DISSERTATION

DODGING WRENCHES IN THE TIME OF COVID: EXPLORING FLAVIVIRUS REPLICATION MECHANISMS AND SARS-COV-2 ANTIBODY DEVELOPMENT

Submitted by

James Steven Terry

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2024

Doctoral Committee:

Advisor: Brian J. Geiss

Jeffrey Wilusz
Gregory Ebel
Christopher Snow

ABSTRACT

DODGING WRENCHES IN THE TIME OF COVID: EXPLORING FLAVIVIRUS REPLICATION MECHANISMS AND SARS-COV-2 ANTIBODY DEVELOPMENT

Flaviviruses pose a significant threat to global health, threatening hundreds of millions of people who live in endemic areas. Infection with flaviviruses such as dengue virus (DENV), Zika virus (ZIKV), and West Nile virus (WNV) can trigger symptoms ranging from a mild cold-like illness to microcephaly, encephalitis, hemorrhagic fever, and death. As climate change alters global temperature ranges, habitable environments for the flavivirus arthropod vectors are expanding into previously unexposed regions. Due to a lack of flavivirus vaccines and antivirals, most efforts to combat infection fall within vector population control and palliative care for infected individuals. To develop antivirals and vaccines against flaviviruses, we need to better understand the fundamental mechanisms through with the viruses replicate. By investigating incompletely understood processes in the replication cycle, new antiviral targets can be identified and pursued. This dissertation investigates two components of the flavivirus replication cycle to better understand the key processes necessary for successful flavivirus infection. Additionally, this dissertation reports on efforts during the SARS-CoV-2 COVID-19 pandemic to develop novel reagents to assist in research and diagnostic development.

An important determinant of successful flavivirus infection is the generation of subgenomic flavivirus RNA (sfRNA). This RNA is composed of exoribonuclease resistant RNA (xrRNA) structures in the flavivirus 3’ untranslated region (3’UTR). These structures allow the flavivirus 3’UTR to withstand degradation by stalling the host Xrn1 exoribonuclease, halting viral RNA degradation and creating sfRNA. The production of sfRNA is critical for flavivirus replication success as the new RNA entity actively suppresses the host cell immune response to viral infection. There are blind spots in our understanding of the key stages of sfRNA generation, namely how the Xrn1 substrate is produced from the flavivirus genome. It has previously been postulated that host decapping enzymes remove the flavivirus Type 1 cap structure, allowing Xrn1 to bind to the 5’ monophosphate and degrade the viral RNA. The enzyme responsible for decapping has not yet been identified. Following preliminary evidence from the Geiss Lab, we investigated the host decapping enzyme Dcp2 as the protein responsible for priming flavivirus RNA for Xrn1 degradation and sfRNA production. We developed a pipeline using splint-ligation to specifically label monophosphorylated WNV RNA with an RNA adapter at the 5’ end. Following this the ratio proportion of viral RNA that is monophosphorylated is revealed using a qRT-PCR reporting system. With this pipeline, it was determined that suppressing Dcp2 expression increased the proportion of monophosphorylated WNV RNA in infected cells while having no significant effect on monophosphorylated RNA in newly produced virions. Additionally, northern blot analysis revealed that sfRNA generation was not reduced by Dcp2 knockdown. From this study we determined that Dcp2 is not necessary for sfRNA generation, and thus other

processes are responsible for the generation of monophosphorylated viral genomes for Xrn1 degradation.

One hole in our understanding of flavivirus replication concerns the viral replication compartment. The compartment is an invagination in the endoplasmic reticulum membrane that is formed through viral protein manipulation. This environment then hosts the viral replication machinery, protecting the vulnerable viral RNA from host cell immune detection as new viral genomes are produced. Proper viral protein-protein interactions are critical for the successful formation of this viral RNA factory. While studies have been conducted to determine the replication compartment location and some interactions between nonstructural proteins, our understanding of how these proteins interact with each other *in situ* is limited. To address this, we employed crosslinking mass spectrometry. First, a flavivirus replication compartment purification and crosslinking pipeline underwent a series of evolutions and significant optimizations followed mass spectrometry data acquisition. Then, a crosslinked protein analysis pipeline using the Bonvin Lab programs DisVis/HADDOCK was validated with crosslinked bovine serum to ensure its utility with crosslinked viral compartment samples. MaxQuant analysis revealed some viral protein crosslinks while highlighting areas for improvement in our methodology. Nevertheless, the identified intramolecular crosslinks within NS1, NS3, and NS5 hint at potential dimer interfaces. An intermolecular crosslink between NS3 and NS4b was identified that supports the observations of previous studies while establishing *in situ* evidence for interactions along an NS3 N-terminus and NS4b residue K172 interface. The results provided intriguing preliminary evidence for future investigations into the replication compartment protein-protein interactions and established a protocol for analyzing viral proteins with crosslinking mass spectrometry.

In addition to chronicling on flavivirus replication cycle studies this dissertation includes a chapter chronicling work during the COVID-19 pandemic on SARS-CoV-2. Monoclonal antibodies targeting the SARS-CoV-2 nucleocapsid protein were generated, characterized, and sequenced during the height of the pandemic. These antibodies were the first of their kind to be published and were made available for use during the global SARS-CoV-2 research effort. This chapter also reports on collaborative efforts surrounding the use of antibodies for diagnostics and predictive computational pipelines. Work was done to assist the Henry Lab in developing inexpensive electrochemical and colorimetric ELISA devices targeting SARS-CoV-2 NP for bedside diagnostic use. Lastly, wet lab verification was performed to validate Jacob Deroo's epitope-predicting PAbFold AlphaFold2 pipeline.

The work covered in this dissertation spans five years, two viruses, and three separate target areas. These projects, while varied, are all bound together by the common goal of contributing to the advancement of knowledge and techniques for stopping viral threats to global health. Knowing how a virus creates a safe environment for genome replication or identifying which host proteins help create an immune-modulating viral RNA molecule is important for identifying new paths towards intelligently designed antivirals. Similarly, developing and characterizing antibodies to supply a global research effort and validate cutting-edge computational tools is necessary for actively combatting a global pandemic and preparing for the next one. With this work, scientific inquiry ranging from foundational knowledge to translational science is explored.

ACKNOWLEDGEMENTS

A lot can happen during a PhD. In the last five-years I have experienced biblical plagues ripped right out of the Old Testament, a life-threatening illness that redefined my baseline, and set-back after set-back in my scientific pursuits. But I am immensely blessed to have been surrounded by people who have helped me persevere and succeed.

First, I must thank Brian. One could not ask for a better PI and mentor to shepherd me through the janky roller coaster ride of graduate school. Everyone lucky enough to work for Brian can attest to the patience he has and the wisdom he provides. Brian has taught me to be a good scientist, one who not only relishes their scientific curiosity but possesses the skills required to pursue their ideas. At my darkest point he showed tremendously more understanding and patience than was required of someone in his position, allowing me to take the time I needed to recover. I would not be the scientist I am today without his help.

I am incredibly appreciative of the people who helped shape me during my professional development. Thank you to my committee members Dr. Greg Ebel, Dr. Chris Snow, and Dr. Jeff Wilusz for helping guide my graduate school journey. Their technical advice, enthusiasm, and mentorship significantly contributed to the success of my PhD studies. Thank you to Dr. Florian Busch for taking the time to assist me in the mass spectrometry data analysis pipeline. Thank you to the scientists at the NeuroMab UC Davis/NIH facility, especially Camelia Dumitras and Dr. Belvin Gong. These people helped me get my start in science and taught me the technical skills I would later use every day. Thank you to the professors of the 2018-2019 MIP Masters B Program who gave me a strong foundational education into the higher levels of microbiology. Additionally, thank you to Glenn Telling and the Telling Lab for helping me get my first taste of academic research, ultimately leading me to where I am today.

One of the strengths of the CSU MIP department is the collaborative and helpful nature of its members. This departmental idiosyncrasy helped my work significantly, as various labs were critical for the success of my projects. Thank you to Dr. Karen Dobos and the Dobos Lab for their training and support during the mass spectrometry portion of my dissertation. Karen swooped in to save the day when it seemed like the project was dead in the water, and I am forever grateful for her generosity. Members of the Dobos lab kindly aided this process, including Megan Lucas and Kimberly Shelton. I would especially like to thank Dr. Luisa Nieto-Ramirez who provided invaluable expertise, advice, and assistance. Additionally, thank you to John Anderson and the Wilusz Lab for training me on radioactivity assays and assisting in the safe and proper completion of such work. These labs provided key resources and training that fostered my development as a scientist, and for that I am truly grateful.

I also must thank various CSU colleagues for their help during my PhD. Thank you to the past and current members of the Geiss Lab for their support over the years. Especially, I would like to thank Denis for his preliminary work for my project and Kelly (Du Pont) Zuccaro for her mentorship and friendship. Additionally, thank you to the other members of the lab have shaped the direction of my PhD, such as Michael Mingroni, Chaisty Trammell, Handly Mayton, Lauren Malsik, Marina Fuji, Kristen Denale, Brooke

Enney, Amanda Bartels, Loran Anderson, and Davis Moline. To my fellow 2019 MIP PhD cohort, we had a unique experience with our PhD. Thank you for helping me get through it, especially Emily “Fitz” Fitzmeyer, Sarah Cooper, Darby Gilfillan, and Katherine Kokkinias.

Thank you to the administrators who helped guide this clueless punk through the bureaucracy of graduate school, especially Heidi Runge, Stephie Zellner, Kailee Martin, and Christina Charles. Special thanks to Dr. Zaid Abdo, who worked tirelessly during his time as the Associate Department Head of Graduate Education to enact reforms that improved the quality of life for MIP students. The people in the MIP administration team do a lot of thankless work to make people’s lives better in the department, especially the students, and for that I thank you all from the bottom of my heart.

Thank you to the Optima XPN-80 Ultracentrifuge for rescuing me from having to operate your predecessor, the sulfurous, screeching ultracentrifuge known only as the Devil’s Washing Machine. Additionally, to the vending machine in the Microbiology Building loading dock, thanks for keeping me alive during the lockdown.

The various families I belong to have been instrumental in helping me succeed during this work. Thank you, mom, dad, Anna, Mikey, Shane, and Heath, (yes you, ya little munchkin) for your love and support throughout this journey. To the Fletcher family, words can’t express my appreciation for y’all welcoming me in as one of your own and encouraging me in my studies. To my California family (Brian, Riley, Weili, Meredith, Karan, Sarah, and Matt) I appreciate your love, friendship, and company over the last six years since I left. Additionally, thank you to my hānai family Doug, Crissy, Madi, and the rest of the Gayagas-Tostrud clan for accepting me and teaching me to be a better man who could succeed in the difficult task that is graduate school.

A huge thank you to my “Fort Collins family”, a group who over the last years have shared in the good times and helped me survive the bad. To my D&D groups, I appreciate you being the best players a DM could ask for- creating stories with you has kept me sane and helped me maintain perspective on life outside of the grind. Pete Justice is truly a one-of-a-kind man who mentored me over the years and has inspired me to be a better person that leaves a lasting, positive effect on the community around me. August Luc is my serendipitous COVID lockdown accounta-bila-buddy and has since become one of my best friends, helping me see the forest for the trees and become a more active thinker about the world outside of science. Thank you to Amy Fox, Camron Pearce, and Charlotte Avanzi for welcoming me and providing a strong network of support that eventually led to me meeting my future wife. To Katherine, Ethan Dunn, Greg Pugh, Angie Pugh, and Sylvie Pugh, thank you for providing that sense of home and belonging that has always eluded this Army brat- y’all are my family, and I cannot describe how much your love and support over the years has meant to me.

Lastly, thank you to McKenzie Lee Fletcher, without whom nothing in this dissertation would have been possible. For the last three years McKenzie has been a sympathetic ear, a trusted confidant, and at times a much-needed reality check. McKenzie has been a constant reminder of what is important in life and has helped me make it through the bad times when it seemed like I was not going to. I am forever grateful for her support, love, and terrible taste in men.

TABLE OF CONTENTS

CHAPTER 1 — A REVIEW OF THE CRITICAL VIRAL NONSTRUCTURAL PROTEINS AND SUBGENOMIC FLAVIVIRUS RNA NECESSARY FOR THE FLAVIVIRUS REPLICATION CYCLE

1.1 Epidemiology and Transmission

Flaviviruses are a genus of the *Flaviviridae* virus family that consists of positive-sense single-stranded RNA arboviruses that rely upon an arthropod vector for transmission, often mosquitos and ticks. These viruses are responsible for significant morbidity and mortality the world over; with the viruses' endemic regions corresponding to their vectors' survivable habitats, a significant portion of the world's population is under threat of at least one of the known 53 flavivirus species. The significant health cost along with the considerable economic cost of flavivirus infections, disease, and control methods adding up to billions of US dollars every year in loss of productivity, output, and time [1-3].

The vast majority of flaviviruses utilize an arthropod vector for transmission to a new host. In an avian or mammalian host, the virus replicates and is introduced to a new arthropod host, often in the form of a bloodmeal. In the mosquito or tick host, the virus replicates and migrates to the salivary glands, ready for the next round of infection with the arthropod vector's next bloodmeal. The vectors responsible for transmitting the majority of flaviviruses tend to fall into two categories- ticks or mosquitos. Flavivirus-carrying mosquitos primarily hail from the *Anopheles, Aedes,* and *Culex* genera*,* and are the most prolific of the flavivirus vectors [4-6]. Their habitat distribution encompasses most of the human-occupied regions of the globe, enabling effective human infection due to high incidences of human-vector contact. Most mosquito-borne flaviviruses utilize a zoonotic reservoir consisting of local avian and mammalian fauna, complicating viral control and eradication methods [7]. Many third-world nations lack the infrastructure to effectively employ mosquito control measures and are often hit hardest by flavivirus infections, where standing water in urban environments, domicile permeability to vectors, and proximity to avian hosts contributes to high infection rates [8].

Several flaviviruses have been of importance for global health, both historically and in the modern age. Dengue virus (DENV) threatens a significant portion of the world's population, with almost 400 million symptomatic and asymptomatic cases occurring every year [4]. Primarily targeting tropical and subtropical regions, DENV is shepherded from host to host by *Aedes albopictus* and *Aedes aegypti.* DENV infection manifestations can range from the relatively mild symptoms of dengue fever to the traumatic and life-threatening dengue hemorrhagic fever and dengue shock syndrome, depending on sequential infection by DENV's five different serotypes [9-11]. These repeated infections increases the risk of poor patient outcomes by antibody-dependent enhancement (ADE), where increased antibody titers from previous infections results catastrophic damage to the host upon new infection [12]. Zika virus (ZIKV) is spread by mosquitos from the *Aedes* genus, tracing its origins from the Zika Forest in Uganda [13]. Originally identified in 1947, Zika virus has since spread steadily across the globe thanks to steadily increasing shipping and air travel [14]. In 2015, ZIKV became widely known in the western hemisphere with an epidemic in central and south America. Like the other flaviviruses, Zika disease symptoms range from the mild cold-like symptoms and the severe Guillain-Barre Syndrome and fetal microcephaly [15-18]. West Nile virus (WNV) is spread by *Aedes* and *Culex* mosquito species using avian hosts as a reservoir, while humans and horses are dead-end hosts [19]. Originally found only in the eastern hemisphere, WNV made landfall in the United States in 1999 [20-22]. Since 1999 WNV has spread across the North American continent, causing febrile illness and, in extreme cases, encephalitis [23]. Today WNV is the most prevalent flavivirus in the continental United States with many research studies into surveillance and vector control being funded by the federal government.

Compounding the threat of these viruses is the lack of effective vaccines, antivirals, and therapeutic compounds. A vaccine exists for humans against yellow fever virus (YFV) infection and another for WNV in horses, but the remaining major flaviviruses have limited prophylactic measures to prevent infection [24-26]. For active infections, palliative care is the only option for patients, as therapeutics to deter viral replication are absent. Some strides have been made in the last few years in developing new countermeasures, with some holding true promise. The inhibitor JNJ-1802 entered clinical trials in 2023, offering pan-serotype prophylactic protection from viral replication and suppression of active viral replication [27]. Still, the JNJ-1802 inhibitor is only in the early clinical trial stages and may prove ineffective following human testing as so many potential therapeutics have done before.

With a lack of currently available therapeutics and vaccines, most protocols for combatting flavivirus infections center around controlling interactions with arthropod vectors. Most vector control methods rely on the use insecticide spraying, habitat reduction, and prophylactic measures combatting blood feeding such as window screens and bed netting. New developments in vector control have been seeing increased research activity in the past few decades. Mosquito population control programs using ivermectin mass drug administration are currently undergoing trials due to the lethal effect ivermectin has on arthropod hosts of *Plasmodium* species and arboviruses [28, 29]. The biological agent *Wolbachia pipientis* has also been investigated for its unique abilities in regard to arthropod vectors. This bacterial endosymbiont infects invertebrate host cells and has shown promise in controlling arbovirus transmission and propagation. Infection with the *Wolbachia* endosymbiont is passed on from female mosquito to its offspring via vertical transmission, granting protection from specific insect-targeting viruses while simultaneously causing infertility between mating pairs. This occurs when one member of a mating pair lacks *Wolbachia* infection or differs in which strain it possesses from its reproduction partner- this ensures both fitness-driven propagation of the control method and a reduction in vector populations [30-32]. Additionally, numerous studies have been reported chronicling the inhibitory effect *Wolbachia* has had on vertebrate-targeting pathogen propagation in the invertebrate host, whether they be viruses or parasites such as *Plasmodium* species [33-37].

Altogether, the flavivirus species and their vectors pose a significant threat for global health. With few effective vaccine regimens to protect from infection and even fewer patient treatment options, it is important to look deeper into the foundational mechanisms these pathogens use to replicate themselves in a mammalian host. In understanding how flaviviruses operate in these hosts to establish a successful infection and replicate their genomes, we can identify new targets for antivirals or vaccines to alleviate the considerable burden they have on global health. In this review, we will offer an overview of flavivirus biology, specifically the multifunctional viral nonstructural proteins, the protective replication compartments, and the subgenomic flavivirus RNA necessary for proviral immunomodulation. These agents of flavivirus replication are critical for successful infection, and thus their understanding is critical for paving new paths towards therapeutic development.

1.2 Flavivirus Biology

1.2.1 The Flavivirus Genome

Flaviviruses are single stranded, positive sense enveloped viruses with an ~11 kilobase genome. The viral genome encodes all the necessary components to kickstart an infection event. The viral RNA (vRNA) resembles a host messenger RNA (mRNA), with a methylated guanine cap and a single long reading frame that generates a polypeptide with the precursors to all the virus' proteins Although, the vRNA 3' end lacks the characteristic polyadenylated tail found on mRNA [38]. The type 1 cap structure possesses a guanine N-7 methylation and 2'-O methylation on the first nucleotide base position [39-41]. The 5'RNA cap structure is critical for translation initiation and aids in host innate immune evasion, as host pattern recognition receptors (PRRs) and immune enzymes are able to identify uncapped or improperly-capped RNA to trigger the immune response signal cascade and RNA degradation pathway [41-43]. The viral RNA exists in three primary forms: the single stranded, capped genome that is packaged in virions or translated, a double-stranded RNA intermediate, and a partially nuclease-resistant subgenomic flavivirus RNA (sfRNA) form that participates in host immune modulation [44-50]. The 5’ and 3’ untranslated regions (UTR) adopt specific structures that aid in translation, transcription, and immune modulation. Structures within the 5’ UTR such as stem loop A (SLA) act to promote translation by enhancing ribosome binding, as well as associating with the viral polymerase to facilitate circularization and replication [51]. The 3’ UTR contains structural motifs that are present in all the arthropod-borne flaviviruses [52, 53]. These manifest in forms known as pseudoknots and dumbbells based on their three-dimensional folding patterns, and these structures are sometimes duplicated in flaviviruses to bestow additional fitness that is normally blunted when switching between vertebrate and invertebrate hosts [54-56]. These structures are required for viral genome transcription and efficient protein translation [57].

1.2.2 The Flavivirus Proteins

The flavivirus genome encodes a single, long polypeptide that contains all ten main flavivirus proteins (**Figure 1.1**). The first three proteins produced are structural proteins used to construct new virions prior to genome packaging: the capsid protein (C), the pre-membrane protein (prM), and the envelope protein (E). The remaining proteins are heavily involved in genome replication, immune modulation, and virion formation.

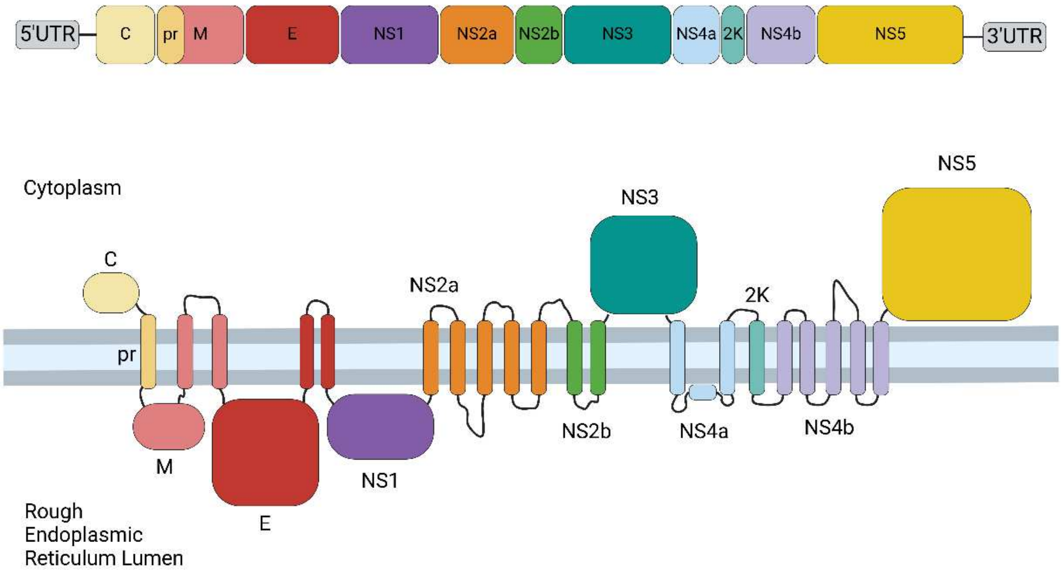


Figure 1.1. Generic flavivirus genome and polyprotein. (**Top**) Flavivirus genomes consist of an 11kb single-stranded positive-sense RNA strand. The single open reading frame of the genome produces a long polypeptide containing the flavivirus structural and nonstructural proteins (**Bottom**). Created with BioRender.com.

1.2.2.1 The Flavivirus Nonstructural Proteins

The seven main nonstructural proteins (NS) are responsible for rearranging host membranes, replicating the viral genome, assembling new virions, and running interference against the host immune response. In translation order on the single polyprotein, they are NS1, NS2a, NS2b, NS3, NS4a, NS4b, and NS5. There is another small peptide formed from the polyprotein; 2K is the lesser known and often ignored 8th nonstructural protein nestled between NS4a and NS4b. The nonstructural proteins are tasked with producing new viral genomes and inhibiting the host's ability to mount an effective defense against a developing infection. Many adhere to endoplasmic reticulum (ER) and Golgi structures, although some have been seen translocating toother organelles such as the nucleus [58]. These proteins all carry multiple functions and contribute towards the generation of new flavivirus genome copies or virion assembly.

NS1 exists on the luminal side of the ER and has been observed manipulating host membranes prior to genome production to facilitate replication and virion formation [59-62]. NS1 will quickly dimerizes following cleavage, form tetramers, or bind into a hexamer when secreted from a lysed cell [61, 63]. The self-association is regulated by environmental conditions, namely pH [63]. The protein's dimer form is known to associate heavily with lipids and its role in affecting host membrane remodeling has been observed; exposure of liposomes to solubilized NS1 results in rapid protein coating and the breakup of the liposomes into smaller lipid particles [60]. This in conjunction with early subcellular localization experiments supports the theory that NS1 participates in the formation of RER invaginations where flavivirus genome replication occurs, also known as replication compartments [62]. Additionally, studies concerning

NS1 supplied in *trans* to different flavivirus species showed that the NS1 and NS4a proteins possess a necessary interaction, as DENV NS1 was unable to rescue replication in NS1-deficient YFV [59, 64]. NS1 interactions with NS4b have critical implications in the replication process, as a complementary mutation in NS4B rescued replication in WNV NS1 mutants possessing amino-acid substitutions in the N-terminal protein targeting motif. NS1 has also been identified as binding to prM and E proteins when assembling new virion particles [65, 66], with human antibodies targeting WNV NS1 protecting from infection. The repeated observations of NS1's necessity for replication and its relatively accessible nature indicates the value NS1 holds as a potential therapeutic target in further studies.

NS2a is one of the flavivirus transmembrane proteins that works to both modulate the host immune response and assist in genome replication. This protein targets the innate host immune response to flavivirus infection by triggering STAT1/2 signaling proteins for proteasomal degradation and by suppressing RIG-I and IRF3 [67-69]. For viral replication, NS2a helps mediate membrane reorganization along with NS1 [70] and binds to assembling structural proteins and the 3' UTR of the flavivirus genome [71, 72]. The sequence can vary significantly between flavivirus species, primarily with the membrane-transversing segments [73]

Critical for polyprotein processing is NS2b, a transmembrane viral protein that is vital for the infection cycle. Arguably the most studied function of NS2b is its ability to pull down and bind NS3 from the cytoplasm to produce the NS2b-NS3 viral protease complex that is responsible for processing the cytoplasmic side of the flavivirus polyprotein [74-81]. The three highly conserved transmembrane regions have been studied for their importance in RNA production and virion assembly via an NS2a interface [82]; manipulation of these regions interferes with the ability for the virus to produce new RNA and infectious particles, although the protein processing abilities of NS2b are not affected by these selective mutations. Additionally, this protein complex targets innate immune proteins such as eIF4G1 and RIG-1 . Due to its role in viral polyprotein processing, NS2b has been the target of extensive antiviral research and development [76, 83, 84].

The structure of NS3 has been elucidated by x-ray crystallography [78, 85], and these findings show the two known domains of NS3, each with separate functions corresponding to both protein and genome generation, respectively. The protease domain works in tandem with NS2b to produce a viral protease, capable of cleaving cytoplasmic regions of the viral polyprotein to produce the individual nonstructural proteins. Interruption of the NS2b-NS3 interface is catastrophic for the viral replication process. The C-terminal domain has 3 subdomains that act as a magnesium ion dependent helicase and triphosphatase [86]. The helicase is responsible for unwinding the double stranded RNA intermediate and the latter processing the 5’ end of new positive-sense RNA for cap addition [87-91]. NS3 and NS5 work to enhance the enzymatic activity of each other [89, 92].

Like the other transmembrane nonstructural proteins, NS4a plays a role in host membrane manipulation and participating in critical binding interactions with the other flavivirus nonstructural proteins. Embedded in the ER membrane, the cytosol-facing N-terminus possesses alpha helices comprised of 48 cytosolic residues that are significantly disordered, providing a wide range of functional interaction with proteins, viral or otherwise [93-97]. This region contributes to the NTPase activity necessary for NS3 processing of the 5' end of the flavivirus genome, indicating a direct role in enhancing replication outside of membrane manipulation [95]. The C terminus is anchored on the luminal side by three transmembrane domains. There are chronicled interactions between NS4 and NS4b between their respective N-terminal regions, with both working in tandem for host lipid manipulation [58, 98-100]. NS4a also ventures away from the canonical location of flavivirus replication to modulate host processes. NS4a and NS5 directly interact with PAF1C in the nucleus, interfering with H3 promoter recruitment and modifying the host's ability to methylate histones [101]. Along with NS4b, NS4a will suppress neurogenesis in fetal stem cells by inhibiting Akt-mTOR signaling, deregulating the resulting signal cascade [102], and also interact with non-ubiquitinated AUP1 in lipid droplets to relocate them to autophagosomes and boost DEN infection [103]. The NF-kB pathway is also targeted by NS4a, although the suppression of the proinflammatory pathways is significantly less than the effect NS2a has [69]. In the same vein of immune modulation, NS4a has been identified as an antagonist of STAT1 and STAT2 binding, inhibiting phosphorylation and assembly of downstream signaling complexes needed for the expression of anti-viral interferon-stimulated genes (ISG) [104].

The 2K peptide is formed through the cleavage of a 23 amino acid long transmembrane domain between NS4a and NS4b and operates as a signaling peptide for the latter [105]. It is necessary for the cytoplasmic cleavage to occur first on the C-terminus of NS4a, allowing the NS4a protein to properly manipulate host membranes and the 2K-NS4b moiety to translocate properly where 2K is fully cleaved from the NS4b N-terminus [96]. 2K and its sequential cleavage is necessary for proper membrane rearrangement, with its cleavage even being postulated as the key trigger for membrane rearrangement [58]. Mutations have been identified that confer resistance to flavivirus infection infections, indicating the importance of this small peptide for the replication cycle via protein interactions [106, 107].

NS4b is the largest of the membrane-associated flavivirus nonstructural proteins that has canonically been described as having three transmembrane domains, an N terminal region within the ER lumen, and a cytosolic loop [108]. The canonical topology has recently come into question, as a study in Li et al. identified two potential transmembrane domains near the N-terminus as indeed crossing the ER membrane, resulting in another cytosolic loop [109]. Newly found interactions between both the new cytoplasmic loop and the supposed lumen-dwelling first 30 amino acids of the N-terminus with the NS3 helicase have been identified as beneficial for enzyme activity [95, 109-111]. To further complicate matters, the recently identified JNJ-1802 DENV inhibitor that interrupts NS3-NS4b binding is rescued by mutations near the N-terminus after the second transmembrane domain in the lumen, suggesting it is facing the cytoplasm at the time of NS3 association [27]. The identification of other critical NS3-

NS4b interaction points on the NS4b cytoplasmic loop from previous years complicates the notion that NS4b adheres to the canonical organization identified by Miller et al. in 2006 [112, 113]. This suggests either two different states of NS4b that bind differently to NS3 or an incomplete flipping of NS4b, with the N terminus coming over to the cytoplasm. No matter where this interaction forms, NS4b has been heavily studied for its interactions with NS3, with the enhancement of NS3 helicase activity being well documented following the binding of NS4b to the D2 and D3 helicase domains [110, 111, 113, 114]. NS4b has been implicated in membrane reorganization, as well as interactions with NS2A and NS1, suggesting its role in assembling the flavivirus replication compartments [98, 112]. As with other nonstructural flavivirus proteins, NS4b plays a role in immune modulation, downregulating type I interferon response genes through STAT1 phosphorylation suppression and TRIF degradation [67, 115].

Like NS3, NS5 has several roles in the flavivirus replication process. Due to its direct role in genome synthesis, extensive studies have been performed investigating NS5 structure, function, expression, and inhibition. Crystal structures are available and have been used in tandem with the NS3 crystal structure for hand-modeling protein interactions [116-119]. The C-terminus of this relatively massive viral protein acts as an RNA-dependent RNA polymerase that can produce both the negative-sense template strand of the viral RNA and new positive-sense copies from said template [88, 116, 120-125]. The N-terminus acts as a guanylyltransferase and a methyltransferase, with the former attaching a guanine monophosphate entity to the dephosphorylated 5’ end of a new genome and the latter methylating the N-7 methyl of said guanine and the ribose

2’-OH from the 5’ adenosine [92, 120, 126-132]. The active sites of these enzymatic domains are highly conserved across flavivirus species [121, 133]. Interestingly, NS5 is known to assemble into a dimer with enzymatic entry and exit points that suggest potential oligomerization in biologically active samples [120]. In addition to producing and capping viral RNA, NS5 binds to a multitude of host immune factors (primarily belonging to the type-1 interferon pathway) to modulate the interferon response triggered by viral infection. NS5 can suppress the STAT1/STAT2 mediated signaling cascade that activates type-1 interferon stimulated genes by suppressing IFNAR1 maturation through sequestration of the transport factor prolidase [134], inhibiting phosphorylation of STAT2 through binding to interrupt ISG activation, and triggering UBR4 mediated degradation of STAT2 [135-139]. Unsurprisingly, the key functions of a polymerase and capping enzyme are sought after targets for antiviral development [128, 133, 140].

1.2.3 The Flavivirus Replication Cycle in Mammalian Hosts

1.2.3.1 Virion Entry and Translation

For a mammalian cell, the flavivirus infection cycle event begins with a virion binding to a cell (**Figure 1.2a**). The E protein has been studied extensively for its receptor-binding capabilities, having been observed binding to glycosaminoglycans, glycosphingolipids, and surface-presenting GRP78 proteins [141-144]. For dendritic cells and macrophages, DENV virions will bind to mannose receptors and DC-SIGN

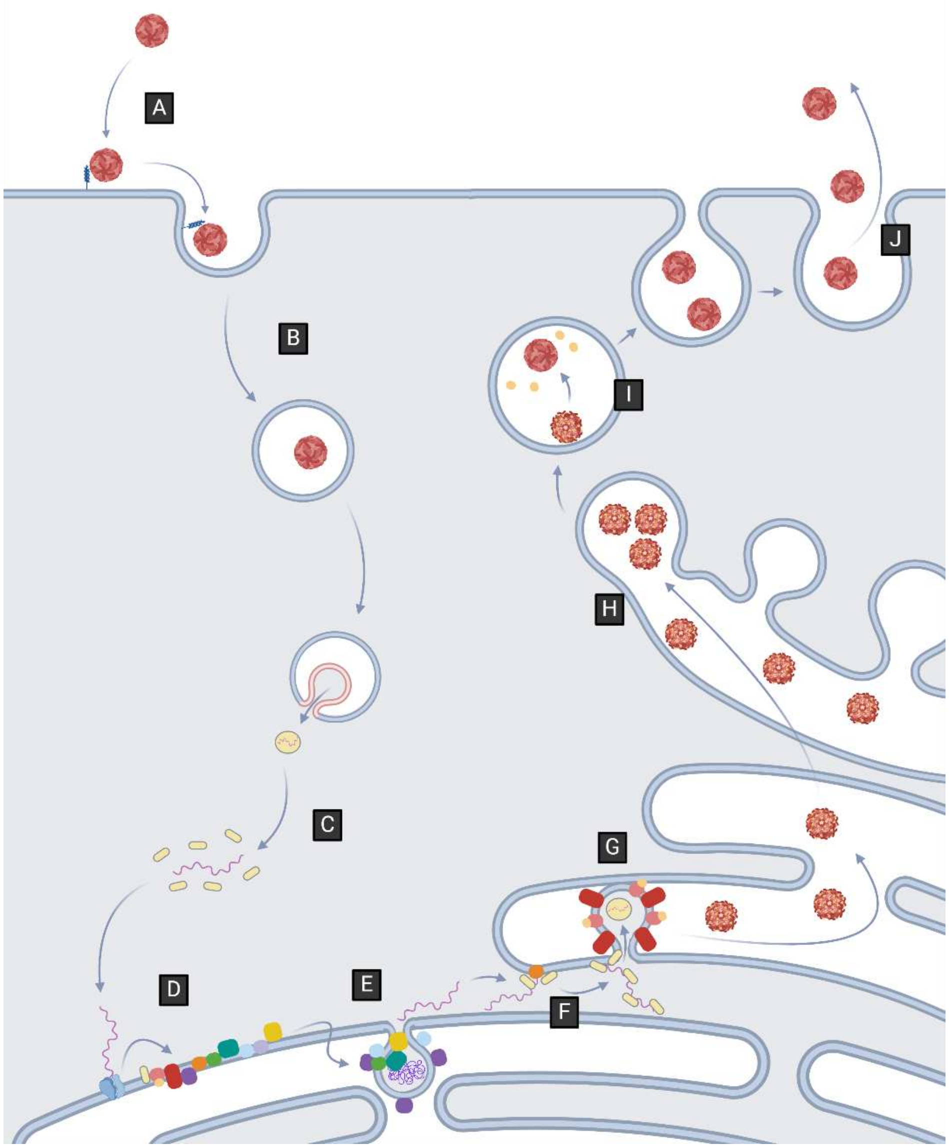


Figure 1.2. Generic flavivirus replication cycle. A) Virion binds to a cell surface receptor, triggering endocytosis. **B)** Cell traffics the engulfed virion into cell via a lysosome. **C)** Lysosome and virion membranes merge as pH drops, expelling the genome that promptly loses its capsid shell. **D)** Viral genome translated on the ER to produce the polyprotein **E)** Polyprotein is cleaved into structural and nonstructural proteins, nonstructural proteins assemble the replication compartment and replicate the viral genome off of a negative-sense strand template. **F)** New viral RNA is sequestered by NS2a and alpha-carbon bound C proteins on the ER membrane, genome is coated in more C protein **G)** The capsid-coated genome is brought into structural protein molded ER vesicle and new virion buds into the ER lumen. **H)** Virions packaged into endosomes destined for the cell membrane in the Golgi network **I)** E protein conformationally shifts and host furin protease cleaves pr portion of prM, creating a mature virion. **J)** Mature virions are released from the infected cell. Created with BioRender.com.

molecules on the cell surface [145-147]. TIM and TAM act as entry factors for the phagocytosis of the virion, bringing the virions in via clathrin-mediated endocytosis [148-151]. Once inside the host endosomes, the maturing endosome drops in pH change and alters its lipid content (**Figure 1.2b**). These conditions allow the flavivirus E protein to undergo a conformational shift and fuse with the late endosome membrane, expelling the viral genome into the cytoplasm (**Figure 1.2c**) [152, 153]. The viral genome is then translated by host ribosomes on the ER membrane into a single polyprotein (**Figure 1.2d**). Here, host signal peptidase and viral NS2b-NS3 proteases cleave the polyprotein into the ten structural and nonstructural flavivirus proteins while the ER membrane protein complex ensures proper folding of the transmembrane proteins [81, 154-158].

During post-translational processing, a series of cleavages occur on structural proteins that are critical for later assembly of new virions. A portion of translated C proteins are cleaved by NS2b-NS3 from the capsid anchor (Ca) transmembrane helix connecting the C and prM proteins while the rest remain bound to the ER membrane [159]. Signal peptidase then cleaves the Ca peptide from the prM protein, freeing it assemble into trimers. prM and E proteins associate into trimers that then combine into the prM-E heterodimer [160-162].

1.2.3.2 The Flavivirus Replication Compartment

Following translation and proteolytic processing, the flavivirus nonstructural proteins manipulate ER membranes to form a replication compartment (**Figure 1.2e**).

This compartment is comprised of lipids originating from the ER with NS1, NS2a, NS2b, NS4a, and NS4b having all been identified as directly contributing to this rearrangement through interactions with host lipids, often working together [58, 60, 65, 72, 93, 94, 98-100, 163-165]. This invagination balloons into the ER luminal space with a pore at the opening to the cytoplasm, allowing the entry and exit of substrates or replication products [163, 165-167]. A simplified theoretical model of the flavivirus replication compartment based on known protein-protein interactions can be seen in **Figure 1.3**, along with a few of the many unanswered questions concerning its organization. In this environment the nonstructural proteins work to replicate the flavivirus genome.

1.2.3.3 Flavivirus Genome Replication

In conjunction with compartment formation, viral nonstructural proteins assemble into a replication complex that uses the helicase and triphosphatase from NS3 and the RdRp, methyltransferase, and guanylyltransferase of NS5 to produce new viral genomes [122, 163, 164, 168, 169]. First, a negative-sense template copy of the flavivirus genome is produced from an existing positive sense genome following the circularization of the positive-sense genome. NS5 binds with the stem loop A (SLA) structural motif in the 5' UTR and the genome circularizes by interacting with structural elements in the 3' UTR and capsid-coding region [51, 170-172]. From here, the NS5 RdRp is free to polymerize the negative sense strand traveling from the 3' to 5' direction on the original positive-sense genome. Interestingly, the 3' UTR contains structural elements that deter polymerase activity without hybridizing with 5' UTR structures,

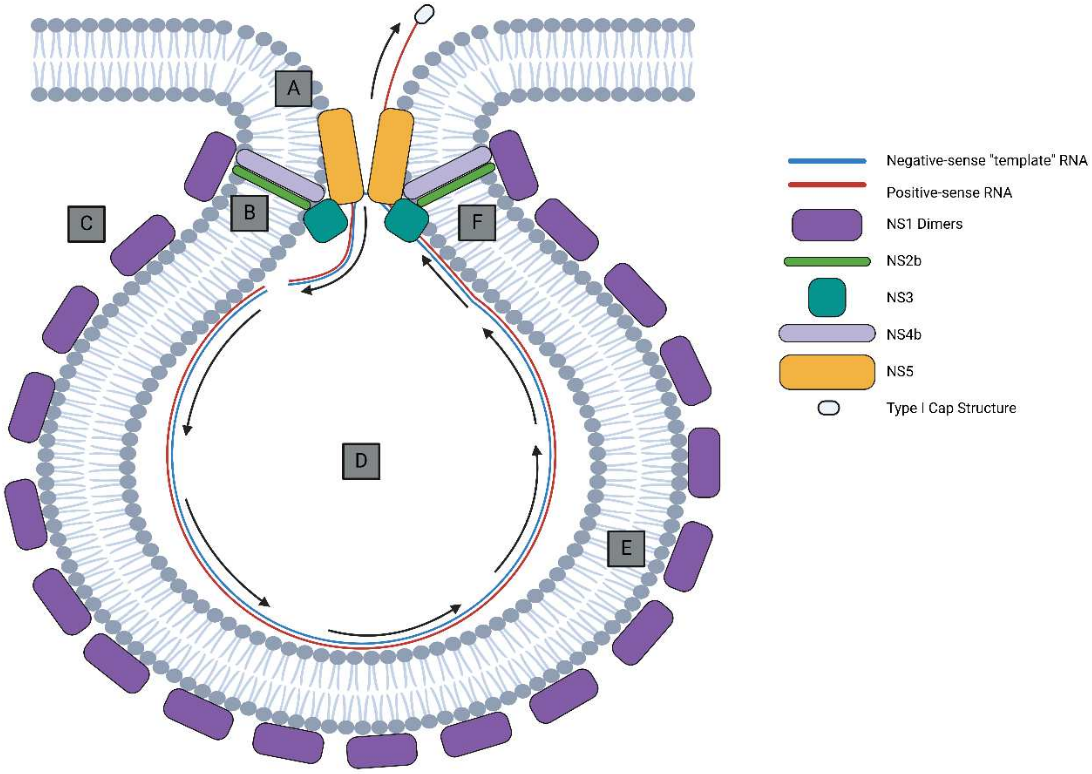


Figure 1.3. Hypothesized flavivirus replication compartment organization. A simplified hypothetical flavivirus replication compartment in the ER demonstrating the cyclical travel of the dsRNA intermediate through the compartment. The dsRNA is unwound by NS3, the negative-sense template passes through an NS5 RdRp domain to produce a new positive-sense vRNA copy, and the dsRNA intermediate returns to the compartment. The positive-sense vRNA copy separated from the template strand by NS3 is capped with a Type I cap structure and is released into the cytoplasm. NS2b and NS4b anchor NS3 to the membrane and NS1 dimers exist in the ER lumen, bound to the membrane. Listed are some of the unanswered questions concerning the structure and organization of the replication compartment:

A) Do the reported NS5 dimers form a pore structure that participates in replication?
B) Does NS2b, NS4b, and NS3 bind together near the neck of the pore or elsewhere in the compartment? Do these interactions hold NS3 in a single location?
C) Do NS1 dimers fully coat the lumen side of the replication compartment? And in what stoichiometry?
D) How is the double stranded vRNA intermediate form organized in the replication compartment space? Or are the nonstructural proteins intermixed?
E) Are the known NS1 and NS4a interactions present within the replication compartment?
F) Do other viral or host proteins participate in protein-protein interactions within the replication compartment?

Created with BioRender.com.

suggesting a control mechanism for generation of new template strands [173]. This template strand is protected in the replication compartment with the replication complex, hidden away from host immune proteins that can identify uncapped or double-stranded RNA [42, 43, 174].

Following the generation of the template viral RNA (vRNA) strand, the replication complex continues its work in generating new, capped viral RNA. How the sequential mechanisms fit together for this process is still unclear, but the major events have been well documented *in vivo* and *in vitro* research [163, 169]. The NS5 RdRp runs along the template strand in the 3’ to 5’ direction, producing a double-stranded RNA intermediate [88, 121, 124, 132, 175]. These two hybridized RNA strands are unzipped by the NS3 helicase domain with the 3’ end of the negative sense strand returning to the polymerase for further RNA production and the new positive-sense strand being processed by the NS3 triphosphatase domain [86, 90, 125]. While this is happening, the NS5 N-terminal capping domain binds to a GTP molecule and dephosphorylates it, with the methyltransferase binding the new GMP moiety covalently [92]. The diphosphate 5’ end of the new genome then enters the capping domain of NS5 where the guanylyltransferase affixes the loaded GMP and the methyltransferase methylates the N-7 of the guanine and the 2’OH of the 5’ adenosine ribose [39, 175-178]. The newly capped viral RNA genome then exits the replication compartment where its fate is then decided.

1.2.3.4 Fates of Viral RNA: Translation, Packaging, Detection, and Degradation

New copies of the flavivirus genome spread out into the cytoplasm for various purposes. Some reengage with the translation process to produce more structural and nonstructural proteins to facilitate transcription and new virion formation, while others are packaged into new viral particles with assistance of NS2A, NS2B, and NS1 [66, 71, 72, 82]. A significant portion of viral RNA engages with the host cell immune proteins. As mentioned previously, host pattern recognition receptors (PRRs) and toll-like receptors (TLRs) recognize uncapped and double stranded RNA as pathogen-associated molecular patterns (PAMPs). This will eventually occur despite the protection replication compartments bestow dsRNA; in Overby et al., the authors indicate that the cloaking of dsRNA in replication compartments can only delay IFN induction (as indicated by eIF2α phosphorylation levels) for 24 hours in tick-borne encephalitis virus (TBEV) [179]. After a certain point the host starts recognizing dsRNA, and even then, nonstructural proteins work to downregulate the triggered interferon response [42, 43, 68, 69, 104, 115, 134, 136-139, 174, 180, 181].

In addition to these fates, copies of viral RNA enter the host mRNA decay pathway. A key member of this pathway is the enzyme Xrn1, a host exoribonuclease that degrades monophosphorylated RNA molecules from the 5' to the 3' direction [182-184]. When monophosphorylated flavivirus genomes are degraded by Xrn1, the enzyme degrades the strand until it reaches the 3' UTR. Here, Xrn1 stalls on one of the three helix structures stabilized by pseudoknots just downstream of the 3' coding terminus [48, 49, 52, 185, 186]. This effectively sequesters Xrn1, reducing its efficacy at degrading transcripts and deregulating the host mRNA degradation pathway [49, 187]. The remaining ~500 nucleotides are termed subgenomic flavivirus RNA (sfRNA) and play a role in modulating the host immune response to viral infection. Schnettler et al. showed that sfRNA, not viral nonstructural proteins, were responsible for suppressing the short hairpin (shRNA) RNA decay mechanisms normally employed by both vertebrate and insect hosts to combat infection [188]. sfRNA actively blocks the activating phosphorylation of IRF-3 and IRF-7, transcription factors that are critical regulate the expression of antiviral ISGs [189]. A study into a prolific DENV strain showed that tripartite motif 25 (TRIM25) ubiquitylation is blocked by sfRNA binding, further suppressing the host IFN response [46]. sfRNA generation by Xrn1 degradation is necessary for deregulating the host antiviral response and supporting an effective viral infection.

1.2.3.5 Viral Packaging and Virion Exocytosis

Many genome copies are destined for future infections and packaged into new virions structures with the necessary structural proteins. The 3'UTR of the newly capped vRNA binds with NS2a on the cytoplasmic side of the ER membrane (**Figure 1.2f**). Here, dimers of Ca-anchored C proteins bind to the vRNA as cytoplasmic C protein bind to the genome. The capsid dimers bring the vRNA to the cytoplasmic side of the prM-E heterodimer to begin forming the new virion [190, 191]. Stabilized by NS1, the structural proteins work in tandem to manipulate the ER membrane into a new virion that “buds” into the ER lumen with the C protein and vRNA inside (**Figure 1.2g**) [61, 66]. As new virions are shuttled towards the cell membrane pH drops as cis-Golgi endosomes mature into trans-Golgi vesicles (**Figure 1.2h**). This pH drop results in a conformational shift in the E protein that lays the angled proteins flat against the membrane, turning two prM-E heterodimer trimer subunits into three dimers [192, 193]. This results in the cleavage of the PrM protein by the host serine protease furin at the newly revealed cleavage site (**Figure 1.2i**) [194, 195]. This cleavage acts as a safeguard to prevent the fusion of the new virion membrane with the host membrane before exocytosis. New virions are then released into the extracellular space as the trans-Golgi vesicle fuses with the cellular membrane and the infection cycle repeats itself (**Figure 1.2j**).

1.3 Holes in Knowledge

While there is progress being made in discovering treatments for flaviviruses, an effective, readily available therapeutic is still far off. To develop a robust arsenal of antivirals, our understanding of the critical facets of viral infection and replication needs to be fleshed out. Here we have primarily covered two critical areas of the flavivirus replication cycle; the assembly of nonstructural proteins into a working replication compartment and complex, and the degradation of viral RNA to produce host immune modulating sfRNA. There are still significant gaps in our understanding of viral replication and host immune evasion that, if illuminated, would provide new avenues for antiviral or prophylactic development against these destructive arboviruses.

1.3.1 The Replication Compartment and Its Contents

The protective role the compartment serves for housing the flavivirus double-stranded RNA intermediate and the replication complex is key in ensuring successful infection and virion propagation. The nonstructural proteins serve a multitude of roles, and their construction of and presence within the compartment are arguably among their most necessary contribution to the continued cycle of virion generation and infection. The successful association of the nonstructural proteins is necessary for forming this compartment, and thus disturbing these interactions can result in the collapse of the infection cycle.

For the replication compartment and complex, there is much still to be learned about the associations between specific proteins and the organization of the replication compartment contents. All nonstructural proteins participate in the formation of the compartment or the enzymatic activity of the complex. NS2b binds with NS3 to create the NS2b-NS3 protease complex that prepares all nonstructural proteins for work in making the compartment and complex [80, 196]. NS1, NS2a, NS4a, and NS4b have all been implicated in replication compartment formation through the reorganization of host membranes [60, 65, 66, 72, 96, 98, 100]. NS4b modulates NS3's helicase activity while NS5 and NS3 provide the primary machinery for the replication complex, creating and capping new viral genomes. Our knowledge of how these proteins bind together and along what interfaces is limited- previous experiments have revealed some nonstructural protein binding interactions depending on the stage of infection. The transmembrane nonstructural proteins NS2a, NS4a, and NS4b work with NS1 to reorganize the host cell membranes for replication compartment and virion formation, with notable interactions between these proteins being chronicled in the past through co-immunoprecipitation and mutagenesis analysis [65, 98, 100, 197]. NS2b binds with NS3 along the C-terminal protease domain to create the NS2b-NS3 protease complex, and NS4b binds to NS3 with both its N-terminus and cytosolic loop for efficient helicase activity. NS3-NS5 improves the enzymatic efficacy of both proteins, boosting their respective roles in processing viral RNA [89, 122, 125]. And even with all this information, we still do not know which host proteins are interacting with viral nonstructural proteins to create the replication compartments and complex, if any.

Previous investigations into flavivirus replication compartments have been limited by the complexity of the samples. Cryo-electron tomography has provided an exciting look at the viral replication compartments superstructures and provide answers into how the virus manipulates host membranes [164, 165]. These studies have revealed the location and dimensions of the flavivirus replication factories, as well as the pore channels linking compartments to the cytoplasm and, in some cases, other compartments. Unfortunately, inferences on the organization of proteinaceous components of these compartments are not attainable from these tomograms outside of the electron density changes from to the coiled genetic material contained inside. *In silico* modeling has been performed to provide theoretical methods for understanding protein interactions in the replication machinery using known active sites and the crystal structures of NS3 and NS5 [118]. Unfortunately, a recent study has disproven this model, demonstrating the need for caution when utilizing a model without supporting experimental evidence [119]. *In vitro* cryo-electron microscopy analysis of the protein interactions between NS3 and NS5 on an RNA construct featuring the 5'UTR SLA has given an idea of the structural organization and interactions between these two proteins, but only at the two specific conformations at certain points in time, while also on a controlled environment devoid of other environmental proteins or multimerization present within the replication compartment [198]. From what we know from mutagenesis experiments, *in silico* studies, and cryo-electron microscopy we can postulate theoretical organization structures for the replication compartment, although these are still held back by unanswered questions (**Figure 1.3**). Our knowledge of how the building blocks of the replication compartment and replication complex associate together withing the compartment are still far from complete, illustrating an important blind spot in the virology's understanding of the flavivirus replication cycle

1.3.2 Enzymatic Processing of Flavivirus Genomes into Xrn1 Substrate

With sfRNA being necessary for successful infection and having a significant effect on host immune modulation, significant effort has gone into understanding the mechanisms behind its generation with Xrn1 [50]. Despite this, there are still unknowns when it comes to the source of the Xrn1 monophosphorylated substrate. Due to the specificity of the monophosphate on the Xrn1 substrate, target RNAs need to be processed by decapping or dephosphorylating enzymes prior to digestion, as most host and viral RNAs are at the very least triphosphorylated. It has been postulated that the monophosphorylated viral RNA is generated by host decapping and dephosphorylating enzymes.

If viral RNAs were from the replication compartment uncapped with a triphosphorylated end, then DUSP11 or the other Nudix hydrolases may be a source of Xrn1-ready RNA. The Nudix family of hydrolases can be found in species ranging from bacteria to humans, all sharing characteristic metal-ion dependent catalytic mechanisms for dephosphorylating RNA molecules at the 5' end. When it comes to viral infection, some Nudix hydrolases may have enhancing effects towards viral replication while others, such as NUDT2, have been shown to expressly interfere with viral replication by processing viral transcripts into Xrn1 digestible molecules [199]. Dusp11 is another phosphatase enzyme that is involved in mediating the concentration of triphosphorylated RNAs in the cytoplasm, dephosphorylating host transcripts to keep PRRs such as RIG-I from over-stimulating the type I interferon response [200]. Additionally, DUSP11 is critical for regulating RNA polymerase III noncoding transcripts in metazoans that participate in the Argonaute protein assembly and thus RNA-induced silencing mediated transcription control process [201]. On the antiviral front, DUSP11 regulates both host antiviral mRNAs to control immune response while also processing triphosphorylated RNAs prior to Xrn1 degradation, as seen with HCV in Kincaid et al. [201, 202].

Another option for monophosphorylated Xrn1 substrate generation would be for the decapping and dephosphorylation of viral RNA. DXO, also known as Dom3Z, appears to be a quality-control mechanism for host transcripts, as it is blocked from effectively binding to RNA transcripts by type 1 caps with a proper 2'O methylation [41]. It was identified as a Swiss-army knife of an RNA degradation enzyme; it is capable of decapping transcripts, dephosphorylation the products, and even acting as an exoribonuclease that breaks down in the 5’ to 3’ direction [203]. Dcp2, a member of the previously highlighted Nudix enzyme family, has been heavily studied for its role in host transcript degradation. This enzyme removes 5’ guanine caps from transcripts, leaving a 5’ monophosphate that is ripe for Xrn1 degradation [204, 205]. This process is critical for maintaining proper mRNA levels in infected cells, as Dcp2 is a key regulator in proinflammatory ISG promoter IRF-7 and OAS2 [206]. Dcp2 has been observed colocalizing with Xrn1 during infection in what are known as processing, or P, bodies during viral infections [182, 184, 207]. It has been postulated that Dcp2 is the source of monophosphorylated flavivirus RNA for Xrn1 breakdown. These enzymes have been theorized for being the first step in generating substrate for Xrn1, but as of yet none have been confirmed as being the source of the monophosphorylated flavivirus RNA.

1.3.3 Filling These Holes

Much like repairing potholes in a road to increase the pace of traffic, filling the gaps in our understanding of these critical players in the flavivirus infection cycle can accelerate the timeline for discovering new treatments and vaccines. Interrupting viral protein-protein interactions for deterring the proper formation of the replication compartment or complex would stop an infection dead in its tracks, but this requires detailed knowledge of those interactions. A thorough *in situ* method such as crosslinking mass spectrometry would be able to identify critical protein interfaces in the replication compartment, allowing us to build a model for the compartment and intelligently design interaction inhibitors. Knowing the whole picture in sfRNA generation can assist in generating compounds to impede the generation of this immunomodulating molecule. Specifically, identifying which of the host proteins involved in RNA decapping and dephosphorylation acts on vRNA to make Xrn1 substrate would provide avenues of attack for disrupting the monophosphorylated RNA generation pipeline. Knowing the basic science behind flavivirus infection is critical for taking steps to combating the threat these pathogens pose to global health.

