

May 2018

Interactions of Viral and Cellular Helicases

Megan Josephine Corby
University of Wisconsin-Milwaukee

Follow this and additional works at: <https://dc.uwm.edu/etd>

 Part of the [Biochemistry Commons](#), [Biology Commons](#), and the [Biophysics Commons](#)

Recommended Citation

Corby, Megan Josephine, "Interactions of Viral and Cellular Helicases" (2018). *Theses and Dissertations*. 1777.
<https://dc.uwm.edu/etd/1777>

This Dissertation is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

INTERACTIONS OF VIRAL AND CELLULAR HELICASES

by

Megan Josephine Corby

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Chemistry

at

The University of Wisconsin at Milwaukee

May 2018

ABSTRACT

INTERACTIONS OF VIRAL AND CELLULAR HELICASES

by

Megan Josephine Corby

The University of Wisconsin-Milwaukee, 2018
Under the Supervision of Professor David Frick

The innate immune system is a part of the first line of defense against virus infection. An important subset of the innate immune system consists of a group of intracellular pattern recognition receptors (PRRs) which recognize conserved features of bacteria and viruses and initiate an interferon response. The RIG-I like receptors (RLRs) are PRRs that bind to RNA viruses (such as hepatitis c virus) and signal through the adaptor mitochondrial anti-viral signaling protein (MAVS).

Hepatitis C virus (HCV) is a small enveloped RNA virus that belongs to the *flaviviridae* family of viruses. HCV infects hepatocytes and can cause a persistent infection. If a chronic infection is established, progressive liver damage along with cirrhosis and sometimes hepatocellular carcinoma may occur. The multi-functional HCV non-structural-3 (NS3) protein is essential for HCV replication and contains covalently linked protease and helicase/ATPase domains. A covalently linked protease and helicase is unique to the *flaviviridae* family of viruses and it is unclear why the two functions are linked. There are multiple effective direct acting anti-virals which target the protease, but none currently approved which inhibit the helicase. In addition to aiding in viral replication, the NS3 protease assists HCV in establishing a persistent infection through cleaving the innate immune RIG-I adaptor protein, MAVS. The purpose of the

studies contained in this thesis is to gain a greater understanding of the function and purpose of the covalently linked HCV NS3 protease and helicase. Förster resonance energy transfer (FRET) is used to explore the interaction of NS3 with RIG-I like receptor proteins and to use FRET to look at the interaction of the RLRs with themselves.

The interaction of the NS3 protease and helicase domain was probed through the exploration of the mechanism of action of a NS3 inhibitor (HPI) which is able to inhibit both the protease and helicase functions of NS3, while not disrupting the ATPase activity. The activity of HPI was determined *in vitro* using a fluorescent protease cleavage assay and a fluorescent helicase unwinding assay. HPI can inhibit both functions with low micro-molar EC_{50} . Next, analysis of HPI to inhibit peptide hydrolysis by wild-type NS3 and a set of NS3 mutants with mutations in the protease domain, helicase domain, and the allosteric groove between the protease and helicase domain suggested that HPI forms a bridge between the NS3 helicase RNA-binding site and the allosteric groove between the protease and helicase domains. The activity of HPI was measured in cells using an HCV sub-genomic replicon tagged with a luciferase reporter. The inhibition of HPI alone and in the presence of other protease inhibitors was tested. HPI can inhibit the HCV genotype 1b sub-genomic replicon and when applied in conjunction with first generation protease inhibitors, telaprevir and boceprevir, the inhibition was additive, as defined by the Bliss Independence Model of additive inhibition. However, when HPI was used in conjunction with macro-cyclic protease inhibitors, danoprevir and grazoprevir, modest synergy was observed.

To look at the protein:protein interactions of the NS3 helicase and the RIG-I like receptor helicases in live cells, a series of quantitative FRET spectrometry studies were employed. Quantitative micro-spectroscopic imaging (Q-MSI) is a technique which uses a fluorescent dye or fluorescent protein to identify sub-cellular regions and then calculates Förster Resonance Energy Transfer (FRET) efficiency and the concentrations of the donor and acceptor proteins. The technique was first applied *in vitro* with a fluorescently tagged NS3 helicase and fluorescently tagged DNA molecules. Next, the technique was applied to combinations of recombinant fluorescently tagged helicases expressed in HEK293T cells. The NS3 helicase, RIG-I like receptor helicases, DDX1, DDX3, and DDX5 helicases, and MAVS were all designed to express off plasmids which also encode and attach a fluorescent protein. The fluorescent proteins used were either cyan fluorescent protein (CFP), enhanced green fluorescent protein-2 (GFP₂), yellow fluorescent protein (YFP) or Venus fluorescent protein and each combination included a donor (CFP or GFP₂) and an acceptor (YFP or Venus) fluorescent protein. The combinations were tested in presence or absence of polyinosinic-polycytidylic acid (poly I:C) which is a synthetic RNA analog capable of eliciting an RLR response. To localize the interaction to the mitochondria, the mitochondrial stain, Mito-Tracker-Red, was used in some experiments.

The experiments revealed a previously unknown interaction between NS3 and the RLR protein, laboratory of genetics and physiology protein-2 (LGP2) which may be biologically relevant. In addition, the relocation of LGP2 cytoplasmic foci in cells over-expressing DDX3 was observed. Q-MSI was used to visualize previously known interactions of RLRs at the mitochondria and in conjunction with MAVS.

Dedicated to Tate, Nora, and, soon to be, new baby.

TABLE OF CONTENTS

Chapter 1 Introduction	1
Chapter 2 Literature review	15
2.1 Helicases	17
2.2 Hepatitis C virus	20
2.2.1 HCV viral lifecycle	22
2.2.2 Viral Proteins	23
2.2.3 Inhibitors	27
2.2.4 HCV NS3:NS4A.....	30
2.3 Innate Immune System.....	34
2.3.1 RIG-I like receptors (RLRs).....	35
2.3.2 RIG-I.....	39
2.3.3 MDA5.....	41
2.3.4 LGP2.....	42
2.3.5 HCV and detection by RIG-I like receptors.....	43
2.4 Other DExD/H-box proteins.....	45
2.4.1 DDX3.....	46
2.4.2 DDX1.....	47
2.4.3 DDX5.....	48
2.5 FRET Microscopy.....	48
2.5.1 Principles	48
2.5.2 FRET calculation using filter based measurements	51
2.5.3 Acceptor photo-bleaching.....	56
2.5.4 Fluorescence lifetime measurements.....	57

2.5.5 Optical Microspectroscopy	58
2.5.7 FRET Spectrometry	59
Chapter 3 Methods	65
3.1 Molecular Cloning.....	65
3.1.1 pCMV-YFP-LGP2	65
3.1.2 PCMV-CFP-LGP2	65
3.1.3 PCMV-YFP-DDX1.....	66
3.1.4 PCMV-YFP-DDX3.....	66
3.1.5 PCMV-YFP-DDX5.....	67
3.1.6 PCMV-CFP-NS3:NS4A	67
3.1.7 pCMV-CFP-SCNS3	67
3.1.8 pCMV-CFP-NS3-HEL1B	68
3.1.9 pCMV-CFP-NS3	68
3.1.10 PCMV-CFP-NS3-PRO	68
3.1.11 P24-CFP-NS3HEL.....	69
3.1.12 P24-YFP-NS3HEL.....	69
3.1.13 pCMV-YFP-RIG-I.....	70
3.1.14 pCMV-CFP-RIG-I	70
3.1.15 pCMV-YFP-MDA5	70
3.1.16 pCMV-CFP-MDA5	70
3.1.17 pCMV-GFP ₂ -MAVS.....	70
3.1.18 EF1-mVenus-MDA5	70
3.1.19 pCMV-GFP ₂ -LGP2	70
3.1.20 pCMV-GFP ₂ (T203I)-MAVS	71

3.1.21 pCMV-Sapphire	71
3.1.22 pCMV-GFP ₂	71
3.2 Protein Purification	71
3.3 Enzyme Assays	72
3.3.1 Helicase	72
3.3.2 Protease	72
3.3.3 Protein:DNA titrations.....	73
3.4 Cell-based Assays.....	78
3.4.1 HCV subgenomic replicon assays.....	78
3.4.1.1 Combination studies	79
3.4.1.2 In-vivo assay data analysis	79
3.4.2 Cell Viability Assays.....	80
3.4.3 Transfection with plasmids	80
3.4.4 Intracellular protease assays.....	81
3.4.5 Transfections with polyinosinic–polycytidylic acid (Poly I:C).....	81
3.4.6 ISRE Assays	81
3.5 Microscopy	82
3.5.1 Choosing FRET pairs	82
3.5.2 Imaging set up and Mito-Tracker staining	84
3.5.3 Collection of microscopic images.....	85
3.5.4 Un-mixing of spectral images.....	86
3.5.5 E _{app} (FRET) calculation	87
3.5.6 Donor and acceptor concentration calculation	90
3.5.7 Pixel level FRET calculation	91

Chapter 4 A Novel Inhibitor Can Inhibit Both Helicase and Protease Activities of HCV NS3 Protease/helicase.....	92
4.1 HPI is specific to HCV genotypes 1b, 3a, and 4a.....	93
4.2 Effect of HPI on NS3 harboring mutations between the protease and helicase domains .	95
4.3 Effect of HPI on the activity of scNS3 and NS3:NS4A	96
4.4. Synergy between HPI and macrocyclic protease inhibitors in sub-genomic replicons.....	97
4.5 Significance of the experimental results on the mechanism for HPI mode of inhibition .	100
Chapter 5 Interactions between NS3 helicase and nucleic acids	104
5.1 Measuring the interaction of HCV NS3 with oligonucleotides using FRET: A comparison of three different FRET calculation techniques	106
Chapter 6 The N-terminal membrane anchor domain of NS4A is needed to localize the hepatitis C virus NS3 protease to biologically active foci	111
6.1 The NS4A membrane anchor is necessary to localize the NS3 complex to sub-cellular foci	113
6.2 The NS3 protease co-factor domain of NS4A is necessary for proteolytic cleavage of an NS3 substrate but the portion encoding the membrane anchor is not necessary for protease activity.	115
6.3 Significance of the requirement of the NS4A co-factor for proteolytic cleavage in cells and the necessity of the NS4A anchor domain to localize NS3	117
Chapter 7 The Hepatitis C virus helicase interacts with the human pattern recognition receptor LGP2	118
7.1 Recombinant fluorescent RNA helicases are biologically active when expressed in HEK293T cells.....	121
7.2 FRET analysis of interaction between HCV NS3 helicase domain and RLRs	125
7.3 Q-MSI analysis of the interaction of NS3 with LGP2 in the presence and absence of the PAMP poly I:C	135
7.4 Significance of LGP2 and NS3 interaction	137
Chapter 8 DDX3 causes LGP2 to re-localize into foci.....	139
8.1 Evidence that DDX3 recruits the RIG-I-like receptor LGP2 to Foci.....	140
Chapter 9 Homo and hetero interactions of RIG-I like receptors in response to poly I:C.....	144

9.1 Sub-cellular analysis of the homo-interactions of the CFP and YFP tagged RIG-I like receptors in response to poly I:C.....	146
9.2 Sub-cellular analysis of the hetero-interactions of the CFP and YFP tagged RIG-I like receptors in response to poly I:C.....	152
9.3 Analysis of the hetero-interactions of RIG-I receptors at the pixel level.....	154
9.3.1 Pixel level FRET analysis of the interaction of GFP ₂ -LGP2 and Venus-MDA5	156
9.3.2 Pixel level FRET analysis of the interaction of MAVs with RIG-I like receptors	161
9.4 Significance of the interactions of the RIG-I like receptors in response to poly I:C.....	173
Chapter 10 The two excitation peaks of the wild-type GFP protein and the development of a novel GFP variant.....	176
10.1 Protein concentration calculations and the gamma ratio.....	179
10.2 Development of a novel GFP variant.....	188
Chapter 11 Conclusions and Future Directions	192
References.....	203
Curriculum Vitae.....	223

LIST OF FIGURES

Figure 2-1 The structure of an enveloped virus.....	15
Figure 2-2: Helicases belong to several super families (SF)..	18
Figure 2-3 The HCV genome is comprised of ~ 9.6 kb genome which encodes for both structural and non-structural proteins... ..	20
Figure 2-4: The HCV genome is cleaved by both host and viral proteases... ..	21
Figure 2-5 Approved inhibitors currently in use for the treatment of HCV listed in conjunction with the protein they target.....	29
Figure 2-6: HCV helicase translocates unidirectionally along single stranded HCV RNA displacing the duplex strand.....	32
Figure 2-7: Inchworm model of HCV NS3 helicase motion.....	33
Figure 2-8: Schematic of RIG-I like receptor function.....	37
Figure 3-1: Comparison of fluorescent protein spectrums.....	82
Figure 3-2: Alignment of wild-type GFP (avGFP), Clontech's eGFP, and GFP ₂ protein sequences.	83
Figure 3-3 Fluorescence intensity map of the donor (A), acceptor (B), and Mito-Tracker (C).	88
Figure 3-4: Calculation of the gamma ratio for CFP and YFP.	89
Figure 4-1 HPI is specific to HCV genotypes 1b, 3a, and 4a.....	93
Figure 4-2 Evaluation of putative HPI binding site.....	95
Figure 4-3: Interactions between HPI and other protease inhibitors.....	98
Figure 4-4 HCV NS3 can cleave substrates <i>in cis</i> and <i>in trans</i>	100
Figure 4-5 Proposed hypothesis for the mechanism of action of HPI.....	102
Figure 5-1 NS3 helicase binding monitored through the quenching of the intrinsic protein fluorescence.....	104

Figure 5-2: Titration of a YFP labeled NS3 helicase with a MAX fluorophore labeled single-stranded DNA.....	109
Figure 6-1: Alignment of cyan fluorescent protein tagged NS3 recombinant protein sequences.	113
Figure 6-2: CFP-NS3 variants expressed in HEK 293T cells.. ..	114
Figure 6-3 : CFP-NS3 encodes a functional protease when expressed in HEK 293T cells.....	116
Figure 7-1 Quantitative Micro-spectroscopic imaging of donor and acceptor labeled recombinant fluorescent fusion proteins in live cells.	120
Figure 7-2: Protein sequence alignment of YFP RIG-I, YFP MDA5, and YFP LGP2.	122
Figure 7-3: Recombinant fluorescent fusion proteins retain biological activity when expressed in HEK293T cells.	124
Figure 7-4 Analysis of FRET helicase interactions using whole cell regions of interest.....	127
Figure 7-5: Sub-cellular average concentrations over selected regions of interest in the analysis of helicase interactions.	128
Figure 7-6 HEK 293T cells were transfected with each of the NS3 domain deletion mutants. ...	129
Figure 7-7: Concentrations of CFP tagged NS3 variants and YFP-LGP2 when co-expressed in HEK293T cells.	131
Figure 7-8: Representative images of CFP-NS3-protease expressed in cells.	132
Figure 7-9: Sub-cellular NS3:NS4A region analysis of NS3 helicase interaction with RIG-I like receptors.	134
Figure 7-10 Analysis of NS3-4A interaction with LGP2 in the presence and absence of poly I:C.	135
Figure 8-1 Co-overexpression of DDX3 and LGP2 causes DDX3 and LGP2 to accumulate in cytoplasmic foci	141
Figure 9-1 Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing donor, acceptor, and the sub-cellular mitochondrial tag, MitoTracker.....	147
Figure 9-2 Response of mitochondrial localized RIG-I, MDA5, and LGP2 to poly I:C exposure.	149
Figure 9-3: Response of mitochondrial localized LGP2 and MDA5 to poly I:C.. ..	152

Figure 9-4: Unmixed k^{DA} , k^{AD} , MitoTracker and E_{app} (FRET) 2D intensity maps.	157
Figure 9-5: Pixel level FRET analysis of sub-cellular regions of cells co-expressing GFP ₂ -LGP2 and Venus-MDA5.	158
Figure 9-6: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP ₂ -MAVS and YFP-RIG-I.	163
Figure 9-7: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP ₂ -MAVS and YFP-RIG-I.	164
Figure 9-8: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP ₂ -MAVS and YFP-MDA5.	164
Figure 9-9: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP ₂ -MAVS and YFP-MDA5.	165
Figure 9-10: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP ₂ -MAVS and Venus-MDA5.	166
Figure 9-11: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP ₂ -MAVS and Venus-MDA5.	168
Figure 9-12: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP ₂ -MAVS and YFP-LGP2.	169
Figure 9-13: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP ₂ -MAVS and YFP-LGP2.	171
Figure 10-1: The GFP chromophore is 4-(<i>p</i> -hydroxybenzylidene)imidazolidin.	176
Figure 10-2: Emission of GFP ₂ overlaps YFP and Venus absorption spectra.	178
Figure 10-3: GFP ₂ expressed in HEK293T cells generates a wide range of gamma ratio values.	181
Figure 10-4: eGFP does not change much with pH but GFP ₂ shows a dramatic change in excitation spectrum with pH change.	183
Figure 10-5: Effect of multiple laser excitations on eGFP and GFP ₂ at varying pH.	184
Figure 10-6: Alignment of the protein sequences of wild type GFP, eGFP, GFP ₂ , GFP ₂ (T203I), and Sapphire.	188

Figure 10-7: Excitation scans of GFP₂(T203I)-MAVS, Sapphire, and GFP₂ expressed in cells. 190

Figure 11-1: Model of HCV NS3 interacting with the RLR oligomer to aid in the cleavage of
MAVS..... 201

Figure 11-2: Model for LGP2:MDA5 oligomer.. 202

LIST OF TABLES

Table 4-1 The ability of HPI and other direct acting antivirals to inhibit the HCV genotype 1b subgenomic replicons and Huh7.5 cell viability.. 97

Table 5-1: Results of FRET calculation using each FRET calculation technique for the YFP-NS3h to MAX-DNA titration. 110

Table 9-1: Distances calculated using molecular dynamic simulations between fluorescent proteins in the RIG-I:MAVS complex..... 162

Table 10-1: Summary of different amino acid substitutions in wild-type GFP (WT GFP), eGFP, GFP₂, GFP₂(T203I), and mT-Sapphire. 189

LIST OF ABBREVIATIONS

ATP	Adenosine tri-phosphate
CARD	Capsase activation and recruitment domain
CFP	Cyan fluorescent protein
CTD	C-terminal domain
DNA	Deoxyribonucleic acid
dsRNA	Double-stranded RNA
E1	Envelope-1
E2	Envelope-2
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein
HCV	Hepatitis C virus
HEK293	Human embryonic kidney cells 293
HPI	Helicase-protease inhibitor
IFN- β	Interferon- β
IRES	Internal ribosomal entry site
IRF-3	Interferon regulatory factor-3
LGP2	Laboratory of genetics and physiology-2
MAVS	Mitochondrial anti-viral signaling protein
MDA5	Melanoma differentiated antigen-5
mRNA	Messenger-ribonucleic acid
NS2	Non-structural protein 2
NS3	Non-structural protein 3

NS4A	Non-structural protein 4A
NS4B	Non-structural protein 4B
NS5A	Non-structural protein 5A
NS5B	Non-structural protein 5B
PAMP	Pattern associated molecular pattern
Poly I:C	Polyinosinic:polycytidylic acid
PRR	Pattern recognition receptor
Q-MSI	Quantitative microspectroscopic imaging
RIG-I	Retinoic inducible gene-I
RLR	RIG-I like receptor
RNA	Ribonucleic acid
ROI	Region of interest
SF1	Super-family 1
SF2	Super-family 2
siRNA	Small interfering ribonucleic acid
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
SVR	Sustained virologic response
TBK1	Tank-binding kinase 1
TLR4	Toll-like receptor 4
TRIF	TIR-domain-containing adapter-inducing interferon- β
tRNA	Transfer-ribonucleic acid
UTR	Un-translated region
YFP	Yellow fluorescent protein

ACKNOWLEDGEMENTS

I would like to acknowledge my mentor, Dr. David Frick, without whom this thesis would not be possible. His help and support throughout my degree program has been invaluable and he has taught me how to be a rigorous and critical scientist. He taught me about the world of an academic scientist and components of pursuing a career in academia.

I would like to acknowledge the help of Dr. Valerica Raicu in understanding FRET and fluorescent microscopy. His willingness to let me participate in his group meetings and to let me help organize the Microspectroscopy conference has increased my understanding of the field of fluorescent microscopy and my understanding of FRET spectrometry.

A special thank you to the rest of my committee. I feel like I received help from each member of my committee at some point in my degree program and I am very grateful for their support. In alphabetical order: Dr. Alexander Arnold, Dr. Nicholas Silvaggi, and Dr. Doug Steeber.

I would like to thank the following scientists for their help (both personally and professionally) throughout my degree program. In alphabetical order: Dr. Gabriel Biener, Dr. Sergei Kuchin, Dr. Sourav Mukherjee, Dr. Jean Ndjomou, Joel Paprocki, Alica Schultz, and Dr. Mike Stoneman.

A special thank you to University of Wisconsin-Milwaukee (UWM) and the UW-Milwaukee graduate school advanced opportunity fellowship (AOP). The AOP fellowship was instrumental in allowing me to pursue my studies and my research free from the demands of outside work.

I would like to acknowledge my parents for their continued support academically and personally. I would like to acknowledge my partner, Tate, for always standing by me and helping me to think positive when things are hard. Lastly, I would like to acknowledge my daughter Nora and her (soon-to-be) little sister for helping me to prioritize the things that matter. You are my everything.

Chapter 1 Introduction

Hepatitis C virus (HCV) is a member of the *flaviviridae* family of RNA viruses. It can replicate without producing symptoms causing a persistent infection which results in progressive liver damage and sometimes liver cirrhosis or hepatocellular carcinoma. Currently, there are up to 103 million people chronically infected world-wide (Ansaldi et al., 2014). The blood supply was first tested for HCV in the early 1990s and therefore risk factors for HCV include a blood transfusion prior to 1992 along with intravenous drug usage and unsterile medical practices. There are six HCV genotypes and more sub-types. HCV genotype 1 is the most common in the United States and one of the harder genotypes to treat. Therefore, in addition to a serum antibody test and HCV RNA measurement, diagnosis also includes determination of viral genotype, sub-type, and resistance-associated mutations. By determining the specific genotype and subtype and the presence of any resistance associated mutations, it is possible to choose the best course of treatment (Ansaldi et al., 2014).

Hepatitis C virus RNA encodes both structural and non-structural proteins (Figure 2-3). Structural proteins are those which make up the HCV virion structure and packaging. The structural proteins include core, envelope 1 (E1), and envelope 2 (E2). The non-structural proteins are involved in HCV replication and establishing an infection. The non-structural proteins include the accessory protein p7, the cysteine protease NS2, the covalently linked protease/helicase NS3, the protease co-factor NS4A, the inter-membrane accessory proteins NS4B and NS5A, and the RNA polymerase NS5B. The current standard of care for HCV infection

involves a combination of direct acting antivirals which target non-structural proteins involved in crucial steps in the HCV lifecycle. Presently, FDA approved direct acting antivirals target the NS3:NS4A protease, the NS5A protein and the RNA-dependent RNA polymerase, NS5B. By combining 2-3 of these anti-viral drugs, over 90% of patients (regardless of genotype) are cured (Ansaldi et al., 2014).

Despite the resounding success of current direct acting anti-virals, it is still not possible to cure 100% of those infected with HCV and an error prone NS5B polymerase means that drug resistant strains are constantly evolving. Additionally, the cost of HCV treatment is high and may be cost prohibitive for some individuals because insurance companies are not always willing to cover direct-acting anti-viral treatments despite proof of long term cost effectiveness. Therefore, many HCV treatments still include the use of interferon which results in un-wanted side-effects and reduces treatment adherence treatment strategies (Elana S Rosenthal and Camilla S Graham, 2016). Furthermore, HCV is closely related to other viruses such as Zika or Dengue fever and therefore understanding more about the HCV life-cycle can impact not only HCV treatment, but also the treatment of like-viruses. One area of viral research, and HCV research specifically, which leaves much to be discovered is the ability of viruses to manipulate and control host immune responses. Gaining a better understanding of the interplay between virus and host could lead to better host-directed therapies and more effective treatments.

Viral research dates back to the late 19th century where Dmitry Ivanovsky demonstrated that a crushed sample of tobacco leaves infected with tobacco mosaic virus retained its ability to infect other tobacco plants even after being filtered with a Chamberlain filter to remove

bacteria (Lechevalier, 1972). An understanding of immunity dates back even farther to the 5th century B.C.E. where Thucydides observed and wrote about individuals who recovered from the plague being able to care for those currently infected with the disease without becoming re-infected (Ranger, 1995). Today, the fields of immunology and virology have progressed, however, the burden of viral illness is ever present. Viruses find ways to mutate and evade the best developed therapies and yet most people recover from viral infections without the aid of drugs or therapies. The immune system itself is incredibly effective at targeting and eradicating viruses. Understanding the immune response to viruses is an important component to viral research and one which has yielded new therapeutic directions such as treatments for Influenza-A which seek to limit the immune system derived inflammation that causes lung damage (Ramos and Fernandez-Sesma, 2015) or host-directed therapies which improve treatment to HIV by delivering recombinant interferons to stimulate a beneficial host response (Kaufmann et al., 2017). The host's own immune system is the best tool for fighting viral replication and a greater understanding of the immune system and its interplay with disease could lead to the development of new therapeutics for treating pathogens (Kaufmann et al., 2017; Ramos and Fernandez-Sesma, 2015; Wodarz and Lloyd, 2004).

The human immune system is complex. Aspects of the immune system from physical and mucosal barriers to specialized pathogen targeting proteins are present in every part of the human body and are indispensable to human survival. The immune system recognizes and neutralizes known targets and those which it has never encountered, all with the goal of preserving health and homeostasis. One way to subdivide the components of the immune system is to separate them into the innate immune system and the adaptive immune system.

The innate immune system is a non-specific system present in all classes of plant and animal life and important to human survival. It responds to pathogens in a generic way through physical barriers or through identifying generic features of pathogens not present in the host. But unlike the adaptive immune system, the innate immune system does not provide any immunological memory. Despite its lack of specificity, the innate immune system is essential to survival and includes major functions such as producing cytokines to recruit other immune cells to the area of an infection, activation of the complement cascade to identify and clear foreign or damaged material in the body, white blood cells to engulf and remove pathogens, physical and mucosal barriers to infectious agents such as the skin and lining of the gut, specialized intra-cellular proteins which recognize when a cell has been infected with a pathogen, and a critical connection to the adaptive immune response through antigen presentation (Alberts et al., 2002).

A sub-set of innate immune system intra-cellular proteins are called the RIG-I like receptors (RLRs) (John S. Errett Michael Gale Jr, 2015; Y. M. Loo and Gale, 2011). Members of this family include, RIG-I, MDA5, and LGP2 (Y. M. Loo and Gale, 2011). The RLRs are part of a larger group of innate immune system proteins known as pattern recognition receptors (PRRs) (Yoneyama et al., 2015). PRRs recognize conserved features of pathogens and initiate a signaling response which results in the production of cytokines (Bruns and Horvath, 2012; Marques et al., 2006; Pichlmair et al., 2015). RIG-I like receptors specifically recognize viral RNA in the cytoplasm of cells. One key component to proper RLR function is their ability to recognize only those RNAs associated with a viral infection and not ones which may be a part of normal cellular function (such as mRNAs, tRNAs, or siRNAs). To separate the host RNAs from the viral

RNAs, RLRs only respond to those RNAs which have patterns absent in host RNAs. Some examples include long double-stranded RNAs and RNAs lacking the 5' 7-methyl-guanylate cap. Once the RLRs have recognized their particular ligand, they oligomerize and interact with the adaptor protein, mitochondrial anti-viral signaling protein (MAVS). MAVS transmits the signal to downstream kinases which leads to the translocation of interferon regulatory factor 3 (IRF-3) to the nucleus to up-regulate the production of interferon. Interferon then aids in localizing other immune cells to the area, including those involved in antigen presentation and the adaptive immune system's immunological memory.

The adaptive immune system is a specific system which is highly specialized to identify and kill known pathogens. Whereas the innate immune system can function and/or respond almost immediately to a perceived threat, the adaptive immune system often requires time, sometimes days, to operate at full capacity. This is due to the highly specific nature of the adaptive immune system which responds with a tailored response to a specific pathogen based off of its immunological memory (Owen et al., 2013). The adaptive immune system is constantly changing as the host encounters and survives pathogenic diseases. With each immune challenge, a new set of antibodies are acquired and the adaptive immune system changes to prepare to meet future challenges. Some of the major roles of the adaptive immune system involve recognizing "self" from "non-self" antigens during antigen presentation, developing immunological memory (antibodies) to retain specific information related to the identification of pathogens through memory B and T cells, and mounting specific attacks on pathogens or pathogen-infected cells (Alberts et al., 2002; Owen et al., 2013).

The immune system is effective at combatting pathogenic infection and many pathogens have developed ways to manipulate or alter the immune response in order to facilitate their infection. One example of pathogenic evasion is achieved through antigenic drift whereby a virus mutates in response to immune system antibodies, so that it is now undetectable by currently circulating antibodies. An example of antigenic drift is the flu. Each year new strains of flu emerge which have slightly different surface antigens than those previously seen and every year people are infected or die from the flu, even though they most likely have contracted the flu in the past and have an immunological memory (antibodies) of some influenza virus antigen (Anderson et al., 2017). In addition to evading the adaptive immune response, several pathogens have established ways to evade the innate immune response. Innate immune response evasion is particularly important for establishing an infection since the innate immune system may be the first system encountered by an infectious pathogen.

One of the initial innate immune system barriers encountered by a pathogen are the epithelial cells of the skin and gut (Owen et al., 2013). The epithelial cells secrete antimicrobial peptides such as defensins, gramicidins, and cathelicidins which disrupt bacterial membrane integrity. One way that bacteria are able to avoid these antimicrobial peptides is through a simple avoidance mechanism characterized by electrostatic repulsion. For example, *Staphylococcus aureus* evades cationic defensin molecules by adding a lysine residue to its outer membrane (Peschel et al., 2001). In addition, some bacteria, such as *Bacillus anthracis* secrete proteases which cleave defensins and/or other anti-microbial peptides. If a bacterium successfully breaches the physical barrier, the bacterium must still overcome pattern

recognition receptors in order to survive. In response to these challenges, some bacteria have developed ways to alter their molecular structure such that it is not recognizable to the PRRs. For example, lipopolysaccharide (LPS) is a core component of many bacterial cell walls and it is recognized by the PRR, toll like receptor 4 (TLR4). But some bacteria have chemically modified the LPS in their cell wall, such that it is no longer recognizable by TLR4 (Reddick and Alto, 2014).

Many viruses also employ techniques to evade or disrupt detection by PRRs. For example, by localizing replication to sub-cellular compartments inaccessible to cytosolic PRRs, dengue fever and influenza A are able to replicate undetected (Onomoto et al., 2012; Ramos and Fernandez-Sesma, 2015). In addition, many viruses disrupt PRR signaling through cleaving adaptor proteins. For example, hepatitis C virus (HCV) cleaves the adaptor proteins mitochondrial anti-viral signaling protein (MAVS) and TIR-domain containing adaptor inducing interferon- β (TRIF), thus cutting off the pattern recognition signaling and allowing the virus to establish a persistent infection (Doehle and Michael Gale, 2013; Hoving et al., 2014; Y. Loo and Gale, 2007). To gain a greater understanding of the function of the RLRs, a study of RLR protein:protein interactions is beneficial.

Several techniques can be used to look at protein:protein interactions. Analyzing proteins which co-purify with proteins expressing an epitope tag, such as a glutathione s-transferase (GST) tag, is a commonly used technique (Scopes, 2010). The proteins pulled down with an epitope tag can be identified using activity assays, antibodies, or through protein mass spectrometry. Co-immunoprecipitation is another example of a technique used to identify interacting partners. It involves isolating the protein of interest with a specific antibody and

then identifying the interacting partners via Western blotting or protein mass spectrometry (Ehle and Horn, 1990). Some drawbacks to the technique are that an antibody specific to the protein is needed, weak interactions may disintegrate once the protein is separated from the cell and it gives no information as to the architecture of the protein:protein complex. Another technique used to identify protein:protein interactions is bimolecular fluorescence complementation which involves tagging two proteins of interest with one half of a complementary fluorescent protein. If the two halves of the fluorescent protein come close enough to each other, they interact and spontaneously fold to form a complete and functional fluorescent protein. If the two halves are not near each other, no complementation can occur and the halves remain non-fluorescent. Therefore if there is no interaction between tagged proteins, there is no fluorescence. If there is interaction, a fluorescent signal can sometimes be observed (Miller et al., 2015). This technique can provide information not only on whether or not an interaction occurs but also demonstrate under what conditions the interaction may not take place (i.e. in presence of an inhibitor). It can be performed in live cells but it only gives a yes/no answer. It does not provide any information as to the subtleties of the dynamics or arrangements of protein:protein interactions and it may miss temporary or weak interactions.

Förster resonance energy transfer (FRET) (Lakowicz, 2010) is another technique for exploring protein:protein interaction. FRET involves the non-radiative transfer of energy between a donor fluorescent tag and an acceptor fluorescent tag. The FRET quantity determined from experiments is the FRET efficiency of the number of occurrences of non-radiative energy transfer from donor to acceptor. If molecular complexes being measured are made up of several donors or acceptors (i.e. a mixture of monomers, dimers, trimers, etc) then

an apparent FRET efficiency (E_{app}) is measured (Raicu and Schmidt, 2017). E_{app} includes all the contributions from donors and acceptors within a given system.

The energy transferred between donor and acceptor is dependent on distance. Therefore, the extent to which there is a transfer of energy between the donor and acceptor molecules can give an indication as to the distance between donor and acceptor. By tagging two proteins of interest with a fluorescent FRET pair designed so that the emission of the donor excites the acceptor, it is possible to not only judge whether the fluorescently tagged proteins are interacting but also to get an idea of the distance between the fluorescent molecules (Lakowicz, 2010). There are many different experimental procedures and methods of analysis which can be used in the calculation of FRET and a few of the options will be discussed in more detail in Chapter 2. However, many of them are fraught with the same issues. Specifically, FRET analysis can often be confounded by direct excitation of the acceptor at the peak donor excitation wavelength, giving the appearance of FRET when there is actually none. If the spectrum of the donor is broad, there may also be a contribution of donor to the acceptor emission wavelength. In addition, some FRET methods (such as acceptor-photo bleaching) cause destruction to the acceptor molecule after just one measurement which significantly limits the amount of data which can be collected from any given sample. Other FRET techniques (such as three-filter sensitized emission) involve the use of multiple excitation scans to determine FRET. The time that passes in between excitation scans allows molecular diffusion to occur within the cell. The molecular diffusion which occurs between excitation scans means that the molecules imaged in the first excitation scan are not the same as the molecules measured in the second excitation scan and therefore FRET can only be calculated as an

average over a larger area as opposed to over smaller pixel level areas. (Clegg, 1995; Lakowicz, 2010; Piston and Kremers, 2007; Selvin, 2000).

A FRET method which seeks to address some of the issues with the calculation of FRET is optical microspectroscopy and FRET spectrometry (Raicu et al., 2009; Raicu and Singh, 2013). Optical microspectroscopy uses a single, two-photon excitation, and spectral imaging to collect the full fluorescent emission spectrum of donor and acceptor molecules. FRET spectrometry incorporates fully quantitative analysis to calculate FRET (E_{app}) and look at the interactions of proteins in living cells. Each component of optical microspectroscopy was specifically chosen to maximize the advantages and minimize the disadvantages of FRET measurement and analysis.

Optical microspectroscopy uses carefully chosen donor and acceptor fluorophores and a single excitation scan to collect all the information necessary to calculate FRET. The use of a single excitation scan is important because if multiple images must be acquired to calculate FRET then the molecules have time to diffuse from one image to another. Specifically, the molecules measured in the first image are not the same molecules measured in the second image. The molecular diffusion that occurs between multiple excitation scans means that FRET cannot be calculated at the pixel level and instead average FRET over larger regions of interest is considered. If only one excitation scan is used to calculate FRET, then FRET may be calculated at pixel level. In addition, by only needing to acquire a single image, transient interactions, that might be lost in the time it takes to excite multiple times, are able to be captured. The use of a two-photon excitation (as opposed to single-photon excitation) minimizes damage to the

sample and reduce fluorophore bleaching. Two-photon excitation allows delicate, biological samples to be measured multiple times with minimal damage.

Spectral imaging is beneficial because it means that a full fluorescence spectra at single nanometer resolution is collected for each measurement. The resulting “mixed” spectra are deconvoluted (i.e. unmixed) with spectra obtained from cells expressing either donor alone or acceptor alone. Assuming that the mixed spectra is a linear combination of the contributions of each individual spectra, factors are determined that scale the fluorescence of the donor in the presence of acceptor and acceptor in the presence of donor to fit the measured composite spectra. By separating out the individual component spectrums from an overall composite spectrum, any contribution of the donor to the acceptor spectrum or the acceptor to the donor spectrum is removed. The result of spectrally unmixing fluorescent images in a FRET spectrometry experiment is the generation of 2D fluorescent intensity maps which locate the donor and acceptor signal within the sample being measured. In addition, supplementary spectrum (such as auto-fluorescence, background light, or additional fluorescent markers) may also be unmixed and considered in the analysis. By considering the entire fluorescent spectrum (as opposed to just a peak intensity or band of wavelengths), spectral imaging allows for the calculation of FRET closer to the true theoretical underpinnings of the theory. Specifically, that FRET represents the fraction of excitations transferred from donors to unexcited acceptors per total donor excitations. Using a single, two-photon excitation and spectral imaging, enough information is captured so that FRET (E_{app}) may be calculated for each pixel within a fluorescent microscopic image using FRET spectrometry. The specific equations involved will be discussed in greater detail in Chapter 2.

A variation on FRET spectrometry, termed fully quantitative spectral imaging (FSI), expanded on the FRET spectrometry technique to include the use of a second excitation wavelength (King et al., 2016; Stoneman et al., 2017). Like FRET spectrometry, FSI uses the first excitation wavelength to specifically excite the donor which would then transfer its energy to a nearby acceptor. A second excitation wavelength is then chosen to primarily excite the acceptor. The second excitation allows for the calculation of the fluorescence of the acceptor in the absence of donor. The fluorescence of the acceptor alone can then be used to back-calculate the fluorescence of the donor emission within the sample in the absence of acceptor. The details of this technique will be discussed in greater detail in Chapter 2 and 3. The benefit of obtaining the fluorescence of the acceptor in the absence of donor and the fluorescence of the donor in the absence of acceptor is two-fold. The first advantage is that by assuming that the fluorescence of a fluorescent molecule will increase linearly with concentration, a standard curve can be constructed which relates fluorescence intensity to concentration. This standard curve can be used in every measurement to calculate the concentrations of the donor and acceptors within each sample. In addition, the back-calculation of the donor in the absence of acceptor requires a correction for any direct excitation of the acceptor. This allows for the calculation of E_{app} between a donor and acceptor even when there is significant direct excitation of the acceptor. Although the ability to correct for direct excitation of the acceptor expands the application of FSI to systems where the donor and acceptor spectrum significantly overlap, it also comes with a tradeoff. Specifically, if two excitation wavelengths are used in the calculation of E_{app} then two excitation scans must be collected to calculate FRET. If two excitation scans are needed, then the time it takes to acquire those two scans allows molecular

diffusion to occur such that the molecules imaged in the first scan are different from those measured in the second scan (King et al., 2016; Patowary et al., 2015). Specifics of the theory and calculation of E_{app} will be discussed further in Chapters 2 and 3.

Quantitative microspectroscopic imaging (Q-MSI) (Corby et al., 2017) follows up on the technique of FSI and FRET spectrometry (King et al., 2016; Raicu and Singh, 2013) and incorporates the use of a fluorescent marker to identify sub-cellular structures. The third fluorescent marker is easily unmixed from the composite spectra and generates a 2D fluorescence intensity map of the desired sub-cellular structure. Sub-cellular regions of interest (ROIs) can then be selected within the unmixed 2D fluorescence intensity map representing the location of the sub-cellular structure being analyzed. The ROIs selected on the fluorescent marker intensity map are applied as a mask to the 2D fluorescent intensity maps generated through the unmixing of the composite spectrum in FRET spectrometry to specifically calculate E_{app} within the sub-cellular region of interest. This technique is particularly helpful when the localization of the fluorescent donor and acceptor molecules is not apparent from visual analysis alone.

This thesis seeks to explore the function of the NS3 protease and helicase domain through its interaction with a novel compound which inhibits both functions. Next, fluorescently labeled HCV NS3 and a fluorescently labeled oligonucleotide is used to calculate FRET *in vitro* using several different techniques including Q-MSI. After which Q-MSI is applied *in vivo* to the homo and hetero interactions of HCV NS3, DDX3, DDX1, DDX5, RIG-I, MDA5, LGP2, and MAVS. A greater understanding of the function of the NS3 protein in the maintenance of

HCV and the inhibition of the innate immune system could inspire new therapies aimed at working with the immune system. In addition, a further understanding of the protein:protein interactions of the innate immune system proteins could lead to a greater understanding of the protein's role in preventing disease and allow therapies to be developed which enhance the innate immune system function.

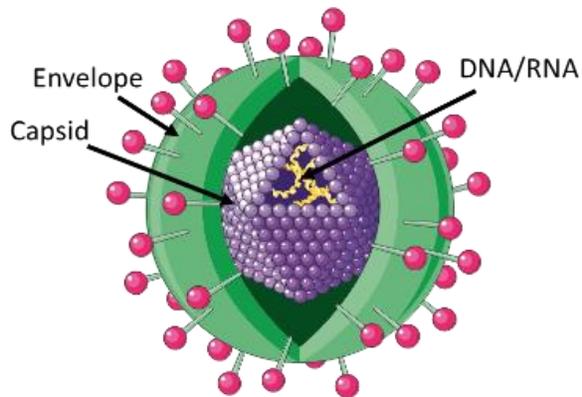


Figure 2-1 The structure of an enveloped virus. The **envelope** coats the virus and aids in transmission. A **capsid** is sometimes present and serves to protect the viral genome. The **viral genome** can be made of DNA or RNA and transmits the genetic information necessary for replication. This figure was adapted with permission. This work is licensed under the Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Viruses are small and lack the necessary components for independent life and therefore viruses are dependent on a host to replicate. Viruses are able to infect all Kingdoms of Life including bacteria, Archaea, plants, and humans (Breitbart and Rohwer, 2005). A viral genome encodes viral proteins which can typically be separated into structural and non-structural components. The structural components provide the packaging for the virus. All viruses have at least a structural capsid (protein coat) which

surrounds and protects the genetic material during transmission. This protein coat can be made up of several different proteins which fit together to form a cohesive layer. Some viruses also have an envelope which surrounds the capsid and aids in transmission (Figure 2-1). There can be several different proteins involved in making one envelope structure.

The structural proteins surround the genetic material and aid in the transmission of the virion. The genetic material contained in the virion may be either DNA or RNA. In viruses like hepatitis C virus (HCV) or human immunodeficiency virus (HIV), the RNA genome encodes for the production of non-structural proteins which make up some of the machinery necessary for

viral replication. This machinery include enzymes such as polymerases, helicases, integrases, and proteases, as well as accessory proteins. The additional enzymes and materials (such as ATP) required for replication are derived from the host cell. Viruses require assistance from the host and many viruses (including HIV and HCV) co-opt host enzymes to aid in viral replication (Gelderblom, 1996; Lodish et al., 2000).

An additional function of some viral non-structural proteins is to evade the host immune response. In order to replicate, many viruses have the ability to mute the host's immune response. Some ways viral proteins achieve this are by direct interaction with intracellular immune system proteins, cleavage of key proteins involved in signal transmission, or through disguising their genomes to better resemble host components (i.e. adding a 5' 7-methyl-guanylate RNA cap or 2' O-methylation) (Hale et al., 2010; Lucas et al., 2001; Picard-Jean et al., 2013)

There are many different viruses which infect humans and cause illness ranging from the common cold to chicken pox and cold sores. Some viruses cause a disruption in homeostasis (and subsequently the onset of sickness) through cell lysis or disruption of cell activities. Lytic viruses rupture and kill the cell host and if enough cells die, the host suffers. Non-lytic viruses can transmit to other organisms without killing the cell, but the presence of viral infection can disrupt cell processes and inhibit normal function by inducing an immune response which shuts down host protein translation and function. The symptoms associated with virus infection; fever, vomiting, diarrhea, malaise, and rash, are often caused by the host immune response attempting to rid the body of the virus. Some viruses are able to replicate without inducing an

immediate immune response and these viruses can establish persistent infections that are difficult or impossible to cure. Persistent viral infections may be harmless or silently deadly. In addition, some viruses (including those that replicate without immune detection) have been linked to cancer (Bowie and Unterholzner, 2008).

2.1 Helicases

Helicases are enzymes indispensable to all living organisms and many viruses. Almost all biological functions involving RNA or DNA include at least one helicase. Helicases are motor proteins fueled by ATP hydrolysis which are able to bind and translocate directionally along nucleic acids. Some helicases also have the ability to unwind and/or remodel the nucleic acid (Patel and Donmez, 2006). Defective helicases are linked to many diseases such as cancer, neurodegenerative diseases, and developmental defects (Abdelhaleem, 2004; Hanada and Hickson, 2007; Nakayama, 2002).

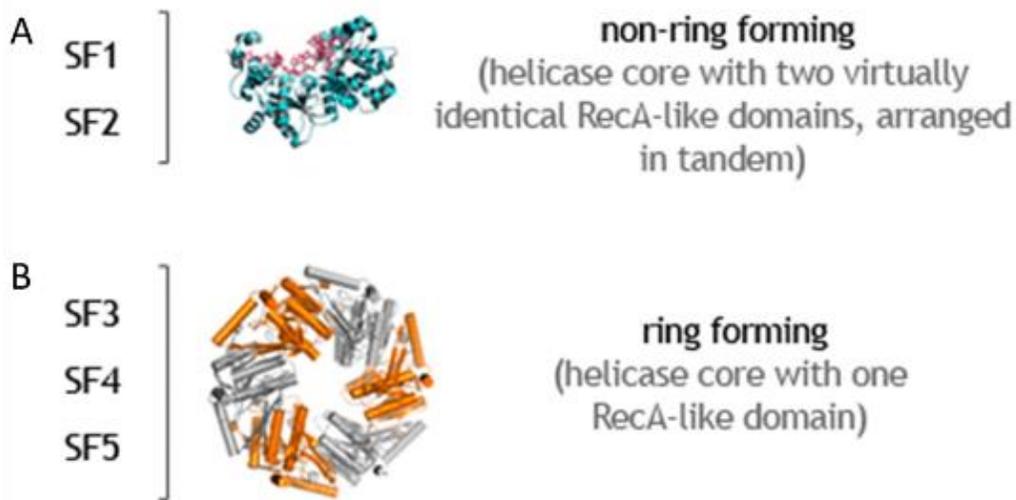


Figure 2-2: Helicases belong to several super families (SF). (A) SF 1 and SF2 contain tandem RecA-like domains and are able to function as monomers. (B) SF3, 4, and 5 are obligate oligomers who form hexameric rings. Altered with permission from the RNA helicase database (Jankowsky et al., 2011).

Helicases are comprised of six super-families. Superfamilies 3-6 form large hexameric rings and are obligate oligomers (Figure 2-2B) whereas superfamilies 1 and 2 can work as monomers (Figure 2-2A). One functional difference between the obligate oligomers and those which are monomers, arises from the presence or absence of a complete ATP binding site. Helicases, such as DnaB, form hexameric rings because each sub-unit has only half of the ATP binding site necessary for ATP hydrolysis. These ATP binding domains are termed RecA-like domains due to their similarity to E.coli's RecA protein which is involved in DNA recombination. Each monomer in a hexameric helicase contains only one RecA-like domain and oligomerization is necessary to create a functional ATP binding site. In contrast, helicases from superfamilies 1 and 2 contain tandem RecA-like domains and therefore are able to function as monomers (Fairman-Williams et al., 2010; Jankowsky et al., 2011; Ye et al., 2004). Despite the ability of SF

1 and 2 helicases to function as monomers, there is evidence that some helicases become better enzymes when they oligomerize. These include UvrD and RecBCD which form a homo or hetero-dimer respectively. Other helicases show functional cooperativity when multiple helicases are loaded onto one strand. These helicases have increased processivity as more helicases bind (Patel and Donmez, 2006).

Helicases can be classified as either DNA or RNA helicases based on their preferential substrate. Some specifically prefer RNA or DNA whereas others target DNA-RNA duplexes. DNA helicases are involved in cellular processes such as DNA replication, recombination and repair and also encoded by DNA viral genomes such as the herpes simplex virus (Chattopadhyay et al., 2006). RNA helicases are involved in RNA metabolism, innate immune signaling, and are encoded by many RNA viruses that replicate in the cytoplasm, such as hepatitis C virus. Many RNA helicases are DExD/H box proteins and belong to superfamily 2. The name derives from the single letter amino acid code describing their highly conserved ATPase site located in motif II. The DExD/H box proteins are further classified into sub-groups based on subtle differences in their conserved amino acid motifs. Examples of the different sub-groups include DEAD-box (DDX) and DEAH-box (DHX) proteins (Bird et al., 1998; Fairman-Williams et al., 2010). Despite the sequence similarity of their catalytic core, the function of these DExD/H box proteins is diverse and the specificity of function determined by sites outside of the amino acid motif after which they are named (Rocak and Linder, 2004; Subramanya et al., 1996).

2.2 Hepatitis C virus

Hepatitis C virus (HCV) is a blood borne pathogen which is able to reproduce after entering hepatocytes. It was identified in 1988 and blood donor testing began in 1990 (Houghton, 2009). By 1992, most of the blood supply was considered free of HCV, yet those who received blood transfusions prior to 1992, were at significant risk for having the disease (Lauren Gravitz, 2011; QL Choo et al., 1989). Today, about 170 million people are infected worldwide with baby boomers and intravenous drug users being the populations most likely to be infected. The hepatitis C virus causes a slowly progressing liver disease that if left untreated causes cirrhosis, hepatocellular carcinoma, and liver failure (Ansaldi et al., 2014).

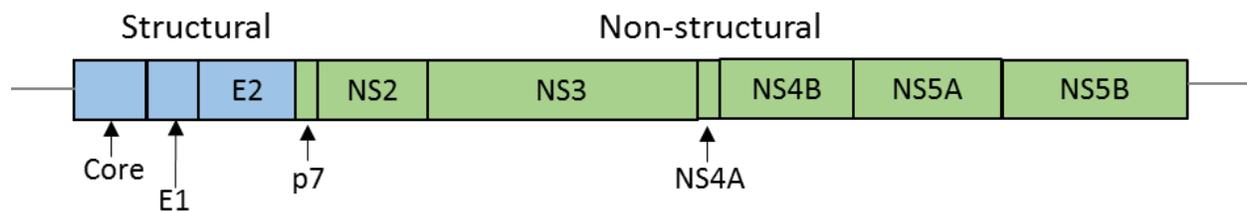


Figure 2-3 The HCV genome is comprised of ~ 9.6 kb genome which encodes for both structural and non-structural proteins. The structural proteins include Core, E1, and E2. The non-structural proteins are p7, NS2, NS3, NS4A, NS4B, NS5A, NS5B.

HCV is a member of the *Flaviviridae* family of viruses and has a positive-sense RNA genome encoding a single long open reading frame. The RNA encodes for a ~3,000 amino acid polyprotein. The poly-peptide is cleaved by host and viral proteases into structural (core, E1,

and E2) and nonstructural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (Figure 2-3) (M. Gu and Rice, 2013).

Structural proteins are processed by host proteases and self-processing begins when NS2 (a cysteine protease) cleaves the junction between NS2/NS3 and releases the NS3 N-terminal serine protease domain to cleave the remaining poly-protein. The NS3 protease cleaves the junction between NS3 and NS4A *in-cis* and then proceeds to cleave the remaining NS4B, NS5A, and NS5B *in trans* (Figure 2-4).

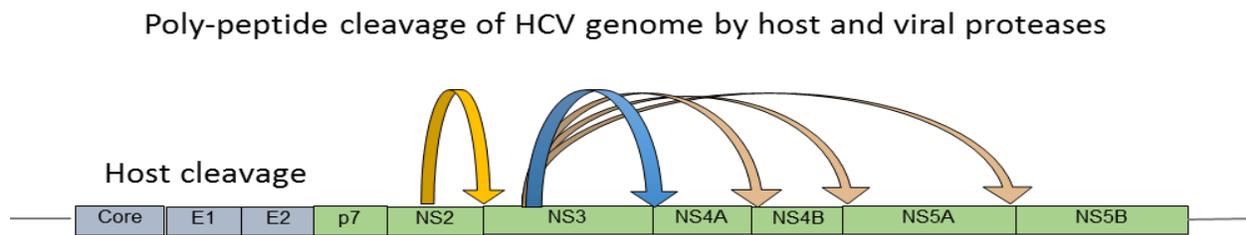


Figure 2-4: The HCV genome is cleaved by both host and viral proteases. Host proteases are responsible for cleaving the junctions between the structural proteins; Core, E1, E2, and p7. NS2 cleaves the junction between NS2 and NS3 *in cis*. The NS3 protease then cleaves the junction between NS3 and NS4 *in cis* and the junction between the remaining non-structural proteins *in trans*.

NS4A is an accessory protein which associates with the NS3 protease and contains the membrane anchors necessary to localize it to the ER membrane bound viral replicase complex. NS4B is involved in the remodeling of the ER membrane and intercalates into the membrane via several trans-membrane regions. NS5A is a zinc-binding phosphoprotein and NS5B is the RNA-dependent RNA-polymerase with C-terminal membrane anchors (Salam and Akimitsu, 2013; Scheel and Rice, 2013). HCV reproduces and evolves rapidly because its RNA-dependent RNA polymerase (NS5B) lacks proofreading activity. Quasi-species evolve in HCV patients, and

hepatitis C viruses are now comprised of a diverse set of 7 genotypes (1-7), and several subtypes (i.e. 1a, 1b), whose sequences differ by typically 20-25% (Simmonds, 2004).

2.2.1 HCV viral lifecycle

Viral entry into the host cell consists of multiple interactions including attachment, entry, and fusion. Viral particle uptake occurs via clathrin-mediated endocytosis (Emmanuelle Blanchard et al., 2006). Once inside the cytoplasm, the HCV virion fuses with cellular endosomes to release the viral RNA into the cell. The IRES (internal ribosomal entry site) found in the 5' UTR (untranslated region) of the HCV genome mediates the translation by recruiting viral and host translation initiation factors in a cap-independent fashion. The resulting polypeptide is about 3,000 amino acids in length and consists of both structural and non-structural proteins. The poly protein is further processed by host and viral factors (Chevaliez and Pawlotsky, 2006; Kim and Chang, 2013).

The poly-peptide processing occurs in a membranous replication web closely associated with the ER membrane and containing vesicles with HCV non-structural proteins, HCV RNA, and lipid droplets. Replication in these webs consists of NS5B making negative sense copies of the positive sense RNA genome to produce further positive sense genomes which are the template for translation (Scheel and Rice, 2013). Several host proteins are also recruited to the membranous webs to aid in HCV replication, including DDX3 which undergoes a dramatic re-localization to HCV replicase complexes mediated by the HCV structural protein core (Angus et al., 2010).

The poly-peptide is cleaved initially by host signal peptidases which targets the poly-protein to the ER membrane and which yields the mature structural viral proteins. The zinc-dependent NS2 cysteine protease cleaves between the NS2-NS3 juncture *in-cis*, freeing the NS3 serine protease to cleave between NS3:NS4A *in-cis* and then the remaining NS4A-NS4B, NS4B-NS5A, and NS5A-NS5B *in-trans* (Scheel and Rice, 2013).

HCV virions are formed with an icosahedral capsid and have with a host-derived lipid membrane embedded with E1 and E2 (envelope 1 and 2) proteins. Inside the lipid bi-layer envelope is a nucleocapsid composed of core protein which surrounds the HCV RNA and protect it before it moves on to the next host cell (Chevaliez and Pawlotsky, 2006; Kim and Chang, 2013).

2.2.2 Viral Proteins

Core Protein

Core protein forms the viral nucleocapsid with the viral RNA in the virion. Core is a 21-kDa protein that has lipid and RNA binding ability. Core behaves as a membrane protein and has two distinct domains, D1 and D2. D1 is a hydrophilic domain with primarily positively charged amino acid residues and D2 is a hydrophobic domain with primarily negatively charged amino acid residues (Angus et al., 2010). Little is known about how core forms the proposed icosahedral nucleocapsid because, so far, it has not been possible to replicate this assembly outside of cells. Reconstituted core particles form irregular structures which are much larger than those isolated from HCV infection (Gawlik and Gallay, 2014; Kunkel et al., 2001).

E1 and E2 Glycoproteins

E1 and E2 are cleaved from the poly-peptide via host signal peptidases. They contain multiple possible glycosylation sites and are often modified by N-linked glycans. These envelope glycoproteins are essential for HCV entry and fusion in a host cell and therefore represent a good target for anti-viral therapies or vaccines. E1 and E2 exist in heterodimers and the formation of this heterodimer is essential for efficient viral entry (Freedman et al., 2016). The E1 protein is indispensable for interacting with cell surface receptors and membrane fusion (Douam et al., 2014).

P7

P7 is a 63 amino acid membrane protein. P7 intercalates into the ER membrane and both the N and C terminus are oriented toward the ER lumen. P7 monomers have been crystalized as hexamers or heptamers and are believed to function as cation-selective channels. Therefore p7 is classified as a viroporin (Gentzsch et al., 2013). Although its role in the life-cycle is unclear, it has been proposed that p7 acts as an ion channel pump which helps to maintain a pH suitable for viral protein development in the membranous web replicase complex. The absence of p7 decreases the infectivity of some HCV strains (Atoom et al., 2014; Dubuisson, 2007).

