2023

Investigating the Viral Nature of Neuronal Capsid Genes

Michael Manuel Long

LOYOLA UNIVERSITY CHICAGO

INVESTIGATING THE VIRAL NATURE OF NEURONAL CAPSID GENES

A THESIS SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

MASTER OF SCIENCE

PROGRAM IN NEUROSCIENCE

BY

MICHAEL M. LONG

CHICAGO, IL

MAY 2023

ACKNOWLEDGEMENTS

I would like to thank my loving family for giving me constant support, especially my parents. I also want to thank my brother and sister for being excellent role models. Thank you to my friends for making sure I always have someone to talk to. Thanks to my dog for giving me periodic memory tests on whether or not she was already fed. Thanks to everyone whom I have hugged and who has hugged me back for sharing a moment of love. Thanks to my mentor, Dr. Ed Campbell, for challenging me and making me a better scientist.

Dedicated to my departed coach Bud Skuba, who advised me to
“Work hard at school as you did for baseball!”

Sometimes you can have a whole lifetime in a day and never notice that this is as beautiful as it gets.

-Robin Williams

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

NT	Neurotransmitter
LTR	Long-terminal repeat
PNMA	Paraneoplastic MA-antigen
MART	Mammalian Retrotransposon
Peg10	Paternally Expressed Gene 10
Ty3	Transposon-Yeast-3
HIV	Human Immunodeficiency Virus
MA	Matrix
CA	Capsid
NC	Nucleocapsid
RT	Reverse Transcriptase
PRO	Protease
POL	Polymerase
ENV	Envelope
IN	Integrase
PI	Protease Inhibitor
MLV	Murine Leukemia Virus
pLVX	Lentiviral Expression Vector
Vsv-G	Vesicular Stomatitis Virus

TBST	Tris-Buffered Saline Tween-20
WAVE1	WASP family verprolin homology isoform 1
hSYN	Human Synapsin
Prox1	Prospero Homeobox 1
Ctip2	Chicken ovalbumin upstream promoter transcription factor-interacting protein 2
NeuN	Neuronal Nuclear Protein
Map2	Microtubule Associated Protein 2
DIV	Day in vitro

ABSTRACT

Memory formation is an essential part of everyday life. Neuronal communication is thought to be a key driver behind memory formation. There are two heavily studied methods of neuronal communication: chemical (neurotransmitter-mediated) and electrical (physically attached). When issues arise environmentally or genetically, people can experience one of many devastating memory disorders i.e. Alzheimer's disease and other Dementias. Due to their prevalence, the treatments for disorders that hinder memory formation are a high priority for pharmaceutical companies. Currently, many of the treatments available focus on neurotransmitters or neuronal activity, and scientists are far from a cure. However, there may be an undiscovered target for the treatment of these disorders.

Recently, an essential memory protein known as the Activity-Regulated Cytoskeleton-associated (Arc) protein was shown to have evolved from the same parent as retroviruses[(1)]. Along this viral trend, purified Arc protein spontaneously aggregates into capsid-like structures[(2, 3)]; this capsid structure and evolutionary history were expanded on by recording Arc's transfer of mRNA from one neuron into another[(3)]. Two other families of endogenous neuronal capsid genes were identified to form capsids, and a gene, Peg10, exhibited RNA transfer between cells[(4)]. Taken together, this information creates the potential for an additional neuronal communication paradigm. If proved true, this would open a new target for therapeutics to treat widespread memory disorders.

Our goal is to test the viral nature of the essential memory protein, Arc. We will examine its viral nature by measuring Arc's release from neuronal and non-neuronal cells. Additionally, we will utilize HIV's well-understood capsid assembly and infectivity to test the viral nature of endogenous capsid genes. By replacing the capsid gene of HIV-1 with the putative CA domains of neuronal CA genes, we will test their ability to infect cells. Thus investigating the infectious nature of endogenous neuronal capsid genes.

CHAPTER ONE

INTRODUCTION

Significance and Innovation

In today's neurological disease treatments, we focus on neurotransmitters (NTs) and NT receptors. However, the neurons adapt to these treatments making them less effective if not ineffective. This limited treatment plan exists for two reasons, the limited state of technology, and our lack of knowledge of neuronal communication. What is known about neuronal communication is that neurons communicate between one another one of two ways: chemical NT release or electrical currents. These signals result in effective communication for all proper motor and cognitive functions. Although when this communication fails, we have drug treatments to increase NT receptor density, increase NT available, or block NT receptors. But what if we found a new target for drug therapies that wasn't reliant on neurotransmitters?

Recently, an essential memory protein (Arc) has shown an ability to form virus-like capsid structures and deliver functional RNA to naïve cells[(3)]. This finding suggests a new route of neuronal communication co-opted from viruses where neurons can send and receive genomic information. This could become a new target for therapeutics to alleviate neurological disorders as Arc is implicated in many of the most prominent neurological diseases.

Arc is one of three families of genes that evolved from long-terminal repeat LTR Retrotransposons. Members of the other families Peg10 and PNMA2 have been implicated in neurological diseases, as well. Members of all three families have been observed forming

capsids. Moreover, Peg10 has the ability to house and transport target RNA into cells like a virus. Hoisting the notion that neurons have an unstudied method of communication that they stole from viruses through evolution. Collectively, there is an urgency to study these endogenous capsid proteins (Arc, Peg10, PNMA2); for, understanding their ability to transport information between neurons could bring about novel therapeutic targets for many difficult-to-treat neurological disorders. This thesis is composed of research focused on the induction of Arc release from cells and the abilities of Arc, PNMA2, and Peg10 to facilitate a true retroviral infection. Research into these neuronal capsid genes may generate a paradigm shift in memory and cognition research as well as potential discussions on the viral implications of cognitive functions on an evolutionary level.

Discovery of *Arc*

The activity-regulated cytoskeleton-associated (Arc) protein is causing a paradigm-shift in the way we think about neuronal communication. *Arc* was first recognized as a gene that is expressed rapidly in response to neuronal stimulation, as produced from high-frequency stimulation and drug-induced seizures [5, 6]. Due to its rapid expression in response to stimuli, *Arc* became classified as an immediate early gene (IEG). Upon imaging of brain slices, it became clear that *Arc* mRNA is predominantly found in the dendritic portions of neurons[5, 6], where it undergoes local translation[7]. The dendrites of neurons receive and propagate signals from other neurons and their proper function is essential for healthy cognitive functions.

Memory Formation

Long-term potentiation (LTP) and Long-term depression (LTD) are states of signaling pathways that result from stimulation patterns, and establish an increase or decrease the strength of a signaling pathway between neurons, respectively. The difference between LTP and LTD is

the frequency of the stimulation. High-frequency stimulation (LTP) establishes a strong connection between neurons. Low-frequency stimulation (LTD) weakens the connection between neurons. On a molecular level, in response to canonical stimulation the excitatory NT, Glutamate, is released from the stimulated neuron's axon terminal to have a stimulatory effect on the downstream neuron's dendritic spines. Glutamate will bind to ionotropic receptors known as alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors (AMPARs), N-methyl-d-aspartate receptors (NMDARs), and Kainate receptors to induce an excitatory effect by allowing positive cations such as sodium and calcium to rush into the neuron's dendritic region. This rush of positive cations into the cell results in a membrane depolarization that will attempt to travel throughout the neuron to propagate the signal. Only large enough depolarizations will successfully reach the threshold to fire an action potential. This action potential will travel the length of the neuron's axon and will signal to neurons which are synapsed the axon terminal of the signaling neuron. Initial glutamate release from a signaling neuron will bind to and open AMPARs, but it will likely not open NMDARs. Glutamate will bind to NMDARs, but they require a larger membrane depolarization to induce the positive flow of cations. This is due to a magnesium block within NMDARs, which requires a significant change in membrane potential, or a high-frequency stimulation like one that causes LTP, to be forced out. Once magnesium is removed, NMDARs will have an additive effect to the AMPARs excitatory effect on that neuron which will result in a large membrane depolarization that may propagate down that neuron. Signals with slow LTD-inducing frequencies will open AMPARs, but will not elicit a large enough membrane depolarization to open NMDARs. However, the number of AMPARs present in that given synapse is a crucial factor. If there are many AMPARs present, a low frequency signal may release enough glutamate to trigger an NMDAR response; likewise, with too few

homeostatic regulation of AMPARs is pivotal in the propagation of LTP and LTD signals. LTP and LTD are the homeostatic signaling mechanisms thought to be the molecular mechanisms of memory formation.

Functions of the Arc Protein

The Arc protein is locally translated in the dendrites, where it interacts with many proteins. Most notably, Arc interacts with the cytoskeleton[8]. Additionally, Arc interacts with dynamin, endophilin, and clathrin-adaptor protein 2[9, 10]. These interactions have been shown to regulate the ability of Arc to impact AMPAR endocytosis. Following *Arc* knockout, the ratio of internalized AMPARs/surface AMPARs is significantly lower[9]. This points to a direct role of Arc in AMPAR endocytosis. AMPARs are essential to normal synaptic signaling and undergo regulated endocytosis to maintain homeostatic synaptic signaling[11]. It appears that Arc plays an important regulatory role in maintaining homeostatic properties of the postsynaptic membrane activity. Thus, proper regulation of *Arc* is vital to healthy cognition functions.

Arc in memory formation. Given the localization of *Arc* mRNA and protein in dendrites and Arc's role in AMPAR endocytosis and the importance of AMPARs in memory signaling processes in the dendrites, rodents were investigated conducting memory tasks. The loss of *Arc* results in severely hindered long-term memory formation[12, 13]. Arc knockout mice show disrupted memory processes on many common memory tasks: Morris water maze, contextualized & cued fear conditioning, novel object recognition, and conditioned taste aversion[13]. When Arc is knocked down before or immediately after learning a Morris water maze platform location, the rodents are unable to store the long-term information of platform location, but in the short-term they are able to recall it[12]. However, when Arc is knocked down

8 hours after the learning task, there is no evidence of hindered memory processes. This signifies the importance of Arc during the transient learning period of a task in order to store long-term memories. Therefore, Arc is essential for long-term memory consolidation processes during learning periods.

