**SANDIA REPORT**
SAND2011-6924
Unlimited Release
Printed September 2011

# Investigation of Type-I Interferon Dysregulation by Arenaviruses: A Multidisciplinary Approach

Bryan D. Carson, Catherine S. Branda, Conrad D. James, Carol L. Kozina, Ronald P. Manginell, Milind Misra, Matthew W. Moorman, Oscar A. Negrete, J. Bryce Ricken, Meiye Wu

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.



Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.



Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2011-6924
Unlimited Release
Printed September 2011

# Investigation of Type-I Interferon Dysregulation by Arenaviruses: A Multidisciplinary Approach

Bryan D. Carson[1], Catherine S. Branda[2], Conrad D. James[3], Carol L. Kozina[4], Ronald P. Manginell[5], Milind Misra[6], Matthew W. Moorman[5], Oscar A. Negrete[7], J. Bryce Ricken[1], Meiye Wu[7]

Bioenergy and Defense Technologies[1], Systems Biology[2], Biosensors and Nanomaterials[3], Biomass Science and Conversion Technologies[4], Microsystems-Enabled Detection[5], Advanced Device Technologies[6], Biotechnology and Bioengineering[7]

Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-1413

**Abstract**

This report provides a detailed overview of the work performed for project number 130781, "A Systems Biology Approach to Understanding Viral Hemorrhagic Fever Pathogenesis." We report progress in five key areas: single cell isolation devices and control systems, fluorescent cytokine and transcription factor reporters, on-chip viral infection assays, molecular virology analysis of Arenavirus nucleoprotein structure-function, and development of computational tools to predict virus-host protein interactions. Although a great deal of work remains from that begun here, we have developed several novel single cell analysis tools and knowledge of Arenavirus biology that will facilitate and inform future publications and funding proposals.

# Acknowledgments

Brad Hance (SNL) – Configuration of temperature controllers and lots of troubleshooting of the iCellator control system

Norbert Herzog (University of Texas, Medical Branch) – Consultation in project design, Arenavirus biology and protocols; live Pichinde virus stocks

Jeffrey Lucero (UNM/SNL) – Student intern, assisted in running/troubleshooting iCellator experiments and developing protocols

Jaewook Joo (University of Tenessee, Knoxville) – Assisted in initial efforts to develop kinetic models for viral interference with NF-KappaB signaling. This work was put on hold upon Dr. Joo's departure from SNL.

Thomas Perroud (Illumina, Inc.) – Drove initial efforts to develop on-chip flow cytometry in conjunction with M. Wu. This work was still in early stages when Dr. Perroud left SNL, so is not detailed in this report.

Daniel Porter (UNM) – Outstanding student intern who wrote the LabView application used to control the iCellator system, as well as designing the initial version of the pinch valve system.

Ronald F. Renzi (SNL) – Excellent work constructing manifolds, valves, and other fluidic control hardware

Nimisha Srivastava (UCSB) – Assisted in initial design of the project and early designs for high-throughput microfluidic cell assay chips

Joanne Volponi – Assisted Dr. Branda in construction of p105 and IL-12 reporters

James Van De Vreugde – Designed and built various electronic components for the fluidic control systems

# Contents

# Figures

# Tables

# Nomenclature

| | |
|---|---|
| DOE | Department of Energy |
| FRET | Forster Resonant Energy Transfer |
| GFP | Green Fluorescent Protein |
| GUI | Graphical User Interface |
| IFN | Interferon |
| IFN-I | Type I Interferon |
| IP | Immunoprecipitation |
| NP | Nucleoprotein |
| PDMS | Polydimethylsiloxane |
| PICV | Pichinde virus |
| PPI | Protein-Protein Interaction |
| SCA | Single Cell Array |
| SNL | Sandia National Laboratories |
| SVM | Support Vector Machine |
| TACV | Tacaribe virus |
| TNFalpha | Tumor Necrosis Factor Alpha |

# 1. INTRODUCTION

Arenaviruses such as Lassa cause lethal hemorrhagic fever in humans and are NIAID category A pathogens of bioterror concern. They may be transmitted by airborne routes and have incubation times under two weeks with mortality up to 30%. A fundamental problem in understanding their pathogenicity is that infected and uninfected cells can exchange information that reciprocally influences their behavior – cell-intrinsic versus extrinsic effects (Figure 1.1).

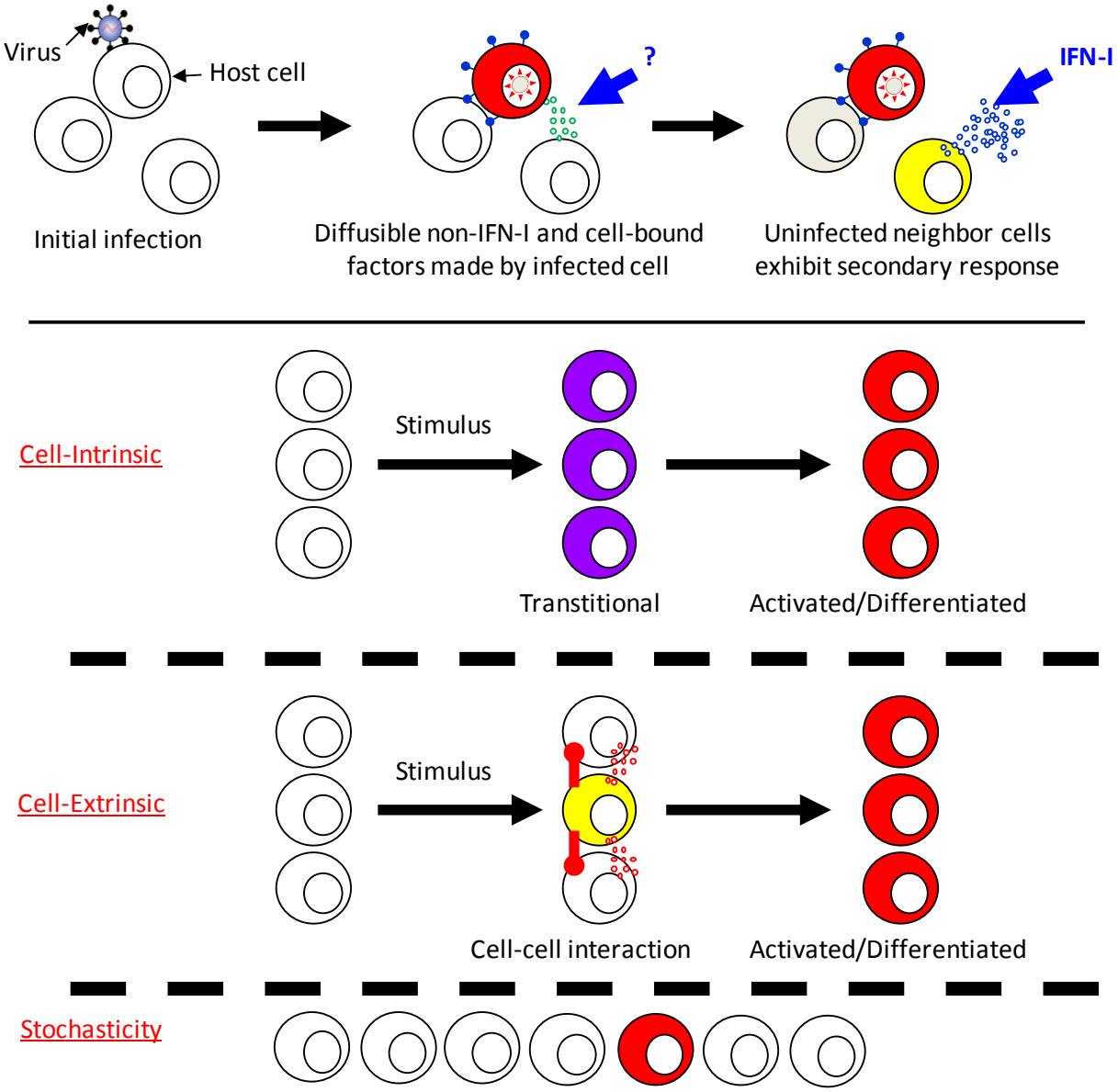



**Figure 1.1 Dissecting primary vs. secondary and intrinsic vs. extrinsic vs. stochastic events in cellular response**

Thus, cell population level experiments will never tell us why some people survive while others die from Lassa fever because many ill effects of viral infection are not mediated directly by the virus but by the response of the immune system. Excessive type I interferons (IFNα/β) drive hemorrhagic symptoms, but arenaviruses paradoxically appear to block production of these cytokines. We hypothesize that this apparent contradiction is due to differential effects of the virus on the infected cell versus uninfected neighboring cells.