1.4 REFERENCES

1. Selck, F.W., A.A. Adalja, and C.R. Boddie, *An estimate of the global health care and lost productivity costs of dengue.* Vector Borne Zoonotic Dis, 2014. **14**(11): p. 824-6.

2. Shepard, D.S., E.A. Undurraga, and Y.A. Halasa, *Economic and disease burden of dengue in Southeast Asia.* PLoS Negl Trop Dis, 2013. **7**(2): p. e2055.

3. Shepard, D.S., E.A. Undurraga, Y.A. Halasa, et al., *The global economic burden of dengue: a systematic analysis.* Lancet Infect Dis, 2016. **16**(8): p. 935-41.

4. Bhatt, S., P.W. Gething, O.J. Brady, et al., *The global distribution and burden of dengue.* Nature, 2013. **496**(7446): p. 504-7.

5. Leta, S., T.J. Beyene, E.M. De Clercq, et al., *Global risk mapping for major diseases transmitted by Aedes aegypti and Aedes albopictus.* Int J Infect Dis, 2018. **67**: p. 25-35.

6. Turrell, M.J., D.J. Dohm, M.R. Sardelis, et al., *An update on the potential of north American mosquitoes (Diptera Culicidae) to transmit West Nile Virus.pdf.* Journal of Medical Entymology, 2005. **42**(1): p. 57-62.

7. D.J., G., *The global emergence-resurgence of arboviral diseases as public health problems.* Archives of Medical Research, 2002. **33**(4): p. 330-342.

8. Girard, M., C.B. Nelson, V. Picot, et al., *Arboviruses: A global public health threat.* Vaccine, 2020. **38**(24): p. 3989-3994.

9. Mustafa, M.S., V. Rasotgi, S. Jain, et al., *Discovery of fifth serotype of dengue virus (DENV-5): A new public health dilemma in dengue control.* Med J Armed Forces India, 2015. **71**(1): p. 67-70.

10. Wang, W.H., A.N. Urbina, M.R. Chang, et al., *Dengue hemorrhagic fever - A systemic literature review of current perspectives on pathogenesis, prevention and control.* J Microbiol Immunol Infect, 2020. **53**(6): p. 963-978.

11. Rothman, A.L. and F.A. Ennis, *Immunopathogenesis of Dengue Hemorrhagic Fever.* Virology, 1999. **257**(1): p. 1-6.

12. Katzelnick, L.C., L. Gresh, M.E. Halloran, et al., *Antibody-dependent enhancement of severe dengue disease in humans.* Science, 2017. **358**(6365): p. 929-932.

13. Hall-Mendelin, S., A.T. Pyke, P.R. Moore, et al., *Assessment of Local Mosquito Species Incriminates Aedes aegypti as the Potential Vector of Zika Virus in Australia.* PLoS Negl Trop Dis, 2016. **10**(9): p. e0004959.

14. Dick, G.W.A., S.F. Kitchen, and A.J. Haddow, *Zika Virus (I). Isolations and serological specificity.* Transactions of The ROyal Society of Tropical Medicine and Hygiene, 1952. **46**(5): p. 509-520.

15. Faye, O., C.C. Freire, A. Iamarino, et al., *Molecular evolution of Zika virus during its emergence in the 20(th) century.* PLoS Negl Trop Dis, 2014. **8**(1): p. e2636.

16. Oehler, E., L. Watrin, P. Larre, et al., *Zika virus infection complicated by Guillain-Barré syndrome – case report, French Polynesia, December 2013.* Eurosurveillance, 2014. **19**(9).

17. Schuler-Faccini, L., E.M. Ribeiro, I.M.L. Feitosa, et al., *Possible Association Between Zika Virus Infection and Microcephaly - Brazil, 2015.* MMWR Morbidity and Mortality Weekly Report, 2016. **65**(3): p. 59-62.

18. Rychlowska, M., A. Agyapong, M. Weinfeld, et al., *Zika Virus Induces Mitotic Catastrophe in Human Neural.* Journal of Virology, 2022. **96**(11).

19. Godsey, M.S.J., M.S. Blackmore, N.A. Panella, et al., *West Nile Virus Epizootiology in the Southeastern United States, 2001.* Vector Borne and Zoonotic Diseases, 2005. **5**(1): p. 82-9.

20. Nash, D., F. Mostashari, A. Fine, et al., *The Outbreak of West Nile Virus Infection in the New York City Area in 1999.* The New England Journal of Medicine, 2001. **344**(24): p. 1807-14.

21. Hurlbut, H.S., F. Rizk, R.M. Taylor, et al., *A Study of the Ecology of West Nile Virus in Egypt.* American Journal of Tropical Medicine and Hygeine, 1956. **5**(4): p. 579-620.

22. Berthet, F.X., H.G. Zeller, M.T. Drouet, et al., *Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses.* The Journal of General Virology, 1997. **78**: p. 2293-7.

23. O'Leary, D., A. Marfin, S. Montogmery, et al., *The Epidemic of West Nile Virus in the United States, 2002.* Vector Borne and Zoonotic Diseases, 2004. **4**(1): p. 61-70.

24. Cendejas, P.M. and A.G. Goodman, *Vaccination and Control Methods of West Nile Virus Infection in Equids and Humans.* Vaccines (Basel), 2024. **12**(5).

25. Naveed, A., L.G. Eertink, D. Wang, et al., *Lessons Learned from West Nile Virus Infection:Vaccinations in Equines and Their Implications for One Health Approaches.* Viruses, 2024. **16**(5).

26. Bovay, A., S.A. Fuertes Marraco, and D.E. Speiser, *Yellow fever virus vaccination: an emblematic model to elucidate robust human immune responses.* Hum Vaccin Immunother, 2021. **17**(8): p. 2471-2481.

27. Goethals, O., S.J.F. Kaptein, B. Kesteleyn, et al., *Blocking NS3-NS4B interaction inhibits dengue virus in non-human primates.* Nature, 2023. **615**(7953): p. 678-686.

28. Foy, B.D., A. Some, T. Magalhaes, et al., *Repeat Ivermectin Mass Drug Administrations for Malaria Control II Protocol for a Double-blind, Cluster-Randomized, Placebo-Controlled Trial for the Integrated Control of Malaria.* JMIR Research Protocols, 2023. **12**.

29. Holcomb, K.M., C. Nguyen, B.D. Foy, et al., *Effects of ivermectin treatment of backyard chickens on mosquito dynamics and West Nile virus transmission.* PLoS Negl Trop Dis, 2022. **16**(3): p. e0010260.

30. Kriesner, P., A.A. Hoffmann, S.F. Lee, et al., *Rapid sequential spread of two Wolbachia variants in Drosophila simulans.* PLoS Pathog, 2013. **9**(9): p. e1003607.

31. Hancock, P.A., V.L. White, S.A. Ritchie, et al., *Predicting Wolbachia invasion dynamics in Aedes aegypti populations using models of density-dependent demographic traits.* BMC Biol, 2016. **14**(1): p. 96.

32. Hedges, L.M., J.C.B. Brownlie, S.L. O'Neill, et al., *Wolbachia and Virus Protection in Insects.* Science, 2008. **322**(5902): p. 702.

33. Amuzu, H.E., K. Tsyganov, C. Koh, et al., *Wolbachia enhances insect-specific flavivirus infection in Aedes aegypti mosquitoes.* Ecol Evol, 2018. **8**(11): p. 5441-5454.

34. Teixeira, L., A. Ferreira, and M. Ashburner, *The bacterial symbiont Wolbachia induces resistance to RNA viral infections in Drosophila melanogaster.* PLoS Biol, 2008. **6**(12): p. e2.

35. Shaw, W.R., P. Marcenac, L.M. Childs, et al., *Wolbachia infections in natural Anopheles populations affect egg laying and negatively correlate with Plasmodium development.* Nat Commun, 2016. **7**: p. 11772.

36. van den Hurk, A.F., S. Hall-Mendelin, A.T. Pyke, et al., *Impact of Wolbachia on infection with chikungunya and yellow fever viruses in the mosquito vector Aedes aegypti.* PLoS Negl Trop Dis, 2012. **6**(11): p. e1892.

37. Glaser, R.L. and M.A. Meola, *The native Wolbachia endosymbionts of Drosophila melanogaster and Culex quinquefasciatus increase host resistance to West Nile virus infection.* PLoS One, 2010. **5**(8): p. e11977.

38. Wengler, G. and G. Wengler, *Terminal sequences of the genome and replicatioe-form RNA of the flavivirus west nile virus absence of poly(A) and possible role in RNA replication.* Virology, 1981. **113**(2): p. 544-55.

39. Ray, D., A. Shah, M. Tilgner, et al., *West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5.* J Virol, 2006. **80**(17): p. 8362-70.

40. Cleaves, G.R. and D.T. Dubin, *Methylation status of intracellular dengue type 2 40 S RNA.* Virology, 1979. **96**(1): p. 159-65.

41. Picard-Jean, F., C. Brand, M. Tremblay-Letourneau, et al., *2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO.* PLoS One, 2018. **13**(3): p. e0193804.

42. Errett, J.S., M.S. Suthar, A. McMillan, et al., *The essential, nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus infection.* J Virol, 2013. **87**(21): p. 11416-25.

43. Fredericksen, B.L., B.C. Keller, J. Fornek, et al., *Establishment and maintenance of the innate antiviral response to West Nile Virus involves both RIG-I and MDA5 signaling through IPS-1.* J Virol, 2008. **82**(2): p. 609-16.

44. Chu, P.W.G. and E.G. Westaway, *Replication strategy of Kunjin Virus: Evidence for recycling role of replicative form RNA as template in semiconservative and asymmetric replication.pdf.* Virology, 1984. **140**: p. 68-79.

45. Pompon, J., M. Manuel, G.K. Ng, et al., *Dengue subgenomic flaviviral RNA disrupts immunity in mosquito salivary glands to increase virus transmission.* PLoS Pathog, 2017. **13**(7): p. e1006535.

46. Manokaran, G., E. Finol, C. Wang, et al., *Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness.* Science, 2015. **350**(6257): p. 217-21.

47. Moon, S.L., B.J. Dodd, D.E. Brackney, et al., *Flavivirus sfRNA suppresses antiviral RNA interference in cultured cells and mosquitoes and directly interacts with the RNAi machinery.* Virology, 2015. **485**: p. 322-9.

48. Roby, J.A., G.P. Pijlman, J. Wilusz, et al., *Noncoding subgenomic flavivirus RNA: multiple functions in West Nile virus pathogenesis and modulation of host responses.* Viruses, 2014. **6**(2): p. 404-27.

49. Michalski, D., J.G. Ontiveros, J. Russo, et al., *Zika virus noncoding sfRNAs sequester multiple host-derived RNA-binding proteins and modulate mRNA decay and splicing during infection.* J Biol Chem, 2019. **294**(44): p. 16282-16296.

50. Pijlman, G.P., A. Funk, N. Kondratieva, et al., *A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity.* Cell Host Microbe, 2008. **4**(6): p. 579-91.

51. Filomatori, C.V., M.F. Lodeiro, D.E. Alvarez, et al., *A 5' RNA element promotes dengue virus RNA synthesis on a circular genome.* Genes Dev, 2006. **20**(16): p. 2238-49.

52. Wengler, G. and E. Castle, *Analysis of Structural Properties which Possibly Are Characteristic for the 3'-Terminal Sequence of the Genome RNA of Flaviviruses.* The Journal of General Virology, 1986. **67**: p. 1183-8.

53. Mohan, P.M. and R. Padmanabhan, *Detection of stable secondary structure at the 3' terminus of dengue virus type 2 RNA.* Gene, 1991. **108**(2): p. 185-91.

54. Villordo, S.M., C.V. Filomatori, I. Sanchez-Vargas, et al., *Dengue virus RNA structure specialization facilitates host adaptation.* PLoS Pathog, 2015. **11**(1): p. e1004604.

55. Chapman, E.G., D.A. Costantino, J.L. Rabe, et al., *The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production.* Science, 2014. **344**(6181): p. 307-310.

56. Chapman, E.G., S.L. Moon, J. Wilusz, et al., *RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA.* Elife, 2014. **3**: p. e01892.

57. Manzano, M., E.D. Reichert, S. Polo, et al., *Identification of cis-acting elements in the 3'-untranslated region of the dengue virus type 2 RNA that modulate translation and replication.* J Biol Chem, 2011. **286**(25): p. 22521-34.

58. Roosendaal, J., E.G. Westaway, A. Khromykh, et al., *Regulated cleavages at the West Nile virus NS4A-2K-NS4B junctions play a major role in rearranging cytoplasmic membranes and Golgi trafficking of the NS4A protein.* J Virol, 2006. **80**(9): p. 4623-32.

59. Lindenbach, B.D. and C.M. Rice, *trans-Complementation of Yellow Fever Virus NS1 Reveals a Role in Early RNA Replication.* Journal of Virology, 1997. **71**(12): p. 9608-9617.

60. Akey, D.L., W.C. Brown, S. Dutta, et al., *Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system.* Science, 2014. **343**(6173): p. 881-5.

61. Edeling, M.A., M.S. Diamond, and D.H. Fremont, *Structural basis of Flavivirus NS1 assembly and antibody recognition.* Proc Natl Acad Sci U S A, 2014. **111**(11): p. 4285-90.

62. Mackenzie, J.M., M.K. Jones, and P.R. Young, *Immunolocalization of the Dengue Virus Nonstructural Glycoprotein NS1 Suggests a Role in Viral RNA Replication.* Virology, 1996: p. 232-240.

63. Poveda-Cuevas, S.A., C. Etchebest, and F.L. Barroso da Silva, *Self-association features of NS1 proteins from different flaviviruses.* Virus Res, 2022. **318**: p. 198838.

64. Lindenbach, B.D. and C.M. Rice, *Genetic Interaction of Flavivirus Nonstructural Proteins NS1 and NS4A as a Determinant of Replicase Function.* Journal of Virology, 1999. **73**(6): p. 4611-4621.

65. Youn, S., T. Li, B.T. McCune, et al., *Evidence for a genetic and physical interaction between nonstructural proteins NS1 and NS4B that modulates replication of West Nile virus.* J Virol, 2012. **86**(13): p. 7360-71.

66. Scaturro, P., M. Cortese, L. Chatel-Chaix, et al., *Dengue Virus Non-structural Protein 1 Modulates Infectious Particle Production via Interaction with the Structural Proteins.* PLoS Pathog, 2015. **11**(11): p. e1005277.

67. Fanunza, E., N. Grandi, M. Quartu, et al., *INMI1 Zika Virus NS4B Antagonizes the Interferon Signaling by Suppressing STAT1 Phosphorylation.* Viruses, 2021. **13**(12).

68. Ngueyen, T.T.N., S.J. Kim, J.Y. Lee, et al., *Zika Virus Proteins NS2A and NS4A Are Major Antagonists that Reduce IFN-beta Promoter Activity Induced by the MDA5/RIG-I Signaling Pathway.* J Microbiol Biotechnol, 2019. **29**(10): p. 1665-1674.

69. Lee, J.Y., T.T.N. Nguyen, and J. Myoung, *Zika Virus-Encoded NS2A and NS4A Strongly Downregulate NF-kappaB Promoter Activity.* J Microbiol Biotechnol, 2020. **30**(11): p. 1651-1658.

70. Leung, J.Y., G.P. Pijlman, N. Kondratieva, et al., *Role of nonstructural protein NS2A in flavivirus assembly.* J Virol, 2008. **82**(10): p. 4731-41.

71. Zhang, X., X. Xie, H. Xia, et al., *Zika Virus NS2A-Mediated Virion Assembly.* mBio, 2019. **10**(5).

72. Xie, X., J. Zou, X. Zhang, et al., *Dengue NS2A Protein Orchestrates Virus Assembly.* Cell Host Microbe, 2019. **26**(5): p. 606-622 e8.

73. Zhang, X., X. Xie, J. Zou, et al., *Genetic and biochemical characterizations of Zika virus NS2A protein.* Emerg Microbes Infect, 2019. **8**(1): p. 585-602.

74. Tseng, A.C., V.R. Nerurkar, K.R. Neupane, et al., *Potential Dual Role of West Nile Virus NS2B in Orchestrating NS3 Enzymatic Activity in Viral Replication.* Viruses, 2021. **13**(2).

75. Xing, H., S. Xu, F. Jia, et al., *Zika NS2B is a crucial factor recruiting NS3 to the ER and activating its protease activity.* Virus Res, 2020. **275**: p. 197793.

76. Brecher, M., Z. Li, B. Liu, et al., *A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease.* PLoS Pathog, 2017. **13**(5): p. e1006411.

77. Hill, M.E., A. Kumar, J.A. Wells, et al., *The Unique Cofactor Region of Zika Virus NS2B-NS3 Protease Facilitates Cleavage of Key Host Proteins.* ACS Chem Biol, 2018. **13**(9): p. 2398-2405.

78. Luo, D., T. Xu, C. Hunke, et al., *Crystal structure of the NS3 protease-helicase from dengue virus.* J Virol, 2008. **82**(1): p. 173-83.

79. Xie, S., Z. Liang, X. Yang, et al., *Japanese Encephalitis Virus NS2B-3 Protein Complex Promotes Cell Apoptosis and Viral Particle Release by Down-Regulating the Expression of AXL.* Virol Sin, 2021. **36**(6): p. 1503-1519.

80. Yamshchikov, V.F. and R.W. Compans, *Formation of the flavivirus envelope: role of the viral NS2B-NS3 protease.* Journal of Virology, 1995. **69**: p. 1995-2003.

81. Chambers, T.J., R.C. Weir, A. Grakoui, et al., *Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein.* Proc Natl Acad Sci U S A, 1990. **87**(22): p. 8898-902.

82. Li, X.D., C.L. Deng, H.Q. Ye, et al., *Transmembrane Domains of NS2B Contribute to both Viral RNA Replication and Particle Formation in Japanese Encephalitis Virus.* J Virol, 2016. **90**(12): p. 5735-5749.

83. Nie, S., J. Zhao, X. Wu, et al., *Synthesis, structure-activity relationship and antiviral activity of indole-containing inhibitors of Flavivirus NS2B-NS3 protease.* Eur J Med Chem, 2021. **225**: p. 113767.

84. Yao, Y., T. Huo, Y.L. Lin, et al., *Discovery, X-ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease.* J Am Chem Soc, 2019. **141**(17): p. 6832-6836.

85. Assenberg, R., E. Mastrangelo, T.S. Walter, et al., *Crystal structure of a novel conformational state of the flavivirus NS3 protein: implications for polyprotein processing and viral replication.* J Virol, 2009. **83**(24): p. 12895-906.

86. Benarroch, D., B. Selisko, G.A. Locatelli, et al., *The RNA helicase, nucleotide 5'-triphosphatase, and RNA 5'-triphosphatase activities of Dengue virus protein NS3 are Mg2+-dependent and require a functional Walker B motif in the helicase catalytic core.* Virology, 2004. **328**(2): p. 208-18.

87. Bartelma, G. and R. Padmanabhan, *Expression, purification, and characterization of the RNA 5'-triphosphatase activity of dengue virus type 2 nonstructural protein 3.* Virology, 2002. **299**(1): p. 122-32.

88. Duan, Y., M. Zeng, B. Jiang, et al., *Flavivirus RNA-Dependent RNA Polymerase Interacts with Genome UTRs and Viral Proteins to Facilitate Flavivirus RNA Replication.* Viruses, 2019. **11**(10).

89. Yon, C., T. Teramoto, N. Mueller, et al., *Modulation of the nucleoside triphosphatase/RNA helicase and 5'-RNA triphosphatase activities of Dengue virus type 2 nonstructural protein 3 (NS3) by interaction with NS5, the RNA-dependent RNA polymerase.* J Biol Chem, 2005. **280**(29): p. 27412-9.

90. Du Pont, K.E., R.B. Davidson, M. McCullagh, et al., *Motif V regulates energy transduction between the flavivirus NS3 ATPase and RNA-binding cleft.* J Biol Chem, 2020. **295**(6): p. 1551-1564.

91. Wengler, G. and G. Wengler, *The NS 3 Nonstructural Protein of Flaviviruses Contains an RNA Triphosphatase Activity.* Virology, 1993. **197**(1): p. 265-273.

92. Issur, M., B.J. Geiss, I. Bougie, et al., *The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure.* RNA, 2009. **15**(12): p. 2340-50.

93. Li, Y., M.Y. Lee, Y.R. Loh, et al., *Secondary structure and membrane topology of dengue virus NS4A protein in micelles.* Biochim Biophys Acta Biomembr, 2018. **1860**(2): p. 442-450.

94. Kumar, A., P. Kumar, and R. Giri, *Zika virus NS4A cytosolic region (residues 1-48) is an intrinsically disordered domain and folds upon binding to lipids.* Virology, 2020. **550**: p. 27-36.

95. Kumar, D., A. Kumar, T. Bhardwaj, et al., *Zika virus NS4A N-Terminal region (1-48) acts as a cofactor for inducing NTPase activity of NS3 helicase but not NS3 protease.* Arch Biochem Biophys, 2020. **695**: p. 108631.

96. Miller, S., S. Kastner, J. Krijnse-Locker, et al., *The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner.* J Biol Chem, 2007. **282**(12): p. 8873-82.

97. Tan, M.J.A., N.G. Brown, K.W.K. Chan, et al., *Mutations in the cytoplasmic domain of dengue virus NS4A affect virus fitness and interactions with other non-structural proteins.* J Gen Virol, 2020. **101**(9): p. 941-953.

98. Zou, J., X. Xie, Q.Y. Wang, et al., *Characterization of dengue virus NS4A and NS4B protein interaction.* J Virol, 2015. **89**(7): p. 3455-70.

99. Li, X.D., H.Q. Ye, C.L. Deng, et al., *Genetic interaction between NS4A and NS4B for replication of Japanese encephalitis virus.* J Gen Virol, 2015. **96**(Pt 6): p. 1264-1275.

100. Cortese, M., K. Mulder, L. Chatel-Chaix, et al., *Determinants in Nonstructural Protein 4A of Dengue Virus Required for RNA Replication and Replication Organelle Biogenesis.* J Virol, 2021. **95**(21): p. e0131021.

101. Shah, P.S., N. Link, G.M. Jang, et al., *Comparative Flavivirus-Host Protein Interaction Mapping Reveals Mechanisms of Dengue and Zika Virus Pathogenesis.* Cell, 2018. **175**(7): p. 1931-1945 e18.

102. Liang, Q., Z. Luo, J. Zeng, et al., *Zika Virus NS4A and NS4B Proteins Deregulate Akt-mTOR Signaling in Human Fetal Neural Stem Cells to Inhibit Neurogenesis and Induce Autophagy.* Cell Stem Cell, 2016. **19**(5): p. 663-671.

103. Zhang, J., Y. Lan, M.Y. Li, et al., *Flaviviruses Exploit the Lipid Droplet Protein AUP1 to Trigger Lipophagy and Drive Virus Production.* Cell Host Microbe, 2018. **23**(6): p. 819-831 e5.

104. Yang, Q., J. You, Y. Zhou, et al., *Tick-borne encephalitis virus NS4A ubiquitination antagonizes type I interferon-stimulated STAT1/2 signalling pathway.* Emerg Microbes Infect, 2020. **9**(1): p. 714-726.

105. Lin, C., S.M. Amberg, T.J. Chambers, et al., *Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 proteinase is a prerequisite for processing at the downstream 4A4B signalase site.* Journal of Virology, 1993. **67**(4): p. 2327-35.

106. Zou, G., F. Puig-Basagoiti, B. Zhang, et al., *A single-amino acid substitution in West Nile virus 2K peptide between NS4A and NS4B confers resistance to lycorine, a flavivirus inhibitor.* Virology, 2009. **384**(1): p. 242-52.

107. Mertens, E., A. Kajaste-Rudnitski, S. Torres, et al., *Viral determinants in the NS3 helicase and 2K peptide that promote West Nile virus resistance to antiviral action of 2',5'-oligoadenylate synthetase 1b.* Virology, 2010. **399**(1): p. 176-185.

108. Miller, S., S. Sparacio, and R. Bartenschlager, *Subcellular localization and membrane topology of the Dengue virus type 2 Non-structural protein 4B.* J Biol Chem, 2006. **281**(13): p. 8854-63.

109. Li, Y., Y.M. Kim, J. Zou, et al., *Secondary structure and membrane topology of dengue virus NS4B N-terminal 125 amino acids.* Biochim Biophys Acta, 2015. **1848**(12): p. 3150-7.

110. Kundharapu, S. and T.K. Chowdary, *Dengue Virus NS4b N-Terminus Disordered Region Interacts with NS3 Helicase C-Terminal Subdomain to Enhance Helicase Activity.* Viruses, 2022. **14**(8).

111. Lu, H., Y. Zhan, X. Li, et al., *Novel insights into the function of an N-terminal region of DENV2 NS4B for the optimal helicase activity of NS3.* Virus Res, 2021. **295**: p. 198318.

112. Chatel-Chaix, L., W. Fischl, P. Scaturro, et al., *A Combined Genetic-Proteomic Approach Identifies Residues within Dengue Virus NS4B Critical for Interaction with NS3 and Viral Replication.* J Virol, 2015. **89**(14): p. 7170-86.

113. Zou, J., T. Lee le, Q.Y. Wang, et al., *Mapping the Interactions between the NS4B and NS3 proteins of dengue virus.* J Virol, 2015. **89**(7): p. 3471-83.

114. Umareddy, I., A. Chao, A. Sampath, et al., *Dengue virus NS4B interacts with NS3 and dissociates it from single-stranded RNA.* J Gen Virol, 2006. **87**(Pt 9): p. 2605-2614.

115. Liang, Y., X. Cao, Q. Ding, et al., *Hepatitis C virus NS4B induces the degradation of TRIF to inhibit TLR3-mediated interferon signaling pathway.* PLoS Pathog, 2018. **14**(5): p. e1007075.

116. Zhao, Y., T.S. Soh, J. Zheng, et al., *A crystal structure of the Dengue virus NS5 protein reveals a novel inter-domain interface essential for protein flexibility and virus replication.* PLoS Pathog, 2015. **11**(3): p. e1004682.

117. Lu, G. and P. Gong, *Crystal Structure of the full-length Japanese encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface.* PLoS Pathog, 2013. **9**(8): p. e1003549.

118. Brand, C., M. Bisaillon, and B.J. Geiss, *Organization of the Flavivirus RNA replicase complex.* Wiley Interdiscip Rev RNA, 2017. **8**(6).

119. Brand, C., B.J. Geiss, and M. Bisaillon, *Deciphering the interaction surface between the West Nile virus NS3 and NS5 proteins.* Access Microbiology, 2024. **6**(6).

120. Klema, V.J., M. Ye, A. Hindupur, et al., *Dengue Virus Nonstructural Protein 5 (NS5) Assembles into a Dimer with a Unique Methyltransferase and Polymerase Interface.* PLoS Pathog, 2016. **12**(2): p. e1005451.

121. Zou, G., Y.L. Chen, H. Dong, et al., *Functional analysis of two cavities in flavivirus NS5 polymerase.* J Biol Chem, 2011. **286**(16): p. 14362-72.

122. Teramoto, T., A. Balasubramanian, K.H. Choi, et al., *Serotype-specific interactions among functional domains of dengue virus 2 nonstructural proteins (NS) 5 and NS3 are crucial for viral RNA replication.* J Biol Chem, 2017. **292**(23): p. 9465-9479.

123. Zhao, B., G. Yi, F. Du, et al., *Structure and function of the Zika virus full-length NS5 protein.* Nat Commun, 2017. **8**: p. 14762.

124. Lee, E., P.J. Bujalowski, T. Teramoto, et al., *Structures of flavivirus RNA promoters suggest two binding modes with NS5 polymerase.* Nat Commun, 2021. **12**(1): p. 2530.

125. Xu, S., Y. Ci, L. Wang, et al., *Zika virus NS3 is a canonical RNA helicase stimulated by NS5 RNA polymerase.* Nucleic Acids Res, 2019. **47**(16): p. 8693-8707.

126. Henderson, B.R., B.J. Saeedi, G. Campagnola, et al., *Analysis of RNA binding by the dengue virus NS5 RNA capping enzyme.* PLoS One, 2011. **6**(10): p. e25795.

127. Klema, V.J., R. Padmanabhan, and K.H. Choi, *Flaviviral Replication Complex: Coordination between RNA Synthesis and 5'-RNA Capping.* Viruses, 2015. **7**(8): p. 4640-56.

128. Geiss, B.J., H.J. Stahla-Beek, A.M. Hannah, et al., *A high-throughput screening assay for the identification of flavivirus NS5 capping enzyme GTP-binding inhibitors: implications for antiviral drug development.* J Biomol Screen, 2011. **16**(8): p. 852-61.

129. Feibelman, K.M., B.P. Fuller, L. Li, et al., *Identification of small molecule inhibitors of the Chikungunya virus nsP1 RNA capping enzyme.* Antiviral Res, 2018. **154**: p. 124-131.

130. Zhao, Y., T.S. Soh, S.P. Lim, et al., *Molecular basis for specific viral RNA recognition and 2'-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5).* Proc Natl Acad Sci U S A, 2015. **112**(48): p. 14834-9.

131. Hall, R.A., S.E. Tan, B. Selisko, et al., *Monoclonal antibodies to the West Nile virus NS5 protein map to linear and conformational epitopes in the methyltransferase and polymerase domains.* J Gen Virol, 2009. **90**(Pt 12): p. 2912-2922.

132. Gullberg, R.C., J. Jordan Steel, S.L. Moon, et al., *Oxidative stress influences positive strand RNA virus genome synthesis and capping.* Virology, 2015. **475**: p. 219-29.

133. Noble, C.G., S.P. Lim, R. Arora, et al., *A Conserved Pocket in the Dengue Virus Polymerase Identified through Fragment-based Screening.* J Biol Chem, 2016. **291**(16): p. 8541-8.

134. Lubick, K.J., S.J. Robertson, K.L. McNally, et al., *Flavivirus Antagonism of Type I Interferon Signaling Reveals Prolidase as a Regulator of IFNAR1 Surface Expression.* Cell Host Microbe, 2015. **18**(1): p. 61-74.

135. Ashour, J., M. Laurent-Rolle, P.Y. Shi, et al., *NS5 of dengue virus mediates STAT2 binding and degradation.* J Virol, 2009. **83**(11): p. 5408-18.

136. Mazzon, M., M. Jones, A. Davidson, et al., *Dengue virus NS5 inhibits interferon-alpha signaling by blocking signal transducer and activator of transcription 2 phosphorylation.* J Infect Dis, 2009. **200**(8): p. 1261-70.

137. Morrison, J., M. Laurent-Rolle, A.M. Maestre, et al., *Dengue virus co-opts UBR4 to degrade STAT2 and antagonize type I interferon signaling.* PLoS Pathog, 2013. **9**(3): p. e1003265.

138. Laurent-Rolle, M., E.F. Boer, K.J. Lubick, et al., *The NS5 protein of the virulent West Nile virus NY99 strain is a potent antagonist of type I interferon-mediated JAK-STAT signaling.* J Virol, 2010. **84**(7): p. 3503-15.

139. Ho, L.J., L.F. Hung, C.Y. Weng, et al., *Dengue virus type 2 antagonizes IFN-alpha but not IFN-gamma antiviral effect via down-regulating Tyk2-STAT signaling in the human dendritic cell.* J Immunol, 2005. **174**(12): p. 8163-72.

140. Saez-Alvarez, Y., A. Arias, C. Del Aguila, et al., *Development of a fluorescence-based method for the rapid determination of Zika virus polymerase activity and the screening of antiviral drugs.* Sci Rep, 2019. **9**(1): p. 5397.

141. Watterson, D., B. Kobe, and P.R. Young, *Residues in domain III of the dengue virus envelope glycoprotein involved in cell-surface glycosaminoglycan binding.* J Gen Virol, 2012. **93**(Pt 1): p. 72-82.

142. Wichit, S., A. Jittmittraphap, K.I. Hidari, et al., *Dengue virus type 2 recognizes the carbohydrate moiety of neutral glycosphingolipids in mammalian and mosquito cells.* Microbiol Immunol, 2011. **55**(2): p. 135-40.

143. Jindadamrongwech, S., C. Thepparit, and D.R. Smith, *Identification of GRP 78 (BiP) as a liver cell expressed receptor element for dengue virus serotype 2.* Arch Virol, 2004. **149**(5): p. 915-27.

144. Upanan, S., A. Kuadkitkan, and D.R. Smith, *Identification of dengue virus binding proteins using affinity chromatography.* J Virol Methods, 2008. **151**(2): p. 325-328.

145. Navarro-Sanchez, E., R. Altmeyer, A. Amara, et al., *Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses.* EMBO Rep, 2003. **4**(7): p. 723-8.

146. Tassaneetrithep, B., T.H. Burgess, A. Granelli-Piperno, et al., *DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells.* J Exp Med, 2003. **197**(7): p. 823-9.

147. Miller, J.L., B.J. de Wet, L. Martinez-Pomares, et al., *The mannose receptor mediates dengue virus infection of macrophages.* PLoS Pathog, 2008. **4**(2): p. e17.

148. Meertens, L., X. Carnec, M.P. Lecoin, et al., *The TIM and TAM families of phosphatidylserine receptors mediate dengue virus entry.* Cell Host Microbe, 2012. **12**(4): p. 544-57.

149. Acosta, E.G., V. Castilla, and E.B. Damonte, *Functional entry of dengue virus into Aedes albopictus mosquito cells is dependent on clathrin-mediated endocytosis.* J Gen Virol, 2008. **89**(Pt 2): p. 474-484.

150. Kalia, M., R. Khasa, M. Sharma, et al., *Japanese encephalitis virus infects neuronal cells through a clathrin-independent endocytic mechanism.* J Virol, 2013. **87**(1): p. 148-62.

151. Suksanpaisan, L., T. Susantad, and D.R. Smith, *Characterization of dengue virus entry into HepG2 cells.* J Biomed Sci, 2009. **16**(1): p. 17.

152. Zaitseva, E., S.T. Yang, K. Melikov, et al., *Dengue virus ensures its fusion in late endosomes using compartment-specific lipids.* PLoS Pathog, 2010. **6**(10): p. e1001131.

153. Modis, Y., S. Ogata, D. Clements, et al., *Structure of the dengue virus envelope protein after membrane fusion.* Nature, 2004. **427**(6972): p. 313-9.

154. Falgout, B. and L. Markoff, *Evidence that flavivirus NS1-NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum.* Journal of Virology, 1995. **69**(11): p. 7232-7243.

155. Lin, D.L., T. Inoue, Y.J. Chen, et al., *The ER Membrane Protein Complex Promotes Biogenesis of Dengue and Zika Virus Non-structural Multi-pass Transmembrane Proteins to Support Infection.* Cell Rep, 2019. **27**(6): p. 1666-1674 e4.

156. Stocks, C.E. and M. Lobigs, *Signal Peptidase Cleavage at the Flavivirus C-prM Junction Dependence on the Viral NS2B-3 Protease for Efficient Processing Requires Determinants in C, the Signal Peptide, and prM.* J Virol, 1998. **72**(3): p. 2141-2149.

157. Bera, A.K., R.J. Kuhn, and J.L. Smith, *Functional characterization of cis and trans activity of the Flavivirus NS2B-NS3 protease.* J Biol Chem, 2007. **282**(17): p. 12883-92.

158. Preugschat, F., C.W. Yao, and S.J. H., *In vitro processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3.* J Virol, 1990. **64**(9): p. 4364-74.

159. Amberg, S.M. and C.M. Rice, *Mutagenesis of the NS2B-NS3-Mediated Cleavage Site in the Flavivirus Capsid Protein Demonstrates a Requirement for Coordinated Processing.* J Virol, 1999. **73**(10): p. 8083-94.

160. Lin, Y.J. and S.C. Wu, *Histidine at residue 99 and the transmembrane region of the precursor membrane prM protein are important for the prM-E heterodimeric complex formation of Japanese encephalitis virus.* J Virol, 2005. **79**(13): p. 8535-44.

161. Lorenz, I.C., S.L. Allison, F.X. Heinz, et al., *Folding and dimerization of tick-borne encephalitis virus envelope proteins prM and E in the endoplasmic reticulum.* J Virol, 2002. **76**(11): p. 5480-91.

162. Kuhn, R.J., W. Zhang, M.G. Rossman, et al., *Structure of Dengue Virus Implications for Flavivirus Organization, Maturation, and Fusion.* Cell, 2002. **108**(5): p. 717-25.

163. Uchil, P.D. and V. Satchidanandam, *Architecture of the flaviviral replication complex. Protease, nuclease, and detergents reveal encasement within double-layered membrane compartments.* J Biol Chem, 2003. **278**(27): p. 24388-98.

164. Welsch, S., S. Miller, I. Romero-Brey, et al., *Composition and three-dimensional architecture of the dengue virus replication and assembly sites.* Cell Host Microbe, 2009. **5**(4): p. 365-75.

165. Gillespie, L.K., A. Hoenen, G. Morgan, et al., *The endoplasmic reticulum provides the membrane platform for biogenesis of the flavivirus replication complex.* J Virol, 2010. **84**(20): p. 10438-47.

166. Mackenzie, J.M., M. Jones, and E.G. Westaway, *Markers for trans-Golgi membranes and the intermediate compartment localize to induced membranes with distinct replication functions in flavivirus-infected cells.* J Virol, 1999. **73**(11): p. 9555-67.

167. Cortese, M., S. Goellner, E.G. Acosta, et al., *Ultrastructural Characterization of Zika Virus Replication Factories.* Cell Rep, 2017. **18**(9): p. 2113-2123.

168. Padmanabhan, R., R. Takhampunya, T. Teramoto, et al., *Flavivirus RNA synthesis in vitro.* Methods, 2015. **91**: p. 20-34.

169. Uchil, P.D. and V. Satchidanandam, *Characterization of RNA synthesis, replication mechanism, and in vitro RNA-dependent RNA polymerase activity of japanese encephalitis virus.* Virology, 2003. **307**(2): p. 358-371.

170. de Borba, L., S.M. Villordo, N.G. Iglesias, et al., *Overlapping local and long-range RNA-RNA interactions modulate dengue virus genome cyclization and replication.* J Virol, 2015. **89**(6): p. 3430-7.

171. Friebe, P., J. Pena, M.O. Pohl, et al., *Composition of the sequence downstream of the dengue virus 5' cyclization sequence (dCS) affects viral RNA replication.* Virology, 2012. **422**(2): p. 346-56.

172. Friebe, P. and E. Harris, *Interplay of RNA elements in the dengue virus 5' and 3' ends required for viral RNA replication.* J Virol, 2010. **84**(12): p. 6103-18.

173. Filomatori, C.V., N.G. Iglesias, S.M. Villordo, et al., *RNA sequences and structures required for the recruitment and activity of the dengue virus polymerase.* J Biol Chem, 2011. **286**(9): p. 6929-39.

174. Fredericksen, B.L. and M. Gale, Jr., *West Nile virus evades activation of interferon regulatory factor 3 through RIG-I-dependent and -independent pathways without antagonizing host defense signaling.* J Virol, 2006. **80**(6): p. 2913-23.

175. Zhang, B., H. Dong, Y. Zhou, et al., *Genetic interactions among the West Nile virus methyltransferase, the RNA-dependent RNA polymerase, and the 5' stem-loop of genomic RNA.* J Virol, 2008. **82**(14): p. 7047-58.

176. Bhattacharya, D., S. Hoover, S.P. Falk, et al., *Phosphorylation of yellow fever virus NS5 alters methyltransferase activity.* Virology, 2008. **380**(2): p. 276-84.

177. Geiss, B.J., A.A. Thompson, A.J. Andrews, et al., *Analysis of flavivirus NS5 methyltransferase cap binding.* J Mol Biol, 2009. **385**(5): p. 1643-54.

178. Zhou, Y., D. Ray, Y. Zhao, et al., *Structure and function of flavivirus NS5 methyltransferase.* J Virol, 2007. **81**(8): p. 3891-903.

179. Overby, A.K., V.L. Popov, M. Niedrig, et al., *Tick-borne encephalitis virus delays interferon induction and hides its double-stranded RNA in intracellular membrane vesicles.* J Virol, 2010. **84**(17): p. 8470-83.

180. Yi, G., Y. Wen, C. Shu, et al., *Hepatitis C Virus NS4B Can Suppress STING Accumulation To Evade Innate Immune Responses.* J Virol, 2016. **90**(1): p. 254-65.

181. Fanunza, E., F. Carletti, M. Quartu, et al., *Zika virus NS2A inhibits interferon signaling by degradation of STAT1 and STAT2.* Virulence, 2021. **12**(1): p. 1580-1596.

182. Chahar, H.S., S. Chen, and N. Manjunath, *P-body components LSM1, GW182, DDX3, DDX6 and XRN1 are recruited to WNV replication sites and positively regulate viral replication.* Virology, 2013. **436**(1): p. 1-7.

183. Delorme-Axford, E., E. Abernathy, N.J. Lennemann, et al., *The exoribonuclease Xrn1 is a post-transcriptional negative regulator of autophagy.* Autophagy, 2018. **14**(5): p. 898-912.

184. Ng, C.S., D.M. Kasumba, T. Fujita, et al., *Spatio-temporal characterization of the antiviral activity of the XRN1-DCP1/2 aggregation against cytoplasmic RNA viruses to prevent cell death.* Cell Death Differ, 2020. **27**(8): p. 2363-2382.

185. Silva, P.A., C.F. Pereira, T.J. Dalebout, et al., *An RNA pseudoknot is required for production of yellow fever virus subgenomic RNA by the host nuclease XRN1.* J Virol, 2010. **84**(21): p. 11395-406.

186. Moon, S.L., J.R. Anderson, Y. Kumagai, et al., *A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability.* RNA, 2012. **18**(11): p. 2029-40.

187. Moon, S.L., J.G. Blackinton, J.R. Anderson, et al., *XRN1 stalling in the 5' UTR of Hepatitis C virus and Bovine Viral Diarrhea virus is associated with dysregulated host mRNA stability.* PLoS Pathog, 2015. **11**(3): p. e1004708.

188. Schnettler, E., M.G. Sterken, J.Y. Leung, et al., *Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells.* J Virol, 2012. **86**(24): p. 13486-500.

189. Chang, R.Y., T.W. Hsu, Y.L. Chen, et al., *Japanese encephalitis virus non-coding RNA inhibits activation of interferon by blocking nuclear translocation of interferon regulatory factor 3.* Vet Microbiol, 2013. **166**(1-2): p. 11-21.

190. Blazevic, J., H. Rouha, V. Bradt, et al., *Membrane Anchors of the Structural Flavivirus Proteins and Their Role in Virus Assembly.* J Virol, 2016. **90**(14): p. 6365-6378.

191. Rana, J., J.L. Slon Campos, G. Leccese, et al., *Role of Capsid Anchor in the Morphogenesis of Zika Virus.* J Virol, 2018. **92**(22).

192. Li, L., S.M. Lok, I.M. Yu, et al., *The flavivirus precursor membrane-envelope protein complex structure and maturation.* Science, 2008. **319**(5871): p. 1830-4.

193. Yu, I.M., W. Zhang, H.A. Holdaway, et al., *Structure of the Immature Dengue Virus at Low pH Primes Proteolytic Maturation.* Science, 2008. **319**(5871): p. 1834-7.

194. Li, M.Y., M. Grandadam, K. Kwok, et al., *KDEL Receptors Assist Dengue Virus Exit from the Endoplasmic Reticulum.* Cell Rep, 2015. **10**(9): p. 1496-1507.

195. Stadler, K., S.L. Allison, J. Schalich, et al., *Proteolytic activation of tick-borne encephalitis virus by furin.* Journal of Virology, 1997. **71**(11): p. 8475-81.

196. Zhang, Z., Y. Li, Y.R. Loh, et al., *Crystal structure of unlinked NS2B-NS3 protease from Zika virus.* Science, 2016. **354**(6319): p. 1597-1600.

197. Yu, L., K. Takeda, and L. Markoff, *Protein-protein interactions among West Nile non-structural proteins and transmembrane complex formation in mammalian cells.* Virology, 2013. **446**(1-2): p. 365-77.

198. Osawa, T., M. Aoki, H. Ehara, et al., *Structures of dengue virus RNA replicase complexes.* Mol Cell, 2023. **83**(15): p. 2781-2791 e4.

199. Laudenbach, B.T., K. Krey, Q. Emslander, et al., *NUDT2 initiates viral RNA degradation by removal of 5'-phosphates.* Nat Commun, 2021. **12**(1): p. 6918.

200. Choi, J.H., J.M. Burke, K.H. Szymanik, et al., *DUSP11-mediated control of 5'-triphosphate RNA regulates RIG-I sensitivity.* Genes Dev, 2020. **34**(23-24): p. 1697-1712.

201. Burke, J.M., R.P. Kincaid, R.M. Nottingham, et al., *DUSP11 activity on triphosphorylated transcripts promotes Argonaute association with noncanonical viral microRNAs and regulates steady-state levels of cellular noncoding RNAs.* Genes and Development, 2016. **30**(18): p. 2076-2092.

202. Kincaid, R.P., V.L. Lam, R.P. Chirayil, et al., *RNA triphosphatase DUSP11 enables exonuclease XRN-mediated restriction of hepatitis C virus.* Proc Natl Acad Sci U S A, 2018. **115**(32): p. 8197-8202.

203. Jiao, X., J.H. Chang, T. Kilic, et al., *A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing.* Mol Cell, 2013. **50**(1): p. 104-15.

204. Li, Y., M.G. Song, and M. Kiledjian, *Transcript-specific decapping and regulated stability by the human Dcp2 decapping protein.* Mol Cell Biol, 2008. **28**(3): p. 939-48.

205. van Dijk, E., N. Cougot, S. Meyer, et al., *Human Dcp2 a catalytically active mRNA decapping.* The EMBO Journal, 2002. **21**(24): p. 6915-24.

206. Li, Y., J. Dai, M. Song, et al., *Dcp2 decapping protein modulates mRNA stability of the critical interferon regulatory factor (IRF) IRF-7.* Mol Cell Biol, 2012. **32**(6): p. 1164-72.

207. Sheth, U. and R. Parker, *Decapping and decay of messenger RNA occur in cytoplasmic processing bodies.* Science, 2003. **300**(5620): p. 805-8.

CHAPTER 2 — GENERATION OF WEST NILE VIRUS MONOPHOSPHORYLATED RNA AND SUBGENOMIC RNA IS NOT DEPENDENT ON HOST DCP2 DECAPPING ACTIVITY

2.1 Overview

Flavivirus proteins and RNA play multiple roles to allow the virus to effectively replicate and spread to other host tissues. This is largely due to the optimization of a genome limited by the relatively small capacity of an ~11kb genome with no alternative reading frames. The products of flavivirus genome translation and transcription diligently perform their numerous roles in infected mammalian cells, helping to facilitate replication and deter detection by host immune responses. Nonstructural proteins such as nonstructural protein (NS) 3 and NS5 both have multiple distinct enzymatic roles in the replication process, and the other nonstructural proteins play multiple roles in membrane rearrangement and supporting the aforementioned enzymes [1-14]. The roles of flavivirus proteins and RNA generated during infection extend into controlling and suppressing host cell innate immunity. Nonstructural proteins such as NS2a, NS4a, and NS5 block the signal transduction of the type 1 interferon response network and interfere with the expression of interferon stimulating genes (ISGs) [15-25]. Even the flavivirus genome itself plays a special role in interfering with the host cell immune response through the unique product of its degradation: subgenomic flavivirus RNA (sfRNA).

sfRNA is comprised of portions of a flavivirus genome’s 3’ untranslated region (3’UTR) following failed 5’ to 3’ degradation by the host exoribonuclease Xrn1 [26-28].

Monophosphorylated Xrn1 substrates that resist degradation are known as xrRNA that possess distinct secondary structures that confer Xrn1 resistance [29, 30]. The most critical of the secondary structure types is a three-way junction that organizes itself into a pseudoknot to block Xrn1 degradation near the 5' end of the 3'UTR [31]. Flavivirus species have anywhere from two to five RNA structures present to block Xrn1 from fully dismantling the sfRNA [32-34].Once the enzyme reaches one of these structures of the 3'UTR the RNA "knot" clogs the entry of RNA into the Xrn1 active site, unable to effectively move through the RNA structures and fully degrade the remainder of the flavivirus genome (**Figure 2.1**) [35, 36]. The newly produced sfRNA proceeds to wreak havoc on the cellular immunity. sfRNA interferes with existing host RNA degradation pathways and blocks phosphorylation activation of the critical ISG transcription factors IRF-3 and IRF-7 [16, 26, 37, 38]. The host interferon response is suppressed through the blocking of TRIM25 ubiquitylation by sfRNA [37]. Micro RNA (miRNA) mediated decay pathways are suppressed, further interrupting the host's ability to modulate its own gene expression [39]. This noncoding RNA decay leftover is so important for successful flavivirus infection that, without it, an infection is unsustainable [40].

For Xrn1 to produce sfRNA from a flavivirus genome, it must first be presented with a suitable substrate for its enzymatic activity. Xrn1 specifically degrades RNA that has a monophosphorylated 5' and is blocked by the presence of more phosphates or a cap structure [41]. The current understanding of flavivirus genome replication depicts new viral RNA (vRNA) being dephosphorylated to a diphosphate by NS3 prior to

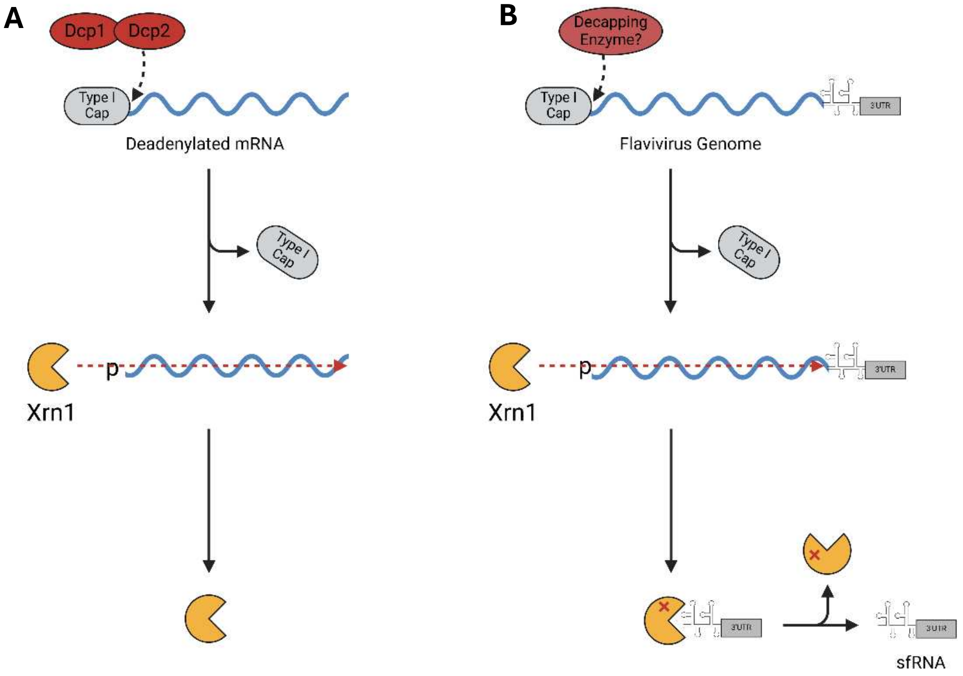


Figure 2.1. Simplified decapping-mediated RNA degradation of host mRNA and flavivirus genomes. A) Traditional decapping and degradation of deadenylated mRNA by the DCP1/DCP2 complex and Xrn1. The host Dcp1/Dcp2 complex decaps the deadenylated mRNA, allowing Xrn1 to degrade the transcript from the 5'→3' direction. **B)** Predicted process of sfRNA generation by decapping and Xrn1 degradation. An unknown host decapping enzyme removes the flavivirus genome cap, allowing Xrn1 to bind to the monophosphorylated 5' end and begin degradation. When Xrn1 reaches the pseudoknots xrRNA structures in the 3'UTR, Xrn1 stalls and fails to degrade the xrRNA. sfRNA is released from the enzyme and proceeds to modulate the host immune response.

attachment of a GMP moiety and methylation by NS5 [6, 42-45]. It is then released from replication compartments embedded in the endoplasmic reticulum as a properly formed flavivirus genome with a 7-methylguanosine cap and a methylated 2'OH on the ribose of the first adenosine in the genome [6, 44, 45]. At this point, the flavivirus is capped or, if capping failed, in a tri- or diphosphorylated when entering the cytoplasm and therefore, not an ideal candidate for Xrn1 degradation. Thus, vRNA needs to undergo modification before Xrn1 can process it into sfRNA.

Normally, Xrn1 substrates are mRNAs that have been decapped and/or dephosphorylated by one of the many enzymes that degrade host transcripts. Phosphatases such as NUDT2 and DUSP11 are known to catalyze the removal of phosphates from triphosphorylated transcripts, leaving monophosphorylated ends open for 5' to 3' nuclease processing [46-49]. If the transcript possesses a guanine cap structure, decapping enzymes can remove said structures and leave behind a monophosphorylated product. Dom3Z, or DXO, is a transcript regulating protein capable of decapping, dephosphorylating, and degrading improperly capped host mRNA that lacks a 2'O methylation [50, 51]. Dcp2 is another key regulator of host gene expression, decapping and dephosphorylating mRNA for subsequent Xrn1 degradation. Xrn1 and Dcp2 are even known to colocalize within processing bodies.[52-54]. NUDT2, DUSP11, DXO, and Dcp2 host proteins are thought to be enzymes that turn triphosphorylated or capped vRNA into the monophosphorylated RNA form needed for Xrn1 degradation and sfRNA generation.

Previously in our lab, we investigated the effect of knocking down the expression of these dephosphorylating and decapping enzymes to assess the impact on the replication of vRNA and viable virion expulsion from infected cells. Dcp2, DXO, DUSP11, and NUDT2 were all suppressed by siRNA knockdown in human Huh7 cells prior to infection with West Nile Virus (WNV), yellow fever virus (YFV), and Zika virus (ZIKV). RNA was extracted from the cells and supernatant following infection and analyzed with reverse transcription (RT) and digital-droplet PCR (ddPCR) for the change in viral copies. Supernatant was assessed by focus forming assays to determine the change in viable virions following the knockdown of these host RNA regulators. Ultimately, it was determined that knocking down Dcp2 resulted in a significant decrease in viral genome copies for the WNV and ZIKV infections, but not the YFV infections (**Figure 2.2a, 2.2b, and 2.3a**). DXO, DUSP11, and NUDT2 knockdowns resulted in no significant change vRNA copy numbers. Dcp2 knockdown also decreased the virus titer levels in all three viruses (**Figure 2.2a, 2.2b, and 2.3b**). These data suggested that Dcp2 plays an important role in effective flavivirus infection.