Arc's implication in neurological diseases. Neuronal signaling is regulated by many factors that must function properly to maintain cognitive homeostasis. It's clear that Arc affects dendritic components essential for neuronal signaling. When a component of the neuronal signaling pathway falters or is influenced by an outside factor, it may result in a neurological disease phenotype. Many studies have suggested that Arc contributes to the pathophysiology of neurological conditions that affect cognition. The gene responsible for causing Fragile X syndrome expresses a protein that suppresses Arc translation[(14)]. A mutation that disrupts the E3 Ubiquitin ligase Ube3a causes the debilitating neurological disease Angelman syndrome. Ube3a directly controls the degradation of Arc, so Arc is overexpressed in Angelman syndrome[(15)]. Overexpression of Arc is observed in a mouse model of Alzheimer's disease, while knockout of Arc results in a decreased amount of disease-associated Amyloid-beta[(16)]. Finally, a recent study showed that plasma levels of Arc protein are significantly higher in individuals with Autism Spectrum disorder versus healthy individuals[(17)]. The *Arc* gene is essential for healthy cognition, but its evolutionary history may surprise you.

LTR Retrotransposons and Retroviruses

Long-terminal repeat (LTR) Retrotransposons and Retroviruses are classes of genes that evolved to move around the genome of organisms. The main difference between the two is that LTR retrotransposons move around the genome within a cell, whereas retroviruses proliferate to transfer genetic material between cells and infect naïve cells with their genetic information. This

difference is highlighted by the envelope domain (ENV) in retroviruses. In LTR retrotransposons, ENV is not present (Fig. 1). Without ENV, located in the plasma membrane responsible for fusion with naïve cells, prevents LTR retrotransposons from fusing, upon a serendipitous release. Besides ENV, LTR retrotransposons and retroviruses share significant homology; specifically, they share the three essential domains: group-specific antigen (GAG), protease (PRO), and polymerase (POL).

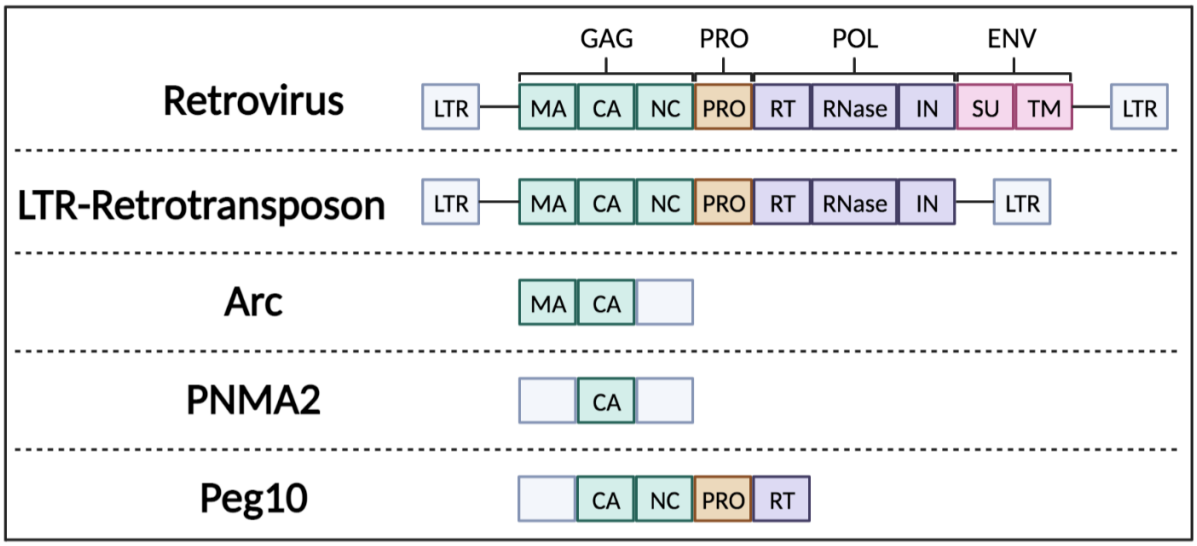

Figure 1. Protein Domains of Retroviruses, LTR Retrotransposons, and Endogenous CA Genes. The evidence for the putative domains of Arc, PNMA2, and Peg10 originated here[(1)]. Abbreviations are as follows: LTR, Long-terminal repeat; GAG, Group-specific Antigen; MA, Matrix; CA, Capsid; NC, Nucleocapsid; PRO, Protease; POL, Polymerase; RT, Reverse Transcriptase; IN, Integrase; ENV, Envelope; SU, Surface; TM, Transmembrane.

The GAG domain is comprised of matrix (MA), capsid (CA), and nucleocapsid (NC). The MA is responsible for targeting the viral particle to the plasma membrane. The CA combines to form the capsid structure. The NC contains an RNA-binding domain to hold the genetic material within the capsid. The PRO domain is comprised of the protease domain responsible for cleaving the numerous domains for them to carry out their proper functions. The POL domain is comprised of reverse transcriptase (RT), RNase, and integrase (INT). RT reverse transcribes the

target RNA into DNA so that it may integrate into the genome using INT. The RNase degrades template RNA following the RT process.

With the GAG and POL domains, both LTR retrotransposons and retroviruses form capsids that house genetic material and are capable of integrating the material in the genome of cells. Recently the scientific community has uncovered many mammalian genes that evolved from LTR retrotransposons who share significant homology with members of the GAG and POL domains. Some of these endogenous genes have been shown to impact cognitive functions. Notably, Arc evolved from LTR retrotransposons[1].

Endogenous Neuronal CA Genes

Arc. Arc's ancestral history provides a unique perspective towards its potential functions. The *Arc* gene evolved from the LTR retrotransposon family Ty3/*Gypsy*[1]. As discussed, both LTR retrotransposons and retroviruses share the GAG domain responsible for viral capsid formation. Indeed, Arc shares structural and sequence homology with the GAG domains of Ty3 and Human Immunodeficiency Virus (HIV), a notable retrovirus[1, 18, 19]. Arc specifically shares significant homology with the CA domain of Ty3 and HIV. The presence of a CA domain in Arc points to a question: does Arc form capsid-like structures? A finding by Myrum et al changed the face of Arc research. While attempting to purify the Arc protein, it continued to form precipitates in solution. Upon further investigation via electron microscopic images, Arc appeared to be oligomerizing into structures resembling viral capsids[2]. This finding has been replicated recently[3], and expanded on by performing structural analyses using high resolution cryo-electron microscopy images of the Drosophila Arc capsids[18]. However, high-resolution structural analysis of mammalian Arc has not been recorded. Similarities between the structures of Arc and HIV have provided clues to Arc's ability to perform viral functions.

It is accepted that mammalian Arc forms structures that resemble viral capsids[2, 3, 20]. Due to this, the viral nature of Arc is beginning to be investigated. Arc is present in the extracellular vesicle (EV) fraction of supernatant collected from cultured rat cortical neurons[3], suggesting Arc exits the cell, similar to HIV. Following RNase treatment on purified recombinant Arc to degrade free RNA, there was no significant difference between the amounts of *Arc* mRNA present in samples treated with RNase versus samples left untreated[3], showing *Arc* mRNA is protected from RNase degradation by being housed within the capsid. Further, the addition of EVs from cultured rat hippocampal neurons onto *Arc*-knockout rat hippocampal neurons resulted in an abundance of *Arc* mRNA and *Arc* protein within the Arc-knockout neurons[3], providing evidence for the intercellular transfer of Arc. Lastly, they determined that the transferred Arc mRNA undergoes translation in the recipient cell[3], contributing evidence to a previously unknown signaling pathway between neurons. Due to the recency of the Arc capsid discovery, studies of the viral nature of Arc are finite, creating a gap in knowledge of this novel therapeutic target for altering communication between neurons.

Few studies have investigated endogenous capsid genes in neurons but a recent paper screened embryonic and fetal rodent brains for highly expressed endogenous capsid genes. They discovered two highly expressed capsid genes: PNMA2 and Peg10[21].

PNMA2 and Peg10. Recently, human endogenous descendants of Ty3/*Gypsy* LTR retrotransposons that share homology with retroviruses have been identified to form capsid-like structures[4]. Of five families of descendants[1], we're examining three that are highly expressed in neurons[5, 21]: The Arc, PNMA, and MART families[1]. There are multiple genes within these families that share homology with certain domains of retroviruses. Namely, the CA domain of retroviruses is the most commonly shared region. Considering the CA domain forms the capsid

structure, a recent study found many of these endogenous capsid genes form oligomeric structures that resemble capsids[4], including PNMA2 and Peg10.

PNMA2 is an endogenous CA gene from the Paraneoplastic MA-antigen (PNMA) family[22]. This name is derived from antibodies against proteins within this family being present in paraneoplastic neurological disorders. These disorders cover broad neurological dysfunctions that occur in conjunction with tumors. Importantly, members of the PNMA family, most of which contain a CA domain, are associated with neurological dysfunction.

Peg10 is an endogenous CA gene belonging to the Mammalian Retrotransposon (MART) family. As it does for Arc, Ube3a also regulates Peg10 degradation[23]. Without Ube3a patients suffer from the neurodevelopmental disorder, Angelman Syndrome. Neurons derived from patients with AS show overexpression of Peg10, and overexpression of Peg10 was show to severely impact neuronal migration[23]. So, Peg10 is associated with neuronal dysfunction that manifests as a neurodevelopmental disorder.

Along with Arc, both PNMA2 and Peg10 have been observed in EVs, highlighting their ability to exit cells[4, 21, 23]. As Arc was shown to share genetic information, Peg10 was recently investigated for the same function. Segel et al. tested whether Peg10 was capable of transferring cargo RNA from one cell to another. A simplified HIV infection from one cell to another involves a packaging apparatus (GAG-POL), cargo RNA to be transferred, and envelope protein. The cargo RNA is housed within the GAG complex, the virus is released from the cell using the cellular machinery, and the virus is coated in envelope proteins to facilitate fusion with naïve cells. In the Segel et al. study, the group designed an artificial system of transfer utilizing Peg10 as the packaging apparatus, while supplying envelope protein and a labeled cargo RNA[4]. As a result, they observed successful transfer of cargo RNA into a mouse neuroblast cell line strictly

employing Peg10 as the “virus” carrier. This result points to the functional capabilities of endogenous CA genes.

As emphasized, endogenous neuronal CA genes such as Arc, PNMA2, and Peg10 may harbor the potential to share genetic material between neurons. If true, these genes may provide valuable targets for therapies in neurological disorders. To recap, Arc is essential for memory formation; the recency of the Peg10 and PNMA2 discoveries has limited the research into their neurological associations, however, it is clear that they are, at minimum, associated with neurological disorders[21-23].

Goal

The goal of this research is to shed light on virus-like communication between neurons as a mechanism of synaptic plasticity to change the dominant thinking in neuronal communication research. Currently, the dominant way of thinking in neuronal communication is NT release facilitates communication. With new studies showing the ability of endogenous neuronal CA genes sharing genetic information between cells, there is the potential for a paradigm shift from only NT release to an additional CA release being a key factor in neuronal communication. If true, this vastly open new therapeutic targets for neurological disorders other than the typical NT receptor and NT targets. To start, we intend to understand what factors influence Arc release, and whether endogenous CA genes can facilitate infection within the context of a retrovirus, such as HIV.