This hypothesis is impossible to test by conventional means which asynchronously infect thousands of cells simultaneously, masking the difference between first and second order effects. Our idea is to deconvolute this system by isolating and infecting individual cells then performing unprecedented measurements of IFNα/β and other response-critical cytokines such as TNFα, IL-8, and IP-10 using novel fluorescent transcriptional reporters and in situ solid phase immunodetection. Using a Sandia-developed microfluidic cell isolation platform, we compared isolated cell to population level infection. We sought to elucidate differences under each of these conditions using both experimental and computational techniques to discover how these viruses provoke lethal cytokine production. We compared lethal and benign strains of the same virus to identify key virulence correlates and confirm our results were linked to virulence. The devices we used were originally developed to study toll-like receptor signaling in response to bacteria. However, with several changes and new experimental protocols, we were able to use these flexible devices to address otherwise impossible fundamentally different biological questions. By setting the stage to elucidate the signals and order of events that induce pathogenic cytokine production, we hope this work will generate science and technology advances to lay the groundwork for improved antiviral therapeutics, vaccines, and biological countermeasures to meet Sandia's core national security missions.

In this report, we describe the progress we made as well as the difficulties and setbacks we encountered. Although some of our initial goals remain incomplete at the end of this project, we had several great successes. The first section describes the designs and advances in microfluidic cell isolation systems we developed. The next describes the molecular tools we developed for measuring responses to infection in live individual cells. Section 5 details some controversial and intriguing findings in our work analyzing the sequence-structure-function relationships in Arenavirus nucleoproteins. The last section describes the methods we are still in the process of developing to predict which host proteins each viral protein may interact with using two distinct types of algorithm and training set.

*Author's Note:* Due to time constraints and some version control difficulties, there may be a number of typographical and grammatical errors in this report. However, every effort has been made to ensure the technical accuracy of the information presented.

# 2. SINGLE CELL ISOLATION SYSTEM

## 2.1 Microfluidic System with iCellator Chip

We developed the Single Cell Array (SCA or iCellator) for monitoring immune response in the P388D1 macrophage-like cell line in response to the dsDNA analog poly(I-C) challenges (a substitute for live virus). The SCA chip architecture is shown in Figure 2.1. The device currently contains two separate chambers for running two experiments simultaneously. In addition, the individual chambers contain multiple inlets and outlets to facilitate the controlled delivery of single host cells to the cell traps. Cell traps are made in silicon wafers with a three-step deep reactive ion etch process to position cells adjacent to the coverslip lid of the SCA chip for optimum optical interrogation. Multiple inlets permit full fluidic-isolation of trapped cells from secretions released by cells located upstream of the device. This is a crucial capability that provides a significant step towards understanding primary and secondary immune signaling.

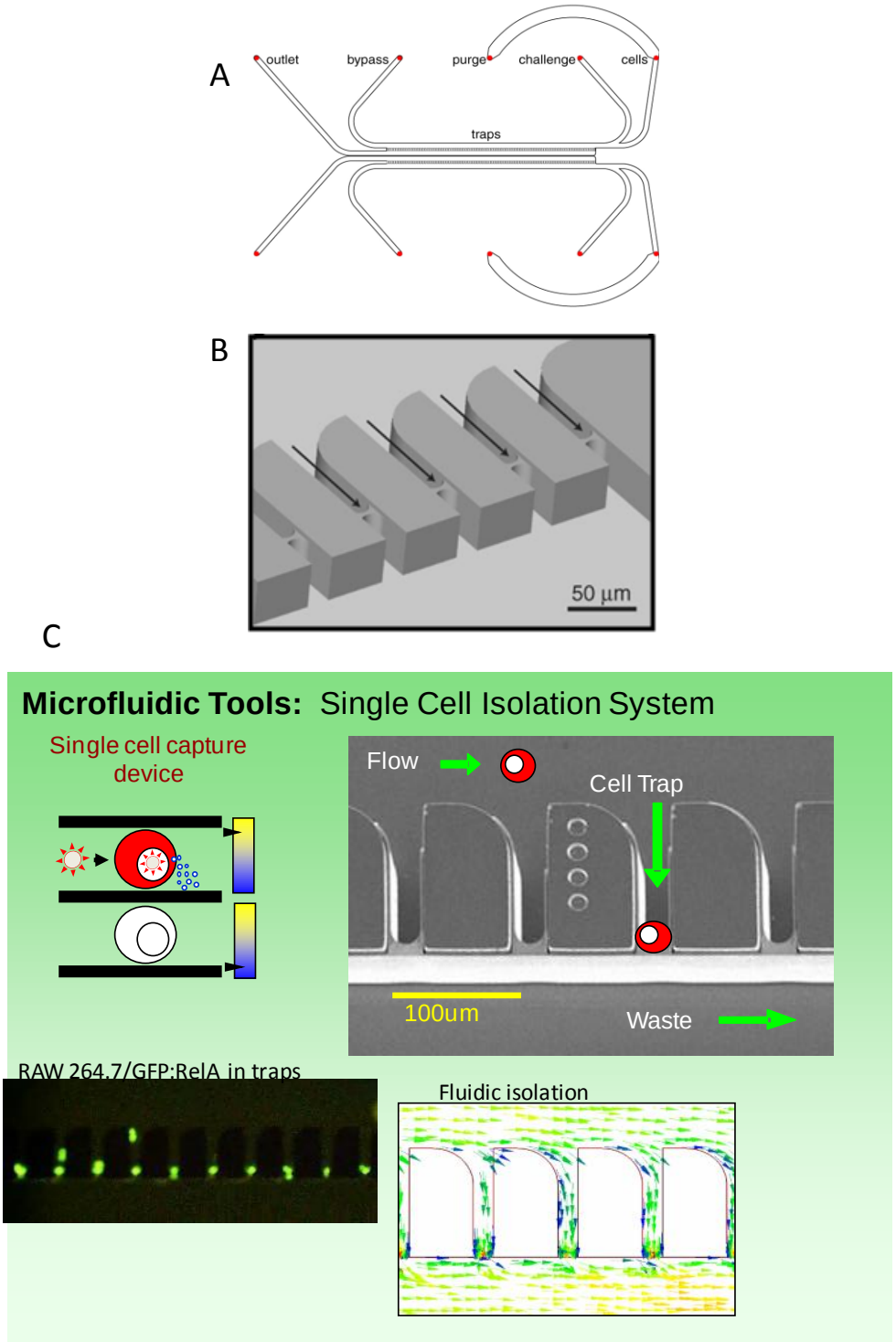



**Figure 2.1 A) SCA chip architecture for running multiple experiments with fluidically-isolated host cells. B) Schematic of the single cell traps. C) Overall scheme showing trapped cells and flow analysis showing fluidic isolation principle**

External hardware for the SCA has also been improved throughout the first year of the project. We are currently finishing up the fluidic control system that houses the pressure controllers, electronic valves, and temperature controllers for running experiments. Figure 2.2 shows the controller box currently near completion including software interfacing.