Wielding this information, we set forth to investigate the specific, mechanistic effect Dcp2 has on flavivirus infection. We hypothesized that Dcp2 was key for dephosphorylating vRNA into a monophosphorylated variant that is broken down by Xrn1 into sfRNA. Identifying Dcp2 as being the first step in generating sfRNA would bring us closer to fully understanding the viral infection event. To do this, we looked at the change in monophosphorylated vRNA levels in infected cells and viral supernatant following Dcp2 knockdown using splint ligation coupled with a tailored qRT-PCR

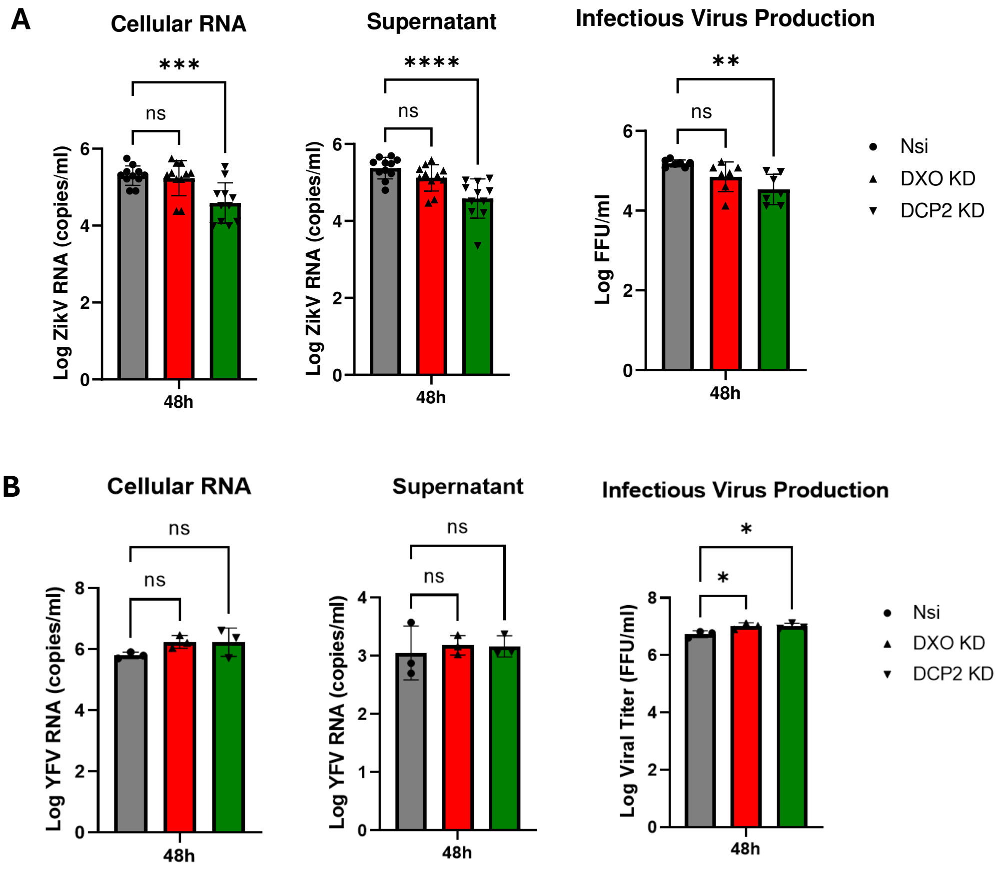


Figure 2.2. Preliminary data showing the effect of Dcp2 and DXO siRNA knockdown on ZIKV and YFV genome copy and virion production. Enzyme knockdowns compared to randomized nonsense siRNA (Nsi) treated samples. **A)** DCP2 but not DXO knockdown reduces the production of Zika virus intracellular RNA, extracellular RNA, and infectious virus production. Huh7 cells were transfected with siRNA against DXO and DCP2. Significant results described as $p<0.05$ (** $p = 0.002$, *** $p = 0.001$, **** $p<0.001$). **B)** DXO and Dcp2 knockdown does not affect YFV RNA production but does increase viral titer generation significantly (* $p<0.05$). Data and graphs generated by Denis.

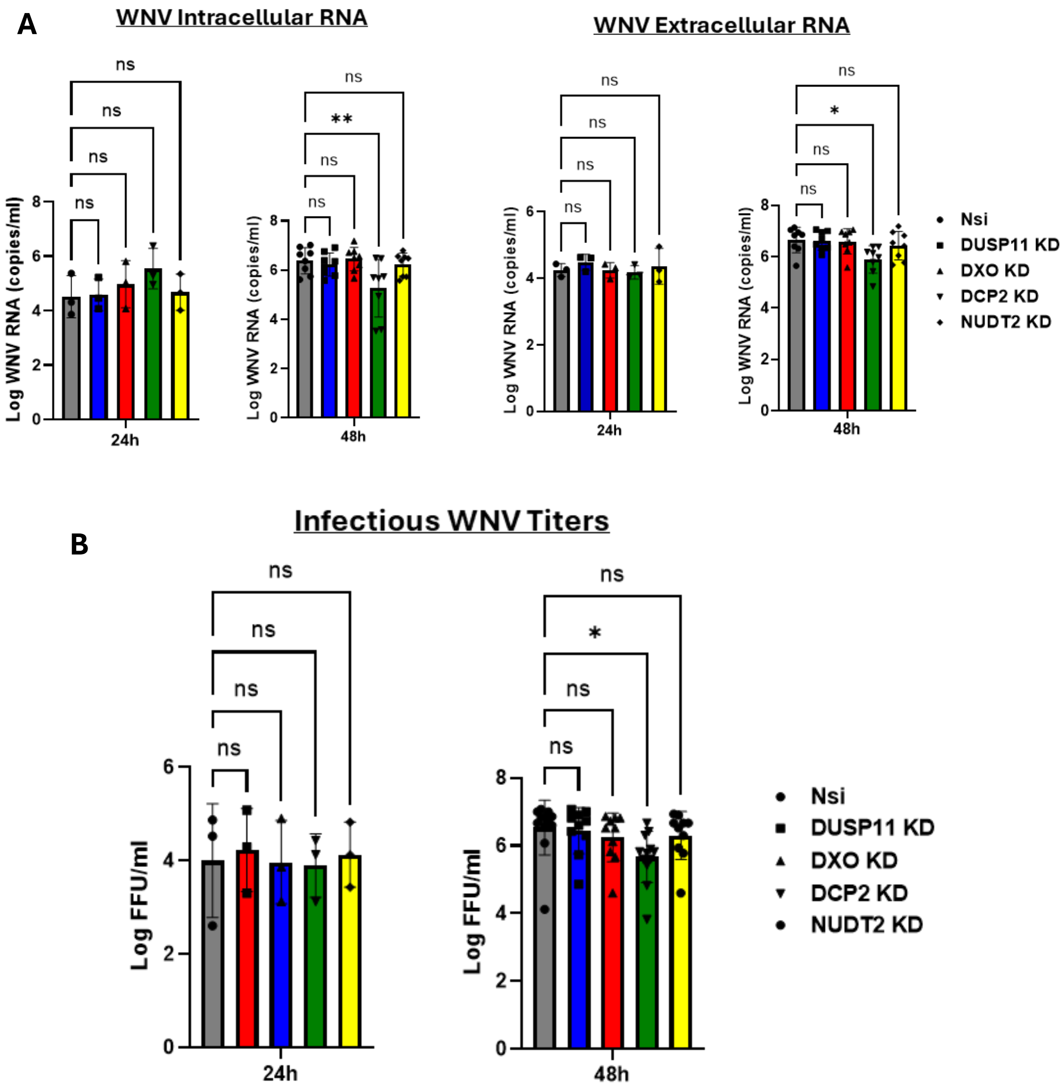


Figure 2.3. Preliminary data showing the effect of Dcp2 and DXO siRNA knockdown on WNV genome copy and virion production. Enzyme knockdowns compared to randomized nonsense siRNA (Nsi) treated samples. **A)** DCP2 but not DXO, DUSP11, or NUDT2 knockdown reduces the production of WNV virus intracellular and extracellular RNA. Statistical analysis was performed using one-way ANOVA test with significant results described as * $p < 0.05$ and ** $p<0.01$. **B)** DCP2 but not DXO, DUSP11, or NUDT2 knockdown resulted in a significant ($p<0.05$) decrease in WNV viral titer after 48 hours. No change was observed in viral titer after 24 hours. Data and graphs generated by Denis.

method. The proportion of capped RNA was assessed by cap pulldown and qRT-PCR analysis while the change in sfRNA from infected cells was assessed with ^{32}P labeled 3'UTR probes on northern blot.

2.2 Materials and Methods

2.2.1 Cell Treatment and Knockdown

Huh7 cells were plated in 6 well plates and allowed to reach 60% confluency before treatment with Dcp2 targeting siRNA delivered by Lipofectamine RNAiMAX Transfection Reagent (Invitrogen #13778075). Three treatment groups were developed: eight Dcp2 siRNA (Invitrogen #AM16708) and RNAiMAX treated replicates, eight randomized nonsense control siRNA (Invitrogen #4390843) treated replicates, and eight untreated replicates. Each biological replicate was comprised of three wells that were pooled at sample collection. 48 hours later, all replicates were infected with a WNV lineage II construct (pBG744) at a multiplicity of infection (MOI) of 0.01 for one hour, followed by a buffer wash, media change, and subsequent 48-hour incubation. Supernatant and cell RNA were collected and purified with TRIzol and phenol/chloroform extraction.

2.2.2 TRIzol and Phenol-Chloroform Extraction of RNA

RNA samples from infected cells and supernatant were isolated by TRIzol extraction followed by phenol-chloroform extraction. 0.66ml of TRIzol Reagent

(Invitrogen #15596026) was added to each of the three wells comprised of a single biological replicate. The solutions were pipetted repeatedly until the sample was homogenized and then combined into a single sample tube. Supernatant from infected cell cultures was combined with TRIzol LS Reagent (Invitrogen #10296028) at a 3:1 ratio as per manufacturer's recommendation. All TRIzol samples were vortexed until clear and allowed to incubate at room temperature for 5 minutes, followed by freezing at -80°C. Samples were thawed and combined with 200µl of chloroform for every 1ml of TRIzol sample. After mixing, the samples were spun at 16,000xg (max speed) for 10 minutes at 4°C. The aqueous layer was carefully aspirated and transferred into a tube with 3µl of glycogen and 500µl of isopropanol for every milliliter in the original TRIzol sample. Samples were spun at max speed for 10 minutes at 4°C and the solution aspirated, leaving a pellet of precipitated RNA, DNA, and glycogen. This pellet was washed with 70% ethanol before air drying for five minutes, covered.

Pellets were resuspended in 160µl RNase-free water and combined with 20µl RQ1 DNase Buffer and 20µl RQ1 DNase (Promega #M6101). Samples were then incubated for 30 minutes at 37°C before being combined with 200µl of phenol-chloroform. Samples were vortexed and centrifuged for 1 minute at max speed. 160µl of the top layer of solution was removed and combined with 66µl 10M ammonium acetate, 4µl of glycogen, and 800µl of 100% ethanol that had been pre-chilled at -20°C. Samples were inverted five times before being incubated at -80°C overnight. The samples were spun at max speed at 4°C for 20 minutes and the resulting pellet washed with 150µl of 70% ethanol. After removing the ethanol wash, the samples were allowed to air dry for

15 minutes, covered, before being resuspended in either 6µl of RNase-free water for splint ligation cleanup or 11µl for original sample purification. If resuspended in 11µl, 1µl of the sample was used to quantify RNA with the Qubit RNA High Sensitivity system.

2.2.3 *In Vitro* Transcription of RNA Adaptor and Controls

To transcribe the RNA adaptor, a DNA oligo template (BG2006) was manufactured by Integrated DNA Technologies that contained an SP6 promoter sequence immediately prior to the adaptor transcription sequence (**Table 2.1**). SP6 RNA polymerase (Thermo Scientific # EP0131) was used to create the adaptor from the template oligo, followed by DNase treatment and phenol-chloroform extraction. Integrated DNA Technologies manufactured a DNA oligo template for the control vRNA (BG1900), containing the first 127 bases of the WNV genome preceded by a modified T7 promoter sequence (**Table 2.1**). To create the pppRNA control molecule, T7 RNA polymerase enzyme (Thermo Scientific #EP0111) was used for transcription as per the manufacturers protocol. To create pRNA and ppRNA control molecules, the same protocol was repeated using 0.5mM ATP and 5mM AMP and 5mM ADP, respectively. The transcription products were followed up by DNase I treatment and phenol-chloroform extraction. To produce the GpppRNA control, pppRNA was treated with the NEB Vaccinia Capping System as per manufacturer protocol and purified by phenol-chloroform extraction.

Table 2.1. Primers, oligos, and templates used for qRT-PCR and reagent transcription

Primer Name	Sequence
BG1816	TCCGCCACAGTCAGCAGTTAT
BG1817	TCTTCACTGGAGCTAGGCAAGTC
BG1886	ACACTACTAAGTTTGTCAGCTCACACAGGCGAACTACTCATCAAAGCCAGCAAACGCAGTGTTCATTCATCGCCAT
BG1900	ATCCATTAATACGACTCACTATAAGTAGTTCGCCTGTGTGAGCTGACAAACTTAGTAGTGTTTGTGAGGATTAACAACAATTAACACAGTGCGAGCTGTTTCTTGGCACGAAGATCTCGATGTCTAAGAAACCAGGAGGGCCCGGTAAAA
BG2006	CACACGCGCGCACACTATTGAATCAAACAGCCGACCAATTGCACTACCATCACAATCCATTAATTTAGGTGACACTATAGGAAGCTGATGGCGATGAATGAACACTGCGTTTGCTGGCTTTGATG
BG2007	ATCGCGTAGCTAGCTAGCTTCGAATGCTGGTTTCTTAGACATCGAGATCTTCG
BG2037	CCTGTGTGAGCTGACAAACTTAGTAGTGTTTGTGAGGATTAACAACAATTAACACAGTGCGAGCTGTTTCTTGGCACGAAGATCTCGATGTCTAAGAAACCAGCATTCGAAGCTAGCTAGCTACGCGAT
BG2038	GATGGCGATGAATGAACACTGCGTTTGCTGGCTTTGATGAGTAGTTCGCCTGTGTGAGCTGACAAACTTAGTAGTGTTTGTGAGGATTAACAACAATTAACACAGTGCGAGCTGTTTCTTGGCACGAAGATCTCGATGTCTAAGAAACCAGCATTCGAAGCTAGCTAGCTACGCGAT
BG2052	GCGATGAATGAACACTGCGTTTGC
BG2053	TGAGCTGACAAACTTAGTAGTGTTTGTGAGG
BG2054	GTAGCTAGCTAGCTTCGAATGCTGG
BG2114	ACCACAGTCCATGCCATCAC
BG2115	TCCACCACCCTGTTGCTGTA

2.2.4 Radiolabeled Vaccinia Capping Protocol

Supernatant and cell RNA were collected from WNV infected HEK 293T cells and purified by TRIzol, DNase I digestion, and phenol-chloroform extraction as described previously. The infected cell RNA, infected supernatant RNA, pRNA control, ppRNA control, and pppRNA controls were treated with the NEB Vaccinia Capping System per manufacturers protocol, with normal GTP being replaced with ^{32}P radiolabeled GTP. Following capping the RNA was extracted using phenol-chloroform extraction before separation on a 5% polyacrylamide gel containing 8M urea at 600V. The resulting gel was dried and exposed on a phosphor screen for 48 hours. Imaging was accomplished using a Sapphire FL by Azure Biosystems.

2.2.5 Splint Ligation and qRT-PCR Pipeline

Splint ligation of the adaptor RNA to sample RNA was adapted from Blewett et al. and the SplintR Ligase manufacturer protocol (New England Biolabs #M0375) [55]. The appropriate mass of sample RNA was combined with 1µl of 0.5µM splint DNA (BG1886) diluted in NEBuffer 2.1r (New England Biolabs #B6002S) and 1µl of the adaptor RNA at 13.2ng/µl (**Table 2.1**). The mixture was heated to 95°C and step-down cooled to 22°C at a rate of 0.5°C per minute to allow components to properly anneal. 25 units of SplintR Ligase were added in addition to the 2µl of 10X ligase buffer and diluted to 20µl, resulting in a final buffer concentration of 1X. Samples were then incubated at 25°C for 1 hour, treated with DNase I and extracted by phenol-chloroform as described previously.

Purified samples then underwent RT using ImProm II Reverse Transcriptase (Promega #A3803) and a reverse primer (BG2007) specific to the 5' UTR starting at nucleotide 86 and containing a 5' tag. The resulting cDNA was then used for both splint-ligation product PCR (APCR) and vRNA PCR (VPCR) reactions. APCR primers (BG2052 and BG2054) targeted both the adaptor sequence and the tag sequence appended to the cDNA by the tagged RT primer, while VPCR primers (BG2053 and BG2054) targeted the 5'UTR starting at nucleotide 16 and the same tag sequence (**Table 2.1**). Both PCR reactions used primers at a concentration of 5mM and PowerUp SYBR Green Master Mix. Reactions were measured using QuantStudio 3 (Applied Biosystems) under fast run conditions and data analyzed with Design and Analysis 2.7.0 software.

A standard curve was established using an APCR product control oligo (BG2038) and VPCR product control oligo (BG2037) and the respective primer pairs for each PCR type (**Table 2.1**). Using the trendline equation for the standard curves, the copies of APCR products and VPCR products were determined and compared in an APCR:VPCR ratio.

2.2.6 Methyl-7 RNA Cap Pulldown

The immunoprecipitation method for pulldown of virion capped vRNA was adapted from a Wilusz lab protocol. 15ng of supernatant RNA was combined with 1µl of Ribolock RNase inhibitor and 1µg of anti-m^7G cap antibody (MBL #RN016M) and incubated for 2 hours at 4°C on an inversion mixer. 25μl of Pierce™ Protein G Magnetic Beads (Thermo Scientific #88847) were washed with 250μl of RIPA buffer (Thermo Scientific #89900) and 1X Halt™ Protease Inhibitor (Thermo Scientific #87786), separated using a magnetic stand, and resuspended in 25μl of RIPA buffer. The 25μl of resuspended beads were added to the antibody-RNA mixture tubes and incubated for 1 hour at 4°C on the inversion mixer. The beads were separated from the supernatant on a magnetic stand and the supernatant was retained as the uncapped vRNA-containing "flow-through" fraction. The beads were washed thrice with 600μl of RIPA buffer and once with 600μl of 1X PBS plus 0.1% Tween20 detergent. Following the washes, the beads were incubated with 100μl TEDS buffer (10mM Tris-Cl, 1mM EDTA, 1% Sodium cholate, 0.1% NP-40) with 1mM DTT at 70°C for 45 minutes. With the captured capped vRNA eluted off the beads, the solution was collected as the capped RNA fraction. Both the capped RNA fraction and "flow-through" fraction were subjected to phenol-chloroform extraction as previously described but without DNase treatment. The samples were then analyzed using the VPCR qRT-PCR protocol from before and the number of copies determined from the established standard curve. The proportion of capped vRNA was assessed as a percentage of total vRNA, obtained by combining the copies from both fractions together.

2.2.7 Northern Blot Method

For detecting sfRNA in knockdown samples, 1μg of Dcp2 knockdown, nonsense siRNA treated, and untreated infected control cell RNA samples were separated on a 5% polyacrylamide gel with 7M urea. Separated RNA was transferred onto a Amersham

Hybond-XL (Cytiva RPN1210S) nylon membrane overnight at 10V and 4°C. Transferred blots were crosslinked by UV exposure and incubated in hybridization buffer as performed in Michalski et al. [26]. Following hybridization, the blot was incubated overnight at 60°C with a transcribed ^{32}P GTP radiolabeled probe corresponding to the sfRNA1 sequence for WNV. The blot was washed twice with low-stringency (0.2X) saline sodium citrate (SSC) buffer with 0.1% SDS followed by two washes with high stringency (2X) SSC buffer with 0.1% SDS, with each wash lasting 15 minutes at 60°C. Probed northern blots were visualized using a 24-hour phosphor screen exposure and read with an Azure Biosystems Sapphire FL imager. Densitometry analysis of sfRAN1 bands was performed using Image-J [56].

2.3 Results

2.3.1 Splint Ligation System, qRT-PCR Protocol, and Control RNA Generation

The hypothesized role of Dcp2 in flavivirus infection revolved around the decapping enzyme producing monophosphorylated RNA for sfRNA generation. From our lab's preliminary experiments, it appeared that cause of reduced infectivity from Dcp2 knockdown may have been the result of a reduced concentration of monophosphorylated vRNA, leading to reduced infection efficiency from the lack of sfRNA. Thus, we would have to look at the amount of monophosphorylated RNA in knockdown samples to determine if Dcp2 was responsible for its generation. To assess the proportion of monophosphorylated RNA in experimental samples, we decided to use a modified splint ligation protocol adapted from Blewett et al. [55]. Briefly, we designed a

DNA oligomer, termed the “splint”, that is complementary to the last 38 bases of an “adaptor” sequence and the first 38 bases of the WNV genome in sequential order. The RNA adaptor sequence is an RNA transcript transcribed off a gBlock using an SP6 promoter. When the adaptor, splint, and monophosphorylated WNV RNA are incubated together, the splint will anneal to both RNA molecules and position the 3’ end of the adaptor immediately before the monophosphorylated 5’ end of the vRNA (**Figure 2.4**). Treatment with SplintR Ligase then ligates the two RNA molecules together into a single strand that can be detected by qRT-PCR. Because of the required monophosphate at the 5’ end of vRNA and complementary sequence on the splint, the ligation process is highly specific and sufficient for reporting the presence of monophosphorylated RNA in our samples. Following this reaction, samples undergo DNase I digestion and phenol-chloroform extraction, leaving a sample with free adaptor, RNA, and ligated product.

To use this system, we needed control molecules for capped, tri-, di-, and monophosphorylated vRNA. We developed a gBlock containing a modified T7 promoter and the first 127 bases of the WNV genome to use as a template for the various phosphorylation state controls. The promoter modification changed the final nucleotide of the T7 promoter sequence to a guanine to enhance transcription. Control RNA (pppRNA) was generated using the Invitrogen ImProm II Reverse Transcriptase following the manufacturers protocol. Monophosphorylated control RNA (pRNA) was made by decreasing the ATP added to the reaction by a factor of 10 (0.5mM) and adding 5mM of AMP, producing RNA with the majority being monophosphorylated. The pRNA control will always have residual quantities of triphosphorylated RNA due to the

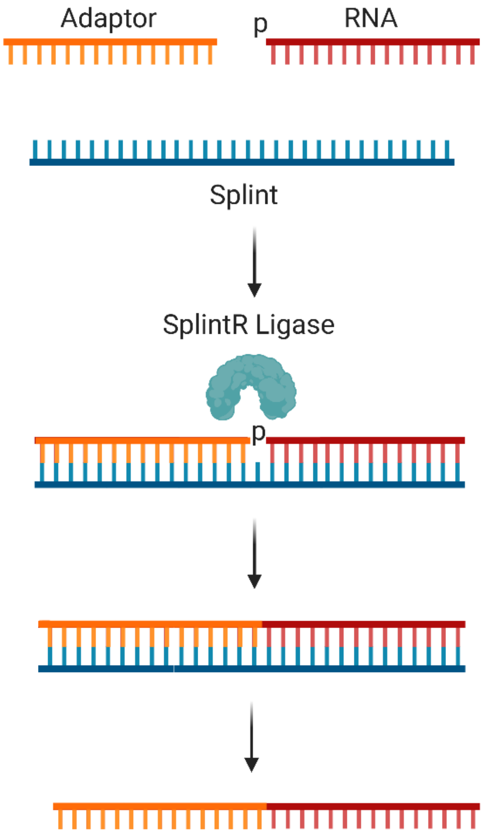


Figure 2.4. Splint Ligation method for labeling monophosphorylated RNA. Process of splint ligation, where a monophosphorylated RNA and adaptor RNA hybridize to a complimentary splint oligo, positioning the 5’ end of the monophosphorylated RNA immediately after the 3’ end of the adaptor RNA. SplintR Ligase (NEB) ligates the adaptor and monophosphorylated RNA together. Created with BioRender.com.

occasional integration of the scarce ATP at the 5’ end of the RNA. Diphosphorylated control RNA (ppRNA) was produced in a similar manner but using 5mM ADP. Capped control RNA (GpppRNA) was produced by treating pppRNA with vaccinia capping enzyme (New England Biolabs #M2080). To determine the relative purity of the pRNA control, the pppRNA, ppRNA, and pRNA controls underwent capping with the same vaccinia capping system using ^{32}P labeled GTP and were subsequently run on an 8M urea gel (**Figure 2.5a**). The radiolabeled gel was exposed to a phosphor screen over 48 hours. The screen was imaged by an Azure Sapphire imager with the upwards shifted bands demonstrating that the pppRNA and ppRNA controls were capable of being capped. This was expected, as pppRNA and ppRNA are suitable substrates for the vaccinia capping enzyme. The pRNA sample showed minimal capping, which is expected from the minute traces of triphosphorylated RNA in the sample from the transcription. With this in mind, we subjected the pRNA control to the splint ligation reaction, using an adaptor that was transcribed with ^{32}P radiolabeled GTP and resolved these samples on a urea gel (**Figure 2.5b**). The lane corresponding to a full reaction of radiolabeled adaptor, splint, and pRNA showed a dramatic band shift upwards, indicating successful ligation of the adaptor to the pRNA.

With the pRNA control produced and validated, we then moved to develop and validate a qRT-PCR reporting system for the splint ligation. One PCR primer set targeted adapter-ligated monophosphorylated vRNA and another amplified all vRNA. With these primer sets we could determine the proportion of monophosphorylated RNA in the sample by comparing the copy numbers. Both PCR reactions used the same

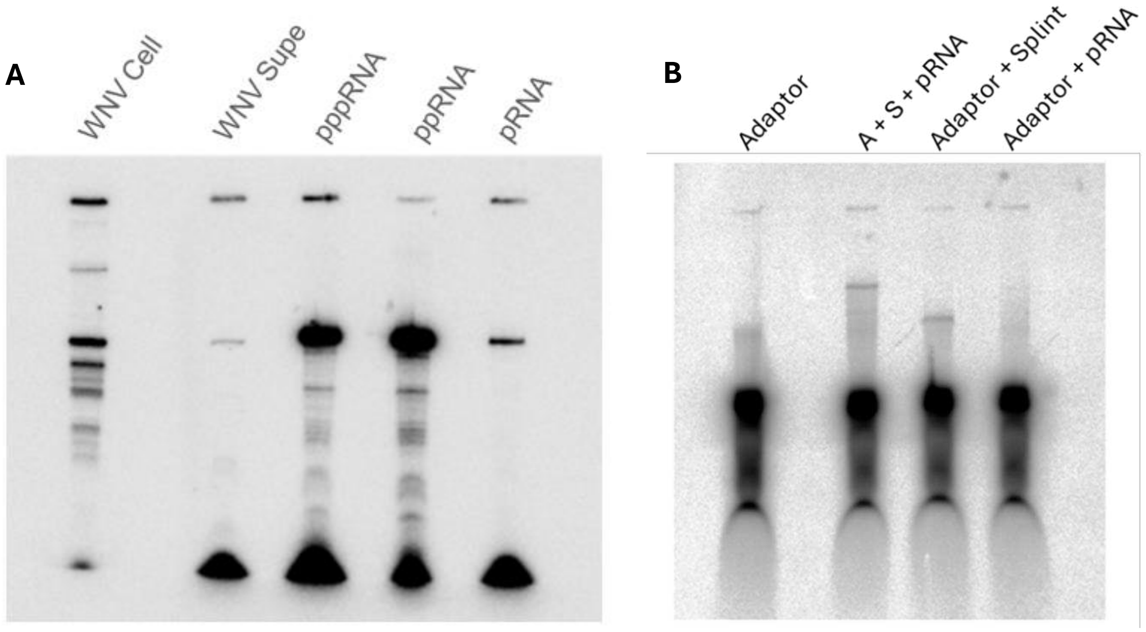


Figure 2.5. Validation of RNA control molecules by vaccinia capping and splint-ligation. Control RNA molecules containing the first 127 bases of the 5’UTR of WNV were generated using *in vitro* transcription and capping protocols: triphosphorylated RNA (pppRNA), dephosphorylated (ppRNA), monophosphorylated RNA (pRNA), and capped RNA (GpppRNA). **A)** Urea gel with control molecules capped with P^{32} radiolabeled GTP (NEB Vaccinia Capping System), where only pppRNA and ppRNA are capable by the enzyme to produced capped RNA (GpppRNA). Residual capping occurs in pRNA sample due to trace amounts of pppRNA being produced in the transcription process. **B)** Urea gel with splint-ligation products following ligation of pRNA with P^{32} radiolabeled adaptor molecules. Gel shift pattern of radiolabeled adaptor indicates successful ligation of adaptor to pRNA.

cDNA made from a single RT primer with a tag overhang. Both reactions also used the same reverse PCR primer that bound to said tag. The PCR primer set targeting vRNA that had successfully ligated to the adaptor was termed APCR and used a forward primer that annealed to the adaptor sequence. The second primer set targeting all vRNA was termed VPCR, with the forward primer annealing just downstream of the 5' end of the genome (**Figure 2.6a**). Copies of ligated RNA and vRNA can be determined from qPCR Cq values using a standard curve, which we developed for APCR and VPCR using serially diluted DNA oligomers corresponding to ligated viral RNA sequences. Together, the copies of APCR products and VPCR products can be compared in a ratio to determine the proportion of viral RNA that is monophosphorylated in a sample. We tested the pRNA, ppRNA, pppRNA, and GpppRNA control strands with both APCR and VPCR. From the standard curve we determined the mass and then copies of the target strands for each PCR reaction. With these viral genome copy counts we determined the ratio of APCR:VPCR, giving the number of monophosphorylated RNA copies per vRNA copy. As seen in **Figure 2.6b**, the pRNA control worked significantly better than the other controls, although efficiency was far from ideal. If the splint ligation reaction was 100% efficient, the APCR:VPCR ratio for the pRNA control would equal 1. The SplintR Ligase enzyme protocol was designed around micrograms of RNA, while we were working with nanogram levels of reagents. From the phosphorylation control splint ligation experiments we determined that the splint ligation reaction was capable of specifically interacting with the pRNA control, allowing the monophosphorylated vRNA proportions to be compared between siRNA treatments. The pRNA control would serve as the ligation efficiency control henceforth.

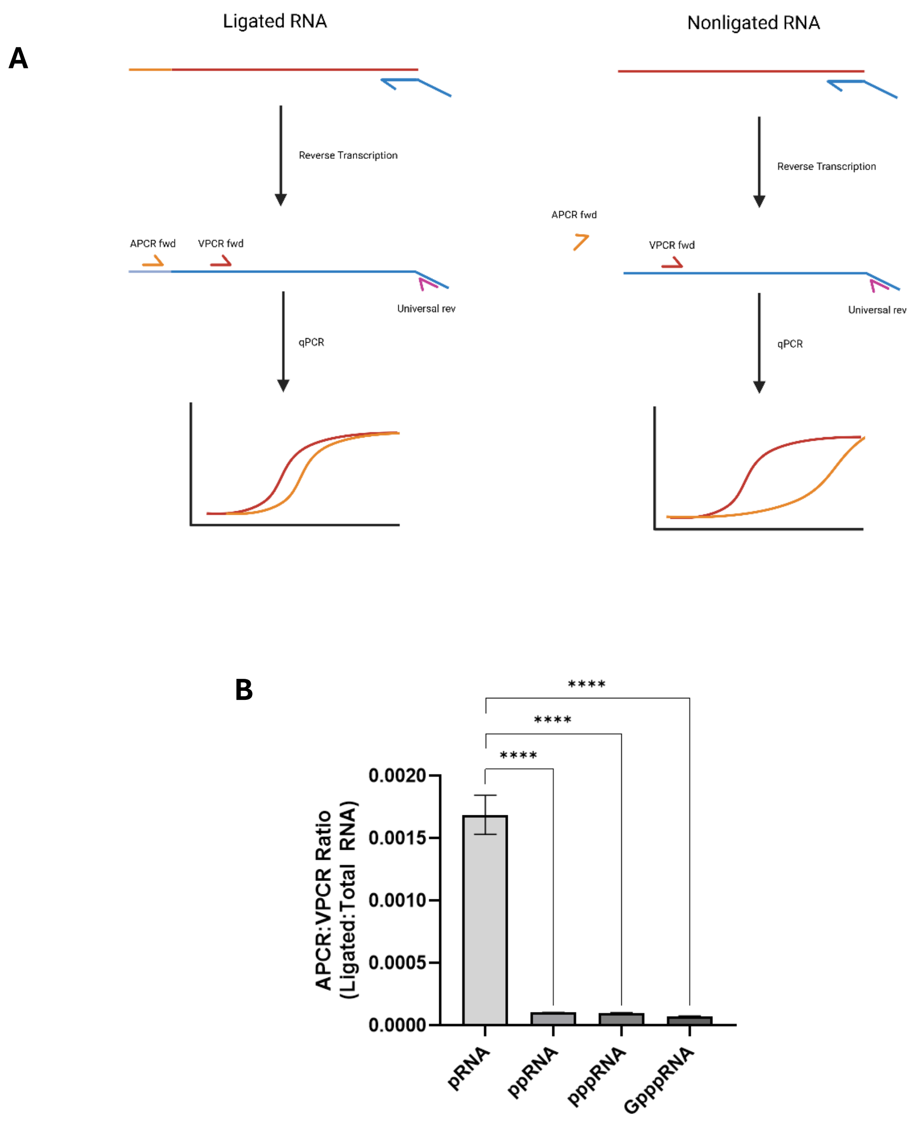


Figure 2.6. qRT-PCR analysis pipeline and control RNA results. A) Potential qRT-PCR outcomes of APCR and VPCR for ligated and nonligated RNA, respectively. Created with BioRender.com **B)** Ratio of ligated RNA copies to total viral RNA copies (APCR:VPCR ratio) for control molecules treated with splint-ligation and qRT-PCR protocol. pRNA APCR:VPCR ratio in **B** is significantly higher compared to other phosphorylation states, indicating more splint-ligation of adaptor to monophosphorylated RNA (**** $p<0.0001$)

2.3.2 Dcp2 Knockdown Sample Generation and Validation

To generate Dcp2 knockdown samples, the protocol for the preliminary knockdown experiment was repeated. Two siRNA constructs were used: one that targeted the Dcp2 mRNA transcript (Invitrogen #AM16708), and another that targeted a randomized nonsense siRNA as a control (Invitrogen #4390843). Both siRNA types were transfected into Huh7 cells using Lipofectamine RNAiMAX (Invitrogen #13778075). There were three designated treatment groups, each with eight biological replicates: the experimental Dcp2 knockdown samples, nonsense siRNA treated control samples, and untreated samples. Huh7 cells at 60% confluency were treated according to their group's designation. All three groups were infected two days later with a WNV lineage 2 replicon from our lab. Following a two-day incubation, cellular and viral RNA were purified by TRIzol and phenol-chloroform for assessment. To confirm Dcp2 knockdown, isolated cell RNA underwent reverse transcription with an oligo-dT primer to generate cDNA from host transcripts. Primers specific to human Dcp2 (BG1816 and BG1817) were developed and premade primers for human GAPDH (BG2114 and BG2115) were purchased from IDT (**Table 2.1**). With these primers, we assessed the relative change in Dcp2 expression by qPCR with PowerUp SYBR Green Master Mix (Applied Biosystems #A25742) and a QuantStudio 3 qPCR machine (Applied Biosystems). Using GAPDH as a housekeeping gene, the relative Dcp2 expression levels were observed once normalized to the untreated control sample. Dcp2 was observed to have been significantly knocked down compared to the nonsense siRNA treated samples with a p-value of 0.0001, confirming that replication of the preliminary experiment's methods was successful (**Figure 2.7**). Confirming the Dcp2 knockdown allowed us to move forward with analyzing the sample vRNA.

2.3.3 Dcp2 Knockdown's Effect on Capped vRNA in Virions

In previously unpublished data from the lab of Dr. Jeff Wilusz at CSU, methyl-7 cap-immunoprecipitation shows that roughly 60% of virions contained uncapped vRNA. We assessed the proportion of capped vRNA in our WNV infected supernatant samples to see if Dcp2 knockdown impacted the ability for infected cells to produce and package capped RNA. Using a protocol adapted from the Wilusz Lab, we incubated supernatant RNA with mouse antibodies targeting the methyl-7 guanine cap structures found on the 5' end of capped flavivirus genomes. Then the samples were incubated with magnetic beads bound to protein G, which we used to pull the capped RNA-antibody hybrids out of solution. Unbound RNA was kept and used for further analysis down the line. After several washes, the capped vRNA was eluted off the magnetic beads. RNA from the eluted volume and the unbound, uncapped RNA in the "flow-through" volume was phenol-chloroform extracted and quantified by Qubit RNA High Sensitivity Kit (Invitrogen #Q32852). For each sample, equivalent masses of eluate RNA and "flow-through" RNA underwent reverse transcription with a WNV specific RT primer. The resulting cDNA then underwent the validated VPCR qPCR protocol. The percentage of viral copies was calculated using the Cq values from the QuantStudio 3 and the established standard curve. The Dcp2 knockdown, nonsense siRNA, and untreated infected groups contained an average of roughly 60% capped vRNA. While slightly elevated, there was no significant difference in the percentage of vRNA that was capped compared to the

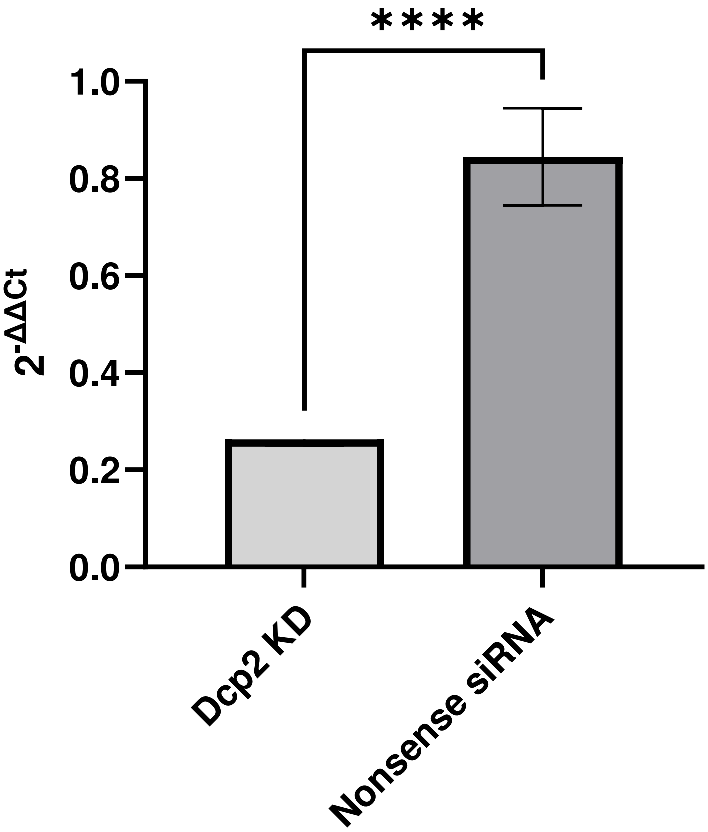


Figure 2.7. Dcp2 expression levels in samples knocked down by Dcp2 targeting siRNA. Huh7 cells were treated with Dcp2-targting siRNA and control randomized nonsense siRNA, along with an untreated control. Groups were infected with WNV and cell RNA extracted. Expression of Dcp2 analyzed by qRT-PCR with GAPDH as a housekeeping gene. Expression levels normalized to untreated control. Significant change in expression between Dcp2 knockdown and nonsense siRNA treated control (p<0.0001)

nonsense siRNA treated and untreated control sample averages (**Figure 2.8a**). Total vRNA in supernatant samples was normalized to the untreated infected control. No significant change in total vRNA was observed in the supernatant, contrary to preliminary data (**Figure 2.8b**). These data indicate that knocking down Dcp2 did not significantly affect total vRNA in virions, nor the proportion of vRNA that was capped in excreted virions during infection.

2.3.4 Splint Ligation of Experimental Samples

The monophosphorylated content of the purified cell and supernatant RNA samples were assessed with the samples subjected to the splint ligation protocol. DNase 1 digestion following splint ligation removed excess splint oligo, and all the purified RNA was reverse transcribed using the 5' tagged reverse transcription primer. Equivalent volumes of cDNA were taken from each sample and added to the APCR and VPCR reactions with PowerUp SYBR Green Master Mix and read on a QuantStudio 3 real-time PCR machine. The resulting Cq values were used to calculate mass of PCR product and thus the number of copies based on the product length. The ratio of APCR copies to VPCR copies was determined for each sample and compared across treatment groups within sample types. Analysis of supernatant samples from the three treatment groups showed that there was no significant change in the proportion of monophosphorylated viral RNA (**Figure 2.9)**. Z -score analysis of the Dcp2 knockdown samples removed two outliers. This information was consistent with the preliminary cap-immunoprecipitation analysis, ultimately confirming that Dcp2 suppression did not

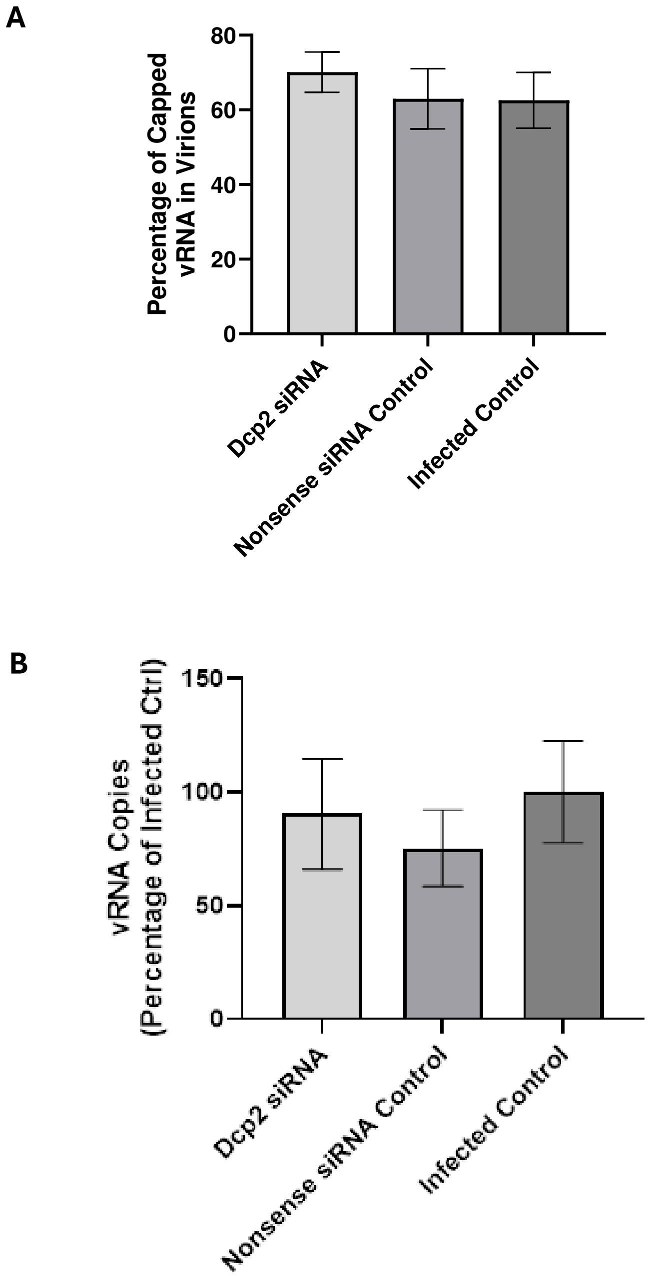


Figure 2.8. WNV RNA content of knockdown treatment supernatant samples. A) Proportion of capped WNV RNA following siRNA treatment in supernatant RNA samples. The observed proportion of capped vRNA is consistent with previous data and not significantly changed by treatment. **B)** Relative WNV RNA copies in supernatant samples compared to untreated infected controls. No significant change ($p>0.05$) was observed between treatment groups and the baseline untreated group.

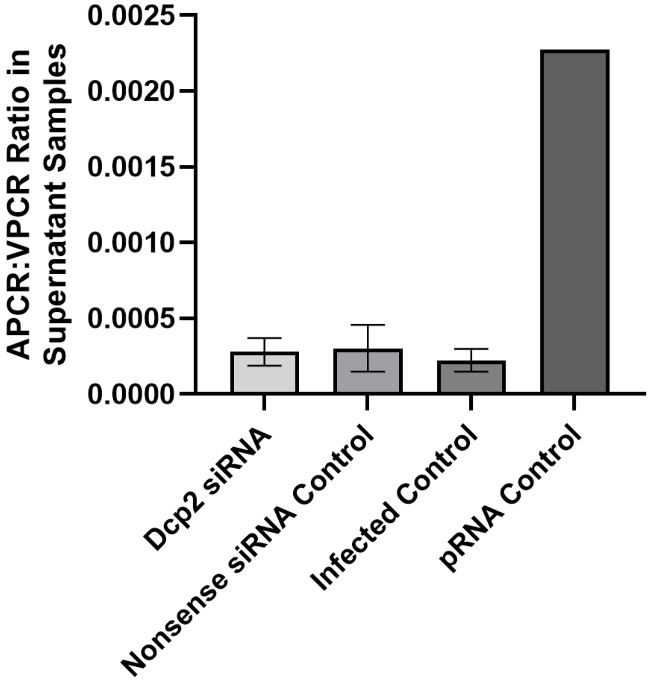


Figure 2.9. APCR:VPCR ratios of Dcp2 knockdown supernatant RNA samples. Ratio of ligated RNA copies to nonligated RNA copies (APCR:VPCR) in supernatant samples following splint ligation and qRT-PCR pipeline. No significant change observed between the Dcp2 siRNA, nonsense siRNA, and untreated treatment groups ($p>0.05$)

significantly affect the proportion of capped vRNA packaged in excreted virions during infection.

Surprisingly, cell samples with knocked down Dcp2 demonstrated an APCR:VPCR ratio that was significantly higher than the nonsense siRNA treated samples and untreated infected control (**Figure 2.10**). This was unexpected and disproved our primary hypothesis by showing that knockdown of Dcp2 did *not* have a deleterious effect on the proportion of vRNA that was monophosphorylated. Instead, the proportion spiked sharply as a result of suppressing the decapping effect of Dcp2.

2.3.5 Northern Blot Analysis of sfRNA

Having observed the increase in monophosphorylated RNA following Dcp2 knockdown, we elected to assess the sfRNA content of our treatment groups. We selected three biological replicates with the highest monophosphorylated RNA proportions and ran their cell sample RNA on 5% urea gels for separation. We used GAPDH to normalize the amount of RNA added for each sample based on the GAPDH content of Dcp2 knockdown sample 3. A marker the same size as the WNV sfRNA 1 was generated using T7 transcriptase for validating the sfRNA 1 band in RNA sample lanes. The samples in the urea gel were transferred to a Hybond-XL nylon membrane, UV crosslinked, hybridized, and probed using a ^{32}P radiolabeled probe corresponding to the same sequence as the marker. Phosphor screen imaging revealed sfRNA 1 and the smaller sfRNA 2 bands in all 9 sample lanes (**Figure 2.11a** demonstrates representative

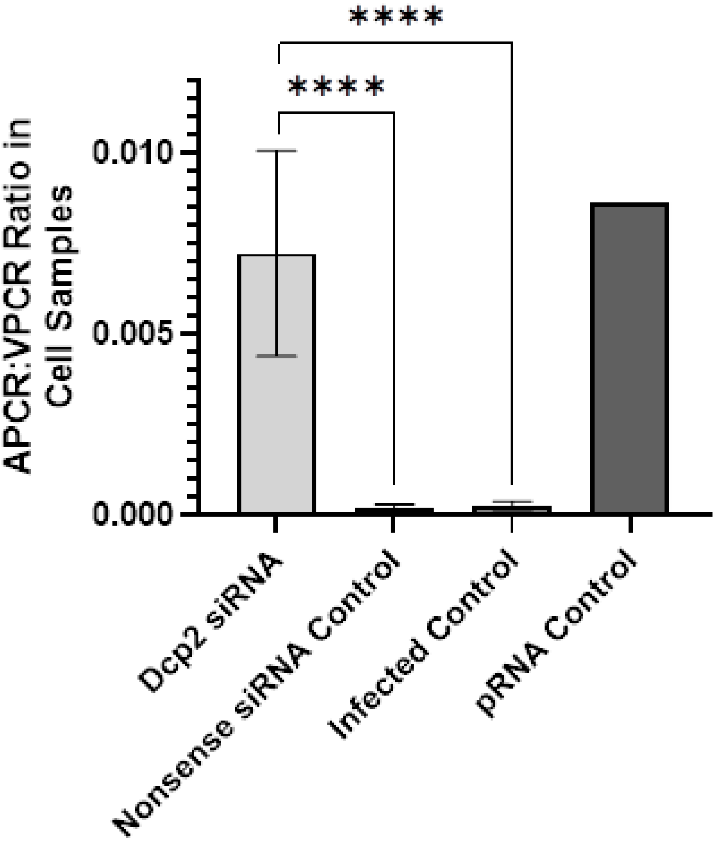


Figure 2.10. APCR:VPCR ratios of Dcp2 knockdown cell RNA samples. Ratio of ligated RNA copies to nonligated RNA copies (APCR:VPCR) in cell samples following splint ligation and qRT-PCR pipeline. Dcp2 siRNA treated samples had a significantly higher APCR:VPCR ratio compared to the other treatment groups (**** p<0.0001)

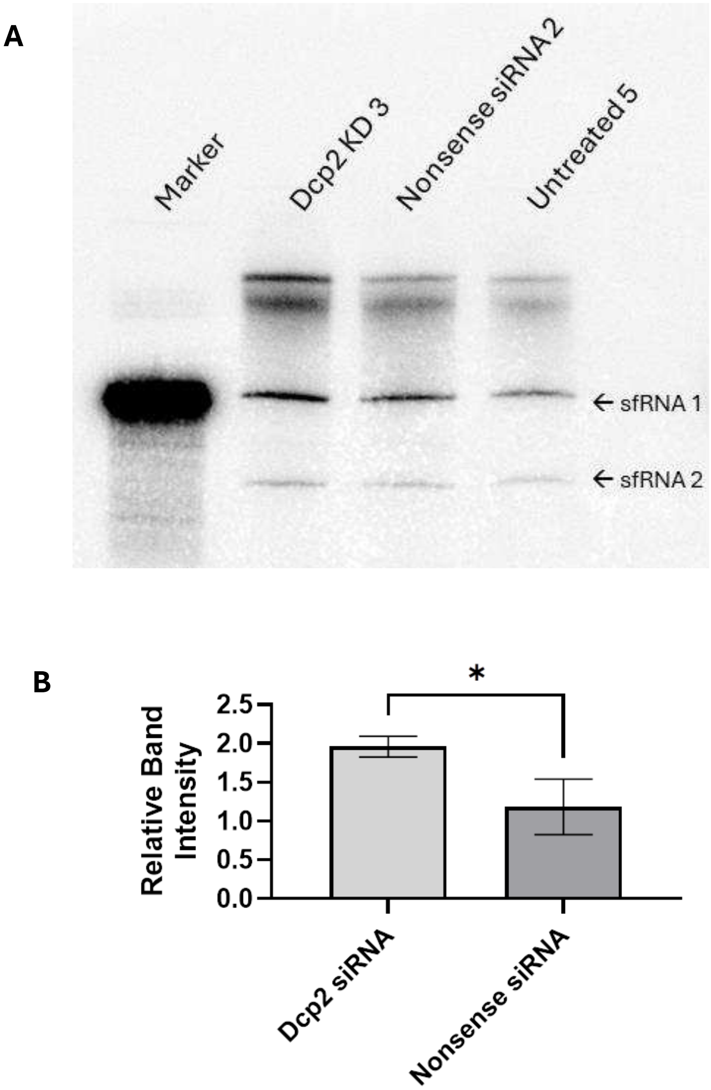


Figure 2.11. Northern blot analysis of cell RNA samples for WNV sfRNA content. A) Cell RNA samples from each treatment group were selected based on their monophosphorylated RNA content within their group. Blot was probed by a P^{32} radiolabeled probe that hybridizes to the full length WNV sfRNA sequence and imaged on a phosphor screen and Sapphire FL imager by Azure Biosystems. Representative image of triplicate comparisons. **B)** Image-J densitometry analysis of the sfRNA1 band normalized to the untreated control sample. A higher band intensity of sfRNA1 was observed in the Dcp2 knockdown samples compared to the nonsense siRNA treated control (* $p<0.05$). Dcp2 knockdown and nonsense siRNA control band intensities normalized to untreated control band intensity.

image). Densitometry analysis of sfRNA1 bands revealed Dcp2 knockdown samples' band were significantly more intense than the nonsense siRNA band following normalization of both groups to the untreated infected control bands (**Figure 2.11b**). This indicated that Dcp2 knockdown was not interfering with sfRNA generation; instead, more sfRNA was present in these samples compared to the other treatments groups.

2.4 Discussion

Viruses, with their quick evolution and harsh natural selection, are experts at redirecting and repurposing host enzymes for their own benefit. Therefore, understanding how host proteins are used by flaviviruses to facilitate their own replication is just as important as knowing the structures and functions of viral proteins and RNA. Flaviviruses use the host Xrn1 enzyme to unleash sfRNA that suppresses the host defense against infection, allowing replication to continue for a while longer. But Xrn1 only works on a specific form of RNA— one with a single phosphate at the 5' end.

As far as we know, the flavivirus replication machinery produces capped RNA, or at the very least triphosphorylated RNA depending on the success of the NS3 triphosphatase and NS5 capping domain. We sought to identify which host proteins participate in the flavivirus replication cycle by producing suitable monophosphorylated viral RNA ends for Xrn1 degradation and sfRNA generation.

Previously our lab investigated host decapping and dephosphorylating enzymes to see if their absence upset viral replication efficiency. The phosphatases DUSP11 and

NUDT2 would be able to convert uncapped vRNA into monophosphorylated RNA. DXO and Dcp2 are known decapping enzymes, with the latter known to work in tandem with Xrn1 to regulate host transcripts. These four enzymes were individually knocked down in Huh7 cells using siRNA transfection followed by WNV, ZIKV, or YFV infection, with the resulting cell and supernatant RNA being tested for viral replication copies. It was concluded that Dcp2 knockdown had a deleterious effect on viral replication for WNV and ZIKV infected cells and all three viruses produced less infectious particles following knockdown of RNase-free water (**Figures 2.2 and 2.3**). It was concluded that Dcp2 was a proviral factor in flavivirus infections and that the decapping abilities of the enzyme were assisting in an efficient infection cycle. With these results, I hypothesized that Dcp2 knockdown results in less monophosphorylated vRNA and interferes with sfRNA generation.

2.4.1 Optimizing the Splint Ligation Protocol

Splint ligation provided an accessible method for determining relative changes in monophosphorylated RNA content across the Dcp2 siRNA, nonsense siRNA, and untreated treatment groups. The specificity of the SplintR Ligase excludes capped RNA, or RNA with other phosphorylation states. We confirmed the specificity using the pRNA, ppRNA, pppRNA, and GpppRNA controls. The main drawback was efficiency— ideally the pRNA control would have had an APCR:VPCR ratio closer to 1. Unfortunately, this was far from the observed ratio for the pRNA control. This was likely due to the protocol and enzyme being tailored to larger quantities of RNA - we were unable to provide the concentrations originally detailed in the cell-RNA oriented original protocol. Potentially, the enzymatic “hook” effect could have decreased efficiency, where an abundance of enzyme compared to substrate results in enzymatic inhibition [57]. Optimization of reagent concentrations took months of effort, where we found that decreasing the concentrations of enzyme or changing the buffer had little to no effect on final output. Nevertheless, the ratio was significantly ($p<0.0001$) higher than those seen with other RNA forms and would work well for our purposes.

The reporting mechanism for ligated, monophosphorylated RNA went through a series of iterations before finally settling on the qRT-PCR method. Initially, urea RNA gel shift patterns were considered as the reporting system; ligation of adaptor to sample RNA changes the travel speed of the adaptor band. For example, ligating a fluorescent or radiolabeled adaptor to a large molecule would result in a slower-traveling product, allowing densitometry analysis to determine the relative proportions of ligated products across the different treatments. Ultimately, reaction efficiency and the limitations of RNA urea gel resolution reduced the viability of this method. The low efficiency of the splint ligation reaction resulted in products that were difficult to observe in experimental samples. As shown in **Figure 2.5b**, even after ligating highly concentrated pRNA control samples the gel shift of the radiolabeled adaptor is miniscule and the intensity weak. The lower concentration of monophosphorylated vRNA in infected cell samples would make any shifted bands even harder to discern on a gel. Additionally, the adaptor-pRNA ligated control molecule had the benefit of being roughly 170 bases and easily observable in the center of the gel. Viral genomes, on the other hand, are around 11 kilobases and any ligated product would be difficult to differentiate from the non-ligated radiolabeled adaptor that failed to travel out of the sample well. After a long period of trying to optimize the process, we instead developed the APCR and VPCR approach for determining the monophosphorylated RNA proportions in RNA samples.

We finally developed the qRT-PCR reporting assay due to its relative ease, sensitivity, and built-in normalization strategy. qRT-PCR assays can handle the multitude of replicates needed for the experiments quickly and at low cost, while also avoiding the use of expensive, delicate reagents such as ^{32}P radiolabeled GTP. qPCR machines can accurately quantify copy numbers of low-concentration RNA molecules. By comparing the ratio of ligated RNA copies determined by APCR to all vRNA observed in VPCR, we can effectively normalize our samples internally using just the virus for a simple comparison between treated samples.

