C., Pacini, L., Migliaccio, G., & La Monica, N. (1997). In vitro study of the NS2-3 protease of hepatitis C virus. *J Virol*, 71(9), 6373-6380.
- Pietschmann, T., Kaul, A., Koutsoudakis, G., Shavinskaya, A., Kallis, S., Steinmann, E., . . . Bartenschlager, R. (2006). Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. *Proc Natl Acad Sci U S A*, 103(19), 7408-7413. doi:10.1073/pnas.0504877103
- Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., . . . Abrignani, S. (1998). Binding of hepatitis C virus to CD81. *Science*, 282(5390), 938-941.
- Pockros, Paul j. (2016). Direct-acting antivirals for the treatment of hepatitis C virus

infection. *UptoDate*.

- Pol, S., Corouge, M., & Vallet-Pichard, A. (2016). Daclatasvir-sofosbuvir combination therapy with or without ribavirin for hepatitis C virus infection: from the clinical trials to real life. *Hepat Med*, 8, 21-26. doi:10.2147/HMER.S62014
- Poynard, T., Yuen, M. F., Ratziu, V., & Lai, C. L. (2003). Viral hepatitis C. *Lancet*, 362(9401), 2095-2100.
- Prince, A. M. (1968). An antigen detected in the blood during the incubation period of serum hepatitis. *Proc Natl Acad Sci U S A*, 60(3), 814-821.
- Prince, A. M., Brotman, B., Grady, G. F., Kuhns, W. J., Hazzi, C., Levine, R. W., & Millian, S. J. (1974). Long-incubation post-transfusion hepatitis without serological evidence of exposure to hepatitis-B virus. *Lancet*, 2(7875), 241-246.
- Putnam, C. D., Hammel, M., Hura, G. L., & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys*, 40(3), 191-285. doi:10.1017/S0033583507004635
- Rambo, R. P., & Tainer, J. A. (2013). Accurate assessment of mass, models and resolution by small-angle scattering. *Nature*, 496(7446), 477-481. doi:10.1038/nature12070
- Raney, K. D., Sharma, S. D., Moustafa, I. M., & Cameron, C. E. (2010). Hepatitis C virus non-structural protein 3 (HCV NS3): a multifunctional antiviral target. *J Biol Chem*, 285(30), 22725-22731. doi:10.1074/jbc.R110.125294
- Reiss, S., Harak, C., Romero-Brey, I., Radujkovic, D., Klein, R., Ruggieri, A., . . . Lohmann, V. (2013). The lipid kinase phosphatidylinositol-4 kinase III alpha regulates the phosphorylation status of hepatitis C virus NS5A. *PLoS Pathog*, 9(5), e1003359. doi:10.1371/journal.ppat.1003359
- Reiss, S., Rebhan, I., Backes, P., Romero-Brey, I., Erfle, H., Matula, P., . . . Bartenschlager, R. (2011). Recruitment and activation of a lipid kinase by hepatitis C virus NS5A is essential for integrity of the membranous replication compartment. *Cell Host Microbe*, 9(1), 32-45. doi:10.1016/j.chom.2010.12.002
- Reviakine, I., Rossetti, F. F., Morozov, A. N., & Textor, M. (2005). Investigating the properties of supported vesicular layers on titanium dioxide by quartz crystal microbalance with dissipation measurements. *J Chem Phys*, 122(20), 204711. doi:10.1063/1.1908500
- Rice, C. (2011). Perspective: miles to go before we sleep. *Nature*, 474(7350), S8. doi:10.1038/474S8a
- Robertson, B., Myers, G., Howard, C., Bretin, T., Bukh, J., Gaschen, B., . . . Weiner, A. (1998). Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. International Committee on Virus Taxonomy. *Arch Virol*, 143(12), 2493-2503.