NS2

The NS2 protein is the HCV N-terminal non-structural protein. It is a cysteine protease which cleaves the junction between NS2-NS3 *in-cis* however it is not required for the replication

of HCV subgenomic replicons (Lohmann et al., 1999). There is some thought that NS2 may also be involved in viral packaging and release. This is due to several mutational studies which show that in the absence of a functional E1 and E2 or core, NS2 is not properly localized to lipid droplet replication complexes. In addition, NS2 shows sequence similarity to other viral proteins (Welbourn and Pause, 2007) which function in packaging and release of viruses. NS2 is a short lived protein that is quickly degraded in the replication complex in a phosphorylation dependent manner (Popescu et al., 2011; Welbourn and Pause, 2007).

NS3

The NS3 protein is comprised of a covalently linked protease and helicase domain, both of which are essential to viral replication (Lin, 2006). The details of NS3 structure and function will be discussed in section 2.2.4.

NS4A

NS4A is a 54 amino acid protein with an N-terminal hydrophobic alpha helix transmembrane segment. The middle part of NS4A is a co-factor for the NS3 serine protease and contributes to the proper positioning of the proteolytic cleavage active site. The C-terminal portion of NS4A forms an alpha helix at low pH. The NS4A protein is critical for HCV replication and localizes the NS3 protease to the ER membrane via the trans-membrane helices. In addition to its activity as a protease co-factor, NS4A also serves to enhance the function of the NS3 helicase by aiding in the proper positioning of the protease catalytic domain (Beran et al., 2009). Further, mutations in NS4A cause problems in virus particle assembly indicating diverse roles for NS4A throughout the viral life-cycle (Phan et al., 2011).

NS4B

NS4B is a highly hydrophobic 27 kDa protein which contains four transmembrane domains. It can exist as a monomer but there is also evidence that it could function as a part of larger, multi-protein complexes with NS4A and NS5A or NS3 and NS5B (Neddermann et al., 1999). NS4B localizes to the ER membrane replication complexes and has been shown to be closely associated with lipid rafts, the putative site for replicase formation. Therefore, it is hypothesized that NS4B may be a critical protein for establishing the replicase complexes. In addition, NS4B has been shown to negatively regulate the NS5B RNA-dependent-RNA (RdRp) polymerase activity *in-vitro* giving further evidence to the idea that NS4B may function in larger multi-protein complexes. Lastly, NS4B is a critical component of the HCV induced interferon effect with the NS4B mutation Q1737H being found in multiple interferon treatment resistant phenotypes (Namba et al., 2004; Sklan and Glenn, 2006).

NS5A

NS5A is an RNA binding protein which contains three distinct domains and integrates into the membrane via a transmembrane helix. Domain 1 is essential for replication and has been crystallized in multiple configurations, including a dimer. Domains 2 and 3 are unfolded and their precise roles have yet to be elucidated (Fridell et al., 2011). NS5A is highly phosphorylated with a basal phosphorylation form and a hyper-phosphorylated form. It is proposed that the difference in phosphorylation may help regulate its function and its interaction with other NS3 non-structural proteins. For example, NS5A can interact with NS5B and change its RNA polymerase activity (Shirota et al., 2002). In addition, multiple mutations

which confer increased replicative efficiency have been found in the NS5A domain and some overlap the phosphorylation sites, inferring that the extent of phosphorylation may control replication (He et al., 2006). The inhibition of NS5A is one of the most effective treatments for HCV and the NS5A inhibitor, Ledipasvir, is included in the currently prescribed and highly effective treatment, Harvoni®.

NS5B

NS5B is an RdRp essential for copying the positive sense HCV genome and making a molar excess of the template strands necessary for translation. Similar to other RdRps, NS5B has a “right hand” formation with a finger, palm, and thumb subdomains. The amino acid residues necessary for nucleotidyl transfer are found in the palm domain and recombinant NS5B was sufficient to catalyze the formation of full length HCV RNA *in vitro*. This implies that NS5B also has the ability to replicate through the stable secondary and tertiary RNA structures found in the untranslated regions of HCV RNA (Ferrari et al., 1999). NS5B lacks proofreading activity and is very error prone, thus contributing to the high mutation rate of HCV (Ranjith-Kumar and Kao, 2006). One of the current HCV treatments, Harvoni®, includes sofosbuvir, a NS5B inhibitor.

2.2.3 Inhibitors

Up until a few years ago, the standard of care for HCV was a combination of pegylated interferon and ribavirin. Although able to completely cure some cases of chronic HCV at an efficiency between 40-90% percent depending on genotype, the side effects made treatment

difficult for the patient and caused low treatment adherence. In particular, genotype 1 (the most prevalent genotype in the United States) is hard to treat with the IFN/RBV regimen and only resulted in about a 40% cure rate, even with extended treatment. Side effects to IFN treatment include flu-like symptoms, bone marrow depression and autoimmune syndromes which contribute to feelings of fatigue, myalgia, depression, loss of appetite, and weight loss. Hemolytic anemia is the most common side effect of ribavirin treatment (Gerlach et al., 2003; Manns et al., 2006; Santantonio et al., 2005).

Currently, the best treatment for HCV is a combination therapy made up of direct acting antivirals (DAAs). DAAs target the HCV non-structural proteins directly and offer the best treatment outcomes (including genotype 1) with the least number of side effects. The first-generation direct acting anti-virals include telaprevir and boceprevir which target the NS3 protease. Combination therapy with a protease inhibitor and PEG-IFN/RBV improves the sustained virological response (SVR) (i.e. an undetectable viral load for 12 weeks) amongst treatment naïve genotype 1 patients to 70%. However, variants resistant to both telaprevir and boceprevir develop rapidly and the presence of side-effects, like fatigue and anemia, reduce patient compliance and success rates (Aghemo and De Francesco, 2013; Kwo and Vinayek, 2011; Welsch et al., 2012).

A second generation of DAAs was approved by the FDA in 2013 and includes simeprevir, a protease inhibitor and sofosbuvir, a nucleotide analog NS5B inhibitor (Figure 2-5). In 2014, a combination therapy including, ledipasvir (NS5A inhibitor) and sofosbuvir was approved and marketed under the name Harvoni®. An additional, interferon-free treatment including DAAs ombitasvir (NS5A inhibitor), paritaprevir (NS3 protease inhibitor), ritonavir (a booster of protease inhibitors which slows their degradation) and dasabuvir (NS5B polymerase inhibitor) was also approved. The future of HCV drug discovery continues to explore DAAs capable of curing even treatment-resistant HCV phenotypes using an interferon-free regimen with reduced side effects. These DAA combinations, in particular those which are able to eliminate the need for interferon, have side-effect profiles which are well tolerated and increase patient treatment adherence and cure rates (Balistreri, 2016; Götte and Feld, 2016; Qian et al., 2016).

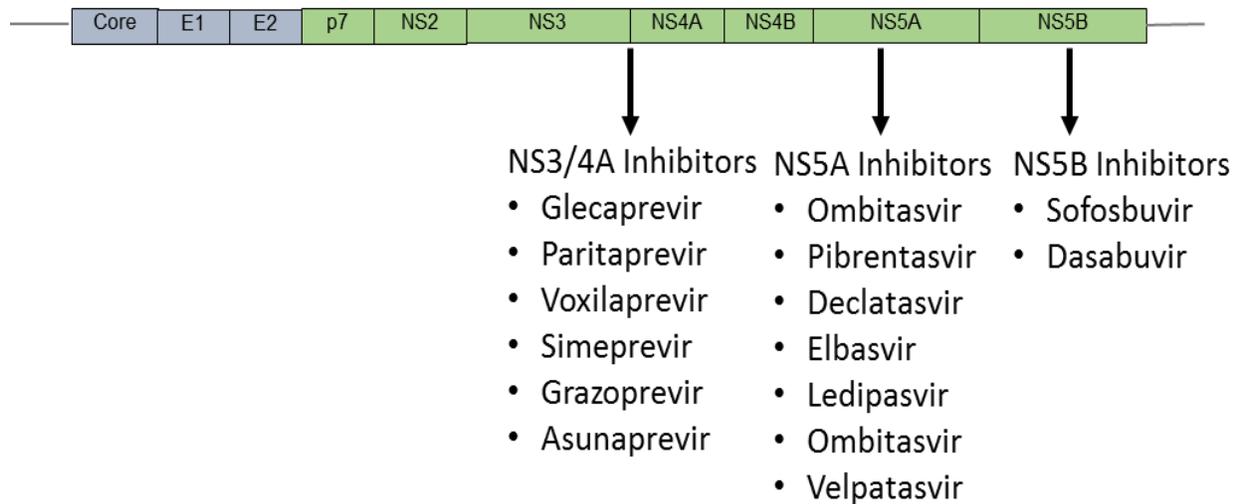


Figure 2-5 Approved inhibitors currently in use for the treatment of HCV listed in conjunction with the protein they target. Combinations of direct acting antivirals are the best treatment currently available for HCV and are able to cure most cases, regardless of genotype.

2.2.4 HCV NS3:NS4A

HCV NS3 is a multi-functional protein with an N-terminal serine protease domain and a C-terminal helicase/ATPase domain. A covalently linked protease and helicase domain is unique to the *flaviviridae* family of viruses and both enzymatic activities are essential for virus propagation, although the functional significance of their linkage is unknown. Helicase and protease enzyme activities are able to function independently, yet the presence of the protease domain has been shown to alter the activity of the helicase domain and the presence of the helicase domain increases the processivity of the protease domain. There is an allosteric domain between the protease and helicase and mutational studies within that domain highlight its importance to overall function but the cross-talk between the two domains leaves much to be discovered (Aydin et al., 2013; Beran and Pyle, 2008; Frick et al., 2004).

NS3 Protease

The NS3 protease is a serine protease which contains the His⁵⁷, Asp⁸¹, and Ser¹³⁹ catalytic triad indicative of the trypsin/chymotrypsin superfamily of proteases. Mutational studies show that the NS3 protease, is indispensable to the cleavage of several sites in the HCV poly-protein including cleavage of the NS3/NS4A *in cis* and the NS4A/NS5B, NS4B/NS5A, and NS5a/NS5B sites *in trans*. Mutations within the catalytic triad of NS3 abolished cleavage at any of these four sites but still allowed for the cleavage of any proteins upstream of NS3 in the HCV poly-peptide (Bartenschlager et al., 1994; Lin, 2006).

The NS3 protease enzyme has two beta-barrel motifs that are capped at either end by short alpha-helices. The overall structure is stabilized by the presence of a Zinc ion that is

coordinated with cysteines. The Zn^{+2} is essential for proteolytic function. The catalytic triad, His⁵⁷, Asp⁸¹, and Ser¹³⁹, is responsible for the proteolytic cleavage and the NS4A protein, which closely associates with the NS3 protease contributes to the proper positioning of the catalytic triad. This explains why NS3 protease function is greatly enhanced when NS4A is present (Raney et al., 2010).

The catalytic mechanism for the HCV NS3 protease is similar to all serine proteases. The His⁵⁷ serves as the base which activates the Ser¹³⁹ nucleophile. The nucleophile may then attack the carbonyl atom of the scissile peptide bond and form a tetrahedral intermediate. Subsequent breakdown of the tetrahedral intermediate leads to the formation of an acyl-enzyme intermediate and the N-terminal peptide product. The N-terminal product then dissociates, allowing for the division of the acyl-enzyme intermediate, the binding of water, and the formation of the C-terminal peptide product. This then allows for dissociation of the product and re-association by the next protease substrate (Di Cera, 2009; Fersht and Sperling, 1973).

Besides aiding in the poly-peptide processing during HCV replication, HCV NS3 protease also has a direct role in inhibiting the innate immune response to HCV infection. HCV infection is sensed through multiple pathways that converge on the adaptor protein, mitochondrial anti-viral signaling protein (MAVS). One way that the NS3 protease is able to disrupt antiviral signaling is to cleave MAVS. This abrogates immune signaling and blunts the innate and adaptive anti-viral response (Horner et al., 2011; Horner et al., 2012).

NS3 Helicase

The NS3 helicase is classified as a super family 2 helicase based on sequence homology. It contains two tandem domains which closely resemble the bacterial recombination enzyme, RecA. ATP binds the NS3 helicase between the two RecA-like domains. Domain 1 contains the phosphate binding loop present in all Walker A and B motif-containing NTPase proteins (Saraste et al., 1990). The phosphate binding loop is responsible for binding the ATP and allowing the

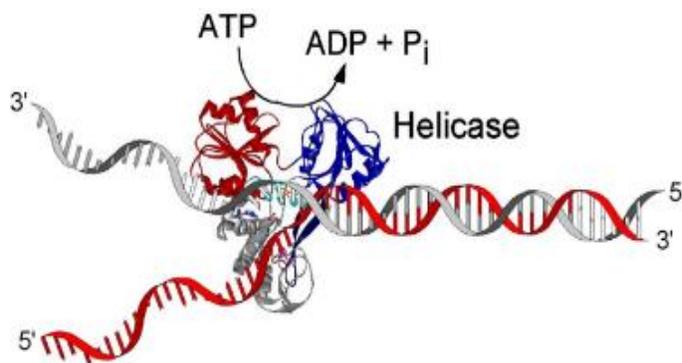


Figure 2-6: HCV helicase translocates unidirectionally along single stranded HCV RNA displacing the duplex strand. The helicase is powered by ATP.

catalytic base (Glu²⁹¹) to activate the water molecule needed for hydrolysis of the ATP. The activated water can then attack the γ -phosphate on the ATP (Belon and Frick, 2009b).

The purpose of the NS3 helicase is to couple the energy liberated through ATP hydrolysis to duplex RNA or DNA strand separation (Figure 2-6). The NS3 helicase will bind to either RNA or DNA with a low nanomolar dissociation constant. Although the binding of the helicase to RNA is not surprising given that HCV is an RNA virus, the helicase's ability to bind and unwind DNA does not have a known function and may indicate a role for NS3 in the cellular nucleus (Kasprzak et al., 2007). The helicase binds with a footprint of about 5-7 nucleotides and can translocate uni-directionally along single stranded RNA/DNA with the ability to disrupt duplex binding. However, the mechanism by which NS3 translocates is unclear and competing theories exist (Frick et al., 2006).

The inchworm model maintains that domains 1 and 2 of the helicase move opposite each other fueled by the energy of ATP. This movement results in a helicase which literally walks along a ssRNA or a ssDNA molecule. Crystal structures done in the presence of a non-hydrolyzable ATP and ATP transition state analogs give credence to the inchworm mechanism (M. Gu and Rice, 2010). Important in these structures are residues W^{501} and V^{432} which form “book-ends” along the DNA, locking the NS3 helicase into a position with a 5 base pair step size. W^{501} sits at the 3’ end of the helicase footprint along the nucleic acid, whereas V^{432} inserts at the 5’ end. ATP hydrolysis results in a conformational shift which causes domains 1 and 3 to rotate relative to one another closing with domain 2. This causes the protein to slide 1 nucleotide. Release of ATP allows the enzyme to swing open and the overall movement is 1 nucleotide per ATP hydrolyzed (Figure 2-7) (Frick et al., 2006; Pyle et al., 2006; Pyle, 2008; Raney et al., 2010; Velankar et al., 1999).

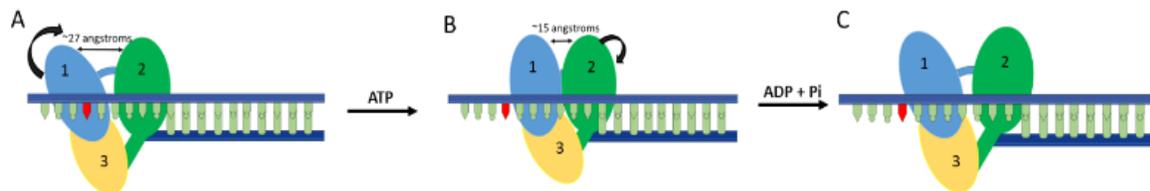


Figure 2-7: Inchworm model of HCV NS3 helicase motion. Prior to binding ATP (A) the helicase is in an open position with domain 1 swung out from domain 2 by about 27 angstroms. After binding ATP, (B) domain 1 swings in to close the gap between domains 1 and 2 to about 15 angstroms and shifting the complex by one nucleotide. Immediately following hydrolysis, domain 2 moves away from domain 1 and the helicase re-assumes an open conformation (C) (Raney et al., 2010).

In contrast to the inchworm model, the Brownian motion ratchet model says that the NS3 helicase exists in two states and that the switch between the two states is powered by ATP. The first state is a high-affinity state whereby the helicase is strongly attracted to the single stranded nucleotide substrate and the second state is a low affinity state, where the helicase is only weakly associated, thus allowing it to slide via Brownian motion. The preference for

forward motion comes from the ratchet effect where W^{501} intercalates into the nucleic acid and prevents the enzyme from sliding back (Levin et al., 2005).

2.3 Innate Immune System

The innate immune system is the first line of defense against invading organisms. Unlike the adaptive immune system which remembers pathogens from previous encounters but requires time, sometimes multiple days, to become effective, the innate immune system is fast. The innate immune system can respond in the first hours after an infection and it doesn't matter if the pathogen is known or completely novel or whether it's bacterial, viral, or fungal. The innate immune system is set up to detect patterns present in pathogens called pathogen associated molecular patterns (PAMPs) and to respond immediately (Alberts et al., 2002).

There are many parts to the innate immune system. One important component is the skin and mucous bilayers present in gut and intestinal linings which provide a physical and chemical barrier to pathogen entry. If the pathogen is able to breach the epithelial wall and reach the tissue or bloodstream, there are several systems including macrophages and the complement system set up to recognize the PAMPs and induce an inflammatory response. If a pathogen enters a cell, there are multiple protein receptors, such as Toll-like receptors and RIG-I like receptors, which will recognize PAMPs such as double stranded RNA or DNA, un-capped RNA, cyclic-diGMP, flagellin or lipopolysaccharides present in the cytoplasm. Once an intracellular innate immune protein has recognized a PAMP, they induce a signaling cascade resulting in the production of interferon and other cytokines (Takeuchi and Akira, 2010).

The function of the innate immune system is important not only for immediate pathogen recognition and neutralization but also for its role in the adaptive immune response and immunological memory (i.e. antibodies). Without the innate immune response, the adaptive immune system is significantly hindered. For example, the release of cytokines initiated by the innate immune response alerts nearby specialized cells, such as dendritic cells (antigen presenting cells), and recruits them to the area of infection. The innate response is so critical that many viruses and other pathogens, such as HCV, seek to disrupt the innate immune signaling. The disruption of innate immune signaling effectively slows the response and can allow the virus to spread undetected until it is well established in the host (Akira et al., 2006; Takeuchi and Akira, 2010).

2.3.1 RIG-I like receptors (RLRs)

A subset of the innate immune system intracellular pattern recognition receptors (PRRs) are known as the RIG-I like receptors (RLRs). The RLRs consist of three receptor proteins retinoic inducible gene I (RIG-I), melanoma differentiation associated protein 5 (MDA5), and laboratory of genetics and physiology protein 2 (LGP2) which belong to the DExD/H family of RNA helicases. RLRs reside at low concentrations in the cytoplasm of most cells in the human body and aid in detecting intracellular RNA virus infection (Jensen and Thomsen, 2012). The RLR's anti-viral function was first revealed by Yoneyama et al when RIG-I was uncovered during a cDNA screen of genes which induced the expression of interferon in response to duplex RNA transfection (Yoneyama et al., 2004a). It was determined that RIG-I's CARDs (capsase activation and recruitment domains) are important for transmitting the down-stream signal, and that the

helicase domain recognizes dsRNA and regulates the signal transduction in an ATPase dependent manner (Yoneyama et al., 2004b). Two additional proteins demonstrating high helicase domain sequence similarity to RIG-I were subsequently identified as MDA5 and LGP2. Both RIG-I and MDA5 possess CARDs and MDA5 was identified as transmitting a signal similar to RIG-I, leading to the activation of interferon (Yoneyama et al., 2005).

The RLR's contain a helicase and ATPase to bind and translocate along RNA, a C-terminal domain (CTD) to modulate and control this interaction, and, in the case of RIG-I and MDA5, two tandem CARDs to transmit the signal downstream. The RLRs are classified as super family 2 helicases and therefore their helicase domains contain two tandem RecA-like domains (Rossmann fold with repeating alpha/beta sheet) (Byrd and Raney, 2012; Rao and Rossmann, 1973). The presence of tandem RecA-like domains means that the super family 2 helicases (like super family 1) form a complete ATP binding site and are able to function as monomers, unlike superfamily 3, 4 or 5 helicases which are obligate oligomers (Fairman-Williams et al., 2010; Hickman and Dyda, 2005; Korolev et al., 1998; Pyle, 2008). Near the tandem RecA-like domains are conserved sequences within the helicase which form a platform for recognition and manipulation of nucleic acids. The helicase domains (hel1 and hel2) are separated by an insertion domain at the N-terminal of hel2, termed hel2i (Luo et al., 2011). The RLRs are not known to unwind RNA, instead they have been shown to translocate along either single or double stranded RNA powered by ATP. The ATPase function is essential for proper downstream signal activation (Errett and Gale, 2015a; Myong et al., 2009).

The C-terminal domain, present in all three RLRs, is a Zn²⁺ containing regulatory domain.

The fold of the CTD is conserved across the RLRs, but the characteristics of the ligand recognition groove vary which seems to account for some of the ligand specificity of each protein (Cui et al., 2008). Binding studies have shown that the RIG-I CTD binds 5' ppp dsRNA, ssRNA, and blunt-ended dsRNA with the highest affinity for 5' ppp dsRNA (Lu et al., 2010). Whereas, MDA5 and LGP2 CTDs preferentially bind dsRNA and the RNA ends are not as important (Bruns and Horvath, 2014; Pippig et al., 2009). The overall CTD binding is primarily electrostatic through a basic binding cleft that is not specific to the internal sequence of the RNA (Luo et al., 2011).

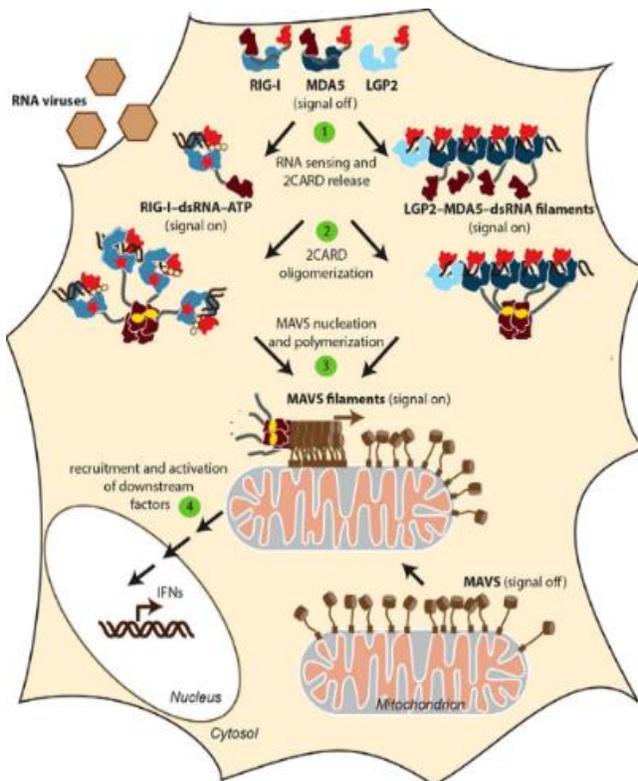


Figure 2-8: Schematic of RIG-I like receptor function. RIG-I, MDA5, and LGP2 exist in an auto-inhibited state in the cytoplasm in the absence of RNA ligand. Upon RNA ligand binding (1), the CARD domains are exposed to allow for oligomerization. RIG-I binds in short filaments whereas MDA5 binds in larger oligomers aided by the presence of LGP2 (2). After oligomerization, localization of the RLR filament to the mitochondrial membrane induces the formation of a MAVS + RLR filament (3) and initiation of downstream signaling (4) resulting in the production of interferon. This research was originally published in *Journal of Biological Science*. 2017; 292: 9000-9009. © The American Society for Biochemistry and Molecular Biology. Adapted for use with permission.

Immediately connected to the CTD, is a “pincer” or “bridge” which consists of two tandem alpha helices which connect the CTD to the hel2 domain. This RIG-I pincer domain has been shown through mutational studies to be imperative for stimulation of interferon in response to RNA binding and elimination of the pincer abolishes signaling altogether. Further, direct study through ATPase assays shows that deletion of even a portion of the pincer significantly reduces the ATPase activity in response to RNA binding. Thus this bridge/pincer, although small, transmits information directly from the CTD to the helicase2 domain (Kowalinski et al., 2011; Luo et al., 2011).

The tandem CARDs are present only in RIG-I and MDA5 and are responsible for the downstream signaling. The CARDs belong to the Death Domain superfamily and the CARD fold consists of six anti-parallel alpha helices. CARD domains are often involved in apoptotic pathways or inflammatory responses and may oligomerize into large complexes. The RLRs, RIG-I and MDA5, which contain CARD domains induce interferons which may lead to inflammation. In addition, the CARDs of MDA5 and RIG-I oligomerize to form larger complexes (Ferraio and Wu, 2012; Wu and Hur, 2015). The CARDs of RIG-I form a helical tetramer with a pitch of a single card. The CARD-CARD interface is mediated via a series of electrostatic and hydrophobic interactions. The RIG-I tandem CARDs have been crystallized in conjunction with MAVS. The model suggests that the tandem CARDs are necessary to create a stable tetramer which seeds the MAVS filament formation necessary for downstream signaling. (Wu et al., 2013; Wu et al., 2014; Wu and Hur, 2015)

The overall functional model of RIG-I and MDA5 could be thought of as the helicase domain forming a base that is covered by two flaps, the CARDs and the C-terminal domain. In the absence of RNA ligand, the protein is in an auto-inhibited state. The CARDs fold to block the helicase domain from interacting with “self” RNAs through a direct interaction. The C-terminal domain then folds on top of the CARD domains and maintains their position blocking the helicase domain. The ligand recognition portion of the CTD is still exposed and available to the cytosol to scan for PAMPS. In the event of PAMP binding, the CTD swings in to bring the ligand into contact with the helicase domain with perfect alignment. The presence of the CTD and ligand in connection with the helicase domain clashes with the CARDs and forces them to swing up and out to initiate downstream signaling (Kowalinski et al., 2011; Luo et al., 2011). The CARDs then oligomerize and re-locate to the mitochondrial membrane where they seed a MAVS filament (Luo et al., 2011) (Figure 2-8). The overall function of LGP2 is thought to be in accessory to MDA5 function instead of through direct signaling with MAVS due to the lack of CARDs (Lässig and Hopfner, 2017).

2.3.2 RIG-I

RIG-I (retinoic acid-inducible gene 1), although similar to MDA5 in structure and overall function, has distinctive ligand preferences and a unique mode of action. During *in vitro* experiments, RIG-I’s ideal ligand is short, di or tri-phosphorylated double stranded RNA (5’ppp-RNA) of around 7-10 base pairs (bps). RIG-I also binds short di or tri-phosphorylated single stranded RNA and RNAs with an m7-guanosine triphosphate cap. In addition, RIG-I can recognize and respond to short U-rich tracts of RNA (Goubau et al., 2014; Schlee et al., 2009).

These ligands represent multiple parts of an RNA virus life cycle including incoming RNA and replication intermediates. In addition, RNA polymerase III can transcribe tracts of viral DNA to 5'-ppp RNA which is also recognized by RIG-I (Sharma and Fitzgerald, 2011). Therefore, RIG-I is able to recognize both RNA and DNA viruses at multiple points in the viral life cycle.

RIG-I discriminates between self and non-self RNA primarily through spatial separation. RIG-I monitors the cytoplasm and 5'ppp-RNAs are not normally present in the cytosol of healthy cells because host mRNAs are capped before exiting the nucleus after transcription. Therefore, the presence of 5'ppp-RNA in the cytoplasm could signal a replicating virus. In addition, it was recently shown that RIG-I can bind to m⁷-guanosine capped RNA but it cannot accommodate the 2'-O-methylation of the first 5' nucleotide and therefore self-capped RNAs are excluded from RIG-I recognition. Any other cytoplasmic RNA is either cleaved to only a 5'-monophosphate (like tRNAs) or masked by ribonuclear complexes, as in ribosomal RNA. RIG-I ligand preference is tightly regulated to prevent self-RNA recognition (Marques et al., 2006).

After RIG-I recognizes a viral RNA ligand, RIG-I becomes ubiquitinated which promotes CARD oligomerization. In the absence of ligand, the RIG-I 2CARD is phosphorylated which sterically prevents ubiquitination further adding to the layers of protection against rogue signaling (Jonathan Maelfait and Rudi Beyaert, 2012). The role of the ATPase domain in RIG-I signaling is controversial. Given that RIG-I prefers ligands as small as 7bp, there is seemingly no need for translocation. However, single molecule studies indicate that RIG-I does translocate along RNA. One hypothesis to explain these findings is that RIG-I dissociates from low affinity substrates at a rate faster than the rate of ubiquitination, oligomerization, and signaling. But

when RIG-I is bound to a high affinity substrate, even in the presence of ATP, it would dissociate slower, allowing time for the signaling complex to assemble and initiate the downstream response. Therefore, the ATPase function would contribute to RIG-I ligand discrimination (Ahmad and Hur, 2015; Myong et al., 2009).

2.3.3 MDA5

MDA5 binds long double stranded RNAs with a length greater than 2,000 base pairs. Unlike RIG-I, MDA5 seems to have no preference for the type of RNA ends and instead freely binds anywhere along the dsRNA. Due to this lack of sequence or end specificity it's possible that MDA5 might be more prone to recognize native cell ligands such as mRNA or tRNA. However, long dsRNAs (>2,000 base pairs) are rare in the cytoplasm and when they do occur, base pairing is reduced by converting some RNPs to inosine by adenosine deaminase (ADAR1). Therefore, because MDA5 binds with highest affinity to RNAs greater than 2,000 base pairs that exhibit perfect complementarity, it is less likely that MDA5 would bind to cell-derived RNAs which, in general, are shorter than 2,000 base pairs or which lack perfect complementarity (Heraud-Farlow and Walkley, 2016; Liddicoat et al., 2015).

The increased binding stability of MDA5 for long dsRNA is mediated via ATP. MDA5 mutants deficient in ATP hydrolysis and ATP binding are constitutively active (Louber et al., 2015). However, it was seen that in the presence of ATP, MDA5 filaments are significantly weakened (i.e. MDA5 dissociates more rapidly), in particular at the ends of the RNA, presumably due to the ATP hydrolysis and MDA5 translocation activity. This shifts the stable MDA5 filament formation towards a preference for long dsRNA and increases the likelihood

that MDA5 will dissociate from the short filaments. Therefore, although ATP is not required for MDA5 signaling, it is required for MDA5 ligand discrimination. The ATP hydrolysis and ATP binding activity limits erroneous MDA5 signaling by decreasing the stability of the MDA5 proteins on a short RNA strand. This establishes the MDA5 ligand preference for longer double stranded RNA (Peisley et al., 2011).

2.3.4 LGP2

LGP2 (laboratory of genetics and physiology 2) is the less-understood RIG-I like receptor. The structure of LGP2 is similar to RIG-I and MDA-5 in the sense that they all share a homologous C-terminal ligand recognition domain and DExD/H box helicase domain. But LGP2 differs from RIG-I and MDA5 in that it lacks CARDS needed to activate MAVS. The lack of CARDS, suggests LGP2 plays a unique role in the innate immune system by signaling PAMP detection through a path that does not involve CARDS, but instead involves an interaction with MDA5 (Bruns and Horvath, 2014; Satoh et al., 2010). It has been shown that LGP2 can form active dimers (Murali et al., 2008) and oligomers that interact with MDA5 filaments (Bruns and Horvath, 2014). In addition, LGP2 has been shown to negatively regulate RIG-I signaling, possibly through substrate sequestration (Rodriguez et al., 2014).

LGP2 favors double stranded RNA ligands and binds with highest affinity to the ends of RNA, however there is no difference in binding affinity for RNA ligands with or without 5'triphosphates. It is possible that LGP2's preference for the ends of dsRNA is in direct competition with RIG-I's RNA end preference and this competition is what causes the observed hindrance of RIG-I signaling when LGP2 is upregulated. LGP2 enhances MDA5 filament

formation and thus MDA5 signaling activity (Bruns et al., 2014; Errett and Gale, 2015b; Uchikawa et al., 2016).

2.3.5 HCV and detection by RIG-I like receptors

HCV is recognized by the intracellular innate immune response once the viral RNA is released into the cytoplasm. Multiple pattern recognition receptors (PRRs), including Toll-like receptors (TLRs) and RIG-I like receptors (RLRs), participate in responding to HCV RNA. Loo et al showed, using immunofluorescence, that IRF-3 accumulation in the nucleus increases after JFH-1 infection in Huh7 cells but not in Huh7.5 cells which have a defective RIG-I pathway (Y. Loo et al., 2006). Therefore, Loo et al concluded that once the RIG-I like receptors have bound to the HCV RNA PAMP (i.e. the 5' triphosphate and the poly-U/UC region (Saito et al., 2008), they initiate a downstream signaling cascade resulting in the production of interferon and other cytokines.

HCV RNA may be recognized at several points in its lifecycle by the RIG-I like receptors. Initially, the newly released HCV RNA may be recognized by RIG-I due to its 5'ppp. Once the HCV ligand binds to RIG-I, it induces a conformational shift which releases the CARD domains for oligomerization and downstream interaction with MAVS. A mutation in the RIG-I CARD (T55I) is present in the HCV cured cell line, Huh 7.5, which is used because it is significantly more permissive to HCV infection. The T55I mutation does not allow for proper ubiquitination of the RIG-I CARD and therefore is able to abrogate RIG-I receptor signaling. The fact that blunting RIG-I signaling makes the cell line much more susceptible to HCV infection lends credence to the importance of RIG-I in HCV detection (Liu and Gale, 2010).

The RLR, MDA5 has also been implicated in HCV PAMP recognition although, due to its preference for longer dsRNA, it's hypothesized to recognize HCV replication intermediates produced later in the life cycle (Errett and Gale, 2015a). When cell lines were produced which knocked out either RIG-I or MDA5 and then looked at the production of interferon- β , interleukin-28, and interleukin-29 in response to HCV RNA transfection, they found that the antiviral response was significantly hampered when MDA5 was knocked out. Further, when siRNA was expressed which targeted MDA5 in primary hepatocytes infected with HCV genotype 2a, the cell's production of interferon- β was decreased. Therefore, it is believed that MDA5 also plays a role in HCV detection and that MDA5 and RIG-I sensing may occur at different points in the viral life cycle (Cao et al., 2015).

LGP2 has also, more recently, been implicated in HCV RNA sensing. Knockdown of LGP2 significantly hampered the interferon response to HCV infection and therefore LGP2 must be a positive regulator of HCV infection. Mechanistic studies showed that LGP2 exerted its effect somewhere upstream of MDA5 in the interferon signaling cascade and that HCV RNA promoted interaction between MDA5 and LGP2, which also enhanced MDA5 interaction with RNA (Hei and Zhong, 2017).

Although HCV RNA is able to trigger RIG-I like receptor signaling and the importance of this signaling to cell permissiveness is demonstrated with the Huh 7.5 cell line defective RIG-I mutation, most cases of HCV go on to establish a chronic infection. This is possible because HCV is adept at evading the host innate immune mechanisms and thereby elongating the time between infection and a robust immune response (Liu and Gale, 2010).

HCV's regulation of the RIG-I like receptor pathway is mediated by the NS3:NS4A protein. NS3:NS4A contains a covalently linked protease and helicase along with a co-factor protein (NS4A) which greatly enhances proteolytic function. NS3:NS4A recognizes and cleaves the MAVS RLR adaptor protein near its C-terminus, thus liberating it from the mitochondrial surface and abrogating its downstream signaling which leads to activation of IFN and other cytokines. Cleavage of MAVS stops all RIG-I like receptor signaling and the prevention of MAVS cleavage through mutation of the MAVS cleavage site (C508Y in MAVS) restores the RIG-I like receptor function. In addition, the HCV NS3-4A protease is able to cleave the adaptor protein for TLR signaling, TRIF, and inhibit detection through the TLR pathway. Taken together, this innate immune signaling disruption is enough to effectively block early HCV detection, giving the virus a chance to establish a chronic infection in hepatocytes (Horner et al., 2012; Horner, 2014).

2.4 Other DExD/H-box proteins

Proteins in the DExD/H box helicase family are known to be involved in many forms of RNA processing such as splicing, transcriptional and translational regulation, as well as mRNA export and decay. All members contain a highly conserved RNA binding site and ATPase. However, their functions are diverse and, in some cases, do not require nucleic acid unwinding (Frick et al., 2004). Thus, it has been proposed that while sometimes these helicases are performing the function of unwinding and remodeling nucleic acids, other times, they may be rearranging protein complexes or using nucleic acid binding activity to sequester nucleic acids. DExD/H box helicases have been discovered that play a role in the host response to pathogenic

infection. DDX1, DDX3, DDX5, DDX 6, DDX9, DDX17, DDX41, DDX60, DHX 9, and DHX36 comprise a short list of the host DExD/H box helicases which also function as PRRs (Jankowsky et al., 2011).

2.4.1 DDX3

DDX3 has the ability to enhance interferon production in response to viral invasion and works downstream of the RLRs. For example, DDX3 was identified in a yeast two-hybrid screen as an interacting partner to MAVS and it was shown using an IRF-3 reporter assay that the DDX3:MAVS interaction upregulates interferon production (Oshiumi et al., 2010). Therefore, DDX3 is a prime target for viral manipulation by human immunodeficiency virus, hepatitis C virus, and poxviruses (Oda et al., 2009; Owsianka and Patel, 1999; Yedavalli, Venkat S R K et al., 2004). Interestingly, DDX3 is also hijacked by HCV and HIV to participate in viral replication (Ariumi et al., 2007; Yedavalli, Venkat S R K et al., 2004). For example, DDX3 binds to the HCV core protein and when cells harboring an HCV replicon or infected with full length HCV-1b expressed a lentivirus vector encoding a short hairpin RNA against DDX3, both the replicon replication and the accumulation of full length HCV RNA were suppressed (Ariumi et al., 2007). DDX3 plays contradictory roles in viral replication and host viral detection/signaling by both detecting viral invasion and initiating immune signaling and by upregulating production of some viruses, like HCV.

DDX3's involvement in the IFN production pathway was discovered when Schroder et al searched for targets of vaccinia virus protein K7 in order to understand its ability to reduce interferon production. Using K7 antibodies to pull down factors in HEK293 cells led to the

identification of an unknown protein which was later identified, using peptide mass fingerprinting, as DDX3. Two groups simultaneously reported that DDX3 interacts with tank binding kinase 1 (TBK1) and I-kappa kappa epsilon (IKKε) in response to virus infection and encourages the phosphorylation of IRF3 and induction of the interferon promoter. When DDX3 was silenced via a small interfering RNA, interferon signaling upon vaccinia virus infection was reduced (Decker et al., 2008; Schröder, 2011). TBK1 and IKKε are at a critical point in the IFN induction pathway where signaling from diverse PRRs converge with signaling through MAVS, TRIF, and STING. DDX3 association with TBK1 and Ikkε, makes it a part of this key step.

Later Schroder et al elaborated on their original findings to show that DDX3 not only interacts with TBK1 and IKKε, but is also phosphorylated by them. They also showed that there is a direct interaction between DDX3 and IRF3 (L. Gu et al., 2013; Schröder, 2011). Soulat et al used chromatin immunoprecipitation to demonstrate that DDX3 binds the interferon promoter directly when stimulated by virus infection. In all cases, DDX3 had a positive effect on interferon induction (Decker et al., 2008). Lastly, Oshiumi et al found, while studying the interaction of HCV core protein and DDX3, that DDX3 also associates with MAVS. Reporter assay showed that the association of DDX3 and MAV enhanced IFN-β promoter activation and knockdown of DDX3 with a small interfering RNA decreased the IFN-β promoter activation (Oshiumi et al., 2010).

2.4.2 DDX1

DDX1 is an RNA helicase able to unwind duplex RNA and RNA-DNA complexes which may play a role in RNA clearance at double-stranded breaks, thereby aiding in template guided repair of the genome. More recently, DDX1 has been implicated in innate immunity as an

intracellular receptor for dsRNA. DDX1 has been shown to bind poly I:C RNA and knockdown of DDX1 decreases interferon signaling (Fullam and Schröder, 2013; Zhang et al., 2011).

2.4.3 DDX5

DDX5 is involved in multiple cellular processes which involve alteration of RNA secondary structure, including translation initiation, nuclear splicing, and spliceosome assembly (Fuller-Pace, 2013). DDX5 has also been implicated in HCV replication where it is shown to interact with NS5B. Initially it was revealed that over-expression of NS5B caused a re-localization of DDX5 from the nucleus to the cytoplasm and that infection with infectious JFH1 particles also caused the same re-localization. Binding studies showed that DDX5 interacts with NS5B in two locations on the protein. Further, in cells where DDX5 was knocked down, production of the negative sense strand of HCV RNA was reduced, implying a possible role for DDX5 in HCV RNA replication (Upadya et al., 2014).

2.5 FRET Microscopy

2.5.1 Principles

The ability to look at sub-cellular structures by labeling structures or proteins of interest using fluorescent dyes or fluorescent proteins is useful to understanding the cellular environment under different conditions. This may be achieved by labeling multiple cellular components in fixed cells and taking several images targeting a different range of emission wavelengths each time to isolate the signal from each fluorescent tag. The resulting images are then overlaid and have contributed some beautiful and informative images to the

understanding of the sub-cellular environment. However, fixed cells are not alive and therefore cannot contribute information as to the dynamic and sometimes transient interactions occurring in live cells. In addition, observing the co-localization of proteins using a fluorescence microscope does not provide the resolution necessary to confirm or deny an interaction because proteins are on the scale of a few nanometers wide and the resolution of a fluorescent microscope is on the order of hundreds of nanometers (Piston and Kremers, 2007; Raicu and Singh, 2013). Instead, FRET microscopy provides an ideal framework to learn more about dynamic, intracellular protein:protein interactions.

Fluorescence resonance energy transfer (or Förster resonance energy transfer) (FRET) (Lakowicz, 2010) can be measured over distances of 0-10 nm (1-100 Angstroms) (Clegg, 1995). The fundamental principle of FRET is that a fluorescent donor molecule is put into an excited state through absorption of light. That donor molecule may then transfer its energy non-radiatively to a nearby fluorescent acceptor molecule through long-range dipole-dipole interactions. The non-radiative transfer is affected by orientation, environmental factors, and distance. Once the acceptor molecule has absorbed the energy from the donor molecule, the light emitted by the acceptor molecule can be measured as FRET (Clegg, 1995; Piston and Kremers, 2007; Selvin, 1995; Selvin, 2000). This is because the light emitted from the acceptor will be directly proportional to the energy transferred from the donor. The transfer efficiency (E) can be measured using the relative fluorescence intensity of the donor in the absence (F^D) and presence (F^{DA}) of the acceptor as:

$$E = 1 - \frac{F^{DA}}{F^D} \quad (\text{Equation 2-1})$$

One of the most important aspects to a successful FRET experiment is selection of an appropriate FRET pair. Ideally, the donor should have a high quantum yield so that it gives off the most energy possible from each excitation photon and the acceptor should have the highest extinction co-efficient possible to ensure that it is able to absorb the maximum number of photons emitted by the donor. Choosing a donor with a high quantum yield and an acceptor with a high extinction co-efficient increases the likelihood of a successful energy transfer. Additionally, the spectral overlap between the donor emission and the acceptor absorption should be at a maximum while minimizing any direct excitation by the acceptor at the donor excitation wavelength (Bajar et al., 2016; Shaner et al., 2005).

In addition, the location and orientation of the fluorescent tags on the proteins or other sub-cellular regions of interest need to be positioned in such a way that productive FRET can occur between the fluorescent tags. Since FRET is only easily measurable within a small range of 1-10 nm, if two proteins interact, but form a complex where the orientation of the fluorescent tags are diametrically opposed, no FRET may be observed. Further, the tagging of a protein surface may block an association that would otherwise take place under native conditions. Therefore, great care must be taken not to obscure the putative interaction site. The interaction site may be known through x-ray crystallography, through known mutations which cause a loss of interaction or it may be predicted through molecular modeling. The fluorescent tags should then be placed outside of the interaction site so as not to block an interaction (Jares-Erijman and Jovin, 2006; Snapp, 2005).

The above mentioned obstacles also highlight an important property of all FRET measurements. Specifically, the absence of FRET cannot conclusively say that two proteins do not interact because the selection of the fluorescent tags may not be optimal or the location of the tags may be obscuring interaction sites. Only the presence of FRET and the intensity of FRET can render information as to the interaction of two proteins. Each FRET experiment must be optimized to achieve the best signal (Piston and Kremers, 2007).

When FRET occurs the donor is excited and transfers energy to an acceptor. There are several different approaches to calculating FRET which rely on three basic components of fluorescence resonance energy transfer. The first approach focuses on the decrease in donor emission intensity when some of the donor emission energy is non-radiatively transferred to a nearby acceptor. The second approach focuses on the increase (sensitization) of acceptor emission intensity when an acceptor receives energy from a nearby donor. The third approach focuses on the property that, when a donor interacts with an acceptor, this causes a decrease in the fluorescence lifetimes of the interacting donors because FRET provides additional pathways through which the donor may lose energy and therefore the total fluorescence lifetime of the donor is decreased. These three effects decrease of the donor, enhanced (sensitized) emission of the acceptor, and reduced lifetime of the donor, make up the basis for several FRET analysis techniques which will be briefly discussed (Raicu and Singh, 2013).

2.5.2 FRET calculation using filter based measurements

Two ways to calculate FRET, as adapted by Lakowicz (Lakowicz, 2010) and Hoppe (Hoppe et al., 2002) are based on whether they depend on the decrease of the donor intensity

due to quenching or the increase of the acceptor intensity due to sensitized emission. The first method is defined based on a decrease of the donor intensity due to fluorescence quenching (Equation 2-2), and the second is defined based on the increase of the acceptor emission intensity due to sensitized emission (Equation 2-3).

$$E_{app} = \left(\mathbf{1} - \frac{F_{DA}(\lambda_D^{ex} \lambda_D^{em})}{F_D(\lambda_D^{ex} \lambda_D^{em})} \right) \quad (\text{Equation 2-2})$$

$$E_{app} = \frac{\varepsilon_A(\lambda_D^{ex})}{\varepsilon_D(\lambda_D^{ex})} \left(\frac{F_{AD}(\lambda_D^{ex} \lambda_A^{em})}{F_A(\lambda_D^{ex} \lambda_A^{em})} - \mathbf{1} \right) \quad (\text{Equation 2-3})$$

In equation 2-2, $F_D(\lambda_D^{ex} \lambda_D^{em})$ is the donor fluorescence of the sample in the absence of acceptor and $F_{DA}(\lambda_D^{ex} \lambda_D^{em})$ is the donor fluorescence in the presence of acceptor. In equation 2-3, the $\varepsilon_A(\lambda_D^{ex})$ is the extinction coefficient of the acceptor at the donor excitation wavelength and $\varepsilon_D(\lambda_D^{ex})$ is the extinction coefficient of the donor at the donor excitation wavelength. $F_{AD}(\lambda_D^{ex} \lambda_A^{em})$ is the acceptor fluorescence in the presence of donor and $F_A(\lambda_D^{ex} \lambda_A^{em})$ is the acceptor fluorescence in the absence of donor.

The application of equation 2-2 is confounded by determination of the fluorescence of the donor in the absence of acceptor ($F_D(\lambda_D^{ex} \lambda_D^{em})$). Some techniques (i.e. acceptor photo-bleaching and FRET spectrometry) are able to calculate ($F_D(\lambda_D^{ex} \lambda_D^{em})$) indirectly and therefore may be applied to fluorescent microscopic images where the direct measurement of ($F_D(\lambda_D^{ex} \lambda_D^{em})$) is impossible. These techniques will be discussed later. However, in the specific case of an *in vitro* titration of a donor with an acceptor that is measured using a spectrophotometer, the starting fluorescence intensity of the donor (before any titrant is added) is a close approximation to the $F_D(\lambda_D^{ex} \lambda_D^{em})$ throughout the titration if the starting

volume is large relative to any titrant added. In this specific case, the starting point in the titration is defined as $I_{Alone}^D(\lambda_D^{ex} \lambda_D^{em})$ which is analogous to $F_D(\lambda_D^{ex} \lambda_D^{em})$ and is directly measured. Each subsequent addition of titrant is defined as $I^{DA}(\lambda_D^{ex} \lambda_D^{em})$ which is analogous to $F_{DA}(\lambda_D^{ex} \lambda_D^{em})$ or donor in the presence of acceptor when excited at the donor excitation wavelength and measured at the donor emission wavelength. If the quantities $I_{Alone}^D(\lambda_D^{ex} \lambda_D^{em})$ and $I^{DA}(\lambda_D^{ex} \lambda_D^{em})$ are substituted into equation 2-2, then equation 2-4 results:

$$F_{FRET} = 1 - \frac{I^D(\lambda_D^{ex} \lambda_D^{em})}{I_{Alone}^D(\lambda_D^{ex} \lambda_D^{em})} \quad (\text{Equation 2-4})$$

Another way to calculate FRET as defined by Hoppe (Hoppe et al., 2002) (equation 2-3) is sometimes called three-filter acceptor “sensitized emission” because it focuses on the energy gained by the acceptor rather than energy lost by the donor. Acceptor sensitized emission is due to energy transfer from the donor. This is determined using equation 2-3 from the ratio of the acceptor fluorescence intensity in the presence of donor $F_{AD}(\lambda_D^{ex} \lambda_A^{em})$ and the acceptor fluorescence intensity in the absence of donor $F_A(\lambda_D^{ex} \lambda_A^{em})$ and the extinction coefficients of the donor and acceptor respectively $\epsilon^A(\lambda_D^{ex})$ and the $\epsilon^D(\lambda_D^{ex})$.

Three-filter sensitized emission involves the use of three filter sets which collect three images. The three images are donor excitation and donor emission, $F(\lambda_D^{ex} \lambda_D^{em})$ or I_D , acceptor excitation and acceptor emission, $F(\lambda_A^{ex} \lambda_A^{em})$ or I_A , and donor excitation and acceptor emission, $F(\lambda_D^{ex} \lambda_A^{em})$, or I_F . The collection I_D , I_A , and I_F assumes that there is negligible transmission of acceptor emission into the donor emission filter when excited at the donor

excitation wavelength and negligible transmission of the donor emission into the acceptor emission filter when excited at the acceptor excitation wavelength or that the following conditions are met:

$$F_A(\lambda_D^{ex} \lambda_D^{em}) = 0 \quad (\text{Equation 2-5})$$

$$F_D(\lambda_A^{ex} \lambda_A^{em}) = 0 \quad (\text{Equation 2-6})$$

Another assumption made in three-filter sensitized emission is that the contributions of excitation light or fluorescence emission for any given fluorophore are linearly related and can be propagated from one filter set to another using scalar factors.

To solve for equation 2-3 using a microscope, a correction must be applied for any donor emission which may overlap with the acceptor emission spectrum. Specifically, $F_{AD}(\lambda_D^{ex} \lambda_A^{em})$ and $F_A(\lambda_D^{ex} \lambda_A^{em})$ must be corrected for any donor spectral contamination. In addition, $F_A(\lambda_D^{ex} \lambda_A^{em})$ (fluorescence of the acceptor in the absence of donor) when measured in a microscope, must be determined in the presence of donors. Assuming the acceptor fluorescence is not modified by the presence of the donor when the acceptor is excited at the acceptor's excitation then,

$$F_{AD}(\lambda_A^{ex} \lambda_A^{em}) = F_A(\lambda_A^{ex} \lambda_A^{em}) \quad (\text{Equation 2-7})$$

If it is assumed that the emission of the acceptor at one wavelength is proportional to the emission of the acceptor at another wavelength and equation 2-7 is used then the fluorescence of the acceptor alone in the presence of donor can be determined using equation 2-8.

$$F_A(\lambda_A^{ex} \lambda_A^{em}) = \alpha F_{AD}(\lambda_A^{ex} \lambda_A^{em}) = \alpha I_A \quad (\text{Equation 2-8})$$

The α co-efficient is measured in a sample containing only acceptor. The α is defined as:

$$\alpha = \frac{F_A(\lambda_D^{ex} \lambda_A^{em})}{F_A(\lambda_A^{ex} \lambda_A^{em})} \quad (\text{Equation 2-9})$$

With any given FRET pair, the emission of the donor often overlaps with the emission of the acceptor such that the signal collected at the acceptor emission upon donor excitation ($F(\lambda_D^{ex} \lambda_A^{em})$, or I_F) contains fluorescence from both the donor and the acceptor (equation 2-10).

$$F(\lambda_D^{ex} \lambda_A^{em}) = F_{AD}(\lambda_D^{ex} \lambda_A^{em}) + F_{DA}(\lambda_D^{ex} \lambda_A^{em}) = I_F \quad (\text{Equation 2-10})$$

To correct for the donor emission within the acceptor emission wavelength another correction (β) can be determined such that:

$$F_{DA}(\lambda_D^{ex} \lambda_A^{em}) = \beta F_{DA}(\lambda_D^{ex} \lambda_D^{em}) = \beta I_D \quad (\text{Equation 2-11})$$

Where the β is determined from a sample expressing solely donor. The β is defined as:

$$\beta = \frac{F_D(\lambda_D^{ex} \lambda_A^{em})}{F_D(\lambda_D^{ex} \lambda_D^{em})} \quad (\text{Equation 2-12})$$

By substituting the equations for the α and β corrections into the equation for sensitized emission, equation 2-3 results in equation 2-13, as defined by Hoppe:

$$E_{app} = \frac{\varepsilon^A(\lambda_D^{ex})}{\varepsilon^D(\lambda_D^{ex})} \left(\frac{F(\lambda_D^{ex} \lambda_A^{em}) - \beta F_{DA}(\lambda_D^{ex} \lambda_D^{em})}{\alpha F_A(\lambda_A^{ex} \lambda_A^{em})} - 1 \right) \quad (\text{Equation 2-13})$$

The fluorescence intensities can be collected using three filter sets, I_D , I_A , and I_F . Therefore equation 2-13 becomes:

$$E_{app} = \frac{\varepsilon^A(\lambda_D^{ex})}{\varepsilon^D(\lambda_D^{ex})} \left(\frac{I_F - \beta I_D}{\alpha I_A} - 1 \right) \quad (\text{Equation 2-14})$$

2.5.3 Acceptor photo-bleaching

One of the challenges which arises with the calculation of E_{app} using any of the methods which rely on the donor decrease upon energy transfer (equation 2-2) is how to calculate the fluorescence of the donor alone in a sample containing both donor and acceptor. Acceptor photo-bleaching is one method which attempts to solve this problem by bleaching the acceptor irreversibly (Jovin et al., 2005). If the acceptor molecule is bleached irreversibly, the donor will no longer be quenched and its fluorescence restored. The ratio between the donor before quenching and the donor after quenching is a measure of how much energy was transferred between the donor and acceptor or FRET.

$$\mathbf{FRET} = \frac{F_{post}^D - F_{pre}^D}{F_{post}^D} \quad (\text{Equation 2-15})$$

Where F_{post}^D is the fluorescence of the donor after acceptor photo-bleaching and F_{pre}^D is the fluorescence of the donor before photo-bleaching.

Several problems exist for acceptor photo-bleaching, particularly in live cell measurements. Photo-bleaching the acceptor takes minutes, and therefore can only be done in fixed cells or with immobilized proteins. The time it takes to bleach the acceptor is ample time for molecular diffusion in a cell and the structure of live dynamic complexes will shift in that time. In addition, the process is irreversible, meaning that each cell can only be measured once. Lastly, it can be difficult to specifically target the acceptor and not also cause some bleaching to the donor (Karpova et al., 2003).

2.5.4 Fluorescence lifetime measurements

Fluorescence lifetime imaging (FLIM) is based off the model that all fluorescent molecules exhibit an exponential decay of intensity and that the rate of this decay is affected by the surrounding environment of the fluorescent molecule. Interactions with nearby quenchers will change the decay rate of the molecule. When the donor molecule is quenched through interaction with an acceptor, the rate of the donor decay will be decreased (Berezin and Achilefu, 2010).

$$\mathbf{FRET} = \mathbf{1} - \frac{\tau^{DA}}{\tau^D} \quad (\text{Equation 2-16})$$

Where T^{DA} is equal to the fluorescence of the donor in the presence of acceptor and T^D is equal to the fluorescence of the donor in the absence of acceptor (Hoppe et al., 2002).

Therefore, FRET can be measured through the decrease in fluorescence decay time in the presence of acceptor. One advantage of FLIM is that it relies only on the fluorescence of the donor and therefore is immune to some of the problems discussed above such as the effect of direct excitation of the acceptor at the donor excitation wavelength. Limitations to FLIM include the nanosecond time scale over which images need to be collected which requires expensive equipment not available to many labs. In addition, the fluorescence lifetimes of fluorescent molecules are also affected by factors such as pH and the changing environment of a cell can cause pH fluctuations which shorten fluorescent lifetimes or local environmental conditions can cause fluorescent molecules to exhibit multiple fluorescent decays. All of these factors can induce artifacts which make data interpretation difficult (Becker, 2012).

2.5.5 Optical Microspectroscopy

Optical microspectroscopy is similar to sensitized emission in that it relies on the calculation of donor and acceptor contributions to a composite donor and acceptor spectrum. A major component of optical microspectroscopy is that it employs spectral imaging. Spectral imaging means that the whole emission spectrum is collected allowing for separation of the donor and acceptor spectrum not just at the peak of emission, but also at the tails of the fluorescent spectra. This allows for better discrimination between donor and acceptor signal and may be able to pick up on more subtle spectra shifts which are missed when only the peak values are collected. In addition, the technique allows for the calculation of the relative amount of donor and acceptor fluorescence. Furthermore, it is possible to take images at two excitation wavelengths and to accurately correct for direct excitation of the acceptor or any donor bleed through in the measurement. One of the drawbacks to the method using two successive excitations is the time it takes to switch from one wavelength to another, which causes molecular diffusion to scramble the information between scans. In other words, the molecules imaged in the first scan are distinct from those measured in the second scan (Chen et al., 2007; Leavesley et al., 2013; Raicu et al., 2009; Raicu et al., 2005; Raicu and Singh, 2013).

