Drew University

College of Liberal Arts

UNDERSTANDING THE INNATE IMMUNE RESPONSE TO SINGLE-STRANDED DNA AND PARVOVIRUSES

A Thesis in Biochemistry and Molecular Biology

By

Marina Hahn

Advisor: Dr. Brianne Barker

Submitted in Partial Fulfillment of the

Requirements for a Degree in Bachelor of Science

with Specialized Honors in Biochemistry and Molecular Biology

May 2020

Abstract

The innate immune system utilizes proteins known as pattern recognition receptors (PRRs) to detect microbe-associated molecular patterns and to lead to the production of signaling molecules known as cytokines. DNA-sensing PRRs cyclic GMP-AMP synthase (cGAS) and interferon-inducible protein 16 (IFI16) bind to viral double-stranded DNA (dsDNA) and lead to the production of pro-inflammatory cytokines known as type I interferons (IFNs). Previous studies have not shown the production of type I IFNs in response to single-stranded viral DNA (ssDNA). However, due to the existence of ssDNA viruses, the nature of ssDNA viral replication resulting in the production of double-stranded viral DNA, and evidence presented in the literature of IFI16 colocalizing and associating with ssDNA, we hypothesized that single-stranded viral DNA should induce a type I IFN response. Furthermore, we hypothesized that adeno-associated virus (AAV), a small ssDNA virus that delivers its DNA into the nucleus, should induce a type I IFN response due to the fact that both IFI16 and cGAS can translocate into nucleus, though this virus is typically thought to induce no immune response. We investigated type I IFN responses to ssDNA by transfecting mature THP-1 cells with ssDNA from vaccinia virus (VAC70). We also investigated type I IFN responses to AAV by infecting mature THP-1 cells with different serotypes of AAV2. IFN responses were determined by measuring transcripts of interferon stimulated genes (ISGs) utilizing RT-qPCR. These findings show that ssDNA does induce an IFN response whether it is delivered through transfection of ssDNA VAC70 or through infection with AAV. These data have implications regarding the presence of unknown ssDNA sensors and clinical relevancy for gene therapy delivery techniques utilizing AAV.

Table of Contents

List of Abbreviations

AAV: Adeno-Associated Virus

cGAMP: Cyclic Guanosine Monophosphate-Adenosine Monophosphate

cGAS: Cyclic GMP-AMP Synthase

dsDNA: double-stranded DNA

dsVAC70: double-stranded Vaccinia Virus 70

HIN: Hemopoietic Expression - Interferon-Inducibility-Nuclear Localization

IFI16: Interferon-Inducible Protein 16

IFN: Interferon

IFNAR: Interferon α/β receptor

IRF3: Interferon Regulatory Factor 3

ISG: Interferon-Stimulated Genes

ITR: Inverted Terminal Repeats

MAMPS: Microbial-Associated Molecular Patterns

NF-kB: Nuclear Factor kB

PMA: Phorbol Myristate Acetate

Poly I:C: Polyinosinic-polycytidylic acid

PRR: Pattern Recognition Receptors

PQBP1: Polyglutamine binding protein 1

ssDNA: single-stranded DNA

ssVAC70: single-stranded Vaccinia Virus 70

STING: Stimulator of Interferon Genes

TBK1: TANK-binding kinase 1

rAAV: Recombinant Adeno-Associated Virus

RIG-I: Retinoic Acid-Inducible Gene I

RPL37a: Ribosomal Protein L37a

RT-qPCR: Reverse Transcriptase Quantitative Polymerase Chain Reaction

VAC70: Vaccinia Virus 70

Introduction

The immune system is a host defense mechanism comprised of complex biological structures and processes. Among its many functions, the immune system can protect a host against potentially pathogenic microbes. The immune system can be subdivided into different layers as described in Figure 1 (Parham 2015). The first layer consists of anatomical and physiological barriers. These barriers include components such as the skin and mucosal areas. Some structures act by physically inhibiting pathogens from colonizing and infecting the host, while others such as antimicrobial peptides, enzymes, and low pH, chemically fight off pathogens (Punt et al. 2019).

If a pathogen is able to overcome these physical and chemical barriers, it will encounter the next layer of immunity known as the innate immune system (Parham 2015). The innate immune system is relatively fast-acting and aims to control and eliminate pathogens as quickly as possible. The innate immune system accomplishes this by reacting to broad classes of pathogens rather than specific microbes. Once the innate immune system has identified the presence of a pathogen, it can then initiate an inflammatory response by recruiting inflammatory cells known as neutrophils and macrophages. In response to an infection, these innate immune cells will produce signaling molecules known as cytokines (Chen et al. 2010). Inflammatory cytokines such as interferons (IFNs) will lead to physiological changes such as increases in temperature, blood flow,

and permeability of blood vessels around the infected area. These innate immune cells also produce chemokines, a type of cytokine. Chemokines will recruit other important immune cells to the site of infection. These newly recruited immune cells are a part of the next layer of immunity: the adaptive immune system (Parham 2015).

Unlike the innate immune system, the adaptive immune system is slow acting but highly specific. In its exquisite specificity, adaptive immunity adapts its response during the course of infection (Parham 2015). This allows the overall population of immune cells to improve their effectiveness in efforts to clear the pathogen as quickly as possible. Another quality unique to the adaptive immune system is immune memory, in which the adaptive immune system produces a more rapid and effective response with each subsequent exposure to a particular pathogen. The innate immune response does not demonstrate immune memory, resulting in a similar primary response with repeated exposure to the same pathogen.

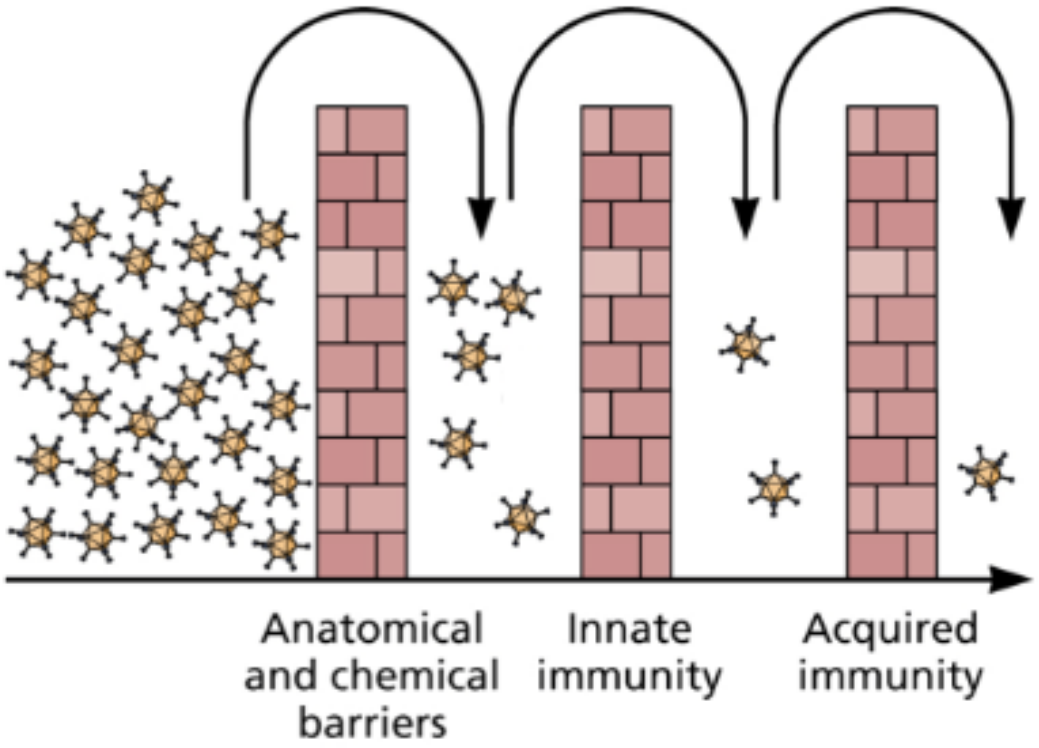


Figure 1: Layers of Immunity

The immune system can be divided into several different layers. Anatomical and chemical barriers inhibit potential pathogens from colonizing and infecting the host. The innate immune system is activated within minutes of infection and aims to eliminate pathogens as quickly as possible. The adaptive immune system (depicted here as acquired immunity) is slow acting but has exquisite specificity. Figure adapted from Flint et al. 2015.

Though the adaptive immune system is an integral part of the immune system, recently it has been made clear that the innate immune system plays an integral part in combating microbes. The innate immune system works to slow down infection while recruiting cells from the adaptive immune system to fight off infection (Iwasaki et al. 2010). Some researchers have even ventured to speculate that there would be no survival without an innate immune system (Parham 2015). It is hypothesized by many that without an innate immune response, pathogen replication would not be controlled, and this would lead to the death of the host, making the innate immune system an important area of study.

The innate immune system consists of a fast and broad response to microbial-associated molecular patterns (MAMPS). MAMPS consist of clearly non-self molecular structures that are essential for microbes (Parham 2015). MAMPS are generally highly conserved and found in various microbes. Targeting MAMPS is evolutionarily advantageous because it allows the innate immune system to produce a broad response to a wide array of microbes. In order to effectively detect pathogenic material, the innate immune system utilizes a class of proteins known as pattern recognition receptors (PRRs). PRRs bind to MAMPS such as lipopolysaccharides and other components of the bacterial cell wall, which would never normally be found on human cells (Walsh et al. 2013). Once a PRR binds to this clearly foreign material, it triggers a specific signal

cascade that leads to the activation of transcription factors and ultimately the production of pro-inflammatory cytokines as seen in Figure 2 (Parham 2015).

There are several broad classes of pro-inflammatory cytokines. The type of cytokine produced in response to infection will depend on the PRR activated (Yoshimura et al. 1997, Perry et al. 2005, Punt et al. 2019). Although there are many broad classes of pro-inflammatory cytokines, one well studied group of inflammatory cytokines activated in response to viral infection are the Type I interferons (IFNs). Type I IFN transcription is controlled by IFN regulatory factor 3 (IRF3) (Honda et al. 2006). Production of IFNs induce an antiviral state and prevent further spread of a viral infection by inducing expression of enzymes that block viral replication (Parham 2015). In the antiviral state, the host cell may inhibit host translation machinery, increase mRNA degradation, and inhibit virus transcription in order to slow viral replication (Parham 2015). Though these strategies successfully slow down viral replication, they cannot be maintained due to the fact that these same strategies hinder the host cell from engaging in functions necessary for its health such as the production of necessary host proteins. Though an IFN-induced antiviral state cannot be maintained, it is effective in slowing viral spread due to virus' reliance on host machinery.

Type I IFNs, consisting of IFN-alpha and IFN-beta, are the first to appear after viral infection and are produced by any nucleated cell infected with a virus. Once secreted, Type I IFNs bind to the interferon α/β receptor (IFNAR) in a

paracrine and autocrine manner as seen in Figure 2 (Piehler et al. 2012). When IFN binds to IFNAR, several host signaling pathways are affected, leading to changes in transcription and the production of IFN-stimulated genes (ISGs). ISGs help establish the antiviral state by targeting pathways and functions required for the microbe's life cycle (Schneider et al. 2014). Kinetically, IFNs are typically measured earlier during viral infection compared to ISGs (Parham 2015). However, ISG production is not always dependent on IFNs production. Some PRRs can activate IRF7, a transcription factor that directly controls the production of ISGs (Barber et al. 2011). Regardless, type I IFNs' secretion signals neighboring cells and leads to antiviral effects such as inhibition of viral gene expression (Bowie et al. 2008). This allows the antiviral response to be amplified and spread to surrounding uninfected cells.