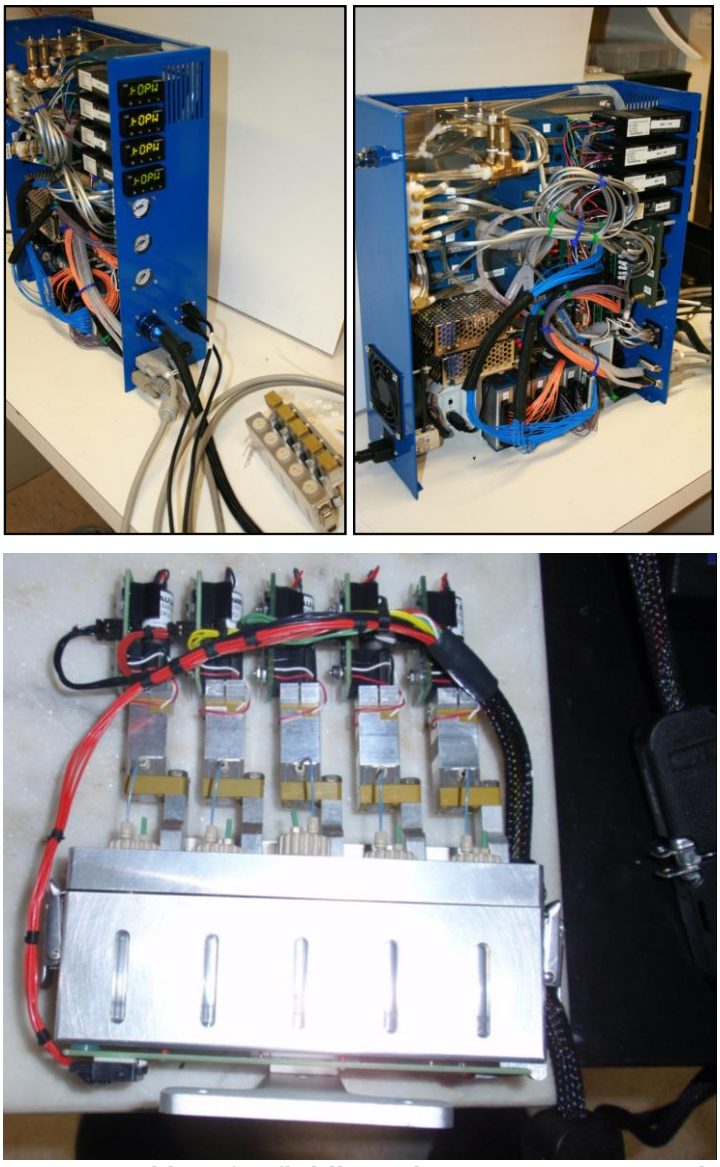


**Figure 2.2 (Top) SCA control box for fluidic and temperature control of the SCA during experiments. (Bottom) Valves and heated fluid reservoirs.**

The controller box currently contains 10 pressure controllers for delivering reagents to the SCA device with 8 electronically controlled valves and 2 manual valves for routing reagents throughout the device. Four temperature controllers are also included within the system for

maintaining physiological temperature at the reservoir block where reagents are stored and at the SCA chip. Currently, ten reservoirs, 5 in each reservoir block, can be implemented and independently controlled for experiments. USB cabling connects the National Instruments data acquisition module to the PC where the software interface resides. Recently, we upgraded the interface from the valves (which had originally been configured through a breakout box) to consist of a single serial cable.

## 2.2 Disposable Microfluidic Valves and LabView GUI

We developed a user-friendly software interface for running the microfluidic system. Figure 2.3 shows the interface created in Labview and which currently exists as an executable that can be run on any personal computer. All ten electronically-controlled valves and pressure controllers are controlled and have their status monitored on the graphical user interface. The temperature controllers for the chip and the two reservoirs are also controlled on this interface.

We have focused on developing disposable microfluidic valves for use with the pressure controller system used for on-chip immune response experiments (Figure 2.4). This will permit the use of this system for BSL3 and BSL4 applications where cleaning and decontamination of particular system components is not possible. Our strategy has been to utilize low-cost silicone microchannels and mechanical pressure to seal off channels. After use, the silicone microchannels can be disposed and a new set of channels placed into the mechanical apparatus for the next experiment. And due to the fact that the mechanical interface does not come into contact with the biological sample, the expensive components of the valving system are protected from contamination. The silicone microchannels are made by micromolding photolithographically-defined features. The original master is fabricated in SU8 and consists of 100 micron wide and 100 micron thick line features. After lithography, the uncured silicone is poured onto the master and cured at 60 C for 2 hours. The replica is then removed, plasma-treated, and then placed into contact with a flat sheet of silicone to create microfluidic channels with all four walls made out of silicone. Once bonded, PEEK tubing is interfaced to either end of the channels for use.

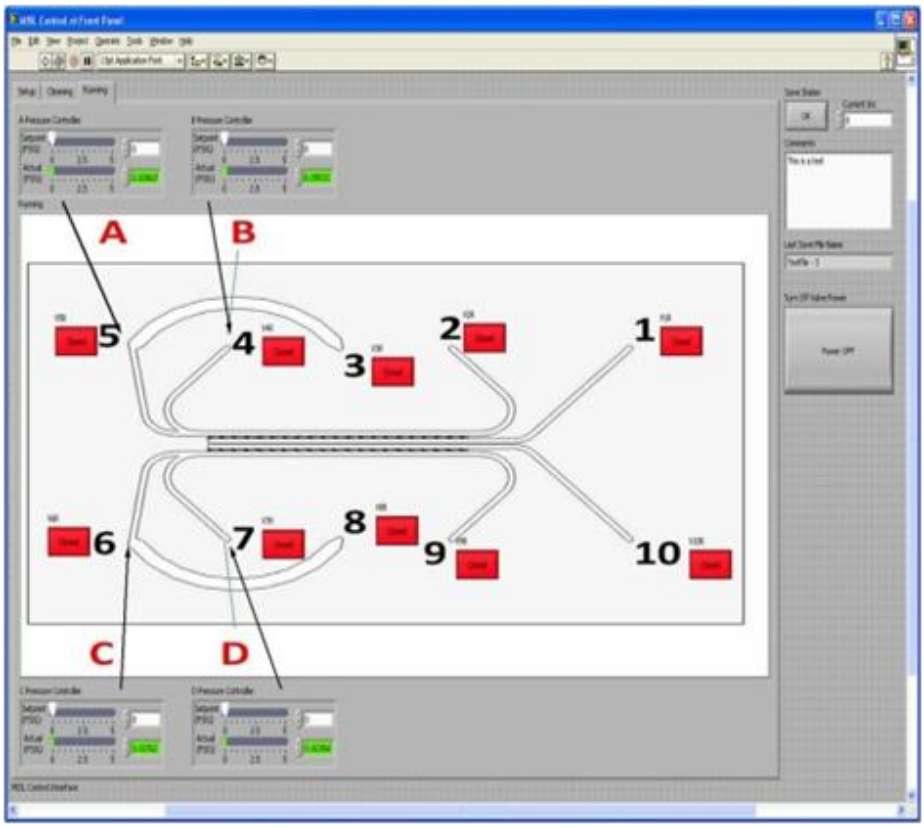



**Figure 2.3 Software interface for running the microfluidic system**

For valving the channels, the silicone device was placed onto a glass substrate and interfaced to a motorized piston. The rotary motor rotates upon actuation and a threaded piston moves perpendicular to the channel surface. We built a current sense circuit for the rotary motor in order to dynamically sense the current being drawn by the motor. This enables the motor to be shut off once a threshold current is reached – this occurs when the piston is compressing the silicon channel and begins to feel the resistance of the glass substrate under the channel. Once shut off, the piston does not rotate counter-clockwise but remains in position to keep the channel closed. We also developed a graphical user interface to monitor the current sense capability and to tune the shutoff threshold current level for particular silicone channels. Currently, the system is capable of shutting off flows of up to 100 microliters per minute using a 10-1 gear ratio motor.