In a typical fluorescent optical microspectroscopy experiment, there are several fluorophores present in the sample being measured. In order to separate the component spectra from the overall composite spectrum, spectral unmixing must be employed. Spectral unmixing separates the relative contribution from each fluorophore for every pixel of an image. In order to achieve this, elementary spectrum of each fluorescent spectrum contributing to the

overall composite spectrum are first collected using samples containing each fluorophore individually. The acquisition of elementary spectra must be obtained using the same conditions employed for the measurement of the composite sample.

The underlying assumption of spectral unmixing is that the composite spectrum is made up of a linear combination of each contributing elementary spectra as defined below:

$$S^m(\lambda_{ex,i} \lambda_{em}) = \sum_l [k^l(\lambda_{ex,i}) s^l(\lambda_{em})] \quad (\text{Equation 2-17})$$

Where i (=1,2) is the excitation wavelength index and l is the summation index for the different spectra present in the composite spectra. The overall spectral contributions from each fluorophore in the composite spectra can then be determined using a linear algebra matrix to calculate the individual contributions of the elementary spectra to each point in the measured composite spectra.

The coefficients, defined as k^x (where x equals the contributing elementary spectra (i.e. donor in the presence of acceptor (DA), acceptor in the presence of donor (AD), etc)) equal to the fluorescent contributions of each fluorescent species present in the composite spectra can then be extracted using a least squares minimization method previously described (Biener et al., 2013; Raicu et al., 2005; Raicu and Singh, 2013).

2.5.7 FRET Spectrometry

FRET spectrometry uses the technique of spectral imaging and applies it to the problem of calculating not only the presence or absence of FRET but also the spatial dynamics of a supermolecular protein complex in living cells. In one version of FRET spectrometry, unlike

other techniques, spatial information can be gained without the burden of a second excitation wavelength (Raicu and Singh, 2013). To achieve this, the selection of fluorescent FRET pairs is of paramount importance because it is essential that there is no direct excitation of the acceptor at the excitation wavelength. Wildtype GFP and a yellow fluorescent protein are one example of a pair which meet this criteria. To eliminate the dependence on acceptor emission, FRET spectrometry relies on the measurement of the donor quenching using equation 2-2 for the calculation of E_{app} .

$$E_{app} = 1 - \frac{F^{DA}}{F^D} \quad (\text{Equation 2-2})$$

In equation 2-2, F^D is defined as the total fluorescence of the donor in the absence of acceptor and F^{DA} is defined as the fluorescence of the donor in the presence of acceptor. To solve for F^{DA} , deconvolution (unmixing) of the emission spectra for the fluorescence of the donor is necessary (equation 2-17). When the composite spectrum is deconvoluted, it yields two quantities (k^{DA} and k^{AD}) which are equal to the maximum emission intensity for donor in the presence of acceptor and for acceptor in the presence of donor respectively (Biener et al., 2013). The quantity k^{DA} can then be used to determine the total number of donor photons emitted per unit time by integrating the fluorescence spectra, $I^{DA}(\lambda_{em}) = k^{DA}i^D(\lambda_{em})$ over all emission wavelengths (equation 2-18).

$$F^{DA} = k^{DA} \int i^D(\lambda_{em}) d\lambda_{em} = k^{DA}w^D \quad (\text{Equation 2-18})$$

where $i^D(\lambda_{em})$ is the emission intensity of the donor spectrum measured in a sample only expressing donor and normalized to the maximum value and w^D is the integral of the elementary spectra of the donor.

Similarly, the k^{AD} value extracted from unmixing can be used to determine the total number of acceptor photons emitted per unit time by integrating the fluorescence spectrum $I^{AD}(\lambda_{em}) = k^{AD}i^A(\lambda_{em})$ over all emission wavelengths.

$$F^{AD} = k^{AD} \int i^A(\lambda_{em}) d\lambda_{em} = k^{AD} w^A \quad (\text{Equation 2-19})$$

Where $i^A(\lambda_{em})$ is the emission intensity of the acceptor spectrum measured in a sample only expressing acceptor and normalized to the maximum value and w^A is the integral of the elementary spectra of the acceptor. The term F^{AD} will be used at a later point.

The term F^D in equation 2-2 is calculated by assuming that light energy is conserved. In other words, F^D is equal to the fluorescence of the donor in the presence of the acceptor (F^{DA}) and the energy lost to FRET through the interaction of the donor and the acceptor ($F^D(\text{RET})$).

$$F^D = F^{DA} + F^D(\text{RET}) \quad (\text{Equation 2-20})$$

To solve for $F^D(\text{RET})$, a relationship between $F^D(\text{RET})$ to $F^A(\text{RET})$ must be derived. The quantum yield of the donor (Q^D) is a ratio of the number of donor emissions per total donor excitations. If the total number of excitations is termed N^{RET} , then the quantum yield defines how many of those total excitations would be emitted by the donor if no RET occurred according to equation 2-21.

$$N^{\text{RET}} Q^D = F^D(\text{RET}) \quad (\text{Equation 2-21})$$

However, once the photons have transferred from donor to acceptor, only a fraction of those (defined as the quantum yield of the acceptor which is the total number of photons emitted by the acceptor per total number of acceptor excitations, Q^A) will be emitted by the acceptor as photons according to equation 2-22.

$$N^{RET} Q^A = F^A(RET) \quad (\text{Equation 2-22})$$

If N^{RET} from equation 2-22 is substituted into equation 2-21, equation 2-23 results (Raicu et al., 2009):

$$F^D(RET) = \frac{Q^D}{Q^A} F^A(RET) \quad (\text{Equation 2-23})$$

To solve for $F^A(RET)$ in equation 2-21, the following equation which relates the fluorescence of the acceptor in the presence of donor is equal to the direct excitation of the acceptor plus the energy gained from FRET.

$$F^{AD} = F_{\lambda_{ex}}^A + F^A(RET) \quad (\text{Equation 2-24})$$

where F^{AD} is defined by equation 2-19.

If there is no direct excitation of the acceptor (as would be the case for carefully chosen FRET pairs) then equation 2-24 may simplify to:

$$F^{AD} = F^A(RET) \quad (\text{Equation 2-25})$$

Since F^{AD} is a measurable quantity (equation 2-19) and $F^{AD} = F^A(\text{RET})$, then $F^D(\text{RET})$ becomes a measurable quantity through its relationship to $F^A(\text{RET})$ shown above in equation 2-23.

By substituting equation 2-18 in for F^{AD} and equation 2-23 in for $F^D(\text{RET})$ in equation 2-20, the following equation 2-26 results.

$$F^D = k^{DA}w^D + \frac{Q^D}{Q^A} F^{AD} \quad (\text{Equation 2-26})$$

Next equations 2-18, 2-19, and 2-26 may be substituted into equation 2-2 to yield equation 2-27.

$$E_{app} = 1 - \frac{k^{DA}w^D}{k^{DA}w^D + \frac{Q^D}{Q^A} k^{AD}w^A} \quad (\text{Equation 2-27})$$

Equation 2-27 can be rearranged to yield equation 2-28.

$$E_{app} = \frac{1}{1 + \frac{Q^D k^{AD} w^A}{Q^A k^{DA} w^D}} \quad (\text{Equation 2-28})$$

The term Q^D is defined as the quantum yield of the donor and the term Q^A is defined as the quantum yield of the acceptor. Both known quantities which can be found in the literature for commonly used fluorescent probes. The term w^D and w^A are the integrals of the donor and acceptor respectively and these are measured and calibrated for each experiment.

By calculating FRET using a single wavelength, the FRET spectrometry technique (Raicu and Singh, 2013) eliminates the need for a second excitation. By eliminating the second excitation wavelength, the molecular diffusion which may occur between scans does not occur and the analysis can be completed on a pixel level as opposed to an average over an area. This

pixel level information gives better spatial resolution of the complex(es) by allowing the creation of an image which shows a pixel by pixel distribution of E_{app} values throughout the cell. In addition, it produces many more data points in one area because now instead of averaging the E_{app} over a region of interest, you can extract a wide number of E_{app} values over the region of interest and thus get a better idea as to the full dynamic range of complexes present within the area.

Chapter 3 Methods

3.1 Molecular Cloning

3.1.1 pCMV-YFP-LGP2

LGP2 (DHX58) human cDNA ORF clone (Origene NM_024119) was amplified with PCR from the pCMV3-RC204837 (Origene, Inc) backbone using primers (Integrated DNA Technologies, Coralville, IA) which contain HindIII and XhoI restriction sites.

Forward 5'-CGC GCG AAG CTT ATG GAG CTT CGG TCC TAC CAA TGG GAG G-3'

Reverse 5'-GCG CGC CTC GAG TCA CTC CAG GGA GAG GTC CGA CAA GTT C-3'

The PCR product was then cut with HindIII and XhoI restriction enzymes and ligated into a similarly treated pCMV6-AN-mYFP (Origene-PS100050) plasmid.

The resulting plasmid expresses a YFP protein connected to LGP2 via a two amino acid linker.

3.1.2 pCMV-CFP-LGP2

LGP2 (DHX58) human cDNA ORF clone (Origene NM_024119) was amplified with PCR from the pCMV3-RC204837 (Origene, Inc) backbone using primers (Integrated DNA Technologies) which contain HindIII and XhoI restriction sites.

Forward 5'-CGC GCG AAG CTT ATG GAG CTT CGG TCC TAC CAA TGG GAG G-3'

Reverse 5'-GCG CGC CTC GAG TCA CTC CAG GGA GAG GTC CGA CAA GTT C-3'

The PCR product was then cut with HindIII and XhoI restriction enzymes and ligated into a similarly treated pCMV6-AN-mCFP (Origene-PS100052) plasmid.

The resulting plasmid expresses a CFP protein connected to LGP2 via a two amino acid linker.

3.1.3 pCMV-YFP-DDX1

DDX1 human cDNA ORF clone (Origene NM_004939) was amplified using primers (Integrated DNA Technologies) which contain HindIII and XhoI restriction sites.

Forward: 5'-CGC GCG AAG CTT ATG GCG GCC TTC TCC GAG AT-3'

Reverse: 5'- GCG CGC CTC GAG TCA GAA GGT TCT GAA CAG CTG GTT AG-3'

The PCR product was then cut with HindIII and XhoI restriction enzymes and ligated into a similarly treated pCMV6-AN-mYFP (Origene-PS100052) plasmid.

The resulting plasmid expresses a YFP protein connected to DDX1 via a two amino acid linker.

3.1.4 pCMV-YFP-DDX3

DDX3 human cDNA ORF clone (Origene NM_001356) was amplified with using primers (Integrated DNA Technologies) which contain HindIII and XhoI restriction sites.

Forward: 5'CGC GCG AAG CTT ATG AGT CAT GTG GCA GTG GAA AAT 3'

Reverse: 5'GCG CGC CTC GAG TCA GTT ACC CCA CCA GTC AAC CCC CTG 3'

The PCR product was then cut with HindIII and XhoI restriction enzymes and ligated into a similarly treated pCMV6-AN-mYFP (Origene-PS100052) plasmid.

The resulting plasmid expresses a YFP protein connected to DDX3 via a two amino acid linker.

3.1.5 PCMV-YFP-DDX5

DDX5 human cDNA ORF clone (Origene NM_004939) was amplified with using primers (Integrated DNA Technologies) which contain HindIII and XhoI restriction sites.

Forward: 5'- CGC GCG AAG CTT ATG TCG GGT TAT TCG AGT GAC C- 3'

Reverse: 5'- GCG CGC CTC GAG TTA TTG GGA ATA TCC TGT TGG C-3'

The PCR product was then cut with HindIII and XhoI restriction enzymes and ligated into a similarly treated PCMV6-AN-mYFP (Origene-PS100052) plasmid.

The resulting plasmid expresses a YFP protein connected to DDX5 via a two amino acid linker.

3.1.6 PCMV-CFP-NS3:NS4A

NS3:NS4A was cut from p28-His-NS3:NS4A-His (Frick et al., 2004) plasmid using NheI and XhoI.

The resulting DNA fragment was gel purified and ligated into PCMV-AN-CFP-scNS3 which was also cut with NheI and XhoI. The resulting plasmid expresses a CFP protein connected to NS3:NS4A via an eight amino acid linker.

3.1.7 pCMV-CFP-SCNS3

SinglechainNS3 (scNS3) was amplified from clone p28-scNS3-1b(con1) (Frick et al., 2010) plasmid using primers (Integrated DNA Technologies) which contain HindIII and XhoI restriction sites.

Forward: 5'- CGC GCG AAG CTT ATG GGC AGC AGC CAT CAT – 3'

Reverse: 5'-CGC GCG CTC GAG GAC GGA GCT CGA ATT CGG-3'

The PCR product was then cut with HindIII and XhoI restriction enzymes and ligated into a similarly treated pCMV6-AN-mCFP-LGP2 vector.

The resulting plasmid expresses a CFP protein connected to scNS3 via an eight amino acid linker.

3.1.8 pCMV-CFP-NS3-HEL1B

The NS3 helicase domain from genotype 1b open reading frame was cut from p24a-NS3h1b(con1) (Belon and Frick, 2009a) plasmid using NheI and XhoI. The resulting DNA fragment was gel purified and ligated into pCMV-AN-CFP-scNS3 which was also cut with NheI and XhoI. The resulting plasmid expresses a CFP protein connected to NS3h via an eight amino acid linker.

3.1.9 pCMV-CFP-NS3

The open reading frame encoding the NS3 protease and helicase domain from genotype 1b was cut out of p280FL-NS3(1b-con1) (Neumann-Haefelin et al., 2008) plasmid using the restriction enzymes NheI and XhoI. The resulting full length NS3 DNA fragment was gel purified and ligated into pCMV-AN-CFP-scNS3 which was also cut with NheI and XhoI. The resulting plasmid expresses a CFP protein connected to NS3FL via an eight amino acid linker.

3.1.10 pCMV-CFP-NS3-PRO

A stop codon was inserted in the linker after protease and before the helicase domain in PCMV-CFP-scNS3 (above) using the following primers.

Forward: 5'-GCTGGGGCATCTAGCGGGCTGCCGTGTGCACCCGAGGGGTTGCG-3'

Reverse: 5'-CGCAACCCCTCGGGTGCACACGGCAGCCCGCTAGATGCCACAGC-3'

Agilent-QuickChange II Site Directed Mutagenesis Kit was used to make the mutation. After amplification, the parental methylated strand was degraded using DpnI. The resulting amplicon was transformed into competent E.coli and transformants expressing the pCMV-CFP-NS3-Pro were selected for using the ampicillin resistance marker. The resulting transformants were purified and sequenced to confirm presence of the stop codon.

3.1.11 P24-CFP-NS3HEL

The open reading frame encoding CFP-NS3h was excised from pCMV-CFP-NS3h (created above) using EcoRI. The resulting DNA fragment was gel purified and ligated into a similarly treated Pet24d vector (Novagen, Inc.)

3.1.12 P24-YFP-NS3HEL

The YFP open reading frame was amplified via PCR primers which contained NdeI and NheI restriction sites from the vector PCMV6-AN-mYFP (Origene-PS100052). The amplicon was digested with NdeI and NheI and ligated into a plasmid p24a-NS3h(con1) (Belon and Frick, 2009a) which was similarly treated.

3.1.13 pCMV-YFP-RIG-I

Vector was purchased from GeneCopoeia (Rockville, MD, Product # EX-T0237-M15).

3.1.14 pCMV-CFP-RIG-I

Vector was purchased from GeneCopoeia (Rockville, MD, Product # EX-T0237-M32).

3.1.15 pCMV-YFP-MDA5

Vector was purchased from GeneCopoeia (Rockville, MD, Product # EX-T0707-M150).

3.1.16 pCMV-CFP-MDA5

Vector was purchased from GeneCopoeia (Rockville, MD, Product # EX-T0707-M32).

3.1.17 pCMV-GFP₂-MAVS

Vector was created and purchased via VectorBuilder with a two amino acid linker between the fluorescent protein and MAVS.

3.1.18 EF1-mVenus-MDA5

Vector was created and purchased through VectorBuilder and contains a 15 amino acid linker between the venus and MDA5.

3.1.19 pCMV-GFP₂-LGP2

LGP2 was cut from pCMV-YFP-LGP2 plasmid using HindIII and XhoI. The resulting DNA fragment was gel purified and ligated into pCMV-AN-GFP₂ which was also cut with HindIII and XhoI. The resulting plasmid expresses a GFP₂ protein connected to LGP2 via a two amino acid linker.

3.1.20 pCMV-GFP₂(T203I)-MAVS

A point mutation was made in the GFP₂-MAVS vector (above) purchased through GenScript such that the residue T203 in GFP₂ was mutated to T203I. GenScript made the mutation.

3.1.21 pCMV-Sapphire

The vector expressing pCMV-Sapphire was designed and purchased through VectorBuilder.

3.1.22 pCMV-GFP₂

The vector expressing pCMV-GFP₂ was designed and purchased through GenScript.

3.2 Protein Purification

YFP-NS3h and CFP NS3h purified proteins used for calibration curves and FRET binding assays were purified from BL21(DE3) *E. coli* containing either pET24-CFP-NS3h or pET24-YFP-NS3h. HCV NS3h, scNS3, and NS3FL were purified from BL21(DE3) *E. coli* containing either pET24-NS3h, pET24-scNS3, or pET24-NS3FL. The method used for purification was exactly as previously described in detail (Frick et al., 2010; Ndjomou et al., 2015).

3.3 Enzyme Assays

3.3.1 Helicase

Molecular beacon-based helicase assays (Belon and Frick, 2008) were performed in the presence of 50 nM NS3, 12.5 nM helicase substrate, and a dilution series of 12 different HPI concentrations from 1 to 100 μ M. The slopes immediately following ATP addition were analyzed to determine unwinding rates per the method previously described (Hanson et al., 2012; Ndjomou et al., 2015).

3.3.2 Protease

HCV NS3 decapeptide cleavage assays were performed with the RET-S1 substrate (Anaspec), which is based on the NS4A/NS4B cleavage site. Each 20 μ l assay contained 50 mM Tris-HCl pH 7.0, 50% glycerol, 2% CHAPS, 30 mM DTT, 5% DMSO, and 0.5 μ M RET-S1 substrate. When NS3 lacking a covalently tethered NS4A was used, reactions were supplemented with 6 μ M pep4AK (Anaspec). Reactions were first performed with various concentrations of each NS3 protein to determine its specific activity. A unit of enzyme was defined as the amount needed to cleave 0.05 μ M of substrate/min. To determine the inhibitory potential of HPI with each enzyme, reactions were performed with 1 unit of each enzyme with various concentrations of HPI. Reactions containing HPI were performed in the presence of a 16-pt two-fold dilution series of HPI starting at 100 μ M. Assays were performed in duplicate, and rates normalized to reactions performed in the absence of HPI. Initial rates of substrate cleavage were then plotted versus HPI concentration and fitted to a concentration response equation to calculate the

concentration of HPI needed to reduce initial rates by 50% using GraphPad Prism (v.6).

(Ndjomou et al., 2015)

3.3.3 Protein:DNA titrations

The interaction between DNA oligonucleotides and fluorescent fusion proteins was monitored *in vitro* using a Cary Eclipse Spectrophotometer (Agilent). The fluorescent moieties used to label DNA were “MAX” (N-hydroxysuccinimide ester) and hexachlorofluoroscein (HF). Oligonucleotides were obtained from Integrated DNA Technologies (IDT). The sequences for the fluorescently labeled DNAs are below.

5'-MAX-ATG GTT CTG AGG GTG GCG GTA CTA-3'

5'- HF-TGG CGA CGG CAG CGA GGC-3'

The fluorescently tagged protein (YFP-NS3h) was purified as described above in section 3.2.

Titration were performed in 1.25 mM MgCl₂, 12.5 mM MOPS (pH 6.5), and with various concentrations of fluorescent proteins which ranged from 50 to 300 nM per titration. Starting volumes were 500 μL and titrations were performed in a cuvette. The fluorescent protein was titrated with fluorescent DNA and a full emission spectrum was collected for 2 excitation wavelengths upon each addition of DNA.

The data for the titration between YFP-NS3h and HF-DNA was analyzed assuming that observed fluorescence intensity at a given set of wavelengths is equal to the concentration of the free acceptor [A], the free donor [D], and the donor acceptor complex [AD] times co-

efficients (b^A , b^D , or b^{AD}) describing a linear relationship between the fluorescent brightness (where brightness is equal to the extinction co-efficient of the fluorescent molecule (ϵ) multiplied by the quantum yield of the fluorescent molecule (Q)) and the concentration of a species as described by equation 3-1.

$$I_{obs} = b^A[A] + b^D[D] + b^{AD}[AD_n] \quad (\text{Equation 3-1})$$

Equation 3-1 has too many unknown variables to be solved directly and therefore the application of equation 3-1 requires that simplifying assumptions be made dependent on specifics of each experiment. The simplifying assumptions will be discussed as they apply to the application of equation 3-1 to the analysis of the titration of YFP-NS3h with HF-DNA in Chapter 5.

In equation 3-1, the A:D complex is described using the stoichiometry coefficient “n,” where n is the number of helicase (donors) bound to the DNA (acceptor), as indicated by the chemical equation below:



It is important to note here that the above model makes the simplifying assumption that there is only one form of donor acceptor complex formed when “n” donors bind to one acceptor (AD_n) without positive or negative cooperativity. Specifically, this is a 1:1 binding model where the acceptors and donors are assumed to be bound as a dimer made up of one acceptor and any “n” number of bound donors.

The equilibrium dissociation constant describing equation 3-2 (K_d) is defined as the concentration of the free acceptor [A] and the free donor [D] (modified by the coefficient “n” to represent that multiple donors may be bound to one acceptor) divided by the total concentration of bound donor and acceptor [AD_n] at equilibrium.

$$K_D = [A][nD]/[AD_n] \quad (\text{Equation 3-3})$$

The total free donor and total free acceptor can be calculated by assuming the following relationships:

$$[D] = [D]_T - [AD_n] \quad (\text{Equation 3-4})$$

$$[A] = [A]_T - [AD_n] \quad (\text{Equation 3-5})$$

Where [D]_T and [A]_T are the total amount of donor and acceptor present at any point in the titration, which are known quantities.

These assumptions allow for the calculation of the concentration of the A:D complex at each point in the titration from the K_d and the n, using equation 3-6:

$$[AD_n] = \frac{([D]_T + n[A]_T + K_d) - \sqrt{([D]_T + n[A]_T + K_d)^2 - 4*[D]_T*n[A]_T}}{2} \quad (\text{Equation 3-6})$$

Substituting equations 3-4 through 3-6 into equation 3-1 yields:

$$i_{obs} = i^A \left[A - \frac{([D]_T + n[A]_T + K_d) - \sqrt{([D]_T + n[A]_T + K_d)^2 - 4*[D]_T*n[A]_T}}{2} \right] + i^D \left[D - \frac{([D]_T + n[A]_T + K_d) - \sqrt{([D]_T + n[A]_T + K_d)^2 - 4*[D]_T*n[A]_T}}{2} \right] + i^{AD} \left[\frac{([D]_T + n[A]_T + K_d) - \sqrt{([D]_T + n[A]_T + K_d)^2 - 4*[D]_T*n[A]_T}}{2} \right] \quad (\text{Equation 3-7})$$

The titration data for the YFP-NS3h titrated with HF-DNA were fit to equation 3-7 using non-linear regression in GraphPad Prism V.5.

The data for the titration of YFP-NS3h with MAX-DNA was fit using the following equation previously defined in Raicu et al (Raicu and Singh, 2013).

$$E_{app} = \frac{1}{D_T} \cdot ([D] \cdot 0 + [AD] \cdot E_{AD}) \quad (\text{Equation 3-8})$$

where D_T is the total donors present, $[D]$ is the concentration of total free donors, $[AD]$ is the concentration of donor:acceptor complexes and E_{AD} is the pairwise FRET efficiency of the donor:acceptor complex.

Equation 3-8 assumes there is no fluorescence of the free monomers and therefore the equation can be simplified to:

$$E_{app} = \frac{[AD]}{D_T} \cdot E_{AD} \quad (\text{Equation 3-9})$$

The concentration of complex ($[AD]$) can be solved for using equation 3-6 and the titration data for the YFP-NS3h titrated with MAX-DNA were fit to equation 3-9 using non-linear regression in GraphPad Prism V.5.

To calculate FRET efficiency between the fluorescently labeled protein (YFP-NS3h) and the fluorescently labeled oligonucleotide (MAX-DNA), three different FRET techniques were used. The three methods were designed to analyze the energy lost by the donor in the presence of acceptor. In the first method, peak intensity values at donor emission (*i.e.* intensity at 525 nm recorded with the donor alone ($I_{Alone}^D (\lambda_D^{ex} \lambda_D^{em})$) and the intensity at 525 nm recorded in the presence of both donor and acceptor ($I^D (\lambda_D^{ex} \lambda_D^{em})$) were used. For a full explanation of the

terms used please see section 2.5.3. Equation 2-4 (reproduced below) was used to calculate F_{FRET} .

$$F_{FRET} = 1 - \frac{I^D(\lambda_D^{ex} \lambda_D^{em})}{I_{Alone}^D(\lambda_D^{ex} \lambda_D^{em})} \quad (\text{Equation 2-4})$$

Where $I^D(\lambda_D^{ex} \lambda_D^{em})$ was the fluorescence intensity recorded at 525 nm (slit width 10 nm) upon excitation at 480 nm (slit width 10 nm) and recorded in the presence of acceptor and $I_{Alone}^D(\lambda_D^{ex} \lambda_D^{em})$ is the same value recorded in the absence of acceptor.

The second method was designed to separate the contributions of the donor and acceptor spectra from the composite spectra and to account for any background fluorescence (see section 2.5.6). Using spectral unmixing, F^D is obtained from unmixing the spectrum of donor alone to obtain k^D . Then the k^D value was multiplied by the normalized spectral integral of the spectrum of the donor alone to account for all the photons contained within the full donor spectrum. The F^{DA} values were obtained by unmixing spectra recorded in the presence of various amounts of acceptor to obtain k^{DA} . Then the k^{DA} values were also multiplied by the normalized spectral integral of the spectrum of the donor alone to account for all the photons contained within the full donor spectrum. This generated F^{DA} values for each point in the titration. E_{app} was then calculated using equation 2-1:

$$E_{app} = 1 - \frac{F^{DA}}{F^D} \quad (\text{Equation 2-1})$$

The third method was used to simulate what would need to be done in a microscope when imaging cells expressing fluorescent proteins. It is important to note that direct measurements of F^D are not possible to obtain when cells expressing fluorescent proteins

are imaged. F^D values were therefore back calculated from F^{AD} and F^{DA} values obtained from spectral unmixing using gamma ratios and the quantum yields of the donor and acceptor per equation 3-10. Equation 3-10 will be discussed in section 3.5.5.

$$F^D(\lambda_{ex,1}) = F^{DA}(\lambda_{ex,1}) + \frac{Q^D}{Q^A} F(\lambda_{ex,1}, \lambda_{ex,2}) \quad (\text{Equation 3-10})$$

E_{app} was then calculated using the calculated F^{DA} values and the back-calculated F^D values using equation 2-1:

$$E_{app} = 1 - \frac{F^{DA}}{F^D} \quad (\text{Equation 2-1})$$

3.4 Cell-based Assays

3.4.1 HCV subgenomic replicon assays.

All subgenomic replicon containing Huh-7.5 cells were isolated and maintained as previously described (Ndjomou et al., 2012). To assess the ability of each compound to inhibit HCV replication, Huh7.5 cells harboring HCV *Rluc* subgenomic replicons were seeded at 10×10^3 cells per well in 96-well plates and incubated 4-5 hours to allow the cells to attach to the plate. Two-fold serial compound dilutions were made in dimethyl sulfoxide (DMSO), and diluted into media, such that the DMSO final concentration was 0.5% after adding dilutions to cells. Compounds and cells were incubated at 37 °C in 5% CO₂. After 3 days, *Renilla* luciferase was measured using Promega's *Renilla* luciferase assay kit. Quantitative reverse transcriptase PCR using Taqman probes specific to conserved sequences in the HCV 5'UTR was used to measure

relative RNA levels as previously described (Ndjomou et al., 2012). PCR data were normalized to RNA levels observed in cells incubated with DMSO only (Ndjomou et al., 2015).

3.4.1.1 Combination studies

Two-fold serial dilutions, starting at ~4 times the EC₅₀ values obtained in concentration-response assays with each compound alone, were prepared in 96-well plates. 50 µl of each dilution of the first compound to be tested (at twice the final compound concentration) was added horizontally and 50 µl of each dilution of the second compound was added vertically to make a 1x final concentration of each compound. After 3 days at 37 °C, luciferase activity was measured. At least four separate experiments (in triplicate) were performed for each combination (Ndjomou et al., 2015).

3.4.1.2 In-vivo assay data analysis

Half-maximal inhibitory concentrations were calculated from concentration response curves using non-linear regression to fit data to a log(inhibitor) vs. normalized response equation with variable slope. The effects of drug-drug combinations were evaluated using the MacSynergy II (<http://www.uab.edu/medicine/peds/macsynergy>) (Prichard and Shipman, 1990). MacSynergy II uses the Bliss independence model to estimate synergy and/or antagonism. Additive inhibition is calculated in MacSynergy using the equation $Z = X + Y(1-X)$, where X and Y correspond to the inhibitory effects of compound 1 and 2 respectively, and Z is the effect produced when the two compounds are combined. Resulting values are then subtracted from the normalized inhibition observed at each drug combination to estimate

synergy (positive values) or antagonism (negative values). Synergy/antagonism values calculated using the MacSynergy II Excel spreadsheet at the 95% confidence levels were plotted in three dimensions using DeltaGraph 6 (Redrock Software) (Ndjomou et al., 2015).

3.4.2 Cell Viability Assays.

To assess compound effects on Huh-7.5 or BHK DenV-Rluc cell viability, cells were plated and treated as above for compound inhibitory activity and the effect of compound on cell viability was tested using the CellTiter-Glo luminescent cell viability kit (Promega) (Ndjomou et al., 2015).

3.4.3 Transfection with plasmids

HEK293T cells were grown to 90% confluence in 10-cm cell culture dishes in 37 °C and 5% CO₂ in complete Hyclone (Hyclone labs, Logan, UT) Dulbecco's Modified Eagle Medium (DMEM)/High Modified Media with 4,500 mg/l glucose, 110 mg/l sodium pyruvate (ThermoFisher Scientific). The media was supplemented with 10% fetal bovine serum (FBS), 1X non-essential amino acids, and 1X penicillin/streptomycin (ThermoFisher Scientific).

The HEK293T cells were trypsinized, and diluted to 1.0×10^5 cells per mL. Then 2 mL (2.0×10^5) cells were seeded to 35 mm petri dishes, with a 14 mm diameter Microwell, No. 1.5 coverglass (0.16-0.19mm) poly-D-lysine coated glass bottom culture dishes (MatTek Corp., Ashland, MA) and allowed to grow for 24 hours in 37C under 5% CO₂.

Aliquots of 1.5 ug of each plasmid was diluted in 600 uL of serum free DMEM growth media with 5 uL of Turbofect Transfection Reagent (Thermoscientific #R0531). The mixture sat at 23 °C for 1 hour. The mixture was then added to the cells in the glass bottom culture dishes and the dishes were returned to 37°C and 5% CO₂ for 24 hours to express the recombinant fluorescent protein construct (Corby et al., 2017).

3.4.4 Intracellular protease assays

A MAVS-based reporter assay was used for cell-based protease assay. 293T cells were co-transfected with plasmids encoding NS3 and the plasmid TRIP-RFP-NLS- IPS (Charles M. Rice, Rockefeller University), which encodes the “TagRFP” red fluorescent protein (RFP, Evrogen, Inc.) and a SV40 nuclear localization signal (NLS) fused to residues 462–540 of MAVS (aka IPS). Cells were imaged as described below (Corby et al., 2017).

3.4.5 Transfections with polyinosinic–polycytidylic acid (Poly I:C)

Four hours prior to imaging, either 200 µl of poly I:C (5 ng/µl) in DMEM plus TurboFect (1:100 dilution) or 200 µl of serum free DMEM plus TurboFect (1:100 dilution) was added to the dishes containing transfected cells. The plates were then returned to the 37 °C, 5% CO₂ incubator for 4 hours (Corby et al., 2017)

3.4.6 ISRE Assays

3.0 x10⁴ of HEK 293T cells per well were plated per well of a 24-well plate. After 12 h at 37 °C, 0.5 µg of the pISRE-luc (Agilent Technologies, Inc., product #219092) plasmid expressing

firefly luciferase under control of and interferon stimulated regulatory element (IRSE) was co-transfected along with either 0.5 μ g of plasmid expressing YFP alone or YFP-RIG-I, YFP-MDA5, or YFP-LGP2 using TurboFect (as above). After 24 h at 37 $^{\circ}$ C, luciferase activity was measured using the Firefly Luc One-Step Flow Assay Kit (Pierce) (Corby et al., 2017).

3.5 Microscopy

3.5.1 Choosing FRET pairs

Initial experiments used the cyan fluorescent protein (CFP) and the yellow fluorescent protein (YFP) as the chosen FRET pair. This pair requires the use of a correction for direct excitation (explained below) due to the excitation of YFP at the 840 nm CFP peak excitation wavelength (Figure 3-1). This correction for direct excitation requires two excitation wavelengths and the time required between the two measurements allows enough time for

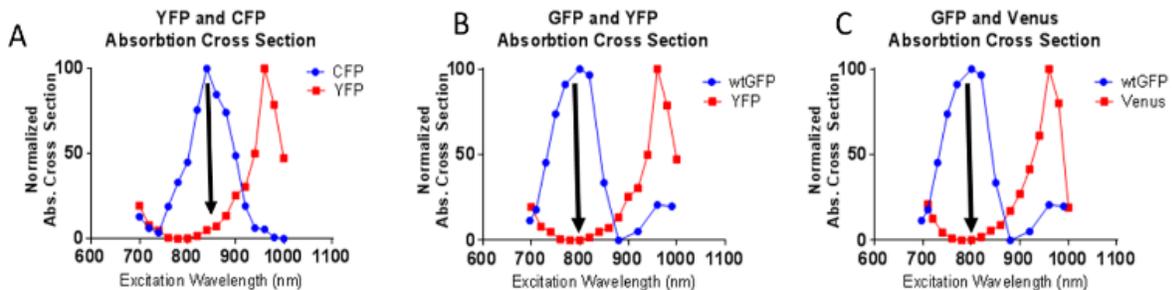


Figure 3-1: Comparison of fluorescent protein spectrums. (A) Normalized absorption cross section of YFP and CFP with arrow indicating the YFP absorption present at the CFP excitation peak. This YFP absorption requires a correction for direct YFP excitation whenever a YFP and CFP pair is used. (B) Normalized absorption cross section of YFP and wild-type GFP (which has a similar absorption spectrum to GFP₂). Arrow indicates that the peak GFP absorption corresponds to no YFP absorption, therefore negating the need for a correction of direct excitation of the acceptor at the donor peak excitation wavelength. (C) Normalized absorption cross section of Venus and wild-type GFP (which has a similar absorption spectrum to GFP₂). Arrow indicates that the peak GFP absorption corresponds to no Venus absorption, therefore negating the need for a correction of direct excitation of the acceptor at the donor peak excitation wavelength.

molecular diffusion. Therefore, only average E_{apps} (FRET efficiencies) over the regions of interest were considered in the analysis of this data.

To move away from the calculation of average E_{apps} and toward the generation of pixel level E_{apps} , the tags were switched to enhanced green fluorescent protein 2 (GFP₂) and YFP or Venus. GFP₂ is very similar to Clontech's enhanced green fluorescent protein (eGFP) molecule which incorporates mutations which allow it to fold at 37 °C and increases overall stability. The difference between GFP₂ and Clontech's eGFP exists in just two amino acid changes. The GFP₂ protein incorporates an A206K mutation which prevents the inherent dimerization which all fluorescent proteins are prone to (von Stetten et al., 2012; Zacharias et al., 2002) and it also returns the eGFP S65T mutation to the wild-type S65 which gives the GFP₂ molecule an

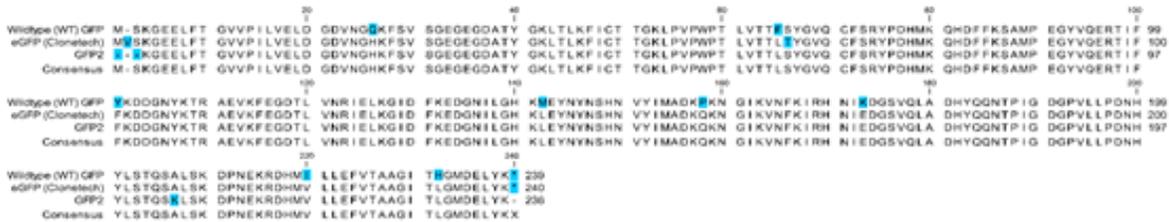


Figure 3-2: Alignment of wild-type GFP (avGFP), Clontech's eGFP, and GFP₂ protein sequences. The GFP₂ protein retains all the characteristics of Clontech's eGFP molecule with the exception of two mutations. Clontech eGFP includes a mutation at position 66 (in this alignment, S66T)) which alters eGFP excitation spectra to prefer a predominantly 475 nm (960 nm-two photon) excitation. Additionally, eGFP retains the wildtype A207 (in this alignment) which means that eGFP is prone to dimerization. Both mutations are indicated by arrows in the alignment. The GFP₂ protein retains the original wild-type GFP S66, thereby giving the GFP₂ protein an excitation spectra very similar to wild-type GFP and a preference to the 395 nm (800 nm-two photon) excitation. The GFP₂ protein also incorporates an A207K mutation, which significantly reduces protein dimerization.

excitation spectra practically identical to wild-type GFP with a peak excitation at about 395 nm (Heim et al., 1995) (Figure 3-2).

By using the GFP₂ protein, which exhibits a large Stokes shift, in conjunction with either YFP or Venus fluorescent protein, the donor can be excited at about 395 nm and completely avoid direct excitation of the acceptor (Figure 3-1). This eliminates the need for a direct excitation of the acceptor correction and allows for the calculation of E_{apps} (FRET efficiencies) at pixel level instead of an average over a larger region of interest. This is preferable because it allows for the generation of pixel level FRET maps which can more precisely locate the origin of FRET interactions and increases the amount of FRET data present in each region of interest from one averaged value to tens or hundreds of values depending on the size of the region of interest.

3.5.2 Imaging set up and Mito-Tracker staining

The cell medium contains the pH indicator, phenol red. To remove phenol red prior to imaging, the cell medium was removed using aspiration and replaced with pre-warmed serum-free DMEM lacking phenol red. To stain mitochondria, cell medium was again removed, and cells were stained in 1 ml of 25 ng/ml MitoTracker-Red (ThermoFisher Scientific) in pre-warmed serum-free DMEM lacking phenol red for 15 minutes. Cell medium was removed using aspiration and replaced with pre-warmed serum-free DMEM lacking phenol red (Corby et al., 2017).

3.5.3 Collection of microscopic images

Plates containing cells were removed from the incubator and washed once with pre-warmed serum-free DMEM lacking phenol red. Then two milliliters of pre-warmed serum-free DMEM lacking phenol red was added to each plate for imaging. The images were collected at room temperature using a two-photon inverted microscope (Nikon Eclipse TiTM (Nikon Instruments Inc., Melville, NY). The microscope is equipped with an OptiMiS scanning/detection head (Aurora Spectral Technologies, Grafton, WI) which was modified to include a line scan protocol (Biener et al., 2013) which allows for the detection of larger signals. In addition, the OptiMiS system includes the ability to switch between excitation wavelengths while maintaining the same power (Stoneman et al., 2017). Fluorescent excitation was achieved through use of a mode-locked laser (MaiTaiTM, Spectra Physics, Santa Clara, CA). The mode-locked laser is tunable between 690 nm and 1040 nm and produces 100 femtosecond pulses with a 7 nm full-width-half-maximum.

The laser was focused using an infinity-corrected oil-immersion objective with 100X magnification (Nikon Instruments, Inc.) The fluorescence detection occurred through a transmission grating whereby the light was collected on a cooled electron multiplying CCD (EMCCD) camera (iXon X3 897; Andor Technologies, Belfast, UK). The spectrum is separated as a function of pixel position on the EMCCD array allowing for the collection of spatial, physical dimensions (X and y) and 200 wavelengths simultaneously. The 200 wavelengths collected range from approximately 200 nm to 600 nm with a maximum spectral resolution of about 1

nm. Custom software was written in C++ and developed by the Raicu lab at University of Wisconsin-Milwaukee.

Before each experiment, a spectral calibration is performed using 10 μM uranine and rhodamine standards which assigns each wavelength to a known camera pixel. Two-photon images were collected for each plate of 293T cells expressing either singly transfected CFP or YFP tagged proteins or a combination of CFP (donor) and YFP (acceptor) tagged proteins.

3.5.4 Un-mixing of spectral images

The first step in spectral un-mixing requires the generation of elementary spectra. The spectra obtained from those cells singly expressing either CFP or YFP tagged proteins was used to generate the elementary donor and acceptor spectrums, respectively. Cells expressing no fluorescently tagged protein but stained with a MitoTracker-red dye were used to determine the elementary spectra of the MitoTracker.

To generate the composite spectrum, a “FRET” scan was performed on the cells expressing the combination of CFP and YFP tagged proteins using the peak excitation of the donor ($\lambda_{\text{ex1}} = 840 \text{ nm}$). To generate an image primarily exciting and targeting the acceptor fluorescence in the presence of donor, a second scan is taken at the acceptor peak excitation ($\lambda_{\text{ex2}} = 960 \text{ nm}$). This second scan can be termed the “acceptor scan.” About 15-30 field of views at 100X magnification were collected for each dish analyzed. The FRET scan and the acceptor scan are taken sequentially for each field of view and the image acquisition time is about 10 seconds for each excitation scan and about 40 seconds to change the wavelength between the

FRET and acceptor scans. The total image acquisition time is therefore about 60 seconds to collect two sets of images each containing 440 x 300 pixels and 20 wavelengths per pixel. The average laser power used during the measurements was about 200 mW per line for both the FRET and acceptor scans and the power was held constant through the modified OptiMiS system previously described in section 3.5.3. (Corby et al., 2017; Stoneman et al., 2017)

The composite spectra was deconvoluted (un-mixed) (Raicu et al., 2005) at pixel level through the assumption that the composite spectra is a linear combination of each of the elementary spectra per equation 2-17. In these experiments, the contributing spectra are donor in the presence of acceptor ($I=DA$), acceptor in the presence of donor ($I=AD$), and in some cases, Mitotracker ($I=M$). In addition, a background, wavelength-independent component which is nearly constant is also included ($I=B$).

The coefficients equal to the fluorescent contributions of each of the above mentioned fluorescent species (k^{DA} , k^{AD} , k^M) were extracted using a least-squares minimization method previously described (Biener et al., 2013).

3.5.5 E_{app} (FRET) calculation

The calculation of E_{app} (FRET) begins with the calculation of the total number of photons detected for donors, acceptors, and (in some experiments) MitoTracker (Raicu et al., 2005). The abundance coefficients (k^{DA} , k^{AD} , k^M) for the donor, acceptor, and MitoTracker are multiplied by the spectral integral (denoted by w with various superscripts) for each donor, acceptor, and MitoTracker signal collected over the wavelengths analyzed via the following equations:

$$F^{DA}(\lambda_{ex,i}) = k^{DA}(\lambda_{ex,i}) * w^D \quad (\text{Equation 3-11})$$

$$F^{AD}(\lambda_{ex,i}) = k^{AD}(\lambda_{ex,i}) * w^A \quad (\text{Equation 3-12})$$

$$F^M(\lambda_{ex,i}) = k^M(\lambda_{ex,i}) * w^M \quad (\text{Equation 3-13})$$

The calculation of F^{DA} , F^{AD} , and F^M allow for the generation of fluorescence intensity maps where the total fluorescent intensity at each pixel is equal to the total number of photons of the specific spectra contribution at each pixel. These maps were then used to generate regions of interest focused on the fluorescence intensity of the protein or marker of interest (Figure 3-3).

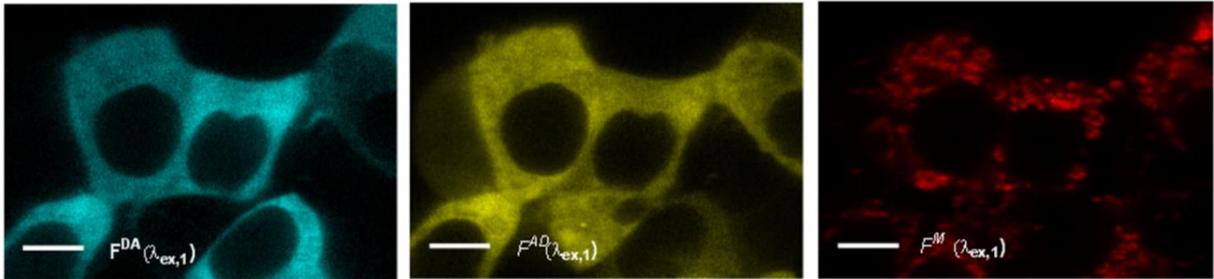


Figure 3-3 Fluorescence intensity map of the donor (A), acceptor (B), and Mito-Tracker (C). When whole cells were selected, the donor intensity map (A) was used in conjunction with the acceptor intensity map (B) to select whole cell regions of interest expressing both donor and acceptor proteins. When the mitochondria were targeted, the Mito-Tracker intensity map (C) was used to circle mitochondrial regions of interest.

The F^{DA} , F^{AD} , and F^M values were used to calculate the average donor fluorescence in the absence of FRET and the average acceptor fluorescence in the absence of FRET (i.e. direct excitation of the acceptor) using the previously derived (Mishra et al., 2016; Patowary et al., 2015; Stoneman et al., 2017) equation:

$$F^D(\lambda_{ex,1}) = F^{DA}(\lambda_{ex,1}) + \frac{Q^D}{Q^A} F(\lambda_{ex,1}, \lambda_{ex,2}) \quad (\text{Equation 3-10})$$

$$F^A(\lambda_{ex,2}) = \frac{F(\lambda_{ex,1}, \lambda_{ex,2})}{\rho^A - \rho^D} \quad (\text{Equation 3-14})$$

Where

$$\rho^A = \frac{\Gamma^{\lambda_{ex,1,A}}}{\Gamma^{\lambda_{ex,2,A}}} \quad (\text{Equation 3-15})$$

$$\rho^D = \frac{\Gamma^{\lambda_{ex,1,D}}}{\Gamma^{\lambda_{ex,2,D}}} \quad (\text{Equation 3-16})$$

$$F^A(\lambda_{ex,1}) = \frac{\Gamma^{\lambda_{ex,1,A}}}{\Gamma^{\lambda_{ex,2,A}}} * F^A(\lambda_{ex,2}) \quad (\text{Equation 3-17})$$

$$F^D(\lambda_{ex,2}) = \frac{\Gamma^{\lambda_{ex,2,D}}}{\Gamma^{\lambda_{ex,1,D}}} * F^D(\lambda_{ex,1}) \quad (\text{Equation 3-18})$$

$$F(\lambda_{ex,1}, \lambda_{ex,2}) = F^{AD}(\lambda_{ex,1}) - \rho^A F^A(\lambda_{ex,2}) \quad (\text{Equation 3-19})$$

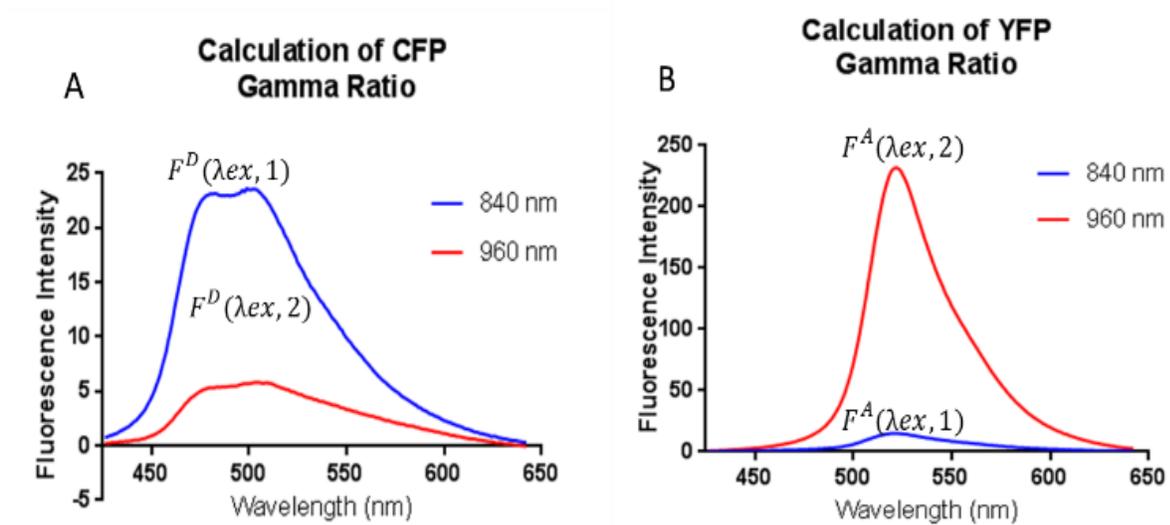


Figure 3-4: Calculation of the gamma ratio for CFP and YFP. The peak emission of CFP when excited at both excitation wavelengths (840 nm and 960 nm) generates the $F^D(\lambda_{ex,1})$ and the $F^D(\lambda_{ex,2})$ used in the gamma ratio calculation for the donor (A). The peak emission of the YFP when excited at both excitation wavelengths generates the $F^A(\lambda_{ex,1})$ and the $F^A(\lambda_{ex,2})$ used in the gamma ratio calculation for the acceptor (B).

The ρ^D is the ratio between the emission peak of the donor at excitation wavelength 1 and excitation wavelength 2 and ρ^A is the ratio between the emission peak of the acceptor at the excitation wavelength 1 and the excitation wavelength 2 (Figure 3-4).

The average FRET efficiency (E_{app}) over the circled regions of interest could then be calculated using the following equation which defines FRET as the ratio of the donor intensity in the presence of acceptor (quenched due to FRET) and the total donor fluorescence in the absence of FRET using equation 2-2 shown below.

$$E_{app} = 1 - \frac{F^{DA}(\lambda_{ex,1})}{F^D(\lambda_{ex,1})} \quad \text{(Equation 2-2)}$$

3.5.6 Donor and acceptor concentration calculation

To calculate the total concentration of the donor-tagged proteins and the acceptor tagged proteins in the dual-expressing cell samples, the values for $F^D(\lambda_{ex,1})$ and $F^A(\lambda_{ex,2})$ were applied to a standard curve of fluorescence intensity. The standard curve was generated using a ten-fold dilution series of purified CFP-NS3h (donor) and YFP NS3h (acceptor) which were imaged under the exact same conditions as the 293T cells. Calibration curves were generated by plotting $F^D(\lambda_{ex,1})$ vs. CFP-NS3h and $F^A(\lambda_{ex,2})$ vs YFP NS3h. The slopes of the calibration curves therefore correspond to the number of fluorescence counts per μM of solution. A value of 958 counts/ μM and 5,317 counts/ μM were found for the donor and acceptor respectively. The pixel-level concentrations were averaged over the whole region of interest to obtain the average concentration of donor and acceptor molecules (Corby et al., 2017).

3.5.7 Pixel level FRET calculation

Pixel level FRET calculation allows for the collection of more E_{app} values per unit area and thus a more detailed analysis of sub-cellular regions where many oligomeric complexes may be interacting and generating a population of different FRET efficiencies. If the selection of FRET pairs is carefully made such that the acceptor is not directly excited by the laser light used to excite the donor then the calculation of pixel level FRET becomes possible because the correction for direct excitation of the acceptor at the first excitation wavelength is no longer necessary and there is no time lost between excitations (Raicu et al., 2009). If there is no direct excitation of the acceptor at the donor excitation wavelength then the fluorescence at excitation wavelength 1 becomes zero and equation 2-28 (derived in Section 2.5.6) may be used to calculate E_{app} .

$$E_{app} = \frac{1}{1 + \frac{Q^A k^{DA} \omega^D}{Q^D k^{AD} \omega^A}} \quad (\text{Equation 2-28})$$

Once pixel level E_{app} values are calculated, the E_{app} values are organized into histograms based on selected regions of interest. The histograms may have several unique peaks which can indicate the presence of varying combinations of donor and acceptors within a larger oligomeric complex. To understand the predominant peaks and features over all pixels collected into histograms by selected regions of interest, the predominant peak from each histogram was recorded and a meta-histogram of all the histogram peaks was generated. This meta-histogram represents the dominant oligomeric configurations within the larger population (Stoneman et al., 2017).

Chapter 4 A Novel Inhibitor Can Inhibit Both Helicase and Protease Activities of HCV NS3 Protease/helicase

The hepatitis C virus (HCV) NS3 protein is a protease and helicase/ATPase. All three activities are essential for HCV replication and therefore all three activities are potential drug targets but currently, only protease inhibitors are approved as drugs. Previously, a set of HCV helicase inhibitors were designed based on the yellow dye, primuline. One of these inhibitors (HPI PubChem CID #50930749) had the ability to inhibit the helicase and the protease domains but not the ATPase domain *in vitro*. Further, HPI was able to inhibit HCV sub genomic replicons in Huh 7.5 cells (Ndjomou et al., 2012). Based on the ability of HPI to inhibit both sub-genomic replicons in cells and the protease domain and helicase domain in *in vitro* assays, it is hypothesized that HPI interacts in the allosteric groove between the NS3 protease and helicase domains and therefore, is able to make contacts which inhibit both activities.

This study strives to understand the mechanism of action of HPI through both *in vivo* and *in vitro* assays. We examine how HPI inhibits HCV sub-genomic replicons of differing genotypes and how HPI inhibition is affected by the presence of first generation linear protease inhibitors and second generation macro-cyclic protease inhibitors in HCV sub-genomic replicons. To precisely target the putative HPI binding site, HCV full length proteins harboring several mutations in both the NS3 protease domain and the allosteric domain between the protease and helicase were tested in an *in vitro* proteolytic cleavage assay in the presence of HPI. In addition, full length HCV NS3 and NS3 with a covalently linked 4A protein were tested in the *in vitro* protease assay with HPI.

4.1 HPI is specific to HCV genotypes 1b, 3a, and 4a

HPI was previously shown to inhibit both the protease and helicase activities of NS3 *in vitro* and was shown to have activity against Huh 7.5 cells stably expressing HCV sub-genomic replicons with a renilla luciferase reporter from genotype 1b (Ndjomou et al., 2012). One explanation for this broad inhibition would be if HPI non-specifically bound to HCV NS3. To test this, an HPI dilution series was applied to cells stably expressing an HCV subgenomic replicon with renilla luciferase reporter from genotype 2a, cells stably expressing a renilla luciferase tagged dengue virus replicon and to cells stably expressing the HCV genotype 1b sub-genomic

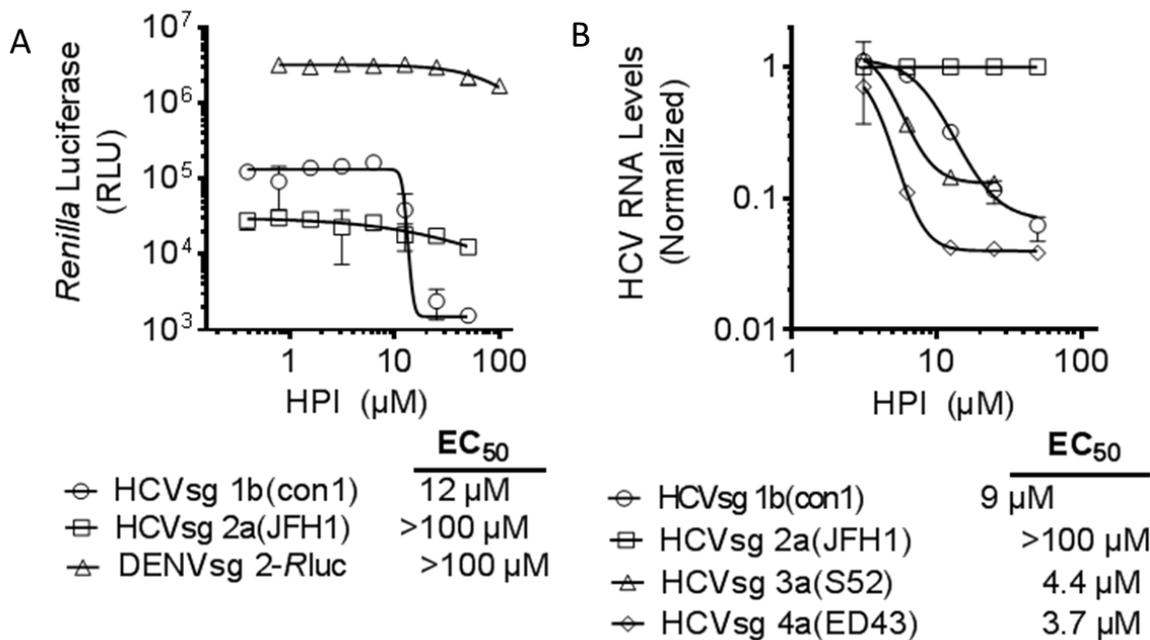


Figure 4-1: HPI is specific to HCV genotypes 1b, 3a, and 4a. A sub-genomic HCV replicon for genotypes 1b and 2a with a Renilla luciferase (R-luc) reporter and a Dengue virus sub-genomic replicon with an R-luc reporter was expressed in cells. The cells were treated with HPI and the R-luc activity measured after treatment (A). The EC_{50} was calculated and HCV 2a and Dengue virus were insensitive to HPI. HCV 1b was sensitive to HPI with an EC_{50} of 12 μM (A). Cells were also transfected with sub-genomic replicons from HCV 1b, 2a, 3a, and 4a. RT-qPCR was performed and RNA levels quantified. Again, HCV genotype 2a was insensitive to HPI, whereas HCV genotypes 1b, 3a, and 4a were sensitive to the drug (B).

replicon. Similar to previous experiments, HPI showed activity against HCV sub-genomic replicon 1b with an EC_{50} of about $12\mu\text{M}$ with no effect on cell viability. However, it showed no activity against genotype 2a or dengue virus even in concentrations up to $100\mu\text{M}$ (Fig 4-1A) (Ndjomou et al., 2015).