The secretion of pro-inflammatory cytokines such as IFNs and the process of inflammation play a role in the progression of various diseases. Though localized inflammation can help clear infections, overwhelming production of pro-inflammatory cytokines can be very dangerous. Excess cytokines destroy the normal regulation of the immune response and induce pathological inflammatory disorders, such as capillary leakage, tissue injury and lethal organ failure (Ulloa et al. 2005). Therefore, the process of MAMP recognition and cytokine production must be heavily regulated. However, understanding this process can be problematic due to the fact that nucleic acids are sensed as MAMPS during viral infection (Rotem et al. 1963, Unterholzner et al. 2010).

Though MAMPS tend to be clearly foreign material, foreign nucleic acids are also sensed by PRRs (Rotem et al. 1963, Unterholzner et al. 2010). This can seem problematic because during infection, foreign nucleic acid is not the only nucleic acid found in the cell. PRRs must be able to differentiate between foreign nucleic acid and host nucleic acid. Some PRRs can use post-transcriptional modifications in order to discriminate between self and foreign RNA. For example, retinoic acid-inducible gene I (RIG-I) protein can detect foreign RNA by sensing the blunt end of double-stranded RNA and those with a 5'-triphosphate cap (Jiang et al. 2011). These structures allow RIG-I to distinguish between host mRNA and viral mRNA due to the fact that host mRNA is protected with a 5’ methylated cap (Munoz et al. 1976). However, details on how PRRs distinguish self from non-self double-stranded DNA are not as clear cut. Cytosolic DNA should be sensed because DNA is not normally found in the cytoplasm while the nuclear membrane would protect host DNA from PRR detection. However, several DNA-sensing PRRs have been found in both the cytoplasm and the nucleus and have been shown to lead to the production of pro-inflammatory cytokines in response to a nuclear-replicating double-stranded DNA virus (Unterholzner et al. 2010, Orzalli et al. 2013).

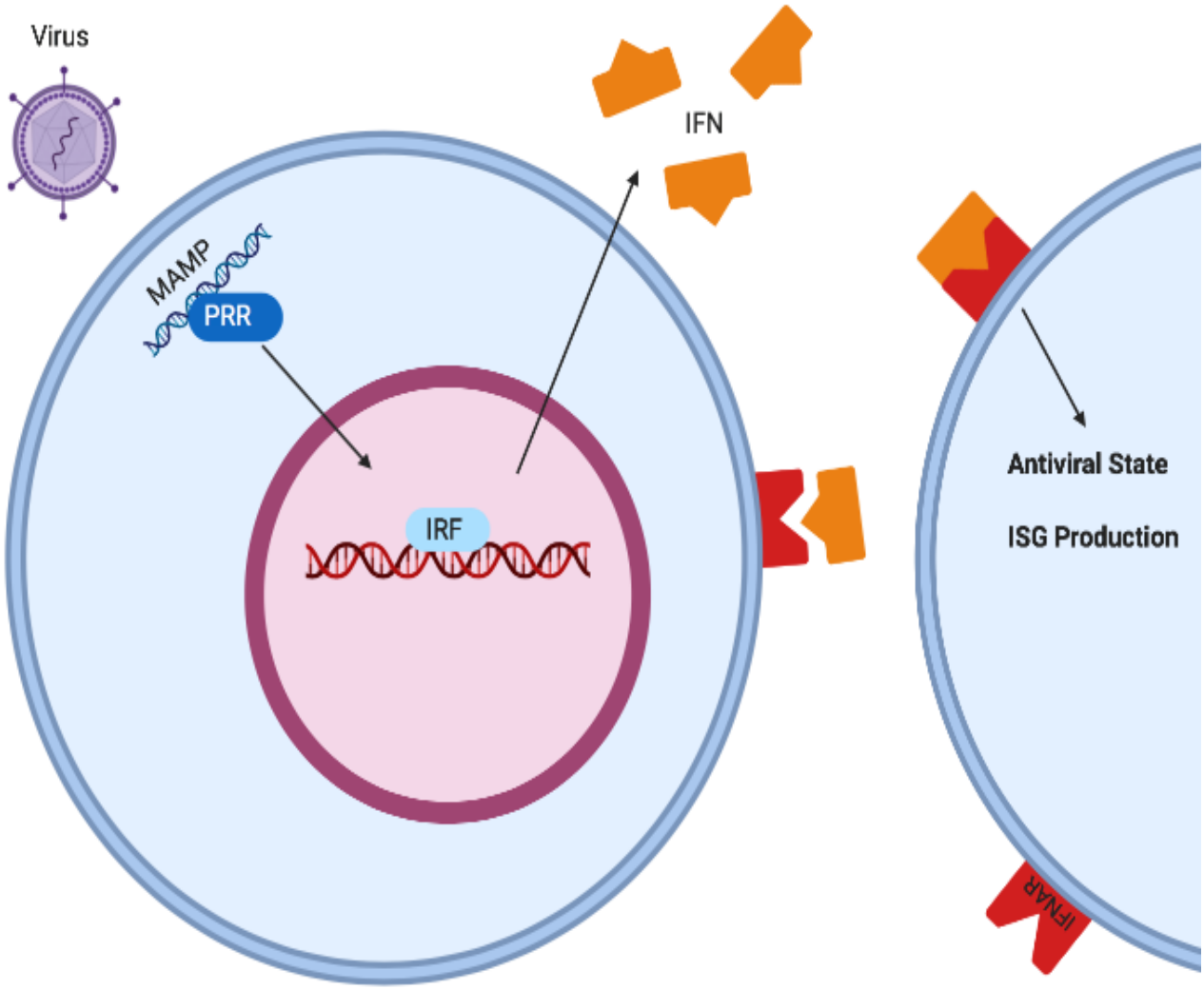


Figure 2: PRR Recognition

PRRs recognize MAMPs such as foreign nucleic acids or components of the bacterial cell wall. Once activated, PRRs will induce a signal cascade leading to production of pro-inflammatory cytokines. Certain PRRs can detect double-stranded viral DNA leading to the activation of transcription factor IRF3 and the production of IFNs. These IFNs can warn neighboring cells and induce an antiviral state.

One important PRR responsible for the detection of viral DNA is interferon-inducible protein 16 (IFI16). IFI16 detects double-stranded viral DNA in a sequence-independent manner (Unterholzner et al. 2010, Jin et al. 2012). However, DNA detection does seem to be dependent on the length of the DNA strand (Unterholzner et al. 2010). IFI16 is classified as a PYHIN protein (Brunette et al. 2012). PYHIN proteins contain a PYRIN domain and a HIN-200 domain. In particular, IFI16 contains two HIN-200 (hemopoietic expression -interferon-inducibility-nuclear localization) domains at the C-terminus that bind DNA and a PYRIN domain at the N-terminus, involved in protein-protein interactions (Brunette et al. 2012, Haronikova et al. 2016). Unexpectedly for a PRR that recognizes double-stranded DNA, studies have shown that IFI16 can be detected in both the nucleus and the cytoplasm (Dawson et al. 1995, Unterholzner et al. 2010, Diner et al. 2015). In fact, IFI16 has been shown to function in the nucleus to sense herpesviruses, double-stranded DNA (dsDNA) viruses that replicate in the nucleus (Orzalli et al. 2012).

IFI16 is not the only PRR known to detect dsDNA as a MAMP. Cyclic GMP-AMP synthase (cGAS) is also responsible for detecting viral DNA and inducing a type I IFN response (Figure 3). In fact, data have shown that IFI16 may cooperate with cGAS in the production of type I IFNs and that these two pathways may not be separate (Almine et al. 2017, Shannon et al. 2018). cGAS, though mainly cytoplasmic, has been shown to translocate to the nucleus (Gentile et al. 2019). cGAS binds double-stranded DNA by interacting with the

phosphodiester backbone (Civril et al. 2013). This allows cGAS to detect foreign DNA in a sequence-independent manner. Unlike IFI16, cGAS does not have HIN or PYRIN domains. Instead, cGAS utilizes a different nucleic acid-sensing motif, a zinc thumb, to bind DNA (Civril et al. 2013). Once bound to viral DNA, cGAS catalyzes the cyclization of AMP and GMP, leading to the formation of the secondary messenger cyclic guanosine monophosphate-adenosine monophosphate (cGAMP) (Wu et al. 2013, Hall et al. 2017). cGAMP, along with activated IFI16, then binds to an adaptor endoplasmic reticulum membrane protein called stimulator of interferon genes (STING) as seen in Figure 3 (Chen et al. 2016, Almine et al. 2017). These binding events cause STING to undergo a conformational change and dimerize. STING will then traffic to the ER-Golgi intermediate complex where it can interact with TANK-binding kinase 1 (TBK1) (Chen et al. 2016). TBK1 phosphorylates STING at its C-terminal tail in order to recruit and phosphorylate transcription factor interferon regulatory factor 3 (IRF3) (Chen et al. 2016). Phosphorylation of IRF3 leads to its dimerization and nuclear translocation. In the nucleus, IRF3 and nuclear factor □B (NF-□B) activate the production of IFNs and ISGs (Cai et al. 2014).

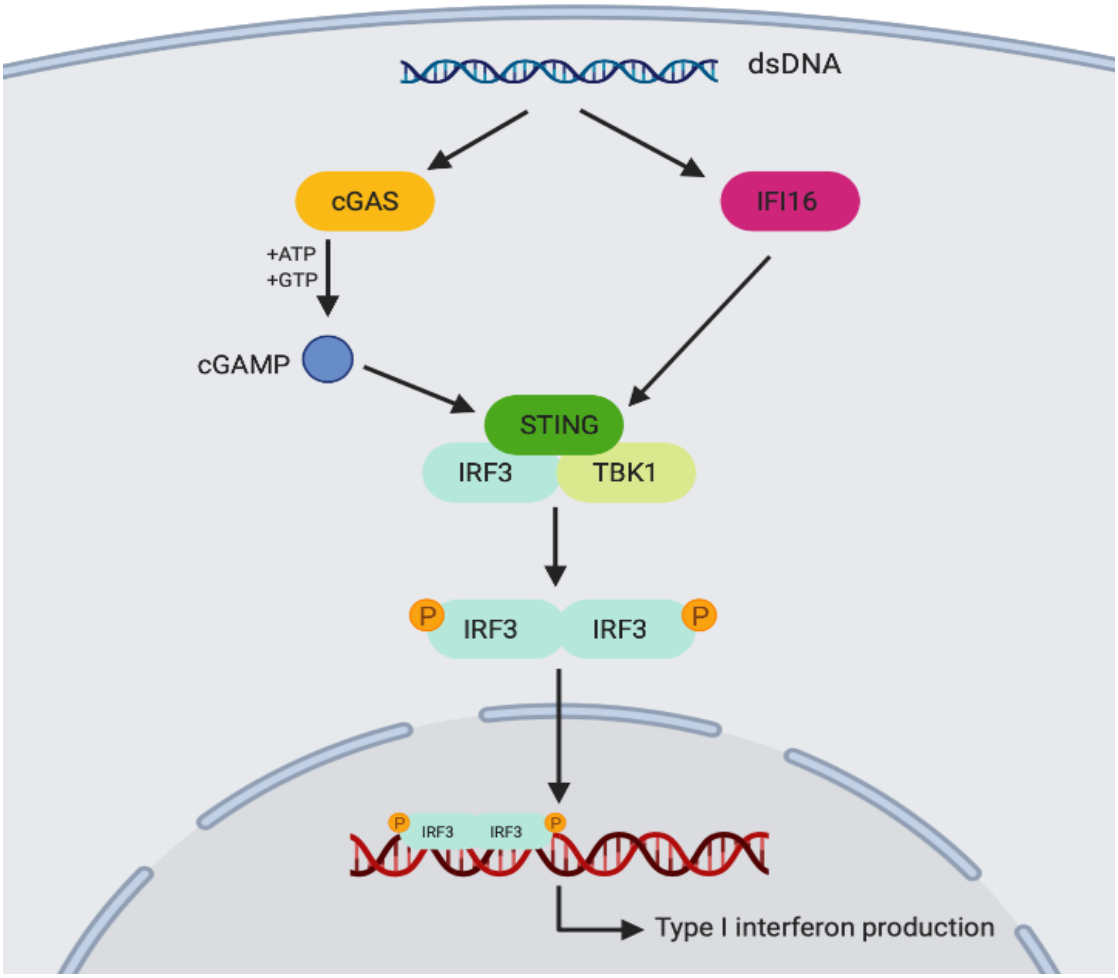


Figure 3: Simplified IFI16-STING and cGAS-STING Signal Cascade

This pathway demonstrates how both IFI16 and cGAS can sense double-stranded DNA leading to a type I interferon response. cGAS binds DNA and produces cGAMP as a result. cGAMP along with activated IFI16 bind to STING, inducing a conformational change. Activated STING dimerizes and recruits TBK1 and IRF3. TBK1 phosphorylates IRF3 which dimerizes and translocates into the nucleus. In the nucleus, phosphorylated IRF3 can act as transcription factor for type I IFNs, along with NF-□B, and initiate an antiviral innate immune response.