Hypotheses

1) Arc will be released from cells in a similar pathway as HIV.

As a result of the data supporting Arc’s similarities to HIV, we believe that Arc functions similarly to retroviruses. Given the lack of data on the mechanisms to support this claim, we will

be investigating the retroviral nature of the Arc protein. Specifically, we will identify factors involved in release.

2) Arc, PNMA2, and Peg10 will facilitate infection in the context of HIV.

To examine the infectious nature of endogenous CA genes, we aim to test whether the CA domains of endogenous CA genes may facilitate a retroviral infection. We will alter HIV such that its CA domain is replaced by the CA domain of PNMA2, Peg10, and Arc, and Full-length Arc independently. Those HIV mutants will be tested on their ability to mature and infect naïve cells.

CHAPTER TWO

MATERIALS AND METHODS

Cell Lines

HEK 293ts and HeLa cell lines (ATCC) were cultured in Dulbecco's Modified Eagle Medium (DMEM) with 10% Fetal Bovine Serum, 1000U/mL Penicillin, and 1000U/mL Streptomycin.

Primary neurons were dissected from brains of Embryonic 17.5-18.5 mice, gifted from the lab of Makio Iwashima. The brains were dissected and neurons were cultured as previously described[24]. To create the hSYN-GFP plasmid, the CMV promoter in our pLVX-GFP plasmid was dropped out using restriction enzymes ClaI and XhoI, and the human synapsin sequence was amplified and inserted using the same restriction enzyme sites. Therefore, the hSYN-GFP plasmid has a green fluorescent protein gene driven by a neuron-specific hSYN promoter. To transduce neurons with hSYN-GFP, 293t cells were PEI transfected with hSYN-GFP as described below in a 10cm dish. Following transfection protocol, the viral supernatant was collected and spun at 5000xg overnight to pellet HIV-1 virus particles as previously described[25]. RT values were measured to quantify infectious units identical to the assay used in Fig. 7[26], and ~0.64 infectious units were added per well to the neurons on DIV7.

For immunostaining validations, primary neurons were plated on glass coverslips on DIV0. On DIV16, neurons were fixed with 3.7% Paraformaldehyde in 0.159M PIPES buffer for 12min. Coverslips were washed 3x with PBS. Neurons were permeabilized with 0.1% Triton for

5 minutes in homemade block solution: PBS with 10% donkey serum and 0.01% NaN_3. The following antibodies were used for staining in homemade block solution: Anti-NeuN 1:1000 (Sigma Aldrich: MAB377); Anti-Map2 1:500 (Santa Cruz Biotechnology: sc-74421); Anti-Prox1 1:200 (Novus Biologicals: NBP1-30045SS); Anti-Ctip2 1:500 (Abcam: ab18465). Primary antibodies were stained for 1 hour, then coverslips were washed 3x with PBS. Donkey Anti-mouse and Donkey Anti-rat secondary antibodies (Jackson ImmunoResearch) were used in block solution with proper controls to verify no crossover (Data not shown). Coverslips were washed 3x in PBS and mounted onto glass slides in Fluoro-Gel (EMS 17985-11) for imaging.

Polyethylenimine (PEI) Transfections

Transfections for stable expression or to produce lentiviral particles were both done by PEI transfection. 293ts are plated at 70% confluency. For stable expression, 100% of the plasmid of interest is added with PEI; to produce lentiviral particles 80% of the HIV-GAG or other plasmid of interest is added while 20% of VSV-g plasmid is added with PEI. These combinations sit at room temperature for 10 minutes before being applied to the cells then incubated overnight at 37˚C 5% CO2. In the morning, the media is removed and replaced with the desired media for the experiment.

Extracellular Vesicle Collection and Immunofluorescent Staining

293ts cells and primary neurons were treated overnight with 100nM Bafilomycin-A1 dissolved in DMSO or equal volume of DMSO. 293ts were treated the morning following PEI transfection when media is changed. Supernatant is collected the following day and put through a 0.45µm filter unit to removal cells and debris. Coverslips stored in 70% ethanol were moved into individual wells in a 24-well plate. Once coverslips were washed 3x with PBS, either 400µL of supernatant or PBS only was added. Plates were spun at 1200xg for 2hr. Following the spin,

supernatant was removed and EVs were fixed with 3.7% Paraformaldehyde in 0.159M PIPES buffer for 12min. Coverslips were washed 3x with PBS. Staining was done using a homemade block solution: PBS with 10% donkey serum and 0.01% NaN_3. EVs were stained with Rabbit Anti-Arc primary antibody 1:2000 (Synaptic Systems: 156003) while simultaneously permeabilized with 0.1% saponin in block solution for 1hr. Coverslips were washed 3x with PBS. EVs were then stained with Donkey anti-Rabbit secondary antibody. Coverslips were washed 3x in PBS and mounted onto glass slides in Fluoro-Gel (EMS 17985-11) for imaging.

Imaging and Extracellular Vesicle Analysis

Images were taken using a DeltaVision wide field fluorescent microscope. Twelve images were taken per coverslip. After deconvolution, images were analyzed using Imaris software (Bitplane). An algorithm was created around puncta of interest by creating surfaces in the Imaris surface function, as previously described[27, 28]. The algorithm will compute the total number of surfaces present per image for each specific condition. Algorithms remained consistent within experimental conditions.

Plasmids and Cloning

A plasmid containing the mArc sequence flanked by its 5' and 3' UTRs, the pLVX-mouse-IRSp53 plasmid, and a plasmid containing the PNMA2 sequence were shared by Jason Shepherd's Laboratory at the University of Utah. We cloned mArc along with its UTR sequences from the Shepherd plasmid into our lab's pLVX backbone using the Gibson Assembly Master Mix (NEB) to create the pLVX-mArc-UTR plasmid.

We cloned the HIV CA mutants into our NL4-3 HIV-1 plasmid with a firefly luciferase reporter in place of the *nef* gene. We identified unique restriction enzyme sites and chose the

closest to the CA domain. The results were a BssHII site ~70 base pairs upstream of the MA domain and an ApaI site within the NC domain (see Figure 5). The goal was to amplify three inserts, then connect them using a sewing PCR protocol. To do this, we developed primers to PCR three inserts that would result in ~30-45 base pairs of overlap with each other: one to copy the DNA sequence containing the first restriction site and the MA domain, a second to contain the new CA domain, and a third to copy the rest of the GAG DNA until the other restriction site is reached (see Figure 5). Once the three inserts are amplified, we run a sewing-like PCR to fuse the PCR products together. This involves rounds of ... without the addition of a primers, then during a rest at 10˚C, we add the primers and the run continues with a normal PCR protocol. The result is a combination of our three inserts as one long PCR product. That product is run on an agarose gel, extracted, digested with restriction enzymes, ligated into our digested NL4-3 HIV-1 backbone, and transformed in a recombinase A deficient competent bacterial cell line. The resulting backbone contained the CA domain you amplified i.e. PNMA2, Peg10, Arc in place of the HIV CA in frame within the HIV-1 genome. The human Peg10 sequence was a gift from Feng Zhang within the plasmid pCMV-HsPeg10rc3 (Addgene plasmid #174859). The rat Arc sequence was ordered as a geneblock fragment (Integrated DNA Technologies) and cloned from there on. All sequences were verified by sequencing analysis at the Genome Research Core at the University of Illinois at Chicago. The MHIV-mMA12CA plasmid was a generous gift from Masahiro Yamashita at Columbia University.

EV-Depleted Media

Place FBS in 38.5mL ultracentrifuge tubes (Beckman Coulter). Spin at 100,000g for 3 hours and sterile filter the supernatant through a 0.2µm filter unit. Combine with media to have 10% serum-containing media.

Supernatant is collected and filtered via a 0.45µm filter. Then supernatant is clarified at 300xg for 5 minutes at 4˚C, transferred to a fresh conical and spun at 3,000xg for 10 minutes at 4˚C[29]. Lastly supernatant is transferred to a 38.5mL ultracentrifuge tube (Beckman Coulter) and spun at 100,000xg for 2.5 hours at 4˚C. Supernatant is poured off and excess is carefully aspirated, while pellet is resuspended in 15µL NP-40 lysis buffer for western blot analysis.

Western Blots

293t cells were plated in 15cm culture dishes and transfected with the HIV plasmid of interest. 24 hours later, culture media was replaced with EV-depleted media either with or without 5µM Indinivir (protease inhibitor). 24 hours later, the supernatant was collected and stored at 4˚C overnight while the media was replenished with or without protease inhibitor, same as before. 24 hours later, the supernatant was collected and pooled with the first collection, filtered through a 0.45µm filter, and run through our EV ultracentrifugation protocol. Using our standard lab SDS-PAGE western blot protocol[27, 28], concentrated virus was lysed using an NP-40 lysis buffer (100mM of Tris pH 8.0, 1% NP-40, 150nM of NaCl) containing a protease inhibitor cocktail (Roche) for 30 minutes on ice. Samples are spun at 13,300 rpm for 5 minutes and supernatants were transferred to fresh tubes. Protein concentrations were assessed using the Pierce BCA protein assay kit (Thermo Scientific). Lysed supernatants were combined with non-reducing 6x SDS and denatured at 100˚C for 5 minutes. Equal protein amounts were added to 4-15% polyacrylamide gel (Bio-Rad) for SDS-PAGE. After separation, proteins were transferred to a nitrocellulose membrane (Bio-Rad) and stained overnight with primary antibodies diluted in a milk block solution 2.5g/50mL of TBST. Antibodies used in the western blots are as follows: Anti-HIV-1 p24 (Santa Cruz sc-69728); Anti-Arc HRP (Santa Cruz sc-17839); Anti-PNMA2

(Protein Tech 16445-1-AP); Anti-Peg10 (Santa Cruz sc-365675). The following morning, membranes are washed in TBST three times for 10 minutes, then incubated with a secondary antibody conjugated to HRP for 30 minutes diluted in the same milk block solution as before. Following the secondary antibody stain the membrane was washed three times for ten minutes each. Signal was detected using the SuperSignal West Femto Chemiluminescent Substrate (Thermo Fisher Scientific). Chemiluminescence was detected using the FluorchemE Imaging System (Protein Simple).