Next to see the effect on a wider range of HCV genotypes, a real-time polymerase chain reaction assay was performed on Huh 7.5 cells with sub-genomic replicons from genotypes 1b, 2a, 3a, and 4a. The replication of HCV RNA was significantly inhibited for genotypes 1b, 3a, and 4a with EC_{50} s of $9\mu\text{M}$, $4.4\mu\text{M}$, and $3.7\mu\text{M}$ respectively. But HCV genotype 2a was again shown to be insensitive to HPI with concentrations up to $100\mu\text{M}$ (Figure 4-1B). HPI is specific to genotypes 1b, 3a, and 4a and therefore it is unlikely that HPI is binding non-specifically to HCV NS3 (Ndjomou et al., 2015).

4.2 Effect of HPI on NS3 harboring mutations between the protease and helicase domains

Due to the ability of HPI to inhibit both the helicase and protease function, it's possible that HPI makes contacts in the allosteric groove between the helicase and protease domain. Molecular modeling further indicated that this may be the case. To test this, HPI was tested in a peptide cleavage assay against the activity of a recombinant NS3 with a covalently tethered N-terminal 4A protein (scNS3). In addition to this wild-type protein, HPI was also tested against the activity of the recombinant scNS3 with mutations in the predicted HPI binding site (S483A, M485A, and V524A) and in or near the binding site for peptido-mimetic inhibitors (F438A, Q526A, and H528A) (Figure 4-2A). The activity of these mutations has previously been

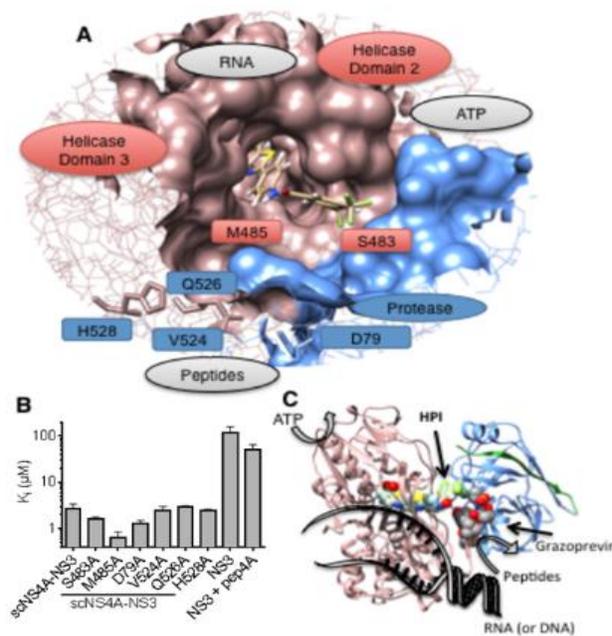


Figure 4-2 Evaluation of putative HPI binding site. The HPI binding site was modeled along with residues in and around that binding site (A). Residues in the HPI binding site and those in the neighboring protease domain were mutated and expressed in a recombinant scNS3 helicase. The mutant scNS3 helicases were tested in a protease cleavage assay to determine the effect of HPI on protease activity when expressing the different binding site mutations. The results indicate that residue M485 is likely blocking the HPI binding site and that mutating that residue to an alanine increases the sensitivity of HCV NS3 to HPI (B). To determine whether or not HCV NS3 could accommodate HPI and a protease inhibitor, a crystal structure of the enzyme bound to Grazoprevir was used and HPI was integrated into the model (C). The model shows that HCV NS3 can accommodate both Grazoprevir and HPI.

characterized (Aydin et al., 2013) and they were all found to retain the ability to bind and unwind nucleic acids as well as cleave peptides and ATP.

The inhibitory potential of HPI was tested with the wild-type scNS3 and the scNS3 proteins with various mutations in an HCV peptide cleavage assay. In this assay, the M485A mutation had the largest effect on the ability of HPI to inhibit NS3 protease catalyzed peptide hydrolysis. When M485 is replaced with a smaller alanine residue, the protein was much more sensitive to inhibition by HPI. A difference of about 5-fold in comparison to wild-type scNS3. It is possible that the smaller alanine allows HPI greater access to the NS3 putative HPI binding site. Both the mutations in the peptido-mimetic inhibitor binding site did not change the sensitivity of NS3 to HPI, implying that HPI binds in a way that is different from peptido-mimetic inhibitors (Figure 4-2B) (Ndjomou et al., 2015). Peptido-mimetic inhibitors have previously been shown to be affected by the Q526A and H528A mutations (Howe et al., 1999).

4.3 Effect of HPI on the activity of scNS3 and NS3:NS4A

To further understand the mechanism of action of HPI against HCV NS3, HPI was tested against different recombinant HCV NS3 proteins using the peptide cleavage assay. The scNS3 protein (discussed above) which contains a covalently linked NS4A peptide on the N-terminal protease, the NS3 full-length version, and the NS3 full-length version with the addition of exogenous NS4A peptide to enhance protease function. Surprisingly, scNS3 was much more sensitive to HPI with a K_i below 10 μM whereas the K_i for NS3 + 4A was closer to 70 μM . The NS3 full-length protein without the 4A peptide was the least sensitive to HPI with a K_i around 100 μM (Figure 4-2B) (Ndjomou et al., 2015).

4.4. Synergy between HPI and macrocyclic protease inhibitors in sub-genomic replicons

The combined effect of HPI and protease inhibitors was tested using HCV sub-genomic replicons stably expressed in Huh 7.5 cells. To determine the EC₅₀ for each compound (HPI, telaprevir, boceprevir, grazoprevir, and danoprevir) a serial dilution of each compound was applied to the HCV replicon cells and after 72 hours of treatment, the *R.luc* activity was measured. The resulting EC₅₀ values are reported in Table 4-1. In addition, cell viability assays were performed and the concentrations of compounds necessary to reduce the cell viability by half (CC₅₀) was always much greater than the concentration required to inhibit replicon activity (*R.luc* signal) by 50% (EC₅₀) (Table 4-1).

Compound	EC ₅₀ (μM)	CC ₅₀ (μM)
HPI	15 ± 4	>100
Telaprevir	0.81 ± 0.4	45 ± 4
Boceprevir	0.64 ± 0.1	>100
Danoprevir	0.002 ± 0.0005	>100
Grazoprevir	0.00076 ± 0.00013	54 ± 5

Table 4-1 The ability of HPI and other direct acting antivirals to inhibit the HCV genotype 1b subgenomic replicons and Huh7.5 cell viability. HPI and other direct acting antivirals were applied to Huh7.5 cells harboring HCV genotype 1b subgenomic replicons. The EC₅₀ values are equal to the concentration of each compound needed to reduce the HCV-1b-subgenomic replicons by 50% relative to cells treated with DMSO alone. CC₅₀ values are the concentration of compound necessary to reduce cell viability by 50%.

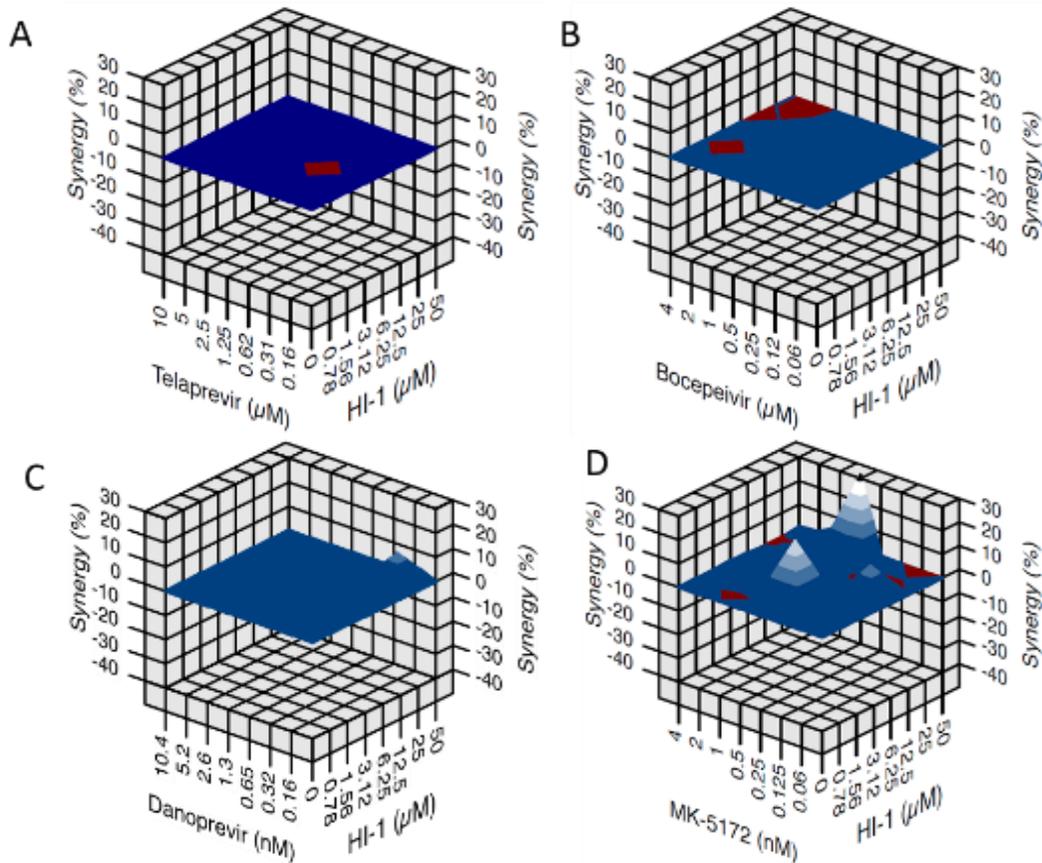


Figure 4-3: Interactions between HPI and other protease inhibitors. Points in line with the plane represent the inhibition which would be expected if the two inhibitors functioned additively as defined by the Bliss independence model. Any signal above the plane in the Z direction represents inhibition seen that is greater than what would be expected from additive inhibition alone. Data were analyzed using MacSynergy II (Prichard & Shipman, 1990).

Using the calculated EC_{50} values, concentrations were determined above and below the EC_{50} for each compound. These concentrations were then applied in different combinations to the HCV sub-genomic replicon cells. The expected additive inhibition from the combination of each drug was determined via the Bliss Independence Model. The Bliss Independence Model assumes that each drug will inhibit according to its individual inhibitory potential at a given concentration and will not be affected by the other drug. Therefore, the combination of two drugs should result in the same total inhibition as if both drugs were applied independently and

each of the drug's inhibitory ability was added. If, there is a significant deviation from this expected additive inhibition, it can be assumed that there is some interaction between the two drugs. If the inhibition with a combination of two drugs is greater than the expected additive inhibition, then the effect is said to be synergistic.

No difference from expected additive inhibition (defined as the plane in Figure 4-3 three-dimensional plots) was observed for the combination of HPI with linear protease inhibitors, telaprevir and boceprevir. Limited synergy was observed when HPI was used in conjunction with danoprevir. Modest synergy was observed when HPI was applied in conjunction with grazoprevir (Figure 4-3) (Ndjomou et al., 2015).

4.5 Significance of the experimental results on the mechanism for HPI mode of inhibition

In this study, HPI (which inhibits both helicase unwinding and protease-catalyzed peptide cleavage) was shown to have activity specific to HCV genotypes 1b, 3a, and 4a (Figure 4-1). HPI appears to behave differently than current peptido-mimetic protease inhibitors

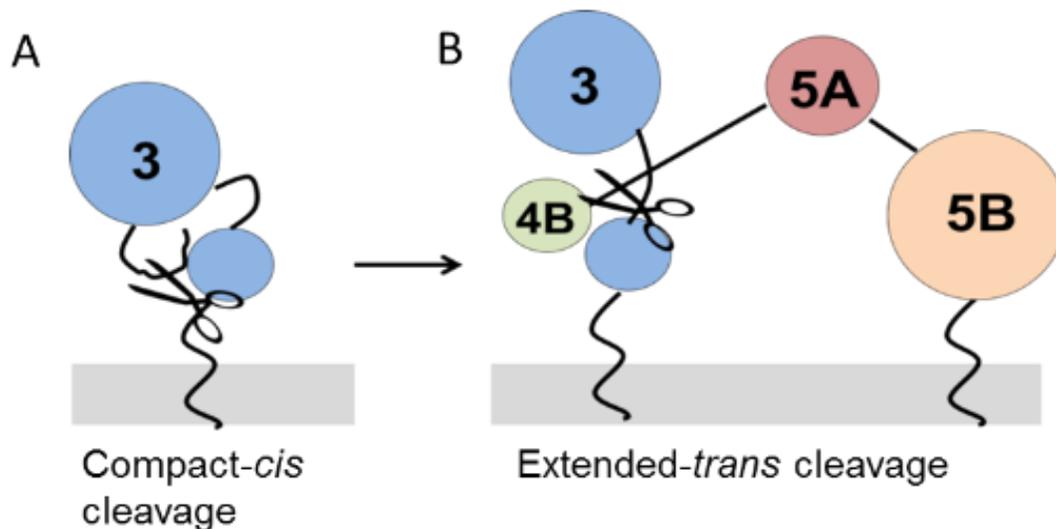


Figure 4-4 HCV NS3 can cleave substrates *in cis* and *in trans*. HCV NS3 helicase first cleaves itself *in cis* to free its C-terminal from the polypeptide (A). It is proposed that this cleavage occurs when the protein is in a closed conformation with the protease domain sitting close to the helicase domain. HCV NS3 assumes an extended conformation to cleave all other downstream targets *in trans* (B). In the extended conformation, the NS3 protease domain is swung away from the helicase domain.

because it is able to inhibit the helicase function in addition to the protease function. Despite the fact that second-generation macrocyclic protease inhibitors were designed to make contacts in both the protease and helicase domain, they are still not able to inhibit helicase function. HPI is proposed to bind in the allosteric site between the protease and helicase domains (Figure 4-2C).

By binding in the allosteric site between the protease domain and helicase domain, HPI may affect the conformation of the NS3 protein. The conformation of the protein is thought to have biological significance in the processing of the HCV polypeptide. Host and viral proteases cleave the HCV polypeptide. HCV NS3 protease is important for cleaving the junction between itself and 4A *in cis* and then cleaving all other downstream non-structural proteins *in trans* (see figure 2-4 in the methods). In order to cleave both *in cis* and *in trans*, HCV NS3 must adopt different conformations. To cleave *in cis* it is thought that the HCV NS3 protease domain swings inward and the overall protein adopts a “closed” conformation, To cleave *in trans*, the NS3 protease domain must swing out away from the helicase domain to reach the additional downstream cleavage sites in the HCV polypeptide (Figure 4-4). Macrocyclic protease inhibitors have been proposed to alter this equilibrium by making contacts in the helicase domain in addition to the protease domain and this characteristic is one of the reasons why the macrocyclic protease inhibitors are much better direct acting anti-virals (inhibit more genotypes and are more active against telaprevir resistant variants) (Schiering et al., 2011).

Modeling and protease cleavage assays with HPI suggest that HPI may also be able to alter the equilibrium between the closed and open conformation by binding in the allosteric site between the helicase and protease domains (Figure 4-2). The M485 residue is located in the allosteric groove between the protease and helicase domains. When it was mutated to alanine, the resulting protein was much more sensitive to HPI implying that the methionine is somehow reducing the interaction of HPI with NS3 and that HPI is indeed binding in this allosteric site. In addition, scNS3 showed much greater sensitivity to HPI in the peptide cleavage assay than either full length NS3 or NS3 with exogenous 4A peptide added. All scNS3 proteins

crystallized have been in a compact conformation (Saalau-Bethell et al., 2012; Yao et al., 1999). In our lab, it has been seen that scNS3 frequently runs faster on a size exclusion column than the smaller protein NS3 and therefore it's possible that scNS3 is favoring the closed conformation and that full length NS3 favors a more open conformation thus slowing its migration. Given that HPI was significantly more potent against the scNS3, it is possible that HPI binds the NS3 protein when it's in the closed conformation and is able to lock it in the close conformation. This hypothesis is modeled in Figure 4-5.

Lastly, the combination of HPI with macro-cyclic protease inhibitors danoprevir and particularly grazoprevir showed a synergistic interaction over and above what would be expected from additive inhibition alone. Since the macro-cyclic protease inhibitors have been shown to make contacts in both the protease domain and the allosteric domain, it's possible

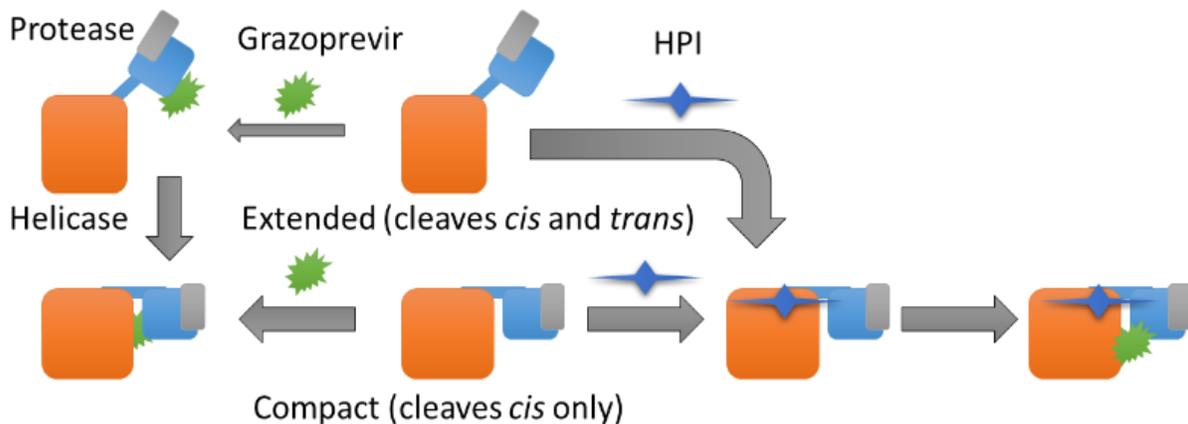


Figure 4-5 Proposed hypothesis for the mechanism of action of HPI. HPI is proposed to bind HCV NS3 in the closed conformation only and to make contacts in the allosteric groove separating the helicase and protease domain effectively shifting the equilibrium towards the closed conformation and limiting the ability of the NS3 protease to cleave downstream substrates. Once HPI has bound to the closed conformation, it may facilitate the binding of the macrocyclic protease inhibitors (such as Grazoprevir) which make contacts in the both the helicase and protease domains, further locking the protein in a closed conformation.

that this synergistic interaction is due to HPI locking NS3 into a closed conformation and

allowing the protease inhibitors better access to both the protease and helicase domains in which to make contacts (Figure 4-4). (Ndjomou et al., 2015) This study is the first time that the two different functions of NS3 were targeted with different compounds simultaneously.

Chapter 5 Interactions between NS3 helicase and nucleic acids

Measuring the interaction of the NS3 helicase with oligonucleotides can be accomplished by monitoring the decrease of the intrinsic protein fluorescence upon oligonucleotide binding because the binding of DNA/RNA to the NS3 helicase causes a concentration dependent decrease in the intrinsic protein fluorescence (340 nm) when the protein is excited at 280 nm. To show this effect, 150 nM YFP labeled NS3 helicase (YFP-NS3h) was titrated with a hexachlorofluoroscein labeled 18-nucleotide DNA strand (HF18). At each point in the titration, the mixture was excited at 280 nm and the fluorescent spectrum collected from 300 nm to 700 nm. The 340 nm fluorescence decreases with each subsequent HF18

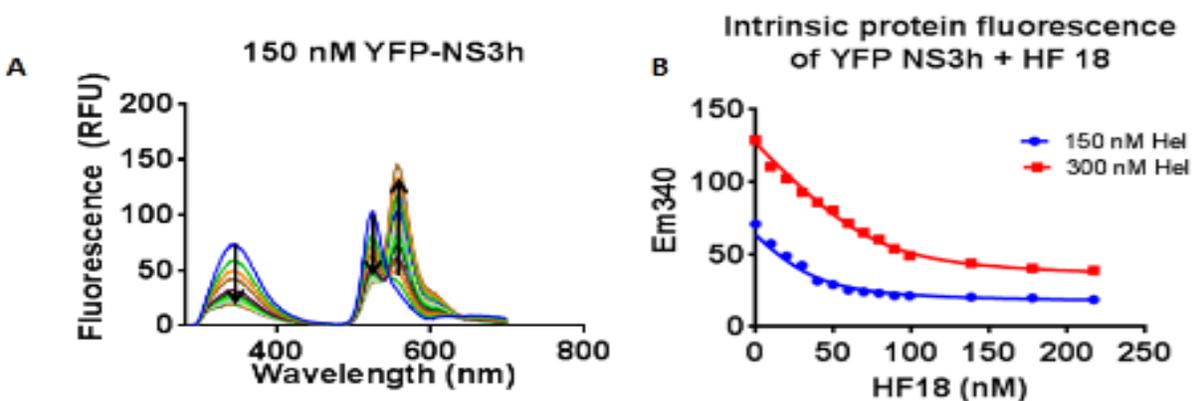


Figure 5-1: NS3 helicase binding monitored through the quenching of the intrinsic protein fluorescence. HCV NS3 binding can be measured through monitoring the quenching of the 340 nm intrinsic protein fluorescence. (A) 150 nM YFP tagged NS3h was titrated with an 18 base nucleic acid tagged with a hexachlorofluoroscein fluorophore (HF-18). The mixtures were excited at 280 nm and the emission spectra recorded. The decrease of the 340 nm fluorescence, along with the decrease of the YFP donor and the increase in concentration of the acceptor HF-18 are noted with arrows. (B) Fluorescence at 340 nm was plotted versus total DNA concentration. Data were fit to equation 3-7 using a K_d of 36 with a 95% confidence interval of 8.7 to 64 and an “n” of 3.9 with a 95% confidence interval of 3.2 to 4.6 and i^D of 0.43 with a 95% confidence interval of 0.41 to 0.43, i^{DA} of 0.11 with a 95% confidence interval of 0.09 to 0.13 and an i^A of 0.

addition (Figure 5-1A). The 280 nm excitation not only excites the intrinsic protein fluorescence but also the YFP and the HF and therefore, it is also possible to see the quenching of the YFP (donor) at 525 nm and the increase of the HF (acceptor) at 556 nm. The decrease of the 340 nm fluorescence (Figure 5-1B) which is caused by the intrinsic protein fluorescence of the protein when excited at 280 nm was fit to equation 3-7:

$$i_{obs} = i^A \left[A - \frac{([D]_T + n[A]_T + K_d) - \sqrt{([D]_T + n[A]_T + K_d)^2 - 4*[D]_T*n[A]_T}}{2} \right] + i^D \left[D - \frac{([D]_T + n[A]_T + K_d) - \sqrt{([D]_T + n[A]_T + K_d)^2 - 4*[D]_T*n[A]_T}}{2} \right] + i^{AD} \left[\frac{([D]_T + n[A]_T + K_d) - \sqrt{([D]_T + n[A]_T + K_d)^2 - 4*[D]_T*n[A]_T}}{2} \right] \quad (\text{Equation 3-7})$$

The observed fluorescence intensity at 340 nm is equal to a linear combination of the free acceptor [A] which is YFP-NS3h, the free donor [D] which is an 18 nucleotide DNA labeled with an HF, and the donor acceptor complex [AD]. Each of these terms is multiplied by a co-efficient (i^A , i^D , or i^{AD}) representing the fluorescence intensity of each observed at 340 nm which should be linearly related to the amount of acceptor, donor, or acceptor donor complex present. The co-efficient i^A represents the fluorescence of the acceptor (HF-DNA) at 340 nm, the co-efficient i^D represents the fluorescence of the donor (YFP-helicase) at 340 nm and the co-efficient i^{AD} represents the fluorescence of the donor: acceptor complex at 340 nm.

In this experimental set up where fluorescence intensity at 340 nm is being measured, it was assumed that i^A is zero because there is no observed fluorescence of the acceptor (HF-DNA) at the 340 nm emission wavelength. For the titration of YFP-NS3h with HF-DNA, the best-

fit K_d value was 36.4 ± 30 nM (8.7 nM to 64.1 at the 95% confidence interval) and the calculated “n” is 3.9 ± 0.7 (3.2 to 4.5 nM at the 95% confidence interval).

5.1 Measuring the interaction of HCV NS3 with oligonucleotides using FRET: A comparison of three different FRET calculation techniques

Another way that HCV NS3 helicase interactions with oligonucleotides can be measured is through fluorescently labeling the helicase and the oligonucleotide with a complimentary FRET pair and measuring FRET between the donor and acceptor fluorophore. A titration of YFP labeled NS3h with a “MAX” fluorophore labeled single stranded DNA (MAX-DNA) was performed. (Figure 5-2A). Elementary spectra of the YFP-NS3helicase and the MAX labeled DNA were collected at 480 nm and 525 nm excitations. Before the addition of any MAX-DNA, 500 μ l of 50 nM of YFP NS3 helicase was excited at 480 nm to generate the emission spectrum for donor alone. Next 500 μ M MAX-DNA was added in 1 μ l increments to a final concentration of 54 nM MAX-DNA. For each point in the titration, emission spectra were collected for both the 480 nm and 525 nm excitations using a slit width of 10 nm. (Figure 5-2B and 5-2C).

To calculate the FRET between YFP-NS3h and the MAX-DNA, three different FRET calculation techniques were employed. The first technique calculated F_{FRET} used the previously defined equation 2-4:

$$F_{FRET} = 1 - \frac{I^D(\lambda_D^{ex} \lambda_D^{em})}{I_{Alone}^D(\lambda_D^{ex} \lambda_D^{em})} \quad (\text{Equation 2-4})$$

To calculate F_{FRET} using equation 2-4, the fluorescence of the titration before any addition of acceptor was measured to obtain the $I_{\text{Alone}}^D(\lambda_D^{\text{ex}} \lambda_D^{\text{em}})$ value. The emission of the donor in the presence of acceptor ($I^D(\lambda_D^{\text{ex}} \lambda_D^{\text{em}})$) was measured at each point in the titration. F_{FRET} was calculated for each point in the titration using the values measured.

A second technique was employed which uses spectral unmixing to separate the contributions of the donor and acceptor spectra from the composite spectra (see section 2.5.6 and section 3.3.4 for further explanation of the technique). Through the use of spectral unmixing, the k^D coefficient was determined from the starting sample which contained only donor (i.e. the starting point in the titration). Then the k^D was multiplied by the spectral integral of the donor alone spectra to obtain an F^D value which represents all the donor photons across the full donor spectrum. This F^D was obtained from the spectrum of the donor alone (i.e. the first point in the titration before any addition of titrant). After each subsequent addition of acceptor, mixtures were allowed to equilibrate, and an F^{DA} value was calculated by using the k^{DA} coefficient generated from unmixing and multiplying it by the spectral integral of the donor only spectrum to obtain all the donor photons across the full donor emission spectra. The constant F^D value and each F^{DA} value was then used in equation 2-2 to calculate an E_{app} value for each point in the titration.

$$E_{\text{app}} = 1 - \frac{F^{\text{DA}}}{F^D} \quad (\text{Equation 2-2})$$

It is important to note that direct measurement of F^D is not possible to obtain in some fluorescence microscopy experiments, such as cells transfected with a plasmid expressing a fluorescent protein. To simulate what would needed to be done in a microscope, the third

FRET calculation technique applied to the titration data back-calculated F^D values from F^{AD} and F^{DA} values obtained from spectral unmixing using gamma ratios and the quantum yields. In this experiment, the gamma ratio of the donor was calculated to be 0.0002 and the gamma ratio of the acceptor was calculated to be 0.0550. The quantum yield of the donor (YFP) is 0.61 and the quantum yield of the acceptor (MAX) is 0.2. When equation 3-19 is substituted into 3-10 the following equation 5-1 results:

$$F^D(\lambda_{ex,1}) = F^{DA}(\lambda_{ex,1}) + \frac{Q^D}{Q^A} F(\lambda_{ex,1}, \lambda_{ex,2}) - \rho^A \frac{Q^D}{Q^A} \cdot F^A(\lambda_{ex,2}) \text{ (Equation 5-1)}$$

where $F^A(\lambda_{ex,2})$ is defined by equation 3-14.

E_{app} was then calculated using the same F^{DA} values calculated in technique two and the back-calculated F^D values using equation 2-2.

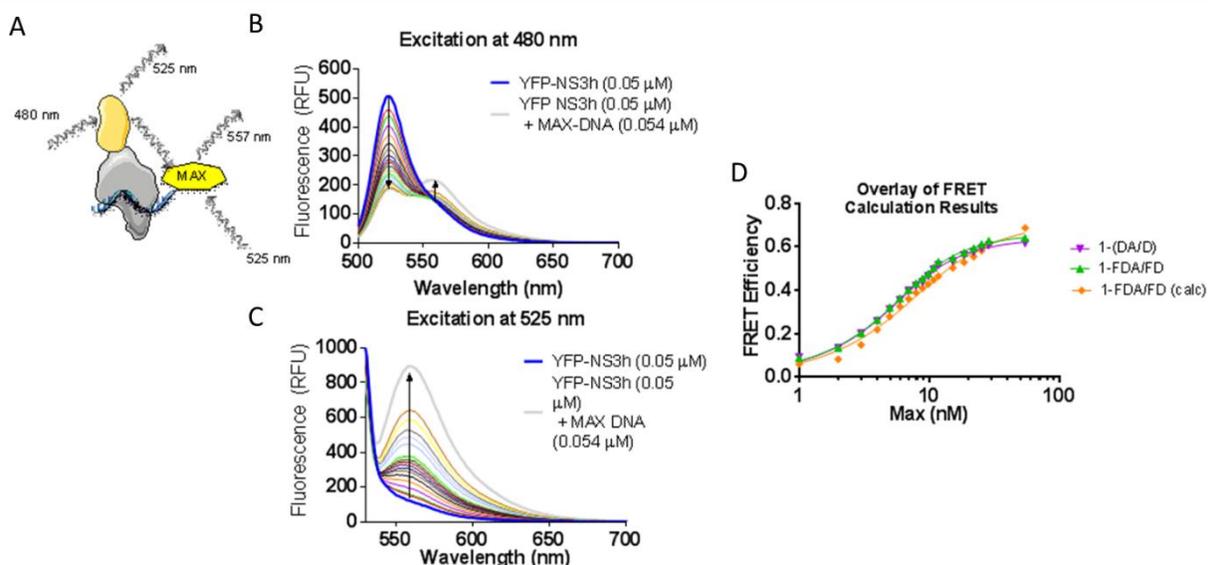


Figure 5-2: Titration of a YFP labeled NS3 helicase with a MAX fluorophore labeled single-stranded DNA. (A) Schematic showing the proposed interaction of YFP NS3helicase with MAX labeled DNA. (B) The 480 nm excitation targets the YFP which is expected to have a peak emission of about 525 nm. Emission spectra for each addition of MAX labeled DNA were collected for the 480 nm excitation up to a final MAX-DNA concentration of 54 nM. (C) The second excitation of 525 nm targets the MAX fluorophore, which is expected to emit at 557 nm. Emission spectra for each addition of MAX labeled DNA were collected for the 525 nm excitation up to a final MAX-DNA concentration of 54 nM. (D) Overlay of FRET calculated for each point in the titration using three different FRET calculation techniques.

All three techniques were used to calculate FRET values for each point in the titration. A graph overlaying the results of the titrations (Figure 5-2D). The curves were fit to equation 3-9. The K_d , E^{AD} , and E_T yielding the best fits are reported in Table 5-1.

All three different FRET calculation techniques employed resulted in similar curve fits.

	Kd	± 95% confidence interval	D_T	± 95% confidence interval	E^{DA}	± 95% confidence interval
$1-(DA/D)$	1.82	1.3 to 2.3	6.7	5.9 to 7.6	0.64	0.63 to 0.66
$1-(F^{DA}/F^D)$	1.9	1.4 to 2.4	6.8	6.1 to 7.6	0.67	0.65 to 0.69
$1-(F^{DA}/F^{Dcalc})$	4.27	1.6 to 6.9	6.4	2.9 to 9.9	0.72	0.66 to 0.78

Table 5-1: Results of FRET calculation using each FRET calculation technique for the YFP-NS3h to MAX-DNA titration.

The first two techniques require that the starting donor concentration be known. This is possible with an *in vitro* titration where the starting point in the titration is known and represents the donor in the absence of any acceptor. However, in cell microscopy where fluorescent proteins are expressed, it is not possible to directly measure the concentration of donor molecules in each cell measured. Therefore, the concentration of the donor needs to be back-calculated. The equation which back-calculates the fluorescence of the donor in the absence of acceptor under circumstances where it cannot be measured directly is equation 3-10. Equation 3-10 was used to back-calculate the fluorescence of donor in the YFP-NS3helicase + MAX-DNA titration in the absence of acceptor (F^D) for each point in the titration. This is the same procedure used to back calculate F^D in FRET spectrometry (discussed in Section 3-5) and applied to live cell fluorescence microscopy. The results are similar to those determined without back-calculating the donor alone fluorescence.

Chapter 6 The N-terminal membrane anchor domain of NS4A is needed to localize the hepatitis C virus NS3 protease to biologically active foci

The hepatitis C virus (HCV) encodes a multifunctional NS3 protein which contains a covalently linked protease and helicase domain. Important to NS3 proteolytic function is the peptide cofactor, NS4A, which is only 54 amino acids in length, yet acts as a key player in HCV replication. One way NS4A aids in viral replication is by tethering NS3 to the endoplasmic reticulum and mitochondrial membranes. NS4A consists of 3 domains; an acidic C-terminal domain, an NS3 cofactor domain, and an N-terminal membrane anchor. Prior research suggests that the anchor domain is required to tether NS3 and form HCV replicase complexes (Andrew Kohlway et al., 2014).

To test the localization of NS3 in cells in the presence and absence of NS4A, four eukaryotic expression plasmids were generated that encoded recombinant proteins in which all or part of NS3 and NS4A were fused to a mCFP fluorescent protein. In the first, only the NS3 helicase domain (mCFP-NS3h) was fused to mCFP. In the second, mCFP was fused with full-length NS3 (mCFP-NS3). In the third, mCFP was fused to NS3 and only the NS3 cofactor domain of NS4A (lacking the membrane anchor and C-terminus, i.e. mCFP-scNS3), and in the fourth, mCFP was fused to NS3 and full length NS4A (mCFP-NS3:NS4A). The aligned protein sequences for the NS3 variants are shown in Figure 6-1. If the NS4A membrane anchor domain is required for the localization of the NS3 protein to foci, then removal of this domain should cause the protein to be diffuse throughout the cell. If the anchor domain is not required for the

localization of NS3 to foci, then removal of the domain should not affect the sub-cellular localization.

To test the activity of the recombinant NS3, scNS3, or NS3:NS4A variants, each variant was co-expressed in HEK293T cells with an NS3 protease substrate tagged with a red fluorescent protein (RFP) connected to a nuclear localization signal (RFP-NLS-MAVS). When the protease substrate is cleaved, the RFP-tagged nuclear localization signal is exposed causing the RFP to re-localize to the nucleus. If the protease substrate is not cleaved, then the nuclear localization signal is not exposed and the RFP remains in the cytoplasm. This system allows for the activity of the NS3 protease to be measured in live cells. If the membrane anchor domain is necessary for the activity of NS3, then removal of this domain should change the biological activity of the NS3 protein. If the membrane anchor domain is not necessary for the activity of NS3, then removal of the domain should not change the biological activity of the protein.



Figure 6-1: Alignment of cyan fluorescent protein tagged NS3 recombinant protein sequences. The following protein sequences were taken from eukaryotic expression vectors expressing variants of CFP tagged HCV NS3. The alignment includes CFP-NS3h, CFP-NS3, CFP-scNS3, and CFP-NS3:NS4A. For a more detailed description of each protein expression vector, see Chapter 3: Methods.

6.1 The NS4A membrane anchor is necessary to localize the NS3 complex to sub-cellular foci

The recombinant CFP tagged NS3 variants were expressed alone in HEK 293T cells for 24 hours after transfection. The cells were excited at 840 nm using a two-photon laser excitation and imaged at 100X magnification. Details of the microscopic image collection can be found in Section 3.5.3. A full 200 wavelength emission spectra was collected for each image but only the image corresponding to the CFP emission peak of 485 nm was used.

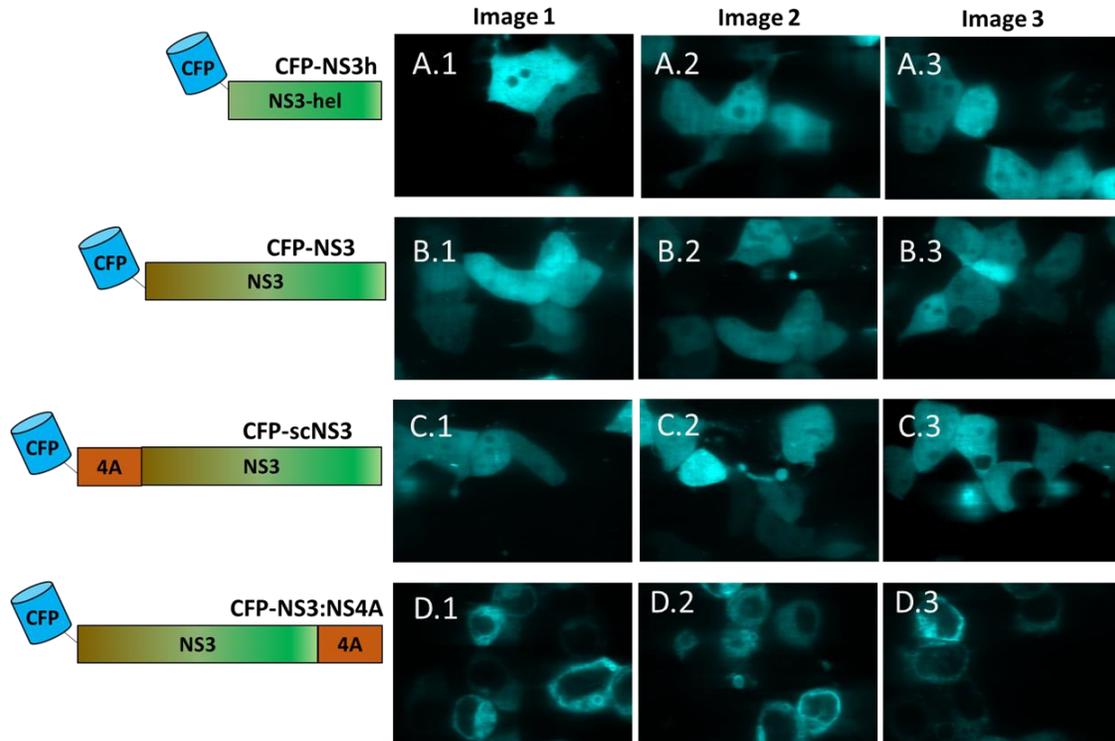


Figure 6-2: CFP-NS3 variants expressed in HEK 293T cells. Recombinant CFP-NS3 variants were expressed in HEK293T cells for 24 hours. Post 24 hours transfection, the cells were imaged by exciting the CFP at 840 nm with a two-photon laser and a magnification of 100X. CFP-NS3h contains only the helicase domain is diffuse throughout the cell (A.1-A.3). CFP-NS3 contains the protease and helicase domain and is diffuse throughout the cell (B.1-B.3). CFPscNS3 contains the protease domain covalently linked to the NS4A protease co-factor domain in addition to a helicase domain and it is diffuse throughout the cell (C.1-C.3). The CFP-NS3:NS4A contains the complete NS3 + NS4A co-factor including the membrane anchors and it is localized to distinct foci in the cell (D.1-D.3).

CFP-NS3h lacks the protease domain and all of the NS4A co-factor. As expected, the images of CFP-NS3h show it diffuse throughout the cytoplasm of the cell (Figure 6-2 A.1-A.3). CFP-NS3 contains both the helicase and the protease domain but lacks any portion of the NS4A co-factor. It was expected that CFP-NS3 would be diffuse throughout the cell as seen in Figure 6-2 B.1-B.3. CFP-scNS3 contains the helicase, protease, and a portion of NS4A but lacks the NS4A membrane anchor and, similar to the other two constructs, CFP-scNS3 is diffuse throughout the cell (Figure 6-2 C.1-C.3). CFP-NS3:NS4A contains the helicase, protease, and the

entire NS4A domain. This construct forms visible foci in the cytoplasmic region of the cell (Figure 6-2 D.1-D.3). It is likely that these foci are targeted to ER or mitochondrial membranes because that is where the HCV replicase complex is known to form (Andrew Kohlway et al., 2014; Lindenbach and Rice, 2005).

6.2 The NS3 protease co-factor domain of NS4A is necessary for proteolytic cleavage of an NS3 substrate but the portion encoding the membrane anchor is not necessary for protease activity.

Each of the recombinant CFP-NS3 constructs were co-expressed in HEK293T cells with a plasmid expressing the RFP-NLS-MAVS protease substrate. The constructs were allowed to express for 24 hours after transfection. Then the cells were excited at 840 nm to target CFP and 980 nm to target RFP using a two-photon laser excitation and imaged at 100X magnification. A full 200 wavelength emission spectra was collected for each image but only the peak emission for CFP (485 nm) and RFP (584 nm) and was used for the image analysis.

CFP-NS3-helicase is lacking the protease domain and the entire 4A co-factor. Therefore, it is not expected to localize in the cell or to be able to cleave the RFP-NLS-MAVs protease substrate. Images show that the RFP tagged substrate appears to be in small groups throughout the cytoplasm and CFP-NS3h is diffuse in the cytoplasm (Figure 6-3 - Row B). The CFP-NS3 contains both the helicase and protease domain but lacks NS4A. The NS3 lacking NS4A can cleave protease substrates *in vitro* as seen in Figure 4-2B. But in cells, CFP-NS3 is diffuse throughout the cell and appears unable to cleave the RFP-NLS-MAVs substrate which remains in the cytoplasm (Figure 6-3 Row C). The CFP-scNS3 contains the helicase and protease domain

along with the protease interacting portion of NS4A. This protein is diffuse throughout the cell but is proteolytically active. In the cells co-expressing CFP-scNS3 and RFP-NLS-MAVS, the RFP is localized to the nucleus (Figure 6-3-Row D). Lastly, the CFP-NS3:NS4A contains the helicase and protease domain and the complete NS4A protein including the membrane anchors. This protein is located in sub-cellular foci and the RFP-NLS-MAVS substrate can be seen localized to the nuclear region of the cell. When CFP-NS3:NS4A was co-expressed with the cleavage resistant mutant (MAVs C508Y), the RFP was located in the cytoplasm. (Figure 6-3 Row E and F).

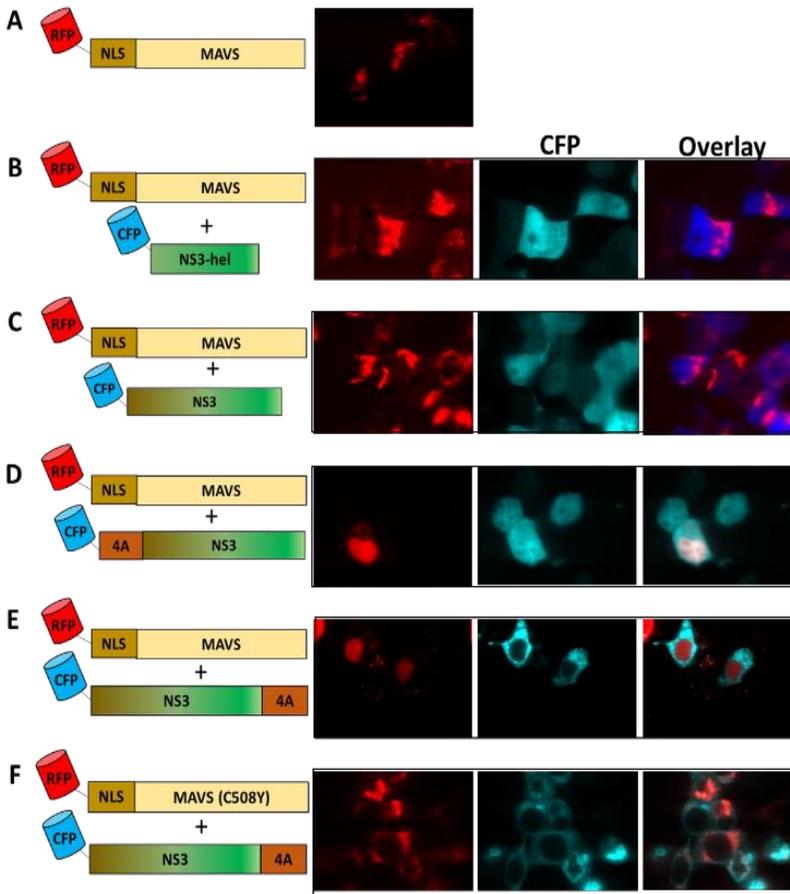


Figure 6-3 : CFP-NS3 encodes a functional protease when expressed in HEK 293T cells.

RFP-NLS-MAVs contains an RFP attached to a nuclear localization signal (NLS) and the NS3 cleavage site of MAVs. When NS3 cleaves MAVs, the NLS is exposed and the RFP molecule moves to the nucleus of the cell. When RFP-NLS-MAVs is expressed alone or in conjunction with CFP-NS3 helicase or CFP-NS3 lacking the 4A co-factor, it is in the cytoplasm (Rows A-C). When RFP-NLS-MAVs is expressed with CFP-scNS3 or CFP-NS3:NS4A which both contain the protease and the 4A co-factor, the RFP localizes to the nucleus (Row D and E). When CFP-NS3:NS4A is expressed in conjunction with a cleavage resistant RFP-NLS-MAVs (C508Y) the substrate is not cleaved and the RFP remains in the cytoplasm of the cells (Row F).

6.3 Significance of the requirement of the NS4A co-factor for proteolytic cleavage in cells and the necessity of the NS4A anchor domain to localize NS3

Previous data has shown that the NS4A anchor domain is necessary to localize the NS4A to replicase complexes in the cell (Andrew Kohlway et al., 2014). The results reported here confirm this finding. Only the NS3 which contained the full NS4A, including membrane anchors, localized to sub-cellular foci in the cell. This means that there are likely no other membrane anchors in NS3 or NS4A which could localize the protein. It also means that this sub-cellular localization is not a requirement for proteolytic cleavage (Figure 6-3 Rows D and E). Given that NS3 protease has multiple sub-cellular targets (NS3 poly-peptide, TRIF, and MAVS) it makes sense that localization to the replicase complexes is not a requirement of protease function.

CFP-NS3 contains the protease domain and *in vitro* experiments with a similar construct lacking NS4A still show protease activity. However, when the NS4A domain is absent in the constructs expressed in cells, the protease is unable to cleave the MAVS substrate. It could be that the NS3 protease lacking NS4A is not a very active protease and the lower concentrations of NS3 protease and substrate in cells do not allow for efficient cleavage. It could also mean that the NS4A molecule somehow affects the local pH environment which is not an issue in a controlled *in vitro* experiment where pH is constant throughout, but is an issue in the sub-cellular environment where local pH values may be different. It's possible that this shift in local pH could affect the ability of the protease to cleave efficiently in cells in the absence of NS4A.

Chapter 7 The Hepatitis C virus helicase interacts with the human pattern recognition receptor LGP2

Hepatitis C virus encodes a multi-functional non-structural protein, NS3, which contains a covalently linked protease and helicase domain. A linked protease and helicase is unique to *flaviviridae* viruses and the purpose behind linking the two activities is unclear. There is some evidence that the activity of the protease is enhanced when the helicase domain is present and the activity of the helicase domain is enhanced when the protease domain is present (Aydin et al., 2013; Beran and Pyle, 2008; Frick et al., 2004). However, this cross-domain activity enhancement does not necessarily explain why their activities are linked.

NS3 plays multiple roles in the HCV viral life cycle. One of the most important functions of NS3 protease is to disrupt innate immune signaling in response to HCV infection (Errett and Gale, 2015b; Horner et al., 2012; Liu and Gale, 2010). The innate immune system is an essential first line of defense against invading organisms. Almost every cell in the body is equipped with a set of innate immune system receptor proteins that recognize and respond to viral nucleic acids. These receptor proteins, including the RIG-I like receptors and the Toll-like receptors, signal through key adaptor proteins which initiate down-stream signaling and the production of interferon. The NS3 protease cleaves the adaptor proteins mitochondrial anti-viral signaling protein (MAVS) and TIR-domain containing adaptor inducing interferon- β (TRIF). This cleavage effectively abrogates RIG-I like receptor signaling and Toll-like receptor signaling in the cell (Akira et al., 2006; Alberts et al., 2002; Horner et al., 2012).

It is possible that NS3 helicase aids in the cleavage of the innate immune adaptor MAVS by helping to localize the NS3 protease domain to MAVS through interaction with one of the RIG-I like receptors. Since all the RIG-I like receptors and HCV NS3 are DExD/H box helicases who are known to oligomerize into higher order complexes and they have the same ligand preference (HCV RNA), it's possible that they could oligomerize with each other, either through a direct protein-protein interaction or through a shared interaction on RNA. This interaction would aid in localizing NS3 near MAVS for cleavage and provide one possible explanation for linking the protease and helicase functions in HCV.

To test for possible interactions of NS3 with the RIG-I like receptors, pairs of fluorescent fusion proteins were designed to be used in live-cell FRET imaging assays. The assays were performed with Cyan fluorescent protein (CFP) as the donor and yellow fluorescent protein (YFP) as the acceptor. These fluorescent proteins were attached to each of the RIG-I like receptor proteins and to domain deletion variants of the HCV NS3 protein. Each recombinant protein was designed to maintain its enzymatic activity and sub-cellular localization.

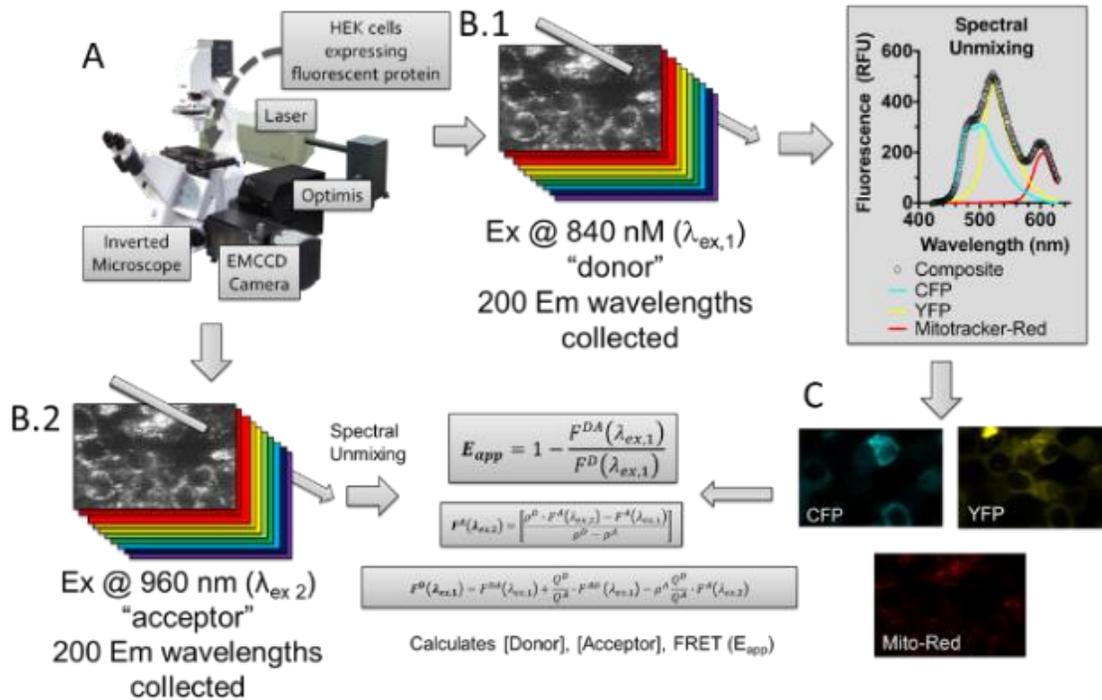


Figure 7-1 Quantitative Micro-spectroscopic imaging of donor and acceptor labeled recombinant fluorescent fusion proteins in live cells. Cells expressing fluorescent fusion proteins were imaged on an inverted microscope and excited using a two-photon laser. Full 200 wavelength spectrums were collected on an EMCCD camera (A). Each plate of cells rendered multiple field of views and each field of view was excited at two wavelengths, one to target the donor (B.1) and one to target the acceptor (B.2). The resulting emission spectra were unmixed to render the contributions of each elementary component to the overall composite spectra (C). Regions of interest were selected and the intensity maps of the individual elementary component intensities (C) were used to compute average FRET and average donor and acceptors concentrations over the regions of interest.

Details of the live cell FRET imaging assay used can be found in section 3.5. Briefly, cells expressing the recombinant fluorescent fusion proteins were imaged on an inverted microscope and excited using a two-photon laser. A full 200-wavelength emission spectra was collected on an EMCCD camera (Figure 7-1A). Two excitation wavelengths were used for each field of view. The first excitation wavelength targets the donor and generates the “FRET” image. The second targets the acceptor and generates the “acceptor” image (Figure 7-1 B.1 and B.2)

The emission spectra were unmixed using a method of least squares to produce the donor and acceptor contributions to the composite spectrum. On some of the images, the cells were also stained with Mito-Tracker-Red. In this case, a third component (Mito-Red) was un-mixed from the composite spectra (Figure 7-1C). These images were then used to compute the average E_{app} (FRET efficiency) for the selected sub-cellular regions of interest. In addition, average donor and acceptor concentrations were calculated for each region of interest.

7.1 Recombinant fluorescent RNA helicases are biologically active when expressed in HEK293T cells

Plasmids were generated which encoded the following fluorescent fusion proteins; CFP-NS3h, CFP-NS3, CFP-scNS3, CFP-NS3:NS4A, YFP-RIG-I, YFP-MDA5, and YFP-LGP2. (Figure 6-1 and Figure 7-2). All of the fluorescent fusion proteins were designed to be expressed under the

control of a cytomegalovirus promoter (CMV) when transfected into HEK 293T cells. Each fluorescent protein tag was attached to the N-terminus of the protein of interest.

HCV NS3:NS4A contains the NS3 protein which has the covalently linked protease and helicase domain along with the NS4A protein. The NS4A protein is a protease co-factor in addition to containing membrane anchors necessary for proper sub-cellular localization to the ER membrane. When the CFP-NS3:NS4A plasmid was expressed in HEK293T cells, the CFP-NS3:NS4A was localized to foci, indicating that the NS4A membrane anchors were functional and securing the protein in specific sub-cellular locations (Figure 6-2 D.1-D.3).

To further assess the function of the fluorescently tagged NS3 protein, the proteins were expressed in conjunction with an RFP tagged protease substrate containing a nuclear localization signal (RFP-NLS-MAVS). Prior to cleavage, the RFP-NLS-MAVS is present in the cytoplasm. When the substrate is cleaved by the functional NS3 protease, the nuclear localization signal is exposed and the RFP re-localizes to the nucleus. The RFP-NLS-MAVS was co-transfected along with the NS3 variants and the location of the RFP molecule was assessed in each case. CFP-NS3:NS4A and CFP-scNS3 both contain the protease domain and the protease co-factor NS4A and both of them were able to cleave the RFP-NLS-MAVS substrate evidenced by the re-localization of the RFP protein to the nucleus of the cell. To confirm that NS3 specifically was cleaving the MAVS substrate, a RFP-NLS-MAVS substrate bearing an NS3 cleavage resistant mutation was also co-transfected with the CFP tagged NS3 proteins (RFP-NLS-MAVS (C508Y)). The cleavage resistant mutant was unable to be cleaved and remained in

the cytoplasm, showing that the cleavage of the MAVS substrate is specific to a functional NS3 protease cleavage site (Figure 6-3 A-E and Figure 7-3 A-B).

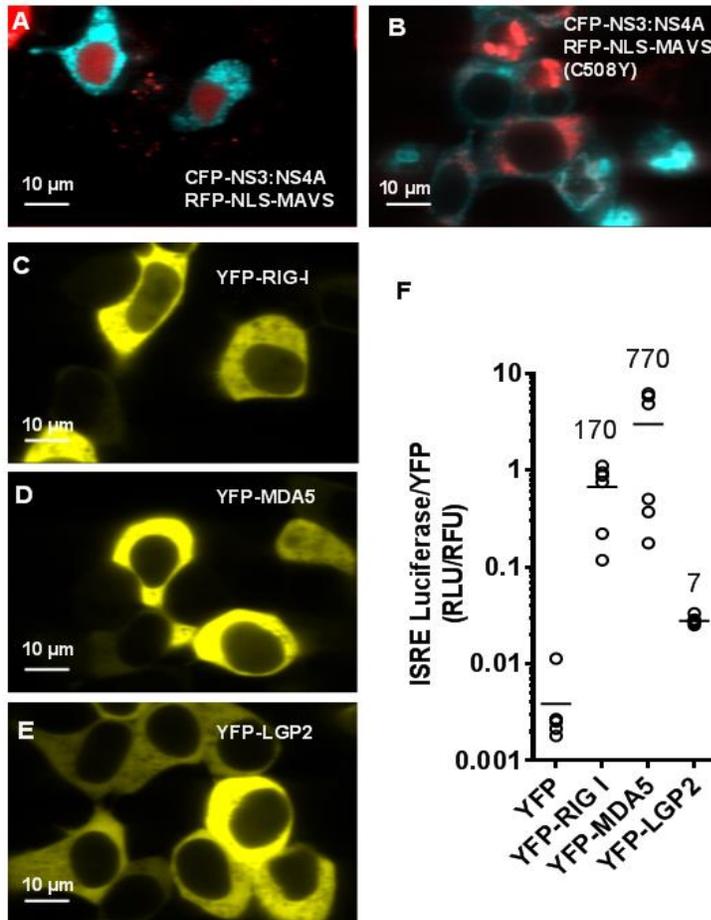


Figure 7-3: Recombinant fluorescent fusion proteins retain biological activity when expressed in HEK293T cells. A plasmid expressing CFP-NS3:NS4A was co-expressed with RFP-NLS-MAVS (A) or RFP-NLS-MAVS expressing a NS3 cleavage site mutation (C508Y) (B) in HEK 293T cells. After 24 hours the cells were imaged at 100X with two excitation wavelengths of 840 nm to target the CFP and 1000 to target the RFP (A and B). Plasmids expressing the YFP tagged RLRs were expressed in HEK 293T cells and imaged after 24 hours (C-E). Plasmids expressing YFP tagged RLRs were co-expressed with a plasmid expressing the ISRE promoter with a luciferase reporter. After 24 hours, the luciferase activity was measured along with the YFP fluorescence in each well. Luciferase activity was normalized to YFP fluorescence.