The ability for IFI16 and cGAS to translocate between the nucleus and the cytoplasm allows these PRRs to detect DNA viruses and retroviruses regardless of their location (Unterholzner et al. 2010, Volkman et al. 2018, Gentile et al. 2019). Data from our lab have demonstrated that the function of a PRR can be dependent on its subcellular localization (Shannon et al. 2018). Polyglutamine binding protein 1 (PQBP1), a binding partner of cGAS, is required for a type 1 interferon response during HIV infection (Yoh et al. 2015). Data presented in the Yoh et al. study demonstrated PQBP1's critical role in response to nuclear viral DNA stimulation. However, when PQBP1's role was examined in response to cytoplasmic viral DNA stimulation, PQBP1 was found to inhibit the type 1 interferon response to cytosolic DNA (Shannon et al. 2018). These data demonstrate the importance that subcellular localization plays in the detection of stimuli and the activation of the innate immune response (Yoh et al. 2015, Shannon et al. 2018). This is of particular importance because different delivery methods of DNA stimuli may lead to a varied response and confounding data. These data also demonstrate that there are a lot of unknowns about how PRRs work.

Single-Stranded DNA Sensing

Though IFI16 and cGAS have been shown to detect double-stranded viral DNA, few studies have examined whether these PRRs can detect single-stranded viral DNA (Monroe et al. 2014, Unterholzner et al. 2010, Jakobsen et al.

2013). Studies describing IFI16 inadvertently demonstrated its ability to bind single-stranded DNA (ssDNA) (Monroe et al. 2014, Unterholzner et al. 2010). Unterholzner and colleagues used a 70 base pair motif of vaccinia virus (VAC70) in order to identify IFI16 as a PRR and further study its ability to signal through the STING pathway. The data presented demonstrated that when mature THP-1 macrophages were transfected with double-stranded VAC70 (dsVAC70) or single-stranded VAC70 (ssVAC70) for 6 hours, samples transfected with dsVAC70 demonstrated significant IFN-beta production (Unterholzner et al. 2010). Samples transfected with ssVAC70 did not show significant IFN-beta production, implying that single-stranded viral DNA was not recognized by IFI16 (Unterholzner et al. 2010). However, in an immunoblot analysis of IFI16 from mature THP-1 macrophages incubated with biotinylated ssVAC70 or dsVAC70, IFI16 was found to bind in complex with both ssVAC70 and dsVAC70 (Unterholzner et al. 2010). These results imply that IFI16 interacts with single-stranded viral DNA. Likewise, Monroe and colleagues also did a pulldown experiment for IFI16 utilizing tonsillar CD4 T cells infected with biotinylated HIV DNA (Monroe et al. 2014). They found that IFI16 could bind this HIV DNA. Monroe and colleagues then ran an SDS-PAGE and silver stain analysis on cytosolic extracts from tonsillar CD4 T cells using biotinylated dsDNA or ssDNA (Monroe et al. 2014). These samples were put into competition with a 10-fold excess of ssDNA or dsDNA. When biotinylated dsDNA was put into competition with unlabeled ssDNA, a slight decrease in labeled IFI16 was seen, suggesting

that ssDNA had successfully competed with dsDNA for IFI16. When biotinylated ssDNA was put into competition with unlabeled dsDNA, IFI16 was no longer detected, suggesting that dsDNA outcompeted IFI16 from ssDNA. Further immunoblotting analysis demonstrated that with high protein input IFI16 binds to biotinylated ssDNA, though more weakly than to dsDNA. These results imply that IFI16 can interact with both dsDNA and ssDNA. Jakobsen and colleagues matured THP-1 cells and then stimulated these cells with ssDNA oligonucleotides derived from HIV-1 (Jakobsen et al. 2013). IFI16 was shown to colocalize to the site of transfected DNA in immunofluorescence data, further implying that IFI16 may interact with ssDNA. These data presented by Unterholzner et al., Monroe et al., and Jakobsen et al. are of great significance because innate immune sensors for ssDNA have not been specifically described in the literature and these data imply that IFI16, an innate immune sensor, may be interacting with and possibly sensing foreign ssDNA.

Though ssDNA sensing has not been fully investigated, the fact that IFI16 has been shown to colocalize and bind to ssDNA seems to suggest that ssDNA may be sensed. This suggestion is further supported by the fact that studies have shown that the IFI16 HIN-A domain demonstrates preferential binding to ssDNA compared to dsDNA (Yan et al. 2008, Haronikova et al. 2016). However, one possibility is that ssDNA is sensed by dsDNA sensors after it has undergone DNA repair into dsDNA. Single-stranded DNA viruses perform DNA processing as the first step of viral replication, resulting in the production of double-stranded

viral DNA (Flint et al. 2015, Vogel et al. 2012, Franzoso et al. 2017). This double-stranded viral DNA should be recognized by PRRs such as IFI16 and cGAS, which are both dsDNA sensors. It is quite possible that IFI16 associates with single-stranded viral DNA weakly, and once it has been turned into dsDNA, IFI16 and cGAS trigger the innate immune response.

One additional piece of evidence linking IFI16 and cGAS to ssDNA is via DNA repair. Previous studies have shown that cGAS and IFI16 may play a role in DNA damage response and repair pathways (Diner et al. 2015, Choubey et al. 2016, Unterholzner et al. 2019). In fact, data have demonstrated that IFI16 may associate with Ku70 and Ku80, important DNA repair proteins (Diner et al. 2015). If ssDNA is being sensed after DNA repair, it is possible to hypothesize that while weakly associating with the single-stranded viral DNA, these PRRs recruit DNA repair proteins and initiate the interferon response once dsDNA has been formed.

Another possibility is that ssDNA is bound by DNA sensors but fails to induce an IFN response. This conclusion is supported by data presented by Unterholzner and colleagues, demonstrating that IFN-beta production was only seen in mature THP-1 macrophages stimulated with dsVAC70 using a 6 hour transfection (Unterholzner et al. 2010). However, these data are not conclusive. The possibility that ssDNA fails to induce an IFN response is also supported by the fact that Anelloviruses, a group of small ssDNA viruses, are found in almost

100% of people worldwide as part of the host virome in metagenomic studies (Freer et al. 2018). More importantly, Anelloviruses are found in healthy people and seem to have no specific pathogenic effect (Freer et al. 2018). However, studies have shown increased viral loads of Anelloviruses in immunosuppressed individuals and in patients suffering from inflammatory diseases. These data may suggest that Anelloviruses are normally kept under immunological control and therefore, should be sensed (Young et al. 2015, Abbas et al. 2016, Freer et al. 2018). If ssDNA truly fails to induce an IFN response, this would lead us to question why and how ssDNA is not being sensed.

Due to the immense number of ssDNA viruses that infect humans, it would be strange if ssDNA was not sensed by PRRs (Freer et al. 2018). Understanding the innate immune response to ssDNA viruses is particularly important due to the fact that adeno-associated virus (AAV), a ssDNA virus, is frequently used for gene therapy (Fisher et. al 1997, Jooss et al. 2003). Therefore, understanding the innate immune response to viral ssDNA is of clinical importance and may have major implications in the development of gene therapy.

Adeno-Associated Virus

AAV is a small (26 nm) non-enveloped virus from the *Parvoviridae* family with a linear ssDNA genome (Senis et al. 2014). Once a single-stranded DNA virus has attached to the host cell, entered, and uncoated it's genome, it cannot begin to replicate its genome and produce viral proteins until it turns its ssDNA

into dsDNA (Flint et al. 2015). Whether ssDNA viruses use the host cell's DNA replication or repair polymerase in order to compliment the ssDNA genome is not fully understood (Vogel et al. 2012, Franzoso et al. 2017). Once there is viral dsDNA, the host cell's replication and transcriptional machinery can proceed with that dsDNA normally. Before viral assembly the double-stranded viral DNA is converted back into ssDNA as seen in Figure 4.

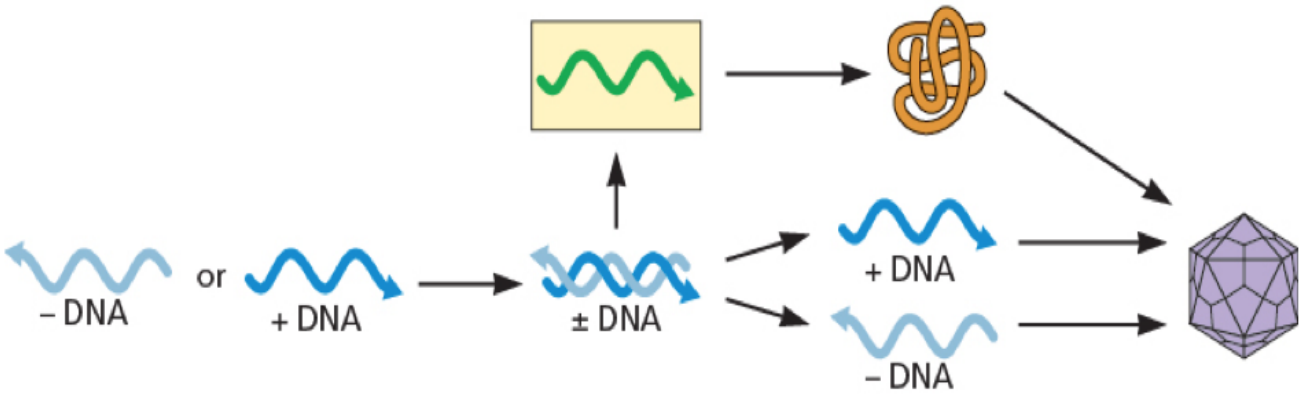


Figure 4: Replication and Transcription Mechanisms of ssDNA Viruses

Single-stranded viruses must first turn their ssDNA into dsDNA before they can produce viral transcripts and proteins. DNA synthesis of the complementary strand can be completed by either DNA repair or replication polymerase. Figure adapted from Flint et al. 2015.

Wildtype AAV has a genome size of 4.8 kb. Its genome consists of three genes, *Rep* (replication), *Cap* (capsid), and *aap* (assembly) (Naso et al. 2017). These three genes are placed between two inverted terminal repeats (ITRs) that span approximately 145 base pairs each. Proteins known as Rep78, Rep68, Rep52, and Rep40 are required for the AAV replication and packaging cycle, while proteins known as VP1, VP2, and VP3 are capsid proteins (Samulski et al. 2014, Naso et al. 2017).