- Romero-Brey, I., & Bartenschlager, R. (2014). Membranous replication factories induced by plus-strand RNA viruses. *Viruses*, 6(7), 2826-2857. doi:10.3390/v6072826
- Romero-Brey, I., Berger, C., Kallis, S., Kolovou, A., Paul, D., Lohmann, V., & Bartenschlager, R. (2015). NS5A Domain 1 and Polyprotein Cleavage Kinetics Are Critical for Induction of Double-Membrane Vesicles Associated with Hepatitis C Virus Replication. *MBio*, 6(4), e00759. doi:10.1128/mBio.00759-15
- Romero-Brey, I., Merz, A., Chiramel, A., Lee, J. Y., Chlanda, P., Haselman, U., . . . Bartenschlager, R. (2012). Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. *PLoS Pathog*, 8(12), e1003056. doi:10.1371/journal.ppat.1003056
- Ross-Thriepland, D., & Harris, M. (2014). Insights into the complexity and functionality of hepatitis C virus NS5A phosphorylation. *J Virol*, 88(3), 1421-1432. doi:10.1128/JVI.03017-13
- Ross-Thriepland, D., & Harris, M. (2015). Hepatitis C virus NS5A: enigmatic but still promiscuous 10 years on! *J Gen Virol*, 96(Pt 4), 727-738. doi:10.1099/jgv.0.000009
- Sackmann, E. (1996). Supported membranes: scientific and practical applications. *Science*, 271(5245), 43-48.
- Sapay, N., Montserret, R., Chipot, C., Brass, V., Moradpour, D., Deleage, G., & Penin, F. (2006). NMR structure and molecular dynamics of the in-plane membrane anchor of nonstructural protein 5A from bovine viral diarrhea virus. *Biochemistry*, 45(7), 2221-2233. doi:10.1021/bi0517685
- Saw, W. G., Tria, G., Gruber, A., Subramanian Manimekalai, M. S., Zhao, Y., Chandramohan, A., . . . Gruber, G. (2015). Structural insight and flexible features of NS5 proteins from all four serotypes of Dengue virus in solution. *Acta Crystallogr D Biol Crystallogr*, 71(Pt 11), 2309-2327. doi:10.1107/S1399004715017721
- Scarselli, E., Ansuini, H., Cerino, R., Roccasecca, R. M., Acali, S., Filocamo, G., . . . Vitelli, A. (2002). The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J*, 21(19), 5017-5025.
- Scheel, T. K., Gottwein, J. M., Mikkelsen, L. S., Jensen, T. B., & Bukh, J. (2011). Recombinant HCV variants with NS5A from genotypes 1-7 have different sensitivities to an NS5A inhibitor but not interferon-alpha. *Gastroenterology*, 140(3), 1032-1042. doi:10.1053/j.gastro.2010.11.036
- Scheel, T. K., & Rice, C. M. (2013). Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nat Med*, 19(7), 837-849. doi:10.1038/nm.3248
- Schlutter, J. (2011). Therapeutics: new drugs hit the target. *Nature*, 474(7350), S5-7. doi:10.1038/474S5a

- Serebrov, V., & Pyle, A. M. (2004). Periodic cycles of RNA unwinding and pausing by hepatitis C virus NS3 helicase. *Nature*, 430(6998), 476-480. doi:10.1038/nature02704
- Sharma, N. R., Mateu, G., Dreux, M., Grakoui, A., Cosset, F. L., & Melikyan, G. B. (2011). Hepatitis C virus is primed by CD81 protein for low pH-dependent fusion. *J Biol Chem*, 286(35), 30361-30376. doi:10.1074/jbc.M111.263350
- Shirota, Y., Luo, H., Qin, W., Kaneko, S., Yamashita, T., Kobayashi, K., & Murakami, S. (2002). Hepatitis C virus (HCV) NS5A binds RNA-dependent RNA polymerase (RdRP) NS5B and modulates RNA-dependent RNA polymerase activity. *J Biol Chem*, 277(13), 11149-11155. doi:10.1074/jbc.M111392200
- Shlomai, A., de Jong, Y. P., & Rice, C. M. (2014). Virus associated malignancies: the role of viral hepatitis in hepatocellular carcinoma. *Semin Cancer Biol*, 26, 78-88. doi:10.1016/j.semcancer.2014.01.004
- Simmonds, P. (2004). Genetic diversity and evolution of hepatitis C virus--15 years on. *J Gen Virol*, 85(Pt 11), 3173-3188. doi:10.1099/vir.0.80401-0
- Skou, S., Gillilan, R. E., & Ando, N. (2014). Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat Protoc*, 9(7), 1727-1739. doi:10.1038/nprot.2014.116
- Smith, D. B., Bukh, J., Kuiken, C., Muerhoff, A. S., Rice, C. M., Stapleton, J. T., & Simmonds, P. (2014). Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology*, 59(1), 318-327. doi:10.1002/hep.26744
- Song, Y., Friebe, P., Tzima, E., Junemann, C., Bartenschlager, R., & Niepmann, M. (2006). The hepatitis C virus RNA 3'-untranslated region strongly enhances translation directed by the internal ribosome entry site. *J Virol*, 80(23), 11579-11588. doi:10.1128/JVI.00675-06
- Stanaway, J. D., Flaxman, A. D., Naghavi, M., Fitzmaurice, C., Vos, T., Abubakar, I., . . . Cooke, G. S. (2016). The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. *Lancet*, 388(10049), 1081-1088. doi:10.1016/S0140-6736(16)30579-7
- Stevens, F. J. (1989). Analysis of protein-protein interaction by simulation of small-zone size exclusion chromatography. Stochastic formulation of kinetic rate contributions to observed high-performance liquid chromatography elution characteristics. *Biophys J*, 55(6), 1155-1167. doi:10.1016/S0006-3495(89)82912-1
- Sun, J. H., O'Boyle, D. R., 2nd, Fridell, R. A., Langley, D. R., Wang, C., Roberts, S. B., . . . Gao, M. (2015). Resensitizing daclatasvir-resistant hepatitis C variants by allosteric modulation of NS5A. *Nature*, 527(7577), 245-248. doi:10.1038/nature15711
- Svergun, D., Barberato, C., & Koch, M. H. J. (1995). CRY SOL - a Program to

- Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *JOURNAL OF APPLIED CRYSTALLOGRAPHY*, 28(6), 768-773. doi:doi:10.1107/S0021889895007047
- Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, 25(4), 495-503. doi:doi:10.1107/S0021889892001663
- Tabaei, S. R., Choi, J. H., Haw Zan, G., Zhdanov, V. P., & Cho, N. J. (2014). Solvent-assisted lipid bilayer formation on silicon dioxide and gold. *Langmuir*, 30(34), 10363-10373. doi:10.1021/la501534f
- Tanaka, T., Kato, N., Cho, M. J., Sugiyama, K., & Shimotohno, K. (1996). Structure of the 3' terminus of the hepatitis C virus genome. *J Virol*, 70(5), 3307-3312.
- Targett-Adams, P., Graham, E. J., Middleton, J., Palmer, A., Shaw, S. M., Lavender, H., . . . Westby, M. (2011). Small molecules targeting hepatitis C virus-encoded NS5A cause subcellular redistribution of their target: insights into compound modes of action. *J Virol*, 85(13), 6353-6368. doi:10.1128/JVI.00215-11
- Tellinghuisen, T. L., Foss, K. L., Treadaway, J. C., & Rice, C. M. (2008). Identification of residues required for RNA replication in domains II and III of the hepatitis C virus NS5A protein. *J Virol*, 82(3), 1073-1083. doi:10.1128/JVI.00328-07
- Tellinghuisen, T. L., Marcotrigiano, J., Gorbalenya, A. E., & Rice, C. M. (2004). The NS5A protein of hepatitis C virus is a zinc metalloprotein. *J Biol Chem*, 279(47), 48576-48587. doi:10.1074/jbc.M407787200
- Tellinghuisen, T. L., Marcotrigiano, J., & Rice, C. M. (2005). Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. *Nature*, 435(7040), 374-379. doi:10.1038/nature03580
- Tellinghuisen, T. L., & Rice, C. M. (2002). Interaction between hepatitis C virus proteins and host cell factors. *Curr Opin Microbiol*, 5(4), 419-427.
- Tokuriki, N., & Tawfik, D. S. (2009). Protein dynamism and evolvability. *Science*, 324(5924), 203-207. doi:10.1126/science.1169375
- Trepo, C. (2014). A brief history of hepatitis milestones. *Liver Int*, 34 Suppl 1, 29-37. doi:10.1111/liv.12409
- Tscherne, D. M., Jones, C. T., Evans, M. J., Lindenbach, B. D., McKeating, J. A., & Rice, C. M. (2006). Time- and temperature-dependent activation of hepatitis C virus for low-pH-triggered entry. *J Virol*, 80(4), 1734-1741. doi:10.1128/JVI.80.4.1734-1741.2006
- Vestergaard, B. (2016). Analysis of biostructural changes, dynamics, and interactions - Small-angle X-ray scattering to the rescue. *Arch Biochem Biophys*, 602, 69-79. doi:10.1016/j.abb.2016.02.029

- Vieyres, G., Dubuisson, J., & Pietschmann, T. (2014). Incorporation of hepatitis C virus E1 and E2 glycoproteins: the keystones on a peculiar virion. *Viruses*, 6(3), 1149-1187. doi:10.3390/v6031149
- Volkov, Vladimir V., & Svergun, Dmitri I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *JOURNAL OF APPLIED CRYSTALLOGRAPHY*, 36(3 Part 1), 860-864. doi:doi:10.1107/S0021889803000268
- Wakita, T., Pietschmann, T., Kato, T., Date, T., Miyamoto, M., Zhao, Z., . . . Liang, T. J. (2005). Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med*, 11(7), 791-796. doi:10.1038/nm1268
- WHO. (2016). *Global Health Sector Strategy on Viral Hepatitis 2016–2021: Towards Ending Viral Hepatitis*. Retrieved from Geneva, Switzerland: <http://apps.who.int/iris/bitstream/10665/246177/1/WHO-HIV-2016.06-eng.pdf?ua=1>
- Wolk, B., Sansonno, D., Krausslich, H. G., Dammacco, F., Rice, C. M., Blum, H. E., & Moradpour, D. (2000). Subcellular localization, stability, and trans-cleavage competence of the hepatitis C virus NS3-NS4A complex expressed in tetracycline-regulated cell lines. *J Virol*, 74(5), 2293-2304.
- Yi, M., & Lemon, S. M. (2003). Structure-function analysis of the 3' stem-loop of hepatitis C virus genomic RNA and its role in viral RNA replication. *RNA*, 9(3), 331-345.
- Zan, G. H., Jackman, J. A., & Cho, N. J. (2014). AH peptide-mediated formation of charged planar lipid bilayers. *J Phys Chem B*, 118(13), 3616-3621. doi:10.1021/jp411648s
- Zuckerman, A. J. (1996). Hepatitis Viruses. In S. Baron (Ed.), *Medical Microbiology* (4th ed.). Galveston (TX).