Future work will include increasing the number of controlled valves to 10 in order to fully operate the pressure controller experimental set-up. We will also focus on developing a disposable silicone manifold for direct interfacing to the silicon microfluidic chip used to house live cells and conduct immune response experiments.

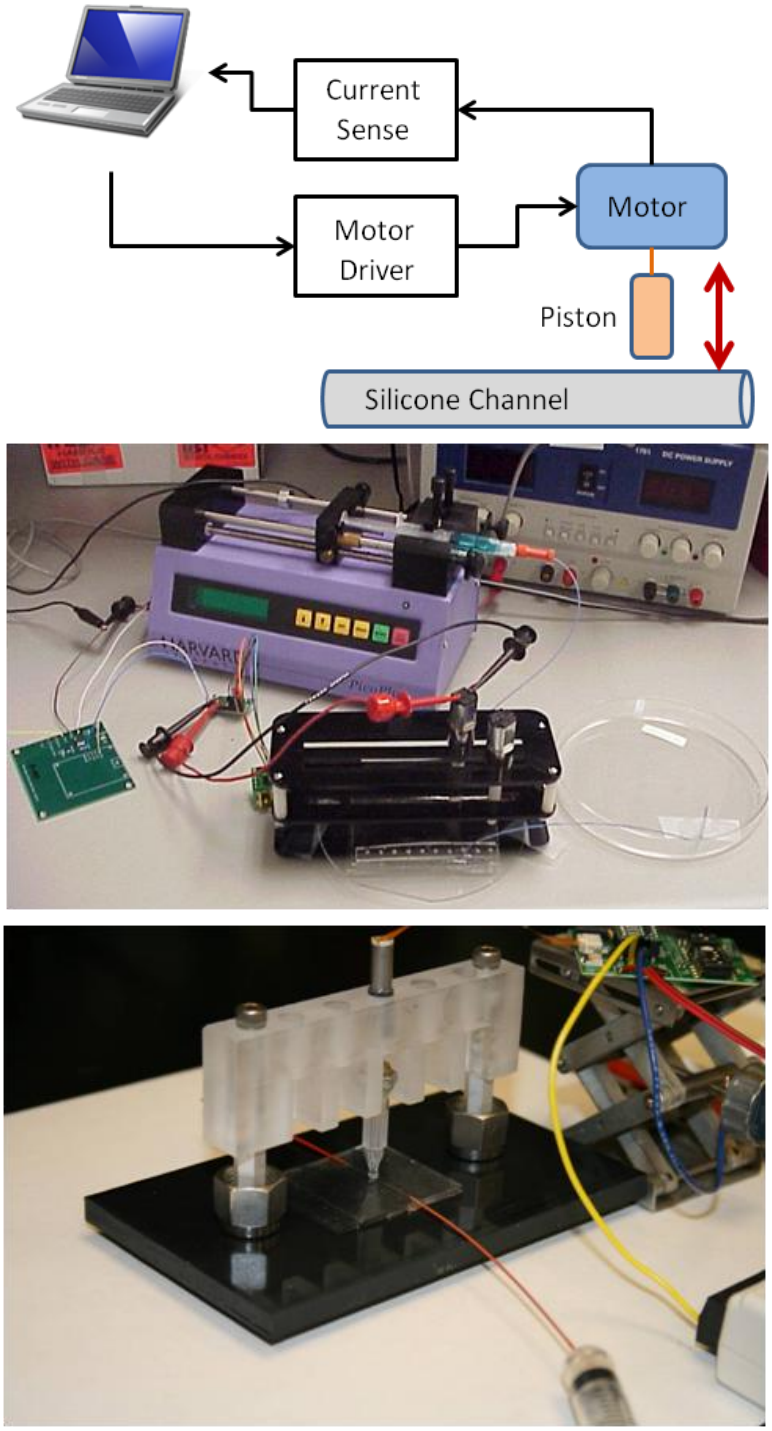



**Figure 2.4 (top) Schematic of the disposable microfluidic valve assembly. (middle) Image of the disposable microfluidic valve assembly. (bottom) Image of a silicone channel interfaced to tubing and a mechanically-actuated piston.**

## 2.3 Experimental Data and Notes

One of the primary objectives of the project is to demonstrate the ability to capture and maintain single cells in fluidic isolation from one another. Figure 2.5 shows an example of a set of RAW cells that have been captured in an iCellator chip.

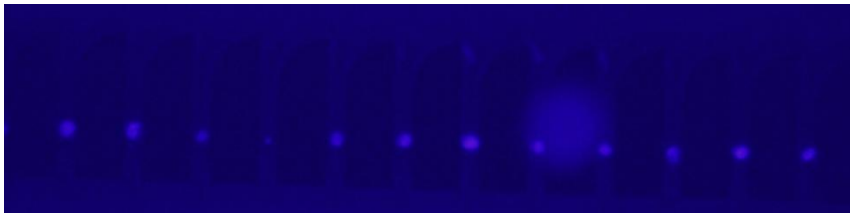


**Figure 2.5 A set of 12 RAW cells individually captured and isolated in microfluidic traps.**

The key to achieving this result was to reduce the cell density in the reservoir vial to approximately $10^4$ cells/ml and to utilize a low flow-rate for delivery of cells to the trap. We have currently demonstrated long-term culturing of macrophages for up to 3 days with no foreseeable difficulty for extending the observation time further. Current host cells contain up to two reporter constructs (described in more detail in Section 3), e.g. GFP-RelA for observing NF-kB translocation kinetics, and mCherry-TNFalpha for observing up- and down-regulation of the synthesis of the TNFalpha cytokine. Figure 2.6 shows an example of an experiment with P388D1 host cells challenged with 100 ug/ml poly(I-C). GFP-RelA translocates into the nucleus within 30 minutes and then returns to the cytoplasm over the course of the next 5 hours. The mCherry-TNFalpha signature is normalized to the pre-challenge fluorescence intensity, and over the course of six hours the fluorescence signal increases by 30%. Over the course of another 24 hours, the mCherry signal for this cell increases to 40% over the pre-challenge intensity. Population statistics for the long-term expression of mCherry-TNFalpha are currently being collected, in addition to validating the PEST sequence fused to the construct that enables tracking of degradation of TNFalpha mRNA.