The RIG-I like receptors, RIG-I and MDA5, exert their signaling capability through their C-terminal tandem CARD domains. To avoid these crucial interacting domains, the RIG-I like receptors were all tagged on the N-terminus. Previous research has shown that RIG-I like receptors tagged on the N-terminus retain their ability to signal and interact with ligand (Rehwinkel et al., 2010; Sanchez David et al., 2016). However, tagging with fluorescent proteins can sometimes cause proteins to enter the nucleus that may not otherwise be able to enter the

nucleus. Un-tagged RIG-I like receptors are known to function in the cytoplasm and to confirm that the fluorescently tagged RIG-I like receptors localize to the cytoplasm each fluorescently tagged receptor was expressed alone in HEK 293T cells. The resulting YFP tagged proteins were all localized to the cytoplasm of the cells (Figure 7-3 C-E). To assay the signaling capability of the YFP tagged RIG-I like receptors, YFP alone, YFP-RIG-I, YFP-MDA5, and YFP-LGP2 were all co-expressed with a second plasmid encoding firefly luciferase under the control of a promoter regulated by an interferon-stimulated response element (ISRE). When YFP alone was co-expressed with the ISRE plasmid, very little luciferase activity was measured. However, when YFP-RIG-I and YFP-MDA5 were co-expressed with the ISRE plasmid, both showed the ability to stimulate the ISRE promoter. YFP-LGP2 also showed the ability to stimulate the ISRE promoter to a greater degree than the YFP vector alone. However, because LGP2 lacks the CARDs necessary for direct down-stream signaling, it is likely that this increase is due to the over-expressed YFP-LGP2 interacting and stimulating the endogenous RIG-I and MDA5 proteins in the HEK293T cells (Figure 7-3 F).

7.2 FRET analysis of interaction between HCV NS3 helicase domain and RLRs

To screen for and analyze the interaction of the RIG-I like receptors with HCV NS3, a fully quantitative FRET technique was employed. This FRET method analyzes the interactions of protein ensembles in live cells and allows for the calculation of donor and acceptor concentrations (for more details on the FRET technique, see methods section 3.5) (Patowary et al., 2015; Stoneman et al., 2017). To specifically target sub-cellular regions of interest (in this study, HCV NS3:NS4A foci and mitochondria), a variation on the fully quantitative FRET

technique was developed which allowed for the selection of sub-cellular regions of interest using one of the fluorescent images or the addition of a third fluorescent tag. This variation on fully-quantitative spectroscopic imaging (FSI) (King et al., 2016; Stoneman et al., 2017) was termed quantitative micro-spectroscopic imaging (Q-MSI) (Corby et al., 2017).

Cells co-transfected with CFP tagged NS3 variants and YFP tagged RLRs were each imaged at two-excitation wavelengths (in this case 840 nm and 960 nm). The elementary spectrum of the CFP (Figure 7-4 A) and the YFP (Figure 7-4 B) were used to de-convolute (un-mix) the composite CFP:YFP spectrum (Figure 7-4 C and D) into the individual CFP and YFP contributions (Figure 7-4 E,F, and G). The second excitation is necessary to calculate donor and acceptor concentrations (Figure 7-5 A-C) and to correct for the direct excitation of YFP at the CFP excitation wavelength of 840 nm. Each excitation scan collects about 200 images with a 1 nm resolution and requires about 10 seconds. To change the excitation wavelength between excitation scans takes about 40 seconds. So the average time to collect two images at two different excitation wavelengths is approximately 60 seconds. This time delay allows for molecular diffusion and limits the analysis of the images to the calculation of FRET averages over specific sub-cellular regions of interest. However, this trade-off is necessary when using CFP:YFP as a FRET pair because YFP is excited at the 840 nm excitation wavelength used to excite CFP. The un-mixed images at each excitation wavelength were used to calculate the average FRET over regions of interest. The average intensities of the un-mixed donor and acceptor images over the same regions of interest were compared to standard curves of CFP and YFP to determine CFP and YFP protein concentrations (Figure 7-5 A-C) (for details see methods section 3.5).

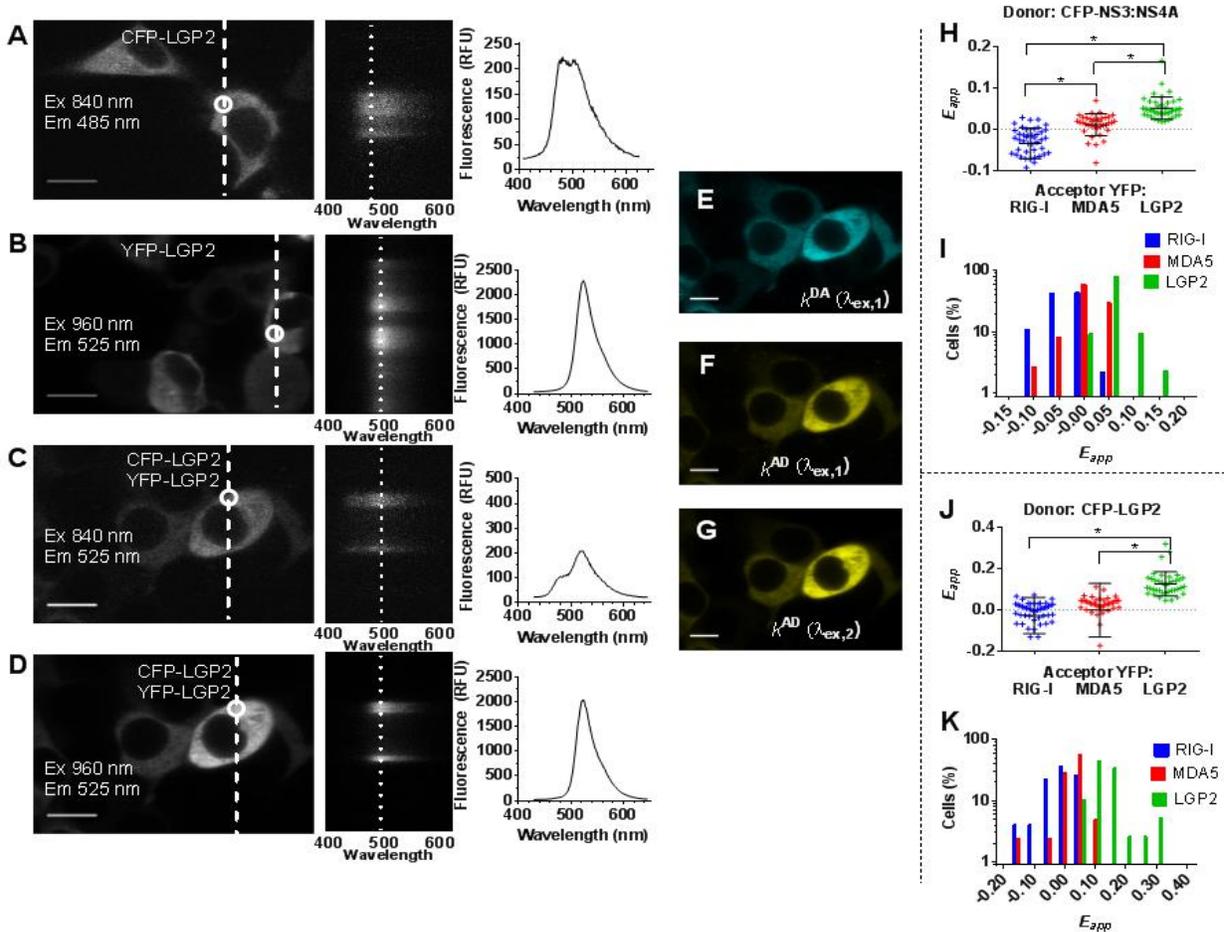
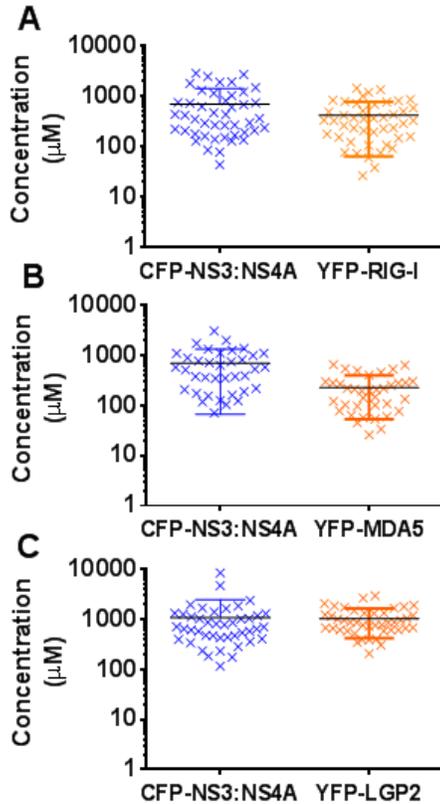


Figure 7-4 Analysis of FRET helicase interactions using whole cell regions of interest. HEK 293T cells were transfected with either CFP-LGP2, YFP-LGP2, or both CFP-LGP2 and YFP-LGP2. After 24 hours of expression, the cells were imaged at two excitation wavelengths (840 nm and 960 nm). The CFP-LGP2 alone cells generated the CFP spectrum (A), the YFP-LGP2 alone cells generated the YFP spectrum (B) and these elementary spectra were then used to deconvolute the composite spectra at 840 nm (C) and at 960 nm (D) excitation. This deconvolution yielded the donor (E) intensity map, the acceptor intensity map (F) and the acceptor concentration intensity map (G). These intensity maps were then used to calculate the average E_{app} for each region of interest. The same technique was applied to the analysis of the interactions of CFP-NS3:NS4A with each of the YFP tagged RIG-I like receptors (H and I). Similarly, the interaction of CFP-LGP2 with each of the YFP tagged RIG-I like receptors was analyzed (J and K).

Donor and Acceptor Concentrations
CFP-NS3:NS4A and YFP-RLRs



Donor and Acceptor Concentrations
CFP-LGP2 and YFP-RLRs

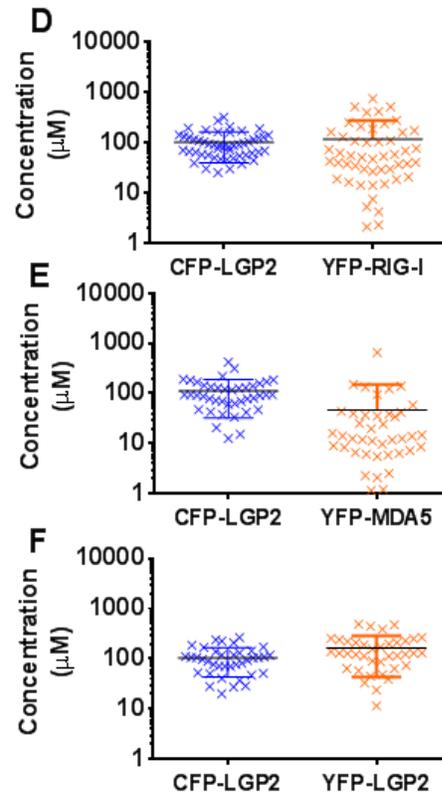


Figure 7-5: Sub-cellular average concentrations over selected regions of interest in the analysis of helicase interactions. The concentrations of the donor and acceptor were calculated for the cells co-expressing CFP-NS3-4A and each of the YFP tagged RIG-I like receptors by using the un-mixed donor and acceptor intensity maps and applying a standard curve either CFP or YFP protein fluorescence respectively. The concentrations of donor and acceptor within any given cell were overall within the same range (A-C). The same technique was used to calculate the donor and acceptor concentrations in the cells co-expressing CFP-LGP2 and each of the YFP tagged RIG-I like receptors. Again the average concentrations over selected regions of interest were roughly equal for each set of cells measured (D-F).

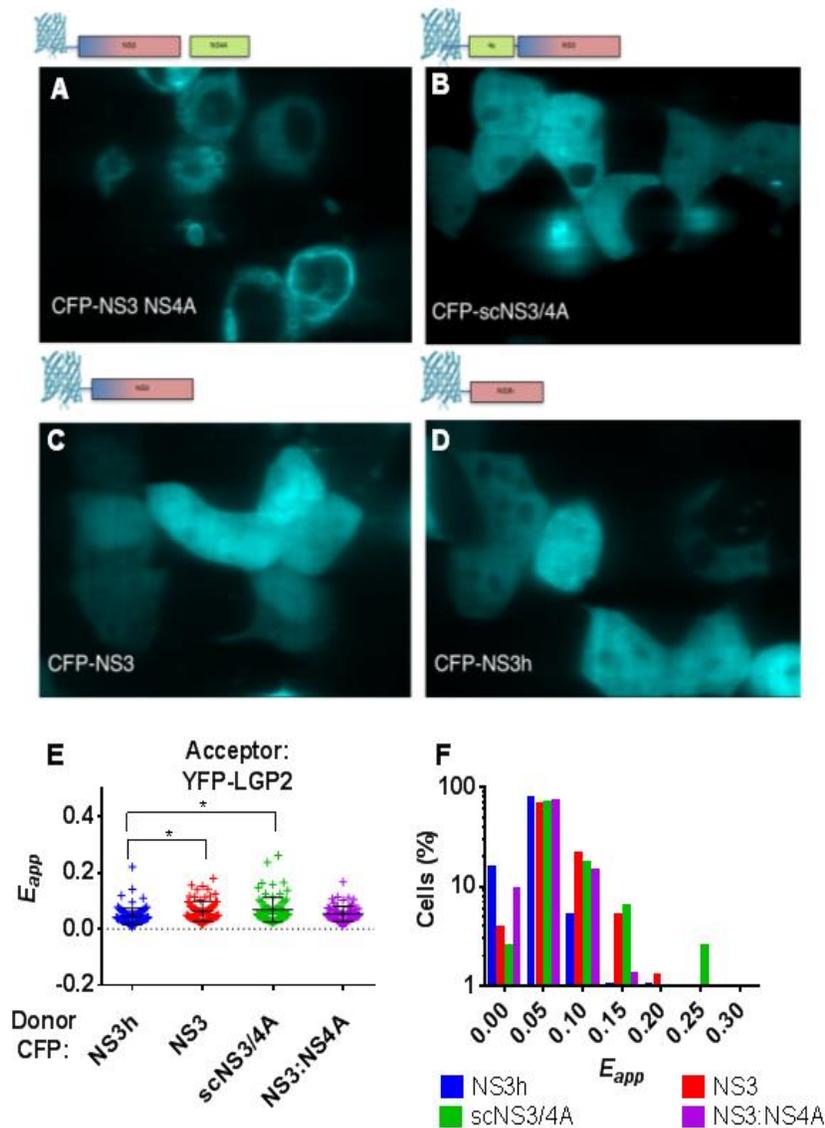


Figure 7-6: HEK 293T cells were transfected with each of the NS3 domain deletion mutants. A selected image of the 485 nm emission from the cells expressing each of the CFP-NS3 variants is shown here (A-D). CFP-NS3:NS4A was localized to foci within the cell (A) whereas each of the other NS3 variants lacking the membrane anchors were diffuse throughout the cell's cytoplasm (B-D). 293T cells were co-transfected with each of the NS3 domain deletion constructs and YFP tagged LGP2. The E_{app} for each combination was determined using Q-MSI and the scatter plot shows the average E_{app} calculated for each cell analyzed. The average $E_{app} \pm$ the S.D. and the number of cells (regions of interest analyzed were: NS3h: 0.045 ± 0.029 (n=96), NS3: 0.063 ± 0.035 (n=78), scNS3: 0.068 ± 0.045 (n=79), and NS3-4A: 0.054 ± 0.027 (n=74). The significance of the difference between groups was calculated using the one-way ANOVA ($F_{3,323}=7.814$, $p<0.0001$). A Tukey post-hoc was used and revealed that the E_{app} of the NS3h group was significantly lower than that of the NS3 or scNS3 groups ($*p<0.004$). A histogram shows the number of cells expressing each E_{app} value (bin width 0.05).

Q-MSI was first used to analyze cells transfected with CFP-NS3:NS4A and either YFP-RIG-I, YFP-MDA5, or YFP-LGP2. In this case, the regions of interest were whole cells that were successfully transfected with both CFP and YFP constructs and therefore expressing both a CFP and YFP signal. When CFP-NS3:NS4A was co-expressed with either YFP-RIG-I or YFP-MDA5, the range of E_{app} values straddled zero with some small positive and negative E_{app} values calculated. However, when CFP-NS3:NS4A was co-expressed with YFP-LGP2, the calculated E_{app} values were all above zero and reached as high as 20% (Figure 7-4 H and I).

A similar set of transfections was performed using CFP-LGP2 as the donor and either YFP-RIG-I, YFP-MDA5, or YFP-LGP2 as the acceptor. Again, whole cell regions of interest were selected where both cells were successfully transfected with both donor and acceptor and expressing both CFP and YFP signal. When CFP-LGP2 was co-expressed with both YFP-RIG-I and YFP-MDA5, the average E_{app} values over whole cells was approximately zero. When YFP-LGP2 was used as the acceptor, most cells had some FRET. The average E_{app} values for CFP-LGP2 and YFP-LGP2 cells was about 12% with values as high as 30% (Figure 7-4 J-K).

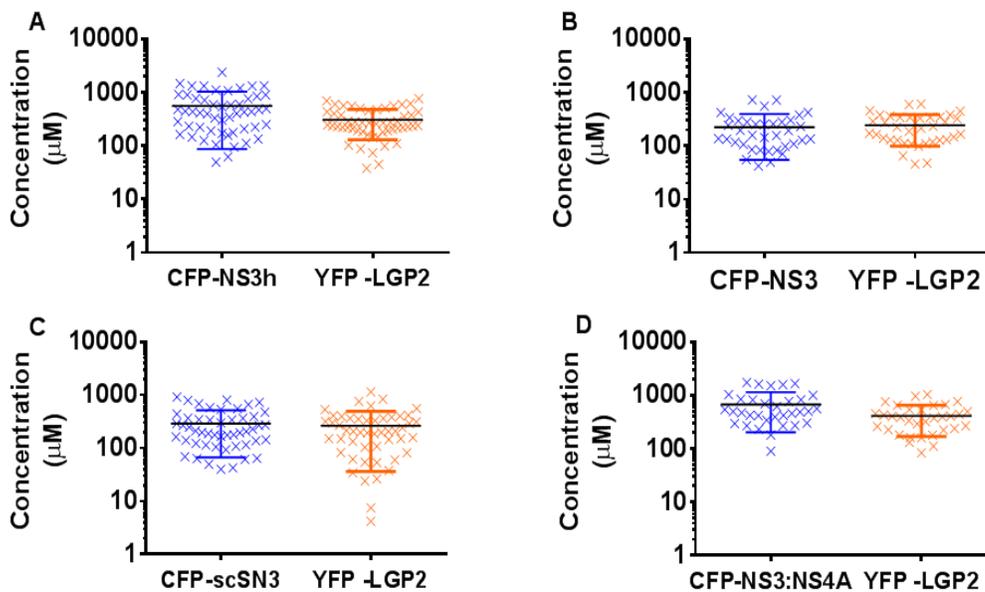


Figure 7-7: Concentrations of CFP tagged NS3 variants and YFP-LGP2 when co-expressed in HEK293T cells. The concentrations for the donor and acceptor in each cell co-expressing either CFP-NS3h, CFP-NS3, CFP-scNS3, or CFP-NS3:NS4A co-expressed with YFP LGP2 was analyzed by using the unmixed donor and acceptor intensity maps and calculating the average donor or acceptor intensity over a region of interest. This intensity was then compared to a standard curve of either CFP or YFP protein to calculate the final concentration of donor or acceptor (respectively) in the cell. (A-D).

To further explore the possible NS3:LGP2 interaction, three domain deletion mutants of NS3:NS4A were used. CFP-NS3:NS4A expresses the full length NS3, including both the helicase and protease domain, and the full length NS4A protein including both the protease co-factor and the membrane anchors (Figure 6-3-row E and Figure 7-6 A). The first domain deletion mutant (CFP-scNS3) eliminates the membrane anchors from the NS4A and covalently attaches the protease co-factor domain to the N-terminal protease domain of NS3. This protein retains the ability to cleave protease substrates but does not localize to foci in the cell due to the lack of membrane anchors (Figure 6-3-row D and Figure 7-6 B). The next NS3 mutant lacks the NS4A

co-factor completely (CFP-NS3) and is therefore not able to cleave MAVs in cells and is distributed diffusely throughout the cytoplasm (Figure 6-3-row C and Figure 7-6 C). The last NS3 domain deletion mutant (CFP-NS3h) only contains the NS3 helicase domain. It is also unable to cleave the MAVs substrate and is diffuse throughout the cytoplasm (Figure 6-3-row B and Figure 7-6 D). A fourth domain deletion mutant (CFP-scNS3pro) which contained only the NS3 protease domain along with the NS4A protease cofactor domain was also generated. However, when it was expressed in cells it not only lacked the ability to cleave MAVs, it was also aggregated in the cytoplasm of the cells. Given that CFP-scNS3pro lacked NS4A membrane anchors, it should have been diffuse throughout the cell and should not have formed sub-cellular foci. Since CFP-scNS3-pro was forming foci in the cell and because it was unable to cleave the RFP-NLS-MAVs substrate, it was concluded that CFP-NS3pro was not folding correctly and therefore not a reliable mutant to consider in this study (Figure 7-7A-B).

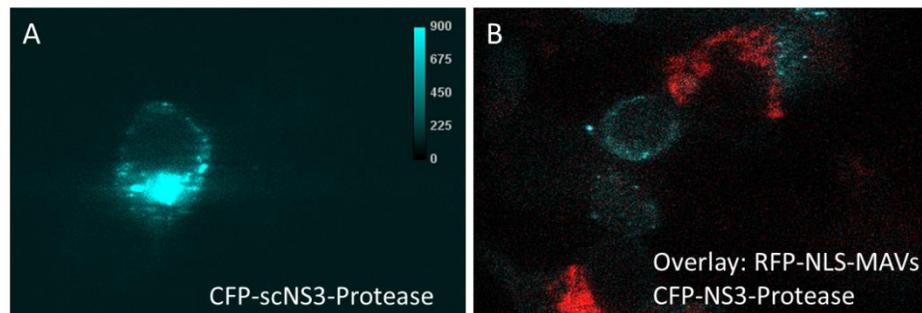


Figure 7-8: Representative images of CFP-NS3-protease expressed in cells. A plasmid expressing CFP-scNS3-pro was generated and expressed in HEK293T cells. CFP-scNS3-pro contains the 4A protease co-factor domain covalently linked to the NS3 protease domain. (A) After 24 hours of expression the cells were imaged by exciting at 840 nm. A representative image of 485 nm emission is shown. (B) CFP-scNS3-pro was co-transfected with RFP-NLS-MAVs. After 24 hours of expression the cells were imaged by exciting to 840 nm. A representative image overlay of the 485 nm emission and the 584 nm emission to target the CFP and RFP respectively is shown.

Each of the different domain deletion mutants of NS3 (excluding CFP-NS3pro) was transfected along with YFP-LGP2 and whole cell regions of cells displaying both CFP and YFP signals were selected. Average calculations of donor and acceptor were calculated for each ROI and found to be roughly equal within a given well (Figure 7-8). Despite the differences in the properties of each of the deletion mutants, they all retained an ability to interact with LGP2. This suggests that the interaction is occurring, at least in part, through the NS3 helicase domain. The range of E_{app} values for the set of NS3 deletion mutants ranged from around 5% to as high as 30%. The CFP-scNS3 construct displayed the highest average E_{apps} (Figure 7-6 E and F).

The previous experiments exploring the interaction of NS3 and LGP2 were all done by averaging E_{app} over a whole cell region. This approach could be lowering the apparent E_{app} values by averaging over regions of the cell where NS3 or LGP2 is not present (i.e. the nucleus) and therefore not able to interact. This is noticeably problematic in the cells expressing CFP-NS3:NS4A because the NS4A membrane anchors tethers the NS3 protein into specific sub-cellular foci. To localize the E_{app} analysis to the NS3:NS4A foci, the CFP-NS3:NS4A foci were selected in cells co-expressing each of the YFP tagged versions of the RLRs. The CFP-NS3:NS4A foci were defined as the bright CFP regions visible in the k^{DA} intensity map of cells co-expressing CFP-NS3:NS4A and any one of the YFP tagged RLRs (Figure 7-9 A and B). This analysis approach yielded similar E_{app} values to those seen in the whole-cell analysis (Figure 7-4 H and I). The E_{apps} calculated for the cells co-expressing CFP-NS3:NS4A and YFP-RIG-I or YFP-MDA5 were essentially zero with small E_{app} values straddling zero. The E_{apps} calculated in the cells co-expressing CFP-NS3:NS4A and YFP-LGP2 were still noticeably higher than those in the YFP-RIG-I or YFP-MDA5 expressing cells. Over 400 sub-cellular regions in cells co-expressing CFP-

NS3:NS4A and YFP-LGP2 were analyzed and results were remarkably reproducible, always yielding an average of about 5% E_{app} (Figure 7-9 C and D). The concentrations of the donor and acceptor were also calculated for the NS3:NS4A regions of interest (Figure 7-9 E-G).

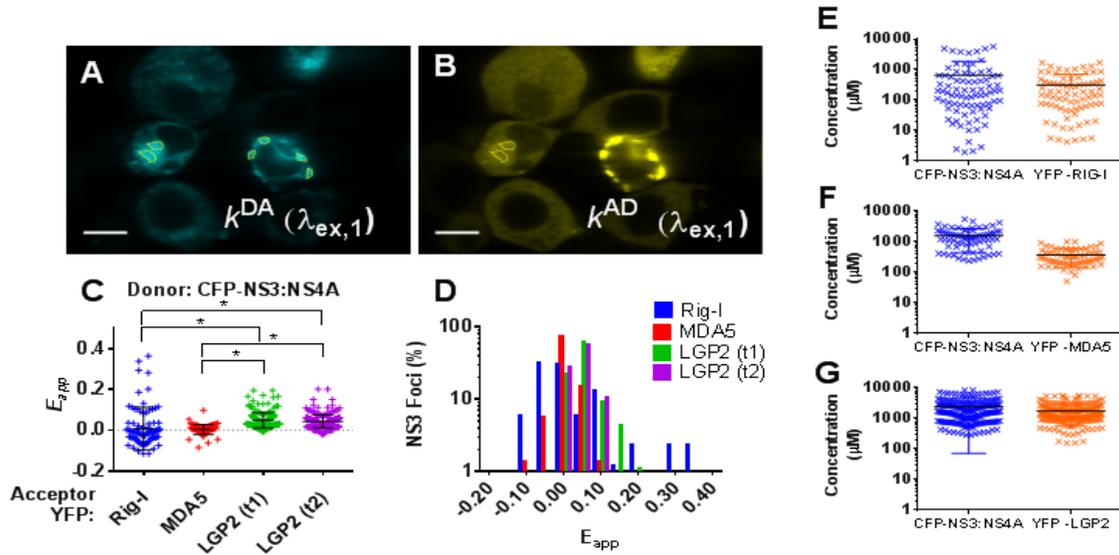


Figure 7-9: Sub-cellular NS3:NS4A region analysis of NS3 helicase interaction with RIG-I like receptors. HEK293T cells were co-transfected with plasmids encoding CFP-NS3:NS4A and either YFP-RIG-I, YFP-MDA5, or YFP-LGP2. After 24 hours, the cells were imaged and analyzed via Q-MSI. The sub-cellular NS3:NS4A regions was selected on the 2D spatial distribution map of CFP-NS3:NS4A fluorescence (A). This selection created a mask which could be applied to the 2D spatial distribution map of the YFP tagged RLRs (B). The average E_{app} values along with average donor and acceptor concentrations were calculated over each NS3:NS4A region. The scatter plot (C) shows the E_{app} value calculated for each ROI. The average $E_{app} \pm$ S.D. and the number of ROIs analyzed (n) are: RIG-I: 0.009 ± 0.10 (n=83), MDA5: 0.004 ± 0.024 (n=71), LGP2 (set 1): 0.050 ± 0.038 (n=180) and LGP2 (set 2): 0.044 ± 0.033 (n=222). A one-way ANOVA showed a significant main effect of group ($F(3,552) = 23.19$, $p < 0.0001$). The Tukey post-hoc showed that both RIG-I and MDA5 were significantly different from each LGP2 set ($*p < 0.0001$). The histogram (D) shows the number of ROIs expressing each E_{app} value (bin width = 0.1). The scatter plots (E-G) show the distribution of average concentrations at each NS3:NS4A region of interest.

7.3 Q-MSI analysis of the interaction of NS3 with LGP2 in the presence and absence of the PAMP poly I:C

LGP2 is known to bind and respond to the pathogen associated molecule pattern (PAMP), poly I:C in cells (Childs et al., 2013). Given that NS3 and LGP2 continue to show an interaction even when only the NS3 helicase domain was present, it is possible that the

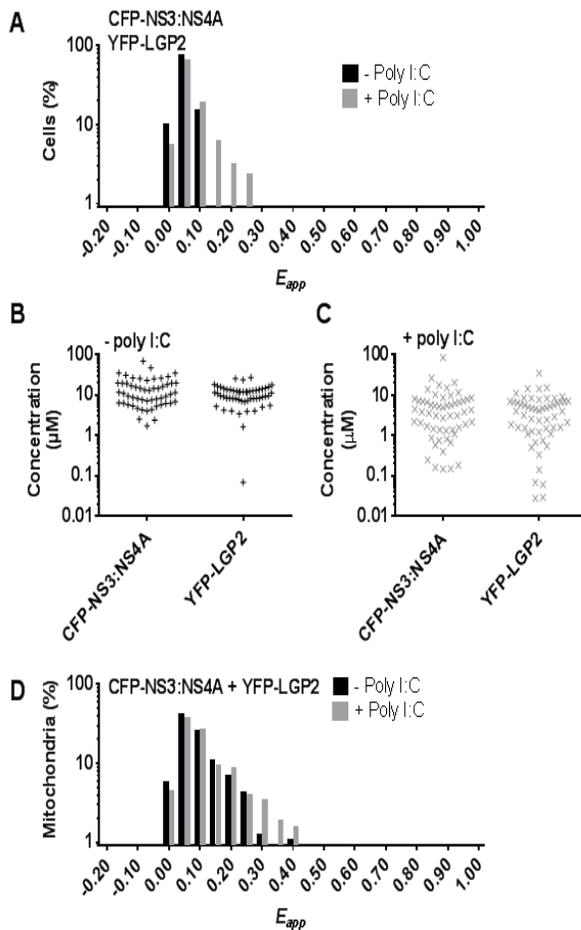


Figure 7-10: Analysis of NS3:NS4A interaction with LGP2 in the presence and absence of poly I:C. HEK293T cells were co-transfected with CFP-NS3-4A and YFP-LGP2 plasmids. The cells were incubated for 20 hours and then transfected with either 0.36 $\mu\text{g}/\text{mL}$ of poly I:C or vehicle. The cells were excited at 840 nm and 960 nm and the composite spectra un-mixed and regions of interest were selected in each image. The average E_{app} for the whole cells with and without poly I:C treatment was calculated along with the donor and acceptor concentrations. The histogram (A) shows the spread of E_{app} values with and without poly I:C. A Brown-Forsythe test was performed and yielded a p value of <0.01 . Concentrations of CFP-NS3-4A (donor) and YFP-LGP2 (acceptor) in cells not treated with poly I:C (B) and concentrations of CFP-NS3-4A and YFP-LGP2 in cells treated with poly I:C (C). To determine if the NS3:NS4A and LGP2 interaction was localized to the mitochondria, cells co-expressing CFP-NS3-4A or YFP-LGP2 were either treated with 0.36 $\mu\text{g}/\text{ml}$ of poly I:C or vehicle. The mitochondria were stained with Mito-Tracker Red and circled. A histogram showing the number of mitochondrial ROIs expressing each E_{app} value (bin width = 0.05) (D). A Brown-Forsythe test was performed and yielded a p value of 0.344.

interaction is, in part, mediated through a shared binding to RNA. To test this hypothesis, the CFP-NS3:NS4A plasmid was co-transfected with YFP-LGP2 in HEK 293T cells. After 24 hours of expression, the cells were treated with either poly I:C (0.36 $\mu\text{g}/\text{mL}$) or vehicle for four hours. The treated cells were imaged using successive 840 nm and 960 nm excitations and analyzed with Q-MSI. Whole cell regions of interest were circled and E_{apps} calculated. The apparent E_{app} values increased slightly with the addition of poly I:C from an average of 5% E_{app} to an average of 7% E_{app} (Figure 7-10 A). In addition, a shift in CFP and YFP protein concentration was observed with the addition of poly I:C. When poly I:C was added, a noticeable population of cells expressed lower overall CFP and YFP proteins, possibly indicating that the poly I:C treatment caused an interferon-induced repression of protein translation (Figure 7-10 B and C).

To isolate the interaction to the mitochondrial regions of the cells, the cells co-transfected with CFP-NS3:NS4A and YFP-LGP2 were stained with Mito-Tracker Red and then imaged. The Mito-Tracker Red dye stained the mitochondria of the cells and the Mito-Tracker signal was unmixed from the composite spectra. This yielded a map of Mito-Tracker intensity across the cells. Sub-cellular Mito-Tracker ROIs were selected using this Mito-Tracker Red intensity map in cells co-expressing CFP and YFP signal. The ROIs selected in the Mito-Tracker image were then applied as a “mask” to the CFP and YFP intensity images to specifically target the mitochondrial regions of the cells. The histogram of E_{app} values at the mitochondria show an increase from 7.7% to 8.2% with poly I:C treatment (Figure 7-10 D). Although there is a small increase in E_{app} at the mitochondria with the addition of poly I:C, that increase is not statistically significant.

7.4 Significance of LGP2 and NS3 interaction

In this study, the potential interaction of HCV NS3 helicase with the RIG-I like receptors was probed using ectopically expressed fluorescent fusion proteins and fully quantitative FRET imaging. The Q-MSI technique allows for the study of protein:protein interactions within sub-cellular regions in living cells. LGP2 was the only RLR which showed any significant interaction with NS3 and experiments with deleted NS3 domains showed that this interaction appears to be due, in part, to an interaction with the NS3 helicase domain. It's possible that the interaction is a direct protein:protein interaction or it may also be partially mediated through a joint interaction with RNA. However, given that the interaction of NS3 and LGP2 increased only slightly with the addition of poly I:C (Figure 7-10), it is unlikely that the interaction is solely mediated through a joint association with RNA.

The interaction of NS3 and LGP2 may be biologically relevant if it aids NS3 in inhibiting the innate immune response by helping NS3 locate cleavage targets, such as MAVS. Previous work has looked at the role of the helicase in NS3-mediated cleavage of MAVs. A mutant of NS3 (W501A) lacks helicase function but still retains the ability to block IRF-3 activation downstream of MAVs in response to Sendai virus infection, meaning that it is still likely able to cleave MAVs. However, NS3 expressing a non-functional protease was able to somewhat inhibit IRF-3 induction as compared to vector alone, implying that other pathways of inhibition may be present (Foy et al., 2003). Another study showed that NS3 lacking the helicase domain is sufficient for MAVs cleavage (Horner et al., 2012). Although some studies show that the helicase domain is not necessary for cleavage of MAVs, it's possible that the helicase domain

aids in the innate immune inhibition and MAVs cleavage. Perhaps the NS3 helicase domain increases the efficiency at which NS3 protease is able to localize to MAVs through interaction with an RLR.

The LGP2:NS3 contact is a curious interaction because LGP2 is not known to make direct contact with MAVs, due to lacking the CARDs. However, it is possible that LGP2 may help shuttle NS3 to areas with MAVs through the known LGP2 interaction with MDA5. LGP2 is known to make an oligomer with MDA5 to aid in its recognition and binding of viral substrates. MDA5 then, in turn, releases its CARDs to interact with the CARDs on MAVs. It's possible that NS3 binds to the LGP2:MDA5 complex and localizes to MAVs through the LGP2:MDA5 interaction with MAVs. One experiment which could test this idea would be to create mutants which prevent the LGP2:MDA5 interaction. Then these mutants could be co-expressed along with fluorescently tagged NS3. Q-MSI could be used to see if the loss of LGP2:MDA5 interaction affects the interaction of NS3 with LGP2 at the mitochondria in response to a PAMP. Since the LGP2:MDA5 interaction is thought to happen in the cytoplasm prior to the LGP2:MDA5 localization to MAVs, the results that poly I:C does not increase the interaction of NS3 with LGP2 at the mitochondria (Figure 7-10 D) but does increase the interaction in the cell as a whole (Figure 7-10 A) supports the idea that the LGP2:MDA5 interaction may be necessary for NS3 re-localization to MAVs.

Chapter 8 DDX3 causes LGP2 to re-localize into foci

The intracellular innate immune system is the cell's first line of defense against invading pathogens. It contains a group of proteins known as pattern recognition receptors (PRRs) which recognize conserved features of pathogens. These conserved features include nucleic acids, bacterial or viral proteins or carbohydrates and are jointly called pathogen associated molecular patterns (PAMPs). Once these PRRs have recognized their cognate ligand, they initiate a signaling cascade which results in the production of interferon and other cytokines (Errett and Gale, 2015b; Horner et al., 2011).

One group of PRRs includes the RIG-I like receptors (RLRs). The RIG-I like receptors are made up of RIG-I, MDA5, and LGP2. The RIG-I like receptors are DExD/H box RNA helicases which specifically recognize the RNA of viral invaders. Once they have bound to their RNA ligand, they signal through the adaptor protein mitochondrial anti-viral signaling protein (MAVs) to initiate the production of interferon. One location of the cell where the RLRs might detect viral RNA PAMPs is in the non-ribosomal cytoplasmic ribonuclear protein (RNP) granules (Protter and Parker, 2016). RNP molecules were first described by developmental biologists studying dark staining "germ" granules that form at one pole of animal embryos which hold maternal mRNA necessary for germ line differentiation (Schisa et al., 2001). Heat shock granules (HSGs), which are similar in composition, were later discovered when cultures of tomato cells were incubated at 40°C. HSGs contain heat shock proteins and "house-keeping" mRNAs sequestered in insoluble aggregates which dissolve once homeostasis has been restored

(Nover et al., 1983). Similar granules appear in mammalian cells in response to oxidative stress, heat shock, UV-radiation, and viral and bacterial infections.

It appears the role RNPs is to function as RNA repositories which sequester RNA when the cell is under stress and then disintegrate, releasing the translation machinery and RNA, once conditions have improved (Buchan et al., 2008; Kwon et al., 2007). These RNPs have been termed “stress granules” and the stress granules formed in response to viral exposure have been specifically called “anti-viral stress granules (Craig McCormick and Denys A Khapersky, 2017; Onomoto et al., 2012; Onomoto et al., 2014). Onomoto et al showed that the RIG-I like receptors localized to stress granules in response to influenza-1 infection (Onomoto et al., 2012) and it’s possible that the RIG-I like receptors are using the RNA harboring stress granules as sites of RNA recognition and anti-viral coordination.

8.1 Evidence that DDX3 recruits the RIG-I-like receptor LGP2 to Foci

This project was initiated to examine the interaction of the RIG-I like receptors with hepatitis C NS3 helicase. To probe these interactions, we constructed a fluorescently tagged LGP2 and several other “control” helicases that are not RLRs (DDX1, DDX3, and DDX5). The fluorescent recombinant LGP2 was tagged with cyan fluorescent protein (CFP) and the DDX1, DDX3, and DDX5 helicases were tagged with yellow fluorescent protein (YFP) (Figure 8-1 B) and co-expressed in HEK293T cells for 24 hours. Cells expressing CFP-LGP2 alone were also imaged

and the image corresponding to the CFP peak emission of (475 nm) is shown (Figure 8-1A).

Samples were excited at 800 nm and emission spectra were collected over 200 wavelengths.

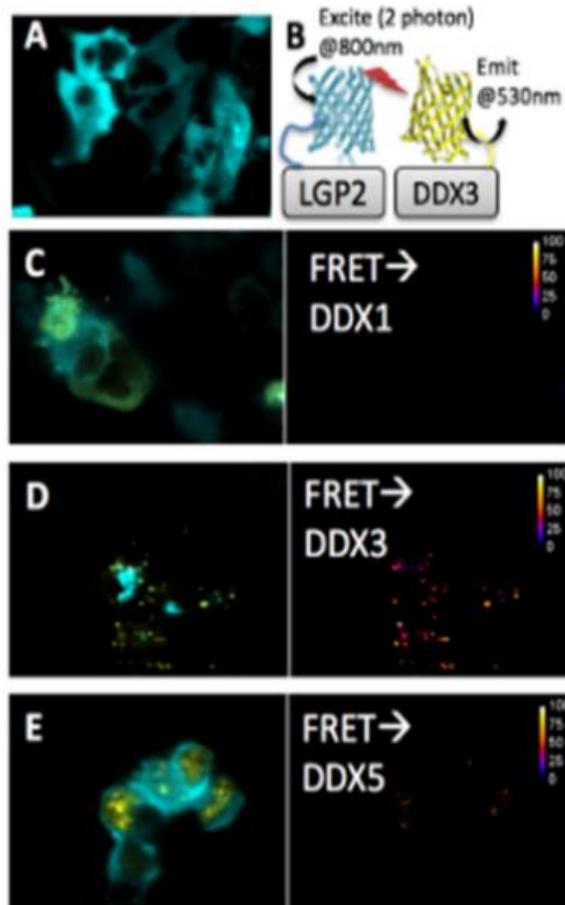


Figure 8-1 Co-overexpression of DDX3 and LGP2 causes DDX3 and LGP2 to accumulate in cytoplasmic foci. A) HEK293T cells transfected with the plasmid pCMV-mCFP-LGP2 which overexpresses the mCFP-LGP2 fluorescent fusion protein. The CFP signal appears to be localized to the cytoplasm of the cells. B) Diagram showing theoretical FRET between a CFP donor excited at 800 nm (two photon excitation) and a YFP acceptor emitting at 530 nm. C) Co-expression of CFP-LGP2 and YFP DDX1 in HEK293T cells. The unmixed CFP intensity map and YFP intensity map taken from the mixed image of the cells co-expressing CFP-LGP2 and YFP-DDX1 were overlaid. Both CFP-LGP2 and YFP-DDX1 are cytoplasmic proteins and pixel level FRET is shown. D) Co-expression of CFP-LGP2 and YFP-DDX3 in HEK293T cells. The unmixed CFP intensity map and YFP intensity map taken from the mixed image of the cells co-expressing CFP-LGP2 and YFP-DDX3 was overlaid. Cytoplasmic foci are seen and FRET is shown. E) Co-expression of CFP-LGP2 and YFP-DDX5 in HEK293T cells. The unmixed CFP intensity map and YFP intensity map taken from the mixed image of the cells co-expressing CFP-LGP2 and YFP-DDX5 was overlaid. DDX5 is a nuclear protein and LGP2 is a cytoplasmic protein. FRET between the two proteins is shown.

The donor and acceptor signal were un-mixed to generate maps of donor and acceptor signal using the FRET spectrometry technique discussed in Methods 3.5. The un-mixed donor (k^{DA}) intensity map and the un-mixed acceptor (k^{AD}) intensity map were used to calculate FRET (See Methods 3.5 for more details) and to generate maps of FRET intensity. The FRET generated here does not include a correction for direct excitation and therefore the intensity of FRET measured is not considered. Rather, the presence or absence of FRET in the each combination is considered only in relation to each other.

The k^{DA} and k^{AD} intensity maps for CFP-LGP2 and each of the YFP tagged DDX proteins were overlaid (Figure 8-1 C-E). While examining the images of these proteins in cells, we made the serendipitous observation that co-expression of DDX3 (but not the homologs DDX1 and DDX5) led to a clear re-localization of LGP2 from the cytoplasm (Figure 8-1A) to distinct cytoplasmic foci (Figure 8-1 D). Stress granules are identified by the consistent presence of molecules such as poly-A binding protein (PABP) and translation initiation factors (eIF3, eIF4E, and eIF4G) along with mRNA and small ribosomal sub-units (Kedersha and Anderson, 2007) and DDX3 has been identified as a key nucleating component of stress granules (Valiente-Echeverría et al., 2015). In addition, the over-expression of DDX3 alone can initiate the formation of stress granules (Shih et al., 2012). Therefore, it's possible that the foci observed in the images containing over-expressed DDX3 are stress granules and that LGP2 is re-localizing to the stress granules. In addition, FRET was observed between CFP-LGP2 and YFP-DDX3 within the cytoplasmic foci whereas no FRET was observed for the combinations of CFP-LGP2 and YFP DDX1 or CFP-LGP2 and YFP-DDX5 (Figure 8-1).

The re-localization of LGP2 to the DDX3 formed foci is reminiscent of what was seen by Onomoto et al who observed using immunofluorescence that the RLRs re-localized to anti-viral stress granules in response to influenza-1 infection. The stress granules were identified using fluorescent antibodies targeted to stress granule markers (such as eIF3) (Onomoto et al., 2012). Given that DDX3 over-expression is able to nucleate the formation of stress granules, it is likely that the foci seen in the images with over-expressed DDX3 are stress granules. It is possible that DDX3 is helping to shuttle LGP2 to stress granules and this could be shown through using immunofluorescence to stain endogenous DDX3, LGP2, and stress granule markers in cells infected with a virus that induces stress granules.

Chapter 9 Homo and hetero interactions of RIG-I like receptors in response to poly I:C

The RIG-I like receptors (RLRs) are a set of sub-cellular DExD/H helicases which recognize the pathogen associated molecular patterns (PAMPs) of RNA viruses in almost every cell in the human body. The RLRs are made up of three different pattern recognition receptors (PRRs); retinoic acid inducible gene -1 (RIG-I), melanoma differentiation associated protein-5 (MDA5), and laboratory of genetics and physiology protein-2 (LGP2). RIG-I and MDA5 both contain tandem CARDs, a helicase domain, and a C-terminal regulatory domain. RIG-I and MDA5 exist in an auto inhibited state in the cytoplasm of cells with their CARD domains swung in towards their helicase domain, preventing un-wanted CARD interaction and signaling. The C-terminal regulatory domain is exposed to the cytosol and is the first domain to recognize the RNA PAMP ligands. Once the C-terminal domain has bound to a prospective RNA PAMP ligand, it swings in to allow the helicase domain to interact with the ligand. Then the CARDs are released to oligomerize and interact with their downstream target, mitochondrial associated anti-viral signaling protein (MAVs).

LGP2 is different from both RIG-I and MDA5 in that it lacks the CARDs necessary for downstream signaling. Instead, LGP2 exerts its effects through interaction with the other RLRs. The effect of LGP2 in innate immune signaling is antithetical with some studies indicating a positive regulatory effect while others indicate a negative regulatory role. Evidence that LGP2 exerts a positive effect on interferon signaling include experiments showing that mice who express mutations in LGP2 are more susceptible to certain virus infections than littermates who

have a functional LGP2. In addition, experiments in cells lacking both LGP2 and MDA5, show a synergistic signal increase of an IRF3 reporter from exogenous co-expression of both MDA5 and LGP2, but not either one expressed independently. This suggests that the interaction of LGP2 with MDA5 induces interferon signaling. Examples of LGP2 negative regulation can be seen from interferon reporter assays that show overexpression of LGP2 from an exogenous vector can downregulate interferon signaling induced via a viral infection (Childs et al., 2013; Pippig et al., 2009; Rodriguez et al., 2014). Alternatively, LGP2 does not enhance RIG-I interaction with MAVs and interferon signaling at low concentrations, but instead seems to only inhibit RIG-I at high concentrations (Bruns and Horvath, 2012). A unifying model has emerged where LGP2 interaction with MDA5 exerts a concentration dependent enhancement of interferon signaling until it reaches a very high concentration. Then LGP2 takes on an inhibitory role for both MDA5 and RIG-I, where its high-affinity binding to RNA PAMPS sequesters the RNA from recognition by RIG-I or MDA5, effectively slowing down signaling (Rodriguez et al., 2014).

This study uses recombinant fluorescent fusion proteins to look at the homo and hetero interactions of the RLRs with other RLRs and with the downstream signaling protein, MAVs. Q-MSI, which employs quantitative Förster Resonance Energy Transfer (FRET) with two-excitation wavelengths and fully quantitative analysis was applied to show that the RLRs form ordered complexes in living cells and that these interactions may be modulated through the presence or absence of a PAMP ligand. Prior studies have shown that RIG-I, MDA5, and LGP2 all form oligomers (Bruns et al., 2013; Wu et al., 2013; Wu et al., 2014) and that the hetero-interaction between MDA5 and LGP2 is important for MDA5 signaling (Bruns and Horvath, 2015; Bruns et al., 2013). But, the interactions have never before been seen in living cells. This study is able to

show the interactions of the RLRs at the mitochondria in living cells and to demonstrate that this interaction shifts with the addition of a PAMP RNA ligand.

9.1 Sub-cellular analysis of the homo-interactions of the CFP and YFP tagged RIG-I like receptors in response to poly I:C

Q-MSI was used to examine the current models which show the RIG-I like receptors responding to viral RNA PAMPs by forming larger oligomeric complexes and interacting with

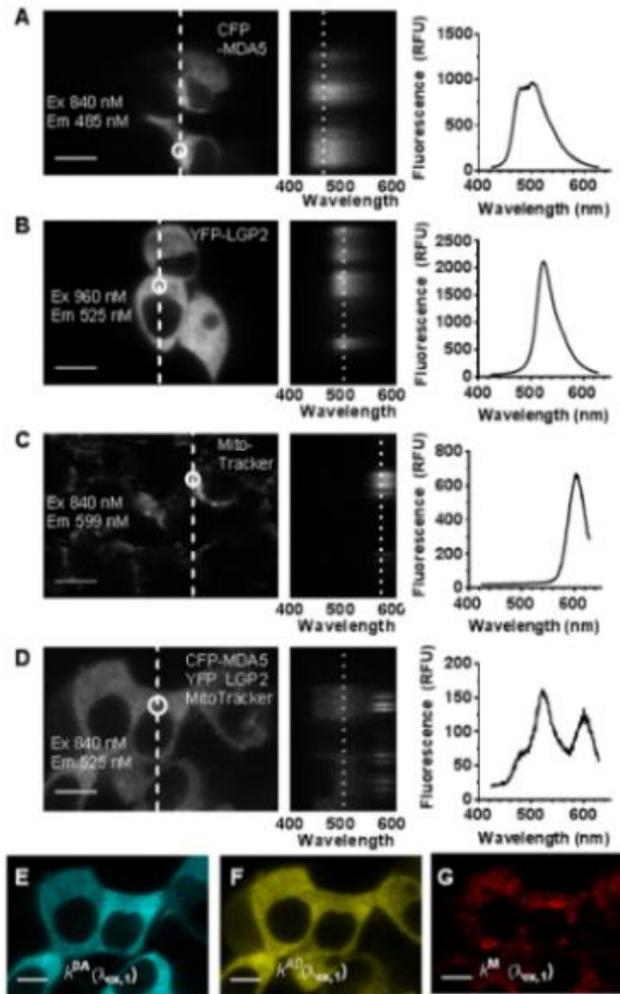


Figure 9-1 Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing donor, acceptor, and the sub-cellular mitochondrial tag, MitoTracker. A) Cells expressing only CFP-MDA5 and excited at 840 nm (two-photon excitation). The peak 485 nm emission is shown as a CFP intensity map of the cells. The orthogonal view of the circled region in the cell intensity map is shown along with the calculated fluorescent spectrum. B) Cells expressing only YFP-LGP2 and excited at 960 nm (two-photon excitation). The peak 525 nm emission is shown as a YFP intensity map of the cells. The orthogonal view of the circled region in the cell intensity map is shown along with the calculated fluorescent spectrum. C) Cells stained only with MitoTracker and excited at 840 nm (two-photon excitation). The peak 599 nm emission is shown as a MitoTracker intensity map of the cells. The orthogonal view of the circled region in the cell intensity map is shown along with the calculated fluorescent spectrum. D) Cells co-expressing CFP-MDA5 and YFP-LGP2 and stained with MitoTracker were excited at 840 nm. The fluorescence intensity map showing the 525 nm emission is shown. The orthogonal view and spectrum of the circled region is also shown. E) The 2D spatial distribution map of the donor fluorescence in the presence of acceptors $k^{DA}(\lambda_{ex1})$. F) 2D map of acceptor fluorescence in the presence of donor $k^{AD}(\lambda_{ex1})$. G) 2D map of MitoTracker $k^M(\lambda_{ex1})$ intensity.

MAVS at the mitochondrial membrane. If the models are correct, then it should be expected that the fluorescently tagged RLRs accumulate at the mitochondria. Since the ISRE promoter is stimulated by the over-expression of the RLRs (Figure 7-3F), it would be expected that this accumulation should occur even in the absence of PAMP but should be enhanced in the presence of PAMP. However, upon visual inspection of the images, there was no sub-cellular localization evident when the RLRs were expressed in HEK293T cells (Figure 7-3 C-E). Instead the proteins looked diffuse throughout the cytoplasm of the cells. Therefore, to localize the analysis to the mitochondria, the Mito-Tracker Red dye was used to label the mitochondria of live cells co-expressing the CFP and YFP tagged RLRs. The Mito-Tracker Red dye was selected because it fluoresces in a region of the spectrum distinct from CFP and YFP. The dye was used to select mitochondrial regions of the cell for analysis with Q-MSI. To identify and separate the mitochondrial localized YFP and CFP tagged proteins, the spectra from cells singly expressing either CFP (Figure 9-1A) or YFP (Figure 9-1B) or cells only stained with the Mito-Tracker Red (Figure 9-1C) were used to unmix the composite spectra from the cells co-expressing CFP and YFP and also stained with Mito-Tracker (Figure 9-1D). The regions of interest (ROIs) were identified on the CFP (k^{DA}) map (Figure 9-1E) and the YFP (k^{AD}) (Figure 9-1F) map by applying the mask created on the Mito-tracker (k^M) map (Figure 9-1G). This allowed for the calculation of FRET (E_{app}) and concentrations of CFP and YFP proteins specifically in the mitochondrial regions of interest. No FRET would be visible if there is no interaction of CFP and YFP at the mitochondria. If there is a small amount of CFP and YFP interacting pairs at the mitochondria or if there is a complex with a lot of donors and only a few acceptors, then a low level of FRET would be detected in the mitochondrial regions. If there are many CFP and YFP interacting pairs

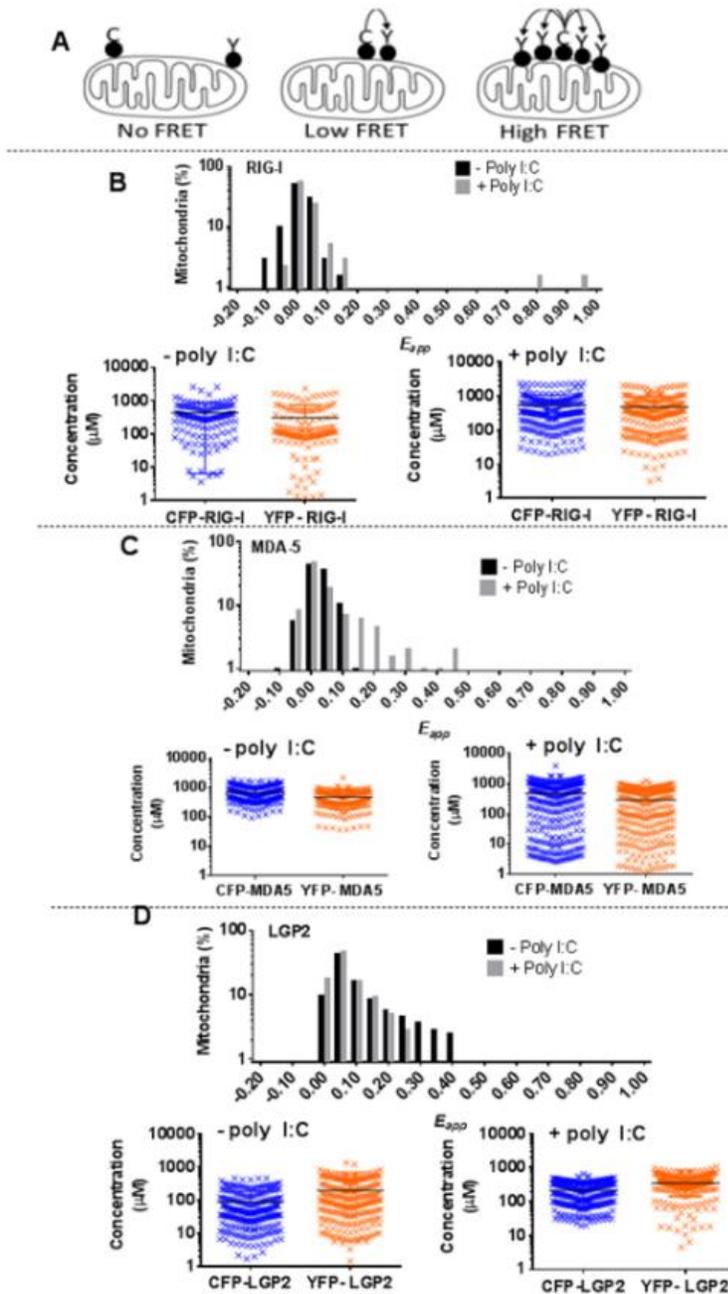


Figure 9-2 Response of mitochondrial localized RIG-I, MDA5, and LGP2 to poly I:C exposure. A) Expected FRET from mitochondrial localized CFP and YFP proteins under different proposed interaction conditions. For each experiment, cells co-expressing a donor and acceptor tagged protein were exposed to poly I:C (0.36 μ g/mL media) for four hours prior to imaging. The cells were stained with MitoTracker and then analyzed with Q-MSI where the mitochondrial regions were selected using the MitoTracker 2D intensity map and the selections applied as a mask to the donor and acceptor images. The average E_{app} for each ROI was calculated. B) Histogram (0.05 bin size) showing the calculated average E_{app} for each ROI in cells co-expressing CFP-RIG-I and YFP-RIG-I. A Brown-Forsythe test was performed and yielded a p value <0.01. Plots of the average concentrations of CFP and YFP within each ROI region measured (with and without poly I:C treatment) is also shown. C) A histogram showing the calculated average E_{app} s for each ROI in cells co-expressing CFP-MDA5 and YFP-MDA5. A Brown-Forsythe test was performed and yielded a p-value of <0.001. Plots of the average concentrations of CFP and YFP within each ROI region measured (with and without poly I:C treatment) is also shown. D) A histogram showing the calculated average E_{app} s for each ROI in cells co-expressing CFP-LGP2 and YFP-LGP2. A Brown-Forsythe test was performed and yielded a p-value of <0.001. Plots of the average concentrations of CFP and YFP within each ROI region measured (with and without poly I:C treatment) is also shown.

or if there is a complex with many more acceptors than donors, then there would be very high FRET at the mitochondria (Figure 9-2A).