Infectivity Assay

293t cells in 10cm plates were PEI transfected by the plasmids of interest to produce lentiviral particles. Supernatant was collected and filtered through 0.45µm filter units (Millipore Sigma). Filtered viral supernatants were spun at 5000xg overnight based on a previously published protocol to pellet viral particles[25]. The following morning, supernatants were aspirated, and concentrated virus was resuspended in cold 10% FBS DMEM. Viral pellets were gently pipetted and placed on ice for 4 hours. 30µL of the resuspended pellets were added to 293ts in a 48-well plate and spun at 1200xg for 2 hour to facilitate infection, with or without 5µM Nevarapine (RT inhibitor) and 10µM Raltegravir (IN inhibitor). 48 hours a luciferase assay was conducted to measure luciferase activity in the 293ts that would be due to expression by the infecting plasmid. Add 100µL of 1x passive lysis buffer (Promega) to each well. Leave plate on the rocker for 40 minutes. Lift cells and transfer 40µL to a luciferase plate per well. Add 100µL firefly luciferase substrate (Promega) to each well of the luciferase plate. Luminescence (reactive light units) was quantified.

CHAPTER THREE

RESULTS

Quantification of Arc Release

Research investigating the mechanism of Arc release from cells is yet to be done and very few have detected Arc release. To examine Arc's release pathway from cells we will use our EV and immunofluorescent staining method followed by Imaris analysis (see Materials and Methods).

To detect Arc in the supernatant of 293t cells, we overexpressed the pLVX-mouse-Arc-UTR plasmid by PEI transfection. This plasmid consists of mArc flanked by its 5' and 3' UTRs. At the time of efficient expression, supernatant is collected and run through spins and filtrations to remove dead cells and debris. Next we spin the EVs onto coverslips, perform a formaldehyde fix, immunofluorescent stain for Arc, and quantify via confocal microscopy (see Materials and Methods). To validate the assay we add proper controls. In PBS and empty-plasmid transfection conditions, Arc release levels are low and considered background (Fig. 2). Whereas when Arc is transfected, levels of Arc in the supernatant are high. After validation, we decided to test whether Arc is released in a pathway similar to HIV. HIV buds off of the plasma membrane with the assistance of its MA and ENV domains. While the drug BAF inhibits vacuolar H^+-ATPase therefore preventing acidification of intracellular compartments and fusion of those compartments with the lysosome[30]. This prevents intracellular compartments like multivesicular bodies (MVBs), autophagosomes, late endosomes, and amphisomes from acidifying and fusing

with the lysosome. As a result, these intracellular compartments would be forced to fuse with the plasma membrane and extrude their contents into the extracellular space. Importantly, this would increase the release of contents within the intracellular compartments.

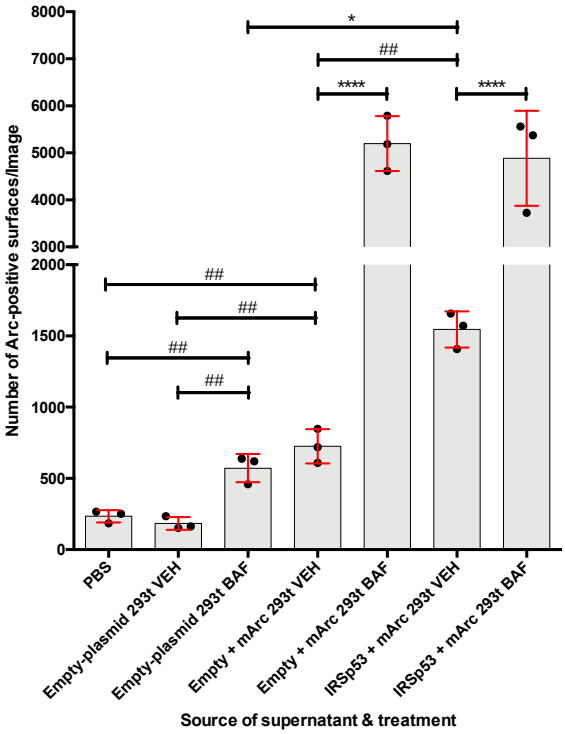

Figure 2. Quantification of Arc Protein in Cell Supernatant. A) In triplicate, 293t cells are PEI transfected with the following combinations of lentiviral plasmids: Empty plasmid; mouse Arc + Empty plasmid; mouse Arc + IRSp53. Cells are treated with 100nM Bafilomycin-A1 (BAF) or DMSO Vehicle (VEH) for 24hrs. Supernatant is collected following treatments, filtered through 0.45µm filter, and spun onto glass coverslips at 1200g for 2hrs. Samples are fixed with 3.7% PFA and stained with Arc primary in 0.1% saponin blocking buffer. Following secondary antibody stain, coverslips are mounted onto glass slides and imaged the following day on Wide-Field fluorescent microscope. Analysis of deconvolved images is performed using Imaris. Values are averaged from 12 images per sample. ## <0.01 by T-Test; * <0.05 or **** <0.0001 by Two-way ANOVA (Tukey HSD).

In our BAF treatment conditions, we consistently see increases in Arc release (Fig. 2). BAF treatment in empty plasmid transfected 293ts reveals a significant increase in released Arc compared to vehicle-treated empty plasmid transfected 293ts signifying BAF increases endogenous Arc release. In Arc transfected 293ts we see an increased release of Arc from

vehicle-treated versus empty plasmid transfected 293ts showing Arc transfection increases basal levels of Arc release (Fig. 2). When Arc-transfected 293ts are treated with BAF, we see a 5-fold increase in Arc release, significant by ANOVA (Fig. 2). BAF treatment significantly increases Arc release demonstrating that Arc exits cells via an intracellular compartment pathway. This intracellular pathway result suggests a few possible release pathways: MVBs or autophagic secretion.

An essential factor for HIV release from cells is the Insulin receptor substrate p53 (IRSp53)[31]. The IRSp53 protein contains an Inverted-BAR (I-BAR) domain. The BAR domain family creates membrane curvature. Specifically, I-BAR domains promote negative membrane curvature, as is necessary for viral budding. The negative membrane curvature from IRSp53's I-BAR domain enhances HIV release, and that HIV-GAG preferentially binds to negatively curved membranes caused by the I-BAR domain of IRSp53[31]. So we looked to see if IRSp53 influences Arc release in our assay. When we co-transfected Arc and IRSp53 into 293ts, we observed a significant increase in Arc release compared to Arc co-transfected with an empty plasmid (Fig. 2). Like HIV, the I-BAR domain containing protein, IRSp53, influences the Arc release mechanism.

The BAF treatment was repeated on hippocampal primary neurons collected and cultured as previously described[24] from E17.5-18.5 mouse brains. To validate that the majority of our cell population was neuronal, we immunofluorescently stained for NeuN, a nuclear mature neuronal marker (Fig. 3b); and Map-2, a dendritic mature neuronal marker (Fig. 3b). Both markers provide evidence that the majority of cells present are neuronal and mature. Additionally, we transduced lentiviral particles expressing GFP downstream from a synapsin-dependent promoter. Synapsins are essential neurotransmitter signaling proteins found mainly in

neurons. Therefore, if the lentiviral particles transduce these neurons, we would observe GFP signal (Fig. 3b). We successfully, eliminated most non-neuronal cell types while providing the neuronal cell types the proper environment to develop into mature neurons.

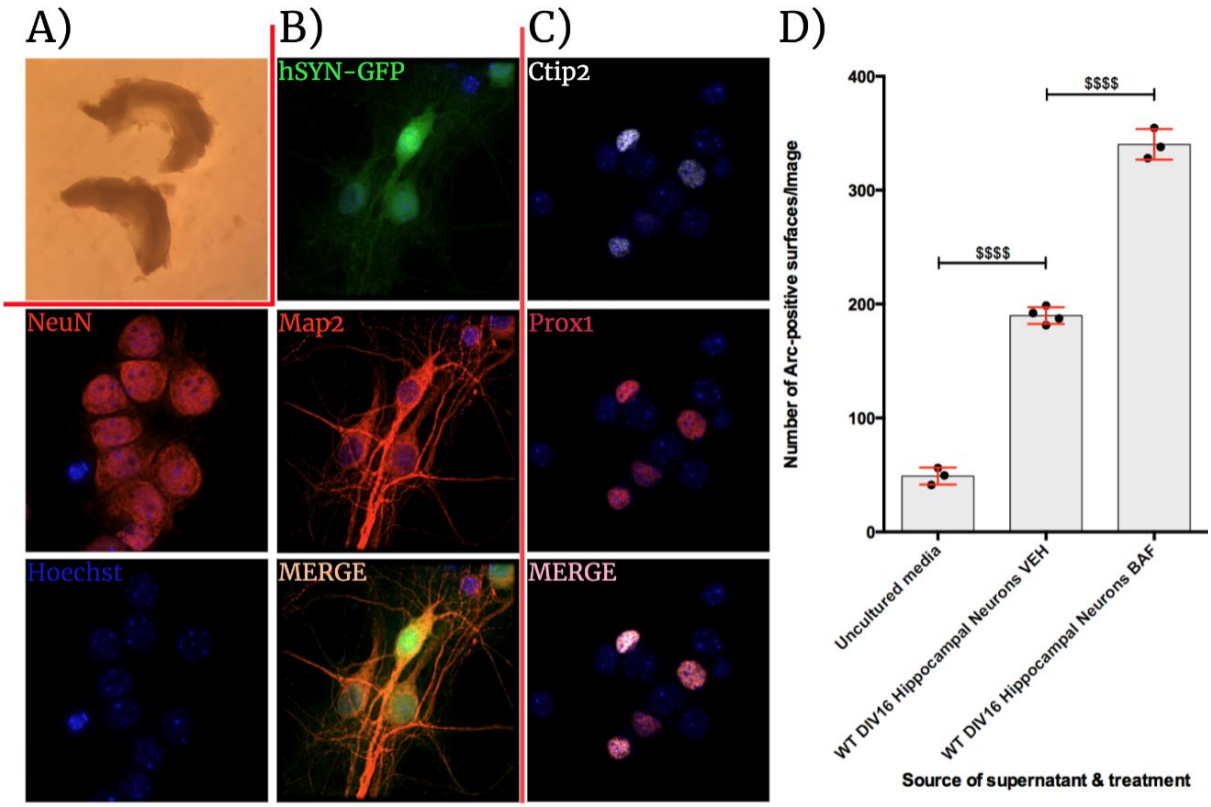

Figure 3. Primary Neuron Validation and Arc Release Experiments. A) Image of dissected right and left hippocampal lobes from a single mouse brain. B) From left to right, immunofluorescent stain of fixed DIV16 primary neurons with NeuN and Hoechst. Produced lentiviral particles with 293ts and transduced DIV7 primary neurons with plasmid containing Synapsin-dependent promoter upstream GFP tag (hSYN-GFP) and immunofluorescent stain with Map2 and Hoechst at DIV16 following fixation. C) Immunofluorescent stain of fixed DIV16 primary neurons with Ctip2, Prox1, and Hoechst. D) WT Primary mouse hippocampal neurons are cultured to DIV16. On DIV15, neurons are treated with DMSO Vehicle or Bafilomycin-A1 for 24hrs. Supernatant is collected following treatments and processed as in Fig. 2. Images are collected and analyzed as in Fig 2. The points represent replicates of supernatant from a single dish of cells. Values are averaged from 12 images per sample. $$$$ <0.0001 by one-way ANOVA (Tukey HSD).