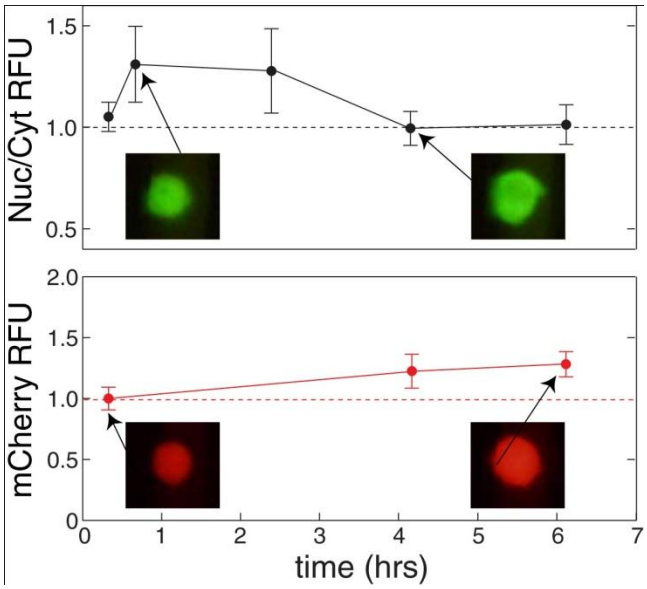



**Figure 2.6 Simultaneous tracking of RelA translocation and TNFalpha synthesis in a single host cell.**

One of our objectives was to facilitate long-term experiments on living cells. Figure 2.7 shows an example of RAW-TAPCP cells cultured in an iCellator chip for 18hours. The visible cell division indicates that cells are healthy and maintained in this environment.

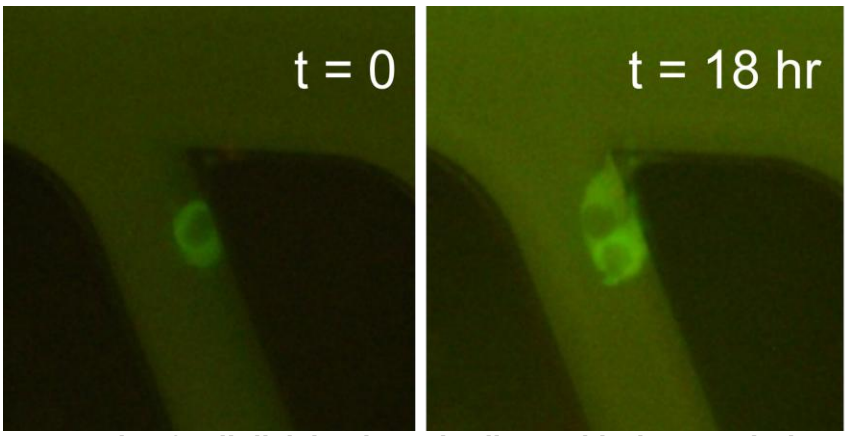



**Figure 2.7 Example of cell division in an iCellator chip (GFP:RelA in RAW 264.7)**

Figure 2.8 illustrates a long-term cell tracking experiment in which P388D1 cells expressing IFNbeta promoter : mCherry were stimulated poly I:C for 46 hours. Increase in red signal (mCherry) over time indicates IFNbeta promoter activity is just beginning to increase in intensity around 5-6hrs.

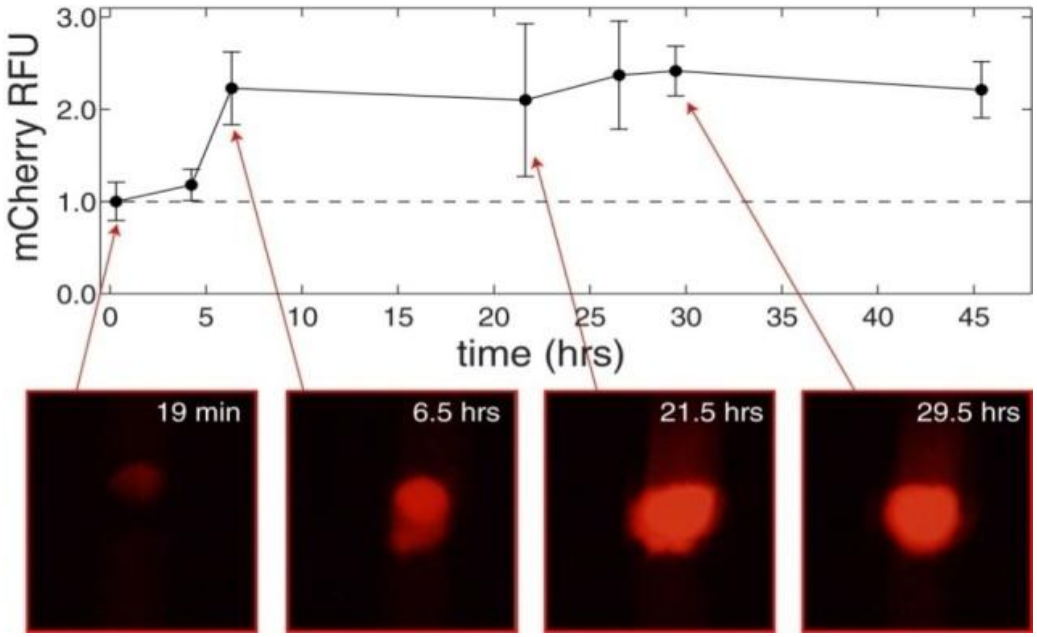



**Figure 2.8 Induction of IFNbeta transcription in an isolated macrophage-like cell by the synthetic viral mimic poly I:C**

In many experiments, we experienced a significant complication. Because there were several cases where cells even in the control channel appeared to die during long time courses, we were unable to unequivocally distinguish between virus-induced death and death from physical factors within the chip. Figure 2.9 shows an example where after 12 hours post infection with Sendai virus (~16 HA units/ml for 60min.), although IFNbeta:mCherry was induced, the cells also stained with Sytox Green, indicating breach of the plasma membrane (death).

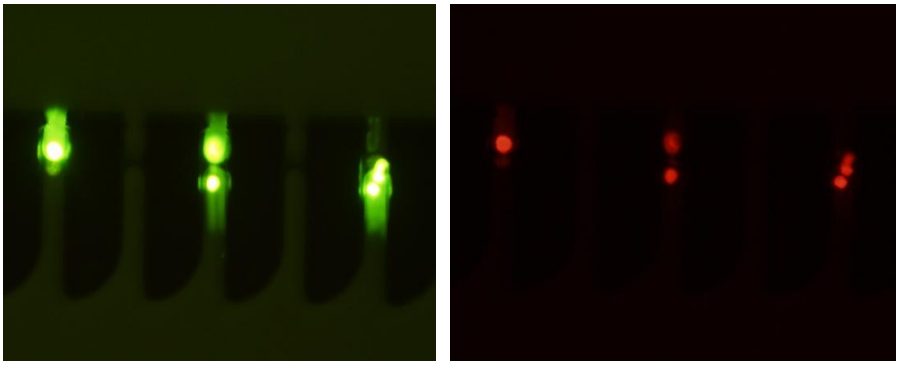


**Figure 2.9 P388D1/IFNb:mCherry cells in traps, 12hrs following SeV infection. Left Sytox Green, right mCherry. Basal levels for both were nearly invisible prior to infection.**

The current iCellator system now integrates automated microscope imaging, user friendly graphical interface of fluidic control and temperature regulation – each of these has improved efficiency of operation. On-chip cell capture has proven to be consistent if particular care is taken processing the cells before loading on-chip, media/stain freshly filtered, and a painstakingly developed proven method is followed in cleaning the iCellator chip. Several experiments have tested the technical assembly to resolve problems with the user interface and fluidic controls, though there have been many delays in the process of upgrading the system, so the most recent version has been tested least.

Initial designs of the SCA/icellator chip fluctuated among several design philosophies. After gaining greater control of the flow with new pressure controllers in the low flow range 0.2 to 2.0 psi, loading cells on to a chip with a greater slant angle and bypass proved to be optimal. Varying the flow in this lower psi range enabled loading of cells into the traps with some predictability (Figure 2.5) previously not achievable without control in that pressure regime. Flow initially was kept at 1.0 psi and gradually decreased to 0.25 psi to allow cells to gently enter traps from beginning to end of each trap array.