2.4.2 Some Surprising Results

With assays established, we set forth to recreate the preliminary Dcp2 knockdown experiments. We sought to keep the knockdown and infection conditions consistent with the preliminary experiments so results would be comparable. The main changes we made to the protocol concerned the number of biological replicates and their composition. Three treatment groups were established - the experimental knockdown group using Dcp2 specific siRNA molecules, a nonsense siRNA negative control group, and an untreated group to establish a baseline for a WNV infection. The low MOI of 0.01 used was in line with the preliminary experiment guidelines and was the cause of the low concentration of viral RNA collected from samples. This made preliminary splint ligation experiments difficult to assess as the monophosphorylated vRNA is a fraction of a fraction of the total RNA collected and thus difficult to resolve. With the quantity of vRNA collected from individual wells being so small, we elected to have our eight biological replicates from each group be comprised of three wells in a six-well plate that were pooled together. This gave us enough RNA to process with splint ligation and receive observable results in qRT-PCR. Infection was followed by a media change to ensure any virus collected from supernatant was newly produced within the cells. 48 hours later the cell and supernatant RNA samples were collected and purified. Using the cell RNA samples we confirmed the knockdown of Dcp2 expression using qRT-PCR targeting Dcp2 and GAPDH mRNA, with the latter acting as a housekeeping gene. The Dcp2 and nonsense siRNA treated group's Dcp2 expression levels were normalized to the untreated infection group and visualized after a log2 transformation (**Figure 2.7**). These results confirmed that Dcp2 was significantly knocked down, inhibiting the decapping activity of Dcp2.

With Dcp2 knocked down, we were interested in the vRNA composition of these virions. The supernatant of infected cells would have possessed only newly formed virions due to the media change performed after infection with WNV. The amount of monophosphorylated vRNA being packaged would hypothetically decrease if Dcp2 was responsible for the decapping of new vRNA in the cytoplasm. Previous experiments in the Wilusz lab have shown that flavivirus virions contain a variable composition of vRNA, with roughly two-thirds being capped. We surmised that the packaging system for flaviviruses was intentionally imperfect, where uncapped vRNA was packed into new virions. This result could be beneficial for future infections, as a packaged monophosphorylated vRNA would enter a new cell and immediately be processed into sfRNA. This sfRNA priming would kickstart the immune modulation of the infected cell while capped vRNA from other virions entered the cell. First, we recreated the Wilusz protocol by using methyl-7 cap-specific immunoprecipitation to assess the change in percentage of capped vRNA in supernatant RNA from the Dcp2 knockdown, nonsense siRNA, and untreated samples. We saw that the vRNA in all three of the treatment groups' supernatant samples, the proportion of virion RNA that was capped was roughly 60%, consistent with the observed proportions from the Wilusz lab experiments (**Figure 2.8a**). While the amount of capped vRNA was slightly elevated in the Dcp2 knockdown samples, the difference was not statistically significant.

To further investigate virion content with a parallel system, we processed the supernatant RNA samples using the splint ligation and qRT-PCR pipeline we established previously. We hoped to directly assess the amount of monophosphorylated vRNA in virions by labeling them directly with the adaptor sequence. Our results showed that the proportion of monophosphorylated RNA in supernatant from the three treatment groups were not significantly different from each other (**Figure 2.9**). When combined with the cap capture assay results, we could not conclude that the suppression of Dcp2 decapping activity altered the capped vRNA proportion in newly expelled virions. This suggests Dcp2 may have less of an effect on the decapping of vRNA destined for packaging into new viral particles. It is possible that the amount of capped vRNA produced significantly outnumbers the processing rate of Dcp2 in a normal infected cell. Thus, the removal of said enzyme has a negligible effect on the amount of capped vRNA packaged and expelled in virions.

After assessing the effect of suppressing Dcp2 activity on new virion vRNA composition, we looked inwards at the vRNA inside infected cells. The idea was that Dcp2 was decapping new vRNA and reducing them to a monophosphorylated form for Xrn1 degradation and sfRNA generation. With Dcp2 expression knocked down, there would hypothetically be less monophosphorylated vRNA and thus less sfRNA being made from Xrn1 processing. Alas, that is not what we saw. In fact, we observed the exact opposite after putting the cell RNA samples through the splint ligation and qRT-PCR pipeline. Knocking down Dcp2 expression in cells and then infecting them with WNV resulted in a staggering increase in the proportion of monophosphorylated vRNA (**Figure 2.10**). Even the relatively low efficiencies of the VPCR and APCR reactions could not deter the clear difference in monophosphorylated vRNA ratios between the treatment's groups. Looking deeper, we analyzed select cell RNA samples for sfRNA content by northern blot. These samples were normalized based on GAPDH content determined by qRT-PCR to ensure the observed bands demonstrated the sfRNA content in a consistent ratio to overall cell RNA. The results were again surprising, as the Dcp2 knockdown samples contained more sfRNA than the nonsense siRNA treated samples and the untreated infected control samples (**Figure 2.11a and b**). This further demonstrated that our initial hypothesis was incorrect, as sfRNA generation was not interrupted by Dcp2 inhibition.

To say this was surprising would be a severe understatement. The preliminary Dcp2 knockdown experiments from our lab showed that WNV genome copies and titers decreased following Dcp2 suppression, and they concluded that Dcp2-mediated decapping has a positive impact on efficient viral replication. Consequently, I predicted that Dcp2 assisted viral replication by driving the creation of monophosphorylated vRNA for Xrn1 digestion and thus facilitating sfRNA generation. After my surprising results and looking deeper, I realized that the conclusion reached in the previous experiments was flawed. Those experiments showed that knocking down Dcp2 had a negative effect on the viral infection event but did not show that Dcp2 influences viral infection *directly*. Dcp2 is a key player in the mRNA decay pathway, and without Dcp2 performing its job as expected, the host cell is unable to properly control the expression of its genes by degrading their transcripts [53]. These genes include ISGs that are activated because of viral infection, including the interferon regulatory factor (IRF) IRF-7 which is directly regulated by Dcp2-mediated degradation [52]. With Dcp2 removed from the equation, host processes concerning homeostasis and immune response to infection are thrown off balance from reduced decapping and degradation. Yes, suppressing Dcp2 influenced viral infection, but not in a direct manner as the downstream effects of host gene expression dysregulation are what likely impacted the efficiency of viral replication and virion formation. The increase in monophosphorylated RNA and sfRNA in the cell RNA samples is due to a currently unknown combination of factors that resulted from Dcp2's inhibition. Thus, with these results and the ambiguity of its causes, my null hypothesis could not be rejected; Dcp2 was not solely responsible for the generation of monophosphorylated vRNA in infected cells.

Even with this surprising result, an interesting conclusion can be formed. Dcp2 is not the primary agent responsible for monophosphorylated vRNA and sfRNA. With 80% of Dcp2 expression suppressed, infected cells still produced monophosphorylated 5' ends on new vRNA and sfRNA is still being made from Xrn1 degradation. While Dcp2 has long been suggested as the creator of the Xrn1 substrate we can now see that it is not the key performer of this role when it comes to flavivirus RNA. Monophosphorylated RNA must instead be produced by other means.

2.4.3 Implications and Future Directions

Our results indicate there must be a different method than Dcp2 decapping to produce monophosphorylated 5' ends of vRNA, especially in the elevated levels seen in the Dcp2 knockdown samples. Other decapping enzymes in a host cell may be acting on vRNA. Or it is possible that the flavivirus replication machinery in NS3 and NS5 are imperfect, naturally producing monophosphorylated viral RNA to help facilitate immune modulation via sfRNA. Similar investigation into the different decapping and dephosphorylating proteins (such as DXO, NUDT2, and DUSP11) is necessary to evaluate their roles in monophosphorylated RNA generation. To assess if the viral replication machinery is producing monophosphorylated RNA naturally, future studies can focus on using replication complexes containing NS3 and NS5 *in vitro* to generate new viral genomes, as performed in Uchil et al. [58]. With proper differential ultracentrifugation, replication complexes can be isolated without cytoplasmic decapping enzymes present. These complexes can then be tested for activity and their products analyzed with splint ligation and qRT-PCR to determine if monophosphorylated RNA is coming fresh off the line.

As for sfRNA, its generation only requires a monophosphorylated 5’ end of RNA, and our efforts have focused on enzymes acting on the 5’ end of vRNA genome. Potentially, one of the endoribonuclease enzymes involved in mRNA decay may be snipping the vRNA somewhere along its 11 kilobase strand resulting in a monophosphorylated end for Xrn1 digestion [59]. APE1 is a known endoribonuclease that cleaves immediately downstream of an abasic site in RNA which could clip vRNA, so long as the vRNA is damaged with missing bases along its strand [60]. The MARF1 protein has binding potential with the Dcp1:Dcp2 decapping complex while also performing as an endoribonuclease capable of kickstarting vRNA degradation [61]. Although, endoribonuclease acting on vRNA does not explain the increase in monophosphorylated 5’ ends of intact flavivirus RNA following Dcp2 knockdown.

The source of monophosphorylated vRNA is a glaring hole in our knowledge of the flavivirus replication cycle. The monophosphorylated RNA is the first step in the generation of sfRNA, a critical component of a successful infection event. Filling this gap in our understanding of a key flavivirus process is important for painting the full picture of viral replication and infection. Without this complete understanding of flavivirus infection there are potential avenues for antiviral development being ignored when we cannot afford to take any tool for granted.

2.5 REFERENCES

1. Assenberg, R., E. Mastrangelo, T.S. Walter, et al., *Crystal structure of a novel conformational state of the flavivirus NS3 protein: implications for polyprotein processing and viral replication.* J Virol, 2009. **83**(24): p. 12895-906.

2. Bartelma, G. and R. Padmanabhan, *Expression, purification, and characterization of the RNA 5'-triphosphatase activity of dengue virus type 2 nonstructural protein 3.* Virology, 2002. **299**(1): p. 122-32.

3. Benarroch, D., B. Selisko, G.A. Locatelli, et al., *The RNA helicase, nucleotide 5'-triphosphatase, and RNA 5'-triphosphatase activities of Dengue virus protein NS3 are Mg2+-dependent and require a functional Walker B motif in the helicase catalytic core.* Virology, 2004. **328**(2): p. 208-18.

4. Chatel-Chaix, L., W. Fischl, P. Scaturro, et al., *A Combined Genetic-Proteomic Approach Identifies Residues within Dengue Virus NS4B Critical for Interaction with NS3 and Viral Replication.* J Virol, 2015. **89**(14): p. 7170-86.

5. Duan, Y., M. Zeng, B. Jiang, et al., *Flavivirus RNA-Dependent RNA Polymerase Interacts with Genome UTRs and Viral Proteins to Facilitate Flavivirus RNA Replication.* Viruses, 2019. **11**(10).

6. Issur, M., B.J. Geiss, I. Bougie, et al., *The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure.* RNA, 2009. **15**(12): p. 2340-50.

7. Kumar, D., A. Kumar, T. Bhardwaj, et al., *Zika virus NS4A N-Terminal region (1-48) acts as a cofactor for inducing NTPase activity of NS3 helicase but not NS3 protease.* Arch Biochem Biophys, 2020. **695**: p. 108631.

8. Kundharapu, S. and T.K. Chowdary, *Dengue Virus NS4b N-Terminus Disordered Region Interacts with NS3 Helicase C-Terminal Subdomain to Enhance Helicase Activity.* Viruses, 2022. **14**(8).

9. Lu, H., Y. Zhan, X. Li, et al., *Novel insights into the function of an N-terminal region of DENV2 NS4B for the optimal helicase activity of NS3.* Virus Res, 2021. **295**: p. 198318.

10. Teramoto, T., A. Balasubramanian, K.H. Choi, et al., *Serotype-specific interactions among functional domains of dengue virus 2 nonstructural proteins (NS) 5 and NS3 are crucial for viral RNA replication.* J Biol Chem, 2017. **292**(23): p. 9465-9479.

11. Umareddy, I., A. Chao, A. Sampath, et al., *Dengue virus NS4B interacts with NS3 and dissociates it from single-stranded RNA.* J Gen Virol, 2006. **87**(Pt 9): p. 2605-2614.

12. Xu, S., Y. Ci, L. Wang, et al., *Zika virus NS3 is a canonical RNA helicase stimulated by NS5 RNA polymerase.* Nucleic Acids Res, 2019. **47**(16): p. 8693-8707.

13. Yon, C., T. Teramoto, N. Mueller, et al., *Modulation of the nucleoside triphosphatase/RNA helicase and 5'-RNA triphosphatase activities of Dengue virus type 2 nonstructural protein 3 (NS3) by interaction with NS5, the RNA-dependent RNA polymerase.* J Biol Chem, 2005. **280**(29): p. 27412-9.

14. Zou, J., T. Lee le, Q.Y. Wang, et al., *Mapping the Interactions between the NS4B and NS3 proteins of dengue virus.* J Virol, 2015. **89**(7): p. 3471-83.

15. Ashour, J., M. Laurent-Rolle, P.Y. Shi, et al., *NS5 of dengue virus mediates STAT2 binding and degradation.* J Virol, 2009. **83**(11): p. 5408-18.

16. Chang, R.Y., T.W. Hsu, Y.L. Chen, et al., *Japanese encephalitis virus non-coding RNA inhibits activation of interferon by blocking nuclear translocation of interferon regulatory factor 3.* Vet Microbiol, 2013. **166**(1-2): p. 11-21.

17. Errett, J.S., M.S. Suthar, A. McMillan, et al., *The essential, nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus infection.* J Virol, 2013. **87**(21): p. 11416-25.

18. Fanunza, E., F. Carletti, M. Quartu, et al., *Zika virus NS2A inhibits interferon signaling by degradation of STAT1 and STAT2.* Virulence, 2021. **12**(1): p. 1580-1596.

19. Fanunza, E., N. Grandi, M. Quartu, et al., *INMI1 Zika Virus NS4B Antagonizes the Interferon Signaling by Suppressing STAT1 Phosphorylation.* Viruses, 2021. **13**(12).

20. Ho, L.J., L.F. Hung, C.Y. Weng, et al., *Dengue virus type 2 antagonizes IFN-alpha but not IFN-gamma antiviral effect via down-regulating Tyk2-STAT signaling in the human dendritic cell.* J Immunol, 2005. **174**(12): p. 8163-72.

21. Lee, J.Y., T.T.N. Nguyen, and J. Myoung, *Zika Virus-Encoded NS2A and NS4A Strongly Downregulate NF-kappaB Promoter Activity.* J Microbiol Biotechnol, 2020. **30**(11): p. 1651-1658.

22. Mazzon, M., M. Jones, A. Davidson, et al., *Dengue virus NS5 inhibits interferon-alpha signaling by blocking signal transducer and activator of transcription 2 phosphorylation.* J Infect Dis, 2009. **200**(8): p. 1261-70.

23. Morrison, J., M. Laurent-Rolle, A.M. Maestre, et al., *Dengue virus co-opts UBR4 to degrade STAT2 and antagonize type I interferon signaling.* PLoS Pathog, 2013. **9**(3): p. e1003265.

24. Ngueyen, T.T.N., S.J. Kim, J.Y. Lee, et al., *Zika Virus Proteins NS2A and NS4A Are Major Antagonists that Reduce IFN-beta Promoter Activity Induced by the MDA5/RIG-I Signaling Pathway.* J Microbiol Biotechnol, 2019. **29**(10): p. 1665-1674.

25. Yang, Q., J. You, Y. Zhou, et al., *Tick-borne encephalitis virus NS4A ubiquitination antagonizes type I interferon-stimulated STAT1/2 signalling pathway.* Emerg Microbes Infect, 2020. **9**(1): p. 714-726.

26. Michalski, D., J.G. Ontiveros, J. Russo, et al., *Zika virus noncoding sfRNAs sequester multiple host-derived RNA-binding proteins and modulate mRNA decay and splicing during infection.* J Biol Chem, 2019. **294**(44): p. 16282-16296.

27. Moon, S.L., J.R. Anderson, Y. Kumagai, et al., *A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability.* RNA, 2012. **18**(11): p. 2029-40.

28. Moon, S.L., J.G. Blackinton, J.R. Anderson, et al., *XRN1 stalling in the 5' UTR of Hepatitis C virus and Bovine Viral Diarrhea virus is associated with dysregulated host mRNA stability.* PLoS Pathog, 2015. **11**(3): p. e1004708.

29. Chapman, E.G., S.L. Moon, J. Wilusz, et al., *RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA.* Elife, 2014. **3**: p. e01892.

30. Chapman, E.G., D.A. Costantino, J.L. Rabe, et al., *The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production.* Science, 2014. **344**(6181): p. 307-310.

31. Funk, A., K. Truong, T. Nagasaki, et al., *RNA structures required for production of subgenomic flavivirus RNA.* J Virol, 2010. **84**(21): p. 11407-17.

32. Roby, J.A., G.P. Pijlman, J. Wilusz, et al., *Noncoding subgenomic flavivirus RNA: multiple functions in West Nile virus pathogenesis and modulation of host responses.* Viruses, 2014. **6**(2): p. 404-27.

33. Yu, L., M. Nomaguchi, R. Padmanabhan, et al., *Specific requirements for elements of the 5' and 3' terminal regions in flavivirus RNA synthesis and viral replication.* Virology, 2008. **374**(1): p. 170-85.

34. Villordo, S.M., C.V. Filomatori, I. Sanchez-Vargas, et al., *Dengue virus RNA structure specialization facilitates host adaptation.* PLoS Pathog, 2015. **11**(1): p. e1004604.

35. Silva, P.A., C.F. Pereira, T.J. Dalebout, et al., *An RNA pseudoknot is required for production of yellow fever virus subgenomic RNA by the host nuclease XRN1.* J Virol, 2010. **84**(21): p. 11395-406.

36. Slonchak, A., R. Parry, B. Pullinger, et al., *Structural analysis of 3'UTRs in insect flaviviruses reveals novel determinants of sfRNA biogenesis and provides new insights into flavivirus evolution.* Nat Commun, 2022. **13**(1): p. 1279.

37. Manokaran, G., E. Finol, C. Wang, et al., *Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness.* Science, 2015. **350**(6257): p. 217-21.

38. Schuessler, A., A. Funk, H.M. Lazear, et al., *West Nile virus noncoding subgenomic RNA contributes to viral evasion of the type I interferon-mediated antiviral response.* J Virol, 2012. **86**(10): p. 5708-18.

39. Schnettler, E., M.G. Sterken, J.Y. Leung, et al., *Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells.* J Virol, 2012. **86**(24): p. 13486-500.

40. Pijlman, G.P., A. Funk, N. Kondratieva, et al., *A highly structured, nuclease-resistant, noncoding RNA produced by flaviviruses is required for pathogenicity.* Cell Host Microbe, 2008. **4**(6): p. 579-91.

41. Delorme-Axford, E., E. Abernathy, N.J. Lennemann, et al., *The exoribonuclease Xrn1 is a post-transcriptional negative regulator of autophagy.* Autophagy, 2018. **14**(5): p. 898-912.

42. Dong, H., S. Ren, B. Zhang, et al., *West Nile virus methyltransferase catalyzes two methylations of the viral RNA cap through a substrate-repositioning mechanism.* J Virol, 2008. **82**(9): p. 4295-307.

43. Geiss, B.J., A.A. Thompson, A.J. Andrews, et al., *Analysis of flavivirus NS5 methyltransferase cap binding.* J Mol Biol, 2009. **385**(5): p. 1643-54.

44. Ray, D., A. Shah, M. Tilgner, et al., *West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5.* J Virol, 2006. **80**(17): p. 8362-70.

45. Zhao, Y., T.S. Soh, S.P. Lim, et al., *Molecular basis for specific viral RNA recognition and 2'-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5).* Proc Natl Acad Sci U S A, 2015. **112**(48): p. 14834-9.

46. Laudenbach, B.T., K. Krey, Q. Emslander, et al., *NUDT2 initiates viral RNA degradation by removal of 5'-phosphates.* Nat Commun, 2021. **12**(1): p. 6918.

47. Burke, J.M., R.P. Kincaid, R.M. Nottingham, et al., *DUSP11 activity on triphosphorylated transcripts promotes Argonaute association with noncanonical viral microRNAs and regulates steady-state levels of cellular noncoding RNAs.* Genes and Development, 2016. **30**(18): p. 2076-2092.

48. Choi, J.H., J.M. Burke, K.H. Szymanik, et al., *DUSP11-mediated control of 5'-triphosphate RNA regulates RIG-I sensitivity.* Genes Dev, 2020. **34**(23-24): p. 1697-1712.

49. Kincaid, R.P., V.L. Lam, R.P. Chirayil, et al., *RNA triphosphatase DUSP11 enables exonuclease XRN-mediated restriction of hepatitis C virus.* Proc Natl Acad Sci U S A, 2018. **115**(32): p. 8197-8202.

50. Jiao, X., J.H. Chang, T. Kilic, et al., *A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing.* Mol Cell, 2013. **50**(1): p. 104-15.

51. Picard-Jean, F., C. Brand, M. Tremblay-Letourneau, et al., *2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO.* PLoS One, 2018. **13**(3): p. e0193804.

52. Li, Y., J. Dai, M. Song, et al., *Dcp2 decapping protein modulates mRNA stability of the critical interferon regulatory factor (IRF) IRF-7.* Mol Cell Biol, 2012. **32**(6): p. 1164-72.

53. Li, Y., M.G. Song, and M. Kiledjian, *Transcript-specific decapping and regulated stability by the human Dcp2 decapping protein.* Mol Cell Biol, 2008. **28**(3): p. 939-48.

54. Sun, Y. and X. Zhang, *Role of DCP1-DCP2 complex regulated by viral and host microRNAs in DNA virus infection.* Fish Shellfish Immunol, 2019. **92**: p. 21-30.

55. Blewett, N., J. Coller, and A. Goldstrohm, *A quantitative assay for measuring mRNA decapping by splinted ligation reverse transcription polymerase chain reaction: qSL-RT-PCR.* RNA, 2011. **17**(3): p. 535-43.

56. Schindelin, J., I. Arganda-Carreras, E. Frise, et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.

57. Barbarakis, M.S., W.G. Qaisi, S. Daunert, et al., *Observation of hook effects in the inhibition and dose-response curves of biotin assays based on the interaction of biotinylated glucose oxidase with (strept)avidin.* Analytical Biochemistry, 1993. **65**(4): p. 457-60.

58. Uchil, P.D. and V. Satchidanandam, *Characterization of RNA synthesis, replication mechanism, and in vitro RNA-dependent RNA polymerase activity of japanese encephalitis virus.* Virology, 2003. **307**(2): p. 358-371.

59. Orban, T.I. and E. Izaurralde, *Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome.* RNA, 2005. **11**(4): p. 459-69.

60. Kim, W.C., B.R. Berquist, M. Chohan, et al., *Characterization of the endoribonuclease active site of human apurinic/apyrimidinic endonuclease 1.* J Mol Biol, 2011. **411**(5): p. 960-71.

61. Nishimura, T., H. Fakim, T. Brandmann, et al., *Human MARF1 is an endoribonuclease that interacts with the DCP1:2 decapping complex and degrades target mRNAs.* Nucleic Acids Res, 2018. **46**(22): p. 12008-12021.

CHAPTER 3 — INVESTIGATING PROTEIN-PROTEIN INTERACTIONS IN THE FLAVIVIRUS REPLICATION COMPARTMENT USING CROSSLINKING MASS SPECTROMETRY

3.1 Overview

Every stage of a flavivirus infection event is important for successful viral propagation. Flaviviruses are single-stranded positive-sense RNA viruses that need a protected environment with newly-translated enzymatic machinery to successfully replicate their genomes. This need is satisfied with the virus replication compartment, also known as replication vesicles. These compartments are an association of host membranes and viral proteins that act as a protective invagination in the endoplasmic reticulum (ER) [1-5]. In the replication compartment the uncapped negative-sense strand template and the double stranded genomic intermediate produced during replication are safe from recognition by host pattern recognition receptors such as RIG-I and MDA5 [6, 7]. Inside this protected environment are viral nonstructural proteins (NS) and the replication complex. This protein complex is primarily comprised of the NS3 helicase and triphosphatase along with NS5 that performs as an RNA-dependent RNA polymerase (RdRp), methyltransferase, and guanylyltransferase [8-18]. This complex works to create newly capped viral genomes based off the template strand for further translation or packaging into virions. Proper organization of the nonstructural proteins in the flavivirus replication compartment is required to successfully replicate viral RNA (vRNA) and propagate infection. Because of their importance in viral replication, these compartments are of great interest to researchers for developing new therapeutics to deter active viral infections.

This key environment for viral replication is held together by specific protein-protein interactions (PPIs). Studies into the flavivirus NS proteins show that formation of replication compartments and enzymatic activity of NS3/NS5 are enhanced by these PPIs [12, 16, 17, 19-23]. Despite how critical the proper association of these proteins are for compartment formation, and thus infection, there is limited knowledge on how the nonstructural proteins are organized in relation to the compartment and each other *in situ*.

Current understanding of the flavivirus replication compartment overall structure is largely the result of cryo-electron tomography studies. These studies used tomograms of infected cells to create stunning three-dimensional structures of the viral replication compartment, revealing details such as the thin channel connecting the main compartment to the cytoplasm [24-26]. These experiments have been invaluable for understanding the flavivirus replication process, namely the extent of host membrane rearrangement and the location of flavivirus genome replication. Unfortunately, cryo-electron tomography is limited in how deeply it can investigate the compartment organization, specifically how proteins interact and how they are oriented to facilitate the replication complex function. Nanogold labeling has been used to identify the location of specific NS proteins, but this protein labeling method is limited due to the size of antibodies, labeling inefficiency, and the inability to multiplex labeled sections of cell tissue [25, 27]. With the limited resolution of cryo-electron tomography, discoveries into the specific structural characteristics of replication compartment proteins have been stymied.

With these limitations on understanding the megastructure of the compartment, researchers have primarily investigated compartment PPIs to build a larger picture from the ground up. The composition of these compartments has been assessed using western blot following experimental treatment and purification. Uchil et al. isolated replication compartments and demonstrated that the compartments are nuclease and nonionic detergent resistant; only after treating with strong ionic detergents and the protease trypsin were they able to successfully degrade viral RNA with nuclease [4]. Co-immunoprecipitation experiments have demonstrated interactions between NS proteins— primarily NS2b, NS3, NS4b, and NS5. With identified co-precipitating protein pairs, selective mutagenesis studies have sought to determine which residues contribute to these interactions [12, 16, 21, 23, 28, 29]. From these efforts we know that certain identified interactions within the compartment can enhance enzymatic activity and membrane rearrangement roles. For instance, removing portions of NS4a, 2K, or NS4b deters membrane reorganization while NS3-NS4b interactions enhance the helicase activity of its helicase domain [3, 21-23, 30, 31]. While providing initial identification of bonded proteins and function-affecting regions, these methods are limited in what they can reveal about quaternary structures. Co-immunoprecipitation from infected cells does not indicate the planes of interactions between two proteins in their natural environments nor the critical residues important to these interactions. Additionally, co-immunoprecipitation does not reveal associations due to a common mediator such as RNA acting as a scaffold between two proteins. Selective mutagenesis does not distinguish between a residue important for interaction and a residue important for proper folding of the protein. Artificial mutations into the protein sequence are capable of triggering a misfolding responsible for reduced interaction. The limitations of co-immunoprecipitation and selective mutagenesis studies restrain discoveries to a level that lacks detail into endogenous interaction planes and the overall structure of multiple protein relationships.

Using *in silico* or *in vitro* methods can be a useful approach to identifying protein interactions. Using crystal structures of NS3 and NS5, Brand et al. reported a theoretical replication complex structure with NS3, NS5, and an actively unwound RNA helix [32]. This model considered known active sites and structural limitations of the represented players. While initially useful, this model lacked supporting experimental evidence and acted as a guide or prediction as opposed to as a discovered structure. Ultimately, a later publication refuted the predicted structure with wet-lab research, demonstrating the limitations of *in silico* predictions [33]. In Osawa et al., the researchers combined an RNA construct of the 5’ UTR stem loop A and a portion of the 3’UTR from the dengue virus (DENV) genome with recombinant NS3 and NS5 to produce snapshots of the three molecules interacting at the beginning of transcription [15]. Analysis of these structures with cryo-electron microscopy resulted in an interesting look into the interactions between NS3 and NS5 when binding to the 5’ end of DENV. This method gave insight into the replication complex structure but was still restricted by the nature of the structure’s generation. The environment where these proteins interacted are patently artificial and devoid of the other factors that can be found in an actual replication compartment, such as full length double or single stranded RNA, intermembrane viral proteins, and other host or viral protein interactions. An artificial environment with *in silico* models or *in vitro* generated proteins can provide useful insights, even while limited in their relevancy to an *in situ* infection event.

With the previously reviewed methods of studying viral structures and PPIs held back by an inability to identify the extent of protein interactions *in situ*, I turned towards crosslinking mass spectrometry (XLMS) for studying replication compartment PPIs. XLMS uses protein crosslinkers that have a spacer of known length connecting two functional groups that bind to specific residues on a protein. These crosslinkers determine distances between pairs of residues on the same protein or on two closely associated proteins. Mass spectrometry analysis of crosslinked peptides then reveals characteristic doublet peaks in MS/MS spectra, known as signature peaks. Computational programs use these signature peaks to identify the sequences of peptides that were bound together by the crosslinker [34-36]. These linked peptide sequences are then matched to proteins allowing interaction planes within or between proteins to be identified.

There are four main types of crosslinks observed in XLMS: monolinks, loop-links, intramolecular crosslinks, and intermolecular crosslinks. Monolinks indicate when a crosslinker only binds only one of its functional groups to a residue while loop-links occur when a crosslink binds to two residues on the same peptide. Intramolecular links bind two separate peptides originating from the same protein together, and intermolecular crosslinks are characterized by the crosslinking of two peptides on separate proteins (**Figure 3.1a and 3.1b**). The various crosslink types can be useful for studying protein structures, but intermolecular crosslinks are of particular interest for identifying structural interactions between flavivirus proteins in the replication compartment. The crosslinking process can occur *in situ*, with infected cells or their lysates being crosslinked to “freeze” the PPIs in their natural states. The nature of a crosslinker’s spacer length provides proximity ranges between two residues- if residues are properly crosslinked, we know those two regions of the protein(s) are within the spacer distance of each other. With enough PPIs identified in a compartment sample, the overall picture of the replication compartment can be reconstructed. The Bonvin Lab’s program DisVis can take identified crosslinks to rebuild interaction interfaces between two proteins while identifying critical residues for the interaction. The HADDOCK program can take these residues and crosslinks and model the docking of two proteins, allowing PPIs to be assembled into an overall protein complex structure [37-40]. The *in situ* capability of XLMS allows for the detailed study of natural protein interactions with unmodified structures, removing the mutagenesis study misfolding concerns.

So, in an ambitious move by a naïve graduate student, I set forth to characterize PPIs found in the flavivirus replication compartment *in situ* using XLMS. Ultimately, the

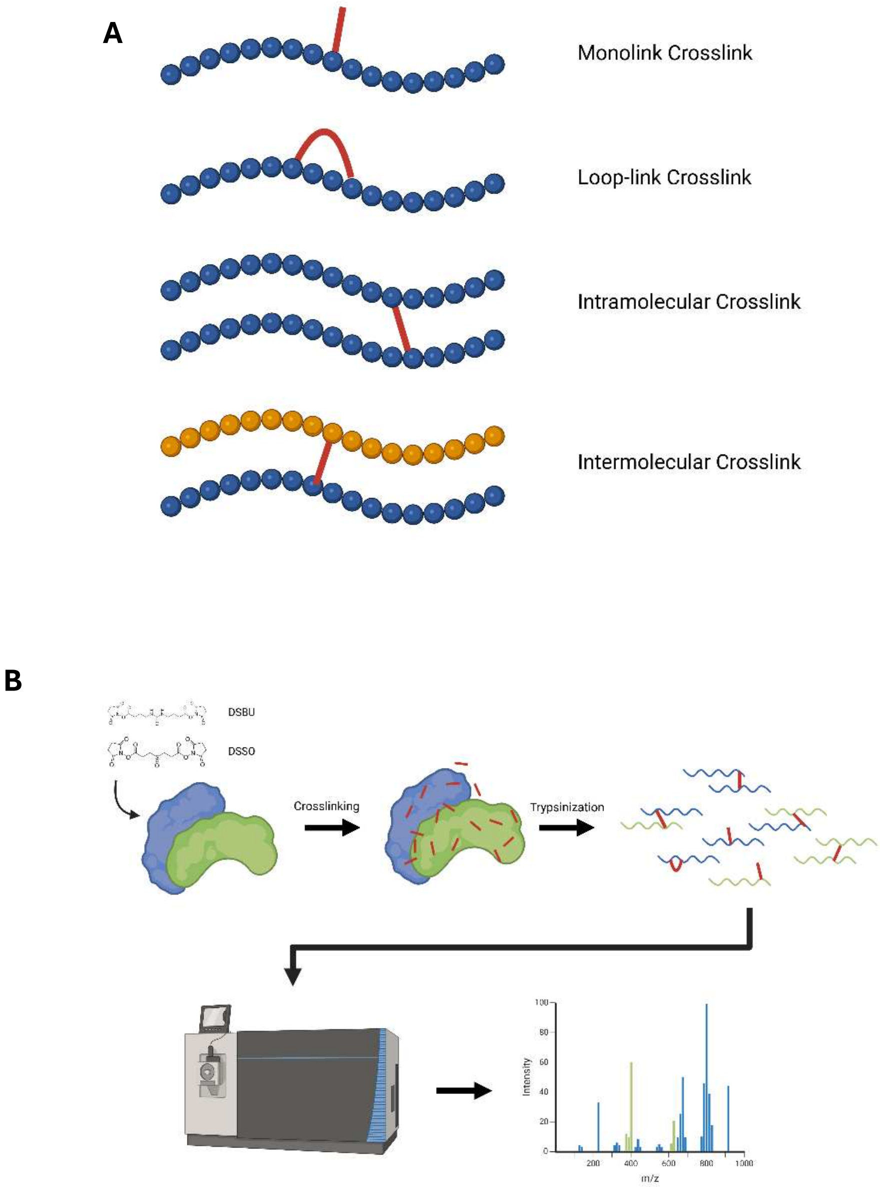


Figure 3.1. Crosslinking mass spectrometry. **A)** Different varieties of crosslinked peptides. Monolink crosslinks bind to a peptide on only one end of the crosslinker. Loop-link crosslinks are characterized by crosslinking two residues on the same peptide. Intramolecular crosslinks crosslink two peptides from the same origin protein sequence. Intermolecular crosslinks link two peptides from separate origin proteins. **B)** Diagram of the crosslinking mass spectrometry (XLMS) workflow. Protein complexes are crosslinked with a homo-bifunctional chemical crosslinker. Following digestion, crosslinked peptides are identified by mass spectrometry using their signature peaks. Created with BioRender.com

goal was to build a detailed picture of the viral protein interactions that could then be used to arrange an accurate model of a flavivirus replication compartment. The project was beset with significant obstacles and unsuccessful in identifying enough PPIs to build a model of the replication compartment. But as is the nature of science, conclusions and evidence could be gleaned from these experiments as well as knowledge on how to enhance the efficacy of XLMS at determining viral protein interactions.

3.2 Materials and Methods

This project exists in two primary phases that are defined by differences in pathogen, purification protocols, and mass spectrometry methods. This separation occurred as the project evolved over five years to accommodate the COVID-19 pandemic, funding constraints, pathogen availability, and instrument access. Ultimately, the results seen in Phase II are inextricably linked to the groundwork laid in Phase I and as such both sections will be presented.

3.2.1 Phase I: KUNV ER Microsome Isolation

Suspension 293F cells were infected at an MOI of 0.1 with Kunjin virus (KUNV) and allowed to incubate for 48 hours before collection. A control sample was uninfected but processed with the same purification steps as the infected samples. The Sigma Endoplasmic Reticulum Isolation Kit protocol was adapted for the isolation of KUNV replication compartments. An initial 500xg centrifuge at 4°C for five minutes pelleted the cells followed by a washing step with 1X PBS and a 20-minute incubation in 1X Hypotonic Extraction Buffer. The swollen cells were pelleted with a 4°C 500xg centrifuge for 5 minutes before being resuspended in 1X Isotonic Extraction Buffer and lysed by 20 strokes of a 7ml Dounce homogenizer. The resulting lysed cells were spun at 1000xg for 10 min at 4°C and the supernatant collected after discarding the thin lipid layer atop the solution. A 12,000xg centrifuge at 4°C for 20 minutes produced a supernatant containing our microsomes, called the post-mitochondrial fraction (PMF). This PMF was then pelleted by centrifuging at 100,000xg for an hour with the pellet resuspended by Dounce homogenization and diluted into a 20% OptiPrep density gradient solution. The 20% sample solution was layered in Beckman Coulter OptiSeal tubes atop 2ml of 30% OptiPrep solution followed by 4ml of 15% OptiPrep solution. The sample tubes were spun at 150,000xg in an Opima XPN-80 Ultracentrifuge (Beckman Coulter) and the resulting solution was fractionated into 400µl fractions beginning at the top of the solution using a long needle and syringe.

3.2.2 Phase I: Characterization of Replication Compartment Genomic Content

A GelGreen staining assay was used to assess the overall nucleic acid content of samples. 20µl of isolated ER fractions, PMF, and whole cell lysates for both infected and uninfected 293F cells were aliquoted into a black, opaque-bottom 384 well plate. 20µl of 10X GelGreen Nucleic Acid Stain was added, and samples were mixed through pipetting. The relative fluorescence readings were measured with a PerkinElmer Victor X5 multilabel plate reader.

RNA from ER fractions and other samples was extracted with Trizol and phenol-chloroform extraction. All samples were mixed in a 3:1 ratio of Trizol LS:sample and vortexed until clear followed by a 5-minute room temperature incubation. Samples were combined with 200μl of chloroform for every 1ml of Trizol sample. After mixing, the samples were spun at max speed for 10 minutes at 4^{o}C. The aqueous layer was carefully aspirated and transferred into a tube with 3μl of glycogen and 500μl of isopropanol for every milliliter in the original Trizol sample. Samples were spun at max speed for 10 minutes at 4^{o}C and the solution was aspirated, leaving a pellet of precipitated RNA, DNA, and glycogen. This pellet was washed with 70% ethanol before air drying for five minutes, covered. Pellets were resuspended in 160μl RNase-free water and combined with 20μl Promega DNase1 Buffer and 20μl Promega DNase 1. Samples were then incubated for 30 minutes at 37^{o}C before being combined with 200μl of phenol-chloroform. Samples were vortexed and centrifuged for 1 minute at 16,000xg. 160μl of the top layer of solution was removed and combined with 66μl 10M ammonium acetate, 4μl of glycogen, and 800μl of 100% ethanol that had been pre-chilled at 20°C to precipitate the RNA and glycogen. Samples were inverted five times before being incubated at -80°C overnight. The samples were spun at max speed at 4^{o}C for 20 minutes and the resulting pellet washed with 150μl of 70% ethanol. After removing the ethanol wash the samples were allowed to air dry for 15 minutes, covered, before being resuspended in either 6µl or 11µl of RNase-free water. If resuspended in 11µl, 1µl of the sample was used to quantify RNA by NanoDrop One.

The positive-sense and negative-sense KUNV vRNA copies were quantified using RT and ddPCR. 200ng of extracted RNA from infected and uninfected samples underwent RT using the Promega ImProm II Reverse Transcriptase System using 10µM of KUNV vRNA primers specific to positive-sense (BG1389) and negative-sense (BG1388) strands. The resulting cDNA was analyzed by ddPCR using the proper sense-specific primers for each cDNA type (BG1391+BG687 for positive-sense and BG1390+BG688 for negative-sense cDNA). Primers are listed in **Table 3.1**. Data was acquired using QX200 ddPCR EvaGreen Supermix in a BIO-RAD QX200 Droplet Reader.

3.2.3 Phase I: Characterization of KUNV Replication Compartment Protein Content

Isolated infected fractions 10-15 demonstrated the highest vRNA content and were chosen for protein content assessment. For silver stain analysis, infected and uninfected fractions 10-15 were run on 10% SDS-PAGE gels at 120V until the dye front reached the bottom of the gel. The resulting gels underwent silver staining using the Pierce Silver Stain Kit and accompanying protocol.

Table 3.1. List of primers used for positive and negative-sense vRNA copy determination.

Primer Name	Sequence
BG688	ACCCCAGCTGCTGCCACCTT
BG687	TGGACGGGGAATACCGACTTAGAGG
BG1962	GCGGGTTCGTGTGCCGATCAAA
BG1961	TGGAAGAGTCATCCTGGGGGAACC
BG1960	GGCCGTCATGGTGGCGAATAACGACGGAAGACATGCTCGCAGT
BG1959	CTGGAGTCGTAGATCCTACCGCTCATTTGTGTGCCCTCCATAGCA
BG1913	GGCCGTCATGGTGGCGAATAA
BG1912	CTGGAGTCGTAGATCCTACCGC
BG1391	GGCCGTCATGGTGGCGAATAA
BG1390	CGGGAAGGCGACTGGAGTGCC
BG1389	GGCCGTCATGGTGGCGAATAAACCCCAGCTGCTGCCACCTT
BG1388	CGGGAAGGCGACTGGAGTGCCTGGACGGGGAATACCGACTTAGAGG

For western blot analysis, the same fractions were run on 10% SDS-PAGE gels and transferred to methanol-activated PVDF membranes for 1 hour at 100V. The blots were blocked with 4% nonfat dry milk powder in 1X TBS-T (4% Blotto) for 1 hour with shaking. Blots were probed with either rabbit anti-WNV NS3 polyclonal antibody (1:2000 in 4% Blotto) or mouse anti-KUNV NS5 monoclonal antibody (1:500 in 4% Blotto) for 1 hour at room temperature, shaking. Blots were rinsed once and washed three times with 4% Blotto before being probed with either HRP-conjugated goat anti-rabbit polyclonal antibody or HRP-conjugated goat anti-mouse polyclonal antibody (both at 1:2000 in 4% Blotto) for 1 hour at room temperature with shaking. After a rinse and three washes with 1X TBS-T the blots were developed with 1-Step TMB-Blotting Substrate Solution.

3.2.4 Phase I: DSSO Crosslinking Protocol

To determine optimal crosslinking concentration, PMF was crosslinked with DSSO. 50mg of DSSO (Thermo Scientific #A33545) was resuspended in 51.5μl DMSO as per manufacturer guidelines. DSSO was added to the PMF solution to achieve a final concentration of 0mM, 0.1mM, 0.5mM, 1mM, 2mM, 3mM, 4mM, 5mM, and 10mM. The samples were incubated on ice for 30 minutes. Crosslinking was quenched with enough 100mM Tris-HCl to achieve a final concentration of 20mM and incubated for 15 minutes on ice. The crosslinked samples were then assessed by SDS-PAGE gel silver stain as performed previously.

3.2.5 Phase I: Mass Spectrometry Analysis of 293F Cell PMF Lysate

Infected cell PMF was purified by ExoQuick TC (System Biosciences #EX0TC10A-1) as per manufacturer protocol. Due to COVID-19 restrictions at Colorado State University, we submitted samples to the University of Colorado School of Medicine Proteomics Core Facility for nontargeted analysis and development of a refined human proteome file. The following methods were derived from the protocol followed by Dr. Anthony Salviola. Samples were subjected to proteolytic digestion using a filter-aided sample preparation (FASP) protocol with 10 kDa molecular weight cutoff filters (Sartorius Vivacon 500 #VN01H02) [41]. Samples were reduced with 5mM tris(2-carboxyethyl)phosphine), alkylated with 50 mM 2-chloroacetamide, and digested overnight with trypsin (enzyme:substrate ratio 1:50) at 37ºC. Peptides were recovered from the filter using successive washes with 0.2% formic acid (FA). Aliquots containing 10 µg of digested peptides were cleaned using Pierce™ C18 Spin Tips (Thermo Scientific) according to the manufacturer's protocol, dried in a vacuum centrifuge, and resuspended in 0.1% FA in mass spectrometry-grade water.

Liquid chromatography tandem mass spectrometry (LC-MS/MS) was performed using an Easy nLC 1200 instrument coupled to a Orbitrap Fusion Lumos Tribrid mass spectrometer (all from ThermoFisher Scientific). Tryptic peptides were loaded on a C18 column (100 µM inner diameter x 20 cm) packed in-house with 2.7 µm Cortecs C18 resin and separated at a flow rate of 0.4 µl/min with solution A (0.1% FA) and solution B (0.1% FA in ACN) and under the following conditions: isocratic at 6% B for 3 minutes, followed by 6%-42% B for 102 minutes, 42%-60% B for 5 minutes, 60%-95% B for 1 min and isocratic at 95% B for 9 minutes. MS/ MS was performed using a data-dependent acquisition (DDA) mode with a cycle time of 3 s between master scans and dynamic exclusion was set to 45 s. Fragmentation spectra were searched against the UniProt human proteome database (Proteome ID # UP000005640 downloaded 1 December 2021) using the MSFragger-based FragPipe computational platform [42]. Contaminants and reverse decoys were added to the database automatically. The precursor-ion mass tolerance and fragment-ion mass tolerance were set to 10 ppm and .2 Da, respectively. Fixed modifications were set as carbamidomethyl (C), and variable modifications were set as oxidation (M) and two missed tryptic cleavages were allowed, and the protein-level false discovery rate (FDR) was ≤ 1%.

3.2.6 Phase I: Mass Spectrometry Analysis at CSU ARC-BIO

The BSA DSSO crosslinked control sample was processed by the CSU ARC-BIO Facility, and the following methods were adapted from the protocol followed by Kitty Brown. A chloroform-methanol extraction was implemented [43]. Briefly, four volumes of methanol are added to one volume of samples, vortexed and centrifuged for 10s at 9,000xg. Then two volumes (relative to initial sample volume) chloroform was added and vortexing/centrifugation repeated. To separate phases, 3 volumes of water were added, and the samples vortexed at high speed for 5 minutes followed by centrifugation at 10,000xg for 15 minutes at room temperature. The top layer was removed leaving the white protein interphase and lower chloroform phases. Methanol (MeOH) was then added at 1:1 v/v to the lower phase followed by high-speed vortex and centrifugation for 5 minutes each. The supernatant was removed and pellets were washed two more times. After the last MeOH supernatant was removed, samples were allowed to air dry under foil. Following reconstitution in 100µl Lysis Buffer (EasyPep Mini MS Sample Prep Kit; ThermoFisher Scientific), an aliquot was diluted 1:2 in Lysis Buffer and measured for total protein content using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific) following manufacturer instructions. Absorbance at 550nm was measured on a BioRad 680 microplate reader and total protein concentrations calculated based on a bovine serum albumin standard curve fit to a quadratic (Microplate Manager 5 software, BioRad). Samples were processed using the EasyPep Mini MS Sample Prep Kit (ThermoFisher Scientific) following the manufacturer's instructions. Briefly, ~25 µg total protein was aliquoted from each sample and raised to 100µl with lysis buffer. Reduction and alkylation solutions were sequentially added followed by incubation at 95˚C for 10 minutes. After cooling to room temperature, 5µg of a Trypsin/LysC mixture was added and samples were digested with shaking at 37ºC for 2.5-hours. The protease enzymes were deactivated with the Digestion Stop Solution and contaminants removed using mixed mode peptide clean up columns. Peptide eluate was dried in a vacuum evaporator and resuspended in 5% acetonitrile/0.1% formic acid. Once resolubilized, absorbance at 205nm was measured on a NanoDrop (Thermo Scientific) and total peptide concentration was calculated using an extinction coefficient of 31 [44].

Reverse phase chromatography was performed using water with 0.1% formic acid (A) and 80% acetonitrile with 0.1% formic acid (B). A total of 1µg of peptides was purified and concentrated using an on-line enrichment column (Thermo Scientific PepMap Neo C18 5μm, 300 um ID x 0.5cm). Subsequent chromatographic separation was performed on a Vanquish Neo (Thermo Scientific) on a reverse phase nanospray column with integrated silica emitter (IonOpticks Aurora Ultimate Gen 3 C18 1.7μm, 75 um ID x 25cm column, 45°C) using a 90 minute method at a flow rate of 300 nanoliters/min: 1-6%B over 3 minutes followed by 6-35%B over 70 minutes, 35-45%B over 5 minutes ending in 12 minutes of washing at 99%B. Peptides were eluted directly into the mass spectrometer (Orbitrap Eclipse, Thermo Scientific) equipped with a Nanospray Flex ion source (Thermo Scientific) and spectra were collected over a m/z range of 375–1600 under positive mode ionization. Ions with charge state +3 or higher were accepted for MS/MS using a dynamic exclusion limit of 1 MS/MS spectra of a given m/z value with an exclusion duration of 60 s. The instrument was operated in FT mode for MS detection (profile; resolution of 60,000) and FT mode for MS/MS detection (profile; 30,000 resolution) with a normalized CID collision energy set to 25%. MS/MS products yielding the targeted mass difference for DSSO were subjected to a third MS event in the ion trap with CID collision energy at 35% and data centroided.

3.2.7 Phase II: WNV Replication Compartment Sample Isolation and Characterization

The WNV replication compartments were isolated using a modified protocol as seen in Phase I for KUNV compartment isolation with the pBG744 WNV-GFP construct being used instead of KUNV and the protocol followed up until the PMF is isolated. This PMF was then used for characterization, crosslinking, and mass spectrometry analysis.

To assess the positive and negative-sense WNV vRNA content of the PMF, RNA was extracted from infected PMF samples using the previously described DNase I digestion and phenol-chloroform extraction protocols. The same RT protocol was performed using positive and negative-sense WNV vRNA specific primers (BG1959 for positive-sense and BG1960 for negative-sense vRNA). 1µl of cDNA was used with the PowerUp 2X SYBR-Green qPCR Mix with 5µM of sense-specific primers (BG1961+BG1912 for positive-sense and BG1962+BG1913 for negative-sense cDNA). Resulting Ct values were obtained using an Applied Biosystems QuantStudio3. The Ct values were converted to copies of viruses using a standard curve created from positive and negative-sense PCR products generated from the pBG744 plasmid. Primers are listed in **Table 3.1**.

The protein content of the PMF and crosslinked samples was assessed using silver stain, western blot, and dot blot assays. The silver stain assay was performed using the PMF and crosslinked RIPA supe and cell samples with the same protocol as in Phase I. The western blot was performed using the isolated PMF and the Phase I NS3 blotting protocol. The dot blot used a vacuum plate apparatus to attach 25µl of the PMF, RIPA supernatant, and RIPA pellet samples to a PVDF membrane followed by western blotting according to the Phase I NS3 blotting protocol.

3.2.8 Phase II: Crosslinking and MS Sample Preparation

Crosslinking in Phase II utilized the crosslinker DSBU at 2mM and the isolated WNV PMF. 50mg of DSBU (Thermo Scientific A35459) was resuspended in 47µl of DMSO. The same Phase I crosslinking protocol was performed with the PMF and the quenched sample was spun down at 100,000xg for 1 hour in a Beckman Coulter Optima XPN-80 Ultracentrifuge. The remaining pellet was resuspended by Dounce homogenizer in RIPA buffer with 1X HALT Protease Inhibitor. The solution was bath sonicated for 5 minutes followed by a 15-minute incubation on ice. The sample was then centrifuged for 10 minutes at 14,000xg and the supernatant collected as the XL RIPA supernatant sample. The pellet was resuspended in RIPA and labeled the XL RIPA pellet sample.

Both RIPA samples were prepared for mass spectrometry using the Dobos Lab In-Solution Digestion protocol. Protein concentrations were determined using the Pierce BCA Assay Kit. 30µg of sample was precipitated in acetone at -80°C. The resulting pellet was washed in acetone once before drying and resuspending with 15µl 8M urea. Following a 5-minute bath sonication 20µl 0.2% ProteaseMAX Surfactant Solution was added and the solution was shaken for 5 minutes. 58.5µl of 50mM ammonium bicarbonate solution was added and 1µl of 0.5M DTT was added for a 20-minute 50°C incubation. Afterwards 2.7µl of 0.55M IAA was added and the solutions incubated at room temperature in the dark for 15 minutes. 1µl of 1% ProteaseMAX and 18µl 0.1µg/µl trypsin was added and the samples incubated at 37°C for 16 hours. The digestion was stopped with 27µl of 2.5% trifluoroacetic acid (TFA), mixed, and then dried in a speed vacuum centrifuge. The dried samples were resuspended in EQ buffer (5% acetonitrile and 0.5% TFA) and bath sonicated for 10 minutes. Pierce C18 spin columns were then used to clean-up the solution and elute them into 70% acetonitrile and 0.1% formic acid. The samples were again dried in a dehydrating centrifuge before being resuspended in acetonitrile and 0.1% formic acid. Protein concentrations were measured with nanodrop (extinction coefficient 31) and normalized to 200ng/µl. Samples were stored at -20°C.

3.2.9 Phase II: timsTOF fleX MALDI-2 Parameters

The mass spectrometry parameters were adapted and modified from Yilmaz et al. [45]. Samples were injected in triplicate 1µl injections at 0.4µl/minute. Column separation was achieved following passage through a 5mm Thermo PepMap Neo Trap Cartridge (cat. 160454) onto an IonOpticks Aurora Ultimate 250mm x 75µm C18 column (25075C18) at 50°C before entering the Nanospray Flex Ion Source. Percent solvent B (acetonitrile with 0.1% formic acid) increased from 2% to 17% over 60 minutes, 17% to 25% over 30 minutes, 25% to 37% over 10 minutes, 37% to 80% over 10 minutes followed by a hold at 80% for 10 minutes. Spectra were collected over an m/z range of 100-1700 under positive mode ionization. Ions with charges +1 to +8 were chosen for MS/MS. PASEF charge states were set to 4 and the selected ions were fragmented using collision energies between 0.73-1.60 V^*s/cm^2 to 23-95 eV. 4 PASEF scans were triggered per 0.53s cycle.

3.2.10 Phase II: MaxQuant Data Analysis

Output files from mass spectrometry data acquisition were submitted to MaxQuant for analysis with parameters being adapted and modified from Yilmaz et al. [45]. No contaminants were included in the search and a peptide-spectrum match (PSM) false discovery rate (FDR) of 0.01 was employed. The minimum peptide length was set to 4 with a maximum mass of 10,000kDa to account for crosslinked entities. Acetylation of N-termini and oxidation of methionine were considered as variable modifications with the only fixed modification being carbamidomethylation of cysteines due to IAA treatment. A maximum of 3 modifications were accepted per peptide and the maximum number of missed cleavages was 5. No second peptides were considered, and peaks were refined in MaxQuant. The MS-cleavable crosslinker was a modified DSBU-2 crosslinker with added reactivity to serine, threonine, tyrosine, and N-terminal amines on both ends of the crosslinker. A maximum of two crosslinkers per peptide were allowed. The instrument selected was the “Bruker TIMS” with a maximum charge of 8 allowed. FTMS MS/MS analyzer higher charges, water loss, ammonia loss, and dependent losses were disabled. The output text files were imported into Microsoft Excel for sorting and analysis.

3.2.11 PyXlinkViewer and PyMOL Analysis

The PyXlinkViewer plugin was used to visualize identified crosslinks using protein structures in PyMOL v2.5.5 [46]. For DSSO, distance constraints were set to 25 angstroms. For DSBU distance constraints were set to 26.5 angstroms (Å), with any satisfied intramolecular and loop crosslinks with measured distances higher than 20 Å be re-examined. Re-examination involved using the PyMOL measurement tool to measure the distance between the ends of serine, threonine, tyrosine, lysine, and primary amines to ensure distances are less than 26.5 Å.

3.3 Results

3.3.1 Phase I: Kunjin Virus Groundwork and Analysis Pipeline Validation

3.3.1.1 Generation and validation of KUNV replication compartment samples

The generation of sufficiently enriched samples for XLMS is important to ensure sufficient crosslinks for structural analysis. We prioritized ER-derived microsomes for purification, as these microsomes contain intact replication compartments so long as the protein structures supporting the compartment maintained their bonds. In Uchil et al., these bonds were strong enough to resist detergent dissociation and were thus anticipated to withstand the mechanical cell lysis and centrifugation used for sample purification [4]. To increase the concentration of cells, we used the HEK 293F suspension culture cells obtained from the Freestyle 293 Expression System. This human cell line eliminated the restrictions of limited horizontal space in an adherent cell flask, allowing cultures to reach cell counts in the high 10^8 range. These cells were infected at an MOI of 0.1 with KUNV and underwent a 48-hour infection incubation. KUNV replication compartments were purified with a protocol adapted from the Sigma Endoplasmic Reticulum Isolation Kit. Following the infection incubation, the cell cultures underwent a series of differential centrifugation steps to produce a post-mitochondrial fraction (PMF). The PMF underwent a 1 hour 100,000xg ultracentrifugation step followed by a density gradient ultracentrifugation step at 150,000xg for 3 hours. The fractions obtained from the density gradient centrifugation contained our replication compartments and underwent characterization assays.