In the first set of experiments, cells were co-transfected with plasmids expressing recombinant fluorescent YFP-RIG-I and CFP-RIG-I or YFP-MDA5 and CFP-MDA5. Some of the cells were treated with poly I:C at a concentration of 0.36 $\mu\text{g}/\text{mL}$ in cell growth media and the rest of the cells were exposed to vehicle (TurboFect) only. Poly I:C is a synthetic double stranded RNA PAMP which both RIG-I and MDA5 recognize. Previous work has shown that RIG-I and MDA5 accumulate at the mitochondria in response to poly I:C (Hou et al., 2011). When Q-MSI was used to calculate the E_{app} at the mitochondria of the cells co-expressing either YFP-RIG-I:CFP-RIG-I or YFP:MDA5:CFP-MDA5, clear differences were apparent between the set treated with poly I:C and the set not treated with poly I:C. Although the average E_{app} for RIG-I only changed slightly from 3% to 4% with the addition of poly I:C, there was a very noticeable shift in overall population distribution. The addition of poly I:C caused the accumulation of a population of YFP-RIG-I and CFP-RIG-I molecules interacting at the mitochondria with high E_{app} values ranging from 50% to 90%. These high FRET values indicate the development of a few large oligomeric complexes at the mitochondria in the presence of poly I:C which were not present in the absence of poly I:C (Figure 9-2B).

When CFP-MDA5 was co-expressed with YFP-MDA5 and the E_{app} at the mitochondria was analyzed via Q-MSI, a shift in the population of mitochondrial ROIs expressing higher E_{apps} in response to poly I:C is visible. Almost identical populations of E_{apps} are visible around 0 to 10 % E_{app} in those cells treated and un-treated with poly I:C. This is expected, because not every

cell expressing the CFP and YFP tagged MDA5 proteins would have necessarily been exposed to the poly I:C application. However, a segment of ROIs with higher E_{apps} develops in response to poly I:C. This indicates the formation of a set of higher order oligomeric complexes at the mitochondria in response to poly I:C (Figure 9-2C). It is important to note that the spread of the data is not error, but instead shows a variety of different combinations of donor and acceptor molecules within a population of oligomeric complexes.

Lastly, the same sub-cellular analysis was performed on cells co-transfected with plasmids expressing CFP-LGP2 and YFP-LGP2. LGP2 lacks the CARD regions necessary for direct interaction with MAVs. Therefore, we did not expect to see much accumulation of LGP2 at the mitochondria in response to poly I:C because any mitochondrial accumulation of LGP2 would most likely occur through interaction with the low levels of endogenous MDA5. Not surprisingly then, the results show that there is not a large difference in the FRET between YFP-LGP2 and CFP-LGP2 at the mitochondria in response to poly I:C and a sub-population of higher FRET efficiencies does not develop, like what was seen with RIG-I and MDA5. In fact, less FRET is seen when poly I:C is added, suggesting that the addition of an RNA PAMP is causing the breakup of some of the LGP2:LGP2 complexes (Figure 36D). One explanation for this decrease in E_{apps} in response to poly I:C, is that LGP2 homo-oligomers may be dissociating to interact with MDA5 on the RNA and to aid in MDA5 RNA recognition. It's possible that the addition of poly I:C stimulated the endogenous expression of more MDA5, and that LGP2 is now interacting with more MDA5:RNA complexes, thus diluting the direct LGP2:LGP2 interaction and reducing FRET efficiency at the mitochondria.

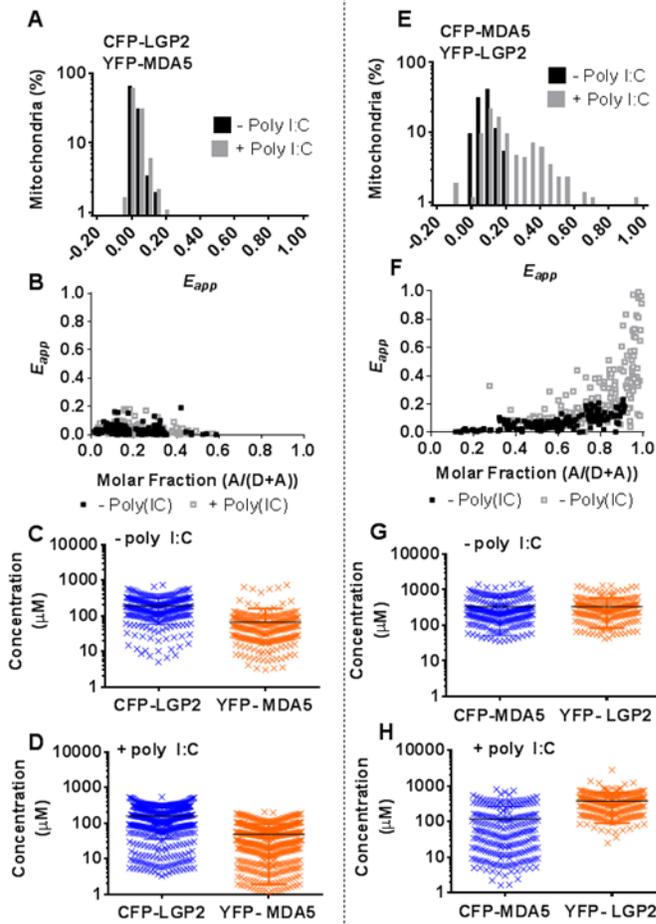


Figure 9-3: Response of mitochondrial localized LGP2 and MDA5 to poly I:C. In both experiments, cells were transfected with plasmids co-expressing a fluorescent recombinant protein construct tagged with either CFP or YFP. The cells allowed to incubate for 20 hours prior to poly I:C (0.36 $\mu\text{g}/\text{mL}$ of media) transfection. After 4 hours of poly I:C exposure, the cells were stained with MitoTracker-Red to identify the mitochondria. Mitochondrial regions were selected using the MitoTracker staining and those regions of interest were then applied to the CFP and YFP expressing images to calculate an average E_{app} and an average concentration of CFP and YFP for each mitochondrial ROI. A) Histogram FRET analysis of cells co-expressing CFP-LGP2 and YFP-MDA5 with and without poly I:C transfection. A Brown-Forsythe test was performed and yielded a p value of 0.03. B) E_{app} values vs. the molar fraction of acceptor molecules with and without poly I:C. C) and D) Average concentrations of CFP and YFP proteins within selected ROIs with and without poly I:C transfection. E) Histogram of the FRET analysis of cells co-expressing CFP-MDA5 and YFP-LGP2 with and without poly I:C transfection. A Brown-Forsythe test was performed and yielded a p value of <0.001. F) A E_{app} values vs. the molar fraction of acceptor molecules with and without poly I:C exposure. G) and H) Average concentrations of CFP and YFP for each ROI with and without poly I:C transfection.

Previous research has shown that there is a special interaction between LGP2 and MDA5 which is not present between RIG-I and LGP2 (Bruns et al., 2014; Bruns and Horvath, 2015; Bruns et al., 2013; Rodriguez et al., 2014). To examine this interaction, cells were co-expressed with plasmids expressing recombinant fluorescent CFP-LGP2 and YFP-MDA5. The cells were then treated with either poly I:C (0.36µg/mL) for four hours or vehicle. After exposure to poly I:C or vehicle the cells were stained with Mito-Tracker and the mitochondrial regions identified and selected. The resulting ROIs were analyzed by Q-MSI. In the absence of poly I:C, these cells showed some LGP2:MDA5 interaction with E_{apps} between 0-20% confirming that simple over-expression of LGP2 and MDA5 can result in the formation of oligomeric interactions at the mitochondria (Figure 9-3A). A small population of higher FRET efficiencies developed in response to the addition of poly I:C. (Figure 9-3A). It was also noted that in all the ROIs measured, the molar fraction of acceptor ($[A]/([D]+[A])$) was shifted towards complexes containing many more donors than acceptors (Figure 9-3B). The overall range of molar fraction of acceptor was 0 to 0.6 with the majority of the ROIs expressing molar fractions around 0.2. This would imply that the mitochondrial ROIs contain many more LGP2 molecules than MDA5 molecules.

To test this notion, a second transfection was performed where the fluorescent tags on the MDA5 and LGP2 molecules were switched. This time, CFP-MDA5 and YFP-LGP2 were co-transfected in HEK293T cells and again some of the cells were treated with poly I:C (0.36µg/mL) and some were treated with vehicle (TurboFect) for four hours. After poly I:C treatment, the cells were stained with Mito-Tracker Red so the mitochondria could be located and selected. The mitochondrial ROIs were analyzed with Q-MSI. The combination of CFP-MDA5 and YFP-

LGP2 yielded similar FRET efficiencies to the previous combination when the tags were switched in the absence of poly I:C. In the absence of poly I:C, the FRET efficiencies ranged from 0-20%. However, when poly I:C was added, there was a dramatic shift to higher FRET efficiencies with a large population of ROIs exhibiting FRET efficiencies between 20%-60% and a few ROIs with even higher FRET efficiencies (Figure 9-3E). In addition, the molar fraction of acceptors shifted to being between 0.4 to 1.0 for both the set with and without poly I:C treatment (Figure 9-3F). However, the highest FRET efficiencies were seen only in the set treated with poly I:C. This confirms the hypothesis that the mitochondrial ROIs contain more LGP2 than MDA5 molecules. It also shows that the interaction of LGP2 and MDA5 is increased at the mitochondria in response to poly I:C stimulation evidence by the dramatic increase in E_{app} with the addition of poly I:C treatment.

9.3 Analysis of the hetero-interactions of RIG-I receptors like receptors in response to poly I:C at the pixel level

Previous experiments were performed by applying Q-MSI to proteins tagged with either CFP or YFP. Although the CFP:YFP FRET pair works well for Q-MSI, it requires a correction for direct excitation of the acceptor at the donor excitation wavelength. This is because the peak excitation for CFP occurs around 840 nm (two photon excitation) and YFP is still easily excited at 840 nm (Figure 3-1 A). To correct for direct excitation of the acceptor, the quantity of acceptor fluorescence in the absence of donor must be calculated. To accomplish this, a second excitation at a longer wavelength targeting the acceptor alone is used (See Methods 3.5 for more details). The time it takes to move between the first (FRET) excitation and the second

excitation is about 40 seconds which is sufficient time to allow for molecular diffusion. Since the proteins have time to diffuse away from their original position between the first and second excitations, only average FRET efficiencies were calculated over selected regions of interest. The correction for direct excitation and the required second excitation prevents the calculation of FRET at pixel level.

To calculate FRET at pixel level, the fluorescent protein pairs used to visualize the RIG-I like receptors were switched to GFP₂ as the donor and either YFP or Venus as the acceptor. The GFP₂ protein excitation spectrum is spectrally very similar to wild-type GFP. Unlike CFP, which excites optimally at 840 nm (two photon excitation), GFP₂ (and wild-type GFP) excite optimally at 800 nm. At 800 nm, both YFP and Venus are not excited at all (Figure 3-1 B and C). This allows the FRET equation to be simplified because the fluorescence of the acceptor at wavelength 1 is now equal to zero (See Methods 3.5 for details). When using either GFP₂ and YFP or GFP₂ and Venus the calculation of E_{app} can be accomplished using only one excitation wavelength. This eliminates the issue of molecular diffusion and allows for a calculation of E_{app} at the pixel level.

Pixel level E_{app} values can then be organized into histograms of E_{app} values contained within selected regions of interest. The pixel level histograms generated for each region of interest contain many characteristic peaks and features which may be attributed to unique combinations of donors and acceptors within the large oligomeric complexes. To gain an understanding of the overall predominant E_{app} combinations within the regions of interest, predominant peaks from each histogram are collected and binned in an overall meta-

histogram. This meta-histogram represents the overall predominant donor and acceptor configurations within the region of interest (Stoneman et al., 2017).

9.3.1 Pixel level FRET analysis of the interaction of GFP₂-LGP2 and Venus-MDA5

Pixel level, fully quantitative FRET analysis was used to examine the interaction between GFP₂-LGP2 and Venus MDA5 at the mitochondrial membrane. Previous experiments showed a small increase in average FRET efficiency when the mitochondrial regions of the cell were analyzed with Q-MSI and LGP2 was tagged with the donor (CFP) and MDA5 was tagged with the acceptor (YFP) (Figure 9-3A). A similar increase in FRET efficiency would be expected when the tags were changed to a FRET pair which does not require a correction for direct excitation. In this case the donor is GFP₂ and is tagged to LGP2 and the acceptor is Venus and is tagged to MDA5.

Plasmids expressing recombinant fluorescent GFP₂-LGP2 and Venus-MDA5 were transfected into HEK 293T cells. The cells were allowed to incubate and express the fluorescently tagged proteins for 20 hours. After 20 hours of protein expression, the cells co-transfected with both GFP₂-LGP2 and Venus-MDA5 were transfected with either poly I:C (0.36µg/mL) or vehicle. The cells were incubated for additional 4 hours. Then, the mitochondria of the cells were stained with Mito-Tracker Red. The cells were excited at 800 nm which excited the donor and did not directly excite the acceptor. The 800 nm excitation also excited the Mito-Tracker Red. Elementary spectra were collected using cells solely expressing GFP₂-LGP2, Venus-MDA5, and only stained with Mito-Tracker Red. The elementary spectra were used to deconvolute the composite image and to calculate E_{app} at every pixel in the image. A map

depicting the 2D spatial distribution of pixel-level E_{app} s was also generated. For more detail on microscopy, un-mixing, or E_{app} calculation, please see Methods 3.5.

Un-mixing of the composite spectra in cells co-expressing GFP₂-LGP2 and Venus-MDA5 yielded the 2D spatial fluorescence intensity maps of k^{DA} (Figure 9-4A), k^{AD} (Figure 9-4B) and of Mito-Tracker, k^M , (Figure 9-4C).

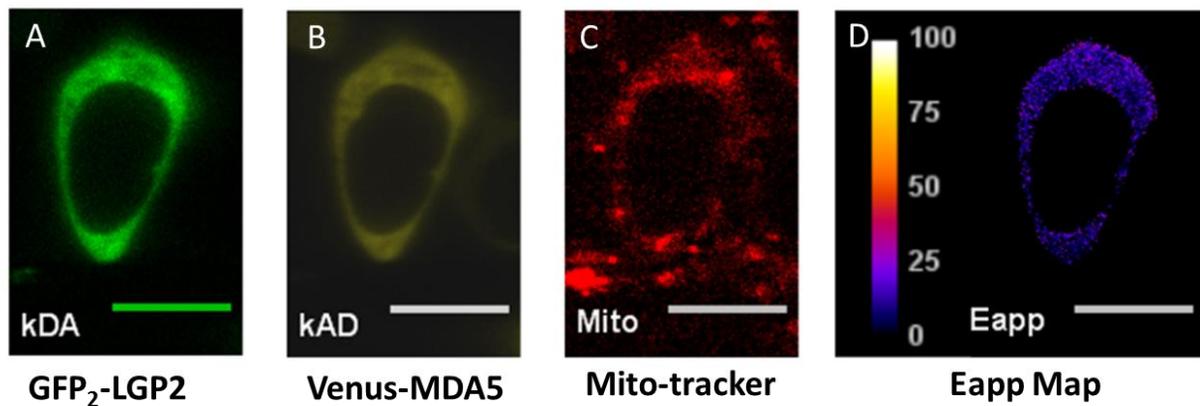


Figure 9-4: Unmixed k^{DA} , k^{AD} , MitoTracker and E_{app} (FRET) 2D intensity maps. Cells were co-transfected with plasmids expressing GFP₂-LGP2 and Venus-MDA5. The mitochondria were stained with MitoTracker-Red and the FRET was calculated at a pixel level. A) The 2D fluorescence intensity map for GFP₂-LGP2 (k^{DA}). B) The 2D fluorescence intensity map for Venus-MDA5 (k^{AD}). C) The 2D fluorescence intensity map for MitoTracker. D) The FRET intensity distribution for cells co-expressing GFP₂-LGP2 and Venus-MDA5.

The k^{DA} and k^{AD} values calculated at each pixel were used to calculate the 2D spatial distribution of E_{app} (Figure 9-4D). The Mito-Tracker map was used to select the mitochondrial regions of the cell (Figure 9-5A) and then this mitochondrial region mask was applied to the E_{app} map to select only the mitochondrial regions (Figure 9-5B). For each mitochondrial region of interest, a histogram of E_{app} peaks was created (Figure 9-5C). The collection of E_{app} histograms was then analyzed to select the predominant peak in each histogram. The predominant peak

was selected from each histogram and compiled into a meta-histogram with a bin width of 2 (Figure 9-5D).

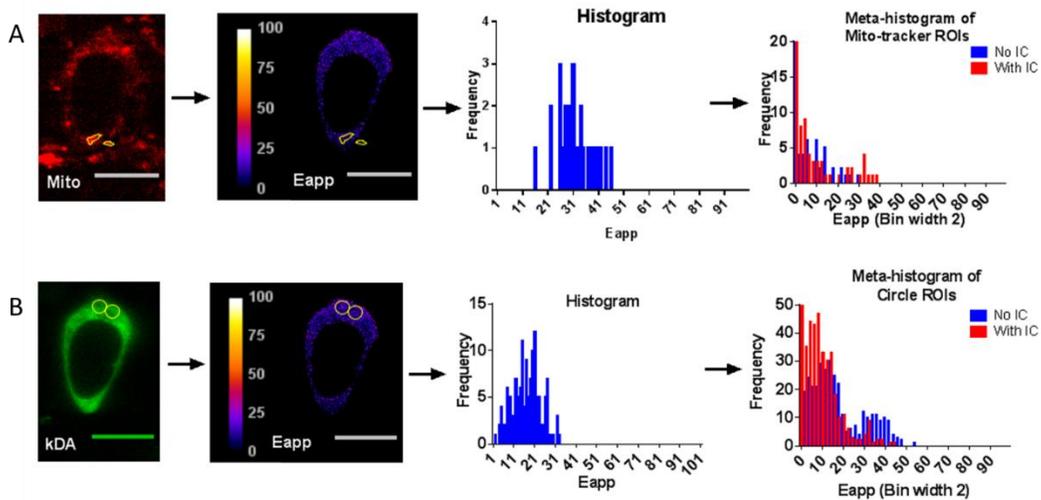


Figure 9-5: Pixel level FRET analysis of sub-cellular regions of cells co-expressing GFP₂-LGP2 and Venus-MDA5. A) Cells co-expressing GFP₂-LGP2 and Venus-MDA5 were stained with MitoTracker to locate the mitochondria. The unmixed MitoTracker image was used to select sub-cellular mitochondrial regions. The selected mitochondrial ROIs were then used as a mask and applied to the E_{app} intensity map to select the E_{app} intensity within each mitochondrial region. The E_{app} values in each mitochondrial region can be graphed as a histogram showing the frequency of each E_{app} value within that ROI. The peaks from these individual ROI histograms were selected to generate a Meta-histogram compiling all the peak E_{app} values across all mitochondrial ROIs selected. B) The same set of cells used for analysis in (A) were re-analyzed by selecting general cytoplasmic regions using a consistent circle comprising 146 pixels. The circle was placed throughout the k^{DA} intensity map such that most of the cell's cytoplasm was selected but none of the circles were over-lapping. The mask created on the k^{DA} intensity map was then applied to the E_{app} intensity map to select random cytoplasmic regions of a consistent size. The E_{app} values from the selected cytoplasmic ROIs can each generate a histogram of E_{app} values within that ROI. The peak E_{app} value from each cytoplasmic ROI histogram was selected to generate a Meta-histogram comprising all peak E_{app} values selected

Most of the predominant pixel-level E_{app} values within the mitochondrial regions were zero (Figure 9-5A). This is not surprising considering that the interacting proteins should exist on the mitochondrial membrane, yet it's impossible to just select the mitochondrial membrane. Therefore, many of the mitochondrial ROIs contain many pixels which would constitute the inside of the mitochondria and not just the mitochondrial membrane. Since the RLRs are not expected to accumulate inside the mitochondria, it would make sense that this region would be largely devoid of FRET. When the mitochondrial regions were analyzed in previous experiments and an average taken over the whole area, this detail was lost in the average. But at pixel level, it is apparent that much of the mitochondrial ROIs do not contain interacting GFP₂ and Venus proteins. However, in some of the histograms, the predominant peak was non-zero and for the set of cells not treated with poly I:C, the E_{app} range was between 0-20%. This is very similar to what was seen with CFP-LGP2 and YFP-MDA5 (Figure 9-3A) which also showed an E_{app} range between 0-20%. When poly I:C was added to the cells, a sub-set of cells exhibited E_{app} values larger than 20% and up to a maximum of 40% (Figure 9-5A). Previous experiments also showed a small increase in E_{apps} within the population of cells transfected with poly I:C, however it never reached 40%. This could be due to the averaging of E_{app} values with regions of zero interaction which is necessary when the tags CFP and YFP are used but irrelevant with GFP₂ and Venus when a pixel level analysis is possible. Pixel level analysis allows these zero pixels to be separated out of the overall analysis instead of averaged.

Another set of ROIs were selected on the same set of cells co-expressing GFP₂-LGP2 and Venus-MDA5. These ROIs were selected using the k^{DA} image (Figure 9-5B). A circle containing

146 pixels was used as the ROI template and was placed throughout the cell selecting circular, non-overlapping cytoplasmic regions. The regions were only selected in cells expressing both GFP₂ and Venus signal. This created a mask of cytoplasmic regions containing the same number of pixels and this mask was applied to the E_{app} map to create E_{app} histograms for each ROI selected. The predominant peak was selected from each of the E_{app} histograms and compiled into a meta-histogram (Figure 9-5 A and B).

The ROIs selected were randomly generated throughout the cell and there are significant regions which do not contain interacting GFP₂-LGP2 and Venus-MDA5. Therefore, the most common E_{app} value in the meta-histogram was zero. The non-zero E_{app} values for the cells not treated with poly I:C were between 0-20%. This is the same range seen when the mitochondrial ROIs were circled (Figure 9-5A) and when the fluorescent tags used were CFP and YFP (Figure 9-3A). However, a small population of higher FRET efficiencies were present in the cells not transfected with poly I:C. In those cells, E_{apps} as high as 50-60% were calculated. The cells transfected with poly I:C also showed a population of ROIs with predominant peak E_{apps} above 20%, but there were not as many of them and they only reached a maximum of 45%. The cytoplasmic ROIs include not only the mitochondria, but also the surrounding cytoplasmic regions. The current model of RLR function predicts that MDA5 and LGP2 interact in the cytoplasm and jointly recognize PAMP RNAs, forming oligomeric complexes before interacting with MAVs at the mitochondria (Bruns et al., 2014; Bruns and Horvath, 2015; Bruns and Horvath, 2014; Rodriguez et al., 2014). Since the oligomeric complexes formed in the cytoplasm between MDA5 and LGP2 is predicted to be solely between those two proteins and the RNA PAMP, it's possible that the E_{app} in the cytoplasm would be overall higher than the when the

complex forms at the mitochondria, where it is also interacting with the CARDS on MAVs. The interaction of the MAVs CARDS may shift the overall E_{app} value for the complex at the mitochondria to be slightly lower and therefore, the overall population displays a slightly lower predominant E_{app} peak.

9.3.2 Pixel level FRET analysis of the interaction of MAVs with RIG-I like receptors

The RIG-I like receptors exert their down-stream signaling effect through the adaptor protein MAVs. MAVs was first identified as a protein essential for NF- κ B and IRF3 activation by RNA viruses (Seth et al., 2005). It contains an N-terminal CARD-like domain and a C-terminal transmembrane domain which localizes MAVs to the mitochondrial membrane. Suppression of MAVs was shown to block interferon production and to increase viral replication. Over-expression of MAVS augmented interferon production and decreased viral replication. Deletion of the CARD domain completely inhibited the MAVs anti-viral effect implying that interaction through the CARD domain was essential for MAVs function (Seth et al., 2005).

The RIG-I like receptors, RIG-I and MDA5, exert their downstream effect directly through MAVs by seeding the formation of a MAVs filament through the interaction of RIG-I or MDA5 CARDS with the MAVs CARD. The formation of a MAVs filament initiates the down-stream signaling cascade which results in the production of interferon and other cytokines (Wu and Hur, 2015). Previous research using acceptor photo bleaching has shown that tagging RIG-I with the donor (GFP2) and MAVs with the acceptor (YFP) yields FRET in the MAVs regions of the cell (Baril et al., 2009). In addition, molecular modeling using a GFP₂ tagged MAVs and YFP tagged RIG-I (Badhu, Raicu and Frick, unpublished) indicated that the MAVS:RIG-I CARD interaction

should be stable and should yield at least one possible GFP₂ and YFP interaction with a distance close enough to achieve productive FRET (Table 9-1).

To study the interaction of RIG-I with MAVs, cells were transfected with plasmids

Distance between GFP ₂ -MAVS and YFP-RIG-I (Å)		
	Time elapsed	Time elapsed
Interaction	0 ns	21 ns
1	81.9	80.6
2	82.3	103.7
3	85.1	93.1
4	92.7	112.3
5	96.3	83.6
6	98.5	99.6
7	99	109.9
8	99.9	98

Table 9-1: Distances calculated using molecular dynamic simulations between fluorescent proteins in the RIG-I:MAVS complex. A theoretical GFP₂-MAVS and YFP-RIG-I oligomer was constructed using known crystal structures of RIG-I and MAVS. The distances between the GFP₂ and YFP protomers were calculated at time zero and after 21 nanoseconds of simulation.

expressing recombinant fluorescent GFP₂-MAVs and YFP-RIG-I. The cells were incubated for 20 hours to allow protein expression and then the cells co-transfected with GFP₂-MAVS and YFP-RIG-I were transfected with either poly I:C (0.36µg/ml) or vehicle. The cells were allowed to incubate for an additional 4 hours and then imaged using two wavelengths, 800 nm and 960nm. Although only the

800 nm wavelength was used for the calculation of E_{app}. The imaging procedure has been described previously in Methods 3.5. Elementary spectra were obtained from cells expressing either solely GFP₂-MAVS (Figure 9-6A) or YFP-RIG-I (Figure 9-6B). The elementary spectra were used to un-mix the

composite spectra from cells co-expressing GFP₂-MAVS and YFP-RIG-I (Figure 9-6C). The E_{app} was calculated by using the k^{DA} (Figure 9-7A) and k^{AD} (Figure 9-7B) 2D spatial distribution maps generated from un-mixing the composite spectra of cells co-expressing GFP₂-MAVS and YFP-RIG-I and excited at 800 nm. The E_{app} values were calculated for each pixel and used to create the 2D spatial distribution of E_{app} values (Figure 9-7C).

Sub-cellular MAVS regions were selected on the k^{DA} map in only those cells co-expressing GFP₂ and YFP signal. The selected MAVS regions were then used to create a mask to apply to the E_{app} map. Each MAVS ROI generated a histogram of E_{app} values within that ROI

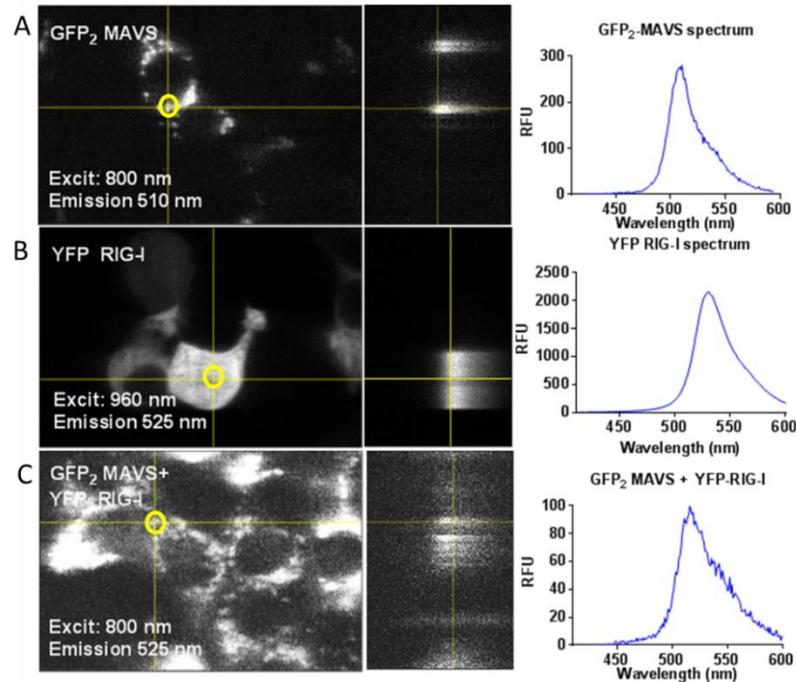


Figure 9-6: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP₂-MAVS and YFP-RIG-I. A) Cells expressing only GFP₂-MAVS and excited at 800 nm (two photon excitation). The peak 510 nm emission is shown as a GFP₂ 2D intensity map of the cells. The orthogonal view of the circled region within the GFP₂ intensity map and the calculated fluorescent spectrum is also shown. B) Cells expressing only YFP-RIG-I and excited at 960 nm (two photon excitation). The peak 525 nm emission is shown as a YFP 2D intensity map of the cells. The orthogonal view of the circled region within the YFP-RIG-I intensity map and the calculated fluorescent spectrum is also shown. C) Cells co-expressing GFP₂-MAVS and YFP-RIG-I and excited at 800 nm (two photon excitation). The orthogonal view of the circled region within the GFP₂-MAVS + YFP-RIG-I intensity map is shown, along with the fluorescent spectrum.

(Figure 9-7D and E). The peak E_{app} value for each histogram was selected and compiled into a meta-histogram containing all the peak E_{app} values from every MAVS region selected. The meta-

histogram shows the frequency at which any given peak E_{app} value appears within the circled MAVS ROI (Figure 9-7F).

The peak E_{app} values calculated were primarily less than 20% with the majority falling around 5% E_{app} . This is not unexpected given the molecular modeling which predicts a distance of about 80 Angstroms between GFP₂ and YFP in the MAVS:RIG-I complex. Interestingly, there was no shift in E_{app} with the addition of poly I:C., implying that the MAVS:RIG-I complex is unchanged with the addition of a PAMP. This is surprising given previous results between RIG-I:RIG-I which suggested that there may be the formation of a few large oligomeric complexes at

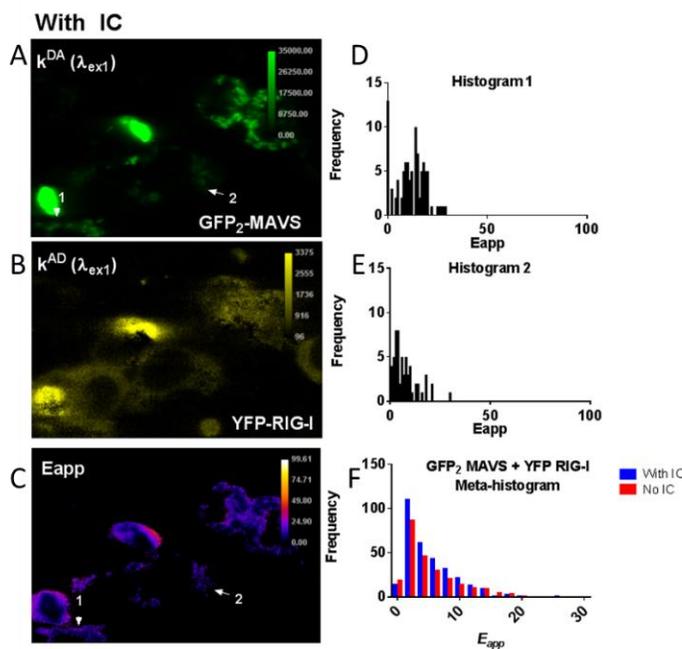


Figure 9-7: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP₂-MAVS and YFP-RIG-I. A) The 2D spatial fluorescence intensity map of donor in the presence of acceptor showing the distribution of GFP₂-MAVS in cells co-expressing YFP-RIG-I. B) The 2D spatial fluorescence intensity map of acceptor in the presence of donor showing the distribution of YFP-RIG-I in cells co-expressing GFP₂-MAVS. C) The calculate pixel level E_{app} map showing the distribution of FRET efficiencies in cells co-expressing GFP₂-MAVS and YFP-RIG-I. D) and E) Histograms of E_{app} values for the MAVS regions of interest indicated in the k^{DA} intensity map. F) Meta-histogram of selected peak FRET efficiencies for every MAVS region of interest selected from the k^{DA} intensity map.

the mitochondria in response to poly I:C addition (Figure 9-2A). However, since it's known that RIG-I signaling occurs even in the absence of PAMP (Figure 7-3F), it's possible that the RIG-

I:MAVS signaling complex is already formed with over-expression and does not shift with the addition of PAMP. Another possibility is that the shift in complex occurs between sub-units too far away to produce productive FRET. Since the molecular modeling between RIG-I and MAVS indicated that most of the interacting sub-units were too far away from each other to FRET, it's possible that any shift in the complex with addition of PAMP occurs between those sub-units and is not visible in this FRET assay with the fluorescent tags in their current positions on the proteins.

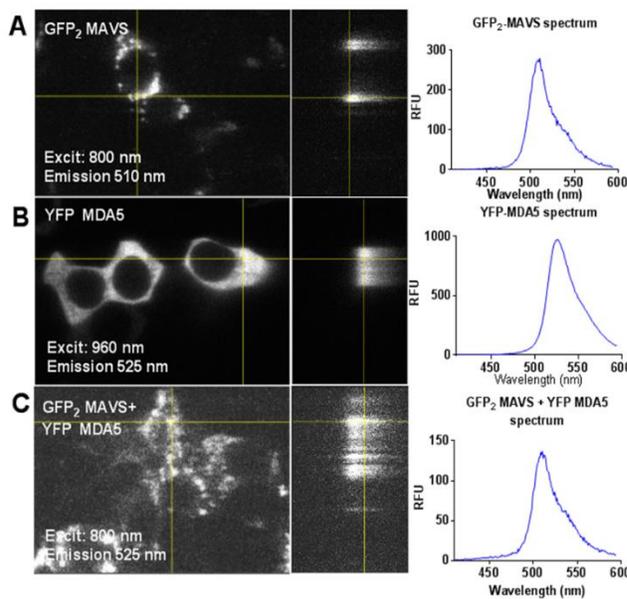


Figure 9-8: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP₂-MAVS and YFP-MDA5. A) Cells expressing only GFP₂-MAVS and excited at 800 nm (two photon excitation). The peak 510 nm emission is shown as a GFP₂ 2D intensity map of the cells. The orthogonal view of the circled region within the GFP₂ intensity map and the calculated fluorescent spectrum is also shown. B) Cells expressing only YFP-MDA5 and excited at 960 nm (two photon excitation). The peak 525 nm emission is shown as a YFP 2D intensity map of the cells. The orthogonal view of the circled region within the YFP-MDA5 intensity map and the calculated fluorescent spectrum is also shown. C) Cells co-expressing GFP₂-MAVS and YFP-MDA5 and excited at 800 nm (two photon excitation). The orthogonal view of the circled region within the GFP₂-MAVS + YFP-MDA5 intensity map is shown, along with the fluorescent spectrum.

In addition to RIG-I, MDA5 is known to interact with MAVS via tandem CARDs. To study the interaction of MDA5 with MAVS, plasmids encoding fluorescent recombinant GFP₂-MAVS

and YFP-MDA5 were co-transfected into HEK293T cells. The cells were allowed to incubate for 20 hours after which they were transfected with either poly I:C or vehicle and allowed to incubate for an additional 4 hours. After incubation, the cells were imaged at 800 nm and 960 nm excitations. The method was the same as described above for RIG-I. Elementary spectra

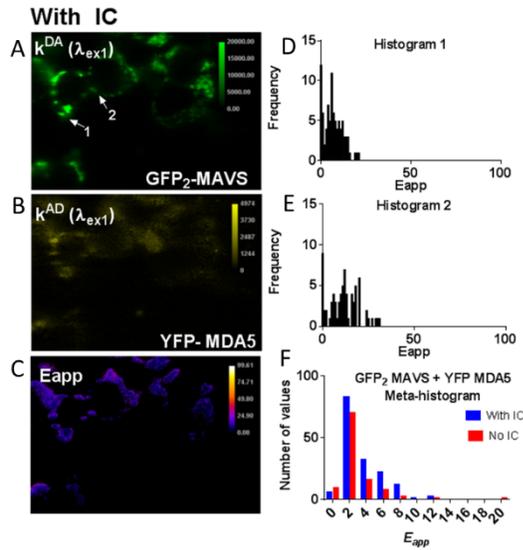


Figure 9-9: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP_2 -MAVs and YFP-MDA5. A) The 2D spatial fluorescence intensity map of donor in the presence of acceptor showing the distribution of GFP_2 -MAVs in cells co-expressing YFP-MDA5. B) The 2D spatial fluorescence intensity map of acceptor in the presence of donor showing the distribution of YFP-MDA5 in cells co-expressing GFP_2 -MAVs. C) The calculated pixel level E_{app} map showing the distribution of FRET efficiencies in cells co-expressing GFP_2 -MAVs and YFP-MDA5. D) and E) Histograms of E_{app} values for the MAVs regions of interest indicated in the k^{DA} intensity map. F) Meta-histogram of selected peak FRET efficiencies for every MAVS region of interest selected from the k^{DA} intensity map.

were obtained for cells solely expressing either GFP_2 -MAVs (Figure 9-8A) or YFP-MDA5 (Figure 9-8B). These elementary spectra were used to un-mix the composite spectra for the cells co-expressing GFP_2 -MAVs and YFP-MDA5 (Figure 9-8C). The E_{app} was calculated by using the k^{DA} (Figure 9-9A) and k^{AD} (Figure 9-9B) 2D spatial distribution maps generated from un-mixing the composite spectra of cells co-expressing GFP_2 -MAVs and YFP-RIG-I and excited at 800 nm. The E_{app} values were calculated for each pixel and used to create the 2D spatial distribution of E_{app} values (Figure 9-9C).

MAVS regions were selected on the k^{DA} map in only those cells co-expressing GFP₂ and YFP signal. The selected MAVS regions were then used to create a mask to apply to the E_{app} map. Each MAVS ROI generated a histogram of E_{app} values within that ROI (Figure 9-9D and E). The peak E_{app} value for each histogram was selected and compiled into a meta-histogram containing all the peak E_{app} values from every MAVS region selected. The meta-histogram shows the frequency at which any given peak E_{app} value appears within the circled MAVS ROI (Figure 9-9F).

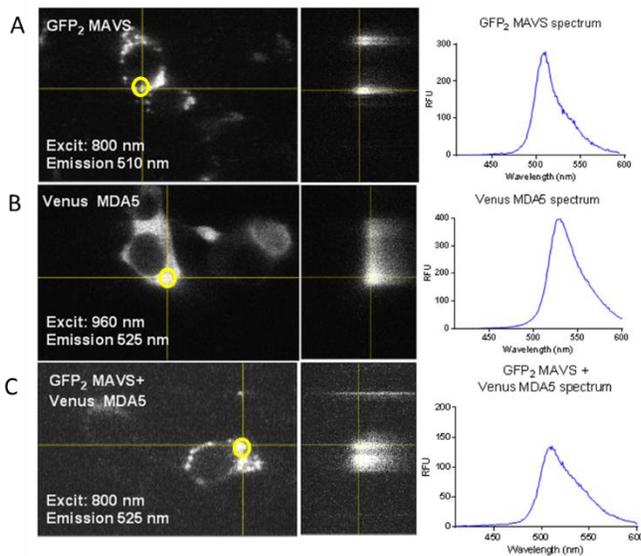


Figure 9-10: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP₂-MAVS and Venus-MDA5.

A) Cells expressing only GFP₂-MAVS and excited at 800 nm (two photon excitation). The peak 510 nm emission is shown as a GFP₂ 2D intensity map of the cells. The orthogonal view of the circled region within the GFP₂ intensity map and the calculated fluorescent spectrum is also shown. B) Cells expressing only Venus-MDA5 and excited at 960 nm (two photon excitation). The peak 525 nm emission is shown as a Venus 2D intensity map of the cells. The orthogonal view of the circled region within the Venus-MDA5 intensity map and the calculated fluorescent spectrum is also shown. C) Cells co-expressing GFP₂-MAVS and Venus-MDA5 and excited at 800 nm (two photon excitation). The orthogonal view of the circled region within the GFP₂-MAVS + Venus-MDA5 intensity map is shown, along with the fluorescent spectrum.

The maximum peak E_{app} calculated for the interaction of GFP₂-MAVS and YFP-MDA5 was slightly over 20%. But the majority of the E_{apps} fell below 10%. It is thought that the MDA5

oligomer is larger than the RIG-I oligomer (Lässig and Hopfner, 2017) and the overall reduced E_{apps} for the MDA5:MAVS interaction compared to the RIG-I:MAVS interaction could be due to shielding from a larger MDA5:MDA5 complex which prevents closer association of the fluorescent tags on MDA5 and MAVS. Similar to the results for the interaction of RIG-I and MAVS, there did not appear to be a shift in interaction upon addition of poly I:C. It was shown that MDA5 over-expression (similar to RIG-I) causes interferon signaling (Figure 7-3F) presumably through MAVS. Therefore, it's possible that the MAVS:MDA5 oligomer may already be in its optimal signaling configuration before PAMP RNA is added. However, it is also possible that the tags are located in regions where the shift in the MDA5:MAVS complex is not readily visible. FRET is never able to prove there is no interaction (or no change in interaction), it can only prove that there is an interaction because multiple other factors affect the success of FRET (discussed in Literature Review and Methods) besides just the interaction of the proteins to which fluorescent tags are attached.

An additional experiment looking at the interaction of MAVs and MDA5 was also performed in the same manner as the previous experiment, except for two changes. The tag on MDA5 was changed from YFP to Venus. Venus is a YFP variant which is more stable and matures faster than YFP but otherwise has the same emission and excitation spectrum as YFP. The additional change was the length of the linker connecting MDA5 to the fluorescent protein. In the previous experiment, the link was only two amino acids long, so MDA5 and the fluorescent protein were very close to each other. The connection between the Venus fluorescent protein and MDA5 is 15 amino acids long. Therefore, there is much more space between the

fluorescent tag and the protein of interest. In contrast, the linker between the GFP₂ and MAVS is only 2 amino acids long.

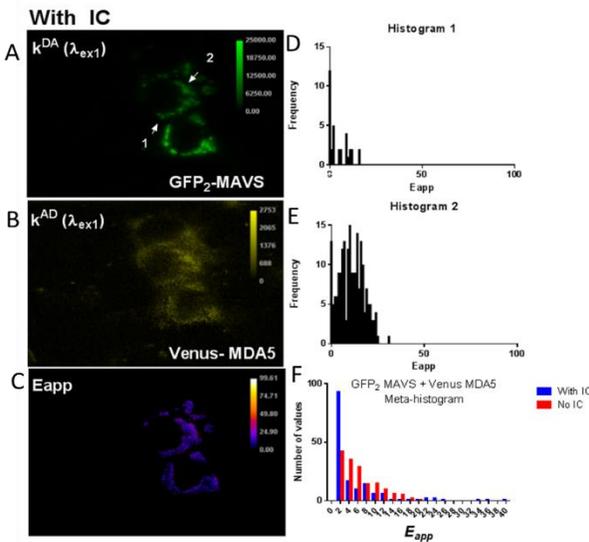


Figure 9-11: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP₂-MAVS and Venus-MDA5.

A) The 2D spatial fluorescence intensity map of donor in the presence of acceptor showing the distribution of GFP₂-MAVS in cells co-expressing Venus-MDA5. B) The 2D spatial fluorescence intensity map of acceptor in the presence of donor showing the distribution of Venus-MDA5 in cells co-expressing GFP₂-MAVS. C) The calculated pixel level E_{app} map showing the distribution of FRET efficiencies in cells co-expressing GFP₂-MAVS and Venus-MDA5. D) and E) Histograms of E_{app} values for the MAVS regions of interest indicated in the k^{DA} intensity map. F) Meta-histogram of selected peak FRET efficiencies for every MAVS region of interest selected from the k^{DA} intensity map.

To collect the elementary spectra from GFP₂ and Venus, HEK293T cells expressing solely GFP₂-MAVS or Venus-MDA5 were imaged at 800 nm excitation and 960 nm excitation. This created the GFP₂ 800 nm elementary spectrum (Figure 9-10A) and the Venus 960 nm elementary spectrum (Figure 9-10B). These elementary spectrum were used to unmix the composite spectra containing both the GFP₂-MAVS and the Venus-MDA5 excited at 800 nm (Figure 9-10 C). This generated the k^{DA} (Figure 9-11A) and k^{AD} (Figure 9-11B) maps. GFP₂-MAVS regions were selected and used to create a mask which was applied to the E_{app} map (Figure 9-11C). Each selected MAVS region in the E_{app} map generated a unique histogram of E_{app} values

(Figure 9-11D and E). The Peaks from these histograms were collected in a meta-histogram (Figure 9-11F).

The meta-histogram of GFP₂-MAVS and Venus-MDA5 shows overall higher E_{app} values than the previous experiment performed with GFP₂-MAVS and YFP-MDA5. This could be due to the stability of the acceptor (Venus). It could also be due to the difference in linker length connecting Venus to MDA5. It's possible that the longer linker length allows for better interaction of the tags connecting MDA5 and MAVS. It is also possible that the increased

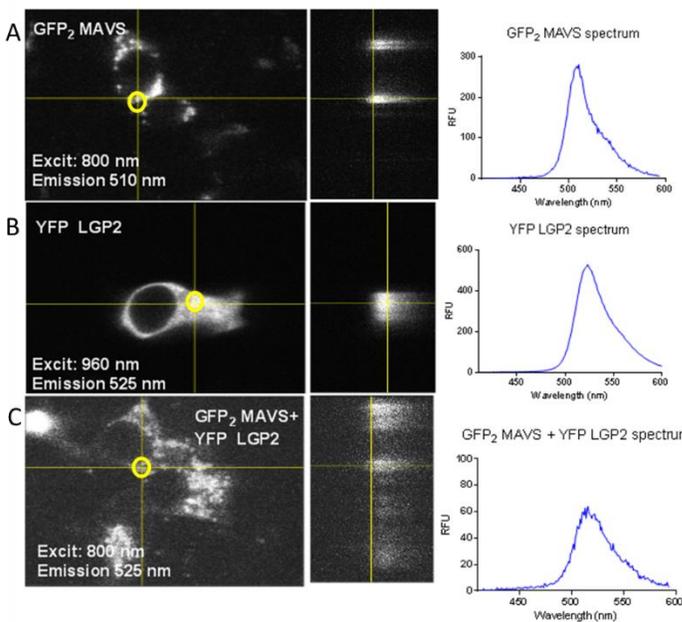


Figure 9-12: Dual excitation fluorescence micro-spectroscopy analysis of cells co-expressing GFP₂-MAVS and YFP-LGP2. A) Cells expressing only GFP₂-MAVS and excited at 800 nm (two photon excitation). The peak 510 nm emission is shown as a GFP₂ 2D intensity map of the cells. The orthogonal view of the circled region within the GFP₂ intensity map and the calculated fluorescent spectrum is also shown. B) Cells expressing only YFP-LGP2 and excited at 960 nm (two photon excitation). The peak 525 nm emission is shown as a YFP 2D intensity map of the cells. The orthogonal view of the circled region within the YFP-LGP2 intensity map and the calculated fluorescent spectrum is also shown. C) Cells co-expressing GFP₂-MAVS and YFP-LGP2 and excited at 800 nm (two photon excitation). The orthogonal view of the circled region within the GFP₂-MAVS + YFP-LGP2 intensity map is shown, along with the fluorescent spectrum.

stability of Venus allowed for more stably expressed acceptors. Since FRET is dependent on the number of acceptors present, if the YFP acceptor is not stable, then less acceptors would be present to receive energy from the donors. It is possible that the GFP₂ and Venus is a better FRET pair. In addition, a small population of ROIs expressing higher FRET efficiencies are visible in the cells treated with poly I:C. These ROIs showed E_{apps} as high as 40%. It's possible that the longer linker length is increasing the sensitivity of the assay by changing how the tags interact and making a shift in complex due to poly I:C which was previously not detected, visible in this FRET assay set up (Figure 9-11).

LGP2 is the third RLR which is not known to directly interact with MAVS due to its lack of CARDs. However, it does interact with MDA5 which in turn interacts with MAVS, and it is known that simply over-expressing LGP2 somewhat increases interferon signaling (Figure 7-3F). Since LGP2 cannot signal directly, this increase in interferon signaling must be through the interaction with endogenous MDA5. To test the interaction of LGP2 and MAVS, cells were transfected with plasmids co-expressing GFP₂-MAVS and YFP-LGP2 and they were treated in the same manner as was described for RIG-I:MAVS and MDA5:MAVS experiments above. The elementary spectra for cells solely expressing GFP₂-MAVS (Figure 9-12A) or YFP-LGP2 (Figure 9-12B). The elementary spectra were used to unmix the composite spectra for those cells co-expressing GFP₂-MAVS and YFP-LGP2 (Figure 9-12C). This unmixing generated k^{DA} (Figure 9-13A) and k^{AD} (Figure 9-13B) 2D spatial intensity maps which were used to calculate pixel level E_{apps} and to generate an E_{app} map (Figure 9-13C) for each field of view. MAVS region ROIs were selected on the k^{DA} map for only those cells co-expressing GFP₂ and YFP signal. This MAVS ROI mask was then applied to the E_{app} map and E_{app} histograms generated for each sub-cellular region (Figure 9-13D and E). The

peak values from each ROI within the E_{app} regions selected were compiled into a meta-histogram (Figure 9-13F).

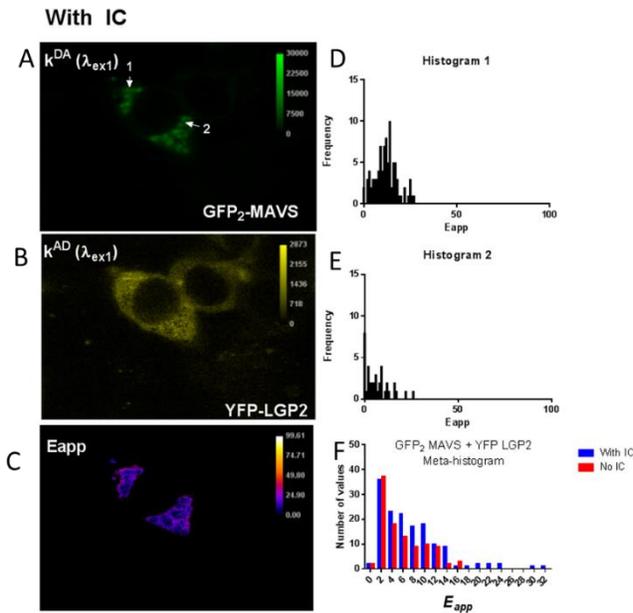


Figure 9-13: Pixel level FRET (E_{app}) calculation and meta-histogram analysis for cells co-expressing GFP₂-MAVS and YFP-LGP2. A) The 2D spatial fluorescence intensity map of donor in the presence of acceptor showing the distribution of GFP₂-MAVS in cells co-expressing YFP-LGP2. B) The 2D spatial fluorescence intensity map of acceptor in the presence of donor showing the distribution of YFP-LGP2 in cells co-expressing GFP₂-MAVS. C) The calculated pixel level E_{app} map showing the distribution of FRET efficiencies in cells co-expressing GFP₂-MAVS and YFP-LGP2. D) and E) Histograms of E_{app} values for the MAVS regions of interest indicated in the k^{DA} intensity map. F) Meta-histogram of selected peak FRET efficiencies for every MAVS region of interest selected from the k^{DA} intensity map.

Surprisingly, the combination of GFP₂-MAVS and YFP-LGP2 yielded high FRET efficiencies of up to 20% for both cells treated and not treated with poly I:C. Given that LGP2 should not interact with MAVS directly, it is unexpected that there be such high FRET efficiencies. It's possible that the interaction of endogenous MDA5 with YFP-LGP2 is positioning LGP2 such that the YFP interacts with the GFP₂ on MAVS. In addition, there is the development of a small population of ROIs with higher FRET efficiencies with the addition of poly I:C. If LGP2 is indeed being recruited by endogenous MDA5 to the MAVS complex, then it makes sense that the addition of poly I:C would increase/change this interaction.

9.4 Significance of the interactions of the RIG-I like receptors in response to poly I:C

The RIG-I like receptors have previously been shown to form oligomers with the use of electron microscopy (Bruns et al., 2014; Bruns and Horvath, 2015; Bruns and Horvath, 2012; Bruns et al., 2013; Bruns and Horvath, 2014). To look at the interactions of RLRs in living cells, ectopically expressed fluorescent fusion proteins were used in conjunction with fully quantitative FRET imaging and analysis (Corby et al., 2017; Stoneman et al., 2017). When whole cells were selected as ROIs, the self-interaction of both RIG-I and MDA5 was very low, whereas the FRET between LGP2:LGP2 self-interactions was noticeable, even in the absence of RNA PAMP (Figure 7-4K). This suggests that only LGP2 forms self-oligomers in the cytoplasm in the absence of PAMPs.

Noteworthy self-interactions between RIG-I:RIG-I and MDA5:MDA5 in the absence of RNA PAMP were visible when the region of interest was switched from the whole-cell to the mitochondria. But the FRET between RIG-I:RIG-I was smaller than the self-interaction between MDA5 (Figure 9-2 B and C). This is consistent with other research which states that RIG-I forms smaller, shorter oligomers and is responsible for detecting shorter RNA PAMPs with different end modifications versus MDA5 which is implicated in sensing longer chain RNA PAMPs (del Toro Duany et al., 2015; Peisley et al., 2014; Peisley et al., 2012). However when poly I:C was added, the RIG-I:RIG-I interaction included the development of a few very large E_{apps} as large as 80-90% whereas the MDA5:MDA5 interaction included a much larger population shift to larger E_{apps} but they only reached a max of 50%. The differences in E_{apps} at the mitochondria could be

indicative of the formation of different oligomers between RIG-I:RIG-I and MDA5:MDA5 at the mitochondria. The differences in MDA5 and RIG-I oligomers could mean that the signal they transmit through MAVS is specific to the type of PAMP which they recognize.

Despite the differences between the RIG-I:RIG-I and MDA5:MDA5 FRET at the mitochondria in response to poly I:C (Figure 9-2 B and C) there do not appear to be large differences caused by the addition of PAMP between RIG-I:MAVS or MDA5:MAVS. Although the overall populations of FRET efficiencies was lower between MDA5: MAVS as compared with RIG-I:MAVS (Figure 9-7F and 9-9F), there was no visible shift with the addition of poly I:C. This may mean that the complex formed between MAVS and either RIG-I or MDA5 is similar with and without the addition of PAMP. Meaning that over-expression alone induces the formation of an oligomer which is the same as the oligomer induced in the presence of RNA PAMP. However, it could also mean that the shift in oligomer is not visible with this FRET assay in the current conditions. An additional experiment performed with MAVS and MDA5 where the fluorescent tag on MDA5, as well as the linker between the fluorescent tag and MDA5 were switched gives credence to the latter explanation. In this experiment, the FRET between MAVS and MDA5 was higher overall and a shift in population to a set of ROIs expressing higher E_{apps} (as high as 40%) was apparent in response to poly I:C transfection (Figure 9-11F). Therefore, it's possible that there is a shift in oligomer occurring with addition of PAMP and that the linker length or fluorescent tags need to be switched to allow that shift to become visible in this FRET assay.

The self-interaction of LGP2:LGP2 at the mitochondria in response to poly I:C contained overall larger E_{apps} than the self-interaction of either MDA5 or RIG-I with E_{apps} both with and without poly I:C reaching as high as 30% (Figure 9-2D). Interestingly, the FRET for the LGP2 self-interaction was highest in the absence of poly I:C and then decreased when poly I:C was added. This possibly indicates that the presence of a PAMP encourages LGP2 interaction with MDA5 to aid in the PAMP recognition, thus decreasing the self LGP2:LGP2 interaction.