In order to effectively replicate, AAV requires co-infection with a helper virus, most commonly adenovirus or herpesvirus (Naso et al. 2017). Without one of these helper viruses, AAV will not be able to replicate. Adenovirus genes *E4*, *E2a,* and *VA* help mediate AAV genome replication and virion packaging, though their exact roles are still under investigation (Matsushita et al 1988, Flint et al. 2015). These genes can be placed on a helper plasmid. In order to produce AAV, a transfer plasmid containing *Rep* and *Cap* and the helper plasmid can be transfected into HEK293 cells (Matsushita et al. 1998). This will result in the production of infectious AAV particles.

Surprisingly, humans infected with AAV do not demonstrate any pathology (Naso et al. 2017). This is particularly interesting because typical innate immune responses to viral infection lead to symptoms such as fever, malaise, fatigue, nausea, loss of appetite due to the production of IFNs (Flint et al. 2015). In order to not induce an innate immune response, a virus must have genes that assist it

to evade immune detection. However, AAV is a small virus with a small genome. It should not have space in its genome to encode for molecules that allow it to evade the immune system and go undetected.

Interestingly, chromosomal integration and the establishment of latent infection has been seen with AAV infection (Kotin et al. 1990). Wildtype AAV typically integrates its genome at AAVS1 on human chromosome 19 (Kotin et al. 1990). Integration occurs at this specific AAVS1 site due to the presence of a *Rep* binding element via homologous recombination. Though it is integrated into the host genome, without a helper virus, AAV will remain replication incompetent. Recombinant AAV (rAAV) is used for gene therapy research instead of wildtype AAV. When using rAAV, genomic integration on chromosome 19 is eliminated because this recombinant virus the Rep proteins (Naso et al. 2017). In fact, it's very unlikely that rAAV will integrate at non-homologous sites in the host genome (Choi et al. 2006) Data has shown that rAAV ssDNA is processed into dsDNA through DNA repair pathways utilizing DNA-PK (Choi et al. 2006, Schwartz et al. 2009).

Eleven serotypes of AAV have been identified. Each serotype differs in transduction efficiency due to slight differences in capsids (Agbandje-McKenna et al. 2011, Merkel et al. 2017). In order to improve transduction efficiency, researchers can create pseudotypes of AAV. Pseudotyping allows for the creation of viral particles that are made up of capsid proteins from one AAV

serotype but carry the genome of a different AAV serotype. Researchers can also experiment with hybrid capsids derived from multiple different serotypes. For example, AAV-DJ is a commonly used hybrid capsid derived from eight serotypes (Grimm et al. 2008). This method has shown to increase transduction efficiency. In fact, AAV-DJ displayed high infectivity across a broad range of cell types (Grimm et al. 2008).

AAV's relatively small size allows it to slip through the nuclear pore and deliver its DNA straight into the nucleus without the risk of engaging cytosolic DNA sensors due to the fact that AAV's capsid will wait until it is in the nucleus to uncoat (Porwal et al. 2013). AAV's ability to deliver its ssDNA into the nucleus makes it useful for gene editing techniques such as delivery of CRISPR/Cas9 (Senis et al. 2014, Chew et al. 2016). Many articles have claimed that AAV induces little to no immune response, making it a useful vector for CRISPR/Cas9 delivery (Fisher et. al 1997, Jooss et al. 2003).

It is possible that AAV's small size allows it to bypass cytoplasmic sensors, but it is unclear why it is not detected by nuclear PPR sensors. AAV's genome, though single-stranded, should become double-stranded while in the nucleus. Though viral DNA is present in the nucleus after AAV infection, no studies have investigated if AAV induces a type I interferon response through a STING-mediated pathway in human cell lines. Though IFI16 and cGAS have been shown to detect both cytosolic and nuclear viral DNA and show some

evidence of ssDNA binding, no studies have investigated whether these PRRs detect AAV. If AAV induces an interferon response, it could affect the efficiency of CRISPR/Cas9 delivery and the symptoms seen in patients treated with gene therapy. If an interferon response is not induced, then it raises questions on how a virus as small as AAV is evading nuclear DNA sensing PRRs when it lacks space in its genome to encode for molecules that allow it to evade the immune system and go undetected.

Experimental Goals

This study investigates whether single-stranded viral DNA induces a type I interferon response. This was done by stimulating THP-1 cells, an immortalized cell line, with single-stranded DNA and comparing it to stimulation with double-stranded DNA or a virus with a single-stranded DNA genome at different time points. The amount of ISG transcription after stimulation was quantified through quantitative polymerase chain reaction (qPCR). If data show the production of ISG transcripts following transfection, then it supports the hypothesis that single-stranded DNA is detected and induces an innate immune response. We hypothesize that single-stranded viral DNA will be sensed in the cytoplasm and lead to an innate immune response kinetically later than double-stranded viral DNA due to the fact that IFI16 has been shown to colocalize and bind to ssDNA, but transfected ssDNA has not been shown to induce an IFN response 6 hours

post-infection according to the literature (Jakobsen et al. 2013, Unterholzner et al. 2010, Monroe et al. 2014).

This study also investigates whether AAV induces a type I interferon response. This was done by stimulating THP-1 cells, an immortalized cell line, with different AAV2 pseudotypes. AAV pseudotypes used in this study differed in capsid type, but all delivered AAV2 genomic ssDNA. The amount of interferon transcription after AAV infection was quantified through quantitative polymerase chain reaction (qPCR). If data show the production of interferon stimulated genes (ISG) following AAV infection, then it will support the hypothesis that AAV is detected and induces an innate immune response. We hypothesize that single-stranded viral DNA delivered by AAV will be sensed in the nucleus and lead to an innate immune response.

Methods

Cell Culture:

THP-1 monocytic cells (obtained from ATCC) were cultured in R10 media. R10 media contained RPMI media (obtained from Invitrogen) supplemented with fetal bovine serum (obtained from Invitrogen, 10%), β-mercaptoethanol (obtained from Invitrogen, 50 nM), non-essential amino acids (obtained from Invitrogen, 1 mL/100 mL), Normocin (obtained from Invivogen, 0.5 mg/mL), sodium pyruvate (obtained from Invitrogen, 1 mM) and penicillin-streptomycin-glutamine (obtained

from Invitrogen, 5 mL/500 mL). Cells were incubated at 37°C in a 5% CO_2 humidified tissue culture incubator.

Cell Stimulation:

THP-1 monocytic cells were plated at the following concentrations: 2 x 10^6 cells in 3 mL of media in a 6-well plate, or 2 x 10^5 cells in 1 mL of media in a 24-well plate. THP-1 cells were treated with 5 ng/mL of phorbol 12-myristate 13-acetate (PMA) to allow for complete maturation into macrophages (Yoh et al., 2015). Cells plated in a 6-well plate were then incubated for 72 hours at 37°C in a 5% CO_2 humidified tissue culture incubator. In order to induce cytoplasmic stimulation after THP-1 maturation, a 250 uL mixture of nucleic acid and Opti-MEM (obtained from Invitrogen) was made. This mixture was incubated at room temperature for 5 minutes. Another mixture of 10 uL of Lipofectamine 2000 (obtained from Invitrogen), a lipid reagent, and 240 uL of Opti-MEM was made. Following the 5-minute incubation, both mixtures were combined and incubated at room temperature for 20 minutes to 6 hours before addition to the cells. This process, known as transfection, allows us to deliver nucleic acid into the cytoplasm of our cells (Figure 5). Lipofectamine 2000 forms a lipid micelle around the nucleic acid stimuli. The micelle then merges with the cell membrane to deliver the stimuli to the cytoplasm. Without the micelle, large and polar stimuli remain outside of the cell.

Nucleic acid used for stimulation include a 70 base pair motif of double-stranded DNA from the Vaccinia virus genome (dsVac70, obtained from Invivogen) and single-stranded DNA of the same sequence from the Vaccinia virus genome (ssVac70, obtained from Invivogen), both of which have been previously examined in experiments related to IFI16 (Unterholzner et al. 2010). For a positive control, a synthetic viral RNA called Poly I:C (obtained from Invivogen) that is known to induce interferon independently of cGAS and IFI16 was used (Jiang et al. 2011). Final concentrations of Poly I:C, ssVac70, and dsVac70 added were 4 μg/mL. Mock samples were generated by combining Lipofectamine 2000 mixture with pure Opti-MEM with no nucleic acid added. Untreated samples were prepared with just Opti-MEM and no Lipofectamine 2000 or nucleic acid. Samples were incubated for 6 hours, 24 hours, or 48 hours following transfection. Cells were then pelleted and stored in RNA lysis buffer (obtained from Zymo Research) at -80°C. Supernatants were also harvested for future examination of secreted cytokines.

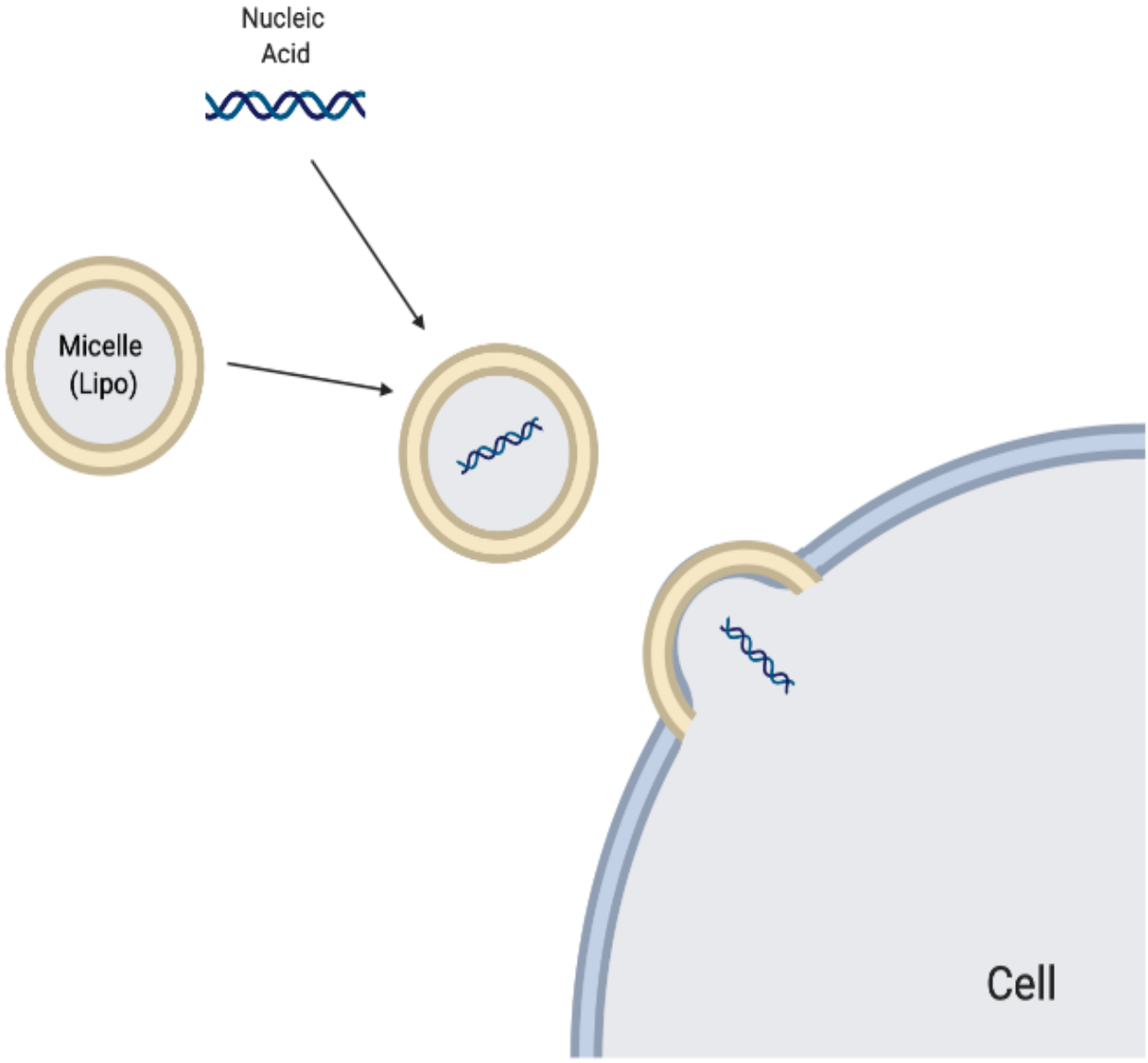


Figure 5: Cytosolic Delivery of Nucleic Acid

By using Lipofectamine 2000 (Lipo), viral and synthetic DNA or RNA can be delivered to the cytoplasm in a process known as transfection. The Lipo forms a micelle around the nucleic acid stimuli. The micelle then merges with the cell phospholipid bilayer and delivers the stimuli to the cytosol.