The ER microsomes containing the KUNV replication compartments would need to contain the two forms of vRNA: the positive-sense genome and the negative-sense template strand used to produce positive-sense copies inside the replication compartment. Overall single-stranded genetic content of the fractions was determined via incubation with GelGreen Nucleic Acid Stain followed by fluorescence reading with a PerkinElmer Victor X5 multilabel plate reader (**Figure 3.2**). Compared to the uninfected fractions, infected fractions 11, 12, and 13 had the highest fluorescent signal, indicating the presence of more single-stranded DNA or RNA. Focusing on this range of fractions, we performed reverse transcription (RT) on extracted vRNA using positive and negative-sense strand specific primers to generate cDNA for digital-droplet PCR (ddPCR) analysis. ddPCR analysis showed that fraction 12 had the lowest ratio of positive to negative-sense strand RNA content, with the uninfected fraction 12 showing a complete lack of viral genome copies (**Figure 3.3**). With these tests, we confirmed elevated genetic material content in the isolated fractions as well as the presence of both negative and positive-sense vRNA, a combination only present in replication compartments.

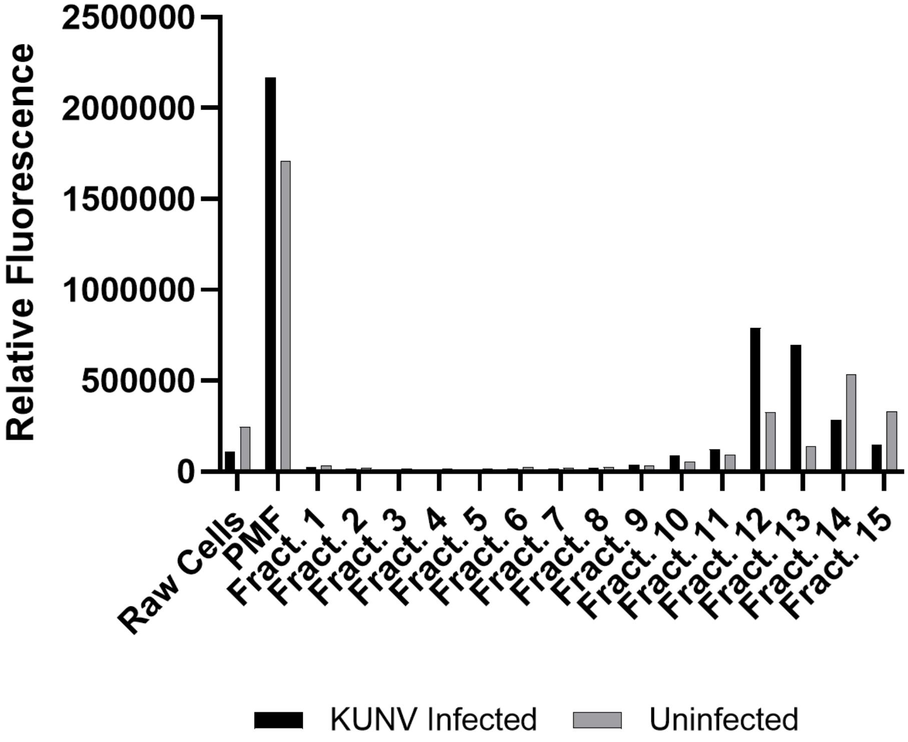


Figure 3.2. Phase I nucleic acid content of KUNV infected density-gradient ultracentrifugation fractions of the endoplasmic reticulum. Addition of GelGreen Nucleic Acid Stain to samples labels DNA or RNA with a green fluorescent moiety that is measured by a PerkinElmer Victor X5 multilabel plate reader. KUNV infected fractions 11, 12, and 13 contained higher concentrations of DNA and RNA than the corresponding uninfected fractions.

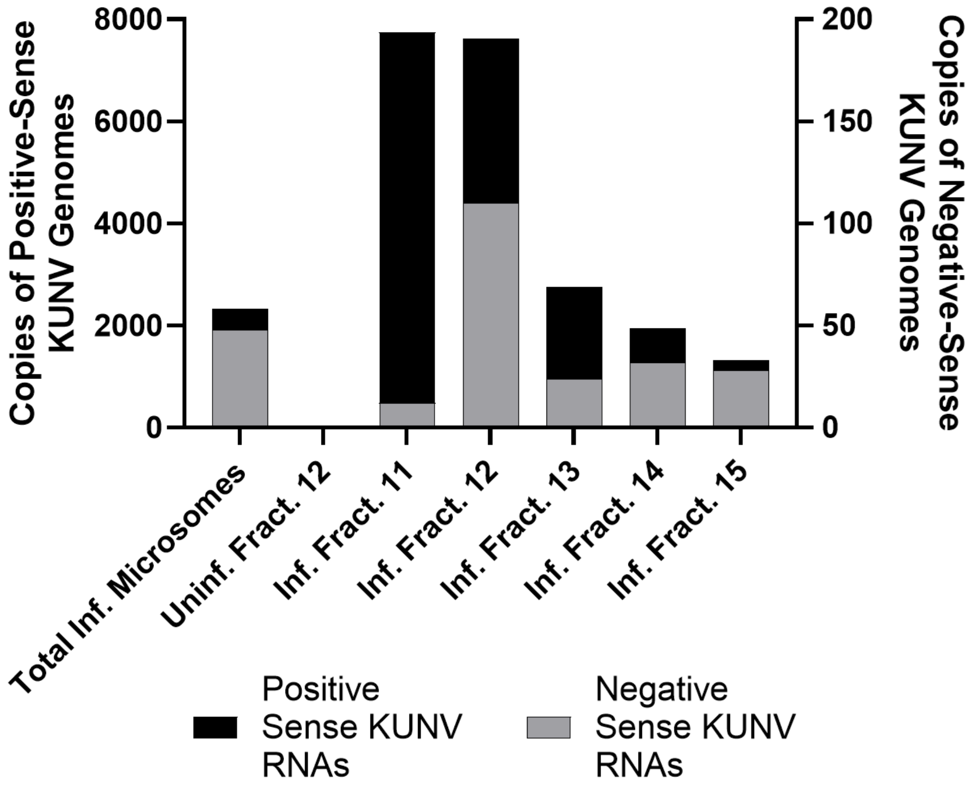


Figure 3.3. Phase I positive and negative-sense RNA copies of KUNV per ng RNA. ddPCR characterization of both vRNA sense strands. The elevated negative-sense content for fraction 12 indicates the presence of KUNV replication compartments as negative-sense flavivirus RNA resides inside these compartments during flavivirus infection. The corresponding uninfected fraction 12 lacked either sense of vRNA.

With the confirmation of both KUNV vRNA forms, we assessed the presence of NS3 and NS5 in the fractions of interest. These viral proteins contain the enzymatic regions necessary for viral replication and are part of the replication compartment-occupying replication complex. To characterize the viral protein content of the microsomes, we performed an SDS-PAGE silver stain and western blot analysis of the isolated fractions. Fractions 10 to 15 of uninfected and infected cell samples were run on a 10% SDS-PAGE gel and stained using the Pierce Silver Stain Kit (**Figure 3.4a**). Infected fractions 11 through 15 showed markedly increased band intensity compared to their uninfected counterparts, indicating an overall higher protein concentration for those ER microsome fractions. Western blot analysis of the same fractions revealed NS3 in infected fractions 10 through 12 and NS5 in fractions 11 through 13 (**Figure 3.4b and 3.4c**). These blots demonstrated the presence of a higher replication compartment concentration in fractions 11 and 12 because both NS3 and NS5 are needed in the compartment to form the replication complex.

We next established an optimal concentration of crosslinker to use on isolated replication compartments. For the KUNV samples we optimized concentrations using disuccinimidyl sulfoxide (DSSO), a homo-bifunctional crosslinker with a 10.3 angstrom crosslinker spacer and primary amine-targeting N-hydroxysuccinimide (NHS) esters at the ends [47]. Like most MS/MS cleavable crosslinkers, collision-induced dissociation (CID) cleaves DSSO on either side of the central sulfur atom, resulting in the distinct mass shift pattern seen in signature peaks allowing for crosslinked peptide identification. To optimize this crosslinker concentration, we treated the PMF from KUNV infected cells with varying concentrations of DSSO ranging from 0 to 10 mM, quenched

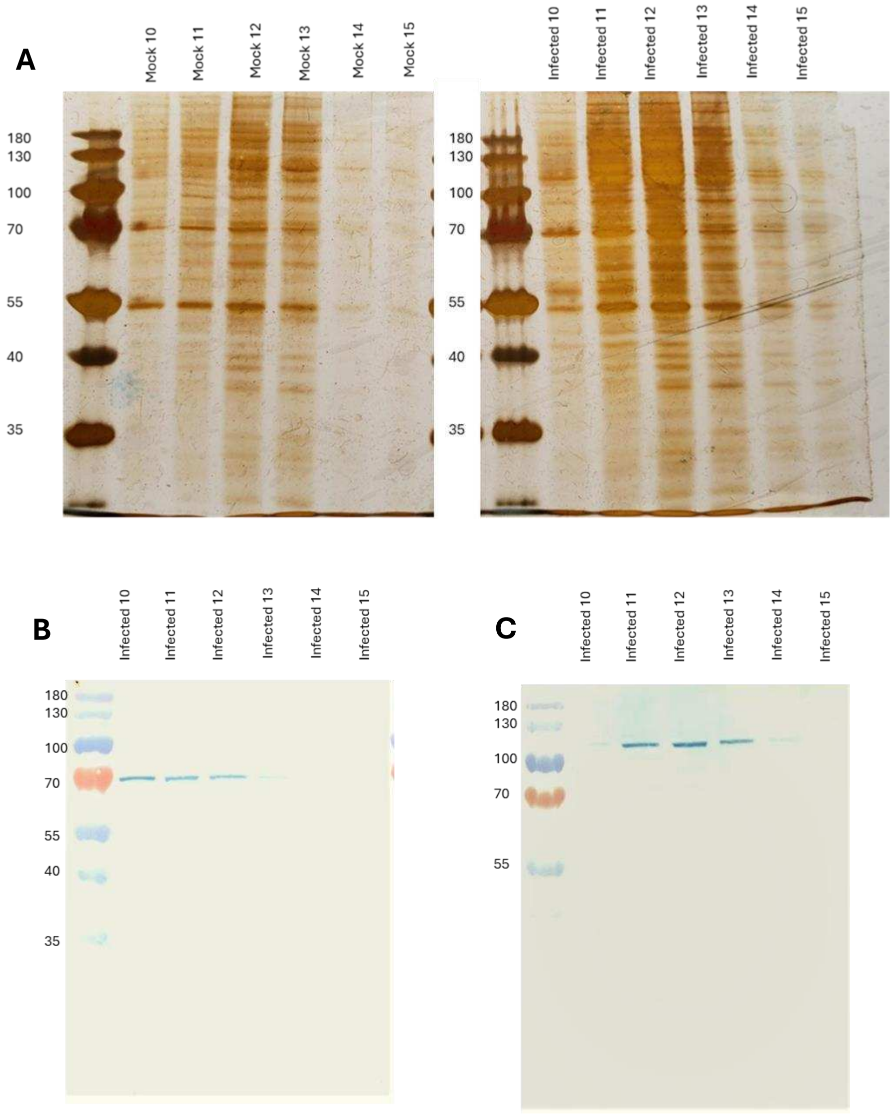


Figure 3.4. Protein characterization of Phase I KUNV replication compartment fractions. A) Silver stain analysis of mock infected and KUNV infected ultracentrifugation fractions. Protein content was elevated in the infected fractions. **B)** Western blot labeling KUNV NS3 content in KUNV infected fractions. Strong bands in infected fractions 10-12 indicate the presence of endoplasmic reticulum microsomes and flavivirus replication compartments. **C)** Western blot indicating KUNV NS5 presence in KUNV infected fractions 11-13. Combining the results from **(B)** and **(C)** reveals colocalization of NS3 and NS5 in fractions 11 and 12, indicating the presence of the flavivirus replication complex in the ER microsome fractions.

with Tris-HCl, and analyzed them by SDS-PAGE gel silver stain. The loss of distinct bands at 50kDa and 75kDa in addition to the strengthening intensity of the bands at the top of the gel lanes indicated an increase in crosslinking efficiency starting at 0.5mM, with sufficient crosslinking occurring at 2mM (**Figure 3.5**). Due to the cost of reagents, we decided that a crosslinker concentration of 2mM would be sufficient moving forward.

3.3.1.2 BSA Control XLMS and DisVis/HADDOCK analysis validation

We next validated a pipeline for using identified crosslinks to generate useful data on PPIs. With the established protocol for crosslinking, 35µg of bovine serum albumin (BSA) was crosslinked, and the samples were submitted to the Colorado State University Analytics Resources Core Bioanalysis and Omics (CSU ARC-BIO) facility for mass spectrometry analysis. There they reduced, alkylated and digested the protein with Trypsin/LysC and purified the peptides with methanol precipitation. A Thermo Scientific Orbitrap Eclipse mass spectrometer was utilized to obtain mass spectra for the sample, with the crosslinks being cleaved by CID after MS detection and MS/MS products demonstrating the targeted mass difference being advanced to a MS^3 event.

The resulting raw files were analyzed by the MaxLynx function of MaxQuant for crosslinks with the limitation of only matching lysine-lysine linkages that DSSO produces. The program identified 306 intramolecular crosslinks and 95 loop-links, with 184 of the intramolecular crosslinks tying together a unique pair of residues. Because

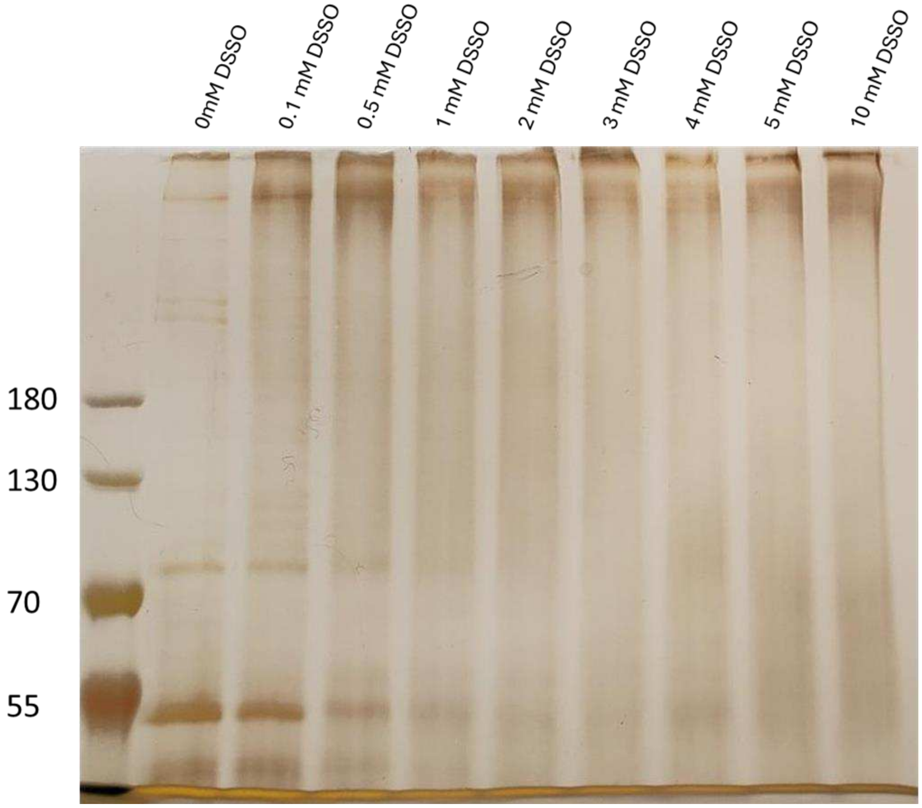
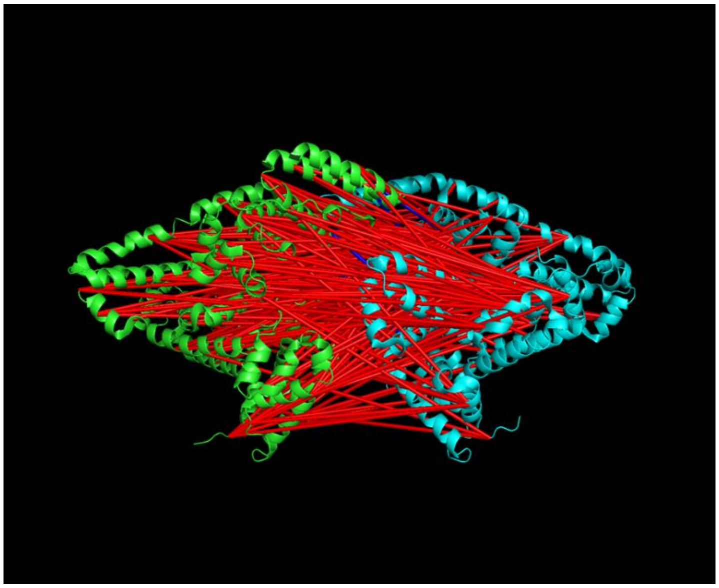


Figure 3.5. Phase 1 silver stain analysis of KUNV infected PMF following DSSO crosslinking. KUNV PMF crosslinked with DSSO and separated by SDS-PAGE. Silver staining reveals ~75 kDa band disappearance starting at 0.5 mM DSSO with full disappearance at 2mM.

MaxQuant determines the type of crosslink based on protein sequence alone, the program is unable to distinguish between loop/intramolecular crosslinks and intermolecular crosslinks when the two proteins are identical and organizing themselves into a multimer. As such, we needed to filter through these intramolecular crosslinks to determine which are derived by crosslinking two peptides from separate proteins. The 184 unique crosslinks were analyzed using the PyMOL plugin PyXlinkViewer. This plugin models the crosslinks on a protein structure and determines if distance constraints are satisfied or rejected based on those constraints [46]. The PDB file 4F5S contained a confirmed BSA dimer and was used as a reference structure for crosslink identification and pipeline verification. PyXlinkViewer modeled the intramolecular crosslinks onto this PDB file with all unique links presented as either intramolecular or intermolecular crosslinks. When treated as intramolecular crosslinks, 81 were confirmed as satisfying the distance constraints of DSSO. When the crosslinks were entered as being intermolecular, only 7 of the crosslinks satisfied the distance constraints and linked the two BSA monomers in the dimer PDB structure (**Figure 3.6**). We moved forward with these crosslinks for further analysis and pipeline verification.

We used DisVis and HADDOCK v2.2 for identifying potential interaction interfaces between the two crosslinked BSA proteins and modeling a dimer structure from distance restraints and residue analysis. Using a BSA monomer structure and the NACCESS protein analysis program, the 40% solvent accessible residues for BSA were identified and compiled into a list file [48]. This list of accessible residues was submitted to the DisVis webserver along with the validated intermolecular crosslinks and two

A



B

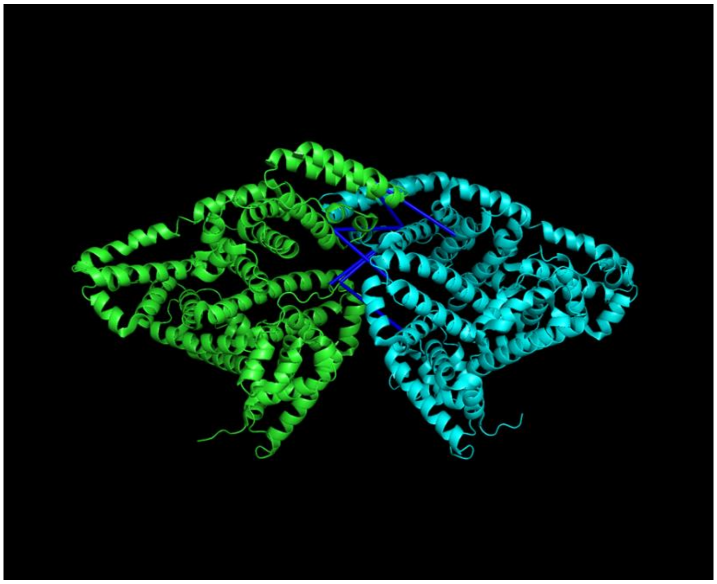


Figure 3.6. Phase 1 PyXlinkViewer modeling of XLMS acquired DSSO crosslinks. Bovine serum albumin was crosslinked with 2mM DSSO and run through mass spectrometry data acquisition at the CSU ARC-BIO facility. MaxQuant identified crosslinks modeled onto a BSA dimer structure to identify distance restraint violating crosslinks (PDB 4F5S). **A)** All MaxQuant intramolecular crosslinks were modeled as intermolecular (7/184 unique crosslinks satisfied). **B)** The 7 BSA crosslinks satisfied by the 26.5 Å distance restraints. Blue bars are satisfied crosslinks, red bars are violated crosslinks

separate BSA monomer PDB files. The output from DisVis provided potential locations of one BSA protein in relation to the other based on the distance restraints from the DSSO crosslinks (**Figure 3.7**). As the number of crosslinks considered by the program increases, the potential area of interaction with the other protein of interest becomes smaller. When all 8 crosslinks were considered, the known interface where one BSA monomer interacts with the other in the BSA dimer structure was visible in the DisVis output file. Additionally, DisVis assessed the accessible residues and their relation to the predicted BSA-BSA interface. The resulting file provided a list of residues likely to participate in the predicted interaction ranked by probability (**Figure 3.8a**). Modeling these residues on the BSA dimer reference structure shows their localization to the known interaction site (**Figure 3.8b**). This list of interaction residues and the known crosslinks were then submitted to the HADDOCK v2.2 web client with the BSA monomer PDB files to generate a series of predicted docking positions (**Figure 3.9**). Overlaying these predictions with the BSA crystal structure used as our reference shows HADDOCK's ability to accurately predict protein dimer models using crosslink restraints and residue analysis from DisVis. This experiment validated the pipeline and provided a path forward once crosslinks were identified for viral proteins in the replication compartments.

3.3.2 Phase II: WNV Sample Purification and Analysis

Following the COVID-19 lockdown, instrument availability, funding constraints, and the reassignment of WNV from a BSL3 to a BSL2+ pathogen, we made significant

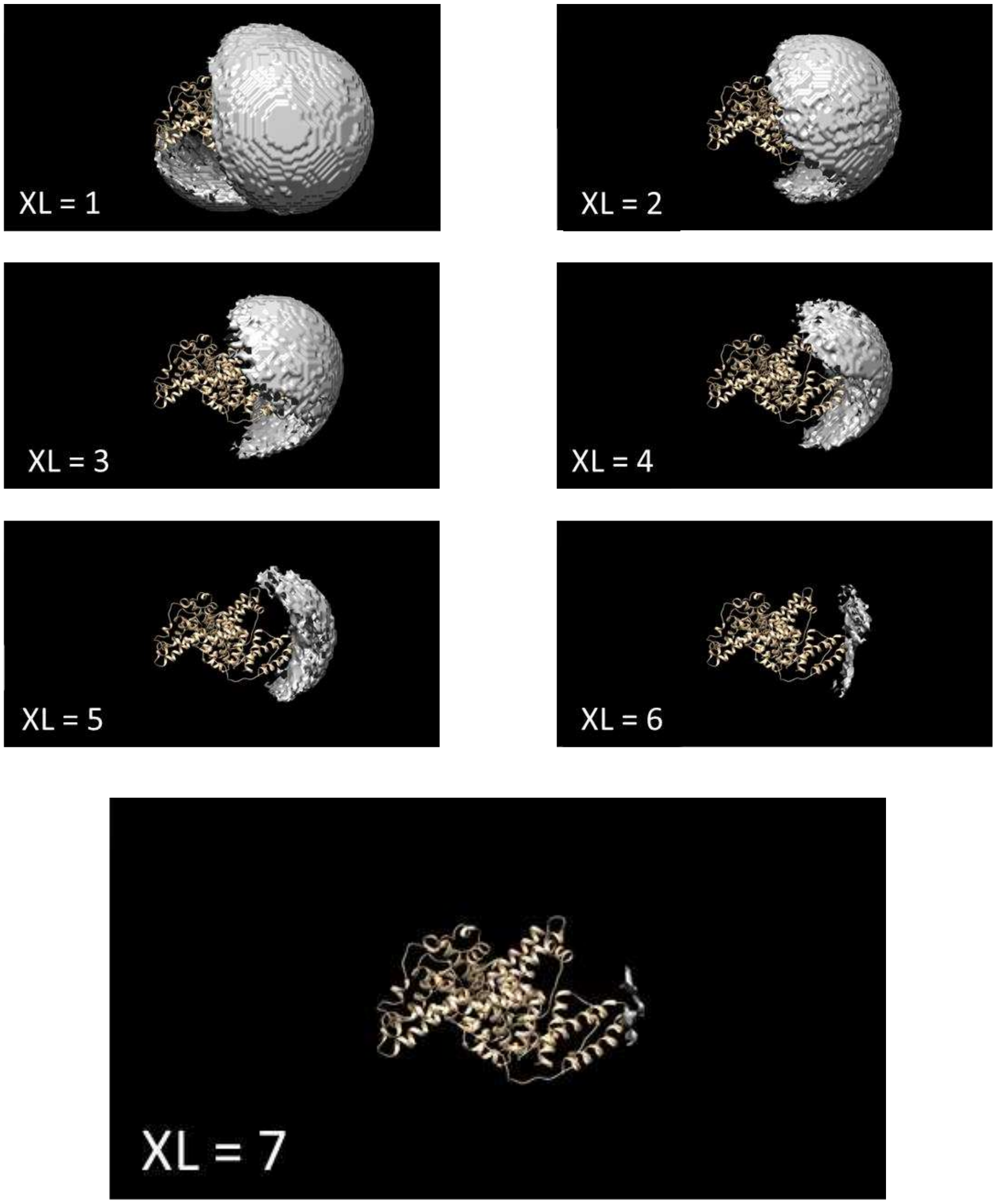
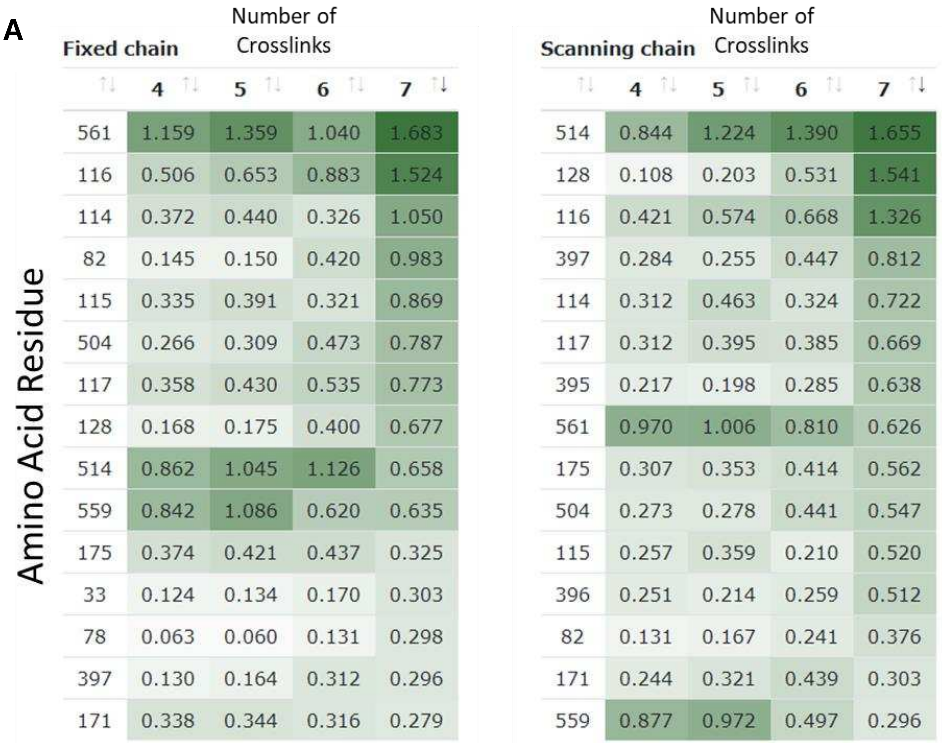


Figure 3.7. DisVis modeling of potential interaction interfaces between two BSA monomers using acquired DSSO crosslinks. DisVis analysis of the acquired DSSO crosslinks between BSA monomers revealed potential interaction interfaces (grey) on one monomer (fixed chain) compared to another (scanning chain). XL number indicates the number of submitted crosslinks DisVis is considering when predicting the interaction interface. As more XL are considered, the potential interface shrinks until it matches the interaction interface shown in the PDB 4F5S reference structure

A



Fixed chain	Number of Crosslinks			
	4	5	6	7
561	1.159	1.359	1.040	1.683
116	0.506	0.653	0.883	1.524
114	0.372	0.440	0.326	1.050
82	0.145	0.150	0.420	0.983
115	0.335	0.391	0.321	0.869
504	0.266	0.309	0.473	0.787
117	0.358	0.430	0.535	0.773
128	0.168	0.175	0.400	0.677
514	0.862	1.045	1.126	0.658
559	0.842	1.086	0.620	0.635
175	0.374	0.421	0.437	0.325
33	0.124	0.134	0.170	0.303
78	0.063	0.060	0.131	0.298
397	0.130	0.164	0.312	0.296
171	0.338	0.344	0.316	0.279

Scanning chain	Number of Crosslinks			
	4	5	6	7
514	0.844	1.224	1.390	1.655
128	0.108	0.203	0.531	1.541
116	0.421	0.574	0.668	1.326
397	0.284	0.255	0.447	0.812
114	0.312	0.463	0.324	0.722
117	0.312	0.395	0.385	0.669
395	0.217	0.198	0.285	0.638
561	0.970	1.006	0.810	0.626
175	0.307	0.353	0.414	0.562
504	0.273	0.278	0.441	0.547
115	0.257	0.359	0.210	0.520
396	0.251	0.214	0.259	0.512
82	0.131	0.167	0.241	0.376
171	0.244	0.321	0.439	0.303
559	0.877	0.972	0.497	0.296

B

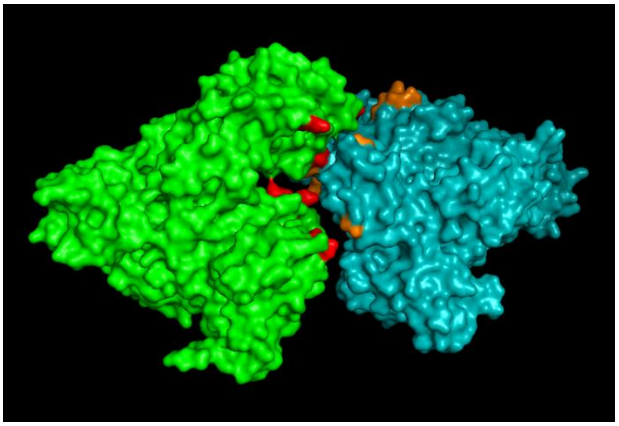


Figure 3.8. DisVis interaction residue analysis. DisVis analysis of crosslinks predicts the probability of specific solvent-accessible residues being involved in the interaction interface. **A)** Output table of DisVis presenting the probability a BSA residue is involved in the PPI based on the crosslink restraints. Residues with probabilities higher than 0.5 are considered participating in the interaction. **B)** Identified interaction residues with a probability higher than 0.5 are modeled onto the reference BSA dimer structure. The red and orange labeled residues localize to the known interaction interface

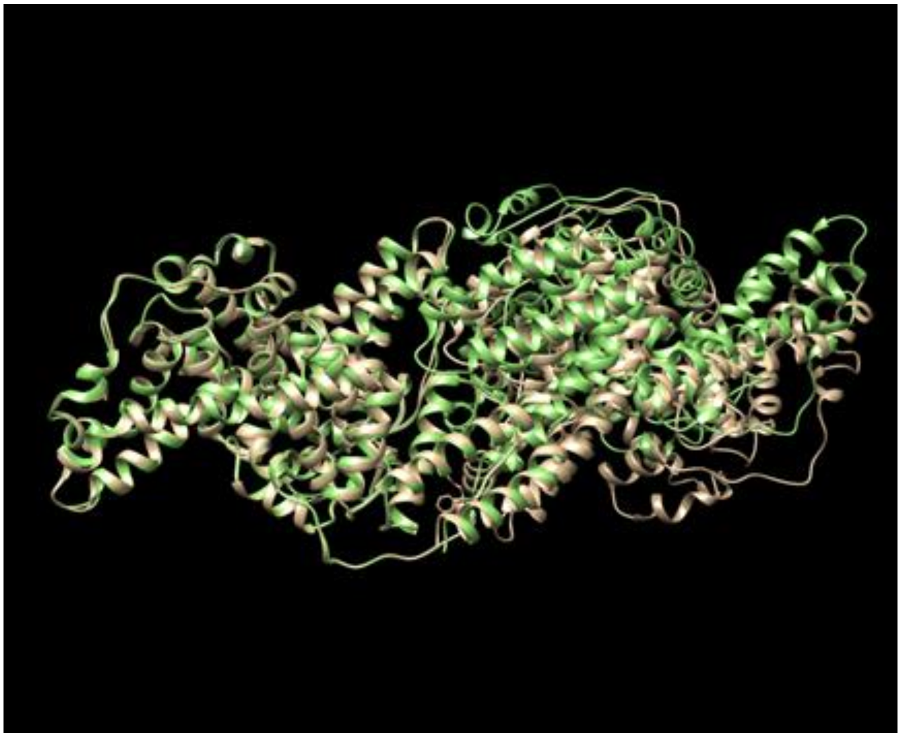


Figure 3.9. HADDOCK modeling of BSA dimer overlayed with known BSA dimer crystal structure. HADDOCK predicted docking structure of two BSA monomers based on crosslink restraints and DisVis predicted interaction interface residues (green) overlayed onto the crystal structure of a BSA dimer (tan).

changes to the established protocol from Phase I. These changes are reflected in the following results sections and in the methods section. Our revised protocol was to crosslink with 2mM of DSBU at the PMF stage and purify with ultracentrifugation and resuspension in lysis buffer. Finally, we transitioned to using the Dobos Lab's sample preparation method and Bruker timsTOF fleX MALDI-2 mass spectrometer for the sake of cost and control over the data acquisition process. Spectra were identified and compiled at the MS2 stage- no MS3 acquisition was performed.

3.3.2.1 WNV sample purification and MS preparation

With WNV no longer being a BSL3 pathogen, we transitioned to using our in-house pBG744 derived WNV for infection, which expresses GPF as a translation indicator to aid in infection visualization. With the PMF isolated, we assessed genetic content using qRT-PCR to measure both positive and negative-sense strands of WNV vRNA. A standard curve was generated using strand-specific primers and controls from a PCR of the pBG744 origin plasmid. The PMF was tested for both forms of RNA. The resulting Ct values were converted to strand copies, allowing for the determination of a positive-sense to negative-sense ratio. We observed 4777 copies of positive-sense viral RNA for every negative-sense strand in the PMF samples (**Figure 3.10**). This ratio is significantly higher than the Phase I ratio due to the reduced purity of the PMF sample compared to the density gradient fractions used in Phase I, with unreleased virions and ribosome-bound vRNA contributing to the positive-strand count. Nevertheless, the

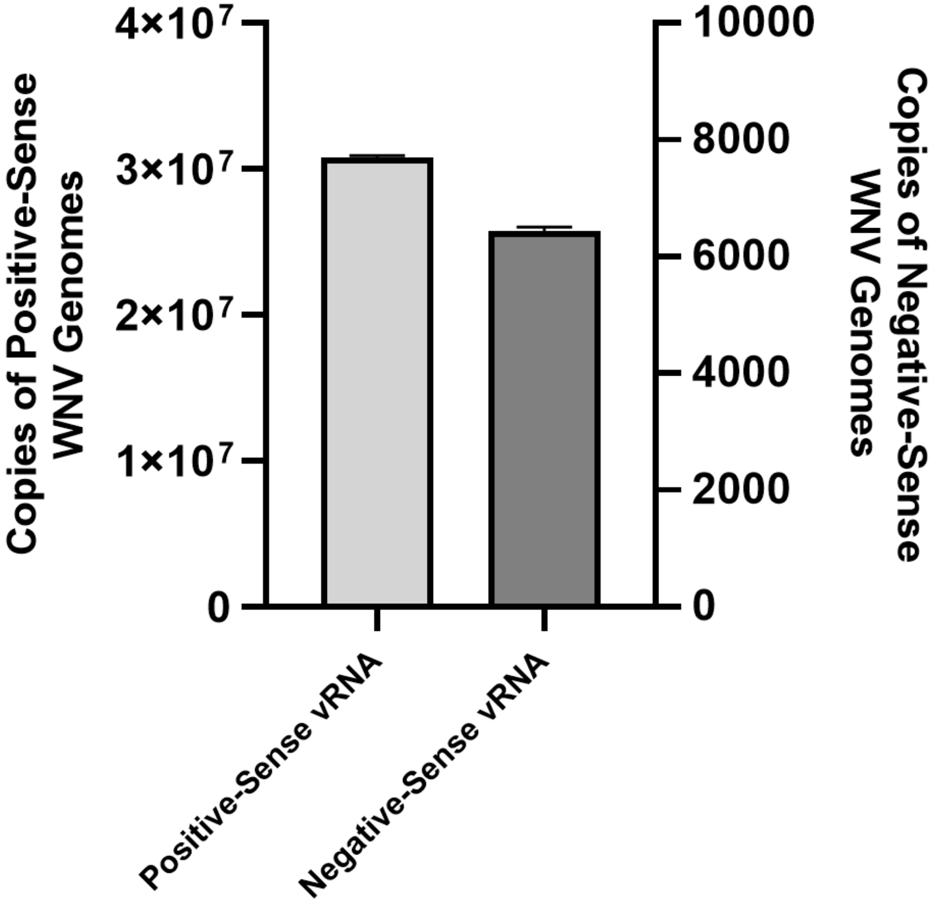


Figure 3.10. Phase II copies of positive-sense and negative-sense WNV vRNA in PMF. qRT-PCR analysis of PMF RNA for positive and negative-sense vRNA copies. Validates the presence of replication compartment-dwelling negative-sense vRNA in the PMF later used for XLMS.

presence of negative-strand viral RNA is a critical indicator of the replication compartment, as this RNA form can only survive unperturbed by host immune factors within the protective compartment. This provided confirmatory evidence for the presence of replication compartments within the WNV PMF samples.

In a series of changes from the Phase I protocol, we decided to use a new crosslinker at an earlier step along with altering the crosslinked-product purification method. The PMF was crosslinked directly with 2mM of DSBU using the same incubation protocol as before. I made the decision to simplify the purification protocol to reduce the time between sample collection and crosslinking, as non-crosslinked samples were at risk of degradation during the lengthy ultracentrifugation steps. The crosslinking between peptides of interest would make band observation difficult on a traditional northern blot. Therefore, following crosslinking and purification, the protein content in these samples was assessed with a dot blot assay. The faint signal in the crosslinked (XL) RIPA supernatant well and the strong signal in the XL RIPA pellet well show that NS3 was retained in the crosslinked samples (**Figure 3.11a**). The lack of signal in the PMF band was likely due to low NS3 concentration in the 9ml of PMF compared to the 150µl XL RIPA samples. Silver stain analysis shows that crosslinking with 1mM of crosslinker was successful, with the deterioration of distinct bands in the crosslinked sample wells compared to the PMF. (**Figure 3.11b**). With the RIPA supe and RIPA pellet samples collected, we performed a BCA and adapted the Dobos Lab In-Solution Digestion protocol to prepare samples for XLMS. The Dobos Lab's Bruker timsTOF flex MALDI-2 mass spectrometer was used for data acquisition.

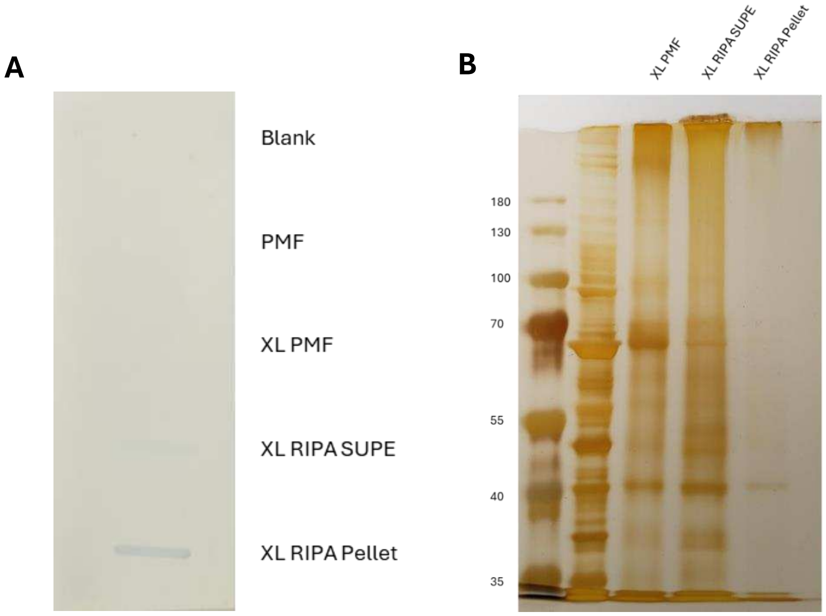


Figure 3.11. Phase II protein analysis of DSBU-crosslinked WNV RIPA supernatant and pellet samples. Samples were crosslinked by DSBU at 2mM followed by ultracentrifugation at 100,000xg for 1 hour. Pellets were resuspended in RIPA buffer, sonicated, spun down, and separated into XL RIPA supernatant (supe) and XL RIPA pellet samples. **A)** Dot blot analysis of WNV NS3 content in DSBU crosslinked samples. Faint band in XL RIPA supe indicates NS3 was retained in the pellet. Low signal in the PMF and XL PMF samples due to NS3 dilution in higher volumes (9ml in PMF compared to 200µl in XL RIPA samples. **B)** Silver stain analysis of DSBU crosslinked samples, demonstrating the loss of distinct bands in XL samples congruent with linkages between proteins

3.3.2.2 XLMS results for WNV crosslinked samples

A combined .fasta file containing the WNV proteome derived from our pBG744 origin plasmid and a reduced human proteome was first used to identify any host-host PPIs or host-virus PPIs from the WNV RIPA supe and pellet samples. Protein identification was filtered by MaxQuant score, with scores higher than 40 being kept for identification. The RIPA supernatant samples had 1,581 protein identifications and the RIPA pellet samples had 1,349 identifications (**Table 3.2a**). Overall crosslink analysis showed a significant number of host-host loop, intramolecular, and intermolecular crosslinks for both RIPA supernatant and pellet samples. Posterior error probability scores of less than or equal to 0.05 was used to remove low-scoring crosslinks. No intermolecular crosslinks between host proteins and viral proteins were identified. In the RIPA supernatant sample, MaxQuant identified 3388 loop links, 301 intramolecular crosslinks, and 119 intermolecular crosslinks, while the RIPA cell sample had 3064 loop-links, 206 intramolecular crosslinks, and 106 intermolecular crosslinks **(Table 3.2a)**. STRING database analysis of both samples' ID data revealed the subcellular localization of the identified proteins, with prominent origins being the nucleus, cytoplasm, and ER (**Table 3.2b**) [49].

The breakdown of identified WNV peptide can be seen in **Table 3.3**, with most viral peptides originating from the envelope (E) protein, NS1, NS3, and NS5. The C, prM, NS2b, and NS4a identified peptides amassed MaxQuant scores less than 40, indicating

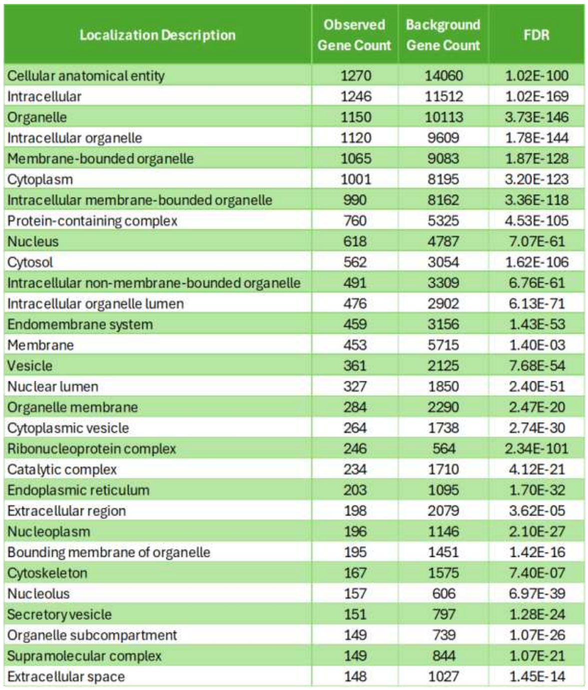
Table 3.2. Phase II protein ID breakdown following XLMS analysis of DSBU-crosslinked RIPA samples. A) Number of identified proteins and crosslinks in RIPA supernatant and pellet samples. Proteins with a MaxQuant score of <40 were excluded from overall protein ID consideration **B)** Top 30 locations for identified protein origins in the RIPA supernatant (left) and RIPA pellet (right) samples. Subcellular localization identified with STRING analysis.

A

RIPA Sample	Protein IDs	Total XL	Loop Crosslinks	Intramolecular Crosslinks	Intermolecular Crosslinks
Supernatant	1581	3808	3388	301	119
Pellet	1349	3376	3064	206	106

B

Localization Description	Observed Gene Count	Background Gene Count	FDR
Cellular anatomical entity	1485	14060	6.22E-111
Intracellular	1463	11512	2.63E-196
Organelle	1339	10113	3.53E-161
Intracellular organelle	1304	9609	1.03E-159
Membrane-bounded organelle	1233	9083	4.04E-138
Intracellular membrane-bounded organelle	1150	8162	7.25E-130
Cytoplasm	1096	8195	1.91E-101
Protein-containing complex	911	5325	7.95E-131
Nucleus	801	4787	2.28E-102
Cytosol	650	3054	7.44E-118
Non-membrane-bounded organelle	609	3321	7.66E-84
Intracellular non-membrane-bounded organelle	608	3309	5.99E-84
Intracellular organelle lumen	565	2902	8.49E-85
Endomembrane system	478	3156	5.78E-40
Nuclear lumen	429	1850	2.39E-80
Vesicle	385	2125	2.53E-46
Cytoplasmic vesicle	292	1738	9.66E-29
Catalytic complex	292	1710	8.78E-30
Ribonucleoprotein complex	288	564	8.15E-116
Organelle membrane	279	2290	2.11E-10
Nucleoplasm	259	1146	1.37E-43
Cytoskeleton	222	1575	1.72E-13
Nucleolus	199	606	7.29E-53
Bounding membrane of organelle	198	1451	8.05E-11
Supramolecular complex	185	844	6.34E-29
Endoplasmic reticulum	168	1095	9.51E-13
Secretory vesicle	161	797	8.91E-22
Cell junction	160	1053	8.04E-12
Secretory granule	156	719	9.19E-24
Nuclear body	142	381	3.59E-42



Localization Description	Observed Gene Count	Background Gene Count	FDR
Cellular anatomical entity	1270	14060	1.02E-100
Intracellular	1246	11512	1.02E-169
Organelle	1150	10113	3.73E-146
Intracellular organelle	1120	9609	1.78E-144
Membrane-bounded organelle	1065	9083	1.87E-128
Cytoplasm	1001	8195	3.20E-123
Intracellular membrane-bounded organelle	990	8162	3.36E-118
Protein-containing complex	760	5325	4.53E-105
Nucleus	618	4787	7.07E-61
Cytosol	562	3054	1.62E-106
Intracellular non-membrane-bounded organelle	491	3309	6.76E-61
Intracellular organelle lumen	476	2902	6.13E-71
Endomembrane system	459	3156	1.43E-53
Membrane	453	5715	1.40E-03
Vesicle	361	2125	7.68E-54
Nuclear lumen	327	1850	2.40E-51
Organelle membrane	284	2290	2.47E-20
Cytoplasmic vesicle	264	1738	2.74E-30
Ribonucleoprotein complex	246	564	2.34E-101
Catalytic complex	234	1710	4.12E-21
Endoplasmic reticulum	203	1095	1.70E-32
Extracellular region	198	2079	3.62E-05
Nucleoplasm	196	1146	2.10E-27
Bounding membrane of organelle	195	1451	1.42E-16
Cytoskeleton	167	1575	7.40E-07
Nucleolus	157	606	6.97E-39
Secretory vesicle	151	797	1.28E-24
Organelle subcompartment	149	739	1.07E-26
Supramolecular complex	149	844	1.07E-21
Extracellular space	148	1027	1.45E-14

Table 3.3. Phase II WNV protein ID breakdown. Number of peptides identified for each of the 10 viral proteins in the **(A)** RIPA supernatant and **(B)** RIPA pellet samples following XLMS data acquisition. All viral proteins were identified despite certain proteins having MaxQuant scores less than 40. Low scores and low presence of transmembrane proteins highlight the difficulty of analyzing these small, membrane bound proteins by XLMS.

A

Protein IDs	Unique Peptide Counts	Sequence coverage [%]	Score	Percent Of Total WNV Proteins
Capsid	4	22	10.4	2.4
prMembrane	2	43.5	7.9	1.2
Envelope	30	68	323.3	18.3
NS1	18	59.9	181.8	11.0
NS2A	6	49.8	80.3	3.7
NS2B	2	19.8	15.6	1.2
NS3	39	78	323.3	23.8
NS4A	3	27.8	2.2	1.8
NS4B	7	40.6	94.2	4.3
NS5	53	65.4	323.3	32.3

B

Protein IDs	Unique Peptide Counts	Sequence coverage [%]	Score	Percent Of Total WNV Proteins
Capsid	10	36.6	80.6	5.6
prMembrane	3	53.3	30.5	1.7
Envelope	34	69.8	323.3	19.2
NS1	21	61.9	323.3	11.9
NS2A	8	55.8	30.2	4.5
NS2B	3	33.6	61.4	1.7
NS3	47	84.3	323.3	26.6
NS4A	3	27.8	9.2	1.7
NS4B	5	29.3	94.1	2.8
NS5	43	57.9	323.3	24.3

low confidence in their identification. Additionally, there were few viral transmembrane proteins observed across the board except for the envelope protein. The underrepresentation of transmembrane protein IDs may be due to their relatively small size and their lack trypsin cleavage sites compared to the other viral proteins.

3.3.2.3 WNV loop and intramolecular crosslinks

We next sought to specifically assess the viral protein and crosslinked PPIs in the samples. We used a .fasta file containing the same WNV Lineage II proteome derived from our pBG744 plasmid as before. The viral crosslink peptide breakdown can be seen in **Figure 3.12**.

As mentioned in Phase I, oligomers complicate MaxQuant determination of crosslink type. Intermolecular crosslinks are identified as loop or intramolecular links when crosslinking two proteins of the same sequence. As such, we investigated all loop and intramolecular crosslinks on single protein structure models to determine if the distance limitations were satisfied. We identified 272 loop links on WNV peptides, 81 of which were unique crosslinks between specific residues. Loop-linked peptides originated from all WNV proteins except for the pre-membrane protein (prM) and NS4a. We observed 32 intramolecular crosslinks that all came from either NS1, NS3, or NS5, with 13 unique crosslink pairs. For PyXlinkViewer we used a distance restraint of 26.5-12.5 Å for the DSBU spacer, 3 Å for the two carbon-nitrogen bonds linking the

A

Protein ID	Total Crosslinks	Loop Crosslinks	Intramolecular Crosslinks
Capsid	2	2	0
prMembrane	0	0	0
Envelope	16	16	0
NS1	3	2	1
NS2A	2	2	0
NS2B	1	1	0
NS3	37	30	7
NS4A	0	0	0
NS4B	5	5	0
NS5	27	22	5
Total Loop and Intramolecular Crosslinks	93	80	13

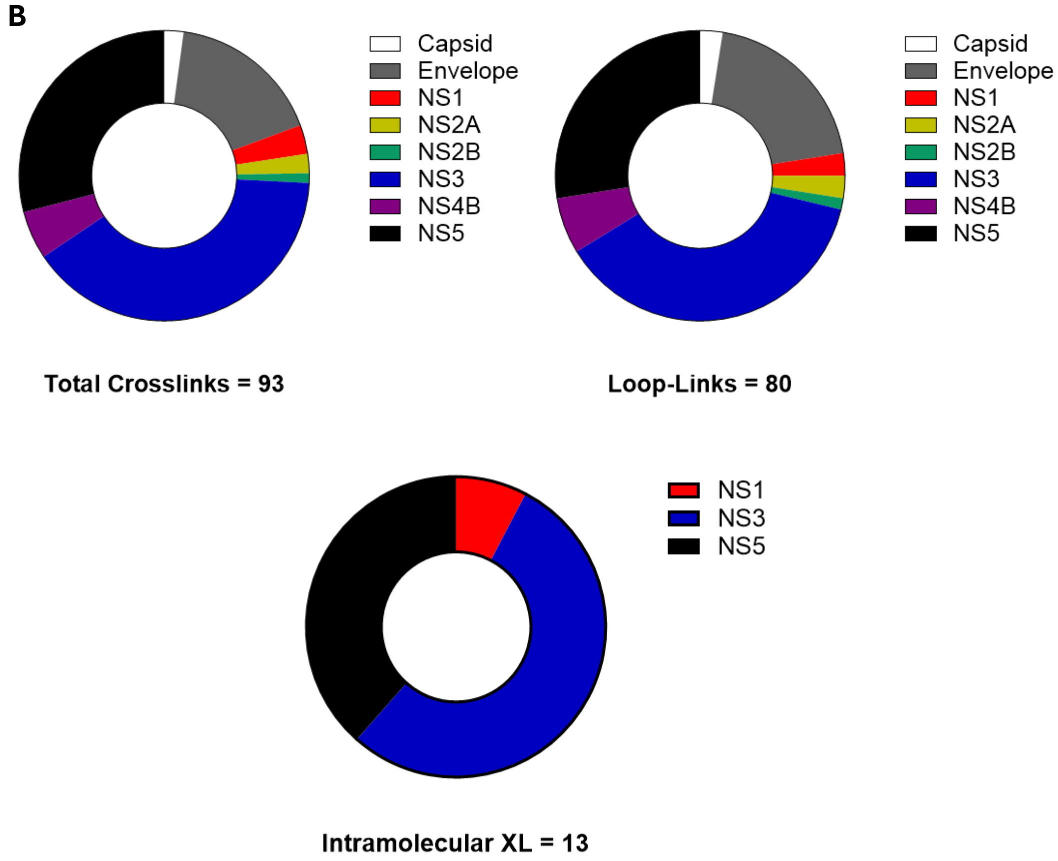


Figure 3.12. Phase II breakdown of WNV replication compartment identified loop or intramolecular crosslinks. A) Table with total breakdown of loop and intramolecular crosslinks, with a total of 93 observed. **B)** Graphical breakdown of total, loop, and intramolecular crosslinks (XL) observed in the RIPA supernatant and pellet samples.

to the two amino-acid residues, and 11 Å for the maximum distance between an alpha-carbon and the end of the side chain as observed with the PyMOL Measurement wizard tool. We mapped the loop and intramolecular crosslinks onto their respective proteins to observe their localization and assess their viability. To ensure accurate distance measurements between sidechains, any alpha-carbon to alpha-carbon distances of 20 Å or higher were investigated further using PyMOL's Measurement wizard to measure distance between the side-chain ends.

Two unique loop-links and one unique intramolecular crosslink were observed for the WNV NS1 protein. The NS1 dimer crystal structure from Akey et al. was used to map the NS1 crosslinks (PDB 4O6D) with PyXlinkViewer [50]. These crosslinks are located on the tips of the "V"-shaped NS1 monomer with no crosslinks observed at the NS1-NS1 interface seen in the published structure (**Figure 3.13**).

The majority of viral protein loop and intramolecular crosslinks were identified in NS3. With no crystal structure for WNV NS3 published, we used AlphaFold2 to generate a structure based on the protein sequence derived from our pBG744 WNV construct plasmid. Modeling was accomplished using the ColabFold webserver [51-54]. All NS3 crosslinks satisfied the distance requirement in PyXlinkViewer on the generated structure, providing further evidence for AlphaFold2's accuracy at predicting structures based on sequences (**Figure 3.14**). Interestingly, the crosslinks tended to favor one side of the NS3 protein.