The largest FRET efficiencies seen in this Q-MSI study were seen between LGP2 and MDA5. This interaction has been studied previously and it is known that LGP2 and MDA5 interact on RNA PAMPs (Bruns et al., 2014; Bruns and Horvath, 2012; Uchikawa et al., 2016; Wu et al., 2014). When YFP-LGP2 was co-expressed with CFP-MDA5 an effect was seen. In the absence of poly I:C, there was a small population of E_{apps} as high as 10%, but with the addition of poly I:C, the E_{app} values shifted above 20%, with some reaching as high as 80-90% FRET efficiency (Figure 9-3D). When the tags were switched, the overall FRET efficiency decreased (Figure 9-3A and B) but there was still an overall shift to higher E_{app} values with the addition of poly I:C. In addition to the shift in E_{app} values, the molar fraction of acceptor shifted when the tags were switched between LGP2 and MDA5. When the acceptor was tagged to LGP2 and the donor attached to MDA5, a shift towards higher molar fractions of acceptor was visible. In contrast, when the donor was LGP2 and the acceptor was MDA5, the molar fraction of acceptor shifted lower. All this data combined suggests that LGP2 does not just prime RNA filaments for MDA5 interaction, but rather is the dominant protomer in a larger MDA5: LGP2 oligomer. Taken together, this data reinforces the idea that addition of an RNA PAMP leads to

accumulation of LGP2 at the mitochondria through interaction with either endogenous or exogenous MDA5.

Chapter 10 The two excitation peaks of the wild-type GFP protein and the development of a novel GFP variant.

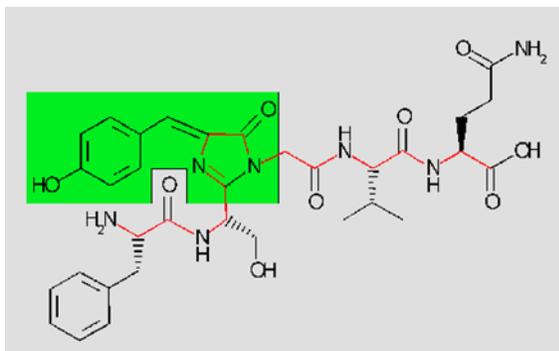


Figure 10-1: The GFP chromophore is 4-(*p*-hydroxybenzylidene)imidazolidin.

The development of the green fluorescent protein, from its isolation in jellyfish to its widespread use today as a fluorescent tag and indicator, involved many groups and over 40 years of work. The work was monumental and culminated in the 2008 Nobel Prize in Chemistry.

The green fluorescent protein (GFP) was first discovered by Shimomura et al as a chemiluminescent protein found in *Aequorea* jellyfish (Shimomura et al., 1962). Soon after, the GFP protein was purified and crystalized and its absorption spectrum and quantum yield measured (Morise et al., 1974). Next it was proposed and later confirmed that the chromophore is 4-(*p*-hydroxybenzylidene)imidazolidin and that the chromophore is attached to the backbone of the protein through the 1 and 2 positions of the ring (Figure 10-1) (Shimomura, 1979). Over the next 20 years, several GFP variants possessing mutations which shifted the excitation and/or emission spectra, increased photo-stability, increased efficient folding at 37°C and increased brightness were developed by Tsien and other groups (Heim et al., 1995; Shaner et al., 2005; Tsien, 1998).

The green fluorescent protein has a spectrum with two excitation peaks (Figure 3-1 B and C). One excitation peak is at about 400 nm (or 800 nm two-photon excitation) and the second excitation peak is at 480 nm (or 960 nm two photon excitation). Shifting the excitation spectra of GFP to favor one excitation over another may be desirable in some cases and is achieved through altering the local environment of the GFP chromophore. If the protonated form is stabilized then the 400 nm excitation is dominant. If the deprotonated anion is stabilized, then the 480 nm excitation is preferred. There are several classes of GFP proteins which can be separated based on their excitation and emission properties. GFP class 1 variants includes wild-type GFP. Class 1 GFP proteins have retained the two excitation maxima (Battistutta et al., 2000). Class 2 GFP variants have a S65T mutation in the chromophore which stabilizes the phenolate anion of the chromophore. This mutation causes the excitation spectra of class 2 variants to shift towards a predominantly 480 nm excitation peak. Class 3 GFP molecules have mutations near the chromophore (such as S202F or T203I) which preference the neutral phenol form of the chromophore. The neutral form shifts toward the 400 nm excitation peak (Zacharias and Tsien, 2006).

Selection of the appropriate fluorescent proteins with the best characteristics for the particular FRET experiment is of the utmost importance. When using pixel level FRET spectrometry to analyze protein:protein interactions there needs to be a sufficient difference between donor and acceptor excitation spectra such that the acceptor is not excited at the donor excitation wavelength. If the acceptor is excited at the donor excitation wavelength, then an additional correction for direct excitation is necessary. This correction employs the use of a second excitation wavelength and the time necessary to switch between two excitation

wavelengths is enough to allow molecular diffusion such that the results can no longer be analyzed at the pixel level. Instead, averages over larger regions of interest must be used. Although averages can still be informative (King et al., 2016), pixel level allows for the acquisition of a much greater amount of information over smaller areas and has been used to determine more information about the specific quaternary arrangement of the protein subunits (see Methods 3.5 for more detail) (Corby et al., 2017; Stoneman et al., 2017).

The fluorescent proteins GFP₂ and YFP or GFP₂ and Venus make ideal FRET pairs for pixel level FRET spectrometry analysis because the GFP₂ has an excitation peak at 800 nm (two photon excitation) and both the YFP and Venus fluorescent proteins do not excite at 800 nm (Figure 3-1 B and C). GFP₂ is very similar in sequence to Clontech's eGFP protein except that GFP₂ converts the S65T mutation, present in the eGFP protein, back to the wild type S65. This has the effect of reverting the GFP₂ back to the wild-type GFP excitation spectra. In addition, GFP₂ incorporates the non-dimerizing mutation, A206K, which decreases the tendency of fluorescent proteins to dimerize. See Figure 3-2 for an alignment of the sequences of wild-type

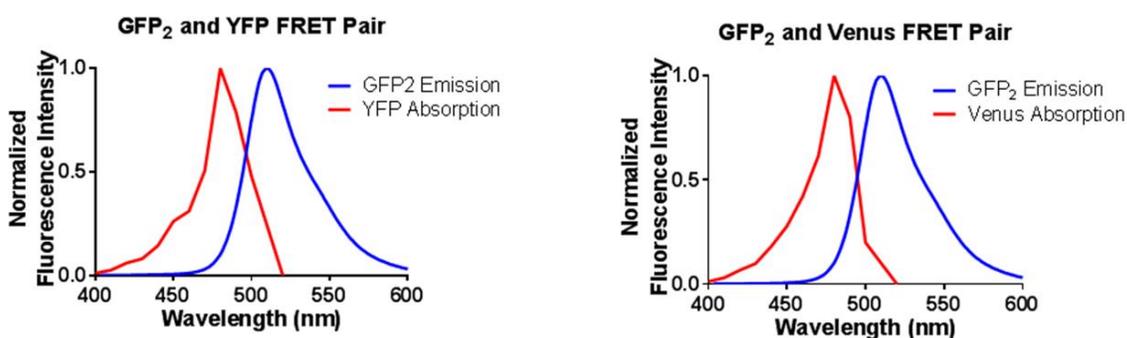


Figure 10-2: Emission of GFP₂ overlaps YFP and Venus absorption spectra. A) Overlay of GFP₂ emission spectra and YFP absorption spectra. B) Overlay of GFP₂ emission spectra and Venus absorption spectra.

GFP, Clontech eGFP, and GFP₂. Both YFP and Venus excite optimally at 960 nm (two photon excitation) and are excited by GFP₂ emission (Figure 10-1) allowing for efficient energy transfer between the fluorescent molecules.

10.1 Protein concentration calculations and the gamma ratio

An additional desirable component of FRET spectrometry is the ability to measure protein concentrations in cells. Although this is not necessary for a pixel level FRET analysis, the calculation of concentration can allow for an analysis of how FRET changes with protein expression (Mishra et al., 2016). In addition, it can provide information on shifts in protein expression or localization (Figure 7-10B and C). To calculate protein concentration in a solution for which the protein concentration cannot be directly measured (i.e. fluorescent proteins expressed in cells) the total fluorescence of the donor (F^D) and the total fluorescence of the acceptor (F^A) must be calculated. To calculate F^A and F^D in cells co-expressing both donor and acceptor molecules, two excitation wavelengths must be used. The calculation for F^A and F^D was previously discussed in detail in Methods 3.5. The F^A and F^D values can then be compared against a standard curves generated for the acceptor and donor respectively and discussed in Methods 3.5.6.

In order to calculate the F^A and F^D values from samples co-expressing donor and acceptor, the gamma ratios of donor and acceptor at each excitation wavelength must be determined. The gamma for the donor is calculated by taking the ratio of the donor emission at each wavelength (Figure 3-4A). The gamma ratio for the acceptor is calculated by taking the ratio of the acceptor emission at each wavelength (Figure 3-4B). The gamma ratio is necessary

when converting between the emission of the donor (or acceptor) at one wavelength to the emission of the donor (or acceptor) at the second wavelength. See Methods 3.5.5 for a more detailed description of the calculation and use of the gamma ratios in the calculation of F^D and F^A . To obtain a gamma ratio for GFP₂ for use in calculation of GFP₂ concentrations, HEK 293T cells were transfected solely with plasmids expressing GFP₂. The cells were allowed to express the GFP₂ protein for 24 hours and then imaged at 100X using 800 nm and 960 nm two-photon excitations. The resulting images were un-mixed (Methods 3.5) to remove any background noise. Then the intensity of the GFP₂ expression at one wavelength was divided by the intensity

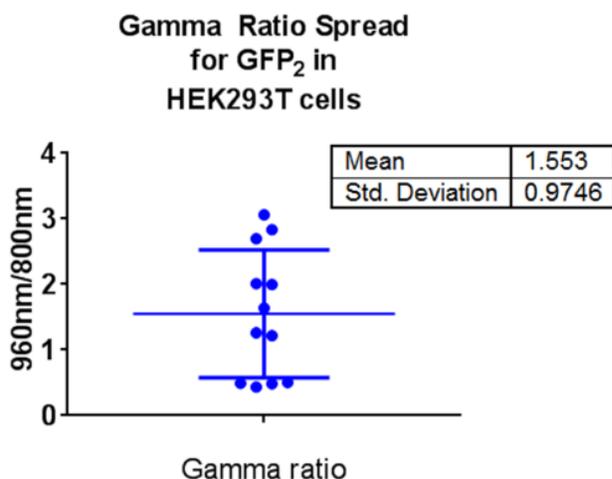


Figure 10-3: GFP₂ expressed in HEK293T cells generates a wide range of gamma ratio values. A plasmid encoding GFP₂ was expressed in HEK293T cells for 24 hours. Images were taken at either the 800 nm or 960 nm excitation wavelengths and unmixed to remove background signal. The ratio was taken of the mean intensity emission at 960nm/800nm for several fields of view. The ratio of 960nm/800nm is termed the gamma ratio and was plotted. The average gamma ratio was 1.553 with a standard deviation of 0.97.

at a second wavelength to obtain the gamma ratio using equations 3.9 and 3.10

shown again below.:

$$F^A(\lambda_{ex,1}) = \frac{\Gamma^{\lambda_{ex,1,A}}}{\Gamma^{\lambda_{ex,2,A}}} \cdot F^A(\lambda_{ex,2})$$

$$F^D(\lambda_{ex,2}) = \frac{\Gamma^{\lambda_{ex,2,D}}}{\Gamma^{\lambda_{ex,1,D}}} \cdot F^D(\lambda_{ex,1})$$

Where $F^A(\lambda_{ex,1})$ is the fluorescence intensity of the acceptor at wavelength 1 and $F^A(\lambda_{ex,2})$ is the fluorescence intensity of the acceptor at wavelength 2. Similarly, $F^D(\lambda_{ex,1})$ is the fluorescence intensity of the donor at wavelength 1 and $F^D(\lambda_{ex,2})$ is the fluorescence intensity of the donor at wavelength 2.

Thus the intensity of the acceptor at wavelength 1 and wavelength 2 is related via the acceptor

gamma ratio $\frac{\Gamma^{\lambda_{ex,1,A}}}{\Gamma^{\lambda_{ex,2,A}}}$. The intensity of the donor at wavelength 1 and wavelength 2 is related

via the donor gamma ratio $\frac{\Gamma^{\lambda_{ex,2,D}}}{\Gamma^{\lambda_{ex,1,D}}}$.

The gamma ratio calculated for cells expressing GFP₂ was indeterminate with an average of 1.5 and a standard deviation of almost 1 (Figure 10-2). The variation of gamma ratio seen in cells solely expressing GFP₂ is evidence that the GFP₂ molecule is existing in two ionization

states. One where the chromophore contains a neutral phenol and prefers the 800 nm excitation and one where the chromophore contains a phenolate anion and prefers the 960 nm excitation. The variation in gamma ratio shows that the GFP₂ molecules preferring the 800 nm excitation versus those preferring the 960 nm excitation is not stable in cells.

To test the factors which contribute to variation in excitation maxima in GFP₂, a series of experiments were conducted with purified eGFP and GFP₂. There are only two mutations that are different between eGFP and GFP₂. One of those mutations is the A206K mutation which affects dimerization and does not affect the GFP chromophore. The other mutation is the S65T mutation. The S65T mutation is present in eGFP and stabilizes the phenolate anion of the chromophore. This causes eGFP to have a much higher 960 nm (two photon) excitation as compared to GFP₂. The GFP₂ variant returns to the wild-type S65 and therefore retains the wild-type GFP's ability to be excited at both 800 nm and 960 nm (two photon excitation).

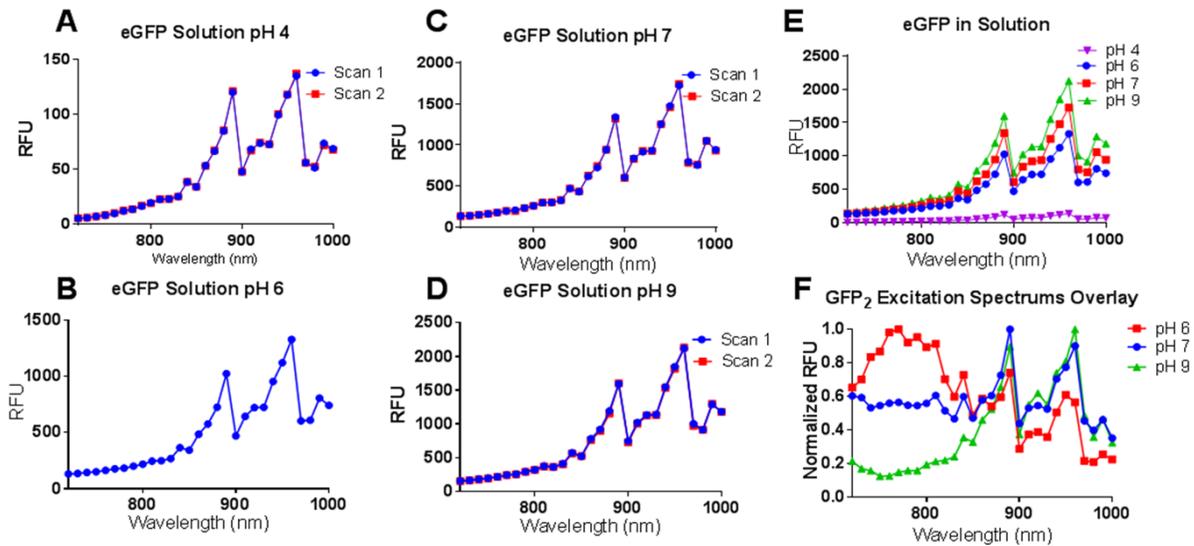


Figure 10-4: eGFP does not change much with pH but GFP₂ shows a dramatic change in excitation spectrum with pH change. A) eGFP excitation spectrum at pH 4 scanned twice successively. B) eGFP excitation spectrum at pH 6. C) eGFP excitation spectrum at pH 7 scanned twice successively. D) eGFP excitation spectrum at pH 9 scanned twice successively. E) Overlay of the GFP₂ excitation spectra in varying pH. F) Normalized excitation spectra for GFP₂ at varying pH.

The first experiment looked at the variation of the excitation spectrum of eGFP in different pHs. The pH variation should affect the protonation state of the chromophore and it would be expected that the excitation spectrum would shift when the proteins were exposed to different pH buffers. Purified eGFP was diluted to a final concentration of 10 μ M in a sodium phosphate buffer with a pH of either 4, 6, 7, or 9. Next, 200 μ L of each solution was imaged in a well plate. An excitation scan was performed on wells containing eGFP at each pH whereby the protein is excited at each excitation wavelength between 720 nm to 1000 nm and the intensity of emission at each wavelength is graphed as a function of the excitation wavelength (Figure 10-4 A-D). The only change in excitation spectra due to pH was a change in intensity. The overall emission intensity increased with pH and pH 4 was significantly less intense than pH 6, 7, or 9 (Figure 10-4E). There was no change in relative excitation peak intensities with pH (Figure

10-4E). This is most likely due to the S65T mutation which stabilizes the phenolate anion on the chromophore and maintains the protein in a consistent ionization state. In addition, successive excitation scans were taken of eGFP suspended in either pH 4, 7, or 9 buffer and no shift in excitation spectrum was observed after a repeat scan was taken in the same region of the plate. If the same proteins were being successively imaged, it would be expected that either a bleaching effect or an excitation spectrum shift due to photo-ionization of the chromophore would be apparent on the second excitation scan. There was no difference between the first and second scans at any pH measured and this is most likely due to the 200 μ L well volume which allows for continual renewal of the eGFP proteins measured via molecular diffusion.

To test the effect of pH on GFP₂, the same sodium phosphate buffers were used to dilute purified GFP₂ to a final concentration of 10 μ M. The solutions were added to a well plate and the excitation spectra were measured for each pH in a method identical to that described above for eGFP. The intensity for GFP₂ at pH 4 did not go above noise and therefore was not included in the final analysis. The effect of pH on the excitation spectra for GFP₂ was remarkable. The peak emission excitation intensity for GFP₂ in pH 6 was 800 nm whereas the 960 nm excitation was significantly lower. Increasing the pH to 7 shifted the spectrum so that the 800 nm excitation was lower and the 960 nm excitation was higher. When the GFP₂ protein was suspended in pH 9 buffer, the 800 nm excitation was practically non-existent and the 960 nm excitation was dominant (Figure 10-3F). This shift in excitation spectra with pH shows how the protonation state of the chromophore of GFP₂ affects the emission spectra. When the phenol is protonated and in a neutral state, the 800 nm excitation is dominant. But as the pH is increased and more and more of the GFP₂ molecules are de-protonated, the 960 nm excitation

increases and the 800 nm excitation decreases. Then at pH 9, most of the molecules are deprotonated and the 960 nm excitation dominates, as is seen in the eGFP variant containing the S65T mutation.

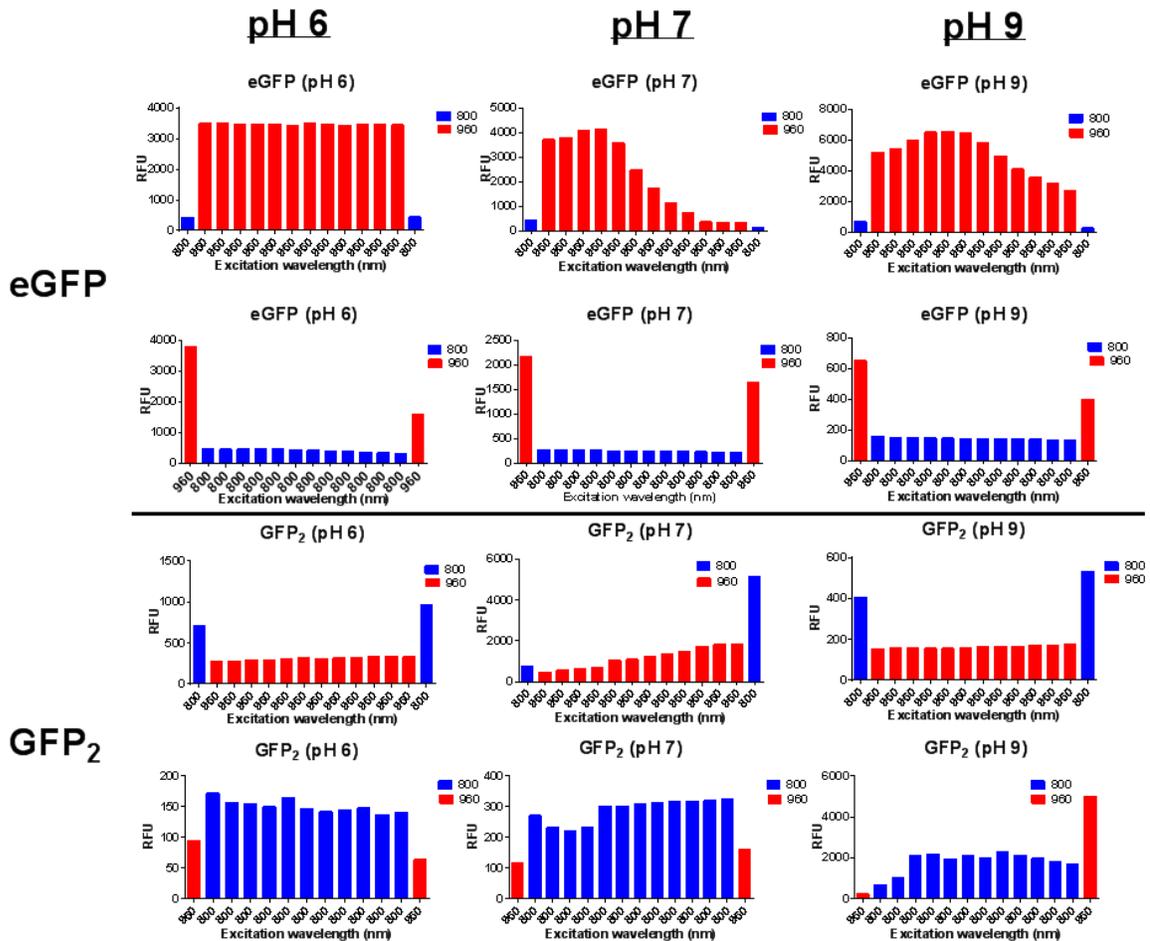


Figure 10-5: Effect of multiple laser excitations on eGFP and GFP₂ at varying pH. A) Effect of successive 960 nm excitations on the intensity of the 800 nm excitation peak of eGFP at pH 6,7, and 9. B) Effect of successive 800 nm excitations on the intensity of the 960 nm excitation peak of eGFP at pH 6,7, and 9. C) Effect of successive 960 nm excitations on the intensity of the 800 nm excitation peak of GFP₂ at pH 6,7, and 9. D) Effect of successive 800 nm excitations on the intensity of the 960 nm excitation peak of GFP₂ at pH 6,7, and 9.

Another way that the GFP chromophore may shift its protonation state is through photo-ionization. To test the photo-ionizability of eGFP and GFP₂ and to see if pH affects the ionization of the chromophore, 10 μL of 10 μM solutions of eGFP and GFP₂ at either pH 6, 7, or 9 were dotted on glass slides. They were covered with a coverslip and allowed to dry for 15 minutes prior to imaging. By imaging the proteins on a slide as opposed to in solution, it is more likely that the same molecules are being successively imaged since most of the fluorescent molecules should be attracted to the glass slide surface and only a thin layer of molecules is spread across the slide under the coverslip. The slides were imaged first at 800 nm to establish a baseline intensity, then successively excited at 960 nm in the same region of the slide. Then imaged again at 800 nm to determine if the intensity of the emission of the 800 nm excitation had changed. This procedure was then altered so that the first excitation was 960 nm and then several successive 800 nm excitations were performed on the same region of the slide. Then the region was again measured at 960 nm to determine if the intensity of emission at the 960 nm excitation had changed by successively exciting at 800 nm.

For eGFP, it is apparent that the molecule excites much better at 960 nm than at 800 nm at all pHs measured, which reinforces the measurements already taken in free moving solution. Successive laser excitations appeared to only exert a small bleaching effect. There was no photo-bleaching effect apparent at pH 6, but pH 7 and pH 9 both exhibited successive decreases in emission intensity when successive 960 nm excitations were used in the same region of the slide. When successive 800 nm excitations were used and then the 960 nm excitation was once again tested, the 960 nm emission intensity at 960 nm excitation was lowered, again indicating a photo-bleaching effect (Figure 10-5).

When GFP₂ was excited at 960 nm and then successively excited at 800 nm, the intensity of the 800 nm emissions remained relatively constant, and the re-measurement at 960 nm excitation yielded an emission intensity which was roughly the same as the original measurement for pH 6 and 7. But when GFP₂ was subjected to repeated 800 nm excitations at pH 9, the emission intensity when the same region was re-measured at 960 nm excitation was increased. This implies that only once the pH was high enough to favor the de-protonation of the chromophore, was the 800 nm excitation able to deprotonate it and thus increase the emission at the 960 nm excitation. Interestingly, when the GFP₂ protein was successively excited at 960 nm, the emission intensity of the 800 nm excitation increased at every pH. However, this increase was small for both pH 6 and pH 9 possibly indicating that the 960 nm excitation is not strong enough to change the protonation state and instead the time between the first and last 800 nm excitations was sufficient to allow recovery of the some of the GFP₂ molecules to a protonated state (Figure 10-5). The measurements at pH 7 show a molecule which is increasing in emission intensity after each successive excitation. It is possible that the slide was not fully dried and new molecules were diffusing into the region being measured over the course of the measurements. However, an alternative explanation would be that somehow the 960 nm excitation is encouraging the protonation of the GFP₂ chromophore and therefore shifting towards a preference for the 800 nm excitation. Whatever the molecular explanation may be, it is apparent that the intensity of the GFP₂ emissions for either the 800 nm excitation or the 960 nm excitation is not consistent between different pHs and that laser light may be able to ionize the chromophore and further shift the equilibrium between emission intensities.

10.2 Development of a novel GFP variant

Over time, many GFP mutations have been made which expand the GFP color tool box to almost every color imaginable. In addition, more subtle mutations have been made which shift the excitation or emission spectra of GFP slightly. One example of an excitation spectrum shifting mutation is the S65T mutation which shifts the excitation peak maxima for the eGFP protein towards the 960 nm peak by stabilizing the phenolate anion form of the GFP chromophore (Heim et al., 1995). Another example of an excitation spectrum shifting mutation is used in the GFP variant H9 and H9-40. H9 and H9-40 incorporate the mutation T203I near the chromophore. The T203I mutation is reported to preference the 800 nm excitation peak and decrease the 960 nm excitation peak by encouraging protonation of the chromophore to a neutral phenol (Tsien, 1998). The H9 GFP variant was further improved upon to incorporate a few additional mutations which improved folding stability at 37°C and named “Turbo-Sapphire” (Zapata-Hommer and Griesbeck, 2003). A GFP protein which preferences and stabilizes the 800 nm excitation and eliminates or greatly reduces the 960 nm excitation would be ideal for pixel

```

                20          40          60          80          100
Wildtype (WT) GFP  M S K G E E L F T  G V V P I L V E L D  G D V N G Q K F S V  S G E G E G D A T Y  G K L T L K F I C T  T G K L P V P W P T  L V T T F S Y G V Q  C F S R Y P D H M K  Q H D F F K S A M P  E G Y V Q E R T I F  99
eGFP (Clonetechn) M V S K G E E L F T  G V V P I L V E L D  G D V N G H K F S V  S G E G E G D A T Y  G K L T L K F I C T  T G K L P V P W P T  L V T T L I Y G V Q  C F S R Y P D H M K  Q H D F F K S A M P  E G Y V Q E R T I F  100
GFP2           M S K G E E L F T  G V V P I L V E L D  G D V N G H K F S V  S G E G E G D A T Y  G K L T L K F I C T  T G K L P V P W P T  L V T T L S Y G V Q  C F S R Y P D H M K  Q H D F F K S A M P  E G Y V Q E R T I F  97
GFP2(T203I)   M V S K G E E L F T  G V V P I L V E L D  G D V N G H K F S V  S G E G E G D A T Y  G K L T L K F I C T  T G K L P V P W P T  L V T T L S Y G V Q  C F S R Y P D H M K  Q H D F F K S A M P  E G Y V Q E R T I F  100
mT-Sapphire   M V S K G E E L F T  G V V P I L V E L D  G D V N G H K F S V  S G E G E G D A T Y  G K L T L K F I C T  T G K L P V P W P T  L V T T F S Y G V M  V F A R Y P D H M K  Q H D F F K S A M P  E G Y V Q E R T I F  100

                120          140          160          180          200
Wildtype (WT) GFP  Y K D D G N Y K T R  A E V K F E G D T L  V N R I E L K G I D  F K E D G N I L G H  K M E Y N Y N S H N  V Y I M A D K F K N  G I K V N F K I R H  N I K D G S V Q L A  D H Y Q Q N T P I G  D G P V L L P D N H  199
eGFP (Clonetechn) F K D D G N Y K T R  A E V K F E G D T L  V N R I E L K G I D  F K E D G N I L G H  K L E Y N Y N S H N  V Y I M A D K Q K N  G I K V N F K I R H  N I E D G S V Q L A  D H Y Q Q N T P I G  D G P V L L P D N H  200
GFP2           F K D D G N Y K T R  A E V K F E G D T L  V N R I E L K G I D  F K E D G N I L G H  K L E Y N Y N S H N  V Y I M A D K Q K N  G I K V N F K I R H  N I E D G S V Q L A  D H Y Q Q N T P I G  D G P V L L P D N H  197
GFP2(T203I)   F K D D G N Y K T R  A E V K F E G D T L  V N R I E L K G I D  F K E D G N I L G H  K L E Y N Y N S H N  V Y I M A D K Q K N  G I K V N F K I R H  N I E D G S V Q L A  D H Y Q Q N T P I G  D G P V L L P D N H  200
mT-Sapphire   F K D D G N Y K T R  A E V K F E G D T L  V N R I E L K G I D  F K E D G N I L G H  K L E Y N E N S H N  V Y I M A D K Q K N  G I K A N F K I R H  N I E D G S V Q L A  D H Y Q Q N T P I G  D G P V L L P D N H  200

                220          240
Wildtype (WT) GFP  Y L S T Q S A L S K  D P N E K R D H M V  L L E F V T A A G I  T H G M D E L Y K  239
eGFP (Clonetechn) Y L S T Q S A L S K  D P N E K R D H M V  L L E F V T A A G I  T L G M D E L Y K  240
GFP2           Y L S T Q S K L S K  D P N E K R D H M V  L L E F V T A A G I  T L G M D E L Y K  236
GFP2(T203I)   Y L S I Q S K L S K  D P N E K R D H M V  L L E F V T A A G I  T L G M D E L Y K  239
mT-Sapphire   Y L S I Q S K L S K  D P N E K R D H M V  L L E F V T A A G I  T S S S S S S S S S  231

```

Figure 10-6: Alignment of the protein sequences of wild type GFP, eGFP, GFP₂, GFP₂(T203I), and Sapphire

level FRET spectrometry while still allowing for the calculation of consistent gamma ratios and the calculation of reliable protein concentrations in cells.

The T203I mutation was made in the pCMV-GFP₂-MAVS vector and the new GFP variant was termed GFP₂(T203I). An alignment of the protein sequences of the newly created GFP₂(T203I) variant along with GFP wild type, eGFP, GFP₂, and Sapphire is shown in Figure 10-6.

Mutation	WT GFP	eGFP	GFP ₂	GFP ₂ (T203I)	mT-Sapphire
26	Q	H	H	H	H
65	F	L	L	L	F
66	S	T	S	S	S
71	C	C	C	C	V
73	S	S	S	S	A
101	Y	F	F	F	F
142	M	L	L	L	L
146	Y	Y	Y	Y	F
164	V	V	V	V	A
173	K	E	E	E	E
176	S	S	S	S	G
203	T	T	T	I	I
220	I	V	V	V	V
232	H	L	L	L	L

Table 10-1: Summary of different amino acid substitutions in wild-type GFP (WT GFP), eGFP, GFP₂, GFP₂(T203I), and mT-Sapphire.

If the T203I mutation functions to favor the 400 nm (800 nm two photon excitation) excitation, then when GFP₂(T203I)-MAVS was expressed in cells, the 400 nm excitation should always be the higher of the two possible GFP excitation maxima. To test the new GFP₂(T203I) variant, a plasmid expressing GFP₂(T203I)-MAVS was solely expressed in HEK 293T cells. After 24 hours, the cells were imaged with fluorescent excitation scans beginning at 720 nm and ending at 1000 nm. A total of three successive scans looking at the same region of interest were collected. The first excitation scan of GFP₂(T203I)-MAVS in cells showed that the 800 nm excitation was stabilized and had a higher emission intensity than the 960 nm excitation scan. However, each successive excitation scan decreased the 800 nm excitation and increased the 960 nm excitation (Figure 10-7).

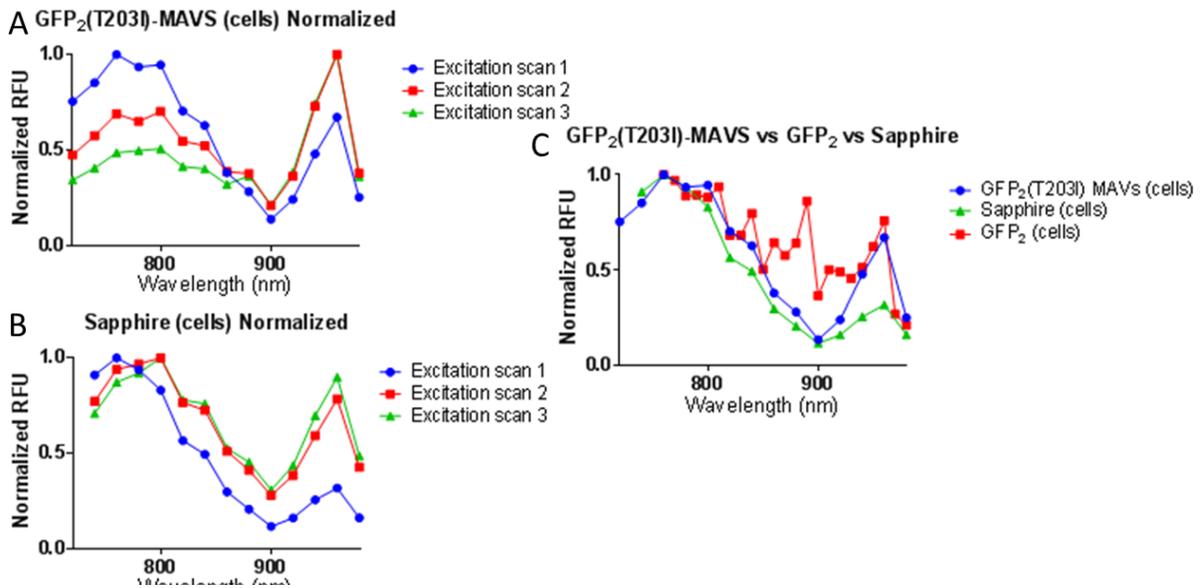


Figure 10-7: Excitation scans of GFP₂(T203I)-MAVs, Sapphire, and GFP₂ expressed in cells. A) Successive excitation scans of GFP₂(T203I)-MAVS expressed in HEK 293T cells. B) Successive excitation scans of Sapphire expressed in HEK293T cells. C) Overlay of the normalized excitation scans of GFP₂(T203I), Sapphire and GFP₂ expressed in cells.

To compare GFP₂(T203I)-MAVS to Sapphire, an additional set of HEK 293T cells were transfected solely with a plasmid expressing Sapphire. After being given 24 hours to express, successive excitation scans were taken of the same region in the cells expressing sapphire. The first excitation scan of sapphire was similar to the first excitation scan of GFP₂(T203I)-MAVS with the 800 nm excitation producing a higher intensity than any other excitation wavelength observed. However, successive excitation scans in the same region exhibit a shift where the emission when the cells were excited at 960 nm is now comparable to the emission when the cells were excited at 800 nm. In contrast to the GFP₂(T203I)-MAVS protein, the emission intensity at 800 nm did not decrease upon successive excitations (Figure 10-6). To compare the initial excitation spectra of GFP₂(T203I)-MAVS, to Sapphire and GFP₂, all three proteins were expressed alone in cells and then the excitation spectra collected from 720 nm to 1000 nm for each one (Figure 10-6). The excitation spectrum for all three proteins contain a peak at 800 nm and one at 960 nm. It was expected that the relative size of the peaks of the 800 nm excitation versus the 960 nm excitation would preference the 800 nm peak in both the sapphire and GFP₂(T203I) variants given that the T203I mutation was present in both and no additional reported chromophore altering mutations were included in sapphire. However, in terms of the relative size of 800 nm excitation peak to 960 nm excitation peak, GFP₂(T203I) was much more similar to GFP₂ than to Sapphire. The excitation peak of sapphire at 960 nm was much lower relative to the 800 nm peak than it was for both GFP₂ and GFP₂(T203I). The major difference between the spectra of GFP₂ and GFP₂(T203I) appears to be in the wavelengths between 800

nm and 960 nm. GFP₂ exhibits a much noisier spectrum between 800nm and 960 nm, whereas GFP₂(T203I) lacks the additional peaks and is smoother, similar to Sapphire.

Chapter 11 Conclusions and Future Directions

This thesis focuses on the interactions of HCV NS3 helicase with itself and with several other cellular helicases. The homo and hetero interactions of multiple cellular helicases such as the RIG-I like receptors was also explored. A new technique, Quantitative Microspectroscopic Imaging (Q-MSI), was developed that allows for the determination of FRET and fluorescent protein concentrations within specific sub-cellular regions in live cells under different conditions (i.e. in the presence or absence of ligand). Lastly, a novel fluorescent GFP molecule (termed GFP₂(T203I) was developed and some of its properties were explored.

Chapter 4 concentrated on the characterization of a new inhibitor, HPI, which has the ability to inhibit both the protease and helicase domains of NS3. HPI was shown, using HCV sub-genomic replicons expressing a renilla luciferase reporter and using rt-PCR, to be a potent inhibitor of HCV genotypes 1b, 3a, and 4a whereas HCV genotype 2a was insensitive to HPI (Figure 4-1). Therefore it was concluded that HPI is specific to the helicase and not binding non-specifically. Currently approved macrocyclic protease inhibitors were designed to make contacts in both the helicase and the protease domains and yet the macrocyclic protease inhibitors are unable to inhibit both functions. HPI appears to behave differently from macrocyclic protease inhibitors because it is able to inhibit not only the helicase but also the protease activity. HPI is proposed to bind to the allosteric site between the helicase and protease domains (Figure 4-2).

If HPI binds to the allosteric site between the protease and the helicase domain, it may be affecting the conformation of the NS3 protein. NS3 can exist in an open conformation with the protease domain swung out from the helicase domain or it can exist in a closed conformation with the protease domain tucked in close to the protease domain. These two conformations have biological significance related to the cleaving of the HCV poly-peptide. Specifically, the NS3 protease must cleave itself *in cis* and *in trans*. In order to cleave itself *in cis*, NS3 must adopt a closed conformation and to cleave itself *in trans* it must adopt an open conformation (Figure 4-4).

Modeling and protease activity assays suggest that HPI may be able to lock NS3 in a closed conformation and limit the ability of NS3 to cleave the HCV poly-peptide *in trans*. Peptide cleavage assays performed on NS3 proteins containing mutations in the helicase active site and the allosteric site between the helicase and protease domains showed that residue M485 when mutated to an alanine, resulted in a protein which was more sensitive to HPI (Figure 4-2). Modeling showed that M485 is located in the allosteric site and that the methionine residue may be blocking HPI access to the allosteric site. This suggests that HPI is binding in the allosteric site between the helicase and protease domains. In addition, the NS3 variant containing a covalently linked NS4A molecule (scNS3) was more sensitive to HPI than a full length NS3 with NS4A added exogenously (NS3:NS4A). The scNS3 molecule has only been crystallized in the compact (closed) conformation and in our lab it migrates faster than full length NS3 on a size exclusion column (despite its larger size). This may mean that scNS3 favors the closed conformation and it's possible that scNS3 is more sensitive to HPI because HPI binds better to the closed conformation than to the open conformation.

Lastly, HPI was tested in combination with linear and macrocyclic protease inhibitors using HCV sub-genomic replicons expressing a renilla luciferase reporter. The combination of HPI with macro-cyclic protease inhibitors danoprevir, and particularly grazoprevir, showed a synergistic interaction. Since the macro-cyclic protease inhibitors have been shown to make contacts in both the protease domain and the allosteric domain, it's possible that this synergistic interaction is due to HPI locking NS3 into a closed conformation and allowing the protease inhibitors better access to both the protease and helicase domains in which to make contacts (Figure 4-4). (Ndjomou et al., 2015).

The study looking at the interaction of HPI with HCV NS3 is the first time that the protease and helicase activities of NS3 have been targeted simultaneously. Future work with HPI could involve crystallizing HPI with the NS3 protein to visualize its mode of binding. This could lead to a greater understanding of the interaction of the NS3 helicase and protease domain during replication and could confirm our proposed model of HPI inhibition (Figure 4-5).

Chapter 5 sought to analyze the binding of a fluorescently tagged HCV NS3 protein with a fluorescently tagged DNA ligand using multiple FRET calculation techniques. The techniques chosen focused on the decrease in the donor emission upon binding to an acceptor (equation 2-2). The first technique used the peak emission intensity of the donor at the first point in the titration as intensity of the donor emission in the absence of acceptor and it was assumed that any change in the donor emission upon acceptor addition would be due to FRET. The peak donor intensity was then recorded at each point in the titration and F_{FRET} was calculated using equation 2-4. The second technique utilized spectral unmixing and determined the peak

emission co-efficient for the contribution of each component spectrum to the overall composite spectrum. The donor emission in the absence of acceptor was again calculated using the first point in the titration when there was no titrant added. Equation 2-2 was used to calculate E_{app} . The last FRET calculation technique mimicked what would be necessary in a microscopy experiment where the emission of the donor in the absence of acceptor cannot be measured directly. Instead, the emission of the donor in the absence of acceptor was back-calculated using equation 3-8 for each point in the titration and FRET was then calculated in the same way as the two previous methods using equation 2-2.

The resulting FRET efficiency values were graphed as a function of concentration of acceptor (Figure 5-2D) and fit to equation 3-8. The k_d , D_T , (total donor concentration) and E_{AD} (pairwise FRET efficiency for a dimer containing both one donor and one acceptor) values yielding the best fits were determined (Table 5-1). The overall titration curves for the FRET calculated using each technique are very similar (Figure 5-2). The results determined using each technique demonstrated that the back-calculation of the donor emission in the absence of acceptor produces results which are all within two fold or less of the values calculated through measurement of the donor emission directly.

Chapter 6 sought to look at the role of the NS4A anchor domain in anchoring a fluorescently labeled NS3 molecule in live cells. Images acquired when fluorescently labeled NS3 variants NS3h, NS3, scNS3, and NS3:NS4A were transfected into HEK293T cells show that only the NS3:NS4A molecule which contains the NS4A membrane anchors was able to localize into foci within the cells. The other proteins were diffuse throughout the cell. This confirms

previous work (Andrew Kohlway et al., 2014) which showed that the membrane anchors are necessary for sub-cellular localization of NS3 to replicase complexes.

In addition, the fluorescently labeled NS3 variants were co-transfected with a fluorescently tagged NS3 substrate containing a nuclear localization signal. When the substrate is cleaved, the nuclear localization signal is exposed and the fluorescent marker moves to the nucleus. As expected, the NS3h (lacking the protease domain) was unable to cleave the substrate. The NS3 molecule (lacking NS4A) was also unable to cleave the substrate. However, both the scNS3 and the NS3:NS4A were able to cleave the substrate, demonstrating that sub-cellular localization is not necessary to cleave protease substrates. Given that the NS3 protease has multiple targets in addition to the HCV poly-peptide, it makes sense that the localization to replicase complexes is not necessary for proteolytic cleavage.

Lastly, the full length NS3 (lacking NS4A) was unable to cleave the fluorescent substrate in cells but experiments *in vitro* using the full length NS3 (lacking NS4A) show that NS3 is able to cleave a protease substrate. Future work could explore why NS4A is necessary for proteolytic cleavage in cells but not *in vitro*. It could be that the pH in cells is different from the pH *in vitro* and it's possible that the pH is affecting the ability of the protease to cleave its substrate. It is also possible that the cell has lower concentrations of both NS3 and substrate than would be present in an *in vitro* experiment and therefore the less active NS3 (lacking NS4A) is ineffective in cells.

Chapter 7 looked at the interaction of HCV NS3 helicase with the RIG-I like receptors using ectopically expressed fluorescent fusion proteins and fully quantitative FRET analysis. Q-MSI was used to localize the interaction to sub-cellular regions within living cells. LGP2 was the only RLR which showed any significant interaction with HCV NS3 and domain deletion mutants of NS3 showed that the interaction of NS3 and LGP2 is due, in part, to the helicase domain. It's possible that this interaction is a direct protein:protein interaction or it's also possible that the interaction is mediated through shared binding to a ligand. However, when NS3 and LGP2 were co-expressed and exposed to the RNA ligand, poly I:C, (Figure 7-10), the FRET between NS3 and LGP2 increased only slightly.

The interaction between NS3 and LGP2 may be biologically relevant if it aids NS3 in localizing to MAVS so that the NS3 protease may cleave MAVS and disrupt RLR signaling. Normal RLR function involves the formation of an oligomer which localizes to the mitochondria to seed the formation of a MAVS filament. The formation of a MAVS filament initiates a signaling cascade which results in the production of interferon and other cytokines (Figure 11-1A). The NS3 protease disrupts RLR signaling through cleaving MAVS and preventing the formation of a MAVS filament. This is one way that HCV is able to establish an undetected

persistent infection. It is possible that NS3 is interacting with LGP2 in order to efficiently localize to MAVS so that the NS3 protease can cleave MAVS and abrogate signaling (Figure 11-1B).

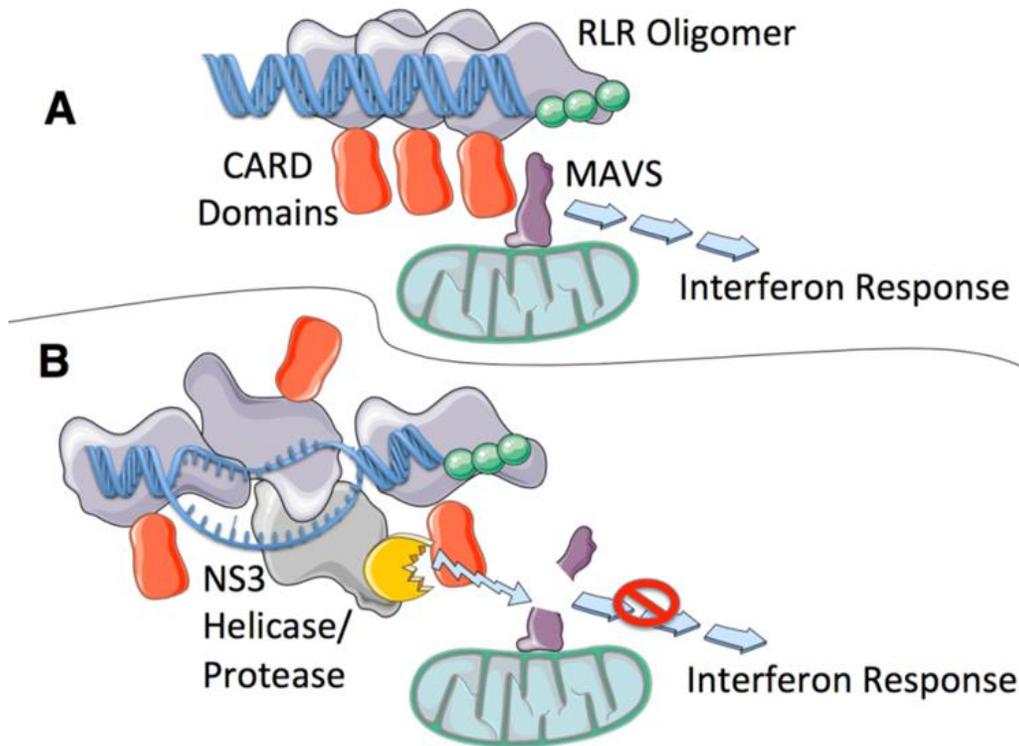


Figure 11-1: Model of HCV NS3 interacting with the RLR oligomer to aid in the cleavage of MAVS. A) In the absence of HCV, the RLR oligomer localizes to MAVS where the RLR CARDs interact with the MAVs CARDs to initiate a downstream signaling cascade. B) Proposed model for HCV NS3 interaction with the RLR filament to aid NS3 in localizing to and cleaving MAVS so as to disrupt the propagation of a downstream signal.

Previous work has looked at the role of NS3 helicase in the cleavage of MAVS. A study using a mutant NS3 (W501A) that lacks helicase function showed that NS3 was still able to block IRF-3 activation in response to Sendai Virus infection, meaning that NS3 was still able to cleave MAVS. However, when a mutant NS3 lacking protease function was used in the same assay, there was still some inhibition of IRF-3 signaling, implying that there may be other pathways of inhibition (Foy et al., 2003). Another study showed that an NS3 lacking the helicase domain was

sufficient for MAVS cleavage (Horner et al., 2012). Despite these findings, it is still possible that an LGP2:NS3helicase interaction is still aiding NS3 in localizing to MAVS. It is possible that although NS3 is still able to cleave MAVs in the absence of the helicase domain, it may be able to find MAVs more efficiently through an interaction with LGP2. Future work could include confirmation of the LGP2:NS3 interaction through alternative techniques such as co-purification using an epitope tag or yeast two-hybrid assays. In addition, confirmation of the interacting using immunofluorescence on endogenous LGP2 and HCV infection would be informative.

Chapter 8 looked at the curious re-localization of LGP2 to sub-cellular foci when it was co-expressed with DDX3. This re-localization is reminiscent of what was seen by Onomoto et al who used immunofluorescence to visualize RLR re-localization to stress granules in response to influenza-1 infection (Onomoto et al., 2012). Since the over-expression of DDX3 is known to induce the formation of stress granules, it is possible that the foci seen (Figure 8-1) are stress granules. DDX3 may help to shuttle LGP2 to stress granules. Future work could involve the use of immunofluorescence to look at the behavior of endogenous LGP2 and DDX3 in response to viral infection.

Chapter 9 looked at the homo and hetero interactions of the RIG-I like receptors. The RIG-I like receptors have previously been shown to form oligomers using electron microscopy (Bruns and Horvath, 2015; Bruns and Horvath, 2014). Q-MSI was used to look at the interactions of RLRs in living cells. When whole cell regions of interest were selected, the self-interaction of RIG-I and MDA5 was very low. But the interaction of LGP2:LGP2 was noticeable,

even in the absence of RNA (Figure 7-4). This may suggest that LGP2 forms homo-oligomers in the cytoplasm in the absence of PAMPs.

The incorporation of a third fluorescent marker to identify the mitochondria showed significant interaction of RIG-I:RIG-I and MDA5:MDA5 at the mitochondria (Figure 9-2). This interaction increased with the addition of a PAMP. However, the increase in interaction (identified by increased FRET efficiency) was much greater overall for MDA5 than for RIG-I. The differences in the range of FRET efficiency values found between RIG-I:RIG-I and MDA5:MDA5 may be indicative of the different oligomers formed at the mitochondria between those proteins. In the electron microscopy experiments, the MDA5 molecules formed much larger oligomers than the RIG-I molecules. It is possible that the overall larger shift in MDA5 FRET efficiency values in the presence of RNA PAMP is indicative of the greater number of larger MDA5 oligomers, as opposed to a smaller number of RIG-I:RIG-I oligomers.

This chapter also sought to look at the interaction of RIG-I and MDA5 with MAVS at the mitochondria. There were no major shift in FRET efficiency between RIG-I:MAVS or MDA5:MAVs in response to the synthetic RNA ligand. This may mean that the complex formed between RIG-I and MAVS and MDA5 and MAVS due to over-expression of MDA5 and RIG-I is the same complex formed when a PAMP is present. However, it could also mean that the shift in oligomer structure is not currently visible in this assay. An additional experiment performed between MAVS and MDA5 where the fluorescent protein on MDA5 and the linker between MDA5 and the fluorescent protein was changed showed overall higher FRET between MAVS and MDA5 and a larger increase in FRET in response to RNA ligand transfection (Figure 9-11).

This implies that the conditions in the previous MAVS:RIG-I and MAVS:MDA5 assays may not be optimal to see that shift in oligomer structure in response to PAMP addition. Future work could include changing the fluorescent proteins, changing the location of the fluorescent protein, or changing the linker between the fluorescent protein and the MAVS, MDA5, and/or RIG-I molecules.

The largest FRET efficiency increase in response to RNA PAMP was seen between LGP2 and MDA5. This interaction was previously studied and it is known that LGP2 and MDA5 interact on RNA PAMPS (Bruns and Horvath, 2015). When fluorescently tagged LGP2 and MDA5 were co-expressed there was a small population of FRET efficiencies greater than 10%, but upon addition of the RNA ligand, the FRET efficiencies increased dramatically, with some reaching as high as 80-90% (Figure 9-3). This suggests that LGP2 may not just “prime” RNA filaments for MDA5 interactions as has been previously proposed. But instead, LGP2 may be the dominant protomer in the LGP2:MDA5 oligomer (Figure 11-2).

Future work should include Q-MSI analysis of endogenous RLRs using immunofluorescence. In addition, the behavior of the RLRs in response to live virus infection could be explored. Lastly, the collection of more fluorescent images using RLRs tagged with carefully chosen fluorescent tags (i.e. GFP₂ and Venus) and the collection of the images using a single wavelength excitation and analysis using FRET spectrometry could lead to more information as to the specific architecture of the RLR complexes.

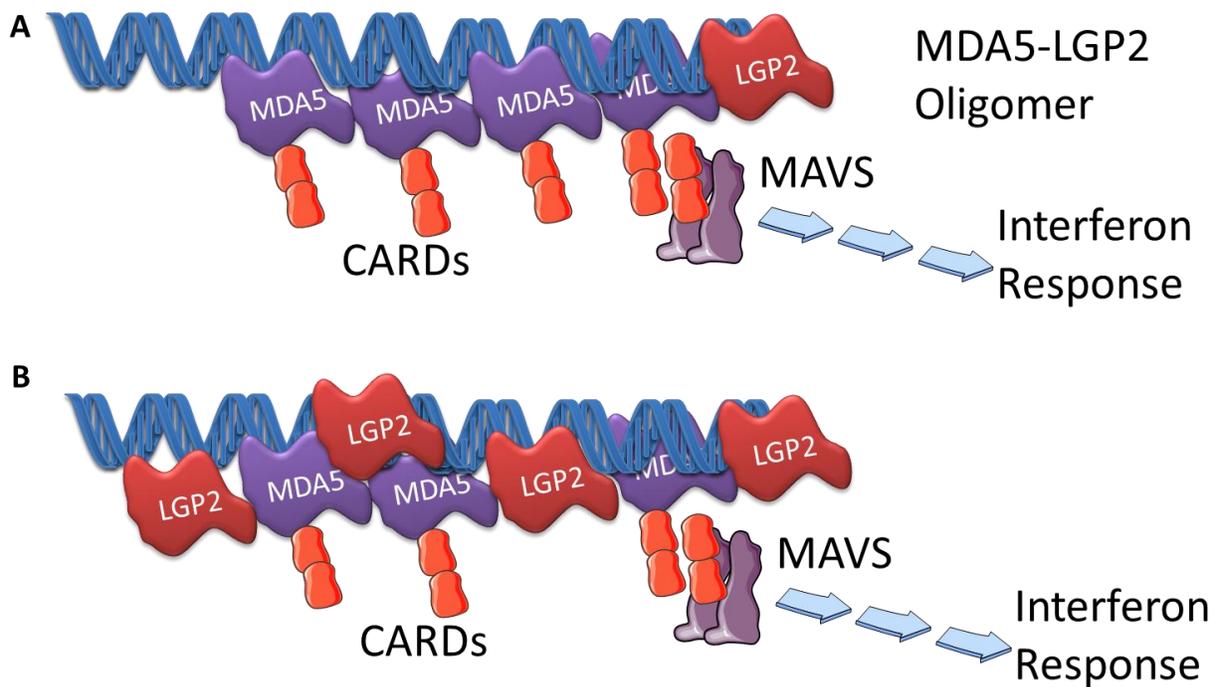


Figure 11-2: Model for LGP2:MDA5 oligomer. A) Proposed model for LGP2: MDA5 oligomer where LGP2 functions as an endcap or primer for the MDA5 filament. B) Proposed model for the LGP2:MDA5 oligomer where LGP2 is the predominant protomer in the overall MDA5:LGP2 oligomer.

Chapter 10 sought to improve future FRET spectrometry experiments through the development of a GFP variant which would be optimally excited at 800 nm (two-photon excitation) and not excited at 960 nm (two-photon excitation). Optical microspectrometry and FRET spectrometry rely on the very careful selection of fluorophores such that the acceptor

molecule is not excited at the donor excitation wavelength. This allows for the use of only a single excitation wavelength and eliminates the need for a second excitation scan to correct for direct excitation of the acceptor at the donor excitation wavelength. Using a single wavelength excitation, FRET spectrometry can generate spectrums of FRET efficiency as a function of number of pixels. These FRET spectrums can be fit to models to determine the overall architecture of the fluorescent protein complex. However, FRET spectrometry may be further refined through the generation of FRET efficiency spectrums as a function of protein concentrations or the molar fraction of protein concentrations. To calculate protein concentrations, a second excitation wavelength must be used and a value called the “gamma ratio” must be employed in the calculation (equations 3-9 and 3-10). A fluorescent protein combination which is optimal for single wavelength FRET spectrometry is GFP₂ and YFP. However, when the gamma ratio for GFP₂ is calculated in cells, there is a wide range of values (Figure 10-3).