Modifying the valve positions from underneath the fluidic reservoirs to above prevented the occasional leaking fluid from damaging the valves. Subsequent upgrades to the user interface and fluidic controls delayed experiments, however the outcome increased ease-of-use and upgrades to the pressure controllers and temperature regulation made cell capture much more reliable.

One of the greatest obstacles to successful experiments was nucleation of air bubbles expanding in tubes and channels of the iCellator system, diminishing flow unpredictably between channels and/or lysing cells in the channel traps. Overcoming this issue was approached from two directions. Initially the process of cleaning the chip did not focus on preventing only removing visible outcroppings of air bubbles by degassing the initial cleaning solution. Subsequent methods involved flushing with degassed isopropanol, water, and media. The key step added was a high pressure flow for extended duration which proved successful. The other approach of ensuring all air bubbles were removed involved heating the reservoir fluid to keep the temperature consistent between chip and reservoir. This proved to be a lengthy undertaking. Nevertheless, the nucleation of bubbles no longer seems to be an issue with the isopropanol flush and heated reservoirs.

We recently transferred the iCellator system to an automated microscope with a higher sensitivity (EMCCD) camera. The combination of the motorized stage and better detector has greatly improved image collection (time resolution) and productivity overall. Hopefully this will improve experiment success rates and also facilitate imaging processes on chip with much faster kinetics.

# 3. FLUORESCENT REPORTER CONSTRUCTS

## 3.1 Overview

In order to visualize cellular responses to virus infection in real time (live cells) within the cell capture and isolation system, we generated a number of fluorescent reporter constructs. These included cytokine promoter activity indicators (TNFalpha, IFNbeta, IL-12; mCherry), transcription factor labels (RelA, IRF3, p105/p50; GFP, YFP, CFP), and non-covalently linked expression indicators (NP:IRES:mCherry or GFP). Of these all were functional to at least some degree except the IL-12 reporter for reasons unknown to us. The p105 constructs were intended to measure (by FRET with GFP:RelA) the relative ratio of transcriptionally active p50:p65 NF-kB heterodimers to inactive/inhibitory p50:p50 homodimers (p105 is proteolytically processed to p50). This parameter is hypothesized to be involved in viral inhibition of cytokine production based on unpublished data from our collaborator Norbert Herzog.

## 3.2 Vector Construction

### *3.2.1 hp105-fluor cloning and transfections*

The $P_{\beta Actin}$-CFP-hp105 construct (Figure 1a) was generated using the $P_{\beta Actin}$-pcDNA3.1 vector as the backbone. The CFP fluor was PCR-amplified from CFP-RelA, courtesy of Dr. Allan Brasier (UTMB), and replaced the region between NheI and AflII of the vector. The IRAT Human clone (Invitrogen) was used as the PCR template for hp105, which replaced the region between AflII and EcoRV in the vector. The $P_{\beta Actin}$-mCherry-hp105 reporter (Figure 3.1) was similarly constructed, using $P_{\beta Actin}$-CFP-hp105 as the parent vector.

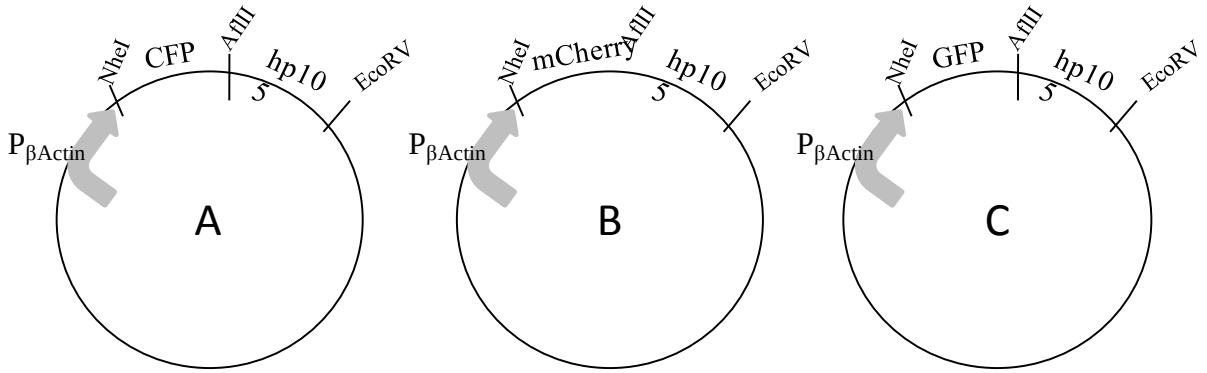



**Figure 3.1 (a) Vector map of PβActin-CFP-hp105-pcDNA3.1. (b) Vector map of PβActin-mCherry-hp105-pcDNA3.1. (c) Vector map of PβActin-GFP-hp105-pcDNA3.1.**

The mCherry fluor was PCR-amplified from the pKs-mCherry vector and replaced the CFP insert between NheI and AflII. The $P_{\beta Actin}$-GFP-hp105 reporter (Figure 1c) was built by replacing mCherry with GFP, which was PCR-amplified from pAcGFP-N1 (Clontech). The

constructs were transformed into One Shot TOP 10 F' *E. coli* competent cells (Invitrogen) and were expressed at 30°C.

The hp105 fusion proteins were successfully expressed in HEK293T cells. The transfection procedure included linearizing the reporter constructs and using the Amaxa transfection system (Lonza). HEK293T cells were grown to 80% confluence prior to transfection and were infected with a 1:1000 dilution of Sendai virus 24 hours post-transfection to initiate nuclear translocation of the fusion protein.

### *3.2.2 TNFalpha, IFNbeta, and IL-12-mCherry Transcriptional Reporters*

The cloned TNFalpha and IFNbeta minimal promoters were kind gifts from Adolfo Garcia-Sastre. Each of these were excised from parental vectors and inserted into pcDNA3.1 (Invitrogen) upstream of the mCherry gene. The CMV promoter was subsequently removed so that only the cytokine promoter remained. mCherry was chosen for its rapid fluor maturation time (~15min). Finally, a PEST sequence was added to mCherry to promote its rapid turnover and allow visualization of both increased and decreased promoter activity.

To construct the IL-12 reporter, an 816bp region of IL-12p40 was PCR-amplified from human genomic DNA. The DNA template was originally purified for use in the RapTOR Grand Challenge and was kindly gifted by Dr. Zachary Bent. Parent vector TNFα-mCherryPEST-pcDNA3.1 was used for the cloning procedure; IL-12p40 replaced TNFα between the NheI and HindIII restriction sites via sequential digestion. Lipofectamine 2000 (Invitrogen) was used to transfect 0, 1, and 2 µg of IL-12p40-mCherry into P388D1 cells. An additional condition was tested in which 2 µg IL-12p40-mCherry were co-transfected with 2 µg control plasmid pMAX-GFP (Lonza). IL-12 was stimulated 24 hours post-transfection with either 1 µg/mL LPS, 1:1000 dilution of Sendai virus, or 100U/mL INF-γ (eBioscience) followed by 1 µg/mL LPS. In all cases, mCherry failed to express while pMAX-GFP expressed well.