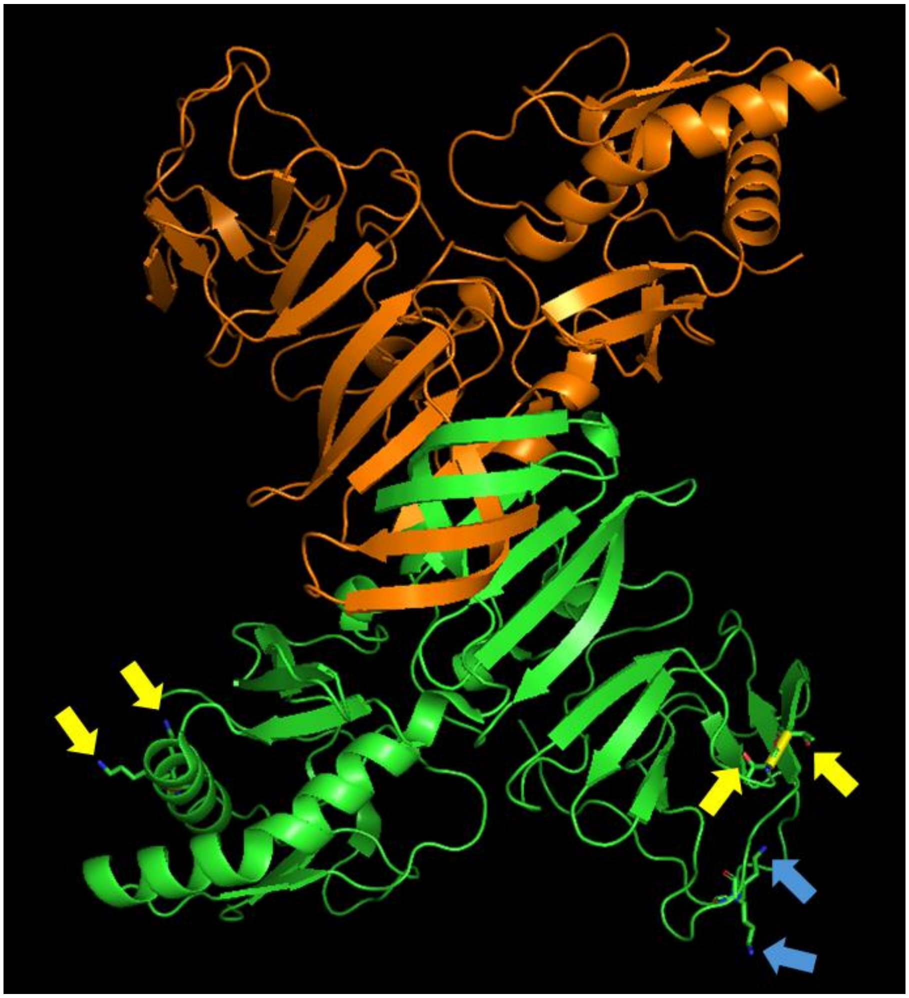


Figure 3.13. Mapped NS1 loop and intramolecular crosslinks. Dimer NS1 file showing two NS1 molecules (green and orange) with the mapped loop (yellow) and intramolecular (blue) crosslinks mapped onto an NS1 monomer (green). Yellow arrows for loop-linked residue emphasis. Blue arrows for intramolecular crosslinked residue emphasis. NS1 dimer structure PDB 4O6D.

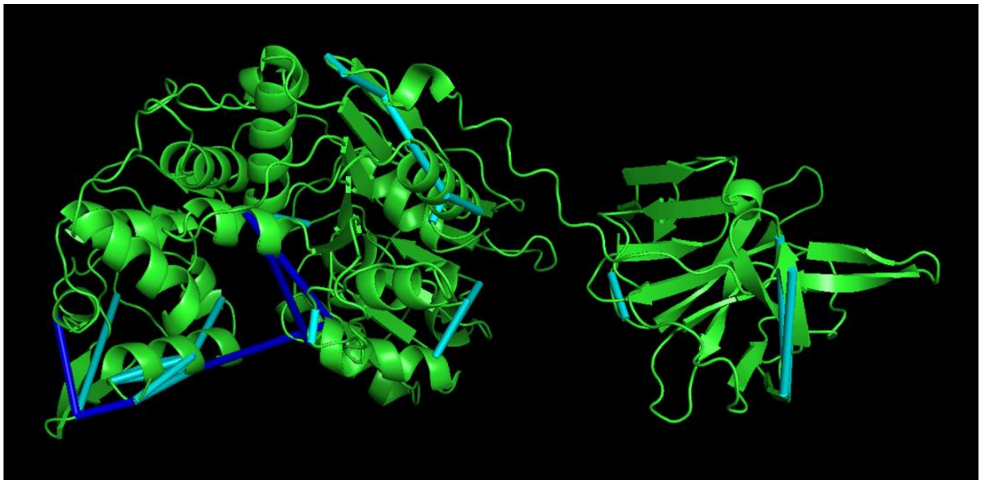


Figure 3.14. Mapped NS3 loop and intramolecular crosslinks. NS3 loop (teal) and intramolecular (blue) crosslinks mapped onto NS3 structure. NS3 structure generated using Alphafold2 and the WNV lineage II sequence for NS3.

As with NS3, a protein structure for NS5 was generated using AlphaFold2 as a published structure for WNV does not yet exist. Modeling of the crosslinks revealed five loop-links that violated the 26.5 Å distance restraint between the ends of the linked residue side-chains (**Figure 3.15**). The distance-restraint violating crosslink pairs were T728-S746, T451-K462, G1-K13, Y374-K391, and K269-K280. Such distance-restraint violations indicate the crosslinks were not intramolecular/loop crosslinks but instead intermolecular, as a crosslink between the two residues on identical proteins in a dimer may be validated instead.

3.3.2.4 WNV NS3-NS4b Intermolecular Crosslink

A single intermolecular crosslink was observed between the NS3 and NS4b viral proteins, with the alpha-amine of the NS3 N-terminus crosslinked to K172 on NS4b (**Figure 3.16**). This crosslinking interaction was observed in both the RIPA supernatant and pellet samples across multiple injections and showed up in both the combined human-WNV proteome and WNV proteome only analysis. This interaction supports known examples of NS3 and NS4b interactions in the replication compartment [21, 22, 28].

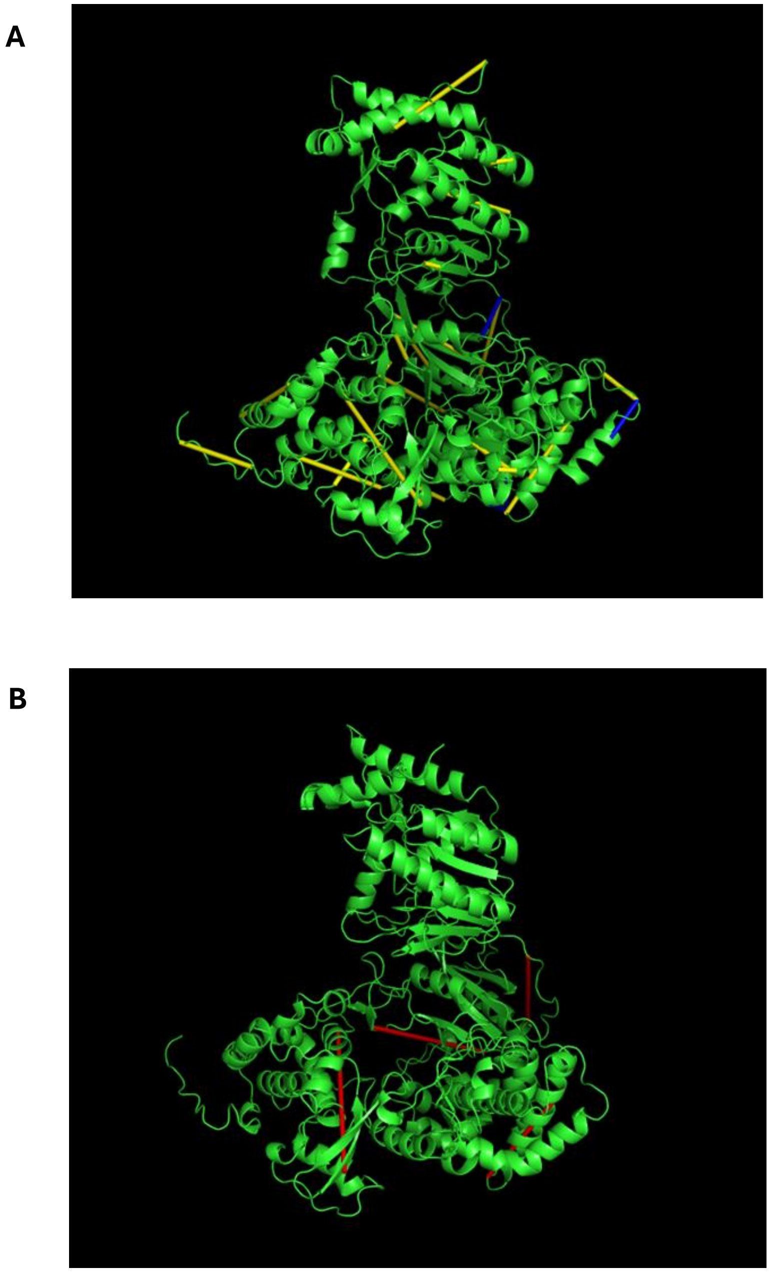


Figure 3.15. Mapped NS5 loop and intramolecular crosslinks. NS5 structure generated using Alphafold2 and the WNV lineage II pBG744 sequence for NS5 **A)** NS5 loop (yellow) and intramolecular (blue) crosslinks mapped onto NS5 structure. **B)** DSBU crosslinks that violated the 26.5 Å distance restraint (red).

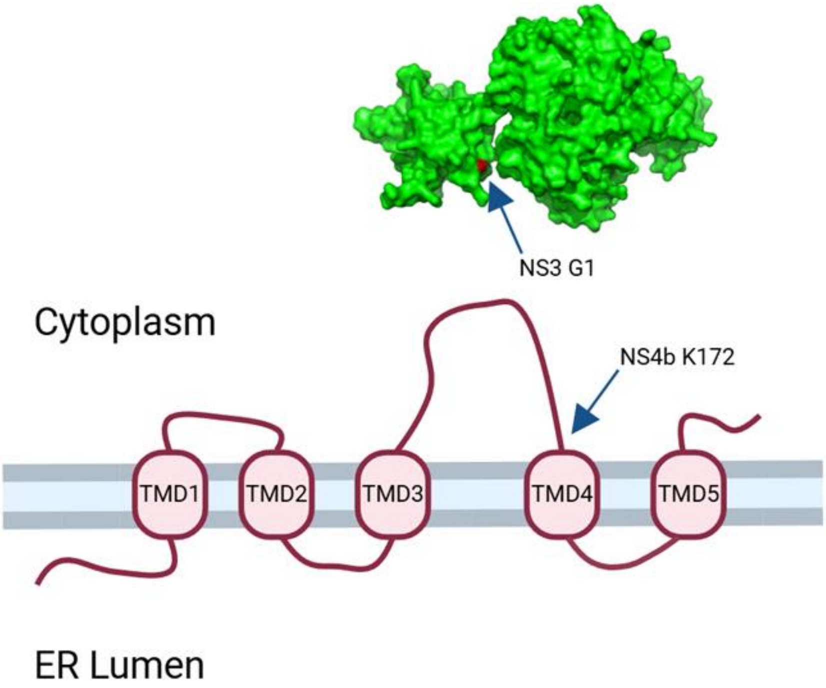


Figure 3.16. Locations of crosslinked residues on NS3 and NS4b. One intermolecular crosslink was observed between NS3 G1 (N-terminus amine) and NS4b K172. Figure indicates the location of the respective residues, with the NS4b topology being base doff of Li et al. 2015. No Structure for NS4b exists. Created with BioRender.com.

3.4 Discussion

Flavivirus proteins are multifunctional contributors to infection, with the largest association of viral proteins occurring in the replication compartment. In these compartments the nonstructural proteins work to manipulate host ER membrane lipids to create protective invaginations. In this environment the negative-sense template vRNA strand and newly generated positive-sense vRNA can evade host immune detection. Inside, the NS3 and NS5-containing replication complex works to produce newly capped viral proteins while other nonstructural proteins form bonds with each other to maintain the stability of the compartment. The proper formation of this compartment is critical for flavivirus infection success, and thus understanding PPIs would open fresh avenues for the development of much-needed antiviral therapeutics.

In addition to discussing the data outlined in the Results section, here I will discuss the challenges this study faced as well as the justification for changing protocols over time as situations evolved over the last five years.

3.4.1 Choosing XLMS to Study PPIs

The challenge in understanding this critical environment for viral infection lies in a lack of knowledge on the organization of this replication compartment and its components. Previous studies have utilized methods such as immunofluorescence or nanogold-labeled electron tomography to identify proteins existing in replication compartments *in situ*. The resolution of these techniques limits our ability to determine the stoichiometry that lends these PPIs their function. Co-immunoprecipitation and selective mutagenesis have also been used to determine if specific residues are critical for function or interaction between two proteins. While providing useful information, mutagenesis has the potential for creating pleiotropic effects on protein structure throughout the infection cycle, making it difficult tie the importance of a mutated residue to the interaction of interest. Finally, *in vitro* experiments with purified protein have been used to study PPIs and protein-RNA interactions. Mutagenesis experiments on these proteins provided insight into pairwise binding interactions and the resulting effects on functionality, while cryo-electron microscopy can look at specific interactions in non-mutated recombinant proteins [15]. These *in vitro* methods have been fruitful for understanding the form and function of specific proteins but still are hampered by the pitfalls of mutagenesis and recombinant expression while simultaneously ignoring the context of a “natural” infection event in a cell.

Ultimately, a method capable of labeling and identifying PPIs within flavivirus replication compartments *in situ* is necessary to understand both their form and function. We identified crosslinking mass spectrometry (XLMS) as a promising method for developing a complete look into the interactions necessary for a successful infection event. This method has been used in the past to map PPIs between host proteins [55-57]. Choosing XLMS provided significant theoretical and practical benefits. With the use of chemical crosslinkers, replication compartments could be labeled *in situ* to provide a snapshot of the organization of proteins within. Mass spectrometry can then identify crosslinked peptides followed by MaxQuant determining crosslinked proteins. The distance constraints between crosslinked proteins at these linked residues are invaluable for developing quaternary structure models for the compartment components. On the practical side, we possessed the necessary facilities for running these experiments, including a BSL2 tissue culture room with incubators, an ultracentrifuge, and access to mass spectrometers for data acquisition. The Bonvin Lab's DisVis and HADDOCK software provided the means to rebuild PPIs and their interfaces. With these tools in hand, we set forth to establish a pipeline for studying PPIs in the flavivirus replication compartments.

3.4.2 The Woes and Evolution of XLMS Method Development

3.4.2.1 Establishing the crosslinking data analysis process

We centered our XLMS study around the use of MS/MS-cleavable crosslinkers for protein linking. These crosslinkers can be cleaved asymmetrically by collision-induced dissociation (CID) as the ions travel from the MS1 and MS2 stages of data acquisition, providing characteristic m/z mass shifts that programs use to identify linked peptides. We first chose the MS/MS-cleavable crosslinker DSSO due to its chemical characteristics. Namely, DSSO has a 10-angstrom spacer and primary amine reactivity. Primary amines are those seen in lysine side chains participating in polar PPIs. DSSO is also membrane permeable, a critical characteristic needed to penetrate the replication compartment. In Phase II, we used the crosslinker disuccinimidyl dibutyric urea (DSBU) instead of DSSO. DSBU has a longer 12.3 angstrom crosslinker arm with reactivity to not only lysine but also serine, tyrosine, threonine, and protein N-terminal amines. This switch was made due to the increased literature support for PPI XLMS studies using DSBU as well as the increased crosslinking potential of a longer spacer and expanded reactivity [36, 45, 58].

The method for mass spectrometry data acquisitions also went through a series of iterations due to evolving circumstances. The original nontargeted mass spectrometry experiment used to identify the host proteins present in microsome samples was performed by the University of Colorado School of Medicine Proteomics Core Facility due to COVID-19 pandemic related restrictions at Colorado State University. We moved on from this process following the onset of a mild heart attack upon receiving the bill for a four-sample run. The optimization of mass spectrometry settings would require a more economical and convenient facility, especially considering we had only guidance from the literature to establish an entire pipeline in a novel system. We next began working with the CSU ARC-BIO facility due to their proximity, cost, and ability to work with us on discussing methods and guiding sample generation. This facility acquired the BSA XLMS data used to validate the DisVis/HADDOCK data analysis pipeline. Unfortunately, using the ARC-BIO core facility resulted in months of delays due to facility upgrades and a significant sample backlog. It took seven months for experimental samples to be analyzed and the resulting data yield was low. For optimizing mass spectrometry parameters, the project needed a faster turnaround between runs. Thankfully, Dr. Karen Dobos saved the project by volunteering the use of her lab's new Bruker timsTOF-fleX MALDI-2. The cost of data acquisition was reduced, and the Dobos Lab protocols allowed me to learn better sample preparation techniques. Dr. Luisa Nieto Ramirez from the Dobos lab provided invaluable guidance on sample preparation and mass spectrometry protocol generation. This final arrangement for mass spectrometry data acquisition in Phase II was significantly more fruitful compared to the previous iterations.

The BSA control XLMS experiment not only validated our crosslinking protocol but also helped to establish a data analysis pipeline. The MaxQuant analysis of the spectra indicated a high number of intramolecular crosslinks, with 184 unique pairs of linked residues. When all intramolecular links are modeled on a monomer structure of BSA using the PyXlinkViewer plugin on PyMOL, only 81 satisfy the DSSO distance constraint. This highlights a limitation of the XLMS process when dealing with homo-multimers: intermolecular crosslinks between proteins with identical sequences would be labeled as loop or intramolecular links instead. BSA aggregates in a variety of protein orientations, resulting in many of these intramolecular crosslinks being improperly labeled. We resolved this issue for our pipeline validation by using a known BSA dimer structure to identify which of the identified intramolecular crosslinks were actually intermolecular for this specific interaction. We relabeled the crosslinks as intermolecular, and, after modeling them on the dimer structure, 7 crosslinks were validated (**Figure 3.6**).

The pipeline then continued on to DisVis and HADDOCK analysis. The program tutorials were mercifully thorough and provided the groundwork for assessing the identified BSA crosslinks. A monomer BSA structure (adapted from the 4F5S BSA dimer PDB file) was analyzed by NACCESS for solvent accessible residues which were submitted to DisVis along with the restraints file. The output was exciting, as it correctly predicted the interaction interface of the dimer and provided a list of residues predicted to participate directly in this interaction, including 10 residues on one BSA chain and 12 residues on the other (**Figure 3.8**). This interaction analysis would be valuable for the replication compartment analysis by providing targets for further antiviral PPI disruption studies. The list of residues and the crosslinks could then be submitted to HADDOCK, which provided a series of predicted dimer structures (**Figure 3.7**). These matched remarkably well with the known dimer crystal structure. With viral protein interaction residues and crosslinks, HADDOCK would be able to generate structures demonstrating their orientation to one another which could then be used to build a larger model of the replication compartment.

Establishing this process also revealed a new challenge to overcome when sufficient crosslinks are identified between replication compartment proteins. All these programs require known structures for each of the identified linked proteins. The majority of flavivirus nonstructural proteins do not have published crystal structures, and there are even fewer crystal structures for WNV proteins specifically. For this reason, we turned to AlphaFold2 to generate our NS3 and NS5 protein structures, as it provided a sequence specific structure for our proteins of interest. Future work into modeling uncharacterized transmembrane nonstructural proteins is necessary for painting a complete picture with PyXlinkViewer, DisVis, and HADDOCK.

The BSA XLMS control experiment validated our downstream analysis pipeline for proteins with a known or modeled structure. Unfortunately, we never used the DisVis/HADDOCK portion of the pipeline, as we failed to identify suitable intermolecular crosslinks for analysis. Due to server-side trouble with DisVis, we were even unable to use the programs for modeling potential NS5 dimers with the validated loop-links. Nevertheless, the BSA experiments revealed the dimer/multimer challenge with XLMS and led to workarounds for missing protein structures. Having this pipeline validated will help to determine critical PPIs in the replication compartment once technical issues with server function and crosslink identification are rectified.

3.4.2.2 Phase I optimizations: an evolving sample purification protocol

Much like the crosslinking and data acquisition process highlighted in the previous section, the approach to obtaining, crosslinking, purifying, and running our WNV replication compartments on a mass spectrometer went through several iterations over the years. There was a significant gap in the optimization process due to the COVID-19 pandemic that introduced new pressure to streamline the process. These changes were fueled by insufficient pipeline results, financial conditions, and detractors from practicality.

The Phase I purification protocol used the ER Isolation Kit instructions as the framework and adapted to fit our needs. This process was very successful at purifying replication compartments. The differential centrifugation of ER microsomes resulted in fractions that were concentrated with both ssRNA and viral proteins. Fraction 12 had a positive-sense genome to negative-sense genome copy ratio of roughly 69:1, which suggests the existence of concentrated replication compartments in the sample (**Figure 3.3)**. Both NS3 and NS5 content were high in fractions 11 and 12, a requirement for intact replication complexes within the compartment (**Figure 3.4b and 3.4c**). For crosslinking, the silver stain gel demonstrating the effect of different DSSO crosslinker concentrations established 1mM as the minimum concentration to use with the microsome samples (**Figure 3.5)**. The presented data showed that replication compartments were preserved during the purification process and were capable of being crosslinked.

The density gradient ultracentrifugation provided a challenge: the amount of effort and time it took to purify the fractions was substantial. The process delayed the characterization assays necessary for identification of replication compartment-concentrated fractions prior to crosslinking. Due to minute variations at any step of the process, the identified replication compartment fraction of interest would vary between fraction 12 to 14 depending on the run. The multi-day process risked sample degradation before crosslinking, which would not be apparent until after mass spectrometry analysis thereby wasting time and resources. To streamline the purification process, we introduced a microsome pulldown step at the PMF stage using ExoQuick TC. ExoQuick TC is a polymer that binds to specific microsome markers that can then be precipitated through a short, low-speed centrifugation. The resuspended pellet would then be crosslinked by DSSO and submitted to ARC-BIO for analysis. This process of purification and crosslinking effectively purified protein-rich samples. However, this method of purification was far from optimal following mass spectrometry sample preparation and analysis. Subsequent mass spectrometry results of ExoQuick purified compartments were lackluster. Sparse crosslinks were identified in our samples with the majority being mono-links and loop-links. This likely stemmed from the loss of compartment-containing microsomes due to failed ExoQuick pulldown. Additionally, compartments were likely further diluted by the added purification steps needed to eliminate traces of the ExoQuick polymer. The ExoQuick polymer had to be removed as it endangered the mass spectrometer instrument, threatening to clog or damage the chromatography columns and hinder data acquisition. These disappointing ExoQuick purification results helped push the evolution of sample purification methods towards the established Phase II methods.

3.4.2.3 Phase II: data acquisition and failed iterations

Phase II was brought on by a series of events that drastically changed many parts of the purification, crosslinking, and mass spectrometry process. We made the change from KUNV as our model pathogen to WNV following the WNV BSL rescheduling from 3 to 2+. Our in-house facilities allowed us to generate the numerous WNV microsome batches needed without the additional training, precautions, or risk associated with a BSL3 environment. As mentioned previously, we transitioned to DSBU crosslinking due to its chemical properties and the availability of literature to develop the mass spectrometry data acquisition protocol. Finally, an eight-month delay in sample analysis by the ARC-BIO core facility prompted us to take sample preparation into our own hands with the Dobos Lab’s timsTOF flex MALDI-2 mass spectrometer.

As with Phase I, the optimization and validation of the Phase II sample generation process went through many iterations. Ultimately, we decided to crosslink the PMF with DSBU at 2mM and purify the crosslinked sample by ultracentrifugation and final RIPA resuspension. The Phase I experiments already demonstrated the presence of intact replication compartments in later fractions, so we inferred that the PMF also contained compartments. This resulted in a diluted sample compared to the fractionation method: there was a higher negative-sense to positive-sense ratio by a factor of 100 and lower viral protein concentration (**Figure 3.10 and 3.11a**). Still, this process significantly accelerated the sample generation and characterization process. Furthermore, this method resulted in an increase in identified viral protein crosslinks compared to Phase I when analyzed with the new mass spectrometry data acquisition system.

Following the establishment of the viral ER isolation and crosslinking method in Phase I, several protocol alterations were employed in Phase II to enrich the intermolecular crosslinks in mass spectrometry samples. We fractionated crosslinked PMF using size-exclusion chromatography (SEC) in a bid to remove smaller or unlinked proteins. This SEC method has been used in the literature to enrich protein samples for

XLMS [59, 60]. This resulted in several fractions possessing NS3, which were pooled, concentrated by ultracentrifugation, and analyzed by mass spectrometry. This resulted in few crosslink identifications and no new intramolecular crosslinks observed at all. Other publications use this method for sample purification and enrichment, but they all investigate general crosslinks in host-protein rich environments, such as proteins in nuclear lysates rather than crosslinks within replication compartments [58, 61].

Crosslink enrichment using tagged crosslinkers was also tested for increasing identified crosslinked viral proteins. Briefly, we switched crosslinkers to an azide-tagged disuccinimidyl bisulfoxide (Azide-A-DSBSO) that operates as an enrichable crosslinker [62, 63]. Magnetic beads conjugated with dibenzocycloctyne (DBCO) were used to pull Azide-A-DSBSO linked peptides out of solution through a click chemistry reaction between the DBCO and azide group. Acid treatment cleaves the azide group from the crosslinker, eluting the crosslinked peptides for C18 column cleanup and mass spectrometry analysis. While promising, this method failed to produce new crosslinks not seen with the RIPA method. This may have been due to loss of crosslinked sample after repeated purification steps to rid the sample of residual magnetic beads that could damage the timsTOF mass spectrometer. Additionally, inefficient acid elution would result in unacceptable peptide loss in an already low-concentration sample.

Different crosslinking and mass spectrometry cleanup methods were also tested in an attempt to enhance the number of identified intermolecular crosslinks in Phase II.

There was concern that not enough crosslinker was being applied to PMF samples to sufficiently breach the replication compartment, which is famously structurally stable and nuclease resistant, only being disrupted after protease treatment [4]. A series of experiments testing the efficacy of crosslinking PMFs with 1mM and 5mM DSBU were performed to compare efficiencies to the established 2mM DSBU method. Neither 1mM nor 5mM crosslinking generated more crosslinks than samples prepared with 2mM DSBU, indicating that the concentration of crosslinker was not the cause of low crosslink identification.

We were also concerned with improper purification of the transmembrane proteins, as failure to precipitate transmembrane proteins would also result in the loss of intermolecular crosslinks. We attempted to remedy the issue using SDS to break up the replication compartments after crosslinking. This method has improved yields of transmembrane protein extractions and mass spectrometry identification in previous studies [41]. We treated crosslinked samples with sodium dodecyl sulfate (SDS) during or after crosslinking to either increase permeability for the crosslinker or free transmembrane proteins from the compartment prior to acetone precipitation. Incubation of crosslinked samples in 2% SDS required an extended acetone precipitation step during mass spectrometry sample preparation to remove residual SDS. This resulted in fewer identified peptides, likely due to sample loss from the stringent purification methods my anxiety deemed necessary to protect the column and mass spectrometer. The use of a new column twice the width as the initial column and adjusted flow rates on the timsTOF mass spectrometer resulted in no new crosslink identifications.

With time running out and funding dwindling, we relented and used the crosslinks obtained in the original Phase II method for downstream analysis. More time and resources would be necessary for further optimization of the sample purification process.

3.4.3 Interpreting the XLMS Results

Phase II analysis of the mass spectrometry data used MaxQuant to search the spectra for proteins and crosslinked peptides. The combined .fasta file with a reduced human proteome and the WNV proteome was necessary for searching for host-viral crosslinks in the human HEK 293F infected cells. The computer used for data analysis was unable to effectively search using a full human proteome, and even the reduced proteome crosslink search took upwards of 280 hours for three samples. Analyzing the RIPA supernatant and pellet samples with the human-WNV proteome file, we identified a total of 6,452 loop links, 507 intramolecular crosslinks, and 225 intermolecular crosslinks (**Table 3.2a**). All the identified intermolecular crosslinks were attributed to host-host protein linkages. The lack of identified viral protein crosslinks using this .fasta file was likely due to how MaxQuant assigns protein IDs. If part of a peptide sequence is shared between two separate proteins sequences, the program labels them as razor peptides and assigns them to the more prevalent protein ID. The prevalence of host peptides was much higher compared to viral proteins, resulting in under-labeling of viral peptides as being of WNV origin. Additionally, the viral proteins prM and NS4a had low MaxQuant scores in both samples and thus low confidence in their accurate identification (**Table 3.3**). STRING analysis of the identified host proteins with a MaxQuant score higher than 40 showed the subcellular localization of proteins in both the RIPA supernatant and pellet samples (**Table 3.2b**). The identified observed genes associated with not only the endoplasmic reticulum but also the cytoplasm and nucleus. Crosslinking and purifying the PMF resulted in samples containing other non-ER organelle and cytoplasm-associated proteins that diluted the WNV protein identifications that apparently reduced WNV peptide identification. Returning to the density gradient ultracentrifugation method in future studies would remedy this issue and provide the necessary enrichment of ER microsomes and WNV replication compartments.

3.4.3.1 WNV crosslinks and the transmembrane protein limitation

After identifying proteins and crosslinks with the combined human-WNV .fasta file, we re-analyzed the mass spectrometry data using only the WNV proteome. Several publications provided insight into the timsTOF mass spectrometer data parameters on MaxQuant, although optimization of settings was needed to account for the unique conditions of the replication compartment target environment [45, 60]. It was necessary to adjust the crosslinker profile to symmetrically link all possible targets of DSBU: N-terminal amines, lysines, threonines, serines, and tyrosines. The maximum number of crosslinks per peptide was increased from 1 to 2 and the maximum mass of a peptide was set to 10 kDa. This is because the close proximity of proteins within the tightly packed, nuclease resistant replication compartment raised concern for “daisy-chain” linkages of a string of peptides crosslinked together with several DSBU molecules.

Unfortunately, no parameter settings could improve the intermolecular crosslink yield we observed, with only one intermolecular crosslink between NS3 and NS4b being observed.

The lack of WNV protein intermolecular crosslinks can be explained by the nature of the viral transmembrane domains. The XLMS sample viral protein content skews heavily against transmembrane nonstructural proteins. They are the least represented proteins in both general identification data and the identified crosslink data (**Figure 3.12)**. This disparity can be traced back to the viral transmembrane proteins' base characteristics and overall propensity to resist acetone precipitation [64]. Among the list of transmembrane proteins are the smallest WNV proteins, with the shortest being prM with 92 amino acids. Most of these small protein sequences reside in hydrophilic transmembrane domains meaning they have fewer lysine or arginine residues accessible for cleaving by trypsin. This results in insufficiently digested proteins that were lost in the subsequent purification steps. Transmembrane protein extraction and digestion difficulties ultimately reduced the number of downstream peptides for the MS data acquisition.

The low number of digested and purified transmembrane peptides created a roadblock in assessing the replication compartment PPIs. But the dearth of transmembrane proteins helps to explain the lack of intermolecular crosslinks in the RIPA supernatant and pellet samples. Crosslinking the tightly packed replication compartment would result in NS1, NS3, and NS5 being tethered to the transmembrane nonstructural proteins such as NS2a, NS2b, NS4a, and NS4b, either directly or through a crosslinked mediator protein. *When the transmembrane proteins fail to precipitate with acetone or digest with trypsin sufficiently, there is significant loss of proteins linked to these membrane-bound proteins with intermolecular crosslinks*. Transmembrane protein loss also explains the lack of NS2b-NS3 intermolecular crosslinks, as the interaction between these two proteins as a protease is well characterized and studied for its potential as an antiviral target [65-69]. The loss of transmembrane proteins and their linked proteins would explain the tendency for identified crosslinks to come from NS1, NS3, and NS5. These proteins are more readily available for crosslinking and may be separate from the transmembrane proteins when crosslinked.

Issues concerning the precipitation and digestion of transmembrane proteins in the viral replication compartment warrant more attention in future studies. Exploring new protein precipitation methods outside of acetone precipitation may help purify the protein prior to digestion. Utilizing a panel of digestion enzymes can improve the yield of transmembrane protein peptides. Proteases such as Asp-N, Glu-C, and ProAlanase target different cleavage sites, allowing for more peptides to be formed for mass spectrometry analysis. Improving the yield of transmembrane peptides would likely enhance the observed intermolecular crosslinks needed to identify PPIs and better understand replication compartment component organization.

3.4.3.2 The NS3-NS4b Intermolecular Crosslink

Only one intermolecular crosslink was observed following mass spectrometry analysis of the RIPA pellet and supernatant samples: the NS3 N-terminal amine was crosslinked to the NS4b K172 residue (**Figure 3.16**). This linked protein pair was a surprise considering the lack of crosslinks for transmembrane proteins in our sample runs. K172 on NS4b is located at the end of the cytoplasmic loop between TMD3 and TMD4 [70]. This cytoplasmic loop has been shown to enhance the helicase activity of the NS3 C-terminus during flavivirus replication after mutagenesis and helicase activity studies [23, 28].This data provides *in situ* evidence that the NS3 N-terminus is within twelve angstroms of the predicted beginning of the NS4b TMD4 and supports the notion that the NS4b cytoplasmic loop directly interacts with NS3.

Recent studies have identified interactions between the ER lumen-residing N-terminus of NS4b with NS3 to enhance the latter’s helicase activity [20, 22]. Additionally, a therapeutic in clinical trials called JNJ-1802 targets NS4b on the luminal side of the canonical NS4b structure to interrupt cytoplasmic interactions with NS3 [71]. These results imply the canonical topography of NS4b is flawed, or at the very least subject to change throughout an infection cycle. It has been suggested that portions of NS4b “flip” across the ER membrane to expose the N-terminus to the cytoplasm for NS3 interactions. A similar event occurs to the C-terminus of NS4b when NS5 is cleaved and TMD5 translocates across the membrane to the ER lumen [70, 72]. While the mechanics for these proposed ideas have not been expanded on, the identified crosslinks between the NS3 N-terminus and the cytoplasmic loop of NS4b shows that the NS4b loop maintains its canonical localization in replication compartments. If this portion of NS4b was “flipped” in the replication compartment, this interaction would not have been repeatedly observable in multiple XLMS samples.

While being the only intermolecular crosslink observed between WNV proteins in these samples, the G1-K172 crosslink between NS3 and NS4b provides a strong starting point with *in* situ experimental evidence. This crosslink is from the replication compartment in its natural composition of RNA and protein components. With more identified crosslinks from enhanced transmembrane protein extraction methods, the critical PPIs between NS3 and NS4b can be further characterized for antiviral-driven interruption.

3.4.3.3 Loop-Link and Intramolecular Crosslink Inferences and Ideas

The sparse identified intermolecular crosslinks combined with lengthy DisVis server-side malfunctions meant the DisVis/HADDOCK analysis pipeline validated in Phase I could not be used for the WNV data. Nevertheless, the identified loop and intramolecular crosslinks still allude to protein-protein interactions. As with the BSA control, intramolecular crosslinks can indicate intermolecular interactions between two identical proteins in a multimer.

With less crosslinks to choose from, a finer tooth comb was employed to sort through crosslinks on PyXlinkViewer. The PyXlinkViewer plugin measures the alpha-carbon to alpha-carbon distances, ignoring the distance between the end of side-chain residues and the alpha-carbon of an amino acid. Thus, the distance measured by PyXlinkViewer may vary based on the orientation of the side chain in reference to the other residue's side-chain. When sorting through loop and intermolecular crosslinks, any measured distances between alpha-carbons above 20 Å was investigated further. Using the PyMOL Measurement wizard, the actual distances between the crosslinked residues' side-chain oxygen or nitrogen atoms was measured for a more accurate length. This new length was assessed for violation of the crosslinker distance restraint. Any violated loop or intramolecular crosslinks provided evidence of potential intermolecular interactions in a multimer.

NS1 was one of the few WNV proteins to have identified intramolecular crosslinks, albeit only possessing one unique crosslink pair and two loop-links. All three crosslinks satisfied distance constraints. The NS1 protein primarily associates with the ER membrane surface from inside the lumen, only existing outside of this environment when expelled by the cell in dimeric or hexameric forms [50, 73, 74]. Crosslinking NS1 was complicated due to this localization and the known interactions with purification-resistant transmembrane proteins [75-77]. The sparse NS1 crosslinks do suggest two potential interaction interfaces for NS1 multimers. The published dimeric crystal structure for WNV NS1 shows two V-shaped NS1 monomers assembling into an “X” shape with the two proteins binding at the vertex of the X [50]. All three identified crosslinks localized to the ends of the two “arms” stretching outwards from the vertex of the X-shaped dimer (**Figure 3.13**). If the arms of one NS1 protein were to interact with the corresponding arm of another NS1 protein, a lattice of X-shaped dimers could form. This membrane surface-associating lattice in the ER lumen would provide support for the ballooning of the replication compartment. Further studies are needed to confirm these interaction points, such as crosslinking recombinant NS1 to observe the formation of multimers larger than two proteins on a lipid scaffold *in vitro*.

Most observed crosslinks connected NS3 peptides, with 30 unique loop-links and 7 unique intramolecular crosslinks being identified. All 37 crosslinks were validated by the fine-tooth distance constraint filtering method. Interestingly, when modeled on the AlphaFold2-generated structure with PyXlinkViewer, the crosslinks primarily linked resides on one side of the NS3 structure (**Figure 3.14**). This could be the result of NS3 dimerization along this interface as we observed with the BSA control in Phase I. The lack of crosslinked residues on the opposite side of NS3 is peculiar and warrants further investigation, as it may indicate a lack of accessibility for crosslink attachment due to polar forces from the membrane, other proteins, or the viral genome itself.

MaxQuant analysis revealed 22 unique loop-links and 5 unique intramolecular crosslinks for NS5. Notably, 5 of the loop-links violated the expanded distance restraint validation method (**Figure 3.15b**). As with the BSA experiment in Phase I, this suggests a PPI between multiple NS5 proteins. NS5 is already known to dimerize in certain instances [78], with the implications being discussed in the next session. Unfortunately, long delays with the DisVis server have made it difficult to test different combinations of violated crosslinks and assess their interactions.

3.4.3.4 The Beginning of an Experimentally Validated Replication Compartment Model

Without an adequate number of protein-protein crosslinks, we are unable to build a detailed model of the flavivirus replication compartment. However, we can still postulate a compartment organization structure with the predictions we developed using the observed crosslinks.

Many of the crosslinks observed were loop and intramolecular linkages that have the potential for being mislabeled intermolecular crosslinks. The distant NS5 loop-links that violated the 26.5 Å distance limit established by DSBU crosslinking implies an intermolecular crosslink between two NS5 proteins. Previously, a DENV crystal structure published in Klema et al. shows specific interactions between the N-terminus domain and RdRp domains in the crystalized NS5 proteins [78]. Future NS5 crosslinking results may provide *in situ* evidence supporting this assertion. Validated intramolecular crosslinks may also provide insight into replication compartment organization. The propensity for the observed NS3 crosslinks to exist on one side of the protein also suggests a potential NS3-NS3 multimer interface. Multimers for viral nonstructural enzyme proteins have been confirmed in other viruses that employ organelle compartments to replicate their genomes. Flockhouse virus (FHV) is another positive-sense single-stranded RNA virus that, unlike flaviviruses, generate their protective replication compartments in infected cell mitochondria. In Ertel et al., researchers isolated FHV replication compartments in purified mitochondria and analyzed them with cryo-electron tomography [2]. Using subtomogram averaging, they were able to identify the 3-dimensional proteinaceous pore structure at the mouth of the replication compartment. This protein structure was comprised of twelve FHV protein A monomers arranged in a ring with twelve-fold symmetry. Interestingly, cryo-electron microscopy has revealed the chikungunya virus nonstructural protein 1 (nsP1) also organizes itself into a dodecamer ring when recombinantly expressed and embedded in a membrane [79]. The alphavirus' nsP1 acts as the virus' methyltransferase and guanylyltransferase, while FHV's Protein A has both RNA-dependent RNA polymerase and capping enzyme functions. These proteins are functionally homologous to the flavivirus NS5 and thus the organization of NS5 into a multimeric structure would be within the realm of possibility.

If flaviviruses were to organize their replication compartments in a similar fashion, NS3 and NS5 would need to associate into multimers at the pore opening of the ER replication compartments. The observed crosslink patterns for NS3 and NS5 opens the possibility for multimer organization along those interfaces. This multimeric pore structure with enzymatic activity would have ample access to genomic building blocks and other substrates needed for replication, such as free nucleotides, magnesium ions, and *S*-adenosyl-*L*-methionine (SAM). If the enzymatic pore structure was anchored, it would explain how the vRNA can cycle through the enzymatic machinery efficiently while simultaneously releasing newly capped vRNA into the cytoplasm. The observed crosslink between NS4b and NS3 could provide *in situ* evidence of a compartment anchor for NS3 and NS5 to attach to the ER membrane. NS3 and NS5 routinely coimmunoprecipitate together, although it is unclear as to whether this is a direct interaction or through a mutually-bound mediator [80]. If future crosslinking studies find NS5 is anchored to the pore via interactions with the NS3-NS4b complex, then the foundation of a solid model for the replication complex and compartment organization can be established.

3.4.4 Conclusion and Future Directions

Understanding the intricacies of the flavivirus replication process is necessary for intelligently developing therapeutics against infection. The flavivirus replication compartment is a critical structure that protects the virus from the host immune recognition sentries and houses the nonstructural proteins responsible for generating new viral genomes. Knowledge of the replication compartment's organization and component interactions would bolster therapeutic discovery efforts. To accomplish this ambitious goal of identifying PPIs between the proteins in the compartment, I turned to XLMS.

Ultimately, I was unsuccessful in generating enough crosslinks between the replication compartment proteins to create a model of the compartment and its components. Delays, funding constraints, and research conditions hindered the efficacy of the XLMS method for tackling this goal. Nevertheless, I firmly believe in XLMS as a useful strategy for studying the organization and components of the flavivirus replication compartment *in situ*. The proteins identified in the purified samples and the few identified crosslinks still provided insight into how to approach this goal in the future. With enough time and funding, I would happily dive back into using XLMS to study the replication compartment.

These observed crosslinks are only the beginning and serve merely as a hint in what can be found with the XLMS technique. Future studies can benefit from the time I spent optimizing the sample preparation and purification methodology. To further investigate these potential interaction interfaces, more crosslinks need to be discovered, especially for linkages with transmembrane proteins. Investigating an anchoring model that ties NS1, NS3, and NS5 to the replication compartment through PPIs necessitates a crosslinking and extraction protocol that efficaciously crosslinks, purifies, and digests transmembrane nonstructural proteins. First, a refined purification protocol is necessary for improving the enrichment of crosslinked viral proteins for XLMS analysis. STRING analysis of the crosslinked PMF data demonstrated the wide range of subcellular localizations that identified proteins possessed, and a more stringent sample generation method would remove these proteins from the mass spectrometry data acquisition (**Table 3.2b**). Revising the Phase I methodology would be a good first step. Density gradient centrifugation provided a higher purity sample than the PMF-centric methods employed in Phase II. The gradient fractionation removed extraneous materials from the sample and the high negative-sense to positive-sense vRNA content demonstrated high replication compartment concentration. Overcoming the variation from the meticulous multi-step process would need to be prioritized. Buffer exchanging the fraction of interest followed by crosslinking would provide ample crosslinked viral compartments. The precipitation method prior to reduction, alkylation, and digestion also needs to be revisited. The freeing of crosslinked transmembrane proteins from the membranes of the replication compartment is necessary for generating intermolecular crosslinks and PPI identification. And finally, a robust panel of different proteases are needed to produce every viable peptide from these transmembrane proteins. Only by enhancing the enrichment of the XLMS samples with transmembrane protein peptides can we discover the crosslinks between all components of the replication compartment.

With more loop, intramolecular, and intermolecular crosslinks, rebuilding PPIs in multimers and between the various nonstructural proteins would be achievable. Combined with the rapidly advancing technology in computational modeling of protein structures, an accurate model of the flavivirus replication compartment and its contents can be generated. This model would be invaluable in driving antiviral development and discovery to combat the threat flaviviruses pose to global health.

3.5 REFERENCES

1. Cortese, M., K. Mulder, L. Chatel-Chaix, et al., *Determinants in Nonstructural Protein 4A of Dengue Virus Required for RNA Replication and Replication Organelle Biogenesis.* J Virol, 2021. **95**(21): p. e0131021.

2. Ertel, K.J., D. Benefield, D. Castano-Diez, et al., *Cryo-electron tomography reveals novel features of a viral RNA replication compartment.* Elife, 2017. **6**.

3. Miller, S., S. Kastner, J. Krijnse-Locker, et al., *The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner.* J Biol Chem, 2007. **282**(12): p. 8873-82.

4. Uchil, P.D. and V. Satchidanandam, *Architecture of the flaviviral replication complex. Protease, nuclease, and detergents reveal encasement within double-layered membrane compartments.* J Biol Chem, 2003. **278**(27): p. 24388-98.

5. Uchil, P.D. and V. Satchidanandam, *Characterization of RNA synthesis, replication mechanism, and in vitro RNA-dependent RNA polymerase activity of japanese encephalitis virus.* Virology, 2003. **307**(2): p. 358-371.

6. Overby, A.K., V.L. Popov, M. Niedrig, et al., *Tick-borne encephalitis virus delays interferon induction and hides its double-stranded RNA in intracellular membrane vesicles.* J Virol, 2010. **84**(17): p. 8470-83.

7. Errett, J.S., M.S. Suthar, A. McMillan, et al., *The essential, nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus infection.* J Virol, 2013. **87**(21): p. 11416-25.

8. Chambers, T.J., R.C. Weir, A. Grakoui, et al., *Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein.* Proc Natl Acad Sci U S A, 1990. **87**(22): p. 8898-902.

9. Bartelma, G. and R. Padmanabhan, *Expression, purification, and characterization of the RNA 5'-triphosphatase activity of dengue virus type 2 nonstructural protein 3.* Virology, 2002. **299**(1): p. 122-32.

10. Benarroch, D., B. Selisko, G.A. Locatelli, et al., *The RNA helicase, nucleotide 5'-triphosphatase, and RNA 5'-triphosphatase activities of Dengue virus protein NS3 are Mg2+-dependent and require a functional Walker B motif in the helicase catalytic core.* Virology, 2004. **328**(2): p. 208-18.

11. Luo, D., T. Xu, C. Hunke, et al., *Crystal structure of the NS3 protease-helicase from dengue virus.* J Virol, 2008. **82**(1): p. 173-83.

12. Xu, S., Y. Ci, L. Wang, et al., *Zika virus NS3 is a canonical RNA helicase stimulated by NS5 RNA polymerase.* Nucleic Acids Res, 2019. **47**(16): p. 8693-8707.

13. Geiss, B.J., A.A. Thompson, A.J. Andrews, et al., *Analysis of flavivirus NS5 methyltransferase cap binding.* J Mol Biol, 2009. **385**(5): p. 1643-54.

14. Issur, M., B.J. Geiss, I. Bougie, et al., *The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure.* RNA, 2009. **15**(12): p. 2340-50.

15. Osawa, T., M. Aoki, H. Ehara, et al., *Structures of dengue virus RNA replicase complexes.* Mol Cell, 2023. **83**(15): p. 2781-2791 e4.

16. Teramoto, T., A. Balasubramanian, K.H. Choi, et al., *Serotype-specific interactions among functional domains of dengue virus 2 nonstructural proteins (NS) 5 and NS3 are crucial for viral RNA replication.* J Biol Chem, 2017. **292**(23): p. 9465-9479.

17. Yon, C., T. Teramoto, N. Mueller, et al., *Modulation of the nucleoside triphosphatase/RNA helicase and 5'-RNA triphosphatase activities of Dengue virus type 2 nonstructural protein 3 (NS3) by interaction with NS5, the RNA-dependent RNA polymerase.* J Biol Chem, 2005. **280**(29): p. 27412-9.

18. Zhao, Y., T.S. Soh, S.P. Lim, et al., *Molecular basis for specific viral RNA recognition and 2'-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5).* Proc Natl Acad Sci U S A, 2015. **112**(48): p. 14834-9.

19. Duan, Y., M. Zeng, B. Jiang, et al., *Flavivirus RNA-Dependent RNA Polymerase Interacts with Genome UTRs and Viral Proteins to Facilitate Flavivirus RNA Replication.* Viruses, 2019. **11**(10).

20. Kundharapu, S. and T.K. Chowdary, *Dengue Virus NS4b N-Terminus Disordered Region Interacts with NS3 Helicase C-Terminal Subdomain to Enhance Helicase Activity.* Viruses, 2022. **14**(8).

21. Umareddy, I., A. Chao, A. Sampath, et al., *Dengue virus NS4B interacts with NS3 and dissociates it from single-stranded RNA.* J Gen Virol, 2006. **87**(Pt 9): p. 2605-2614.

22. Lu, H., Y. Zhan, X. Li, et al., *Novel insights into the function of an N-terminal region of DENV2 NS4B for the optimal helicase activity of NS3.* Virus Res, 2021. **295**: p. 198318.

23. Zou, J., T. Lee le, Q.Y. Wang, et al., *Mapping the Interactions between the NS4B and NS3 proteins of dengue virus.* J Virol, 2015. **89**(7): p. 3471-83.

24. Cortese, M., S. Goellner, E.G. Acosta, et al., *Ultrastructural Characterization of Zika Virus Replication Factories.* Cell Rep, 2017. **18**(9): p. 2113-2123.

25. Welsch, S., S. Miller, I. Romero-Brey, et al., *Composition and three-dimensional architecture of the dengue virus replication and assembly sites.* Cell Host Microbe, 2009. **5**(4): p. 365-75.

26. Gillespie, L.K., A. Hoenen, G. Morgan, et al., *The endoplasmic reticulum provides the membrane platform for biogenesis of the flavivirus replication complex.* J Virol, 2010. **84**(20): p. 10438-47.

27. He, Y., G.J. Jensen, and P.J. Bjorkman, *Nanogold as a specific marker for electron cryotomography.* Microsc Microanal, 2009. **15**(3): p. 183-8.

28. Chatel-Chaix, L., W. Fischl, P. Scaturro, et al., *A Combined Genetic-Proteomic Approach Identifies Residues within Dengue Virus NS4B Critical for Interaction with NS3 and Viral Replication.* J Virol, 2015. **89**(14): p. 7170-86.

29. Youn, S., T. Li, B.T. McCune, et al., *Evidence for a genetic and physical interaction between nonstructural proteins NS1 and NS4B that modulates replication of West Nile virus.* J Virol, 2012. **86**(13): p. 7360-71.

30. Roosendaal, J., E.G. Westaway, A. Khromykh, et al., *Regulated cleavages at the West Nile virus NS4A-2K-NS4B junctions play a major role in rearranging*

cytoplasmic membranes and Golgi trafficking of the NS4A protein. J Virol, 2006. **80**(9): p. 4623-32.

31. Li, Y., Y.M. Kim, J. Zou, et al., *Secondary structure and membrane topology of dengue virus NS4B N-terminal 125 amino acids.* Biochim Biophys Acta, 2015. **1848**(12): p. 3150-7.

32. Brand, C., M. Bisaillon, and B.J. Geiss, *Organization of the Flavivirus RNA replicase complex.* Wiley Interdiscip Rev RNA, 2017. **8**(6).

33. Brand, C., B.J. Geiss, and M. Bisaillon, *Deciphering the interaction surface between the West Nile virus NS3 and NS5 proteins.* Access Microbiology, 2024. **6**(6).

34. Gotze, M., C. Iacobucci, C.H. Ihling, et al., *A Simple Cross-Linking/Mass Spectrometry Workflow for Studying System-wide Protein Interactions.* Anal Chem, 2019. **91**(15): p. 10236-10244.

35. Gotze, M., J. Pettelkau, R. Fritzsche, et al., *Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis.* J Am Soc Mass Spectrom, 2015. **26**(1): p. 83-97.

36. Iacobucci, C., M. Gotze, C.H. Ihling, et al., *A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein-protein interactions.* Nat Protoc, 2018. **13**(12): p. 2864-2889.

37. van Zundert, G.C.P., J. Rodrigues, M. Trellet, et al., *The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes.* J Mol Biol, 2016. **428**(4): p. 720-725.

38. Honorato, R.V., P.I. Koukos, B. Jimenez-Garcia, et al., *Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem.* Front Mol Biosci, 2021. **8**: p. 729513.

39. van Zundert, G.C., M. Trellet, J. Schaarschmidt, et al., *The DisVis and PowerFit Web Servers: Explorative and Integrative Modeling of Biomolecular Complexes.* J Mol Biol, 2017. **429**(3): p. 399-407.

40. van Zundert, G.C. and A.M. Bonvin, *DisVis: quantifying and visualizing accessible interaction space of distance-restrained biomolecular complexes.* Bioinformatics, 2015. **31**(19): p. 3222-4.

41. Wisniewski, J.R., A. Zougman, N. Nagaraj, et al., *Universal sample preparation method for proteome analysis.* Nat Methods, 2009. **6**(5): p. 359-62.

42. Kong, A.T., F.V. Leprevost, D.M. Avtonomov, et al., *MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics.* Nat Methods, 2017. **14**(5): p. 513-520.

43. Wessel, D. and U.I. Flugge, *A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids.* Analytical Biochemistry, 1984. **138**: p. 141-143.

44. Scopes, R.K., *Measurement of protein by spectrophotometry at 205 nm.* Analytical Biochemistry, 1974. **59**(1): p. 277-282.

45. Yilmaz, S., F. Busch, N. Nagaraj, et al., *Accurate and Automated High-Coverage Identification of Chemically Cross-Linked Peptides with MaxLynx.* Anal Chem, 2022. **94**(3): p. 1608-1617.

46. Schiffrin, B., S.E. Radford, D.J. Brockwell, et al., *PyXlinkViewer: A flexible tool for visualization of protein chemical crosslinking data within the PyMOL molecular graphics system.* Protein Sci, 2020. **29**(8): p. 1851-1857.

47. Kao, A., C.L. Chiu, D. Vellucci, et al., *Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes.* Mol Cell Proteomics, 2011. **10**(1): p. M110 002212.

48. Hubbard, S. and T. JM, *NACCESS*. 1993: Department of Biochemistry and Molecular Biology, University College London.

49. Szklarczyk, D., R. Kirsch, M. Koutrouli, et al., *The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest.* Nucleic Acids Res, 2023. **51**(D1): p. D638-D646.

50. Akey, D.L., W.C. Brown, S. Dutta, et al., *Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system.* Science, 2014. **343**(6173): p. 881-5.

51. Mirdita, M., K. Schutze, Y. Moriwaki, et al., *ColabFold: making protein folding accessible to all.* Nat Methods, 2022. **19**(6): p. 679-682.

52. Mirdita, M., M. Steinegger, and J. Soding, *MMseqs2 desktop and local web server app for fast, interactive sequence searches.* Bioinformatics, 2019. **35**(16): p. 2856-2858.

53. Mirdita, M., L. von den Driesch, C. Galiez, et al., *Uniclust databases of clustered and deeply annotated protein sequences and alignments.* Nucleic Acids Res, 2017. **45**(D1): p. D170-D176.

54. Mitchell, A.L., A. Almeida, M. Beracochea, et al., *MGnify: the microbiome analysis resource in 2020.* Nucleic Acids Res, 2020. **48**(D1): p. D570-D578.

55. Back, J.W., A.F. Hartog, H.L. Dekker, et al., *A new crosslinker for mass spectrometric analysis of the quaternary structure of protein complexes.* J Am Soc Mass Spectrom, 2001. **12**: p. 222-227.

56. Matzinger, M. and K. Mechtler, *Cleavable Cross-Linkers and Mass Spectrometry for the Ultimate Task of Profiling Protein-Protein Interaction Networks in Vivo.* J Proteome Res, 2021. **20**(1): p. 78-93.

57. Tang, X., G.R. Munske, W. Siems, et al., *Mass Spectrometry Identifiable Cross-Linking Strategy for Studying Protein-Protein Interactions.* Anal. Chem, 2005. **77**: p. 311-318.

58. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.* Nat Protoc, 2016. **11**(12): p. 2301-2319.

59. Muller, M., D. F., I. C., et al., *Cleavable Cross-Linker for Protein Structure Analysis Reliable Identification of Cross-Linking Products by Tandem MS.* Anal. Chem, 2010. **82**: p. 6958–6968.

60. Ihling, C.H., L. Piersimoni, M. Kipping, et al., *Cross-Linking/Mass Spectrometry Combined with Ion Mobility on a timsTOF Pro Instrument for Structural Proteomics.* Anal Chem, 2021. **93**(33): p. 11442-11450.