Besides identifying the brain region the cells originated (Fig. 3a), we attempted to identify the cell type. To examine the cell type, we conducted fluorescent imaging for following

immunostaining for hippocampal markers Ctip2 and Prox1. Firstly, COUP-TF-Interacting protein 2 (Ctip2), a cortical and hippocampal neuronal marker was present (Fig. 3c). Secondly, immunostaining with a dentate granule cell-specific hippocampal neuronal marker, Prospero Homeobox 1 (Prox1)[32] showed clear and present expression. The consistent colocalization between Ctip2 and Prox1 provides evidence to hippocampal specific cell types as well as the presence of dentate granule neurons. Based on this stain, the majority of neurons in our sample were granule neurons.

Measuring solely endogenous Arc release from the primary neurons, we observed significantly increased levels of Arc in the supernatant above background with vehicle treatment and then additionally following BAF treatment (Fig. 3d). So BAF influences Arc release in an exogenous 293t system and endogenous primary neuron system.

Endogenous Neuronal CA Genes

A new direction of investigation has begun examining the capsid forming abilities of endogenous neuronal genes evolved from LTR retrotransposons. This includes members of the PNMA and MART families as well as Arc. Many of the genes within these families have been observed forming capsid-like structures[3, 4]. Given their relationships with retroviruses along with their publicized ability to form capsid-like structures, we decided to test whether the CA domains of these endogenous neuronal CA genes could replace HIV-1 capsid to facilitate infection by HIV-1. We utilized the HIV capsid's assembly mechanism whereby the capsid is formed following budding into an enveloped vesicle and proteolytic cleavage of the GAG domain. In this way, CA protomers are free to oligomerize in the comfort of an enveloped vesicular environment. Using this mechanism, our goal is to swap the HIV CA domain for the

CA domain of the endogenous neuronal CA genes highlighted in this thesis: PNMA2, Peg10, and Arc (Fig. 4).

To create the endogenous CA mutant plasmids, we first identified the CA domains within the DNA sequence. A bioinformatic tool known as HHpred[33, 34] has the ability to pinpoint structural similarities between amino acid sequences.

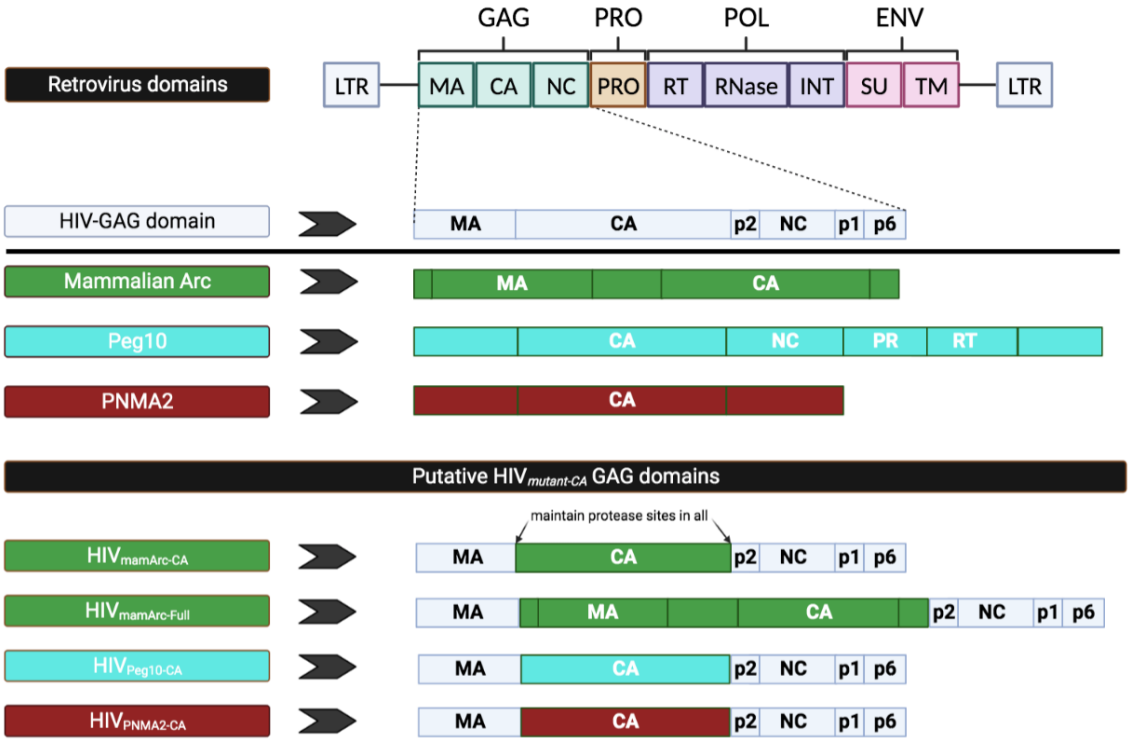

Figure 4. Putative HIV CA Mutant Clone Schematic. Isolating the GAG domains of the retrovirus domains taken from Figure 1. Aligning to the putative[1] domains of mammalian Arc, Peg10, PNMA2 (not to scale). Showing the final HIV CA mutant domains after cloning in the new CA domains.

Using this tool, the well-established CA domain regions of HIV, MLV, and Ty3 were structurally aligned to the amino acid sequences of PNMA2, Peg10, and Arc. Additionally, aligning the established Arc CA domain amino acid sequence to the PNMA2 and Peg10 amino

acid sequences provided rigorous identification of putative CA domains in the endogenous CA genes.

To develop the HIV CA mutant plasmids we used a combination of multiple cloning techniques. First, we conducted three PCR reactions one upstream of the CA domain, a second downstream of the CA domain, and a third on the new CA domain, as seen in Fig. 5.

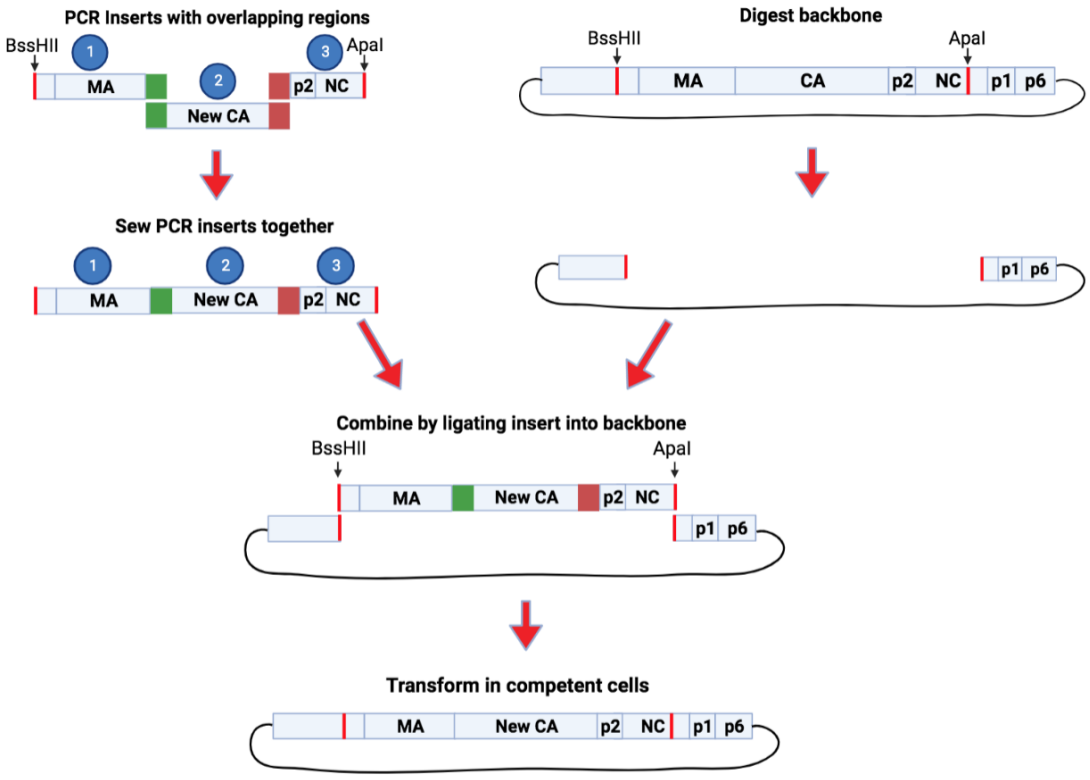

Figure 5. HIV CA Mutant Cloning Method. For detailed methodology, see Plasmids and Cloning section in Materials and Methods

The upstream section contained a 5' restriction enzyme site BssHII while the 3' end had over 30 base pairs of overlap with the CA domain section. The CA domain section has over 30 base pairs of overlap in both the 5' and 3' directions while amplifying the CA domain sequence of PNMA2, Peg10, or Arc. The third segment contains a 30 base pair overlap with the CA domain fragment on the 5' end and has a restriction enzyme site on the 3' end, ApaI. Once all three were

amplified, we ran a specialized amplification program to sew the three DNA fragments together into one long amplification product. Following amplification, we ligate our product into our digested backbone and transformed it into a competent bacterial cell line (Fig. 4). Once completed, we have HIV GAG plasmids containing mutant CA domains from endogenous neuronal CA genes.