## 3.3 Results and Discussion

*hp105 Fusion Proteins*

The cloning process for the hp105-fluor constructs was rife with complications. It is suspected that the hp105 insert is particularly toxic to *E. coli* and is therefore randomly mutated or rearranged during manipulation. In an effort to overcome toxicity to *E. coli*, the constructs were transformed into One Shot TOP 10 F' competent cells, which contains the F' episome to generate single stranded DNA. Additionally, plasmids were amplified in *E. coli* at 30°C, rather than the typical 37°C, in an effort to limit the expression of toxic components. A potential remedy for these and future cloning issues would be to take advantage of a strain of competent cells which has been engineered to clone unstable inserts. Strains MAX Efficiency Stbl2 (Invitrogen) and Copy Cutter EPI400 (Epicentre) have been developed for such purposes.

The hp105-fluor fusion proteins were functionally expressed in HEK293T cells. Figure 3.2 shows the hp105-fluor proteins expressed in the cytoplasm prior to viral infection.

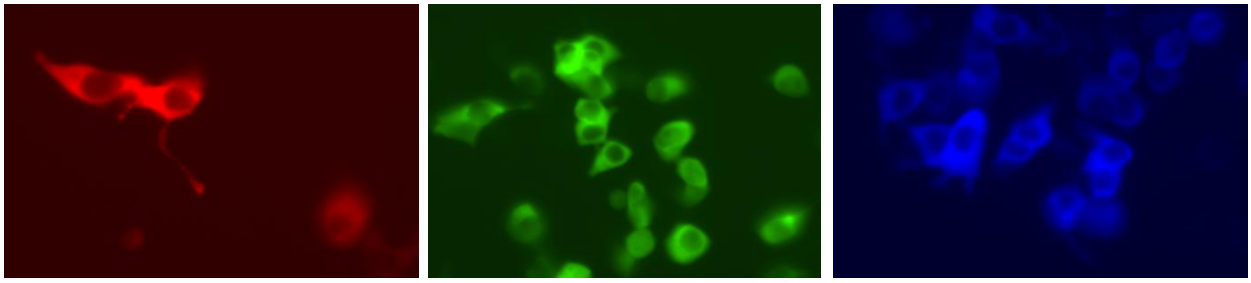


**Figure 3.2 hp105-fluor fusion proteins expressed in HEK293T cells (a) hp105-mCherry. (b) hp105-GFP. (c) hp105-CFP.**

Following 24 hours of infection with Sendai virus, the fluorescent proteins were observed to translocate to the nucleus (Figure 3.3).

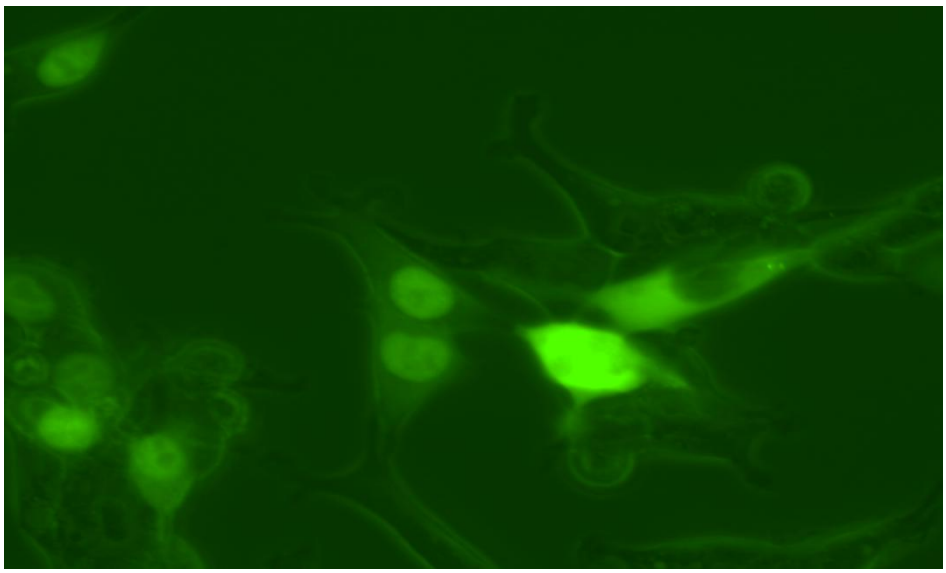


**Figure 3.3 hp105-GFP expressed in HEK293T cells 24 hours post infection with 1:1000 dilution of Sendai virus. The fusion protein is functional and has translocated to the nucleus following viral infection.**

Fusion proteins were never successfully expressed in either RAW 264.7 or P388D.1 macrophages. It is possible the DNA itself was contaminated with residual endotoxin from the plasmid isolation or that traditional Amaxa, Lipofectamine 2000, TransIT (Mirus), jetPEI-Macrophage (Polyplus), or calcium phosphate (Invitrogen) transfections were not optimal for these cells. A different transfection reagent such as FuGENE (Promega) may prove more successful in transfecting macrophages with the hp105-fluor fusions. Endotoxin-free preparations of the plasmids were generated, but were also unsuccessful in transfections.

*TNFalpha and IFNbeta Transcriptional Reporters*

Each of these constructs was initially transfected into RAW264.7 or P388D1 macrophage-like cells using transient methods. However, due to the trauma of transfection, the cells were activated and mCherry was expressed constitutively. Subsequently, stably expressing

lines were cloned based on a low constitutive level of expression then tested for responsiveness to infection or other stimuli. Figure 3.4 shows a diagram of the construct and results of flow cytometric analysis of stably transfected P388D1 cells infected with Sendai virus (positive control virus).