AAV Infection:

THP-1 cells were treated with 5 ng/mL of phorbol 12-myristate 13-acetate (PMA) to allow for complete maturation (Yoh et al., 2015). Cells plated in a 24-well plate (at a concentration of $2x10^5$ cells in 1 mL) were then incubated for 72 hours at 37°C in a 5% CO2 humidified tissue culture incubator. Cells were infected with the following AAV serotypes (obtained from Dr. Alejandro Balazs, Ragon Institute of MGH, MIT, and Harvard): AAV2/1, AAV2, AAV2/5, AAV2/7, AAV2/8, AAV2/9, AAVrh10, AAVDJ, and AAVDJ8. A final concentration of 2×10^{10} genome copies/mL was used for infection. A multiplicity of infection (MOI) of 2 was used (Zaiss et al. 2008). All samples were incubated for 6-8 hours before RNA lysis and extraction. Cells were pelleted and stored in RNA lysis buffer (obtained from Zymo Research) at -80°C. Supernatants were also harvested for future examination of secreted cytokines.

RNA Extraction and cDNA Synthesis:

Following the Quick-RNA MiniPrep Kit protocol (obtained from Zymo Research), samples were lysed and RNA extracted. The recommended DNase I treatment was also performed according to the Quick-RNA MiniPrep Kit protocol. The concentration and purity of extracted RNA samples was confirmed via Nanodrop. Samples were stored at -80°C.

Using the ProtoScript First Strand cDNA Synthesis Kit (obtained from New England Biolabs), cDNA samples were prepared from purified RNA samples. A simplified diagram of the cDNA synthesis process can be seen in Figure 6A. cDNA samples were stored at -20°C.

RT-qPCR:

RT-qPCR samples were prepared by mixing 2 µL of cDNA sample with 10 µL iTaq Universal SYBR Green mix (obtained from BioRad), 6 µL nuclease-free water, and 1 µL each of the forward and reverse primers at a concentration of 10 uM. Amplification primers were obtained from IDT. Primer sequences are as follows:

ISG56 Forward: 5'-CCT CCT TGG GTT CGT CTA CA-3'

ISG56 Reverse: 5'-GGC TGA TAT CTG GGT GCC TA-3'

ISG54 Forward: 5'-CAG CTG AGA ATT GCA CTG CAA-3'

ISG54 Reverse: 5'-GTA GGC TGC TCT CCA AGG AA-3'

RPL37A Forward: 5'-ATTGAAATCAGCCAGCACGC-3'

RPL37A Reverse: 5'-AGGAACCACAGTGCCAGATCC-3'

ISG56 primers were designed based on Jakobsen et al 2013. *ISG54* primers were designed based on Herzner et al 2015. *RPL37A* primers were designed

based on Maess et al 2010. The samples were run on BioRad CFX96 Real-Time PCR Machine. Samples were denatured at 95°C for 3 minutes, annealed at 95°C for 10 seconds and extended at 60°C for 30 seconds for 40 cycles. At the end of each cycle, fluorescence from the SYBR Green dye was read. After 40 cycles, melt curves were generated by increasing the temperature from 65°C to 95°C in 0.5°C increments to confirm products. A simplified diagram of the RT-qPCR process can be seen in Figure 6B. Threshold cycle values (Cq) for *ISG54*, *ISG56*, and *RPL37A* were collected and converted to fold change as follows (Livak et al. 2001):

The samples were first normalized to *RPL37A*, which encodes for a protein of a ribosomal subunit (Maess et al., 2010). This normalization was done by calculating the change (Δ) in Cq value. The ΔCq value for each sample was calculated using the following equation:

$$\Delta Cq = Cq \text{ of ISG} - Cq \text{ of RPL37A}$$

The samples were then normalized to background expression in controls. This was done by first calculating the average ΔCq values for the negative control, which was the mock or the uninfected condition. This value was then used to calculate the ΔΔCq value by utilizing the following equation:

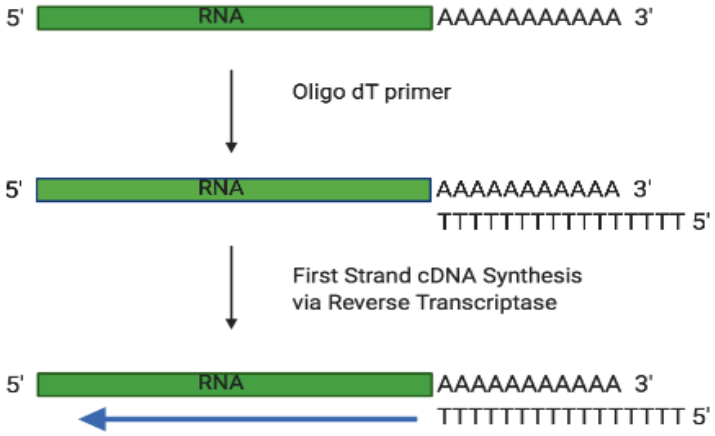
$$\Delta\Delta Cq = \Delta Cq \text{ of experimental sample} - \text{AVERAGE } \Delta Cq \text{ negative control}$$

The fold change was calculated for each sample using ΔΔCq values by utilizing the following equation:

$$Fold\ Change = 2^{-(\Delta\Delta Cq)}$$

Representative graphs of each experiment demonstrate average fold change with error bars representing the standard deviation for each sample.

A.



B.

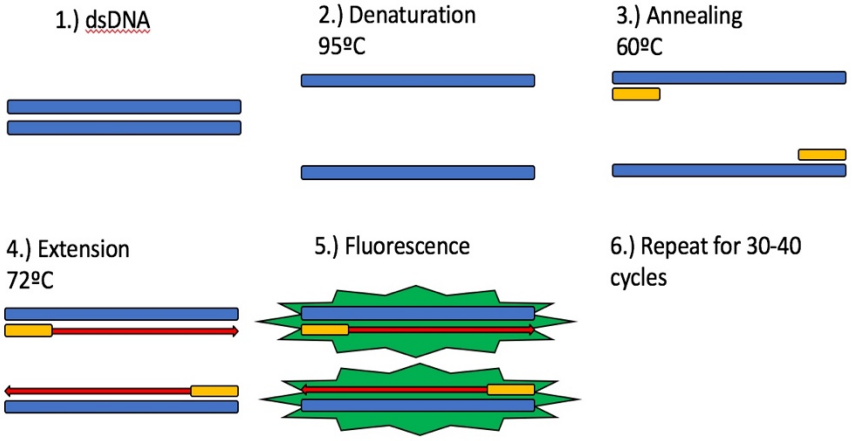


Figure 6: RT-qPCR Overview

This figure shows the steps that occur during a RT-qPCR to measure mRNA levels. A. cDNA is synthesized from extracted RNA using the ProtoScript First Strand cDNA Synthesis Kit (obtained from New England Biolabs). B. dsDNA is first denatured at a high temperature. Primers for the gene of interest then anneal and extend forming the complementary sequence. SYBR Green dye binds to the newly formed dsDNA and is used to measure cDNA levels. Figure adapted from Dorrity, 2018.

Results

In order to determine whether ssDNA delivered to the cytoplasm triggers an interferon response, THP-1 monocytes were matured with PMA for 72 hours. Once matured into macrophages, cells were stimulated with a 70 base pair DNA motif of vaccinia virus (dsVAC70) or a single-stranded version of the same sequence (ssVAC70) to the cytoplasm with the help of Lipofectamine 2000. This motif was chosen due to previous data shown by Unterholzner and colleagues demonstrating IFI16 binding to this motif (Unterholzner et al. 2010). dsVAC70 is a known DNA stimulant that triggers the production of interferon through the STING-mediated pathway (Unterholzner et al. 2010). As negative controls, cells received media with no nucleic acid (untransfected) or a mixture of Lipofectamine 2000 and media (mock) to control for the effect of the Lipofectamine 2000. As a positive control, cells received Poly I:C, a synthetic viral RNA that is known to induce interferon independently of cGAS and IFI16.

ISG56 and *ISG54* are known measures of an interferon response. Using qPCR, *ISG56* transcript levels in response to different stimuli can be measured. Regardless of time allotted for stimulation, untransfected and mock samples were expected to demonstrate insignificant amounts of *ISG56*. As seen in Figure 7, mock and untransfected generally demonstrated background levels of *ISG56*. While untransfected samples only received media, mock samples received both media and Lipofectamine 2000, a delivery agent, for 6 hours. These data demonstrate that Lipofectamine 2000 caused only a minor change in the

production of *ISG56* transcripts. Poly I:C, the positive control, was expected to demonstrate increased levels of *ISG56* transcript compared to mock and untransfected. This increase in production of *ISG56* in response to Poly I:C was seen in both the 6 hour and the 24 hour mark (Figure 7). However, at 48 hours, Poly I:C did not show an increase in production of *ISG56* and reverted to mock and untransfected levels (Figure 7). These data imply that production *ISG56* transcript should be measured prior to 48 hours for these experiments. At 48 hours post-transfection, it is possible that the stimulated cells have died or the response was already over