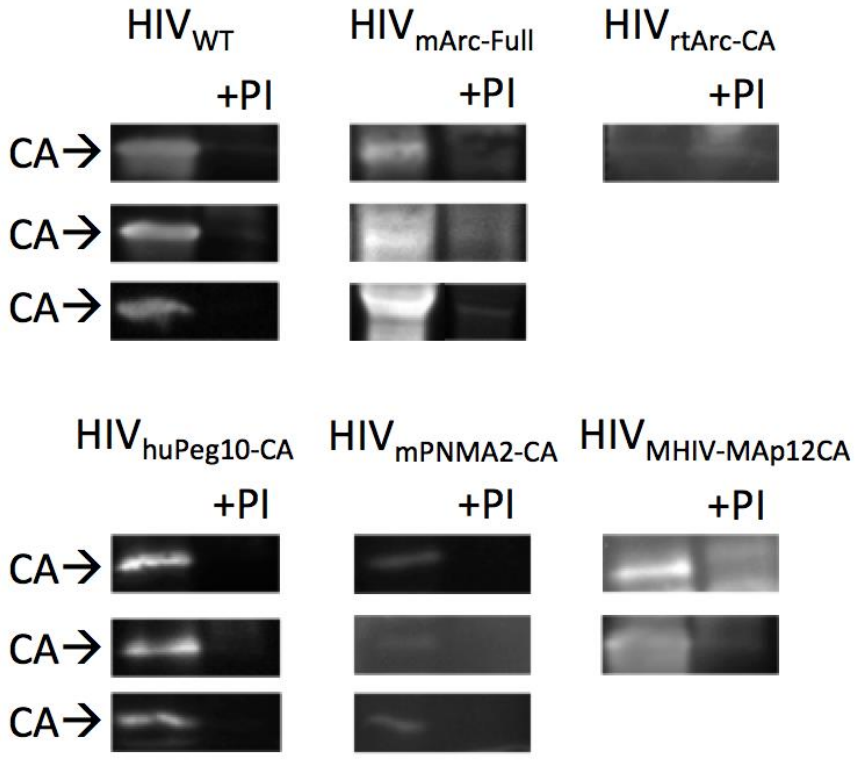

Figure 6. Western Blots of HIV CA Mutant Maturation. Each image represents an individual experiment. Cleavage tests for HIV WT, HIV mArc-Full, HIV huPeg10-CA, HIV mPNMA2-CA. Each HIV plasmid contains a protease inhibitor (PI) treatment group.

Before testing the infectivity of the HIV CA mutants, we had to verify that the HIV CA mutants were maturing as HIV wild type CAs mature. Maturation is the process of the newly translated HIV GAG polyprotein exiting the cell in an enveloped vesicle and undergoing proteolytic cleavage into its respective functional domains. This proteolytic cleavage is facilitated by the PRO domain and is an essential part of HIV virus development. The PRO

domain cleaves the HIV polyprotein at recognized amino acid sequences in between each domain[35, 36]. When designing the HIV CA mutants, the protease cleavage sites remain flanking the newly inserted CA domain. However, we must validate that cleavage of the GAG components is occurring, as it occurs in wild-type HIV maturation, in our HIV system expressing ancestral genes. We ran western blots on supernatant collected from 293t cells expressing components to produce virus particles along with our HIV CA mutants (Fig. 6).

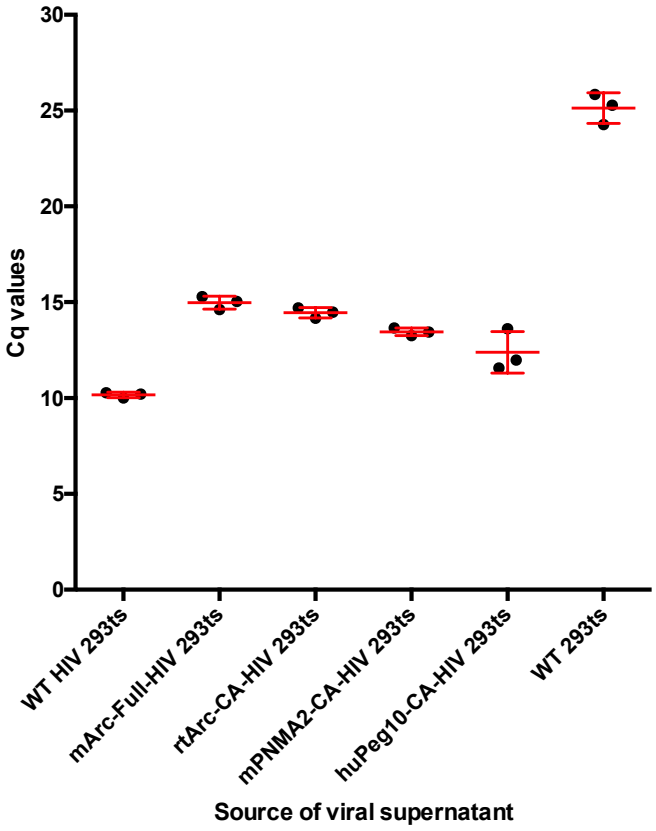

Figure 7. RT activity in the supernatant of 293ts transfected with HIV CA mutants. Each dot represents an individual transfection for each plasmid in 24-well format. Ran each well in triplicate and took the averages for each data point.

Virus particles in the supernatant were collected and concentrated as previously described[29]. Samples were stained with antibodies specific to the protein within the CA domain. Screenshots were taken at the locations where the properly cleaved CA domain is expected based on the

SDS-PAGE ladder spread (Fig. 6). Importantly, we do not see cleaved CA domain when we incubate the cells with a protease inhibitor (PI) (Fig. 6). Our mutants appear to be cleaved normally, so we can move onto further validation steps (Fig. 6).

Another aspect of HIV infectivity is the functionality of it RT enzyme. So, after infecting cells with our HIV CA mutants, we collect the supernatant and run a well-established RT assay[(26)]. Beyond the ability of the western blot to show the CA cleavage, an RT assay informs us of the cleavage of the other constituents of the GAG polyprotein in our large HIV CA mutant complex. In these experiments, we repeatedly see our HIV CA mutants with RT values that are higher, or less effective, than WT HIV levels however, the values are two times closer to WT HIV levels than to background levels (Fig. 7). Given this result, we can move forward with the HIV CA mutant examinations.

Considering the HIV CA mutants show effective GAG polyprotein processing, we decided to determine if they could facilitate HIV-1 infection. To test this, the HIV CA mutants are expressed in 293t cells alongside a plasmid containing the envelope gene, Vsv-G. The supernatant is harvested from these producer 293t cells, filtered, and spinoculated onto HeLa cells. Infectivity is measured 48 hours following spinoculation via a luciferase assay using the firefly luciferase reporter within the HIV CA mutant's plasmid. The presence of luciferase signal in HeLa cells would present proof of the virus mediating HIV-1 infectivity; for, the luciferase signal is the result of successful integration of the viral RNA into the HeLa cell genome.

In our first infection experiment, we collected and filtered virus from 6cm plates of producer cells. The concentrated virus was assessed for RT levels, as in Fig. 7, to determine the infectious units per sample. Using this information we added equal levels of infectious units to each well of HeLa cells (Fig. 8A). Its clear that WT HIV is infectious, validating the experiment,

and that there may be some infection by the mArc full length-HIV mutant as well as the huPeg10-CA-HIV mutant. For validation, we ran a follow-up experiment including additional treatments and a previously published mutant known as MHIV-mMA12CA[37].

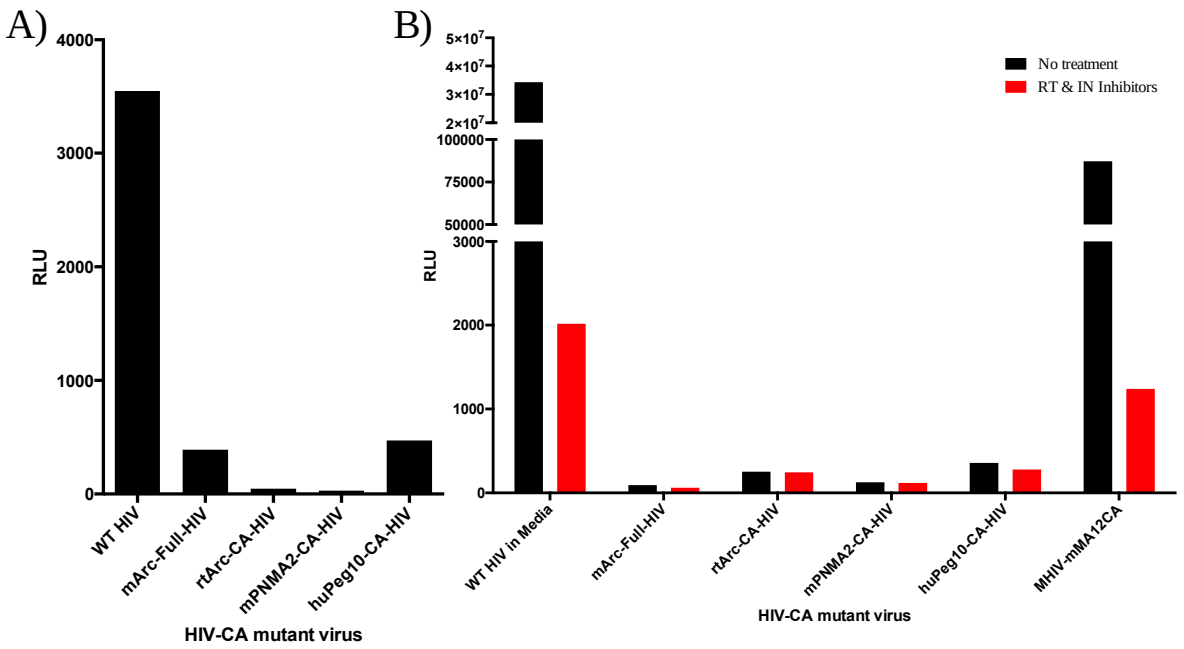

Figure 8. Infectivity of HIV-CA Mutants in HeLa Cells. A) Relative light units from a firefly luciferase infectivity assay, when equal concentrations of viral vector were added to the HeLa cells based off of an RT assay. Samples were run in duplicate and average values of the duplicates are presented n=1 per HIV CA mutant. B) Relative light units from a firefly luciferase infectivity assay when ~30% of the virus collected from a 10cm plate was added to the HeLa cells following concentration procedure. Black bars represent cells that were untreated when they received the virus. Red bars represent cells that were treated with both RT (Nevarapine) and IN (Raltegravir) inhibitors. Each bar represents a single replicate.

MHIV-mMA12CA replaced the MA and CA domains of HIV with the MA, P12, and CA domains of MLV. This mutant was previously shown to be ~100x less infectious than WT HIV[37], and our experiment shows MHIV-mMA12CA is ~400x less infectious than WT HIV (Fig. 8B). The additional treatments in red represent RT and IN inhibitors that should completely block infection and therefore represent any present background signal. The high levels of infection observed in the inhibitor conditions could be attributed to spillover due to their

locations on the luciferase plate adjacent to the wells with the high signal. Given the low signal and lack of difference in luciferase signal upon treatment of infection inhibitors, we do not observe clear infection of HIV being facilitated by the HIV CA mutants in our system.

CHAPTER FOUR

DISCUSSION

In this thesis, we revealed a potential method of quantification and mechanism of exit for the Arc protein from cells. Additionally, we illustrated the lack of infectivity of endogenous neuronal capsid genes in the context of an artificial retroviral infection. Taken together, these findings do not dismiss the potential for a novel neuronal communication paradigm that impacts memory but rather imply the difficulty of studying such a complicated topic.