61. Belsom, A., M. Schneider, L. Fischer, et al., *Blind testing cross-linking/mass spectrometry under the auspices of the 11(th) critical assessment of methods of protein structure prediction (CASP11).* Wellcome Open Res, 2016. **1**: p. 24.

62. Matzinger, M., W. Kandioller, P. Doppler, et al., *Fast and Highly Efficient Affinity Enrichment of Azide-A-DSBSO Cross-Linked Peptides.* J Proteome Res, 2020. **19**(5): p. 2071-2079.

63. Burke, A.M., W. Kandur, E.J. Novitsky, et al., *Synthesis of two new enrichable and MS-cleavable cross-linkers to define protein-protein interactions by mass spectrometry.* Org Biomol Chem, 2015. **13**(17): p. 5030-7.

64. Doucette, A.A., D.B. Vieira, D.J. Orton, et al., *Resolubilization of precipitated intact membrane proteins with cold formic acid for analysis by mass spectrometry.* J Proteome Res, 2014. **13**(12): p. 6001-12.

65. Lin, C., S.M. Amberg, T.J. Chambers, et al., *Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 proteinase is a prerequisite for processing at the downstream 4A4B signalase site.* Journal of Virology, 1993. **67**(4): p. 2327-35.

66. Kumar, A., P. Kumar, M. Aarthy, et al., *Experiments and simulation on ZIKV NS2B-NS3 protease reveal its complex folding.* Virology, 2021. **556**: p. 110-123.

67. Yamshchikov, V.F. and R.W. Compans, *Formation of the flavivirus envelope: role of the viral NS2B-NS3 protease.* Journal of Virology, 1995. **69**: p. 1995-2003.

68. Bera, A.K., R.J. Kuhn, and J.L. Smith, *Functional characterization of cis and trans activity of the Flavivirus NS2B-NS3 protease.* J Biol Chem, 2007. **282**(17): p. 12883-92.

69. Hill, M.E., A. Kumar, J.A. Wells, et al., *The Unique Cofactor Region of Zika Virus NS2B-NS3 Protease Facilitates Cleavage of Key Host Proteins.* ACS Chem Biol, 2018. **13**(9): p. 2398-2405.

70. Miller, S., S. Sparacio, and R. Bartenschlager, *Subcellular localization and membrane topology of the Dengue virus type 2 Non-structural protein 4B.* J Biol Chem, 2006. **281**(13): p. 8854-63.

71. Goethals, O., S.J.F. Kaptein, B. Kesteleyn, et al., *Blocking NS3-NS4B interaction inhibits dengue virus in non-human primates.* Nature, 2023. **615**(7953): p. 678-686.

72. Zou, J., X. Xie, T. Lee le, et al., *Dimerization of flavivirus NS4B protein.* J Virol, 2014. **88**(6): p. 3379-91.

73. Poveda-Cuevas, S.A., C. Etchebest, and F.L. Barroso da Silva, *Self-association features of NS1 proteins from different flaviviruses.* Virus Res, 2022. **318**: p. 198838.

74. Edeling, M.A., M.S. Diamond, and D.H. Fremont, *Structural basis of Flavivirus NS1 assembly and antibody recognition.* Proc Natl Acad Sci U S A, 2014. **111**(11): p. 4285-90.

75. Lindenbach, B.D. and C.M. Rice, *Genetic Interaction of Flavivirus Nonstructural Proteins NS1 and NS4A as a Determinant of Replicase Function.* Journal of Virology, 1999. **73**(6): p. 4611-4621.

76. Lindenbach, B.D. and C.M. Rice, *trans-Complementation of Yellow Fever Virus NS1 Reveals a Role in Early RNA Replication.* Journal of Virology, 1997. **71**(12): p. 9608-9617.

77. Mackenzie, J.M., M.K. Jones, and P.R. Young, *Immunolocalization of the Dengue Virus Nonstructural Glycoprotein NS1 Suggests a Role in Viral RNA Replication.* Virology, 1996: p. 232-240.

78. Klema, V.J., M. Ye, A. Hindupur, et al., *Dengue Virus Nonstructural Protein 5 (NS5) Assembles into a Dimer with a Unique Methyltransferase and Polymerase Interface.* PLoS Pathog, 2016. **12**(2): p. e1005451.

79. Jones, R., G. Bragagnolo, R. Arranz, et al., *Capping pores of alphavirus nsP1 gate membranous viral replication factories.* Nature, 2021. **589**(7843): p. 615-619.

80. Kapoor, M., L. Zhang, M. Ramachandra, et al., *Association between NS3 and NS5 proteins of dengue virus type 2 in the putative RNA replicase is linked to differential phosphorylation of NS5.* J Biol Chem, 1995. **270**(32): p. 19100-6.

CHAPTER 4 — GENERATION AND CHARACTERIZATION OF MOUSE MONOCLONAL ANTIBODIES TARGETING THE SARS-COV-2 NUCLEOCAPSID PROTEIN

4.0 Context

The COVID-19 pandemic struck the United States in the beginning of 2020, throwing a giant bag of wrenches in my research plans while the world worked to handle this health crisis. Labs across CSU transitioned to work on SARS-CoV-2 research or provide services/surveillance to assist public health efforts. My PI Brian Geiss had an idea on how to help: we could develop mouse monoclonal antibodies targeting the SARS-CoV-2 nucleocapsid protein. I jumped into the project, eager to assist in the global effort against SARS-CoV-2 in any way I could. This chapter is adapted from the publication reporting this work [1] in addition to the assistance I provided the Henry Group during their diagnostic device development. While unrelated to the flavivirus work I presented in this dissertation, it is still an important part of my graduate school journey and scientific development.

4.1 Overview

In 2020 the SARS-CoV-2 pathogen spread rapidly across the world and triggered the COVID-19 pandemic. By March 2023 there were 676 million reported cases and 6.9 million deaths attributed to the virus [2]. The relative ease with which the virus transmitted between individuals via respiratory droplet resulted in explosive spread, as asymptomatic patients unwittingly transported the virus to new populations via air travel.

The sudden and explosive spread of SARS-CoV-2 highlighted the dire need for rapid diagnostic testing to help control outbreaks. Like many labs across the world in 2020, the Geiss lab shifted gears, pausing our research into flavivirus replication machinery and instead working towards developing tools to combat this global threat.

In the early days of the pandemic, testing for COVID-19 infection focused primarily on detection of viral genomic RNA present in patient respiratory samples, including nasopharyngeal and nasal swab samples. Because COVID-19 is a respiratory disease, detection of viral genomic RNA in patient nasal samples is a positive indicator of both infection and the potential for an infected individual to spread the virus to others. The only diagnostic for detecting viral genomic RNA in 2020 was quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR), which can sensitively detect the presence of viral RNA in samples and can be automated to test large numbers of samples in parallel [3-5]. This assay provided exquisitely sensitive and specific detection of SARS-CoV-2 infection in the early days of the pandemic but faced challenges. These challenges included the need for a centralized laboratory for testing, significant pre-processing of samples, the high cost of reverse-transcription quantitative PCR reagents, and the need for sophisticated real-time capable thermocyclers for performing the PCR procedure [6]. Fortunately, there are analytes other than RNA that can provide significant clinical value for diagnosing viral infections. We identified the coronavirus nucleocapsid protein one detectable analyte for rapid, specific, and inexpensive diagnostic methods.

Coronavirus RNA genomes are coated with nucleocapsid proteins within virions and infected cells. The nucleocapsid (N) protein is a ~50 kDa protein who's dimers oligomerize onto viral RNA, protecting the viral genome from cellular RNA decay enzymes and compacting the viral genome into a package small enough to fit within virion particles [7-10]. It is estimated that between 720 and 2200 nucleocapsid monomers are present for every viral RNA genome copy within virion particles, making the nucleocapsid protein a potential potent analyte for viral infection [9, 11-15]. Several publications from the original SARS-CoV outbreak in 2003 indicated that detection of nucleocapsid in patient serum samples is diagnostic for early SARS disease, and the amount of detectable SARS-CoV nucleocapsid antigen present in patient samples correlated well with viremia [16-20]. More recent data from the SARS-CoV-2 pandemic indicate that N protein is found in very low but detectable amounts in patient serum [21]. However, N protein has been found in greater amounts in patient nasopharyngeal swabs and anterior nares swab samples [22]. Given the high quantity of the N protein compared to viral genomes and the relative stability of N protein in patient samples, we hypothesized that detection of this antigen could serve as a valuable orthogonal marker in diagnostic assays such as ELISA or point-of-care lateral flow assays compared to genome detection by qRT-PCR.

Detection of protein analytes requires specific antibodies. Early on there were no SARS-CoV-2 specific antibodies reported in the literature. Due to the significant homology between SARS-CoV and SARS-CoV-2, there was a need for new antibodies with increased specificity, utility for detecting SARS-CoV-2 nucleocapsid protein, and potential therapeutic use [23-25]. Additionally, there were no published sequences for SARS-CoV or SARS-CoV-2 specific antibodies. We saw validated and publicly available anti-N antibody sequences as being a good starting point for contributing new molecular tools such as scFv and chimeric antibodies.

With the world shut down and a desire to help the scientific community fight the pandemic, we set out to fulfill this need for a novel monoclonal antibody against the SARS-CoV-2 nucleocapsid protein (NP). What started as an attempt to contribute to the global SARS-CoV-2 research effort expanded into collaborative research in diagnostic and prediction software development. This chapter will discuss the development of several mouse monoclonal cell lines targeting SARS-CoV-2 NP, characterizing these antibodies' molecular qualities and genetic sequences, and the resulting projects that benefited from this work.

The information presented in this chapter is both adapted from and expands upon the work in Terry et al., where the generation and characterization of several monoclonal antibody lines targeting the SARS-CoV-2 N protein were reported [1]. We expressed and purified truncated recombinant N proteins, used the recombinant SARS-CoV-2 antigen to immunize mice and generate a panel of hybridomas, and tested the resulting clones for activity in Western blots, ELISAs, and immunofluorescence assays with N protein from SARS-CoV-2 infected cells. Cross-reactivity of the antibodies against N protein from the five other representative human coronaviruses (SARS-CoV, HuCoV-OC43, HuCoV-HKU1, HuCoV-NL63, and HuCoV-229E) was assessed to determine the antibody line’s viability for specific diagnostic methods. We determined the V_H and V_L sequences of the top 6 antibody clones and performed epitope mapping to identify antigenic regions within the N protein. Thorough investigation into the mBG17 antibody epitope in tandem with predictive software confirmed the antibody’s specific epitope and critical residues. Overall, our data provided strong evidence for the specificity and utility of the SARS-CoV-2 monoclonal antibodies for the study of SARS-CoV-2 NP and COVID-19 diagnostic development.

4.2 Materials and Methods

4.2.1 Expression and Purification of Coronavirus N Proteins

Amino acid sequences for coronavirus N proteins (SARS-CoV-2 (YP_009724397.2), SARS-CoV (ABI96968.1), MERS (YP_009047211.1), HuCoV-NL63 (ABI20791.1), HuCoV-229E (N protein_073556.1), HuCoV-HKU1 (AYN64565.1), HuCoV-OC43 (QBP84763.1) were obtained from GenBank for sequence alignments. For generation of *E. coli* expression plasmids, amino acid sequences from each virus (SARS-CoV-2 (AA133-419), SARS-CoV (AA133-422), MERS-CoV (AA122-413), HuCoV-OC43 (AA147-448), HuCoV-NL63 (AA100-377), HuCoV-HKU1 (AA146-441), and HuCoV-229E (AA102-389), were used to generate bacterial codon optimized DNA gBlocks using the Integrated DNA Technologies web server tool. A list of gBlocks and PCR primers can be found in **Table 4.1**. Each gBlock was cloned into the pET28a T7 expression vector using the New England Biolabs NEBuilder Assembly 2X Master Mix

Table 4.1. DNA gBlocks and oligonucleotide primers

		Protein Expression DNA Oligomers
Oligo	Description	Sequence
BG1464	2019nCoV Coronavirus NP protein (AA133-416) gBlock for NEBuilder assembly of with pET28A	gagatataccatggctagcatgactggtggacagcaaatgggtcgggatccggttgcaactgaaggggccctgaataccccaaaggatcatattggcacacgcaaccctgctaataacgcggccatcgttttacaacttcctcaaggtaccacactgcctaaaggattttacgcagaaggtagtcgtggaggttcacaggctagctcacgctcgtcatcacgttcgcgcaattcctcacgtaattctacgcccggaagctctcgtggtacatcccctgcccgcatggcagggaatggcggggacgccgcccttgctttattgcttcttgatcgtttgaaccaacttgagtcgaagatgtcaggtaaaggacagcagcaacagggtcagacggtcactaagaagtccgccgccgaagcgtcgaagaagccccgccagaaacgtacagcgactaaagcgtataatgtcacacaagcgtttggcgccgcgggcctgagcagactcaagggaattttggggaccaggagttaattcgccaggcaccgattataaacattggcctcagattgcgcagtttgctccgagcgccagcgcattcttcggaatgagccgtattggtatggaggtaaccccatcagggacgtggctgacatatacgggtgctattaagttggatgataaagatcctaactttaaggaccaggtcatcctgttgaacaagcacatcgatgcttataaactttccctcccactgagccgaaaaggataaaaaaagaaagccgatgaaacgcaagcgttgcctcaacgccaaaaaaacaacagacggtgacattactgcctgccgcggatcttgacgatttcagtaaacagcttcaacaaagcatgtcatccgctgattcacttgcggccgcactcgagcaccaccaccaccaccactga
BG1465	pET28a 2019nCoV Coronavirus NEBuilder forward primer (assembly with BG1464)	ctcgagcaccaccaccac
BG1466	pET28a 2019nCoV Coronavirus NEBuilder Reverse primer (assembly with BG1464)	tgctagccatggtatatctccttcttaaagttaaacaaaattatttctagag
BG1514	pET28a Forward for NEBuilder Assembly for fusion between pET28a and 6xHis Tag. Use pBG690 as PCR template with BG1515	cccaccatcaccatcaccactgagatccggctgctaac
BG1515	pET28a Reverse for NEBuilder Assembly for fusion between pET28a and T7 Leader peptide. Use pBG690 as PCR template with BG1514.	ccaccagtcatgctagccatggtatatctccttcttaaagttaaacaaaattatttc
BG1516	gBlock Bacterial codon optimized T7-leader-SARS1 nucleocapsid (ABI96968.1 AA 134-422) - 6xHis	atggctagcatgactggtggacagcaaatgggtcgggatccggttgcaacagagggtgctcttaatactccgaaggatcatattggaactcggaacccaaataataatgcagctacggttctgcaattgccacaaggcactacgctgcctaagggcttttatgccgaaggcagtcgtggtggtagccaggcaagttcccgctctagttcgcgttcacgtgggaactcgcgtaattctactcctggttcctcacgtggtaattctccggctagaatggcttcgggtggtggagagacggcactggccttattattactggatcgccttaatcagttagaaagtaaggtatctggaaaaggccagcagcagcaagggcagactgttacgaagaaaagtgcagccgaggcatcgaaaaaacccagacagaagcggacggcgacaaaacagtataacgtaacacaagcgttcggtcggcgcgggccagaacagacgcaaggcaatttcggtgaccaagaccttattcggcagggaactgactataaacattggccacaaatcgcacagttcgcgccctctgcatctgccttcttcggaatgtcacgcatcggcatggaagtaaccccgtctggcacatggcttacgtaccatggcgcgataaaactggacgataaagacccacaatttaaggacaacgtaatacttctgaacaagcatattgatgcctataaaacctttcgccaaccgagccgaaaaggacaaaaaaagaaaacggacgaggcacaacctctgccccagagacagaagaaacagccgacagtaacgctgttaccggccgcggatatggacgacttctcccgtcaattacagaacagcatgagtggggcctccgcagattccacacaggcccaccatcaccatcaccac
BG1518	gBlock Bacterial codon optimized T7-leader-Human Coronavirus NL63 Nucleocapsid (ABI20791.1 AA 101-377) - 6xHis	atggctagcatgactggtggacagcaaatgggtcgggatccggtagctaaggaaggcgccaagactgtgaacacatcattgggcaacagaaaacggaatcaaaaacctttagagccaaaattttcaattgcattaccgccggaattatccgtggtagagtttgaggaccgcagtaataatagcagtcgggcaagcagtagaagctcgacacgcaacaactcacgtgattcttcacggtctaccagtcgtcaacagagccgtacccgtagtgactccaaccagagcagttccgatctggtagcagcagttacgttagcgttgaaaaacttaggatttgataatcagtcaaagtcgccatcgagttccgggacatcaacgccgaagaaaccaaacaagcctttgtcgcagcctcgtgcggacaaacccagccaactgaagaagccaagatggaaaagagtcccgacccgcgaagagaacgtcattcagtgttttgggccacgggattttaatcacaacatgggggactccgacctggtgcagaatggagtagatgcgaagggctttccacaattggcggagctgatccccaaccaagctgccttattcttcgactcggaagtcagtacagatgaggttggagataatgtacagatcacgtatacctataagatgttagtggcaaaagataacaagaacttacccaaatttatagagcaaatttcggctttcaccaagccatcgtcaatcaaggaaatgcagagtcagtcgtcgcatgtcgtgcagaatacagtgcttaacgctagtatacccgagagtaaaccattagccgatgatgacagtgccatcatagagatcgtcaacgaggtcttacaccatcaccatcaccac
BG1519	gBlock Bacterial codon optimized T7-leader-Human Coronavirus 229E Nucleocapsid (NP_073556.1 AA 103-389) - 6xHis	atggctagcatgactggtggacagcaaatgggtcgggatccggtggccgtcgacggagctaaaactgagccaacgggctatggtgttcgtcgcaagaatagcgagcccgaaatacctcattttaatcaaaagttgccaaacggtgtaacagtggtcgaagaacccgacagtcgggccccttcacgttcacagtcccgtagccaatctcggggtagaggcgagagtaaacgcagtcccgcaacccgtcatccgaccggaatcacaattcgcaggatgatataatgaaagcggtggctgcggcgcttaagtctctgggatttgataagccacaagaaaaggataagaaatcggcgaaaccggaaccccaaaaccttcgcgcaaccagtctccagcgtcttctcaaacatctgcaaagtctttagcacgcagtcagtcgtcagaaaccaaagaacaaaaacacgagatgcaaaagccgcgctggaagcgccagccgaacgatgacgttacaagcaacgttacccaatgcttcggtccccgcgatttggaccacaacttcggaagtgccggcgtggtcgcaaatggtgtaaaggcaaagggatacccacagttcgccgagttggtaccgtctacggcggcgatgttgtttgactcccatattgttagtaaagagtcgggcaacactgttgtattgacattcacaactagagtaacggtgcctaaagaccatccgcatttgggtaaattcttggaggagttgaatgcgtttacccgcgagatgcaacagcaccctttattaaaccctagcgcattagaatttaatccgtcacagacatcacccgccacagcagaacccgttcgggatgaggtgtccatcgagactgatattatagatgaggtaaatcaccatcaccatcaccac
BG1531	SARS-CoV-2 NP Delta 133-183_fwd NEBuilder For deletion of NP AA 133-183)	gcaaatgggttcacaggctagctcacgc
BG1532	SARS-CoV-2 NP Delta 133-183_rev NEBuilder (For deletion of NP AA 133-183)	tagcctgtgaacccatttgctgtccacc
BG1533	SARS-CoV-2 NP Delta 184-223_fwd NEBuilder (For deletion of NP AA 184-223)	tcgtggaggtcttgagtcgaagatgtcag
BG1534	SARS-CoV-2 NP Delta 184-223_rev NEBuilder (for deletion of NP AA 184-223)	tcgactcaagacctccacgactaccttc
BG1535	SARS-CoV-2 NP Delta 224-274_fwd NEBuilder (For deletion of NP AA 224-274)	tttgaaccaagagcagactcaagggaattttg
BG1536	SARS-CoV-2 NP Delta 224-274_rev NEBuilder (For deletion of NP AA 224-274)	gagtctgctcttggttcaaacgatcaagaag
BG1537	SARS-CoV-2 NP Delta 275-320_fwd NEBuilder (For deletion of NP AA 275-320)	ccgcgggcctaccccatcagggacgtggctgac
BG1538	SARS-CoV-2 NP Delta 275-320_rev NEBuilder (For deletion of NP AA 275-320)	ctgatggggtaggcccgcggcgcccaaa
BG1539	SARS-CoV-2 NP Delta 321-371_fwd NEBuilder (For deletion of NP AA 321-371)	tatggaggtacaagcgttgcctcaacgc
BG1540	SARS-CoV-2 NP Delta 321-371_rev NEBuilder (For deletion of NP AA 321-371)	gcaacgcttgtacctccataccaatacggc
BG1541	SARS-CoV-2 NP Delta 372-416_fwd NEBuilder (For deletion of NP AA 372-416)	cgatgaaacgcaccaccaccaccaccac
BG1542	SARS-CoV-2 NP Delta 372-416_rev NEBuilder (For deletion of NP AA 372-416)	ggtggtggtgcgtttcatcggctttcttttttttatcc
BG1506	Template switch oligo (Meyer et al 2019 (PLoS One))	aagcagtggtatcaacgcagagtacatrgrgrg
BG1507	mIGK RT primer (Reverse primer for kappa chain)	ttgtcgttcactgccatcaatc
BG1508	mIGL RT primer (Reverse primer for lambda chain)	ggggtaccatctaccttccag
BG1509	mIGHG TRT Primer (Reverse primer for heavy chain)	agctgggaaggtgtgcacac
BG1510	ISPCR Universial Forward Primer	aagcagtggtatcaacgcagag
BG1511	mIGK PCCR Reverse Primer	acattgatgtctttggggtagaag
BG1512	mIGL Reverse PCR Primer	atcgtacacaccagtgtggc
BG1513	mIGHG Reverse PCR Primer	gggatccagagttccaggtc

(New England Biolabs #E2621) according to the manufacturer’s instructions. The resulting clones (SARS-CoV-2 = pBG690 | SARS-CoV = pBG700 | MERS-CoV = pBG706 | HuCoV-OC43 = pBG701 | HuCoV-NL63 = pBG702 | HuCoV-HKU1 = pBG703 | HuCoV-229E – pBG705) were sequence verified by Sanger sequencing (Genewiz).

Nucleocapsid proteins were expressed in BL21 DE3 pLys *E. coli* and purified by nickel affinity and gel filtration chromatography on a HiLoad 16/600 Superdex 200 pg column (Cytiva #28989335) as previously described [26]with the exception of using 25 mM HEPES (pH 7.5) as the buffer and changing the final NaCl concentration to 500 mM in all of the buffers to reduce N protein oligomerization. N proteins were concentrated to ~2 mg/ml in gel filtration buffer (25 mM Hepes pH 7.5, 500 mM NaCl, 1 mM DTT, 0.1 mM AEBSF, and 10% glycerol (v/v)) using 10 K Centricon concentrators, flash frozen in liquid nitrogen, and stored at −80 °C until use.

4.2.2 N Protein Antibody Generation and Initial Clone Selection

All immunizations were intraperitoneal in a final volume of 200µl containing 25 µg of recombinant N protein antigen. Two 6-week-old female BALB/c mice (Jackson Laboratories) were primed with antigen emulsified in an equal volume of complete Freund's adjuvant (Millipore-Sigma #F5881). The mice were boosted 2 and 4 weeks later with antigen emulsified in incomplete Freund's adjuvant. Mice were bled from the tails at 2 and 6 weeks to verify seroconversion by ELISA. A final boost of antigen in PBS at 8 weeks was followed by euthanasia and spleen retrieval 4 days later. Final antibody titer was determined to be > 6400 at the time of fusion. On the day of splenocyte collection, immunized mice were euthanized by CO_2 asphyxiation followed by exsanguination. Mouse carcasses were sterilized with 70% ethanol before the spleen was removed and eviscerated. After passing through a 16-gauge needle, the splenocytes were washed and fused with Sp2/0 Ag14 myeloma cells using the ClonaCell™-HY Hybridoma Kit (StemCell Technologies #03800). Any ClonaCell™ media referenced was sourced from this kit. Following fusion, hybridoma cells rested for 24 hour before being resuspended and plated in ten 10-cm plates with ClonaCell™-HY Semi-Solid Medium D and allowed to propagate for 10 days at 5% CO_2 at 37 °C. Eleven days after fusion, individual colonies were selected and transferred to individual wells in a 96-well plate using a 10μl micropipette. 920 colonies were harvested and grown for 5 days at 37 °C with 5% CO_2.

4.2.3 Primary Hybridoma Screen

Hybridomas were screened using indirect enzyme-linked immunosorbent assay (ELISA) for activity against SARS-CoV-2 N protein. 96-well plates were coated with 2 ug/ml N protein diluted in 1X PBS and incubated overnight at 4 °C. Plates were blocked with SuperBlock™ T20 (TBS) (Thermo Scientific #37535) for 1 hour shaking at room temperature. After three washes with 0.1% Tween (v/v) in Hyclone 1X PBS (Cytiva #SH30256.02), supernatants from the 920 hybridoma colonies were incubated on the plates for 1 hour with shaking at room temperature. After three additional washes, the plates were incubated with HRP conjugated goat anti-mouse polyclonal antibodies (Abcam ab97023) diluted at 1:10,000 in 1X PBS for 1 hour at room temperature with shaking. The plates were washed three additional times before being developed with 1-Step™ Ultra TMB-ELISA Solution (Thermo Scientific #34029). The reaction was stopped after 4 minutes by adding an equal volume of 2M H_2SO_4. Absorbance at 450 nm (A_{450}) was determined for each well using a PerkinElmer Victor X5 multilabel plate reader. Absorbances were corrected against the PBS negative control and organized by absorbance on Microsoft Excel. Ninety-two colonies with the highest absorbance at 450 nm were selected for further testing.

4.2.4 Bacteria Cross-Reactivity Screen

One 96-well plate was coated with 200ng/well of recombinant N protein and another plate was coated with 200ng/well of *E. coli* lysate (BL21 DE3 pLys) diluted in 1X PBS. The ELISA protocol was performed using the selected 92 parent hybridoma culture supernatants as the primary antibody. Following A_{450} determination and correction in relation to the ClonaCell™ Medium E negative control, any colonies with reactivity towards the bacterial lysate protein were excluded from further screening.

4.2.5 His-tag Cross-Reactivity Screen

96-well plates (Corning 9018) were coated with 200ng/well of recombinant N protein and control plates were coated with 200ng/well of recombinant 6xHis-tagged SARS-CoV-2 spike protein receptor binding domain (RBD) produced using the FreeStyle 293 Expression System (Invitrogen #K900001) from BEI Resources (Cat# NR-52366) as described [27]. The ELISA protocol described previously was performed using the selected 92 parent hybridoma culture supernatants as the primary antibody. Following A_{450} determination and correction in relation to the Medium E negative control, any colonies with reactivity towards the His tags were removed from consideration.

4.2.6 SARS-CoV-2 Virus

SARS-CoV-2 coronavirus (Isolate USA-WA1/2020) was obtained from BEI Resources. Stocks of virus were grown in Vero-E6 cells in DMEM (Gibco #11965092) supplemented with 10% fetal bovine serum (Atlas Biologicals #F-0500-A) and 25 mM HEPES (pH 7.5) in BSL-3 containment at the Colorado State University Regional Containment Biocontainment Laboratory. Virus containing media was stored at −80 °C in single use aliquots. Viral titers were performed using a plaque assay as described [28].

4.2.7 Immunofluorescence Assay

Vero cells were plated at a concentration of 3.5×10^4 cells/well in a 96-well plate. The cells were inoculated with SARS-CoV-2 in BSL-3 at a MOI of 0.1 and allowed to absorb for 1 hour at room temperature. Unabsorbed virus was washed with 1X PBS, cells were overlaid with 1X DMEM supplemented with 2 mM glutamine, non-essential amino acids and 2% fetal bovine serum, and the cells were incubated for 24 hours at 5% CO_2 and 37 °C. Infected cells were briefly washed with 1X PBS then fixed with either 4% paraformaldehyde (w/v) (PFA) or methanol for 10 minutes before being washed three more times. PFA fixed cells were permeabilized by incubating with 0.5% Tween (v/v) in 1X PBS for 20 minutes. Plates of infected fixed cells were blocked with 4% nonfat dry milk powder (w/v) dissolved in 1X PBS + 0.1% Tween 20 (v/v) for 1 hour at room temperature with shaking. Supernatant from the 18 parent clones along with rabbit anti-N protein polyclonal control antibody were then added and incubated for 1 hour with shaking at room temperature. Plates were then washed three times with 1X PBS + 0.1% Tween 20 (v/v) [29]. FITC-conjugated goat anti-mouse secondary antibodies (Abcam #ab6785) or Alexa 488-conjugated goat anti-rabbit secondary antibodies (Abcam #ab150080) at a dilution of 1:2000 in 1X PBS +0.1% Tween 20 (v/v) were added to the appropriate wells. All wells received Hoescht staining at 1:4000. After 1 hour of incubating the secondary antibodies and stains in the dark at room temperature, the plates were washed three times with 1X PBS +0.1% Tween 20 (v/v) with the last wash being 50µl of 1X PBS. Cell labeling and fluorescence were observed with a Celigo high-content imaging cytometer (Nexcelcom) and a Nikon Diaphot 200 fluorescence microscope.

4.2.8 Western Blot

The ability for the initial parental clones to detect linear epitopes was assessed via western blotting. Recombinant N protein and recombinant spike RBD proteins were resolved on a 12% gel in triplicate at 120 V for 1 h before being transferred to PVDF membranes (Sigma-Aldrich #IPVH00010) for 1 h at 100 V. Blots were blocked with 4% nonfat dry milk powder (w/v) blocking solution in 1X PBS. Supernatants were diluted 1:5 in blocking solution (4% non-fat dry milk in 1X PBS+ 0.1% Tween 20 (v/v)) before being applied to blocked blots overnight with shaking at 4 °C. Blots were washed three times in blocking solution for 5 min each before being incubated for 1 h at room temperature with goat anti-mouse HRP-conjugated secondary antibodies diluted in blocking solution. After three additional washes with 1X PBS, the blots were developed with 1-Step™ Ultra TMB-Blotting Solution (Thermo Scientific #37574) before being quenched in deionized water and imaged.

Reactivity towards endogenously produced N protein from an active SARS-CoV-2 infection was assessed with infected cell lysates on western blots. Vero cells were infected at 0.1 MOI with SARS-CoV-2 in BSL-3 and allowed to incubate for 48 h before cells were trypsinized, centrifuged, and resuspended in RIPA buffer (Thermo Scientific #89900). These samples were then diluted in 2X Laemmli Buffer (Bio Rad #1610737) and boiled for 15 min before being resolved on an SDS-PAGE gel alongside uninfected Vero cell lysate and processed for Western blot analysis as described above.

4.2.9 Human Coronavirus Cross-Reactivity Screen

96-well plates were coated with 2 ug/ml of recombinant SARS-CoV-2, SARS-CoV, MERS-CoV, HuCoV-OC43, HuCoV-HKU1, HuCoV-NL63, and HuCoV-229E N protein. These plates were used to test the 18 parent clone antibodies for reactivity using the previously described indirect ELISA protocol. The resulting reactivities were normalized against SARS-CoV-2 reactivity for analysis.

Western blots were performed using the top six mBG clones and the recombinant human coronavirus N proteins to determine the specificity of the antibodies in linearized protein. The previously described Western blot protocol was performed using 10 ug/well of N protein for the mBG67 and mBG86 gels and 1 ug/well of N protein per well for mBG17, mBG21, mBG22, and mBG57 gels due to parent supernatant concentration differences. A control blot from gels loaded with 1 ug/well N protein was labeled using anti-6xHis tag mouse monoclonal antibody (Sigma #SAB2702218).

4.2.10 Isotyping

The isotype of the antibodies produced by parent hybridomas was assessed using the Pierce™ Rapid Antibody Isotyping Kit plus Kappa and Lambda (Pierce #26179) for mouse antibodies according to the provided procedure. Briefly, antibodies were diluted 1:100 in the provided sample diluent before being applied to the lateral flow assay and the corresponding bands observed after 10 min.

4.2.11 Sequencing of Anti-N Protein Monoclonal Antibody Genes

RNA samples from hybridomas 17, 21, 22, 57, 67, and 86 were isolated and stored at −80 °C following Trizol (Invitrogen #15596026) and phenol-chloroform extraction. Isolation of kappa, lambda, and heavy chain RNA sequences for hybridoma colonies was accomplished using a previously reported monoclonal antibody sequencing protocol [30]. Primers specific for the 3' constant region of either the kappa, lambda, or heavy chain RNA sequences outlined in the protocol were used in conjunction with SMARTScribe Reverse Transcriptase (Takara Bio #639538). Amplification of the RT product was accomplished with PCR using primers specific to the universal sequence and a region of the kappa, lambda, or heavy chain sequences that were offset to the primers used during reverse transcription. Successfully produced PCR products were then cloned into the NEB pMiniT 2.0 vector using the NEB PCR Cloning Kit (New England Biolabs #E1202S). Five kappa and heavy chain clones were sequenced for each hybridoma by Genewiz. Clone sequences were analyzed by IgBLAST for antibody framework regions (FMW) and complementary-determining regions (CDR) [31].

4.2.12 Generation of Recombinant Anti-N Protein Monoclonal Antibodies

The mBG17 and mBG86 heavy chain and light chain regions identified by sequencing were selected for cloning into eukaryotic expression vectors. The heavy chain and kappa regions of the antibody sequences were extracted from the NEB pMiniT 2.0 vectors by PCR. The PCR products were then inserted using the NEBuilder

Assembly Kit into heavy chain and light chain expression vectors provided by Dr. Jennifer Deluca at Colorado State University [32]. Following confirmation of successful gene insertion by agarose gel analysis, vectors were transfected into 293F cells using the Freestyle 293 Expression System (Invitrogen K900001). Supernatant was collected and validated for activity against SARS-CoV-2 NP by ELISA.

4.2.13 Virion-derived N Protein Detection by Sandwich ELISA

SARS-CoV-2 virion particles were collected from infected Vero cell supernatant at 5.5×10^5 PFU/ml and inactivated with IGEPAL CA-630 (Sigma-Aldritch #542334) (0.1% (v/v)) for 30 minutes at 4 °C. Inactivated viral stocks were serially diluted from 1:10 to 1:10,000 in 1X PBS. 96 well plates (Corning #9018) were coated with 0.1 ug/ml of anti-N protein rabbit polyclonal antibody overnight and then blocked with 4% (w/v) dry milk powder blocking solution dissolved in 1X PBS + 0.1% Tween 20 (v/v) for 1 h at room temperature. Following three washes with 1X PBS +0.1% Tween 20 (v/v), the diluted stocks of inactivated SARS-CoV-2 were applied to the capture antibody-coated plates in 50µl triplicate aliquots and incubated at 37 °C for 90 minutes. No N protein was applied to control wells with only diluent being applied to control for false positive reporting of monoclonal antibody reactivity. The plate was washed three additional times, and 50µl of mBG17 or mBG86 parent supernatant was applied to each of the deactivated virus dilutions in triplicate and then incubated for 1 h at room temperature. Following this, the remaining steps were derived from the indirect ELISA protocol described earlier. Reported signal was corrected for using the wells in which no inactivated virus was applied.

4.2.14 General Epitope Mapping

Primers described in **Table 4.1** were used to create 50 amino acid deletions from the pBG690 plasmid used to produce SARS-CoV-2 N protein amino acids 133–419. The primers produced recombinant N protein variant sequences corresponding to proteins Δ133–179, Δ180–229, Δ230–279, Δ325–379, and Δ381–419. These PCR products were circularized using the NEBuilder HiFi DNA Assemby kit (New England Biolabs #E5520), cloned into DH5-alpha *E. coli*, and recombinant E. coli clones sequenced. For protein expression, verified plasmids were transformed into BL21 *E. coli* cells with kanamycin and chloramphenicol selection. Cultures were incubated overnight in LB broth with 10% glucose (w/v), and recombinant N protein production was induced with 0.4 mM IPTG alongside uninduced transformed variants for 5 hours. Cells were centrifuged, resuspended in Laemmli sample buffer, boiled for 5 minutes, and then resolved on a 10% SDS-PAGE gel. Protein lysates were then transferred to a PVDF membrane for Western blot analysis using the indicated monoclonal antibody according to the aforementioned Western blot protocol.

4.2.15 Specific Epitope Mapping by Peptide Competition ELISA

The anti-SARS-CoV-2 nucleocapsid protein mBG17 antibody epitope was further characterized using competition enzyme-linked immunosorbent assay (ELISA). Using the previously determined 39 nucleocapsid protein amino acid range for the mBG17 epitope as a starting point, seven overlapping peptides were synthesized (Thermo Scientific) spanning the 39 amino acid regions with overlaps of 5 amino acids. These peptides were termed Fragment 1 through 7. A 96-well ELISA plate was coated with 0.1µg/ml of recombinant SARS-CoV-2 NP (36) overnight at 4oC. The plate was blocked with 4% (w/v) dry non-fat milk in 1X PBS with 0.1% (v/v) Tween-20 for 1 hour with shaking at room temperature. While blocking, recombinant mBG17 antibody samples were inhibited by incubating 40µL of antibody with 40 µg (approximately 30 nmol) of a single peptide fragment for one hour at room temperature. Following this, peptide-incubated mBG17 was applied to the blocked nucleocapsid protein coated plate in triplicate and allowed to incubate for 1 hour at room temperature while shaking. The plate was rinsed with 0.1% (v/v) Tween-20 in 1X PBS and washed three more times for 5 minutes shaking at room temperature. The plate was then incubated with HRP-conjugated goat anti-mouse polyclonal antibody solution diluted at 1:20,000 in 1X PBS for 1 hour shaking at room temperature. After another rinse and three more washes the plate was developed with 1-Step™ Ultra TMB-ELISA Solution (Thermo Scientific #34029) before stopping the reaction with an equal volume of 2M H2SO4. Solution absorbance at 450 nm was measured using a PerkinElmer Victor X5 multilabel plate reader. Absorbances were averaged within fragment-inhibited sample groups and corrected with the average value of the negative control. These absorbances were then normalized against the absorbance from the group with the highest value before multiplying by 100 to obtain percentage of potential signal.

The effect of single alanine substitutions on fragment 5 (DDFSKQLQQS) peptide binding was determined by competition ELISA using a series of ten alanine-substituted peptides (Table 4.1) at a range of concentrations to determine relative competition activity. A modified version of the previously described inhibition ELISA was performed using the unmodified Fragment 5 peptide and the ten alanine-substituted peptides. During the mBG17 inhibition step, the mBG17 antibody solution was incubated with a 4-fold serial dilution of peptides beginning at 40μg and continuing to ~2.5 ng before being applied to the NP coated plates in triplicate. The remainder of the competition ELISA was carried out as described above.

4.3 Results

4.3.1 Generation and Characterization of Anti-SARS-CoV-2 Nucleocapsid Antibodies

This project began with developing novel antibodies that can specifically detect SARS-CoV-2 nucleocapsid protein for use in research and diagnostic testing efforts. Recombinant SARS-CoV-2 N protein needed to be produced as antigen for hybridoma production, but certain cross-reactivity concerns needed to be addressed. The N-terminal domain of human coronavirus N proteins contain several conserved regions that may contribute to monoclonal cross-reactivity (**Figure 4.1a**). Previous work

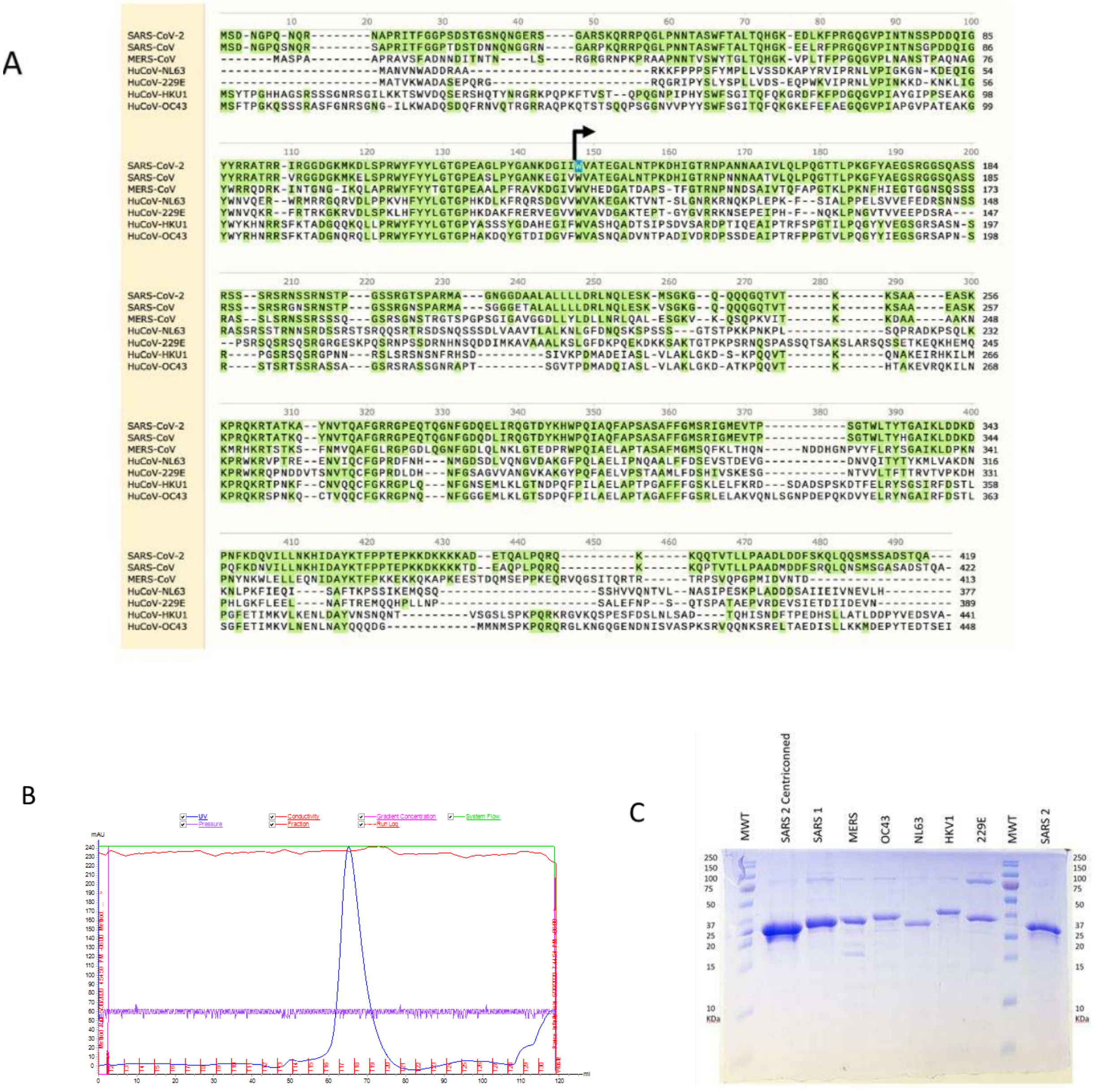
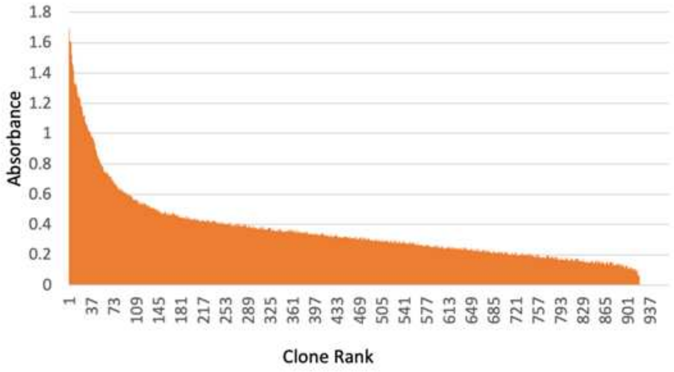


Figure 4.1. Production and characterization of truncated SARS-CoV-2 nucleocapsid protein. A) Sequence alignment of human coronavirus N proteins. **B)** Size exclusion chromatograph of nickel-column purified SARS-CoV-2 NP (133-419). **C)** SDS-PAGE gel of purified Coronavirus nucleocapsid proteins.

developing monoclonal antibodies against MERS coronavirus N protein found that removal of the N-terminal domain improved antibody [33]. Therefore, we developed a bacterial expression plasmid that produces AA133-419 of the SARS-CoV-2 (isolate USA-WA1/2020) N protein with an N-terminal T7 leader sequence to improve translation efficiency (pBG690). Expression of recombinant N protein in BL21 DE3 pLys *E. coli* was relatively robust, but we found that the recombinant protein formed large molecular weight oligomers at NaCl concentrations below 300 mM (data not shown). Purification of nickel-affinity purified N protein on a Superdex 200 gel filtration column in 500 mM NaCl resolved these aggregation issues and the resulting protein migrated as the predicted molecular weight of a truncated N dimer (**Figure 4.1b**). Recombinant SARS-CoV, MERS-CoV, HuCoV-OC43, HuCoV-NL63, HuCoV-HKU1 and HuCoV-229E N protein were produced using the same process for downstream cross-reactivity screening. SDS-PAGE analysis of purified coronavirus N proteins showed that the proteins were >98% pure and migrated at the expected molecular weight (**Figure 4.1c**). This purified SARS-CoV-2 NP protein was found to be antigenic and produced rabbit polyclonal antibodies that showed affinity for NP produced in deer mouse SARS-CoV-2 infections in parallel work [29]. The recombinant NP was used to immunize BALB/c mice according to the protocol described later to generate immunized mice, from which spleens were collected and hybridoma clones were produced.

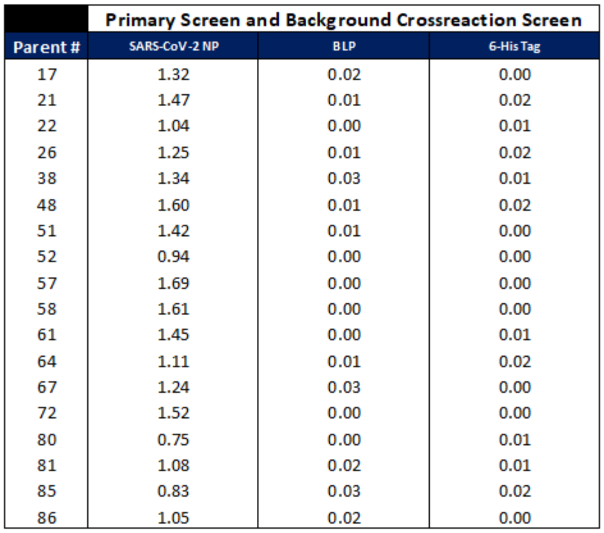
For the primary screening of hybridoma colonies, supernatant was collected from 920 colonies and tested for the presence of N protein specific antibodies by ELISA (**Figure 4.2a**). The highest scoring 92 hybridoma colonies were selected for further testing

A



920 clones tested
Top 18 clones characterized

B



	Primary Screen and Background Crossreaction Screen		
Parent #	**SARS-CoV-2 NP**	**BLP**	**6-His Tag**
17	1.32	0.02	0.00
21	1.47	0.01	0.02
22	1.04	0.00	0.01
26	1.25	0.01	0.02
38	1.34	0.03	0.01
48	1.60	0.01	0.02
51	1.42	0.01	0.00
52	0.94	0.00	0.00
57	1.69	0.00	0.00
58	1.61	0.00	0.00
61	1.45	0.00	0.01
64	1.11	0.01	0.02
67	1.24	0.03	0.00
72	1.52	0.00	0.00
80	0.75	0.00	0.01
81	1.08	0.02	0.01
85	0.83	0.03	0.02
86	1.05	0.02	0.00

Figure 4.2. Screening of Anti-Nucleocapsid Clones. A) Direct ELISA analysis of 920 clones picked from hybridoma fusion. **B)** Verification ELISA of top 18 anti-nucleocapsid monoclonal antibody clones and counter screening against bacterial lysate (BLP) and 6-His-tagged SARS-CoV-2 Spike RBD domain. Averages are presented following background subtraction.

and expansion as “parent” colonies. Due to the bacterial system used for N protein immunogen production, we then screened these 92 parents for reactivity against bacterial proteins. Plates were coated with bacterial lysate protein (BLP) and tested by ELISA in parallel to N protein-coated plates to determine which parents had residual bacterial reactivity. Of these, parents 5,7, 23, 43,53, 54, 65, 77, and 79 demonstrated specificity for bacterial proteins and were thus eliminated from further testing. The remaining parents were then tested for reactivity against the 6x-His tag using 6x-His tagged soluble spike receptor binding domain as the counter screening antigen. Parent 55 showed reactivity to the 6xHis tag and was thus eliminated from further testing. 18 top-performing parents were selected for further testing and given the prefix mBG (**Figure 4.2b**).

Western blots were performed to determine reactivity of the selected 18 parents towards linear or semi-linear epitopes. Western blots comparing reactivity towards SARS-CoV-2 infected Vero cell lysates and uninfected Vero cell lysates (**Figure 4.3**) demonstrated a high specificity towards antigen in infected cells and the endogenously produced N protein for the top 18 clones. To test the capability of these antibodies for identifying subcellular localization of NP in SARS-CoV-2 infected Vero cells, the ability for the top 18 parents to bind to endogenous N protein following PFA or methanol fixation was determined by immunofluorescence assay (IFA). IFA positive clones varied in effectiveness between methanol and paraformaldehyde fixation methods (**Figure 4.4a**). Results indicated that the parent antibodies that demonstrated reactivity to fixed

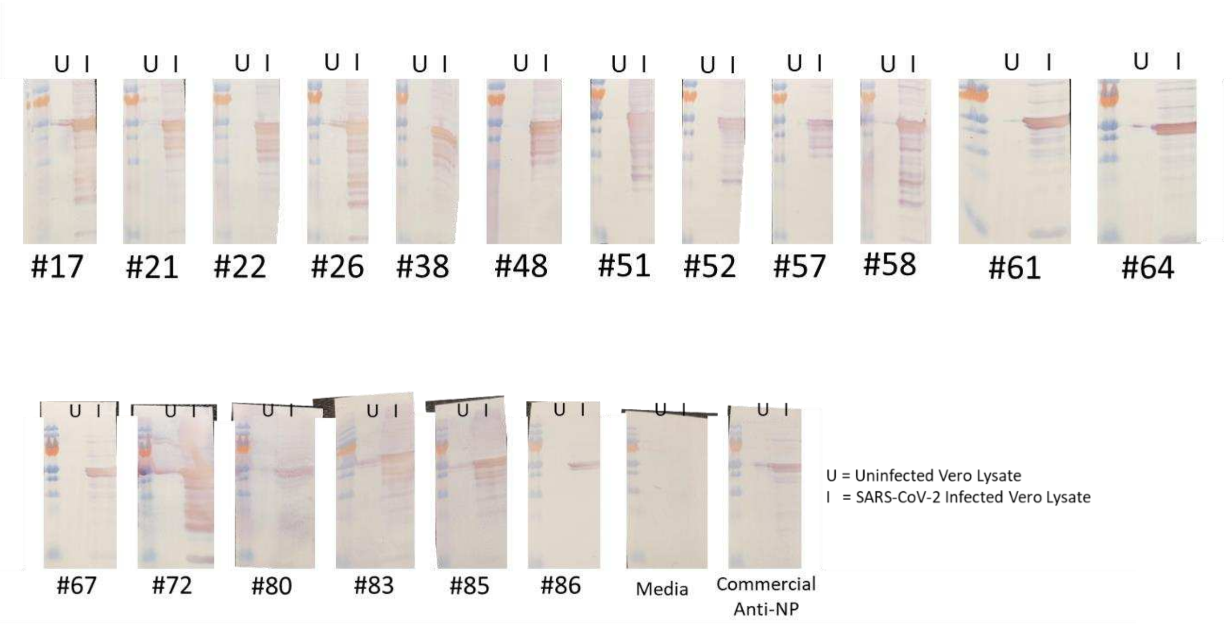
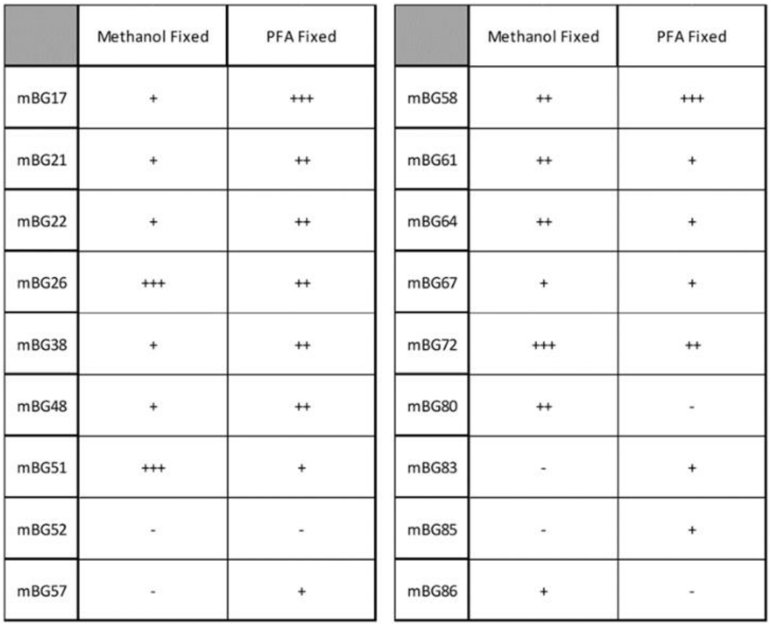


Figure 4.3. Western Blot Analysis of Anti-Nucleocapsid Monoclonal Antibody Clones Against Uninfected (U) or SARS-CoV-2 infected (I) Vero cells.

A



	Methanol Fixed	PFA Fixed		Methanol Fixed	PFA Fixed
mBG17	+	+++	mBG58	++	+++
mBG21	+	++	mBG61	++	+
mBG22	+	++	mBG64	++	+
mBG26	+++	++	mBG67	+	+
mBG38	+	++	mBG72	+++	++
mBG48	+	++	mBG80	++	-
mBG51	+++	+	mBG83	-	+
mBG52	-	-	mBG85	-	+
mBG57	-	+	mBG86	+	-

B

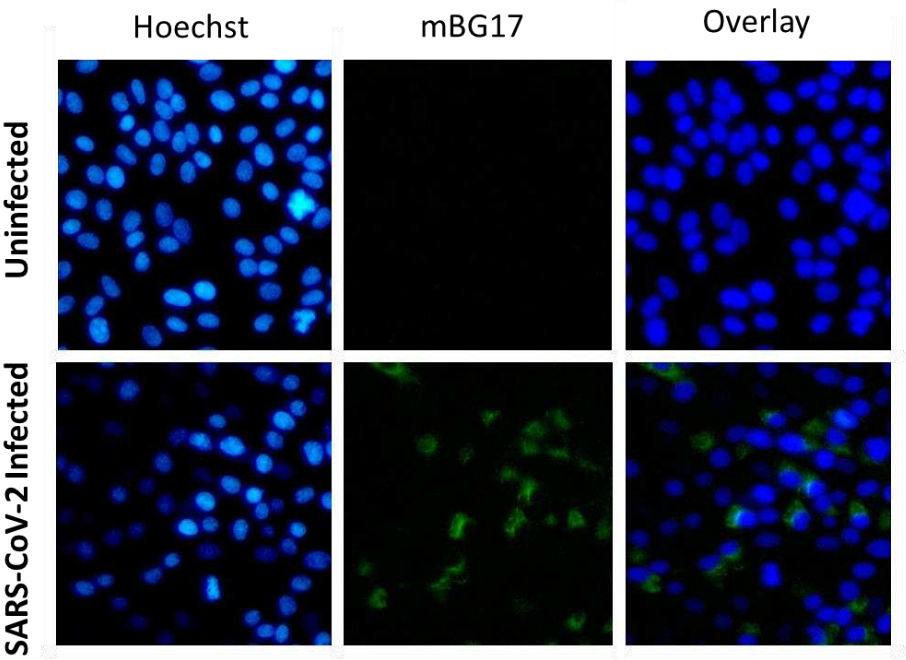


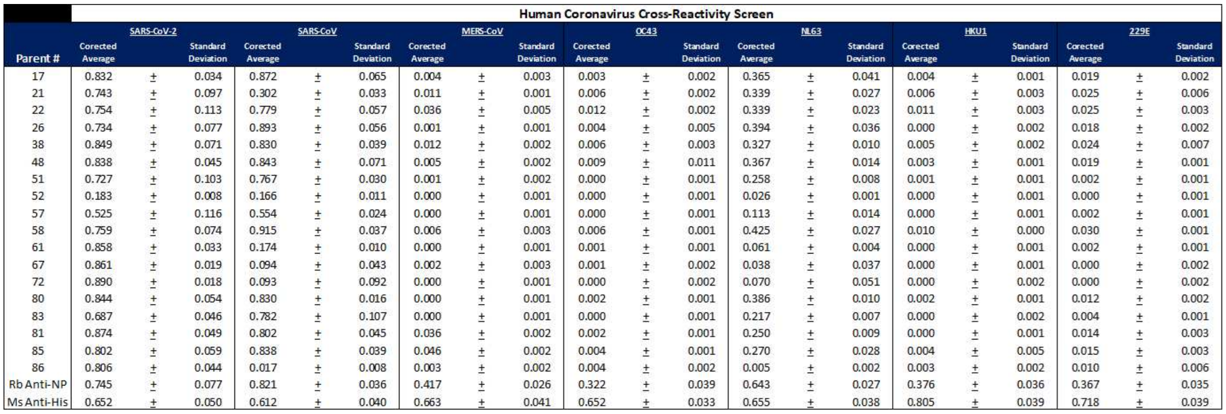
Figure 4.4. Immunofluorescence Analysis of Anti-Nucleocapsid Monoclonal Antibody Clones. A) Relative reactivities of clones in SARS-CoV-2 infected Vero cells fixed with methanol or paraformaldehyde. **B)** representative images of uninfected and SARS-CoV-2 infected Vero cells (paraformaldehyde fixed) processed for immunofluorescence analysis with mBG17.

infected cells were non-reactive towards fixed uninfected cells. N protein was found predominantly SARS-CoV-2 infected cell cytoplasm as expected (**Figure 4.4b**).