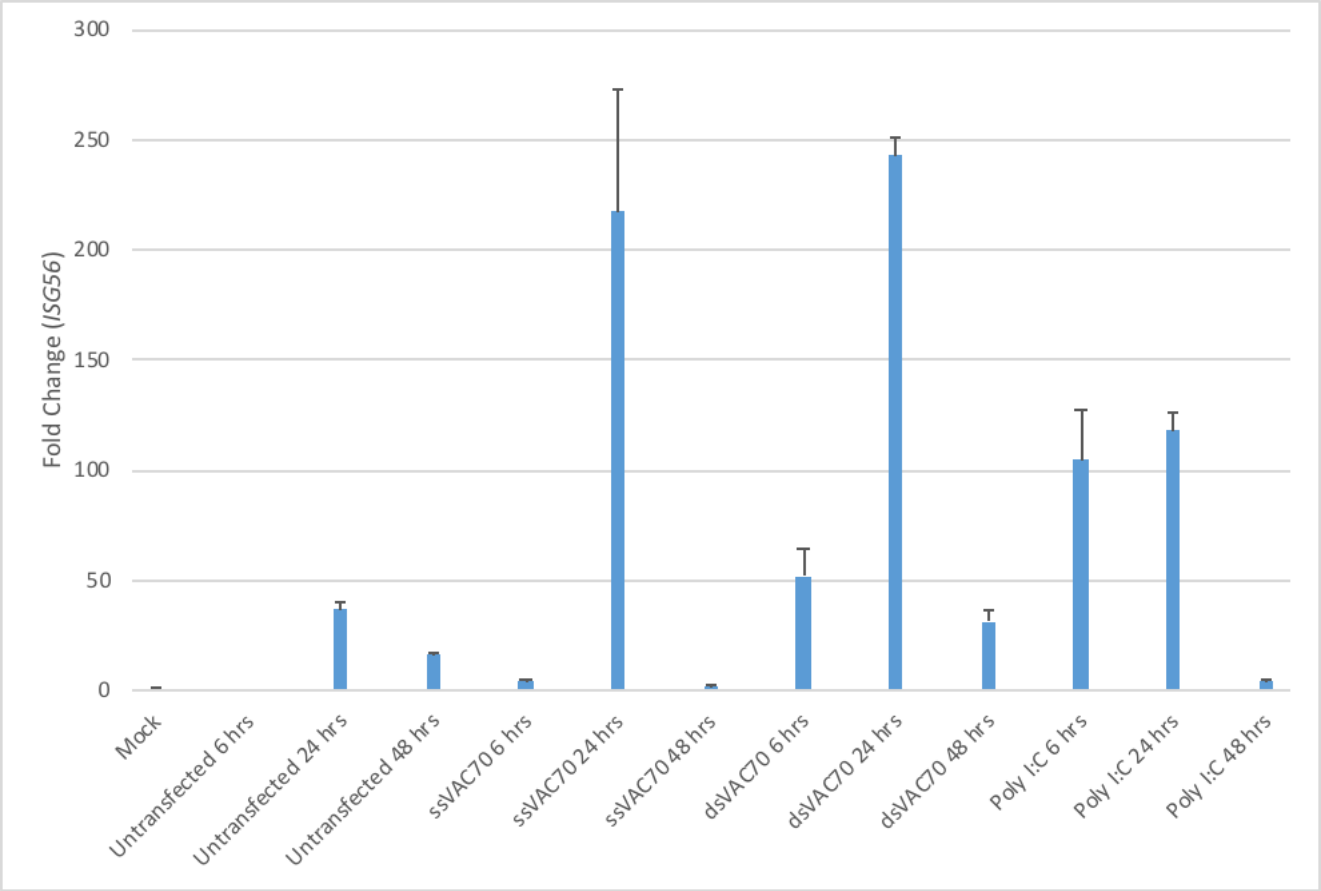


Figure 7: Kinetic Dependence of ssDNA on *ISG56* Production

PMA-treated THP-1 cells stimulated with ssVAC70, dsVac70, or Poly I:C (delivered through transfection). RNA from samples was extracted and cDNA was synthesized. *ISG56* transcripts measured through RT-qPCR. Samples collected 6 hours, 24 hours, and 48 hours post-transfection. Fold changes are all calculated relative to mock. Data are representative of the mean of technical triplicates and error bars represent the standard deviation. These data are representative of 3 experiments.

Double-stranded VAC70 (dsVAC70) was expected to demonstrate increased levels of *ISG56* transcript compared to mock and untransfected as demonstrated in the literature (Unterholzner et al. 2010). Data collected at 6 hours and 24 hours demonstrate increased levels of *ISG56* transcript in samples stimulated with dsVAC70 compared to mock and untransfected as seen in Figure 7. There is an increase in *ISG56* between the 6 hour mark and the 24 hour mark. This trend did not continue at 48 hours (Figure 7). These data further imply that production *ISG56* transcript occurs prior 48 hours and, therefore, should be measured earlier than 48 hours post-transfection. Interestingly, as seen in Figure 7, ssVAC70 did not demonstrate an increased amount of *ISG56* compared to mock and untransfected at 6 hours post-transfection. However, 24 hours post-transfection, ssVAC70 demonstrated 217-fold greater production of *ISG56* transcripts compared to mock (Figure 7). Unsurprisingly based on the pattern seen in our other samples, by 48 hours the production of *ISG56* transcripts had diminished back to mock and untransfected levels as seen in Figure 7 indicating that ssVAC70 is being sensed by the host cell early post-transfection. Overall, these data suggest that ssDNA does trigger an interferon response when transfected into the cytoplasm and that it does so kinetically slower than dsDNA.

In order to determine whether ssDNA delivered to the nucleus by infection with AAV triggers an interferon response, THP-1 monocytes were matured with PMA for 72 hours. Once matured into macrophages, cells were infected with various AAV2 serotypes for 8 hours or 24 hours. Uninfected samples received

media only. As seen in Figure 8, all samples that were infected with AAV for 8 hours demonstrated an increase in *ISG56* transcription compared to uninfected. Production of *ISG56* transcript varied between AAV2 serotypes (demonstrating a fold change range of 3.0 - 5.3). A positive control was not included in this study. Interestingly, AAV infection induced the production of *ISG56* transcript at 8 hours post-infection. Increased production of *ISG56* transcript this early on post-infection was not expected based on data presented in Figure 7. Data presented in Figure 7 suggested that 24 hours post-transfection is the ideal time point to look for an interferon response to AAV infection.

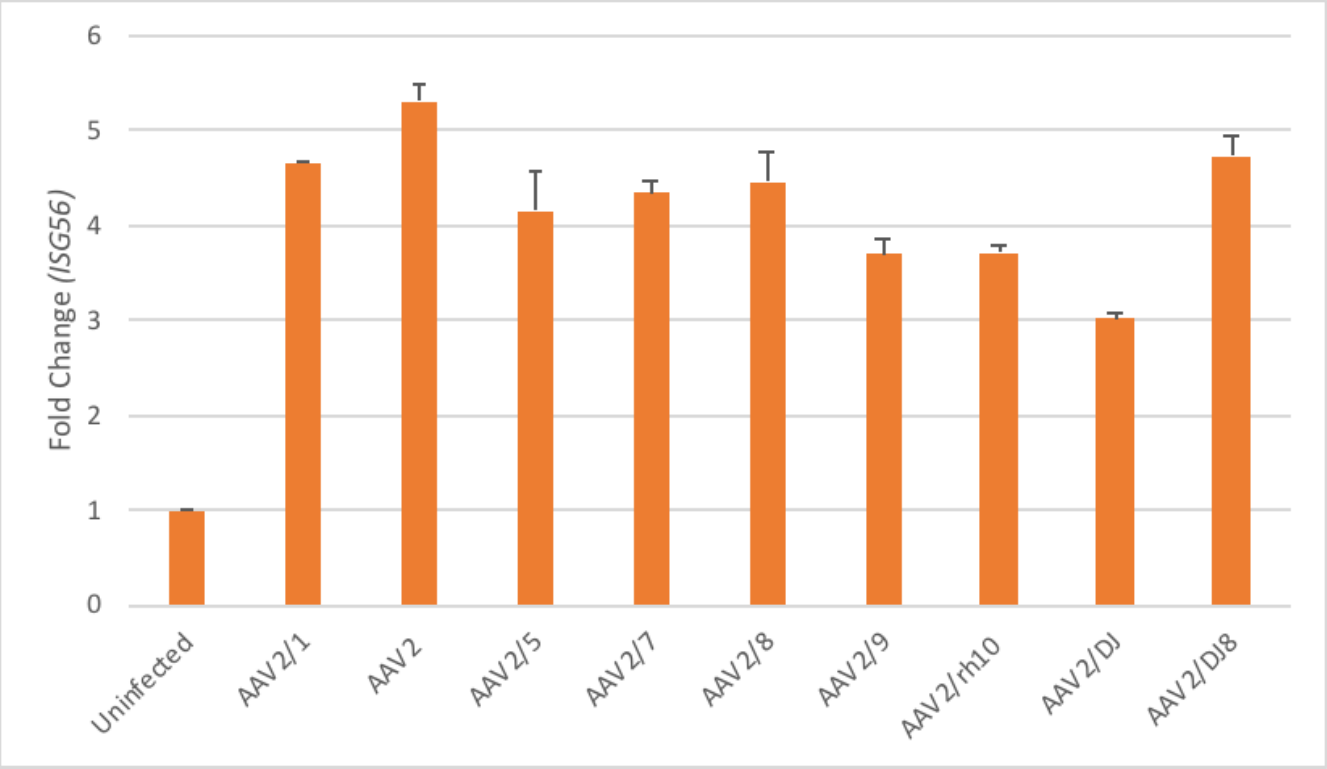
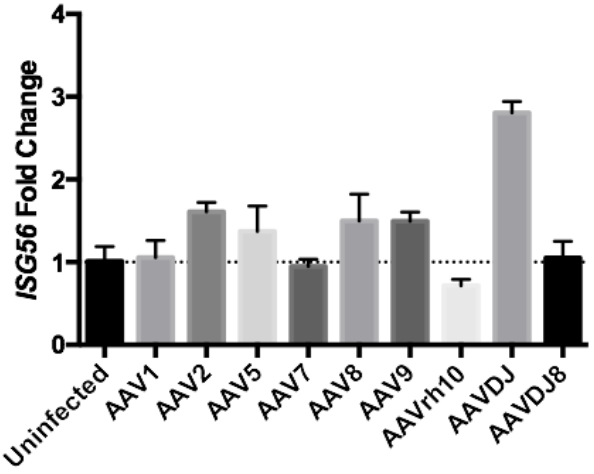


Figure 8: AAV serotypes sensed at 8 hours

PMA-treated THP-1 cells infected with AAV2 serotypes for 8 hours (MOI of 2). RNA samples were collected and cDNA was synthesized. *ISG56* transcripts were measured using RT-qPCR. Fold changes are all calculated relative to uninfected. Data are representative of the mean of technical triplicates and error bars represent the standard deviation. These data are representative of 3 experiments.

As seen in Figure 9a, some samples that were infected with AAV for 24 hours demonstrated an increase in *ISG56* transcription compared to uninfected. Five out of nine samples infected with different AAV2 serotypes demonstrated noticeable increases in *ISG56* transcription compared to uninfected. Production *ISG56* transcripts varied between AAV2 serotypes (demonstrating a fold change range of 0.71-2.8). Likewise, as seen in Figure 9b, some samples that were infected with AAV for 24 hours demonstrated an increase in *ISG54* transcription compared to uninfected. Five out of nine samples infected with different AAV2 serotypes demonstrated noticeable increases in *ISG54* transcription compared to uninfected. Production of *ISG54* transcripts varied between AAV2 serotypes (demonstrating a fold change range of 0.64-1.63). Overall, these data suggest that ssDNA does trigger an interferon response when delivered to the nucleus through AAV infection.

A.



B.

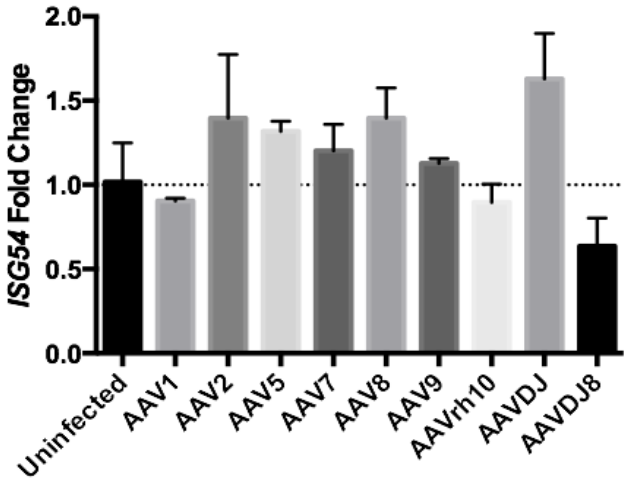


Figure 9: AAV serotypes sensed at 24 hours

PMA-treated THP-1 cells infected with AAV2 serotypes for 24 hours (MOI of 2). RNA samples were extracted and cDNA was synthesized. Transcripts were measured by RT-qPCR. A. *ISG56* transcripts measured B. *ISG54* transcripts measured. Fold changes are all calculated relative to uninfected. Data are representative of the mean of technical triplicates and error bars represent the standard deviation. These data are representative of 3 experiments.