Most research regarding the Arc protein has focused on its memory involvement[13]. Until recently, a group was attempting to purify the Arc protein and noticed it continuously crash out of solution. After observing the precipitate under electron microscopy they observed capsid like structures[2]. Rapidly following this discovery two groups corroborated the capsid structure and took the viral functionality one step further[3, 38]. One group, using Drosophila observed trafficking of Arc across the neuromuscular junction (from axon to muscle) via EVs, without which negatively impacted synaptic plasticity[38]. The other group studied Arc in mammalian cells and quantified trafficking between immortalized cells as well as primary cultured rodent neurons[3]. They collected supernatant from the wild-type mouse primary hippocampal neurons and transferred it to the Arc-knockout neurons' plates; the result was clear transfer of Arc protein and mRNA into the Arc-knockout neurons via EVs. Further, the newly-transferred RNA was accessible for activity-dependent translation[3]. Collectively, these are the two major findings that support the study of the viral nature of the Arc protein and impress the importance of its viral

functions in memory. For this reason, our first aim focused on investigating the mechanism of release of the Arc protein to better understand how it is being trafficked between cells.

Cellular contents are canonically released through a few main methods: Secretory autophagy, exosome secretion, and a combination of the two known as amphisome secretion. Retroviral contents, like HIV, are released by directly budding from the plasma membrane. In secretory autophagy, a signaling pathway is triggered resulting in the formation of an autophagophore, or a double-membraned structure, to surround specific contents to be exported from the cell. In exosome secretion, contents gather in vesicles within an intercellular compartment known as a multivesicular body (MVB). Upon fusion of the MVB with the plasma membrane, the vesicles within the MVB are excreted from the cell. Amphisomes are created when endosomes fuse with autophagosomes, and their secretion occurs when an amphisome fuses with the plasma membrane. In our experiment, Bafilomycin-A1 treatment is used to specifically target the intracellular compartments. Bafilomycin-A1 (BAF) is an inhibitor of the vacuolar ATPase that is responsible for acidifying the intracellular environments within vesicular compartments such as lysosomes and endosomes. Following the addition of BAF on cells, the acidification of the lysosomes is blocked and therefore fusion of intracellular vesicles with the lysosomes does not occur; as a result, there is an increased release of exosomes[39]. There is also evidence supporting an increased shuttling of other intracellular compartments such as autophagophores and amphisomes following BAF treatment[40]. However, HIV which traffics and inserts itself into the plasma membrane using the myristyl sequence on the N-terminal tail of the MA domain, where the GAG domains will line the membrane, package RNA, and bud directly off the plasma membrane. This budding directly off the plasma membrane is thought to be unaffected by BAF treatment as acidification of intracellular compartments will not affect

plasma membrane budding. Considering how BAF treatment affects different pathways of release, we would expect to see no change in Arc release with BAF treatment if Arc is release in a similar pathway as HIV. Alternatively, if Arc is released via an intracellular compartment we would observe an increase in Arc release following BAF treatment. Our immunofluorescent EV quantification experiments revealed a significant increase in Arc release following BAF treatment in both 293ts and primary neurons. Thus, our data suggests Arc is released from an intracellular compartment.

On the other hand, overexpression of IRSp53 resulted in significant increases in Arc release in our experiments (Fig. 2). IRSp53 is an I-BAR domain containing protein. I-BAR domains are known for inducing negative membrane curvature for budding vesicles. A recent publication found that full HIV-1 assembly requires IRSp53[(31)]. Considering our findings that Arc release increased with IRSp53 overexpression, one might postulate that Arc is budding directly off the plasma membrane, like HIV-1. However, increased Arc release due to IRSp53 overexpression was only observed when Arc was overexpressed in 293t cells. It's possible that the increase in Arc release is due to location bias along with significant negative membrane curvature events. Not to mention, IRSp53 interacts with WASP family verprolin homology (WAVE) isoform 1[(41)], which has been shown directly interacting with Arc[(19)], and Arc itself has been shown to induce negative membrane curvature[(42)]. This common interaction could indirectly link Arc and IRSp53 in a negative membrane curvature event and lead to release due to overexpression of both Arc and IRSp53. More research is needed using controlled levels of both proteins to identify the direct pathway by which Arc is released.

Future experiments and controls will be required to focus on specific pathways using targeted knockdowns and other drugs to block release pathways, like calpeptin to block

microvesicle release. To validate that our increased Arc release is a result of BAF and not due to BAF's side-effect preventing endocytosis, an additional experiment using dynasore to inhibit endocytosis is required. Regardless, it is clear that Arc release is occurring in a biologically relevant cell type, and we have begun to understand how that release could lead to a transfer of cognitive information. Future studies must investigate whether release of Arc is due to an autophagic event, a controlled budding, or an exosomal event.

To this day, the infectious nature of endogenous capsid genes remains a minimally researched field. A seminal paper uncovered endogenous LTR retrotransposon offspring that share homology with retroviruses. These were grouped into five families of which we've discussed as the Arc, PNMA, and MART families[1]. When examining members of each family, a few publications have identified that the majority of genes in the three discussed families are expressed in the brain[5, 21-23, 43, 44]. Moreover, a few members of the MART family were detected forming capsid-like structures, as were several members of the PNMA family[4]. At least one member of each family has been noted exiting cells[3, 4, 21]. Two groups recently discovered Arc is capable of trafficking between cells[3, 38], and even more recently Peg10 was shown to be capable of functioning as a packaging plasmid in an artificial system for trafficking materials between cells[4].

The trafficking of these capsid genes is significant given their importance in neuropathology. During Angelman Syndrome pathology, Peg10 is present in ~20% of EVs compared to ~6% in healthy conditions, and assessing the transcriptome of cells during Peg10 knockdown revealed Peg10 is important in neuronal migration[23]. Together, one could postulate that Peg10 traffics between neurons to transfer homeostatic migratory signals. This would open a wide range of targets for neurodegenerative diseases to assist neuronal migration. Another

member of the MART family, named MART4 has had effects related to cognition. In a paper published in 2015, a group achieved a complete knockout of the MART4 gene in mice and observed significantly worse performance on a Y-maze task, and delayed noradrenaline recovery compared to dopamine, following high potassium perfusion of the prefrontal cortex[(44)]. Considering, there are at least 6 other members of the MART family expressed in the brain that share homology with retroviruses, the field is itching for an endogenous capsid neuronal communication story in the brain. All genes in the PNMA family are expressed in the brain as they are specifically related to neuropathology(22). The PNMA, or Paraneoplastic MA antigen, family is a collection of endogenous proteins recognized by autoantibodies released from tumors[(22)]. Several members of the PNMA family share homology with retroviruses[(1)], and six members have been observed forming capsid-like structures[(4)]. A paper published in 2022 chose to investigate Peg10 and PNMA2 of the many endogenous neuronal capsid genes as a result of a transcriptome screening identifying Peg10 and PNMA2 as highly expressed in embryonic and fetal rodent brains. For this reason, along with the findings that Peg10 and Arc can transfer RNA and all genes share homology to retroviruses, we chose to focus this thesis project on *Peg10*, *PNMA2*, and *Arc*.

Given our position as an HIV research lab, we decided to utilize the available tools that our lab specializes in to examine the viral nature of the endogenous capsid genes. Utilizing the method of capsid assembly by HIV-1 where, following viral budding into an enveloped vesicle, the capsid domain is cleaved from the other GAG components and assembles into the capsid structure. We decided to replace the capsid domain of HIV with the capsid domains of endogenous neuronal genes expressed in the brain that evolved from LTR retrotransposons (Fig 2A). By inserting these endogenous capsid domains into the capsid domain of HIV (Fig. 3), we

predicted the mutants would mature properly and facilitate infection. We found that maintaining the protease sites flanking the capsid resulted in successful protease cleavage and maturation of the mutants (Figs. 5&6). We found that viruses produced by our HIV CA mutants maintained RT levels that were consistently closer to wild-type HIV RT levels than background RT levels (Fig. 7). In our first infectivity experiment, where we added equal amounts of infectious units of virus from each sample onto the cells, and measured a strong infectious signal from our wild-type virus. It appeared that the full-length mouse Arc and the huPeg10 HIV CA mutants may be showing some infectivity (Fig. 8A). To validate this finding we repeat the experiment using infectivity inhibitors to gather a clear background, and look for distinct differences between infectivity levels with and without inhibitors. We quantify clear differences between infectivity levels when using the inhibitors only in the wild-type HIV and the previously published MHIV-mMA12CA[37] (Fig. 8). The HIV CA mutants did not have discernable differences in infectivity levels with and without inhibitors (Fig. 8). These findings did not agree with our hypothesis that the endogenous capsid genes *Arc*, *PNMA2* and *Peg10* will facilitate infection in the context of HIV; however, we were able to prove that the HIV CA mutants matured properly and that they retained some of HIVs RT activity.

Importantly, HIV infectivity is a highly controlled process. However, if instead of testing infectivity, we hijacked HIV's capsid assembly using HIV CA mutants to deliver capsids to target cells. To this day Arc capsids have not been identified in cells, nor have Peg10 or PNMA2 capsids. So, we could instead use our HIV CA mutant system to assemble the capsids, fuse into the target cells, and study their location of delivery, interacting partners, and cellular responses. We could open a new field of study where we track the delivery of endogenous capsid genes into neurons. We believe our project is a bridge into research that uses well-known viral assembly

mechanisms to transport mutants into target cells to study these minimally examined endogenous capsid genes.

The future of Arc research can utilize this method to deliver Arc capsids into neurons, and track the behavior of the capsid and cellular effects. This would advance the viral side of Arc research tremendously. The result of those experiments may be the key to deciphering how Arc fits into neuronal communication. It is well established that Arc is essential for long-term memory formation[12, 13]. As it stands neuronal communication occurs electrically, through a physically connection, or chemically, using neurotransmitters across a synapse. Arc is known to affect chemical communication by removing AMPARs from the plasma membrane of the dendrites[9]. Now, we know Arc forms capsids, exits cells, houses RNA, and delivers RNA, and the target cell translates the RNA[3], like viruses. The idea being, if Arc is known to function like a virus in trafficking functional RNA between cells, and is also essential for long-term memory formation, we may be discussing a new paradigm of neuronal communication. So, will the next step in neuroscience be one towards endogenous capsid genes? I hope so. For, if the capsid genes are trafficking information between neurons that's imperative to cognitive functions, we may be opening an entirely untouched realm of therapeutic targets for neurological disorders.

REFERENCE LIST

1. Campillos M, Doerks T, Shah PK, Bork P. Computational characterization of multiple Gag-like human proteins. Trends Genet. 2006;22(11):585-9.