The top 18 parents were tested for cross-reactivity against the six other human coronaviruses known to be represented in the human population. This assessment was performed to determine which antibodies were only reactive towards SARS-CoV-2 N protein. The recombinant SARS-CoV, MERS-CoV, HuCoV-OC43, HuCoV-NL63, HuCoV-HKU1 and HuCoV-229E N proteins were expressed and purified for determining the top 18 parents' specificity compared to SARS-CoV-2 N protein by indirect ELISA (**Table 4.2**). With the exception of mBG86 there was strong cross reactivity with SARS-CoV N protein. Similarly, mBG86 lacked cross reactivity with NL63 N protein while other parents maintained some reactivity, albeit less than they demonstrated for SARS-CoV. Additionally, mBG 21, 22, 57, and 67 all had reduced reactivity towards NL63 N protein. The anti-His control showed equal reactivity towards all N proteins tested. From this Indirect ELISA, we determined that mBG86 was the only clone specific for SARS-CoV-2 N proteins.

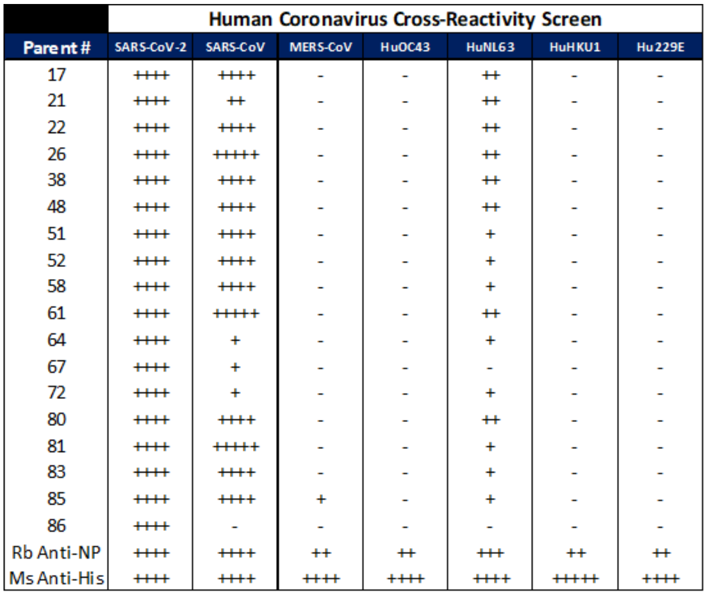
Based on the activity of clones in ELISA, Western blot, and IFA analysis, we chose six colonies? (mBG17, mBG21, mBG22, mBG57, mBG67, and mBG86) for sequence analysis using the method outlined in [30]. Five separate clones for each V_H or V_L were sequenced, and consensus sequences for each clone were assembled, translated, and subjected to IgBLAST analysis to determine framework and

Table 4.2. Cross-Reactivity Screening of Anti-Nucleocapsid Monoclonal Antibody Clones. Top: Direct ELISA analysis of top 18 clones against SARS-CoV2, SARS-CoV, MERS-CoV, OC43, NL63, HKU1, and 229E human coronavirus nucleocapsid proteins. Average signals are corrected against background signal Bottom: Relative binding of monoclonal antibodies to nucleocapsid proteins compared to SARS-CoV-2.



Human Coronavirus Cross-Reactivity Screen

	SARS-CoV-2			SARS-CoV			MERS-CoV			OC43			NL63			HKU1			229E		
Parent #	Corected Average		Standard Deviation	Corected Average		Standard Deviation	Corected Average		Standard Deviation	Corected Average		Standard Deviation	Corected Average		Standard Deviation	Corected Average		Standard Deviation	Corected Average		Standard Deviation
17	0.832	±	0.034	0.872	±	0.065	0.004	±	0.003	0.003	±	0.002	0.365	±	0.041	0.004	±	0.001	0.019	±	0.002
21	0.743	±	0.097	0.302	±	0.033	0.011	±	0.001	0.006	±	0.002	0.339	±	0.027	0.006	±	0.003	0.025	±	0.006
22	0.754	±	0.113	0.779	±	0.057	0.036	±	0.005	0.012	±	0.002	0.339	±	0.023	0.011	±	0.003	0.025	±	0.003
26	0.734	±	0.077	0.893	±	0.056	0.001	±	0.001	0.004	±	0.005	0.394	±	0.036	0.000	±	0.002	0.018	±	0.002
38	0.849	±	0.071	0.830	±	0.039	0.012	±	0.002	0.006	±	0.003	0.327	±	0.010	0.005	±	0.002	0.024	±	0.007
48	0.838	±	0.045	0.843	±	0.071	0.005	±	0.002	0.009	±	0.011	0.367	±	0.014	0.003	±	0.001	0.019	±	0.001
51	0.727	±	0.103	0.767	±	0.030	0.001	±	0.002	0.000	±	0.001	0.258	±	0.008	0.001	±	0.001	0.002	±	0.001
52	0.183	±	0.008	0.166	±	0.011	0.000	±	0.001	0.000	±	0.001	0.026	±	0.001	0.000	±	0.001	0.000	±	0.001
57	0.525	±	0.116	0.554	±	0.024	0.000	±	0.001	0.000	±	0.001	0.113	±	0.014	0.000	±	0.001	0.002	±	0.001
58	0.759	±	0.074	0.915	±	0.037	0.006	±	0.003	0.006	±	0.001	0.425	±	0.027	0.010	±	0.000	0.030	±	0.001
61	0.858	±	0.033	0.174	±	0.010	0.000	±	0.001	0.001	±	0.001	0.061	±	0.004	0.000	±	0.001	0.002	±	0.001
67	0.861	±	0.019	0.094	±	0.043	0.002	±	0.003	0.001	±	0.002	0.038	±	0.037	0.000	±	0.001	0.000	±	0.002
72	0.890	±	0.018	0.093	±	0.092	0.000	±	0.001	0.000	±	0.002	0.070	±	0.051	0.000	±	0.002	0.000	±	0.002
80	0.844	±	0.054	0.830	±	0.016	0.000	±	0.001	0.002	±	0.001	0.386	±	0.010	0.002	±	0.001	0.012	±	0.002
83	0.687	±	0.046	0.782	±	0.107	0.000	±	0.001	0.000	±	0.001	0.217	±	0.007	0.000	±	0.002	0.004	±	0.001
81	0.874	±	0.049	0.802	±	0.045	0.036	±	0.002	0.002	±	0.001	0.250	±	0.009	0.000	±	0.001	0.014	±	0.003
85	0.802	±	0.059	0.838	±	0.039	0.046	±	0.002	0.004	±	0.001	0.270	±	0.028	0.004	±	0.005	0.015	±	0.003
86	0.806	±	0.044	0.017	±	0.008	0.003	±	0.002	0.004	±	0.002	0.005	±	0.002	0.003	±	0.002	0.010	±	0.006
Rb Anti-NP	0.745	±	0.077	0.821	±	0.036	0.417	±	0.026	0.322	±	0.039	0.643	±	0.027	0.376	±	0.036	0.367	±	0.035
Ms Anti-His	0.652	±	0.050	0.612	±	0.040	0.663	±	0.041	0.652	±	0.033	0.655	±	0.038	0.805	±	0.039	0.718	±	0.039



Human Coronavirus Cross-Reactivity Screen

Parent #	SARS-CoV-2	SARS-CoV	MERS-CoV	HuOC43	HuNL63	HuHKU1	Hu229E
17	++++	++++	-	-	++	-	-
21	++++	++	-	-	++	-	-
22	++++	++++	-	-	++	-	-
26	++++	+++++	-	-	++	-	-
38	++++	++++	-	-	++	-	-
48	++++	++++	-	-	++	-	-
51	++++	++++	-	-	+	-	-
52	++++	++++	-	-	+	-	-
58	++++	++++	-	-	+	-	-
61	++++	+++++	-	-	++	-	-
64	++++	+	-	-	+	-	-
67	++++	+	-	-	-	-	-
72	++++	+	-	-	+	-	-
80	++++	++++	-	-	++	-	-
81	++++	+++++	-	-	+	-	-
83	++++	++++	-	-	+	-	-
85	++++	++++	+	-	+	-	-
86	++++	-	-	-	-	-	-
Rb Anti-NP	++++	++++	++	++	+++	++	++
Ms Anti-His	++++	++++	++++	++++	++++	+++++	++++

complementary determining regions (**Table 4.3**). mBG21 and 22 presented identical sequences, suggesting a common B-cell origin. Isotype tests revealed the isotypes for each of the five parents, with the identical isotypes of mBG21 and 22 supporting the prediction of a common origin. Clones mBG21/22, and mBG57 shared identical heavy chains but different kappa chains. Clone mBG17 possessed unique heavy and light chains. mBG67 and mBG86 possessed identical kappa light chains and heavy chains separated by three mutations in the heavy chain framework 3. These mutations may be the source of the improved specificity of mBG86 compared mBG67 observed in **Table 4.2** and merits further study.

With the selected top six hybridoma clones we screened for Western blot cross-reactivity against the heterologous recombinant coronavirus proteins (**Figure 4.5**). The results largely mirrored the indirect ELISA data, with mBG17, 21, 22, and 57 possessing cross reactivity with SARS-CoV and NL63 N protein. mBG67 and mBG86 produced weak additional bands against SARS-CoV N protein that indicate slight reactivity towards the linearized SARS-CoV protein in western blots. Of note, western blots for mBG67 and mBG86 needed more loaded protein sample for bands to be detected than mBG17, 21, 22, and 57. This can be traced back to the propensity for mBG67 and 86 targeting a linear epitope compared to a conformational epitope that provides more epitopes for antibody binding when linearized on a western blot. Overall, these results support that mBG67 and mBG86 appear to have a significant preference for conformational epitopes whereas clones mBG17, 21, 22, and 57 appear to have a strong preference towards linear epitopes.

Table 4.3. Immunoglobulin heavy (A) and light chain (B) amino acid sequences of top 5 hybridoma clones as defined by IgBlast following hybridoma RNA extraction, template switch RT-PCR, bacterial cloning, and sequencing. Key three amino acid mutations between the heavy chains of mBG67 and mBG86 are highlighted in red.

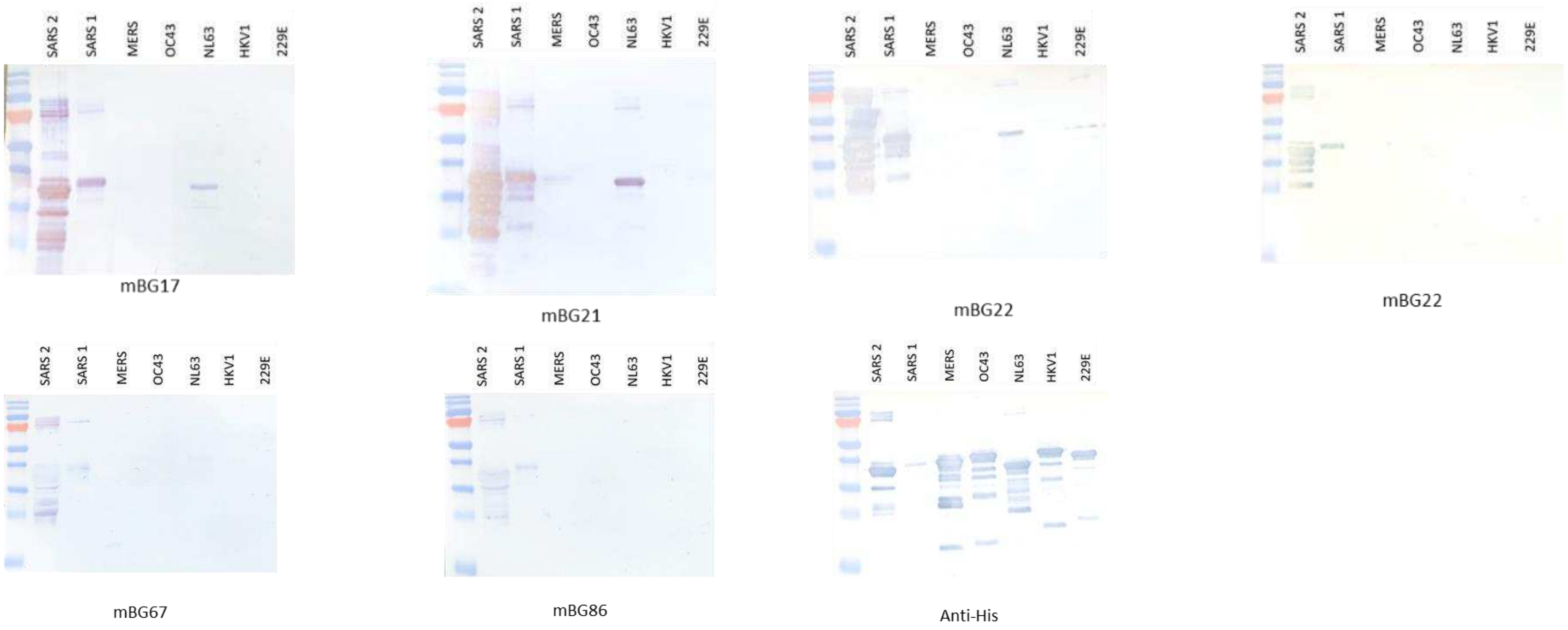
A

Clone	Isotype	Chain	FRW-1	CDR-1	FRW-2	CDR-2	FRW-3	CDR-3	FRW-4
17	IgG1	Heavy	EVKLEESGGGLVQPGGSMKFSCVAS	GFTFSDYW	MNWVRQSPDKGLEWVAE	IRLKSNNYAT	HYAASVKGRFTISRDDSKSSVYLQMNNLRAEDSGIYYC	TRSAMDY	WGQGTSVTVSS
21	IgG2b	Heavy	QIQLVQSGPELKKPGETVKISCKAS	GYTFTDYS	MHWVKQAPGKGSKWMGW	INTETGEP	TYADDFKGRFAFSLETSASTAYLQINNLKNEDTATYFC	ALRR	WGQGTLVTVS
22	IgG2b	Heavy	QIQLVQSGPELKKPGETVKISCKAS	GYTFTDYS	MHWVKQAPGKGSKWMGW	INTETGEP	TYADDFKGRFAFSLETSASTAYLQINNLKNEDTATYFC	ALRR	WGQGTLVTVSA
57	IgG1	Heavy	QIQLVQSGPELKKPGETVKISCKAS	GYTFTDYS	MHWVKQAPGKGSKWMGW	INTETGEP	TYADDFKGRFAFSLETSASTAYLQINNLKNEDTATYFC	ALRR	WGQGTLVTVSA
67	IgG1	Heavy	EVQLVESGGGLVQPGGSLKLSCAAS	GFTFSNYG	MSWVRQTPDKRLELVAT	INRNGGST	YYLDSVKVRFTISRDNAKSTLFLQLSSLKSDDTAMYYC	ARIYDFDEDYFDV	WGAGTTVTVSS
86	IgG1	Heavy	EVQLVESGGGLVQPGGSLKLSCAAS	GFTFSNYG	MSWVRQTPDKRLELVAT	INRNGGST	YYLDSVKVRFTISRDNAKSTLFLQMSSLKSEDTAMYYC	ARIYDYDEDYFDV	WGAGTTVTVSS

B

Clone	Isotype	Chain	FRW-1	CDR-1	FRW-2	CDR-2	FRW-3	CDR-3	FRW-4
17	IgG1	Kappa	DIVMSQSPSSLAVSVGEKITMSCKSS	QSLLYTSDQKNY	LAWFQQKPGQSPKLLIF	WAS	TRDSGVPDRFTGSGSGTDFTLTISSVKAEDLAVYYC	QQFYNYPRT	FGGGTKLEIK
21	IgG2b	Kappa	IVMTQTPKFLLVSAGDRVTITCKAS	QSVSND	VAWFQQKPGQSPKLLIY	FAS	NRYTGVPDRFTGSGYGTDFTFTITTVQAEDLAVYFC	QQDYSSPWT	FGGGTKLEIK
22	IgG2b	Kappa	IVMTQTPKFLLVSAGDRVTITCKAS	QSVSND	VAWFQQKPGQSPKLLIY	FAS	NRYTGVPDRFTGSGYGTDFTFTITTVQAEDLAVYFC	QQDYSSPWT	FGGGTKLEIK
57	IgG1	Kappa	DIVMTQAAPSEPVTPGESVSISCGSS	KSLLHSNDNTY	LYWFLQRPGQSPQLLIY	RMS	NLASGVPDRFTGSGSGTAFTLRISRVEAEDVGVYYC	MQHLENPLG	VRWRHQAGNQT
67	IgG1	Kappa	QIVLTQSPAIMSASLGERVTMTCTAS	SSVSSSY	LHWYQQKPGSSPKLWIY	STS	NLASGVPARFSGSGSGTSYSLTISSMEAEDAATYYC	LQYHRSPWT	FGGGTKLEIK
86	IgG1	Kappa	QIVLTQSPAIMSASLGERVTMTCTAS	SSVSSSY	LHWYQQKPGSSPKLWIY	STS	NLASGVPARFSGSGSGTSYSLTISSMEAEDAATYYC	LQYHRSPWT	FGGGTKLEIK

Figure 4.5. mBG western blot reactivity to linearized recombinant human coronavirus N protein. Recombinant human coronavirus western blots show the reactivity of the selected top mBG antibodies. SDS-PAGE gels for mBG67 and mBG86 were loaded with 10μg/well recombinant N protein while the other gels were loaded with 1μg/well due to concentration differences.



Antibodies from clones mBG17 and mBG86 were selected for use in sandwich ELISA to determine the monoclonal antibodies' effectiveness at detecting captured virion-derived N protein from SARS-CoV-2 virions. The rabbit polyclonal antibody was effective in capturing N protein, and the monoclonal mBG antibodies selectively labeled captured antigen as opposed to binding to the capture antibody. Both mBG17 and mBG86 showed proficiency at reporting captured virion N protein with the former being able to detect 5500 PFU/ml and the latter detecting 55,000 PFU/ml at the undiluted supernatant concentration (**Figure 4.6**). mBG17's higher signal compared to mBG86 is similar to the IFA data in **Figure 4.4**, and may indicate that the epitope for mBG17 (AA381-419) may be more accessible in N protein derived from infection than the mBG67/86 linear epitope (AA133-179). Given the propensity for N proteins to aggregate in vitro and the oligomerization of N protein on viral genomes during replication, additional steps (such as increasing ionic concentration or degrading RNA) may be needed to improve mBG86 accessibility in viral samples before the antibody line can be used in a diagnostic capacity. Future studies will assess how to increase the sensitivity of these antibodies in various applications.

4.3.2 Epitope Mapping and PAbFold Testing

We determined preliminary epitope ranges for the clones mBG17, mBG22, mBG57, and mBG67 using SARS-CoV-2 NP deletion mutants in Western blot analysis. These clones were selected from the top six clones due to mBG22 and mBG67's sequence similarity to mBG21 and mBG86, respectively. Recombinant N protein

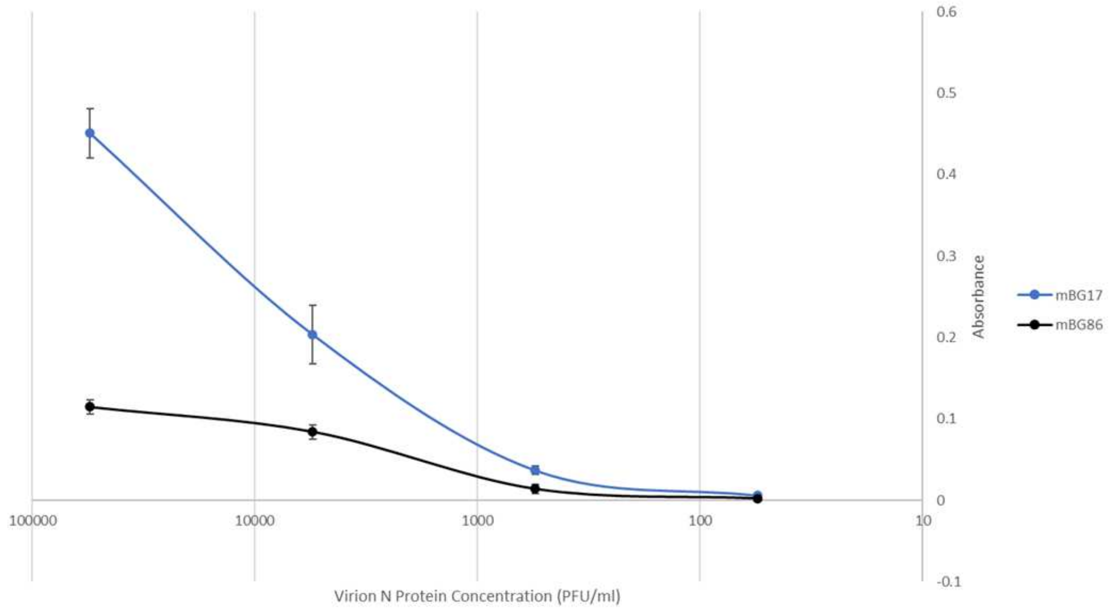
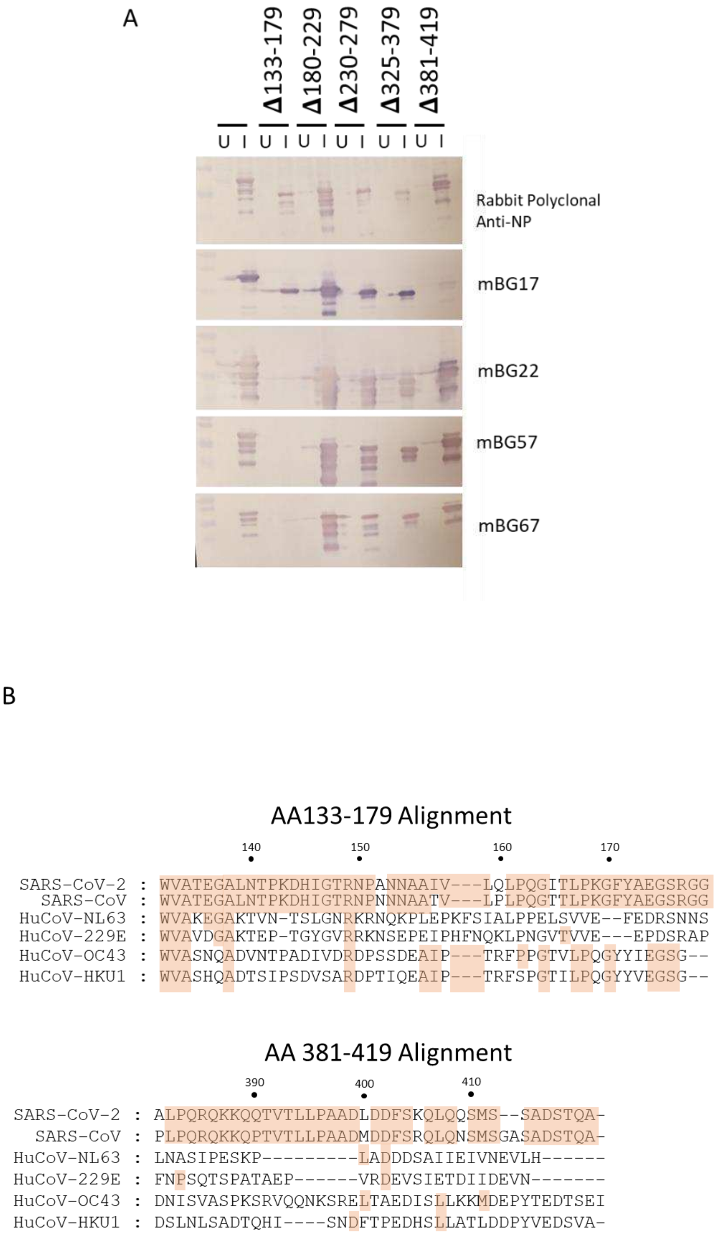


Figure 4.6. Reactivity of mBG17 and mBG86 against captured virion-derived N protein. Virion-derived N protein serially diluted from 5.5x10^5 PFU/ml stock and captured by anti-NP rabbit polyclonal antibody. mBG17 and mBG86 used in tandem with goat anti-mouse HRP-conjugated antibody to report on mBG effectiveness in capture ELISA

expression plasmids, each with a different deletion corresponding to a 50 amino acid region, were constructed and used to narrow down the range of epitope locations for mBG17, mBG22, mBG57, and mBG67 (**Figure 4.7**). mBG22, mBG57, and mBG67 showed strong reactivity in Western blot analysis against all deletion proteins except the Δ133–179 deletion, suggesting that the mBG21/22, mBG57, and mBG67/86 epitopes reside within AA133-179 range. mBG17 showed strong reactivity towards all deletions except Δ381–419, indicating that clone mBG17 is likely within AA381-419 at the C-terminal end of the N protein.

We collaborated with Jacob DeRoo from the Snow Lab and their PAbFold AlphaFold2 pipeline to assess the newly identified, rough epitope region for mBG17 [34]. This program utilizes machine learning to generate single-chain variable fragment antibodies using the known complimentary determining regions and framework regions from a monoclonal antibody [34]**.** We sought to experimentally verify the minimal linear epitope for mBG17 to determine how closely the AlphaFold2 prediction matched our experimental data. Seven 10 amino acid peptides with overlaps of 5 amino acids each were synthesized and used in competition ELISAs with mBG17 monoclonal antibody and recombinant SARS-CoV-2 NP (**Figure 4.8a**). The peptide corresponding to amino acids 401-410 (termed fragment 5) showed significant, nearly complete competition of mBG17 binding to NP following ELISA development, while none of the other peptides



AA133-179 Alignment

```
                   140        150            160        170
                    •          •              •          •
SARS-CoV-2 : WVATEGALNTPKDHIGTRNPANNAAIV---LQLPQGITLPKGFYAEGSRGG
  SARS-CoV : WVATEGALNTPKDHIGTRNPNNNAATV---LPLPQGTTLPKGFYAEGSRGG
HuCoV-NL63 : WVAKEGAKTVN-TSLGNRKRNQKPLEPKFSIALPPELSVVE--FEDRSNNS
HuCoV-229E : WVAVDGAKTEP-TGYGVRRKNSEPEIPHFNQKLPNGVTVVE---EPDSRAP
HuCoV-OC43 : WVASNQADVNTPADIVDRDPSSDEAIP---TRFPPGTVLPQGYYIEGSG--
HuCoV-HKU1 : WVASHQADTSIPSDVSARDPTIQEAIP---TRFSPGTILPQGYYVEGSG--
```

AA 381-419 Alignment

```
                   390        400        410
                    •          •          •
SARS-CoV-2 : ALPQRQKKQQTVTLLPAADLDDFSKQLQQSMS--SADSTQA-
  SARS-CoV : PLPQRQKKQPTVTLLPAADMDDFSRQLQNSMSGASADSTQA-
HuCoV-NL63 : LNASIPESKP---------LADDDSAIIEIVNEVLH------
HuCoV-229E : FNPSQTSPATAEP------VRDEVSIETDIIDEVN-------
HuCoV-OC43 : DNISVASPKSRVQQNKSRELTAEDISLLKKMDEPYTEDTSEI
HuCoV-HKU1 : DSLNLSADTQHI----SNDFTPEDHSLLATLDDPYVEDSVA-
```

Figure 4.7. Epitope mapping using N protein deletions. A) Western blot analysis of selected antibody reactivity against SARS-CoV-2 nucleocapsid protein deletions. Bacterial lysates of non-induced protein expression cultures (U) and induced protein expression cultures (I) were run side by side. **B)** Sequence alignments of NP AA133-179 and AA381-419 regions with heterologous human coronavirus nucleocapsid proteins.

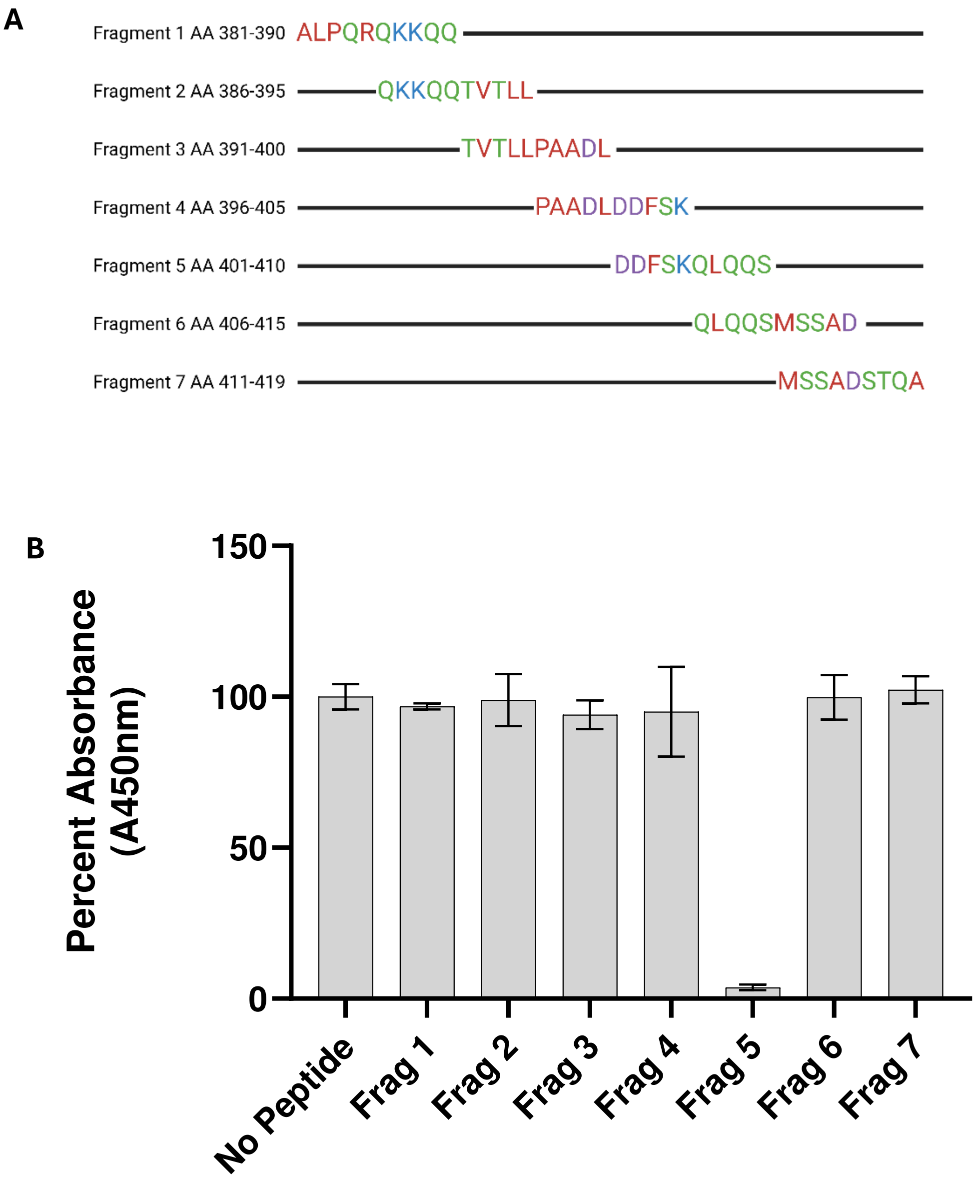


Figure 4.8. mBG17 epitope mapping using fragments of known epitope region in competition ELISA. A) Fragment sequences of peptides derived from known 39 amino acid epitope region for mBG17. Created with BioRender.com. **B)** Competition ELISA results following incubation of mBG17 primary antibody with epitope fragments. Severe loss of signal from fragment 5 incubation indicates the presence of the whole mBG17 epitope

significantly competed for mBG17 binding to wild-type NP. The alignment of all peptides used in the overlapping peptide competition ELISA experiments showed that the fragment 5 peptide sequence DDFSKQLQQS represents the experimentally determined epitope for mBG17 (**Figure 4.8b**). This was nearly identical to the epitope predicted by AlphaFold2.

To further experimentally characterize the binding of the mBG17 to fragment 5 (DDFSKQLQQS) peptide and compare experimental data with the predicted AlphaFold2 model, we designed and synthesized ten additional peptides, each containing an alanine point mutation at one position in the AA 401-410 peptide. The peptides are labeled D1A, D2A, F3A, S4A, K5A, Q6A, L7A, Q8A, Q9A, and S10A. Competition ELISAs were performed using increasing concentrations of each peptide to better assess differential binding (**Figure 4.9**). As expected, unmodified fragment 5 (AA 401-410) peptide showed strong competition, although Q9A competition was slightly increased. D1A showed no change in competition and thus is likely not involved in binding to the antibody CDR. The K5A, Q6A, and S10A substitution peptides demonstrated only minor reductions in competition while S4A showed a moderate reduction on competition. Peptides with the residues D2A, F3A, L7A, and Q8A all showed a strong reduction peptide-incubation competition. These data indicate that residues D2, F3, L7, and Q8 are important for mBG17 binding to NP Fragment 5 with S4 playing a moderate role.

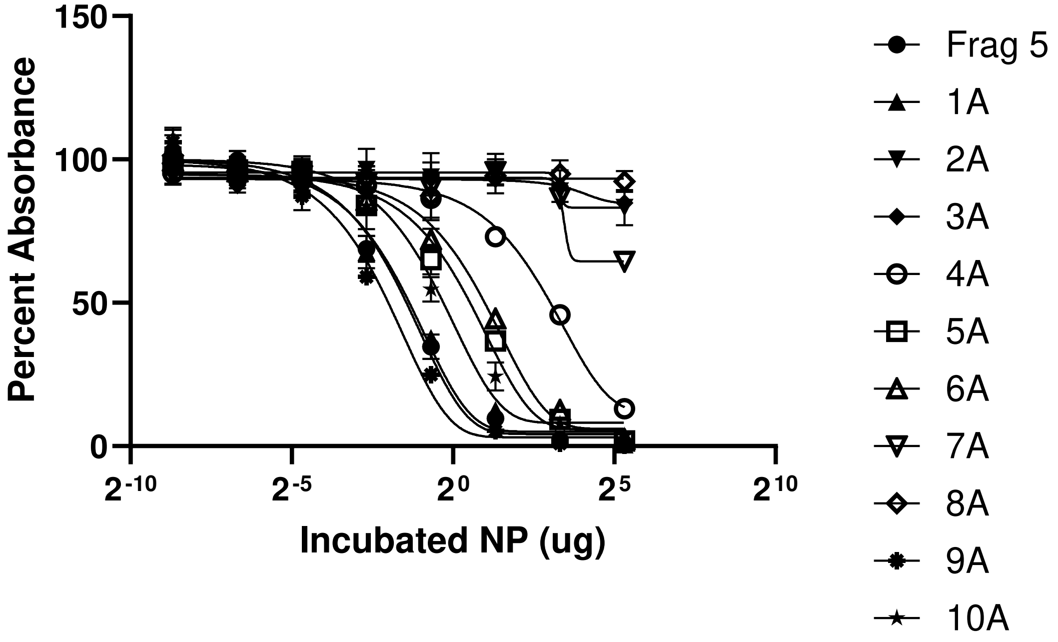


Figure 4.9. Residue analysis of mBG17 epitope sequence by alanine scanning. mBG17 was incubated with differing concentrations of peptides possessing an alanine substitution at each epitope residue and then tested via competition ELISA. Rescue of ELISA signal indicates the interruption of epitope-antibody interaction as a result of alanine substitution and thus that residue is important for antibody binding

4.4 Discussion

4.4.1 mBG Antibody Generation and Characterization

The arrival of the COVID-19 pandemic caught the scientific community off-guard. Despite the 2003 SARS outbreak and the MERS-CoV outbreak in the early 2010s, steps had not been taken at the research level to develop therapeutics or the government level to enact legislature for the protection of the global population from coronavirus outbreaks. When the pandemic hit and the world went into lockdown, there was a dearth of effective diagnostics, treatment, and reagents for tackling coronaviruses and developing new detection and research methods. This project landed in my lap as the world shut down with the COVID-19 quarantine. Before graduate school, I had experience at the UC Davis/NIH NeuroMab facility generating antibodies against mouse neuronal proteins. Along with my advisor Brian Geiss' antibody experience, we saw an opportunity to help move the research field forward in the global scientific effort and even expand further into research tool development.

With SARS-CoV-2 NP identified as a suitable antigen, we generated the antibodies from immunized mice and narrowed down suitable clones from the swirling mix of recently fused hybridomas. Several antibody lines were identified (mBG 17, 21, 22, 57, 67, and 86) that were active against SARS-CoV-2 NP. Characterization of these antibodies was thorough. This included assessing their reactivity to NP by western blot, indirect ELSIA, capture ELISA, and immunocytochemistry. We identified their isotypes using a lateral flow kit and cross-reactivity predisposition against the first SARS-CoV and other human coronaviruses with western blot. With extracted RNA from these identified hybridoma lines, we isolated the antibody-expressing gene sequences using the template-switch reverse transcription sequencing method from Meyer et al. and cloned them into vectors for sequencing. The framework regions and complementarity determining regions were identified by IgBlast and published with the rest of the results.

The resulting antibodies were available for use by the research community during the height of the pandemic, enabling future studies into N protein structure, intracellular interactions, diagnostic development, and potential therapeutics. Due to their single-cell origin, monoclonal antibodies are reliable reagents, capable of being manufactured in identical lots with consistent labeling compared to polyclonal antibodies. Reliable reagents are necessary for reproducible science and efficacy of diagnostic devices such as the diagnostic devices I will discuss in the next section. Thanks to their enhanced specificity, these antibodies can be used in the future for rapid, accurate, and inexpensive diagnostic assays for bed-side diagnosis of COVID patients. The identified sequences allow for the in-lab generation of recombinant antibodies in eukaryotic expression systems, reducing the research reagent cost and the need for new hybridoma generation by animal sacrifice. Additionally, knowledge of these antibody sequences allows for the creation of single-chain variable fragment antibodies that may be used for observing real-time N protein production or used in neutralizing treatments [35, 36].

4.4.2 Epitope Mapping and PAbFold Testing

Western blot analysis using region-deleted recombinant antigens determined rough ~50 amino acid epitopes for all 5 hybridoma lines. Later, a new project arose thanks to Jacob DeRoo in the Snow Lab at CSU. There, he developed PAbFold, a pipeline for the machine learning program Alphafold2 [34]. This pipeline predicts linear antibody epitopes after generating predicted antibody heavy and light chain interactions on a single-chain variable fragment (scFv) antibody, followed by testing interactions with a target antigen sequence. This was shown to work well when testing against HA and Myc tags and their specific, commercially available antibodies, but these interactions were already well characterized and used in machine-learning systems. The next step was to determine PAbFold's ability to predict the epitope of a new antibody when structural information was minimal.

With the epitope for mBG17 being roughly between amino acids 381-419 of NP (**Fig. 7**) we used the PAbFold pipeline to narrow down and specifically identify the 8-11 amino acid epitope window of this antibody. To test this prediction, we conducted more stringent epitope mapping experiments with competition ELISA experiments using overlapping peptide fragments of this region. If ELISA signals decreased following incubation with a peptide, then the fragment contained the epitope as the antibody was sequestered prior to application on the plate. We identified DDFSKQLQQS in fragment 5 as the mBG17 epitope due to the severe decrease in ELISA signal following antibody-fragment incubation (**Figure. 8b**). Even though the peptide fragments 4 and 6 overlap with fragment 5 at the N- and C-terminus respectively, neither were able to reduce ELISA signal, indicating that the mBG17 epitope binds to both halves of fragment 5. This aligned with PAbFold's prediction, demonstrating the pipeline's ability to accurately predict an antibody's binding region to a novel linear epitope that is not included in AlphaFold2's training set. Pleased with this result, we took it a step further to assess PAbFold's ability to determine amino acids in an epitope that are critical for binding. We used fragments with an alanine substitution at each amino acid position in the epitope to recognize which residues were required to rescue activity in the ELISA. A return of signal indicated that the substituted residue was needed for proper binding to the antibody. We identified the substitutions D2A, F3A, L7A, and Q8A as reducing competitive binding significantly while S4A showed a moderate rescuing of ELISA signal (**Figure 4.9**). The best scoring AlphaFold2 mBG17 DDFSKQLQQ model suggests the alanine substitutions disrupted specific antigen-antibody interactions with Fragment 5. Peptide residue D2 forms a hydrogen bond with mBG17 Y34 and residue F3 forms a hydrophobic interaction with mBG17 L185. Residue S4 lacks appears to lack a binding partner while residue L7 engages in a hydrophobic interaction with mBG17 A104 at the base of the binding cleft. Lastly, residue Q8 hydrogen bonds with both the backbone carbonyl group of Y34 and the amide of the W35 backbone [34]. The residues that demonstrated little to no effect on competition are all predicted to primarily interact with the solvent while lacking observable interactions between the NP peptide and scFv sequence. In summary, the AlphaFold2-driven PAbFold prediction was remarkably consistent with the experimental alanine scanning data. Due to the correct prediction for how the linear epitope binds to the antibody, we feel confident that the prediction of the mBG17 linear epitope location is accurate.

Altogether, testing the PAbFold pipeline's novel application for identifying antibody epitopes based on the antigen structure and antibody variable fragment region sequences was fruitful. The pipeline was able to identify specific antibody epitopes in parallel with experimental data that validated these findings without the normally required labor and reagent costs. While we began with a rough epitope region, the program can be applied across a region of interest or the entire sequence of an antigen to identify where antibodies will bind. This development in predictive technology is significant for furthering the study of antibodies, both as research tools and diagnostics. Generating new antibodies is a lengthy process that requires animal sacrifice and a significant amount of testing and vetting. The speed and specificity of machine learning tools such as AlphaFold2's PAbFold pipeline can significantly reduce the time and effort it takes to generate new tools for virus research and treatment. Tailor-made antibodies are no longer a feat of science fiction, and, considering the rate in which the SARS-CoV-2 spike protein mutated during the COVID-19 pandemic, this technology can help the scientific community keep up with viral evolution.

With the efficacy of PAbFold's linear epitope binding capabilities established, more work needs to be done to expand PAbFold's applications. Many antibodies bind to residues brought into close proximity with each other by a protein's tertiary structure. As

PAbFold has only been used for predicting and characterizing linear epitope sequences, further tweaking and testing will be needed to expand its usefulness into predicting nonlinear epitopes. This is especially true considering that it works by chopping the sequence up into short fragments and assessing binding affinity with the generated scFv. The DeRoo Lab and Geiss Lab are looking into using next generation programs such as AlphaFold3 or RF Diffusion and Protein MPNN to computationally predict the conformational epitopes for antibodies. Another future project in the Geiss Lab is to modify the CDR of the mBG17 scFv's to adjust the binding epitope. This pipeline addition would allow us to choose where on the antigen we want a modified scFv to bind, generate predictions of CDRs on PAbFold, and test these predictions experimentally. Such recombinant antibodies would be producible in bacterial expression systems due to the relatively simple folding mechanisms required for an scFv. Expanding on the PAbFold pipeline would be a significant step in studying how to modify existing monoclonal antibodies for increased efficacy in an ethical and reproducible manner.

4.4.3 Translational Science Work

While working on the generation and characterization of the mBG monoclonal antibodies I had the opportunity to work on the development of diagnostic devices with the Henry Lab at CSU. The COVID-19 pandemic demonstrated the need for inexpensive diagnostics that were easy to use from home when medical services were saturated, and these devices aimed at filling that need. These quick, easy-to-use ELISA devices could be used at home or in a clinic for rapid indication of antibody or antigen presence within a sample. I served as a consultant and wet-lab validator during development of these devices. While they originally used the mBG monoclonal antibodies, the projects eventually focused on using polyclonal alternatives due to a polyclonal mixture's varied epitope targets for antigen capture. Nevertheless, modifications to the mBG variable regions predicted by the PAbFold pipeline may provide a beneficial change, as a generated mix of modified monoclonal antibodies would allow for reproducible antibody batches as opposed to the variability of a polyclonal source.

I became a contributing author with my contributions to the work reported in Samper et al. Here, we developed and validated an electrochemical device utilizing a change in current as the reporting mechanism for a capture ELISA targeting SARS-CoV-2 NP in nasopharyngeal samples [37]. In these devices a capture polyclonal antibody targeting NP was attached to screen-printed carbon electrodes. Following sample administration, secondary polyclonal antibodies conjugated to horseradish peroxidase were applied followed by TMB. During this process, an electrochemical potentiostat measured the change in current. If an antigen was present in the sample, it would be pulled out by the capture antibodies and retained on the electrode. The resulting secondary antibody binding and TMB reaction would generate an electrochemically active variant of TMB that the potentiostat would identify. Changes in current were associated with specific quantities of NP and would allow for accurate identification of antigen presence.

Another device created in the Henry lab was the Capillary-Driven Immunoassay (CaDI), an inexpensive microfluidic device with layers of clear plastic and a paper pump to deliver reagents sequentially across an antibody-embedded membrane, replicating the steps of a capture ELISA. Adding a sample to a port and buffer to another would trigger the delivery of reagents through the device, free from an external energy source requirement. Wash buffer, HRP-conjugated secondary, and TMB solution would follow across this membrane and, if NP was present in the sample, a blue band would appear. The diagnostic device design was innovative in its ease of use, affordability, and specificity. Because of my work providing insight into the ELISA methodologies/reagents and wet-lab validation experiments, I became the contributing author on two papers. Cody et al. chronicled the device's generation and efficacy at identifying samples containing host antibodies targeting NP [38]. In Link et al. they worked to apply the same mechanisms for identifying NP within a sample [39]. The CaDI project showed how basic research translates into the industrial realm, as the device became the springboard for the founding of the biotech company Burst Diagnostics by Dr. Chuck Henry, Dr. David Dandy, and Dr. Brian Geiss.

Assisting with the development of these devices provided the opportunity to be a part of the translational science side of research. Applying the discoveries of foundational science to generate tools to help the public during a global health crisis was a rewarding experience that offered a glimpse into the industry side of science.

4.4.4 Reflection

My work on this project and the other related efforts straddled the line between foundational science and applied science. This region is where many microbiologists thrive as our work feels closer to causing direct results that can be seen in enhancing public health. Additionally, the allure of science (at least for me) is discovering new things and sharing them in the hopes that my small contributions can help build a larger breakthrough that makes the world better. With this work I discovered and studied new antibody lines that were specific to this new threat against public health and could be used to study this pathogen. These antibodies are still available to any researcher who needs them. The specific sequences for each antibody were revealed and are something that can be shared with other researchers easily, where they can use them to make homegrown antibodies for experiments at the benchtop or for modeling studies *in silico*. These antibodies and their sequences I discovered are now published and available for use and manipulation in the pursuit of knowing more about what is (hopefully) the biggest threat to global health in my lifetime, and personally I find that quite nifty.

I did not expect to take this detour in my graduate school studies, but the pandemic was one of those times in life where the reins are ripped from our hands. I instead focused on what I could control, and as a result I had the opportunity to participate in research with incredible people while developing novel tools against the SARS-CoV-2 pathogen.

4.5 REFERENCES

1. Terry, J.S., L.B. Anderson, M.S. Scherman, et al., *Development of a SARS-CoV-2 nucleocapsid specific monoclonal antibody.* Virology, 2021. **558**: p. 28-37.

2. *COVID-19 Dashboard*. 10 March 2023 [cited 2024; Available from: https://gisanddata.maps.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6.

3. Corman, V.M., O. Landt, M. Kaiser, et al., *Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR.* Euro Surveill, 2020. **25**(3).

4. Wang, W., Y. Xu, R. Gao, et al., *Detection of SARS-CoV-2 in Different Types of Clinical Specimens.* JAMA, 2020. **323**(18): p. 1843-1844.

5. Zhai, J., T. Briese, E. Dai, et al., *Real-time polymerase chain reaction for detecting SARS coronavirus, Beijing, 2003.* Emerging Infecitous DIsease, 2004. **10**(2): p. 311-316.

6. Li, Y., L. Yao, J. Li, et al., *Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19.* J Med Virol, 2020. **92**(7): p. 903-908.

7. Chang, C.K., Y.L. Hsu, Y.H. Chang, et al., *Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: implications for ribonucleocapsid protein packaging.* J Virol, 2009. **83**(5): p. 2255-64.

8. Saikatendu, K.S., J.S. Joseph, V. Subramanian, et al., *Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-terminal domain of N protein.* J Virol, 2007. **81**(8): p. 3913-21.

9. Zeng, W., G. Liu, H. Ma, et al., *Biochemical characterization of SARS-CoV-2 nucleocapsid protein.* Biochem Biophys Res Commun, 2020. **527**(3): p. 618-623.

10. Zuniga, S., I. Sola, J.L. Moreno, et al., *Coronavirus nucleocapsid protein is an RNA chaperone.* Virology, 2007. **357**(2): p. 215-27.

11. Cavanagh, D., *Coronavirus IBV further evidence that the surface projections are associated with two glycopolypeptides.* The Journal of General Virology, 1983. **64**: p. 1787-91.

12. Hogue, B.G. and D.A. Brian, *Structural proteins of human respiratory coronavirus OC43.* Virus Research, 1986. **5**(2-3): p. 131-44.

13. Neuman, B.W., G. Kiss, A.H. Kunding, et al., *A structural analysis of M protein in coronavirus assembly and morphology.* J Struct Biol, 2011. **174**(1): p. 11-22.

14. Escors, D., E. Camafeita, J. Ortego, et al., *Organization of two transmissible gastroenteritis coronavirus membrane protein topologies within the virion and core.* J Virol, 2001. **75**(24): p. 12228-40.

15. Liu, D.X. and S.C. Inglis, *Association of the infectious bronchitis virus 3c protein with the virion envelope.* Virology, 1991. **185**(2): p. 911-917.

16. Che, X.Y., W. Hao, Y. Wang, et al., *Nucleocapsid protein as early diagnostic marker for SARS.* Emerging Infecitous DIsease, 2004. **10**(11): p. 1947-49.

17. Li, Y.H., J. Li, X.E. Liu, et al., *Detection of the nucleocapsid protein of severe acute respiratory syndrome coronavirus in serum: comparison with results of other viral markers.* J Virol Methods, 2005. **130**(1-2): p. 45-50.

18. Shi, Y., Y. Yi, P. Li, et al., *Diagnosis of severe acute respiratory syndrome (SARS) by detection of SARS coronavirus nucleocapsid antibodies in an antigen-capturing enzyme-linked immunosorbent assay.* J Clin Microbiol, 2003. **41**(12): p. 5781-2.

19. Zhu, H., S. Hu, G. Jona, et al., *Severe acute respiratory syndrome diagnostics using a coronavirus protein microarray.* Proceedings of the National Acadamy of Sciences of The United States of America, 2006. **103**(11): p. 4011-6.

20. Peiris, J.S.M., C.M. Chu, V. Cheng, et al., *Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia a prospective study.* Lancet, 2003. **361**(9371): p. 1767-72.

21. Li, T., L. Wang, H. Wang, et al., *Serum SARS-COV-2 Nucleocapsid Protein: A Sensitivity and Specificity Early Diagnostic Marker for SARS-COV-2 Infection.* Front Cell Infect Microbiol, 2020. **10**: p. 470.

22. Diao, B., K. Wen, J. Chen, et al., *Diagnosis of Acute Respiratory Syndrome Coronavirus 2 Infection by Detection of Nucleocapsid Protein.* MedRxiv, 2020.

23. Lee, H., B.H. Lee, N.K. Dutta, et al., *Detection of antibodies against SARS-Coronavirus using recombinant truncated nucleocapsid proteins by ELISA.* Journal of Microbiology and Biotechnology, 2008. **18**(10): p. 1717-21.

24. Lee, H.K., B.H. Lee, S.H. Seok, et al., *Production of specific antibodies against SARS-coronavirus nucleocapsid protein without cross reactivity with human coronaviruses 229E and OC43.* J Vet Sci, 2010. **11**(2): p. 165-7.

25. Dutta, N.K., K. Mazumdar, B.H. Lee, et al., *Search for potential target site of nucleocapsid gene for the design of an epitope-based SARS DNA vaccine.* Immunol Lett, 2008. **118**(1): p. 65-71.

26. Geiss, B.J., A.A. Thompson, A.J. Andrews, et al., *Analysis of flavivirus NS5 methyltransferase cap binding.* J Mol Biol, 2009. **385**(5): p. 1643-54.

27. Stadlbauer, D., F. Amanat, V. Chromikova, et al., *SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup.* Curr Protoc Microbiol, 2020. **57**(1): p. e100.

28. Case, J.B., A.L. Bailey, A.S. Kim, et al., *Growth, detection, quantification, and inactivation of SARS-CoV-2.* Virology, 2020. **548**: p. 39-48.

29. Fagre, A., J. Lewis, M. Eckley, et al., *SARS-CoV-2 infection, neuropathogenesis and transmission among deer mice: Implications for spillback to New World rodents.* PLoS Pathog, 2021. **17**(5): p. e1009585.

30. Meyer, L., T. Lopez, R. Espinosa, et al., *A simplified workflow for monoclonal antibody sequencing.* PLoS One, 2019. **14**(6): p. e0218717.

31. Ye, J., N. Ma, T.L. Madden, et al., *IgBLAST: an immunoglobulin variable domain sequence analysis tool.* Nucleic Acids Res, 2013. **41**(Web Server issue): p. W34-40.

32. DeLuca, K.F., J.E. Mick, A.H. Ide, et al., *Generation and diversification of recombinant monoclonal antibodies.* Elife, 2021. **10**.

33. Yamaoka, Y., S. Matsuyama, S. Fukushi, et al., *Development of Monoclonal Antibody and Diagnostic Test for Middle East Respiratory Syndrome Coronavirus Using Cell-Free Synthesized Nucleocapsid Antigen.* Front Microbiol, 2016. **7**: p. 509.

34. DeRoo, J., J.S. Terry, N. Zhao, et al., *PAbFold: Linear Antibody Epitope Prediction using AlphaFold2.* bioRxiv, 2024.

35. Zhao, N., K. Kamijo, P.D. Fox, et al., *A genetically encoded probe for imaging nascent and mature HA-tagged proteins in vivo.* Nat Commun, 2019. **10**(1): p. 2947.

36. Nakanaga, K., K. Yamanouchi, and K. Fujiwara, *Protective effect of monoclonal antibodies on lethal mouse hepatitis virus infection in mice.* Journal of Virology, 1986. **59**(1): p. 168-71.

37. Samper, I.C., C.J. McMahon, M.S. Schenkel, et al., *Electrochemical Immunoassay for the Detection of SARS-CoV-2 Nucleocapsid Protein in Nasopharyngeal Samples.* Anal Chem, 2022. **94**(11): p. 4712-4719.

38. Carrell, C., I. Jang, J. Link, et al., *Capillary driven microfluidic sequential flow device for point-of-need ELISA: COVID-19 serology testing.* Anal Methods, 2023. **15**(22): p. 2721-2728.

39. Link, J.S., C.S. Carrell, I. Jang, et al., *Capillary flow-driven immunoassay platform for COVID-19 antigen diagnostics.* Anal Chim Acta, 2023. **1277**: p. 341634.