A GFP₂ variant which incorporates the T203I mutation present in Sapphire was developed with the hope that it would produce an excitation spectrum with a 800 nm two-photon excitation maxima which was stable despite pH fluctuations and repeated fluorescent excitations. However, when the variant GFP₂(T203I) was tested in cells, the excitation spectra displayed properties similar to GFP₂ (Figure 10-4F) in that it fluctuated with pH and with repeated excitation scans (Figure 10-7A). Future work could involve making more mutations near the GFP₂ chromophore and then characterize the properties of those mutations to see if there are any mutations which stabilize the 800 nm excitation and minimize the 960 nm excitation over varying pHs or after repeated fluorescent excitations. The creation of a GFP

fluorescent protein with a stable 800 nm excitation and a minimal 960 nm excitation would aid in the calculation of E_{app} at pixel level.

Q-MSI is a powerful technique which allows for the calculation of protein:protein interactions in live cells and the determination of fluorescent protein concentrations in small sub-cellular regions of interest. Q-MSI was used to show that the interaction and oligomerization of the RIG-I like receptors could be monitored in live cells and it was used to discover a new and potentially biologically relevant interaction between LGP2 and HCV NS3. Continuation of the experiments to include pixel level FRET spectrometry analysis will allow for the elucidation of the specific oligomeric structure of the RIG-I like receptors and the LGP2:NS3 complex in live cells. It is possible that the oligomeric structure of the RIG-I like receptors may be different in response to different RNA PAMPS and an understanding of the structure could lead to a greater understanding as to how the RIG-I like receptors discriminate self from non-self and how the RLRs transmit their signal downstream. In addition, determination of specific oligomeric structure of the LGP2:NS3 complex could elucidate its function and provide information as to the biological relevance of the interaction.

Reference List

- Abdelhaleem, M. (2004). Do human RNA helicases have a role in cancer? *Biochimica Et Biophysica Acta (BBA) - Reviews on Cancer* 1704, 37-46.
- Aghemo, A., and De Francesco, R. (2013). New horizons in hepatitis C antiviral therapy with direct-acting antivirals. *Hepatology* 58, 428-438.
- Ahmad, S., and Hur, S. (2015). Helicases in Antiviral Immunity: Dual Properties as Sensors and Effectors. *Trends Biochem Sci* 40, 576-585.
- Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. *Cell* 124, 783-801.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell* (Garland Science).
- Anderson, C.S., Ortega, S., Chaves, F.A., Clark, A.M., Yang, H., Topham, D.J., and DeDiego, M.L. (2017). Natural and directed antigenic drift of the H1 influenza virus hemagglutinin stalk domain. *Sci. Rep.* 7, 14614.
- Andrew Kohlway, Nathan Pirakitikulr, Francisco N. Barrera, Olga Potapova, Donald M. Engelman, Anna M. Pyle, and Brett D. Lindenbach. (2014). Hepatitis C Virus RNA Replication and Virus Particle Assembly Require Specific Dimerization of the NS4A Protein Transmembrane Domain. *J. Virol.* 88, 628-642.
- Angus, A.G.N., Dalrymple, D., Boulant, S., McGivern, D.R., Clayton, R.F., Scott, M.J., Adair, R., Graham, S., Owsianka, A.M., Targett-Adams, P., *et al.* (2010). Requirement of cellular DDX3 for hepatitis C virus replication is unrelated to its interaction with the viral core protein. *J Gen Virol* 91, 122-132.
- Ansaldi, F., Orsi, A., Sticchi, L., and Bruzzone, B. (2014). Hepatitis C virus in the new era: Perspectives in epidemiology, prevention, diagnostics and predictors of response to therapy. *World J. Gastroenterol.* 20, 9633-9652.
- Ariumi, Y., Kuroki, M., Abe, K., Dansako, H., Ikeda, M., Wakita, T., and Kato, N. (2007). DDX3 DEAD-box RNA helicase is required for hepatitis C virus RNA replication. *J. Virol.* 81, 13922-13926.
- Atoom, A.M., Taylor, N.G.A., and Russell, R.S. (2014). The elusive function of the hepatitis C virus p7 protein. *Virology* 462, 377-387.

Aydin, C., Mukherjee, S., Hanson, A.M., Frick, D.N., and Schiffer, C.A. (2013). The interdomain interface in bifunctional enzyme protein 3/4A (NS3/4A) regulates protease and helicase activities. *Protein Sci.* *22*, 1786-1798.

Bajar, B.T., Wang, E.S., Zhang, S., Lin, M.Z., and Chu, J. (2016). A Guide to Fluorescent Protein FRET Pairs. *Sensors (Basel, Switzerland)* *16*, 1488.

Balistreri, W. (2016). A Focus on Hepatitis C: Novel Therapeutic Regimens. *Medscape*

Baril, M., Racine, M., Penin, F., and Lamarre, D. (2009). MAVS Dimer Is a Crucial Signaling Component of Innate Immunity and the Target of Hepatitis C Virus NS3/4A Protease. *J. Virol.* *83*, 1299-1311.

Bartenschlager, R., Ahlbornlaake, L., and Mous, J. (1994). Kinetic and Structural Analyses of Hepatitis C Virus Polyprotein Processing. *J. Virol.* *68*, 5045-5055.

Battistutta, R., Negro, A., and Zanotti, G. (2000). Crystal structure and refolding properties of the mutant F99S/M153T/V163A of the green fluorescent protein. *Proteins* *41*, 429-437.

Becker, W. (2012). Fluorescence lifetime imaging--techniques and applications. *J. Microsc.* *247*, 119-136.

Belon, C.A., and Frick, D.N. (2009a). Fuel specificity of the hepatitis C virus NS3 helicase. *J. Mol. Biol.* *388*, 851-864.

Belon, C.A., and Frick, D.N. (2009b). Helicase inhibitors as specifically targeted antiviral therapy for hepatitis C. *Future Virol.* *4*, 277-293.

Belon, C.A., and Frick, D.N. (2008). Monitoring helicase activity with molecular beacons. *BioTechniques* *45*, 440, 442.

Beran, R.K.F., Lindenbach, B.D., and Pyle, A.M. (2009). The NS4A Protein of Hepatitis C Virus Promotes RNA-Coupled ATP Hydrolysis by the NS3 Helicase. *J. Virol.* *83*, 3268-3275.

Beran, R.K.F., and Pyle, A.M. (2008). Hepatitis C Viral NS3-4A Protease Activity Is Enhanced by the NS3 Helicase. *J. Biol. Chem.* *283*, 29929-29937.

Berezin, M.Y., and Achilefu, S. (2010). Fluorescence Lifetime Measurements and Biological Imaging. *Chem. Rev.* *110*, 2641.

Biener, G., Stoneman, M.R., Acbas, G., Holz, J.D., Orlova, M., Komarova, L., Kuchin, S., and Raicu, V. (2013). Development and experimental testing of an optical micro-spectroscopic technique incorporating true line-scan excitation. *Int. J. Mol. Sci.* *15*, 261-276.

- Bird, L.E., Subramanya, H.S., and Wigley, D.B. (1998). Helicases: a unifying structural theme? *Curr. Opin. Struct. Biol.* *8*, 14-18.
- Bowie, A.G., and Unterholzner, L. (2008). Viral evasion and subversion of pattern-recognition receptor signalling. *Nat Rev Immunol.* *8*, 911-922.
- Breitbart, M., and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* *13*, 278-284.
- Bruns, A.M., and Horvath, C.M. (2015). LGP2 synergy with MDA5 in RLR-mediated RNA recognition and antiviral signaling. *Cytokine* *74*, 198-206.
- Bruns, A.M., Leser, G.P., Lamb, R.A., and Horvath, C.M. (2014). The innate immune sensor LGP2 activates antiviral signaling by regulating MDA5-RNA interaction and filament assembly. *Mol. Cell* *55*, 771-781.
- Bruns, A.M., and Horvath, C.M. (2014). Antiviral RNA recognition and assembly by RLR family innate immune sensors. *Cytokine Growth Factor Rev.* *25*, 507-512.
- Bruns, A.M., and Horvath, C.M. (2012). Activation of RIG-I-like receptor signal transduction. *Crit. Rev. Biochem. Mol. Biol.* *47*, 194-206.
- Bruns, A.M., Pollpeter, D., Hadizadeh, N., Myong, S., Marko, J.F., and Horvath, C.M. (2013). ATP hydrolysis enhances RNA recognition and antiviral signal transduction by the innate immune sensor, laboratory of genetics and physiology 2 (LGP2). *J. Biol. Chem.* *288*, 938-946.
- Buchan, J.R., Muhlrads, D., and Parker, R. (2008). P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *J. Cell Biol.* *183*, 441-455.
- Byrd, A.K., and Raney, K.D. (2012). Superfamily 2 helicases. *Front Biosci.* *17*, 2070-2088.
- Cao, X., Ding, Q., Lu, J., Tao, W., Huang, B., Zhao, Y., Niu, J., Liu, Y., and Zhong, J. (2015). MDA5 plays a critical role in interferon response during hepatitis C virus infection. *J. Hepatol.* *62*, 771-778.
- Chattopadhyay, S., Chen, Y., and Weller, S.K. (2006). The two helicases of herpes simplex virus type 1 (HSV-1). *Front. Biosci.* *11*, 2213-2223.
- Chen, Y.E., Mauldin, J.P., Day, R.N., and Periasamy, A. (2007). Characterization of spectral FRET imaging microscopy for monitoring nuclear protein interactions. *J. Microsc.* *228*, 139-152.
- Chevaliez, S., and Pawlotsky, J. (2006). HCV Genome and Life Cycle. In *Hepatitis C Viruses: Genomes and Molecular Biology*, Tan, Seng-Lai ed., (Norfolk (UK): Horizon Bioscience)

Childs, K.S., Randall, R.E., and Goodbourn, S. (2013). LGP2 plays a critical role in sensitizing mda-5 to activation by double-stranded RNA. *PLoS One* 8, e64202.

Clegg, R.M. (1995). Fluorescence resonance energy transfer. *Curr. Opin. Biotechnol.* 6, 103-110.

Corby, M.J., Stoneman, M.R., Biener, G., Paprocki, J.D., Kolli, R., Raicu, V., and Frick, D.N. (2017). Quantitative microspectroscopic imaging reveals viral and cellular RNA helicase interactions in live cells. *J. Biol. Chem.* 292, 11165-11177.

Craig McCormick, and Denys A Khapersky. (2017). Translation inhibition and stress granules in the antiviral immune response. *Nat. Rev. Immunol.* 17, 647.

Cui, S., Eisenächer, K., Kirchhofer, A., Brzózka, K., Lammens, A., Lammens, K., Fujita, T., Conzelmann, K., Krug, A., and Hopfner, K. (2008). The C-Terminal Regulatory Domain Is the RNA 5'-Triphosphate Sensor of RIG-I. *Mol. Cell* 29, 169-179.

Decker, T., Soulat, D., Bürckstümmer, T., Stefanovic, A., Goncalves, A., Hantschel, O., Superti-Furga, G., Bennett, K.L., Westermayer, S., and Bauch, A. (2008). The DEAD-box helicase DDX3X is a critical component of the TANK-binding kinase 1-dependent innate immune response. *The EMBO Journal* 27, 2135-2146.

del Toro Duany, Y., Wu, B., and Hur, S. (2015). MDA5-filament, dynamics and disease. *Curr. Opin. Virol.* 12, 20-25.

Di Cera, E. (2009). Serine Proteases. *IUBMB Life* 61, 510-515.

Doehle, B.P., and Michael Gale, J. (2013). Innate Immune Evasion Strategies of HCV and HIV: Common Themes for Chronic Viral Infection (Landes Bioscience).

Douam, F., Dao Thi, V.L., Maurin, G., Fresquet, J., Mompelat, D., Zeisel, M.B., Baumert, T.F., Cosset, F., and Lavillette, D. (2014). Critical interaction between E1 and E2 glycoproteins determines binding and fusion properties of hepatitis C virus during cell entry. *Hepatology* 59, 776-788.

Dubuisson, J. (2007). Hepatitis C virus proteins. *World J. Gastroenterol.* 13, 2406-2415.

Ehle, H., and Horn, A. (1990). Immunoaffinity chromatography of enzymes. *Bioseparation* 1, 97-110.

Elana S Rosenthal, and Camilla S Graham. (2016). Price and affordability of direct-acting antiviral regimens for hepatitis C virus in the United States. *Infect. Agent Cancer* 11, 24.

Emmanuelle Blanchard, Sandrine Belouzard, Lucie Goueslain, Takaji Wakita, Jean Dubuisson, Czeslaw Wychowski, and Yves Rouillé. (2006). Hepatitis C Virus Entry Depends on Clathrin-Mediated Endocytosis. *J. Virol.* *80*, 6964-6972.

Errett, J.S., and Gale, M. (2015a). Emerging complexity and new roles for the RIG-I-like receptors in innate antiviral immunity. *Virol Sin* *30*, 163-173.

Errett, J.S., and Gale, M. (2015b). Emerging complexity and new roles for the RIG-I-like receptors in innate antiviral immunity. *Virol Sin* *30*, 163-173.

Fairman-Williams, M.E., Guenther, U., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.* *20*, 313-324.

Ferrao, R., and Wu, H. (2012). Helical assembly in the death domain (DD) superfamily. *Curr. Opin. Struct. Biol.* *22*, 241-247.

Ferrari, E., Wright-Minogue, J., Fang, J.W.S., Baroudy, B.M., Lau, J.Y.N., and Hong, Z. (1999). Characterization of Soluble Hepatitis C Virus RNA-Dependent RNA Polymerase Expressed in *Escherichia coli*. *J. Virol.* *73*, 1649-1654.

Fersht, A.R., and Sperling, J. (1973). The charge relay system in chymotrypsin and chymotrypsinogen. *J. Mol. Biol.* *74*, 137-149.

Foy, E., Li, K., Wang, C., Sumpter, R., Ikeda, M., Lemon, S.M., and Gale, M. (2003). Regulation of interferon regulatory factor-3 by the hepatitis C virus serine protease. *Science* *300*, 1145-1148.

Freedman, H., Logan, M.R., Law, J.L.M., and Houghton, M. (2016). Structure and Function of the Hepatitis C Virus Envelope Glycoproteins E1 and E2: Antiviral and Vaccine Targets. *ACS Infect. Dis.* *2*, 749-762.

Frick, D.N., Ginzburg, O., and Lam, A.M.I. (2010). A method to simultaneously monitor hepatitis C virus NS3 helicase and protease activities. *Methods Mol. Biol.* *587*, 223-233.

Frick, D.N., Jaeschke, H., Mullen, K., and Moradpour, D. (2006). Step-by-step progress toward understanding the hepatitis C virus RNA helicase. *Hepatology* *43*, 1392-1395.

Frick, D.N., Rypma, R.S., Lam, A.M.I., and Gu, B. (2004). The nonstructural protein 3 protease/helicase requires an intact protease domain to unwind duplex RNA efficiently. *J. Biol. Chem.* *279*, 1269-1280.

Fridell, R.A., Qiu, D., Valera, L., Wang, C., Rose, R.E., and Gao, M. (2011). Distinct Functions of NS5A in Hepatitis C Virus RNA Replication Uncovered by Studies with the NS5A Inhibitor BMS-790052. *J. Virol.* *85*, 7312-7320.

Fullam, A., and Schröder, M. (2013). DExD/H-box RNA helicases as mediators of anti-viral innate immunity and essential host factors for viral replication. *BBA-Gene Regulatory Mechanisms* 1829, 854-865.

Fuller-Pace, F.V. (2013). The DEAD box proteins DDX5 (p68) and DDX17 (p72): multi-tasking transcriptional regulators. *Biochim. Biophys. Acta* 1829, 756-763.

Gawlik, K., and Galloway, P. (2014). HCV core protein and virus assembly: what we know without structures. *Immunol. Res.* 60, 1-10.

Gelderblom, H.R. (1996). Structure and Classification of Viruses. In *Medical Microbiology*, Baron, Samuel ed., (Galveston (TX): University of Texas Medical Branch at Galveston)

Gentsch, J., Brohm, C., Steinmann, E., Friesland, M., Menzel, N., Vieyres, G., Monteiro Perin, P., Frentzen, A., Kaderali, L., and Pietschmann, T. (2013). Hepatitis C Virus p7 is Critical for Capsid Assembly and Envelopment. *PLoS Pathogens* 9,

Gerlach, J.T., Diepolder, H.M., Zachoval, R., Gruener, N.H., Jung, M., Ulsenheimer, A., Schraut, W.W., Schirren, C.a., Waechtler, M., Backmund, M., and Pape, G.R. (2003). Acute hepatitis C: high rate of both spontaneous and treatment-induced viral clearance. *Gastroenterology* 125, 80-88.

Götte, M., and Feld, J.J. (2016). Direct-acting antiviral agents for hepatitis C: structural and mechanistic insights. *Nat. Rev. Gastroenterol. Hepatol.* 13, 338-351.

Goubau, D., Schlee, M., Deddouche, S., Pruijssers, A.J., Zillinger, T., Goldeck, M., Schuberth, C., Van der Veen, Annemarie G, Fujimura, T., Rehwinkel, J., *et al.* (2014). Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates. *Nature* 514, 372-375.

Gu, L., Fullam, A., Brennan, R., and Schroeder, M. (2013). Human DEAD Box Helicase 3 Couples I κ B Kinase ϵ to Interferon Regulatory Factor 3 Activation. *Mol. Cell. Biol.* 33, 2004-2015.

Gu, M., and Rice, C. (2010). Three Conformational Snapshots of the Hepatitis Virus NS3 Helicase Reveal a Ratchet Translocation Mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 107, 521-528.

Gu, M., and Rice, C.M. (2013). Structures of hepatitis C virus nonstructural proteins required for replicase assembly and function. *Curr. Opin. Virol.* 3, 129-136.

Hale, B.G., Albrecht, R.A., and García-Sastre, A. (2010). Innate immune evasion strategies of influenza viruses. *Future Microbiol.* 5, 23.

Hanada, K., and Hickson, I. (2007). Molecular genetics of RecQ helicase disorders. *Cell. Mol. Life Sci* 64, 2306-2322.

Hanson, A.M., Hernandez, J.J., Shadrick, W.R., and Frick, D.N. (2012). Identification and analysis of inhibitors targeting the hepatitis C virus NS3 helicase. *Meth. Enzymol.* *511*, 463-483.

He, Y., Staschke, K.A., and Tan, S. (2006). HCV NS5A: A Multifunctional Regulator of Cellular Pathways and Virus Replication. In *Hepatitis C Viruses: Genomes and Molecular Biology*, Tan, Seng-Lai ed., (Norfolk (UK): Horizon Bioscience)

Hei, L., and Zhong, J. (2017). Laboratory of genetics and physiology 2 (LGP2) plays an essential role in hepatitis C virus infection-induced interferon responses. *Hepatology* *65*, 1478-1491.

Heim, R., Cubitt, A.B., and Tsien, R.Y. (1995). Improved green fluorescence. *Nature* *373*, 663-664.

Heraud-Farlow, J.E., and Walkley, C.R. (2016). The role of RNA editing by ADAR1 in prevention of innate immune sensing of self-RNA. *J. Mol. Med.* *94*, 1095-1102.

Hickman, A.B., and Dyda, F. (2005). Binding and unwinding: SF3 viral helicases. *Curr. Opin. Struct. Biol.* *15*, 77-85.

Hoppe, A., Christensen, K., and Swanson, J.A. (2002). Fluorescence resonance energy transfer-based stoichiometry in living cells. *Biophys. J.* *83*, 3652-3664.

Horner, S.M., Liu, H.M., Park, H.S., Briley, J., and Gale, M., Jr. (2011). Mitochondrial-associated endoplasmic reticulum membranes (MAM) form innate immune synapses and are targeted by hepatitis C virus. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 14590-14595.

Horner, S.M. (2014). Activation and Evasion of Antiviral Innate Immunity by Hepatitis C Virus. *J. Mol. Biol.* *426*, 1198-1209.

Horner, S.M., Park, H.S., and Gale, M. (2012). Control of innate immune signaling and membrane targeting by the Hepatitis C virus NS3/4A protease are governed by the NS3 helix $\alpha 0$. *J. Virol.* *86*, 3112-3120.

Hou, F., Sun, L., Zheng, H., Skaug, B., Jiang, Q., and Chen, Z.J. (2011). MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell* *146*, 448-461.

Houghton, M. (2009). Discovery of the hepatitis C virus. *Liver Int.* *29 Suppl 1*, 82-88.

Hoving, J.C., Wilson, G.J., and Brown, G.D. (2014). Signalling C-Type lectin receptors, microbial recognition and immunity. *Cell. Microbiol.* *16*, 185-194.

- Howe, A., Chase, R., Taremi, S., Risano, C., Beyer, B., Malcolm, B., and Lau, J.Y.N. (1999). A novel recombinant single-chain hepatitis C virus NS3-NS4A protein with improved helicase activity. *Protein Sci.* **8**, 1332-1341.
- Jankowsky, A., Guenther, U.-., and Jankowsky, E. (2011). The RNA helicase database. *Nucleic Acids Research* **39**, D341.
- Jares-Erijman, E.A., and Jovin, T.M. (2006). Imaging molecular interactions in living cells by FRET microscopy. *Curr Opin Chem Biol* **10**, 409-416.
- Jensen, S., and Thomsen, A.R. (2012). Sensing of RNA Viruses: a Review of Innate Immune Receptors Involved in Recognizing RNA Virus Invasion. *Journal of Virology* **86**, 2900.
- John S. Errett Michael Gale Jr. (2015). Emerging complexity and new roles for the RIG-I-like receptors in innate antiviral immunity. *Viol. Sin* **30**, 163-173.
- Jonathan Maelfait, and Rudi Beyaert. (2012). Emerging Role of Ubiquitination in Antiviral RIG-I Signaling. *Microbiol. Mol. Biol. Rev.* **76**, 33-45.
- Jovin, T.M., Lidke, D.S., and Jares-Erijman, E.A. (2005). Fluorescence Resonance Energy Transfer (FRET) and Fluorescence Lifetime Imaging Microscopy (FLIM). In *From Cells to Proteins: Imaging Nature across Dimensions*, Springer, Dordrecht) pp. 209-216.
- Karpova, T.S., Baumann, C.T., He, L., Wu, X., Grammer, A., Lipsky, P., Hager, G.L., and McNally, J.G. (2003). Fluorescence resonance energy transfer from cyan to yellow fluorescent protein detected by acceptor photobleaching using confocal microscopy and a single laser. *J. Microsc.* **209**, 56-70.
- Kasprzak, A., Adamek, A., Biczysko, W., Seidel, J., Przybyszewska, W., Olejniczak, K., Juszczak, J., and Zabel, M. (2007). Intracellular expression of the proliferative marker Ki-67 and viral proteins (NS3, NS5A and C) in chronic, long lasting hepatitis C virus (HCV) infection. *Folio Histochem. Cytobiol.* **45**, 357-366.
- Kaufmann, S.H.E., Dorhoi, A., Hotchkiss, R.S., and Bartenschlager, R. (2017). Host-directed therapies for bacterial and viral infections. *Nat. Rev. Drug Discov.* nrd.2017.162.
- Kedersha, N., and Anderson, P. (2007). Mammalian stress granules and processing bodies. *Meth. Enzymol.* **431**, 61-81.
- Kim, C.W., and Chang, K. (2013). Hepatitis C virus: virology and life cycle. *Clinical and Molecular Hepatology* **19**, 17-25.
- King, C., Stoneman, M., Raicu, V., and Hristova, K. (2016). Fully quantified spectral imaging reveals in vivo membrane protein interactions. *Integr. Biol.* **8**, 216-229.

Korolev, S., Yao, N., Lohman, T.M., Weber, P.C., and Waksman, G. (1998). Comparisons between the structures of HCV and Rep helicases reveal structural similarities between SF1 and SF2 super-families of helicases. *Protein Sci.* 7, 605-610.

Kowalinski, E., Lunardi, T., McCarthy, A., Louber, J., Brunel, J., Grigorov, B., Gerlier, D., and Cusack, S. (2011). Structural Basis for the Activation of Innate Immune Pattern-Recognition Receptor RIG-I by Viral RNA. *Cell* 147, 423-435.

Kunkel, M., Lorinczi, M., Rijnbrand, R., Lemon, S.M., and Watowich, S.J. (2001). Self-Assembly of Nucleocapsid-Like Particles from Recombinant Hepatitis C Virus Core Protein. *J. Virol.* 75, 2119-2129.

Kwo, P.Y., and Vinayek, R. (2011). The Therapeutic Approaches for Hepatitis C Virus: Protease Inhibitors and Polymerase Inhibitors. *Gut Liver* 5, 406-417.

Kwon, S., Zhang, Y., and Matthias, P. (2007). The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. *Genes Dev.* 21, 3381-3394.

Lakowicz, J.R. (2010). *Principles of fluorescence spectroscopy* (New York, NY: Springer).

Lässig, C., and Hopfner, K. (2017). Discrimination of cytosolic self and non-self RNA by RIG-I-like receptors. *J. Biol. Chem.* 292, 9000-9009.

Lauren Gravitz. (2011). A smouldering public-health crisis. *Nature* 474, S2.

Leavesley, S.J., Britain, A.L., Cichon, L.K., Nikolaev, V.O., and Rich, T.C. (2013). Assessing FRET using Spectral Techniques. *Cytometry A* 83, 898-912.

Lechevalier, H. (1972). Dmitri Josifovich Ivanovski (1864-1920). *Bacteriol. Rev.* 36, 135-145.

Levin, M.K., Patel, S.S., and Gurjar, M. (2005). A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nat. Struct. Mol. Biol.* 12, 429-435.

Liddicoat, B.J., Piskol, R., Chalk, A.M., Ramaswami, G., Higuchi, M., Hartner, J.C., Li, J.B., Seeburg, P.H., and Walkley, C.R. (2015). RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* 349, 1115-1120.

Lin, C. (2006). HCV NS3-4A Serine Protease. In *Hepatitis C Viruses: Genomes and Molecular Biology*, Tan, Seng-Lai ed., (Norfolk (UK): Horizon Bioscience)

Lindenbach, B.D., and Rice, C.M. (2005). Unravelling hepatitis C virus replication from genome to function. *Nature* 436, 933-938.

- Liu, H.M., and Gale, M. (2010). Hepatitis C Virus Evasion from RIG-I-Dependent Hepatic Innate Immunity. *Gastroenterol. Res. Pract.* *2010*, 548390.
- Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., and Darnell, J. (2000). *Molecular Cell Biology* (W. H. Freeman).
- Lohmann, V., Körner, F., Koch, J., Herian, U., Theilmann, L., and Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* *285*, 110-113.
- Loo, Y.M., and Gale, M., Jr. (2011). Immune signaling by RIG-I-like receptors. *Immunity* *34*, 680-692.
- Loo, Y., and Gale, M. (2007). Viral Regulation and Evasion of the Host Response. In *Interferon: The 50th Anniversary*, Springer, Berlin, Heidelberg) pp. 295-313.
- Loo, Y., Owen, D.M., Li, K., Erickson, A.K., Johnson, C.L., Fish, P.M., Carney, D.S., Wang, T., Ishida, H., Yoneyama, M., *et al.* (2006). Viral and therapeutic control of IFN- β promoter stimulator 1 during hepatitis C virus infection. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 6001-6006.
- Louber, J., Brunel, J., Uchikawa, E., Cusack, S., and Gerlier, D. (2015). Kinetic discrimination of self/non-self RNA by the ATPase activity of RIG-I and MDA5. *BMC Biol.* *13*, 54.
- Lu, C., Xu, H., Ranjith-Kumar, C.T., Brooks, M.T., Hou, T.Y., Hu, F., Herr, A.B., Strong, R.K., Kao, C.C., and Li, P. (2010). The structural basis of 5' triphosphate double-stranded RNA recognition by RIG-I C-terminal domain. *Structure* *18*, 1032-1043.
- Lucas, M., Karrer, U., Lucas, A., and Klenerman, P. (2001). Viral escape mechanisms – escapology taught by viruses. *Int. J. Clin. Exp. Pathol* *82*, 269-286.
- Luo, D., Ding, S., Vela, A., Kohlway, A., Lindenbach, B., and Pyle, A. (2011). Structural Insights into RNA Recognition by RIG-I. *Cell* *147*, 409-422.
- Manns, M.P., Wedemeyer, H., and Cornberg, M. (2006). Treating viral hepatitis C: efficacy, side effects, and complications. *Gut* *55*, 1350-1359.
- Marques, J.T., Devosse, T., Wang, D., Zamanian-Daryoush, M., Serbinowski, P., Hartmann, R., Fujita, T., Behlke, M.A., and Williams, B.R. (2006). A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nat. Biotechnol.* *24*, 559-65.
- Miller, K.E., Kim, Y., Huh, W., and Park, H. (2015). Bimolecular Fluorescence Complementation (BiFC) Analysis: Advances and Recent Applications for Genome-Wide Interaction Studies. *J. Mol. Biol.* *427*, 2039-2055.

Mishra, A.K., Gragg, M., Stoneman, M.R., Biener, G., Oliver, J.A., Miszta, P., Filipek, S., Raicu, V., and Park, P.S.-. (2016). Quaternary structures of opsin in live cells revealed by FRET spectrometry. *Biochem. J.* *473*, 3819-3836.

Morise, H., Shimomura, O., Johnson, F.H., and Winant, J. (1974). Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* *13*, 2656-2662.

Murali, A., Li, X., Ranjith-Kumar, C.T., Bhardwaj, K., Holzenburg, A., Li, P., and Kao, C.C. (2008). Structure and function of LGP2, a DEX(D/H) helicase that regulates the innate immunity response. *J. Biol. Chem.* *283*, 15825-15833.

Myong, S., Cui, S., Cornish, P.V., Kirchhofer, A., Gack, M.U., Jung, J.U., Hopfner, K., and Ha, T. (2009). Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA. *Science* *323*, 1070-1074.

Nakayama, H. (2002). RecQ family helicases: roles as tumor suppressor proteins. *Oncogene* *21*, 9008-9021.

Namba, K., Naka, K., Dansako, H., Nozaki, A., Ikeda, M., Shiratori, Y., Shimotohno, K., and Kato, N. (2004). Establishment of hepatitis C virus replicon cell lines possessing interferon-resistant phenotype. *Biochem. Biophys. Res. Commun.* *323*, 299-309.

Ndjomou, J., Corby, M.J., Sweeney, N.L., Hanson, A.M., Aydin, C., Ali, A., Schiffer, C.A., Li, K., Frankowski, K.J., Schoenen, F.J., and Frick, D.N. (2015). Simultaneously Targeting the NS3 Protease and Helicase Activities for More Effective Hepatitis C Virus Therapy. *ACS Chem. Biol.* *10*, 1887.

Ndjomou, J., Kolli, R., Mukherjee, S., Shadrack, W.R., Hanson, A.M., Sweeney, N.L., Bartczak, D., Li, K., Frankowski, K.J., Schoenen, F.J., and Frick, D.N. (2012). Fluorescent primuline derivatives inhibit hepatitis C virus NS3-catalyzed RNA unwinding, peptide hydrolysis and viral replicase formation. *Antiviral Res.* *96*, 245-255.

Neddermann, P., Clementi, A., and De Francesco, R. (1999). Hyperphosphorylation of the Hepatitis C Virus NS5A Protein Requires an Active NS3 Protease, NS4A, NS4B, and NS5A Encoded on the Same Polyprotein. *J. Virol.* *73*, 9984-9991.

Neumann-Haefelin, C., Frick, D.N., Wang, J.J., Pybus, O.G., Salloum, S., Narula, G.S., Eckart, A., Biezyński, A., Eiermann, T., Klenerman, P., *et al.* (2008). Analysis of the evolutionary forces in an immunodominant CD8 epitope in hepatitis C virus at a population level. *J. Virol.* *82*, 3438-3451.

Nover, L., Scharf, K.D., and Neumann, D. (1983). Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves. *Mol. Cell. Biol.* *3*, 1648-1655.

- Oda, S., Schröder, M., and Khan, A.R. (2009). Structural basis for targeting of human RNA helicase DDX3 by poxvirus protein K7. *Structure* 17, 1528-1537.
- Onomoto, K., Jogi, M., Yoo, J., Narita, R., Morimoto, S., Takemura, A., Sambhara, S., Kawaguchi, A., Osari, S., Nagata, K., *et al.* (2012). Critical role of an antiviral stress granule containing RIG-I and PKR in viral detection and innate immunity. *PLoS ONE* 7, e43031.
- Onomoto, K., Yoneyama, M., Fung, G., Kato, H., and Fujita, T. (2014). Antiviral innate immunity and stress granule responses. *Trends Immunol.* 35, 420-428.
- Oshiumi, H., Sakai, K., Matsumoto, M., and Seya, T. (2010). DEAD/H BOX 3 (DDX3) helicase binds the RIG-I adaptor IPS-1 to up-regulate IFN-beta-inducing potential. *Eur. J. Immunol.* 40, 940-948.
- Owen, J., Punt, J., Stranford, S.A., and Kubly, J. (2013). *Immunology* (New York, NY: Freeman).
- Owsianka, A.M., and Patel, A.H. (1999). Hepatitis C virus core protein interacts with a human DEAD box protein DDX3. *Virology* 257, 330-340.
- Patel, S.S., and Donmez, I. (2006). Mechanisms of helicases. *J. Biol. Chem.* 281, 18265-18268.
- Patowary, S., Pisterzi, L., Biener, G., Holz, J., Oliver, J., Wells, J., and Raicu, V. (2015). Experimental Verification of the Kinetic Theory of FRET Using Optical Microspectroscopy and Obligate Oligomers. *Biophys. J.* 108, 1613-1622.
- Peisley, A., Wu, B., Xu, H., Chen, Z.J., and Hur, S. (2014). Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I. *Nature* 509, 110-114.
- Peisley, A., Hyun Jo, M., Lin, C., Wu, B., Orme-Johnson, M., Walz, T., Hohng, S., and Hur, S. (2012). Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments. *Proc. Natl. Acad. Sci. U. S. A.* 109, E3340.
- Peisley, A., Lin, C., Wu, B., Orme-Johnson, M., Liu, M., Walz, T., and Hur, S. (2011). Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition. *Proc. Natl. Acad. Sci. U. S. A.* 108, 21010-21015.
- Peschel, A., Jack, R.W., Otto, M., Collins, L.V., Staubitz, P., Nicholson, G., Kalbacher, H., Nieuwenhuizen, W.F., Jung, G., Tarkowski, A., van Kessel, Kok P M, and van Strijp, Jos A G. (2001). Staphylococcus aureus Resistance to Human Defensins and Evasion of Neutrophil Killing via the Novel Virulence Factor Mprf Is Based on Modification of Membrane Lipids with L-Lysine. *J. Exp. Med.* 193, 1067-1076.

Phan, T., Kohlway, A., Dimberu, P., Pyle, A.M., and Lindenbach, B.D. (2011). The Acidic Domain of Hepatitis C Virus NS4A Contributes to RNA Replication and Virus Particle Assembly. *J. Virol.* *85*, 1193-1204.

Picard-Jean, F., Tremblay-Létourneau, M., Serra, E., Dimech, C., Schulz, H., Anselin, M., and Bisaillon, V. (2013). RNA 5'-end Maturation: A Crucial Step in the Replication of Viral Genomes. *Curr. Issues Mol. Biol*

Pichlmair, A., Schulz, O., Tan, C., Rehwinkel, J., Kato, H., Akira, S., Way, M., and Reis e Sousa, C. (2015). Activation of MDA5 Requires Higher-Order RNA Structures Generated during Virus Infection.

Pippig, D.A., Hellmuth, J.C., Cui, S., Kirchhofer, A., Lammens, K., Lammens, A., Schmidt, A., Rothenfusser, S., and Hopfner, K.P. (2009). The regulatory domain of the RIG-I family ATPase LGP2 senses double-stranded RNA. *Nucleic Acids Res.* *37*, 2014-2025.

Piston, D.W., and Kremers, G. (2007). Fluorescent protein FRET: the good, the bad and the ugly. *Trends Biochem. Sci.* *32*, 407-414.

Popescu, C., Callens, N., Trinel, D., Roingeard, P., Moradpour, D., Descamps, V., Duverlie, G., Penin, F., Héliot, L., Rouillé, Y., and Dubuisson, J. (2011). NS2 Protein of Hepatitis C Virus Interacts with Structural and Non-Structural Proteins towards Virus Assembly. *PLOS Pathogens* *7*, e1001278.

Prichard, M.N., and Shipman, C. (1990). A three-dimensional model to analyze drug-drug interactions. *Antiviral Res.* *14*, 181-205.

Protter, D.S.W., and Parker, R. (2016). Principles and Properties of Stress granules. *Trends Cell Biol.* *26*, 668-679.

Pyle, A.M. (2008). Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.* *37*, 317-336.

Pyle, A.M., Dumont, S., Serebrov, V., Cheng, W., Tinoco, I., Beran, R.K., and Bustamante, C. (2006). RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP. *Nature* *439*, 105-108.

Qian, X., Zhu, Y., Zhao, P., and Qi, Z. (2016). Entry inhibitors: New advances in HCV treatment. *Emerg. Microbes Infect.* *5*, e3.

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

Raicu, V., Jansma, D.B., Miller, R.J.D., and Friesen, J.D. (2005). Protein interaction quantified in vivo by spectrally resolved fluorescence resonance energy transfer. *Biochem. J.* **385**, 265-277.

Raicu, V., and Schmidt, W. (2017). Advanced Microscopy Techniques. In *G-protein-Coupled Receptor Dimers*, (Cham: Humana Press) pp. 39-75.

Raicu, V., and Singh, D.R. (2013). FRET Spectrometry: A New Tool for the Determination of Protein Quaternary Structure in Living Cells. *Biophys. J.* **105**, 1937-1945.

Raicu, V., Stoneman, M.R., Pisterzi, L.F., Fox, M., Wells, J.W., Jansma, D.B., Melnichuk, M., and Fung, R. (2009). Determination of supramolecular structure and spatial distribution of protein complexes in living cells. *Nat. Photonics***3**, 107-113.

Ramos, I., and Fernandez-Sesma, A. (2015). Modulating the Innate Immune Response to Influenza A Virus: Potential Therapeutic Use of Anti-Inflammatory Drugs. *Front Immunol.* **6**,

Raney, K.D., Sharma, S.D., Moustafa, I.M., and Cameron, C.E. (2010). Hepatitis C Virus Non-structural Protein 3 (HCV NS3): A Multifunctional Antiviral Target. *J. Biol. Chem.* **285**, 22725-22731.

Ranger, T. (1995). *Epidemics and Ideas: Essays on the Historical Perception of Pestilence* (Cambridge University Press).

Ranjith-Kumar, C.T., and Kao, C.C. (2006). Biochemical Activities of the HCV NS5B RNA-Dependent RNA Polymerase. In *Hepatitis C Viruses: Genomes and Molecular Biology*, Tan, Seng-Lai ed., (Norfolk (UK): Horizon Bioscience)

Rao, S.T., and Rossmann, M.G. (1973). Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **76**, 241-256.

Reddick, L.E., and Alto, N.M. (2014). Bacteria Fighting Back – How Pathogens Target and Subvert the Host Innate Immune System. *Mol. Cell* **54**, 321-328.

Rehwinkel, J., Tan, C.P., Goubau, D., Schulz, O., Pichlmair, A., Bier, K., Robb, N., Vreede, F., Barclay, W., Fodor, E., and Reis e Sousa, C. (2010). RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell* **140**, 397-408.

Rocak, S., and Linder, P. (2004). DEAD-box proteins: the driving forces behind RNA metabolism. *Nat. Rev. Mol. Cell Biol.* **5**, 232-241.

Rodriguez, K.R., Bruns, A.M., and Horvath, C.M. (2014). MDA5 and LGP2: Accomplices and Antagonists of Antiviral Signal Transduction. *J. Virol.* **88**, 8194-8200.

Saalau-Bethell, S.M., Woodhead, A.J., Chessari, G., Carr, M.G., Coyle, J., Graham, B., Hiscock, S.D., Murray, C.W., Pathuri, P., Rich, S.J., *et al.* (2012). Discovery of an allosteric mechanism for the regulation of HCV NS3 protein function. *Nat. Chem. Biol.* **8**, 920-925.

Saito, T., Owen, D.M., Jiang, F., Marcotrigiano, J., and Gale, M. (2008). Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. *Nature* **454**, 523-527.

Salam, K.A., and Akimitsu, N. (2013). Hepatitis C virus NS3 inhibitors: current and future perspectives. *BioMed Res. Int.* **2013**, 467869.

Sanchez David, R.Y., Combredet, C., Sismeiro, O., Dillies, M., Jagla, B., Coppée, J., Mura, M., Guerbois Galla, M., Despres, P., Tangy, F., and Komarova, A.V. (2016). Comparative analysis of viral RNA signatures on different RIG-I-like receptors. *Elife* **5**, e11275.

Santantonio, T., Fasano, M., Sinisi, E., Guastadisegni, A., Casalino, C., Mazzola, M., Francavilla, R., and Pastore, G. (2005). Efficacy of a 24-week course of PEG-interferon α -2b monotherapy in patients with acute hepatitis C after failure of spontaneous clearance. *J. Hepatol.* **42**, 329-333.

Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**, 430-434.

Satoh, T., Kato, H., Kumagai, Y., Yoneyama, M., Sato, S., Matsushita, K., Tsujimura, T., Fujita, T., Akira, S., and Takeuchi, O. (2010). LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 1512-1517.

Scheel, T.K.H., and Rice, C.M. (2013). Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nat. Med.* **19**, 837-849.

Schiering, N., D'Arcy, A., Villard, F., Simic, O., Kamke, M., Monnet, G., Hassiepen, U., Svergun, D.I., Pulfer, R., Eder, J., Raman, P., and Bodendorf, U. (2011). A macrocyclic HCV NS3/4A protease inhibitor interacts with protease and helicase residues in the complex with its full-length target. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 21052-21056.

Schisa, J.A., Pitt, J.N., and Priess, J.R. (2001). Analysis of RNA associated with P granules in germ cells of *C. elegans* adults. *Development* **128**, 1287-1298.

Schlee, M., Roth, A., Hornung, V., Hagmann, C.A., Wimmenauer, V., Barchet, W., Coch, C., Janke, M., Mihailovic, A., Wardle, G., *et al.* (2009). Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity* **31**, 25-34.

Schröder, M. (2011). Viruses and the human DEAD-box helicase DDX3: inhibition or exploitation? *Biochem. Soc. Trans.* **39**, 679-683.

- Scopes, R.K. (2010). Protein purification (New York, NY: Springer).
- Selvin, P.R. (1995). Fluorescence resonance energy transfer. *Meth. Enzymol.* 246, 300-334.
- Selvin, P.R. (2000). The renaissance of fluorescence resonance energy transfer. *Nat. Struct. Mol. Biol.* 7, 730-734.
- Seth, R.B., Sun, L., Ea, C., and Chen, Z.J. (2005). Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF- κ B and IRF3. *Cell* 122, 669-682.
- Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide to choosing fluorescent proteins. *Nature Methods* 2, 905-909.
- Sharma, S., and Fitzgerald, K.A. (2011). Innate Immune Sensing of DNA. *PLoS Pathogens* 7, e1001310.
- Shih, J., Wang, W., Tsai, T., Kuo, C., Li, H., and Wu Lee, Y. (2012). Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. *Biochem. J.* 441, 119-129.
- Shimomura, O. (1979). Structure of the chromophore of Aequorea green fluorescent protein. *FEBS Letters* 104, 220-222.
- Shimomura, O., Johnson, F.H., and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J. Cell Comp. Physiol.* 59, 223-239.
- Shirota, Y., Luo, H., Qin, W., Kaneko, S., Yamashita, T., Kobayashi, K., and Murakami, S. (2002). Hepatitis C virus (HCV) NS5A binds RNA-dependent RNA polymerase (RdRP) NS5B and modulates RNA-dependent RNA polymerase activity. *J. Biol. Chem.* 277, 11149-11155.
- Simmonds, P. (2004). Genetic diversity and evolution of hepatitis C virus--15 years on. *J. Gen. Virol.* 85, 3173-3188.
- Sklan, E.H., and Glenn, J.S. (2006). HCV NS4B: From Obscurity to Central Stage. In *Hepatitis C Viruses: Genomes and Molecular Biology*, Tan, Seng-Lai ed., (Norfolk (UK): Horizon Bioscience)
- Snapp, E. (2005). Design and use of fluorescent fusion proteins in cell biology. *Curr. Protoc. Cell Biol.* Chapter 21, Unit 21.4.
- Stoneman, M.R., Paprocki, J.D., Biener, G., Yokoi, K., Shevade, A., Kuchin, S., and Raicu, V. (2017). Quaternary structure of the yeast pheromone receptor Ste2 in living cells. *Biochim. Biophys. Acta* 1859, 1456-1464.

Subramanya, H.S., Bird, L.E., Brannigan, J.A., and Wigley, D.B. (1996). Crystal structure of a DExx box DNA helicase. *Nature* *384*, 379-383.

Takeuchi, O., and Akira, S. (2010). Pattern Recognition Receptors and Inflammation. *Cell* *140*, 805-820.

Tsien, R.Y. (1998). The Green Fluorescent Protein. *Annu. Rev. Biochem.* *67*, 509-544.

Uchikawa, E., Lethier, M., Malet, H., Brunel, J., Gerlier, D., and Cusack, S. (2016). Structural Analysis of dsRNA Binding to Anti-viral Pattern Recognition Receptors LGP2 and MDA5. *Mol. Cell* *62*, 586-602.

Upadya, M.H., Aweya, J.J., and Tan, Y. (2014). Understanding the interaction of hepatitis C virus with host DEAD-box RNA helicases. *World J. Gastroenterol.* *20*, 2913-2926.

Valiente-Echeverría, F., Hermoso, M.A., and Soto-Rifo, R. (2015). RNA helicase DDX3: at the crossroad of viral replication and antiviral immunity. *Rev. Med. Virol.* *25*, 286-299.

Velankar, S.S., Soutanas, P., Dillingham, M.S., Subramanya, H.S., and Wigley, D.B. (1999). Crystal Structures of Complexes of PcrA DNA Helicase with a DNA Substrate Indicate an Inchworm Mechanism. *Cell* *97*, 75-84.

von Stetten, D., Noirclerc-Savoye, M., Goedhart, J., Gadella, T.W.J., and Royant, A. (2012). Structure of a fluorescent protein from *Aequorea victoria* bearing the obligate-monomer mutation A206K. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* *68*, 878-882.

Welbourn, S., and Pause, A. (2007). The hepatitis C virus NS2/3 protease. *Curr. Issues. Mol. Biol.* *9*, 63-69.

Welsch, C., Jesudian, A., Zeuzem, S., and Jacobson, I. (2012). New direct-acting antiviral agents for the treatment of hepatitis C virus infection and perspectives. *Gut.* *61 Suppl 1*, 36.

Wodarz, D., and Lloyd, A.L. (2004). Immune responses and the emergence of drug-resistant virus strains in vivo. *Proc. Biol. Sci.* *271*, 1101-1109.

Wu, B., and Hur, S. (2015). How RIG-I like receptors activate MAVS. *Curr. Opin. Virol.* *12*, 91-98.

Wu, B., Peisley, A., Richards, C., Yao, H., Zeng, X., Lin, C., Chu, F., Walz, T., and Hur, S. (2013). Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. *Cell* *152*, 276-289.

Wu, B., Peisley, A., Tetrault, D., Li, Z., Egelman, E.H., Magor, K.E., Walz, T., Penczek, P.A., and Hur, S. (2014). Molecular imprinting as a signal-activation mechanism of the viral RNA sensor RIG-I. *Mol. Cell* *55*, 511-523.

- Yao, N., Reichert, P., Taremi, S.S., Prosis, W.W., and Weber, P.C. (1999). Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* 7, 1353-1363.
- Ye, J., Osborne, A.R., Groll, M., and Rapoport, T.A. (2004). RecA-like motor ATPases—lessons from structures. *BBA - Bioenergetics* 1659, 1-18.
- Yedavalli, Venkat S R K, Neuveut, C., Chi, Y., Kleiman, L., and Jeang, K. (2004). Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. *Cell* 119, 381-392.
- Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004a). The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat. Immunol.* 5, 730-737.
- Yoneyama, M., Kikuchi, M., Matsumoto, K., Imaizumi, T., Miyagishi, M., Taira, K., Foy, E., Loo, Y., Gale, M., Akira, S., *et al.* (2005). Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *J. Immunol.* 175, 2851-2858.
- Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004b). The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat. Immunol.* 5, 730-737.
- Yoneyama, M., Onomoto, K., Jogi, M., Akaboshi, T., and Fujita, T. (2015). Viral RNA detection by RIG-I-like receptors. *Curr. Opin. Immunol.* 32, 48-53.
- Zacharias, D.A., and Tsien, R.Y. (2006). Molecular biology and mutation of green fluorescent protein. *Methods Biochem. Anal.* 47, 83-120.
- Zacharias, D.A., Violin, J.D., Newton, A.C., and Tsien, R.Y. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* 296, 913-916.
- Zapata-Hommer, O., and Griesbeck, O. (2003). Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP. *BMC Biotechnol.* 3, 5.
- Zhang, Z., Kim, T., Bao, M., Facchinetti, V., Jung, S.Y., Ghaffari, A.A., Qin, J., Cheng, G., and Liu, Y. (2011). DDX1, DDX21, and DHX36 helicases form a complex with the adaptor molecule TRIF to sense dsRNA in dendritic cells. *Immunity* 34, 866-878.

CURRICULUM VITAE

Megan Josephine Corby

EDUCATION

University of Wisconsin – Milwaukee

Milwaukee, WI

Ph.D. Chemistry and Biochemistry

Graduation: Spring 2018

GPA: 3.95

Areas of study: High throughput drug discovery, protein biochemistry, protein-protein interactions, and multi-photon FRET microscopy

Post-baccalaureate – Focus on Chemistry and Biochemistry

2009 - 2012

GPA: 3.91

Dean's Honor List

Relevant coursework: Organic Chemistry 343 and 345, Inorganic Chemistry 311, Genetics 325, Intro to Biochemistry 501, Biophysical Chemistry 560

Portland State University

Portland, OR

Post-baccalaureate – Focus on Chemistry and Biology

2008 - 2009

University of Wisconsin – Madison

Madison, WI

Bachelor of Arts

2002 - 2006

GPA: 3.35

Areas of study: Women's health and LGBT studies

RESEARCH EXPERIENCE

Graduate Research Assistant

Milwaukee, WI

Department Chemistry and Department of Physics – UW-Milwaukee

2012 – 2018

- Developed expression system for fluorescently tagged human and hepatitis C virus recombinant proteins in human cell lines
- Designed Quantitative Micro-Spectroscopic Imaging: a multi-photon excitation FRET assay with spectral resolution to analyze oligomeric interactions of fluorescently tagged human and hepatitis C virus recombinant proteins expressed in human cells
- Performed high-throughput in-vitro protease and helicase assays to screen for drugs targeting the hepatitis C viral helicase/protease
- Identified lead compounds from in-vitro assays and screened them in cell based assays against a hepatitis C virus subgenomic replicon for efficacy and for cytotoxicity
- Purified proteins using BioRad and Akta protein purification systems
- Experience with RNA extraction and RT-qPCR to study effects of lead compounds on cell RNA expression
- Designed and sub-cloned genes into human and protein expression vectors

- Supervised, mentored, and taught over twenty graduate students, undergraduate students, and high school students in the lab

Undergraduate Research Assistant

Milwaukee, WI

Department of Biology – UW-Milwaukee

2010 - 2012

- Performed metabolism assays on yeast and extracted protein for Western blots
- Investigated growth and heat tolerance of quiescent yeast

Great Lakes WATER Institute – UW- Milwaukee

2009 - 2010

- Set up laser and camera imaging system to capture the dynamics of microscopic water organisms

PUBLICATIONS

- **Corby, M.J.**, Stoneman, M., Beiner, G., Paprocki, J.D., Kolli, R., Raicu, V., Frick, D.N. (2017) Quantitative Micro-Spectroscopic Imaging Reveals Viral and Cellular RNA Helicase Interactions in Live Cells. *Journal of Biological Chemistry*. *In press*
- Ndjomou, J., **Corby, M.J.**, Sweeney, N.L., Hanson, A.M., Aydin, C., Ali, A., Schiffer, C.A., Li, K., Frankowski, K.J., Schoenen, F.J., Frick, D.N. (2015) ***Simultaneously targeting the NS3 protease and helicase activities for more effective hepatitis C therapy***. *ACS Chem Biol*. 10(8), 1887-96.

TEACHING AND MENTORING EXPERIENCE

Mentor to high school SMART teams

Milwaukee, WI

Milwaukee School of Engineering Center for Bio-molecular Modeling

2014 – 2017

- Advisor for SMART (Students Modeling A Research Topic) teams
- Assist high school SMART teams in identifying and carrying out research on the structure and function of a protein of interest to aid in the preparation of physical bio-molecular models and a poster and presentation at a national conference.

Teaching assistant, Chemistry 603 – Introduction to Biochemistry Lab

Milwaukee, WI

Department of Chemistry – UW-Milwaukee

2014 - 2015

- Prepared laboratory experiments for a class of ten undergraduate and graduate students including PCR, restriction enzyme digests and ligation, E.coli transformation, protein expression and purification, and in-vitro fluorescent enzyme assays
- Received exemplary student and faculty reviews

Teaching assistant, Chemistry 100 – Chemical Science

Milwaukee, WI

Department of Chemistry – UW-Milwaukee

2012 - 2014

- Prepared discussion lectures for over 100 students each semester
- Covered basic chemical concepts including electronic structure of an atom, chemical bonding, reaction rates and equilibrium, REDOX reaction, and chemical naming

Teaching assistant, Chemistry 103 – Survey of Biochemistry for Nurse Milwaukee, WI
Department of Chemistry – UW-Milwaukee 2014

- Prepared discussion lectures for 50 students
- Covered identifying, naming, and drawing basic organic functional groups, chemical reactions and equilibria, basic organic chemistry as it applies to the healthcare field, metric conversions, chemical bonding
- Ran a laboratory where classroom principals were tested in a way applicable to healthcare

Research Assistant, Classroom Salon – Online social learning forum Milwaukee, WI
Department of Chemistry – UW- Milwaukee 2012

Carnegie Mellon University

The International Society for Technology in Education

- Project entitled “A Socially-Centric Blended Learning Model for At Risk Youth in an Urban Institution.”
- Tested Classroom Salon and communicated with principal investigator to improve student outcomes
- Integrated insights gained from online classroom learning into classroom discussions

Mentor to incoming chemistry department teaching assistants Milwaukee, WI
University of Wisconsin – Milwaukee 2012 – 2013

- Provided mentor support to new teaching assistants
- Observed classroom management, communication skills, and level of preparedness
- Offered feedback throughout the new teaching assistant’s first semester
- Formally reviewed each teaching assistant at the close of the semester

Teaching assistant, Chemistry 104 – General Chemistry II Milwaukee, WI
Department of Chemistry – UW-Milwaukee 2012

- Prepared discussion lectures for over 50 students
- Taught chemistry concepts such as: chemical kinetics, equilibria, acid-base chemistry, and chemical thermodynamics
- Ran laboratory
- Created and administered quizzes to gauge student learning

HONORS AND AWARDS

Advanced Opportunity Fellowship
Graduate School, University of Wisconsin - Milwaukee 2015 – 2018

- Twenty-five awards given to the top students pursuing graduate degrees

Alice C. Helland Fellowship
Greater Milwaukee Foundation 2015

- Honors top-performing, non-traditional female students pursuing an undergraduate or graduate degree at a University of Wisconsin school.

McFarland Graduate Research AwardDepartment of Chemistry – UW-Milwaukee

2014

- Exceptional poster and presentation

First Place in Liberal Arts Essay ContestUniversity of Wisconsin System

2012

- Top essay regarding the benefits of a liberal arts education.

NATIONAL CONFERENCE PRESENTATIONS

- Corby, M.J., Stoneman, M., Biener, G., Raicu, V., Frick, D.N. Ligand induced changes in the quaternary structure of the LGP2/MDA5 signaling complex. In: Midwest Enzyme Chemistry Conference.; 2016; University of Illinois-Chicago, Chicago, IL.
- Corby, M.J., Biener, G., Raicu, V., Frick, D.N. FRET studies examining the interaction of hepatitis C NS3 helicase and human DExD/H box helicases . American society for biochemistry and molecular biology national meeting; 2016; San Diego, CA.
- Corby, M.J., Biener, G., Raicu, V., Frick, D.N. The hepatitis C virus helicase interacts with human pattern recognition receptor LGP2. In: 3rd Biennial symposium: Optical micro-spectroscopy and molecular imaging.; 2015; University of Wisconsin Milwaukee, Milwaukee, WI.
- Mahim, A., Corby, M.J., Rajesh, K., Biener, G., Raicu, V., Frick, D.N. Probing the possible interaction between hepatitis C virus NS3 helicase and host DExD/H box helicases using fluorescent fusion proteins. In: Midwest Enzyme Chemistry Conference.; 2014; Northwestern University, Evanston, IL.
- Corby, M.J., Ndjomou, J., Hanson, A., Frick, D.N. Effect of combining helicase and protease inhibitors on hepatitis C NS3-4A. In: American Chemical Society 247th National Meeting; 2014; Dallas, TX.
- Corby, M.J., Ndjomou, J., Hanson, A., Frick, D.N. Targeting a multifunctional protein: Effects of protease and helicase inhibitors on hepatitis C NS3 protease/helicase. In: Midwest Enzyme Chemistry Conference; 2013; Loyola University, Chicago, IL.

UNIVERSITY ORGANIZATIONS

Bio-physics Graduate Student Council**Co-chair**

2016 - 2018

- Coordinate UWM -Biophysics 2017 Conference in conjunction with local and national faculty.

University of Wisconsin-Milwaukee – Media Services 2017
Intern
• Interviewing and writing scientific stories to promote the university and its faculty and students

Chemistry and Biochemistry Graduate Student Council
Co-chair 2014 - 2015
• Designed and coordinated annual Chemistry Awards Day
• Grant writer to fund awards day, travel awards, and student council operation costs

UNIVERSITY SERVICE

Co-chair: 4th Biennial Symposium: Optical Microspectroscopy and Molecular Imaging 2017
• Coordinate and organize conference
• Invite and host speakers from around the world

UW-Milwaukee Graduate School Ambassador 2015 - 2017
• Present research to potential donors