2. Myrum C, Baumann A, Bustad HJ, Flydal MI, Mariaule V, Alvira S, et al. Arc is a flexible modular protein capable of reversible self-oligomerization. Biochem J. 2015;468(1):145-58.

3. Pastuzyn ED, Day CE, Kearns RB, Kyrke-Smith M, Taibi AV, McCormick J, et al. The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. Cell. 2018;172(1-2):275-88 e18.

4. Segel M, Lash B, Song J, Ladha A, Liu CC, Jin X, et al. Mammalian retrovirus-like protein PEG10 packages its own mRNA and can be pseudotyped for mRNA delivery. Science. 2021;373(6557):882-9.

5. Lyford GL, Yamagata K, Kaufmann WE, Barnes CA, Sanders LK, Copeland NG, et al. Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. Neuron. 1995;14(2):433-45.

6. Link W, Konietzko U, Kauselmann G, Krug M, Schwanke B, Frey U, et al. Somatodendritic expression of an immediate early gene is regulated by synaptic activity. Proc Natl Acad Sci U S A. 1995;92(12):5734-8.

7. Dynes JL, Steward O. Arc mRNA docks precisely at the base of individual dendritic spines indicating the existence of a specialized microdomain for synapse-specific mRNA translation. J Comp Neurol. 2012;520(14):3105-19.

8. Nair RR, Patil S, Tiron A, Kanhema T, Panja D, Schiro L, et al. Dynamic Arc SUMOylation and Selective Interaction with F-Actin-Binding Protein Drebrin A in LTP Consolidation In Vivo. Front Synaptic Neurosci. 2017;9:8.

9. Chowdhury S, Shepherd JD, Okuno H, Lyford G, Petralia RS, Plath N, et al. Arc/Arg3.1 interacts with the endocytic machinery to regulate AMPA receptor trafficking. Neuron. 2006;52(3):445-59.

10. DaSilva LL, Wall MJ, L PdA, Wauters SC, Januario YC, Muller J, et al. Activity-Regulated Cytoskeleton-Associated Protein Controls AMPAR Endocytosis through a Direct Interaction with Clathrin-Adaptor Protein 2. eNeuro. 2016;3(3).

11. Wang G, Gilbert J, Man HY. AMPA receptor trafficking in homeostatic synaptic plasticity: functional molecules and signaling cascades. Neural Plast. 2012;2012:825364.

12. Guzowski JF, Lyford GL, Stevenson GD, Houston FP, McGaugh JL, Worley PF, et al. Inhibition of activity-dependent arc protein expression in the rat hippocampus impairs the maintenance of long-term potentiation and the consolidation of long-term memory. J Neurosci. 2000;20(11):3993-4001.

13. Plath N, Ohana O, Dammermann B, Errington ML, Schmitz D, Gross C, et al. Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. Neuron. 2006;52(3):437-44.

14. Niere F, Wilkerson JR, Huber KM. Evidence for a fragile X mental retardation protein-mediated translational switch in metabotropic glutamate receptor-triggered Arc translation and long-term depression. J Neurosci. 2012;32(17):5924-36.

15. Greer PL, Hanayama R, Bloodgood BL, Mardinly AR, Lipton DM, Flavell SW, et al. The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. Cell. 2010;140(5):704-16.

16. Wu J, Petralia RS, Kurushima H, Patel H, Jung MY, Volk L, et al. Arc/Arg3.1 regulates an endosomal pathway essential for activity-dependent beta-amyloid generation. Cell. 2011;147(3):615-28.

17. Alhowikan AM. Activity-Regulated Cytoskeleton-Associated Protein Dysfunction May Contribute to Memory Disorder and Earlier Detection of Autism Spectrum Disorders. Med Princ Pract. 2016;25(4):350-4.

18. Erlendsson S, Morado DR, Cullen HB, Feschotte C, Shepherd JD, Briggs JAG. Structures of virus-like capsids formed by the Drosophila neuronal Arc proteins. Nat Neurosci. 2020;23(2):172-5.

19. Zhang W, Wu J, Ward MD, Yang S, Chuang YA, Xiao M, et al. Structural basis of arc binding to synaptic proteins: implications for cognitive disease. Neuron. 2015;86(2):490-500.

20. Hallin EI, Eriksen MS, Baryshnikov S, Nikolaienko O, Grodem S, Hosokawa T, et al. Structure of monomeric full-length ARC sheds light on molecular flexibility, protein interactions, and functional modalities. J Neurochem. 2018;147(3):323-43.

21. Pinson MR, Chung DD, Mahnke AH, Salem NA, Osorio D, Nair V, et al. Gag-like proteins: Novel mediators of prenatal alcohol exposure in neural development. Alcohol Clin Exp Res. 2022;46(4):556-69.

22. Schuller M, Jenne D, Voltz R. The human PNMA family: novel neuronal proteins implicated in paraneoplastic neurological disease. J Neuroimmunol. 2005;169(1-2):172-6.

23. Pandya NJ, Wang C, Costa V, Lopatta P, Meier S, Zampeta FI, et al. Secreted retrovirus-like GAG-domain-containing protein PEG10 is regulated by UBE3A and is involved in Angelman syndrome pathophysiology. Cell Rep Med. 2021;2(8):100360.

24. Moutin E, Hemonnot AL, Seube V, Linck N, Rassendren F, Perroy J, et al. Procedures for Culturing and Genetically Manipulating Murine Hippocampal Postnatal Neurons. Front Synaptic Neurosci. 2020;12:19.

25. Qing E, Hantak MP, Galpalli GG, Gallagher T. Evaluating MERS-CoV Entry Pathways. Methods Mol Biol. 2020;2099:9-20.

26. Vermeire J, Naessens E, Vanderstraeten H, Landi A, Iannucci V, Van Nuffel A, et al. Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. PLoS One. 2012;7(12):e50859.

27. Burbidge K, Zwikelmaier V, Cook B, Long MM, Balva B, Lonigro M, et al. Cargo and cell-specific differences in extracellular vesicle populations identified by multiplexed immunofluorescent analysis. J Extracell Vesicles. 2020;9(1).

28. Dharan A, Bachmann N, Talley S, Zwikelmaier V, Campbell EM. Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. Nat Microbiol. 2020;5(9):1088-95.

29. Qing E, Kicmal T, Kumar B, Hawkins GM, Timm E, Perlman S, et al. Dynamics of SARS-CoV-2 Spike Proteins in Cell Entry: Control Elements in the Amino-Terminal Domains. mBio. 2021;12(4):e0159021.

30. Yamamoto A, Tagawa Y, Yoshimori T, Moriyama Y, Masaki R, Tashiro Y. Bafilomycin A1 prevents maturation of autophagic vacuoles by inhibiting fusion between autophagosomes and lysosomes in rat hepatoma cell line, H-4-II-E cells. Cell Struct Funct. 1998;23(1):33-42.

31. Inamdar K, Tsai FC, Dibsy R, de Poret A, Manzi J, Merida P, et al. Full assembly of HIV-1 particles requires assistance of the membrane curvature factor IRSp53. Elife. 2021;10.

32. Iwano T, Masuda A, Kiyonari H, Enomoto H, Matsuzaki F. Prox1 postmitotically defines dentate gyrus cells by specifying granule cell identity over CA3 pyramidal cell fate in the hippocampus. Development. 2012;139(16):3051-62.

33. Gabler F, Nam SZ, Till S, Mirdita M, Steinegger M, Soding J, et al. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. Curr Protoc Bioinformatics. 2020;72(1):e108.

34. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, et al. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol. 2018;430(15):2237-43.

35. Larrouy L, Chazallon C, Landman R, Capitant C, Peytavin G, Collin G, et al. Gag mutations can impact virological response to dual-boosted protease inhibitor combinations in antiretroviral-naive HIV-infected patients. Antimicrob Agents Chemother. 2010;54(7):2910-9.

36. Pettit SC, Sanchez R, Smith T, Wehbie R, Derse D, Swanstrom R. HIV type 1 protease inhibitors fail to inhibit HTLV-I Gag processing in infected cells. AIDS Res Hum Retroviruses. 1998;14(11):1007-14.

37. Yamashita M, Emerman M. Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. J Virol. 2004;78(11):5670-8.

38. Ashley J, Cordy B, Lucia D, Fradkin LG, Budnik V, Thomson T. Retrovirus-like Gag Protein Arc1 Binds RNA and Traffics across Synaptic Boutons. Cell. 2018;172(1-2):262-74 e11.

39. Villarroya-Beltri C, Baixauli F, Mittelbrunn M, Fernandez-Delgado I, Torralba D, Moreno-Gonzalo O, et al. ISGylation controls exosome secretion by promoting lysosomal degradation of MVB proteins. Nat Commun. 2016;7:13588.

40. Minakaki G, Menges S, Kittel A, Emmanouilidou E, Schaeffner I, Barkovits K, et al. Autophagy inhibition promotes SNCA/alpha-synuclein release and transfer via extracellular vesicles with a hybrid autophagosome-exosome-like phenotype. Autophagy. 2018;14(1):98-119.

41. Goh WI, Lim KB, Sudhaharan T, Sem KP, Bu W, Chou AM, et al. mDia1 and WAVE2 proteins interact directly with IRSp53 in filopodia and are involved in filopodium formation. J Biol Chem. 2012;287(7):4702-14.

42. Hedde PN, Malacrida L, Barylko B, Binns DD, Albanesi JP, Jameson DM. Membrane Remodeling by Arc/Arg3.1. Front Mol Biosci. 2021;8:630625.

43. Brandt J, Schrauth S, Veith AM, Froschauer A, Haneke T, Schultheis C, et al. Transposable elements as a source of genetic innovation: expression and evolution of a family of retrotransposon-derived neogenes in mammals. Gene. 2005;345(1):101-11.

44. Irie M, Yoshikawa M, Ono R, Iwafune H, Furuse T, Yamada I, et al. Cognitive Function Related to the Sirh11/Zcchc16 Gene Acquired from an LTR Retrotransposon in Eutherians. PLoS Genet. 2015;11(9):e1005521.

VITA

Michael Long was born in Park Ridge, IL. He received his High-School education from Adlai E. Stevenson located in Lincolnshire, IL. After graduating, Michael continued his education at the University of Kansas, where he worked as a tutor and spent his final year conducting research on Drosophila *virilis* under the mentorship of Dr. Justin Blumenstiel. In May of 2017, Mr. Long received his Bachelor's degree in Human Biology with a concentration in Psychology.

In the fall of 2017, Michael began his studies in the Neuroscience Graduate Program at Loyola University Chicago. Michael chose to investigate the Arc protein in the laboratory of Dr. Ed Campbell. Michael will continue his passion for educating as a scientific communicator and writer.