Discussion

The aim of this study was to investigate whether or not single-stranded viral DNA is sensed and initiates an interferon response. We hypothesized that due to previous data alluding to IFI16's ability to colocalize and bind single-stranded DNA (Jakobsen et al. 2013, Unterholzner et al. 2010, Monroe et al. 2014), that single-stranded DNA would be sensed and lead to a type I interferon response. Additionally, we hypothesized that single-stranded viral DNA would be sensed later than double-stranded viral DNA due to previous information on double-stranded DNA synthesis during single-stranded viral replication (Flint et al. 2015). In order to understand single-stranded viral DNA sensing in the cytoplasm, we examined the type I interferon response of matured THP-1 macrophages stimulated with single-stranded and double-stranded viral DNA from vaccinia virus at various different time points. Type I IFN responses to single-stranded and double-stranded viral nucleic acid were measured by quantifying the transcription of *ISG56*, an antiviral gene upregulated during a type I IFN response. Overall, these data demonstrate that ssDNA in the cytoplasm is sensed and leads to an innate immune response, though kinetically later than dsDNA.

We also hypothesized that single-stranded viral DNA should be sensed in both the cytoplasm and the nucleus due to IFI16's and cGAS' ability to translocate between the nucleus and the cytoplasm (Unterholzner et al. 2010, Gentile et al. 2019). In order to understand single-stranded viral DNA sensing in

the nucleus, we examined the type I interferon response of matured THP-1 macrophages stimulated with AAV2 serotypes, a ssDNA virus known to deliver its genome directly to the nucleus. These data are also of interest due to their clinical relevance for gene therapy studies that assume AAV infection does not result in an immune response. Type I IFN responses to viral nucleic acid were measured by quantifying the transcription of *ISG56* and *ISG54* through. Overall, these data suggest that ssDNA delivered to the nucleus by AAV2 serotypes is sensed and initiates an interferon response.

Single-Stranded Viral DNA in the Cytoplasm leads to a Type I IFN Response

The addition of transfected single-stranded viral DNA in the form of ssVAC70 resulted in an increased IFN response after 24 hours of stimulation compared to mock (Figure 7). However, this pattern was not seen after 6 hours or 48 hours of stimulation (Figure 7). This IFN response to transfected single-stranded viral DNA resulted kinetically later than the IFN response of transfected double-stranded viral DNA in the form of dsVAC70 (Figure 7). Double-stranded VAC70 has been shown to induce IFN and ISG responses as seen in the literature (Unterholzner et al. 2010). These data were confirmed by the increased level of *ISG56* transcription in dsVAC70 transfected cells in Figure 7. IFN response to dsVAC70 was detected at 6 hours post-transfection and peaked at 24 hours post-transfection (Figure 7). Single-stranded viral DNA sensors have yet to be described in the literature. However, these data provide some insight on

how ssDNA is being sensed. Transfection of ssVAC70 demonstrated an increased level of *ISG56* transcripts compared to mock (Figure 7). This IFN response to ssVAC70 was only seen at 24 hours post-transfection (Figure 7). This observation is interesting because unlike ssVAC70, dsVAC70 demonstrated IFN response as early as 6 hours. These data suggest that overall single-stranded viral DNA is sensed kinetically later than double-stranded viral DNA.

These data support our hypothesis that ssDNA may be sensed as dsDNA after DNA processing has taken place in the cell. These data may also suggest ssDNA is being sensed by an alternative pathway. A recent study has identified a STING-independent DNA sensing pathway that leads to an IFN response (Burleigh et al. 2020). Burleigh and colleagues demonstrated that DNA-PK, a DNA repair protein, functioned as a sensor for this newly described pathway (Burleigh et al. 2020). According to data, this STING-independent pathway seems to sense DNA much later than the STING-dependent pathway (Burleigh et al. 2020). It is quite possible that ssDNA is being sensed by this STING-independent pathway. ssDNA has already been shown to be acted upon by DNA-PK in the literature (Choi et al. 2006, Schwartz et al. 2009). These data, however, do not differentiate whether DNA-PK is involved in sensing ssDNA in Burleigh's alternative sensing pathway or if DNA-PK is repairing ssDNA into dsDNA that can be sensed by PRRs that sense dsDNA.

AAV2 Serotypes lead to a Type I IFN Response

Infection of mature macrophages with AAV2 serotypes for 8 hours or 24 hours demonstrated an IFN response (Figure 8, Figure 9). Surprisingly, infection for 8 hours demonstrated a strong IFN response regardless of serotype (Figure 8). Increased production of *ISG56* transcript this early post-infection was not expected based on data presented in Figure 7. Data presented in Figure 7 suggested that 24 hours post-transfection is the ideal time point to look for an interferon response to AAV infection. Samples that were infected for 24 hours demonstrated a wide range of IFN responses (Figure 9). All AAV serotypes delivered AAV2 single-stranded genomic DNA. Therefore, there should not be a difference in the type of sensing if not for differences in the ability to uncoat viral capsids and ability to deliver the genome seen in each serotype. Overall, these data suggest that AAV2 serotypes do induce an innate immune response. This is significant due to the fact that AAV is used for gene therapy techniques due to the fact that it is generally described as leading to relatively low to no levels of immune activation, contrary to our data.

It is also important to note that differences in transduction efficiency may have altered IFN production in this study (Figure 8, Figure 9). Transduction efficiency describes how well a virus delivers the gene of interest to its cell. In previous literature, it was found that some AAV serotypes seem to display better transduction efficiency than others (Agbandje-McKenna et al. 2011, Merkel et al. 2017). However, the reason for this distinction remains largely unknown. There is

a possibility that certain AAV serotypes induce larger amounts of IFN and are inhibited from delivering their gene of interest. AAV serotypes that induce less amounts of IFN would be able to deliver their gene of interest and, therefore, have a better transduction efficiency. Understanding whether there is a role for serotype in IFN production and whether it is related to transduction efficiency would be particularly important in the study of AAVs in order to find the optimal gene delivery *in vitro* or *in vivo*.

Single-Stranded DNA in the Nucleus leads to a Type I IFN Response

Through the use of AAV2 infection, we are able to study the innate immune response to single-stranded viral DNA in the nucleus without any contribution of cytoplasmic DNA sensing. This can be done because AAV waits until it is inside the nucleus before it uncoats its capsid and delivers its genomic DNA into the nucleus (Porwal et al. 2013). The presence of single-stranded viral DNA in the nucleus delivered through AAV infection resulted in an increased IFN response after 8 hours of stimulation compared to mock (Figure 8). This indicates that single-stranded viral DNA is being sensed in the nucleus to lead to an innate immune response. The variation in IFN response after 24 hours of infection (Figure 9) should not be due to differences in the sequence of the single-stranded viral DNA since all AAV serotypes delivered AAV2 genomic DNA. We hypothesize that the difference in *ISG56* and *ISG54* transcript production is due to differences in the ability to uncoat viral capsids and ability to deliver the

genome seen in each serotype (Agbandje-McKenna et al. 2011, Merkel et al. 2017).

These experiments (Figure 8, Figure 9), however, are not without their limitations. The AAV used for this study could not replicate due to the lack of a helper virus (Naso et al. 2017). Therefore, there was no amplification of the viral ssDNA in the cell to induce further detection. Furthermore, in these experiments, THP-1 cells were infected with AAV at a concentration of 2×10^{10} genome copies/mL, which may be less than 4 ug/mL of ssDNA delivered through transfection. This may be why the fold change differences in these experiments (Figure 8, Figure 9) differ greatly from the fold changes seen when ssDNA is transfected into the cell (Figure 7). If we were to use comparable amounts of ssDNA, we hope to get comparable sizes of responses. In order to determine the amount of DNA in the AAV dose, one would have to develop a standardized qPCR assay (Dobnik et al. 2019). Due to time constraints, we could not quantify the dose of ssDNA being delivered by AAV and therefore, we could not include a positive control in our experiments at the same dose (Figure 8, Figure 9). A positive control was also not included due to the fact that it would have to be delivered through transfection with Lipofectamine, which has its own effect on IFN responses seen in Figure 7. Future studies should develop a standardized qPCR assay for AAV quantification. This would allow for our lab to quantify the AAV genome copies and develop a positive control to make data comparable.

Future Directions and Relevance

Future studies are necessary to identify how single-stranded viral DNA is being sensed. In order to do so, one could use THP-1 cell lines that have been modified with CRISPR to produce *IFI16*, *cGAS*, and *STING* knockout cell lines. These knockout cell lines could then be stimulated with single-stranded viral DNA either through transfection into the cytosol or infection into the nucleus through AAV delivery. These samples would then be monitored for an IFN response. This set of experiments will help us identify whether single-stranded viral DNA is being sensed by the IFI16-cGAS-STING-mediated pathway. This would be interesting to investigate especially since IFI16 has been shown to colocalize with and bind ssDNA (Unterholzner et al. 2010, Monroe et al. 2014).

Future studies can also investigate whether ssDNA is being sensed by the STING-independent pathway described by Burleigh and colleagues (Burleigh et al. 2020). This can be done by stimulating THP-1 cells that have been treated with a DNA-PK inhibitor with single-stranded viral DNA. These samples can then be monitored for an IFN response. IFN secretion can also be measured directly from supernatants saved throughout past experiments using a HEK Blue Assay or ELISA.

Overall, the data presented in this study have demonstrated that ssDNA sensors should be investigated. Whether transfected into the cytosol or delivered into the nucleus, ssDNA induces an innate immune response. Therefore, it must be sensed by some PRR, though there are currently no ssDNA sensors

described in the literature. These data are also of particular interest due to their clinical relevance in gene therapy research. Contrary to the literature, AAV infection does cause an innate immune response. The fact that AAV was found to induce an innate immune response may have many implications on gene therapy delivery techniques.

References

Abbas AA, Diamond JM, Chehoud C, Chang B, Kotzin JJ, Young JC, et al. 2016. The perioperative lung transplant virome: torque teno viruses are elevated in donor lungs and show divergent dynamics in primary graft dysfunction. Am. J. Transplant. 17:1313–1324.

Agbandje-McKenna M, Kleinschmidt J. 2011. AAV capsid structure and cell interactions. Methods Mol Biol. 807:47-92.

Almine J, O'Hare C, Dunphy G, Haga IR, Naik RJ, Atrih A, Connolly DJ, Taylor J, Kelsall IR, Bowie AG, Beard PM, Unterholzner L. 2017. IFI16 and cGAS cooperate in activation of STING during DNA sensing in human keratinocytes. Nat Commun 8:1-15.

Bandera A, Masetti M, Fabbiani M, Biasin M, Muscatello A, Squillace N, Clerici M, Gori A, Trabattoni D. 2018. The NLRP3 inflammasome is upregulated in HIV-infected antiretroviral therapy-treated individuals with defective immune recovery. Frontiers in immunology 12:1-7.

Barber GN. 2011. Cytoplasmic DNA innate immune pathways. Immunol. Rev. 243:99–108.

Bowie AG, Unterholzner L. 2008. Viral evasion and subversion of pattern-recognition receptor signalling. Nature Reviews Immunology 8(12):911-922.

Brunette RL, Young JM, Whitley DG, Brodsky IE, Malik HS, Stetson DB. Extensive evolutionary and functional diversity among mammalian AIM2-like receptors. J Exp Med. 2012; 209(11):1969–83.

Burleigh K, Maltbaek JH, Cambier S, Green R, Gale M, James, RC, Stetson DB. 2020. Human DNA-PK activates a STING-independent DNA sensing pathway. Sci. Immunol. 5:eaba4219

Cai X, Chiu YH, Chen ZJ. 2014. The cGAS-cGAMP-STING pathway of cytosolic DNA sensing and signaling. Mol Cell 54:289-296.

Chen GY, Nunez G. 2010. Sterile inflammation: sensing and reacting to damage. Nat Rev Immunol 10:826-837.

Chen Q, Sun L, Chen Z. 2016. Regulation and function of the cGAS-STING pathway of cytosolic DNA sensing. Nat Immunol 17:1142-1149.

Chew WL, Tabebordbar M, Chng JK, Mali P, Wu EY, Ng AH, Zhu K, Wagers AJ, Church GM. 2016. A multifunctional AAV-CRISPR-Cas9 and its host response. Nat Methods. 13(10):868-74.

Choi VW, McCarty DM, Samulski RJ. 2006. Host cell DNA repair pathways in adeno-associated viral genome processing. J. Virol. 80(21):10346-10356

Choubey D, Panchanathan R. 2016. IFI16, an amplifier of DNA-damage response: Role in cellular senescence and aging-associated inflammatory diseases. Ageing Research Reviews 28:27-36.

Civril F, Deimling T, de Oliveira Mann CC, Ablasser A, Moldt M, Witte G, Hornung V, Hopfner K-PP. 2013. Structural mechanism of cytosolic DNA sensing by cGAS. Nature 498:332-337.

Dawson MJ, Trapani JA. The interferon-inducible autoantigen, IFI 16: localization to the nucleolus and identification of a DNA-binding domain. Biochem Biophys Res Commun. 1995; 214(1):152–62.

Diner BA, Li T, Greco TM, Crow MS, Fuesler JA, Wang J, Cristea IM. 2015. The functional interactome of PYHIN immune regulators reveals IFIX is a sensor of viral DNA. Molecular Systems Biology 11:1-22.

Dobnik D, Kogovsek P, Jakomin T, Kosir N, Znidaric MT, Leskovec M, Kaminsky SM, Mostrom J, Lee H, Maja R. Accurate quantification and characterization of Adeno-Associated viral vectors. 2019. Front Microbiology. 10:1-13.

Fisher KJ et al. 1997. Recombinant adeno-associated virus for muscle directed gene therapy. Nat Med. 3:306-312.

Franzoso FD, Seyffert M, Vogel R, et al. Cell Cycle-Dependent Expression of Adeno-Associated Virus 2 (AAV2) Rep in Coinfections with Herpes Simplex Virus 1 (HSV-1) Gives Rise to a Mosaic of Cells Replicating either AAV2 or HSV-1. 2017. J Virol. 91(15):e00357-17.

Freer G, Maggi F, Pifferi M, Di Cicco ME, Peroni DG, Pistello M. The Virome and Its Major Component, Anellovirus, a Convoluted System Molding Human Immune Defenses and Possibly Affecting the Development of Asthma and Respiratory Diseases in Childhood. 2018. Front Microbiol. 9:686.

Gentile M, Lahaye X, Nadlin F, Nader GPF, Lombardi EP, Herve Solene, De Silva NS, Rookhuizen DC, et al. 2019. The N-terminal domain of cGAS determines preferential association with centromeric DNA and innate immune activation in the nucleus. Cell Reports. 26(9):2377-2393

Grimm D, Lee JS, Wang L, Desai T, Akache B, Storm TA, Kay MA. 2008. In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. J Virol. 82(12):5887-5911.

Hall J, Ralph EC, Suman S, Wang H, Byrnes LJ, Horst R, Wong J, Brault A, Dumlao D, Smith JF, Dakin LA, Schmitt DC, Trujillo J, Vincent F, Griffor M, Aulabaugh AE. 2017. The catalytic mechanism of cyclic GMP-AMP synthase (cGAS) and implications for innate immunity and inhibition. Protein Sci. 26(12):2367-2380.

Haronikova L, Coufal J, Kejnovska I, Jagelska EB, Fojta M, et al. 2016. IFI16 preferentially binds to DNA with quadruplex structure and enhances DNA quadruplex formation. PLOS ONE 11(6):1-19.

Herzner AM, Hagmann CA, Goldeck M, Wolter S, Kubler K, Wittmann S, Gramberg T, Andrea L, Hopfner KP, Martens C et al. 2015. Sequence-specific activation of the DNA sensor cGAS by Y-form DNA structures as found in primary HIV-1 cDNA. Nat Immunol 16(10):1025-33.

Honda K, Takaoka A, Taniguchi T. 2006. Type I Interferon Gene Induction by the Interferon Regulatory Factor Family of Transcription Factors. Immunity. 25:349-360.

Iwasaki A, Medzhitov R. Regulation of adaptive immunity by the innate immune system. Science. 2010;327:291–295.

Jakobsen MR, Bak RO, Andersen A, Berg RK, Jensen SB, Jin T, Laustsen A, Hansen K, Ostergaard L, Fitzgerald KA, Xiao TS, Mikkelsen JG, Mogensen TH, Paludan SR. 2013. IFI16 senses DNA forms of the lentiviral replication cycle and controls HIV-1 replication. Proceedings of the National Academy of Sciences 110(48):E4571-E4580.

Jiang F, Ramanathan A, Miller MT, Tang GQ. 2011. Structural basis of RNA recognition and activation by innate immune receptor RIG-I. Nature 479(7373):423-427.

Jin, T. et al. Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor. *Immunity* 36, 561–571 (2012).

Jooss K, Chirmule N. 2003. Immunity to adenovirus and adeno-associated viral vectors: implications for gene therapy. Gene Therapy 10:955-963.

Kotin RM, Siniscalco M, Samulski RJ, Zhu XD, Hunter L, Laughlin CA, McLaughlin S, Muzyczka N, Rocchi M, Berns KI. 1990. Site-specific integration by adeno-associated virus. Proc Natl Acad Sci U S A. 87(6):2211-5.

Livak KJ, Schmittgen TD. 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C}$T Method. Methods. 25**:**402–408.

Maess MB, Sendelbach S, Lorkowski S. 2010. Selection of reliable reference gene during THP-1 monocyte differentiation into macrophages . BMC Mol. Bio. 11, 90.

Matsushita T, Elliger S, Elliger C, Podsakoff G, Villarreal L, Kurtzman GJ, Iwaki Y, Colosi P. 1998. Adeno-associated virus vectors can be efficiently produced without helper virus. Gene Therapy 5:938-945.

Merkel SF, Andrews AM, Lutton EM, Mu D, Hudry E, Hyman BT et al. 2017. Trafficking of adeno-associated virus vectors across a model of the blood-brain barrier; a comparative study of transcytosis and transduction using primary human brain endothelial cells. J Neurochem. 140(2):216-230.

Monroe KM, Yang Z, Johnson JR, Geng X, Doitsh G, Krogan NJ, Greene WC. 2014. IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV. Science. 343:428-432.

Munoz RF, Darnell JE. Structural difference between the 5' termini of viral and cellular mRNA in poliovirus-infected cells: possible basis for the inhibition of host protein synthesis. 1976. J Virol. 18(2):719–726.

Naso MF, Tomkowicz B, Perry WL, Strohl WR. 2017. Adeno-Associated Virus as a vector for gene therapy. BioDrugs. 31(4):317-334.

Orzalli MH, Conwell SE, Berrios C, DeCaprio JA, Knipe DM. 2013. Nuclear interferon-inducible protein 16 promotes silencing of herpesviral and transfected DNA. Pro Natl Acad Sci USA 110:E4492-E4501.

Parham P. 2015. The immune system. 4th edition. New York (NY) Garland Science 532p.

Perry AK, Gang C, Zheng D, Hong T, Cheng G. 2005. The host type I interferon response to viral and bacterial infections. 15(6):407-422.

Piehler J, Thomas C, Garcia KC, Schreiber G. 2012. Structural and dynamic determinants of type I interferon receptor assembly and their functional interpretation. Immunol Rev. 250(1):317-334.

Porwal M, Cohen S, Snoussi K, Popa-Wagner R, Anderson F, Dugot-Senant N, et al. 2013. Parvoviruses cause nuclear envelope breakdown by activating key enzymes of mitosis. PLoS Pathog 9(10): e1003671.

Punt J, Owen JA, Stanford SA, Jones PP, Kuby J. Kuby immunology. 8th ed. New York: W.H. Freeman/Macmillan Learning; 2019.

Rotem Z, Cox RA, Isaacs A. 1963. Inhibition of virus multiplication by foreign nucleic acid. Nature 197:564-566.

Samulski RJ, Muzyckzka N. 2014. AAV-mediated gene therapy for research and therapeutic purposes. Annu Rev Virol. 1(1):427-451.

Schneider WM, Chevillotte MD, Rice CM. 2014. Interferon-Stimulated Genes: A Complex Web of Host Defenses. Annu Rev Immunol. 32: 513–545.

Schwartz RA, Carson CT, Schuberth C, Weitzman MD. 2009. Adeno-associated virus replication induces a DNA damage response coordinated by DNA-dependent protein kinase. J. Virol. 83:6269-78.

Senis E, Fatouros C, Grobe S, Liedtke E, Niopek D, Mueller A, Borner K, Grimm D. 2014. CRISPR/Cas9-mediated genome engineering: An adeno-associated viral (AAV) vector toolbox. Biotechnology Journal 9:1402-1412.

Shannon JL, Murphy MS, Kantheti U, Dorrity TJ, Burnett JM, Hahn MG, Bacas CJ, Mattice EB, Corpus KD, Barker BR. 2018. Polyglutamine binding protein 1 (PQBP1) inhibits innate immune responses to cytosolic DNA. Molecular Immunology 99:182-190.

Ulloa L, Tracey KJ. The “cytokine profile”: A code for sepsis. 2005. Trends Mol Med. 11(2):56–63.

Unterholzner L, Keating SE, Baran M, Horan KA, Jensen SB, Sharma S, Sirois CM, Jin T, Latz E, Xiao TS, et al. 2010. IFI16 is an innate immune sensor for intracellular DNA. Nat Immunolo. 11(11):997-1004.

Unterholzner L, Dunphy G. 2019. cGAS-independent STING activation in response to DNA damage. Mol Cell Oncol. 6(4):e1558682

Vogel R, Seyffert M, Strasser R, et al. Adeno-associated virus type 2 modulates the host DNA damage response induced by herpes simplex virus 1 during coinfection.2012. J Virol. 86(1):143–155.

Volkman HE, Cambier S, Gray EE, Stetson DB. 2018. cGAS is predominantly a nuclear protein. bioRxiv [Internet]. Available from: https://www.biorxiv.org/content/biorxiv/early/2018/12/04/486118.full.pdf

Walsh D, McCarthy J, O’Driscoll C, Melgar S. 2013. Pattern recognition receptors-molecular orchestrators of inflammation in inflammatory bowel disease. Cytokine Growth Factor Rev. 24:91-104.

Wu J, Sun L, Chen X, Du F, Shi H, Chen C, Chen ZJ. 2013. Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. Science. 339:826-830.

Yan H, Dalal K, Hon BK, Youkharibache P, Lau D, Pio F. 2008. RPA nucleic acid-binding properties of IFI16-HIN200. Biochem Biophys Acta. 1784:1087-1097.

Yoshimura A, Hara Y, Kaneko T, Kato I. 1997. Secretion of IL-1B, TNF-a, IL-8 and IL-1 by human polymorphonuclear leukocytes in response to lipopolysaccharides from periodontopathic bacteria. Journal of periodontal research. 32(3):279-286.

Young JC, Chehoud C, Bittinger K, Bailey A, Diamond JM, Cantu E, et al. 2015. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. *Am. J. Transplant.* 15:200–209.

Zaiss A, Cotter MJ, White LR, Clark SA, Wong NCW, Holers VM, Bartlett JS, Muruve DA. 2008. Complement Is an Essential Component of the Immune Response to Adeno-Associated Virus Vectors. J Virol. 82(6):2727-40.