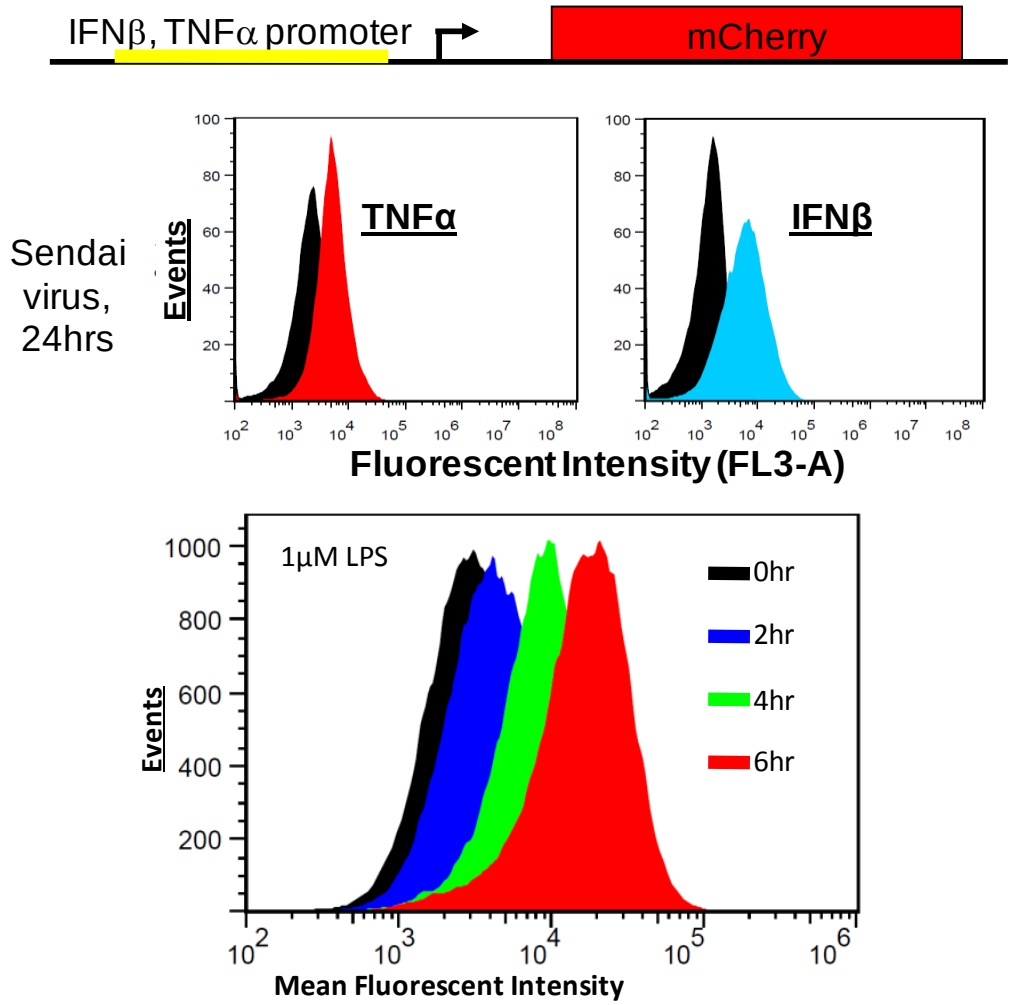



**Figure 3.4 (Top) Diagram of cytokine transcriptional reporters and flow cytometric analysis of stably expressing P388D1 cells. The data show strong induction of both reporters in response to infection with Sendai virus (~100 HA Units/ml). Black curves show uninfected levels. (Bottom) TNFalpha reporter shows time (and dose-not shown) responsiveness to lipopolysaccharide (a TLR4 agonist and potent inducer of TNFalpha in responsive cells).**

*IL-12-mCherry Transcriptional Reporter*

IL-12 is required for the induction of INF-γ production and is composed of two subunits, IL12p35 and IL-12p40, which form active IL-12p70 when co-expressed (1). After construction of the IL-12p40 plasmid, it was determined through literature searches that IL-12p35 is the rate-limiting factor of IL-12P70 formation (2) and therefore a better candidate than IL-12p40 for the transcriptional reporter. Using a construct with IL-12p40, rather than p35 may have been the cause of the lack of expression post-transfection. New PCR primers have been designed to

amplify two different lengths of sequence upstream of IL-12p35: (-1585→+66bp) and (-648→+66bp). The longer sequence encompasses the entire length of the promoter while the shorter sequence corresponds to relatively higher promoter activity in a previous study (3). Cloning of these constructs has progressed to the point where the PCR product replaced TNFα in parent vector TNFα-mCherryPEST-pcDNA3.1. There is a short sequence (approximately 50bp) downstream of IL-12p35 and upstream of mCherryPEST that may have an adverse effect on the functionality of these reporters. Since this sequence is not flanked by restriction sites, it must be excised via overlapping PCR. However, since the construction of these reporters has not been completed, it may be far more efficient to simply use a different parent vector that does not contain this problematic sequence upstream of the mCherryPEST region. Additionally, there is the possibility of that, as with the hp105 reporters, transfection will not be possible in macrophages. To determine if this is true, the IL-12 reporters should be tested in HEK293T cells. If expression works well in a non-macrophage cell line, the DNA should possibly be isolated with an endotoxin-free kit and linearized prior to transfection.

## 3.4 References


1. K. Gee, C. Guzzo, N.F.C. Mat, W. Ma, and A. Kumar, *Inflammation & Allergy*, 2009, **8**, 40-52.
2. A. Snijders, C.M.U. Hilkens, T.C.T.M van der Pouw Kraan, M. Engel, L.A. Aarden, and M.L. Kapsenberg, *J Immunology*, 1996, **156**, 1207-1212.
3. S. Goriely, D. Demonté, S. Nizet, D. De Wit, F. Willems, M. Goldman, and C. Van Lint, *Blood*, 2003, **101**, 4894-4902.

# 4. ON- AND OFF-CHIP CHARACTERIZATION OF CYTOKINE REPORTER CELL LINES

## 4.1 TNFalpha Reporter

To generate stable cell lines that display fluorescent fusions of signaling proteins as well as fluorescent fusions of cytokines for direct imaging and flow cytometric analysis, the RG16 RelA-GFP RAW macrophage cells used in the MISL project were used to generate mCherry-TNFa stable cell lines. These mouse macrophage cells contain both GFP-labeled RelA/p65 and mCherry –labeled TNFa cytokine. Four clones of the stably transfected RG16-mCherry TNFa cells were stimulated with LPS – TLR4 agonist or Poly I:C – TLR3 agonist to characterize their responses.

### *4.1.1 Materials and methods*

Cells were seeded at 2 x $10^6$ /well in 6 well ultra low attachment plates (Corning Cat #3471) in RAW growth media. E. Coli smooth LPS (O55:B5) (Sigma) and Poly I:C (Invivogen) were resuspended in DI water. LPS was added at the final concentration of 100nM and PolyI:C was added at 10 µg/ml. The cells were fixed 12 hours after stimulation and the mCherry fluorescence was detected using BD FACScan.

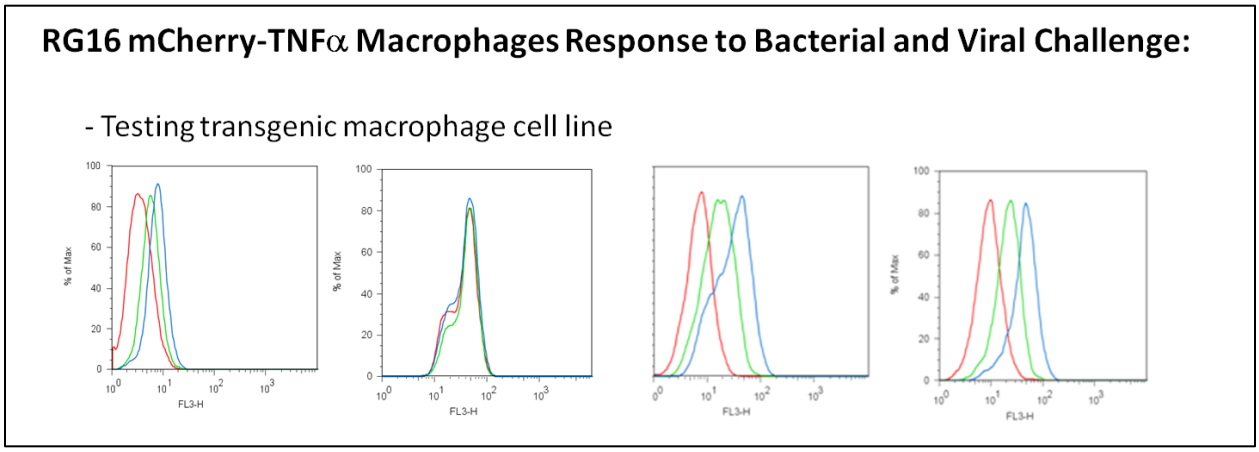



**Figure 4.1 Characterization of stable cell clones by monitoring amount of mCherry-TNFa fluorescence using flow cytometry**

### *4.1.2 Results*

Of the four clones tested, the one labeled "clone 5" shown as the furthest overlay on the right in figure 4.1, showed most intense mCherry fluorescence. Future experiments using RG16-mCherry-TNFa cells should be done on Clone 5.

### *4.1.3 TLR signal degradation studies:*

To check for the ability of macrophages to recover from TLR activation, a series of withdrawal studies were conducted using LPS and polyI:C on Clone 5. Cells were seeded at

0.75 x $10^6$ / well into 12 well cell culture treated plates. 1μM LPS or 10 μg/ml polyI:C were added to the designated wells for 24 hours, or 4 hours followed by 3 sequential washes using media. The washed wells were harvested at 18 hours or 48 hours post wash to check for degradation of TLR induced TNFa sign

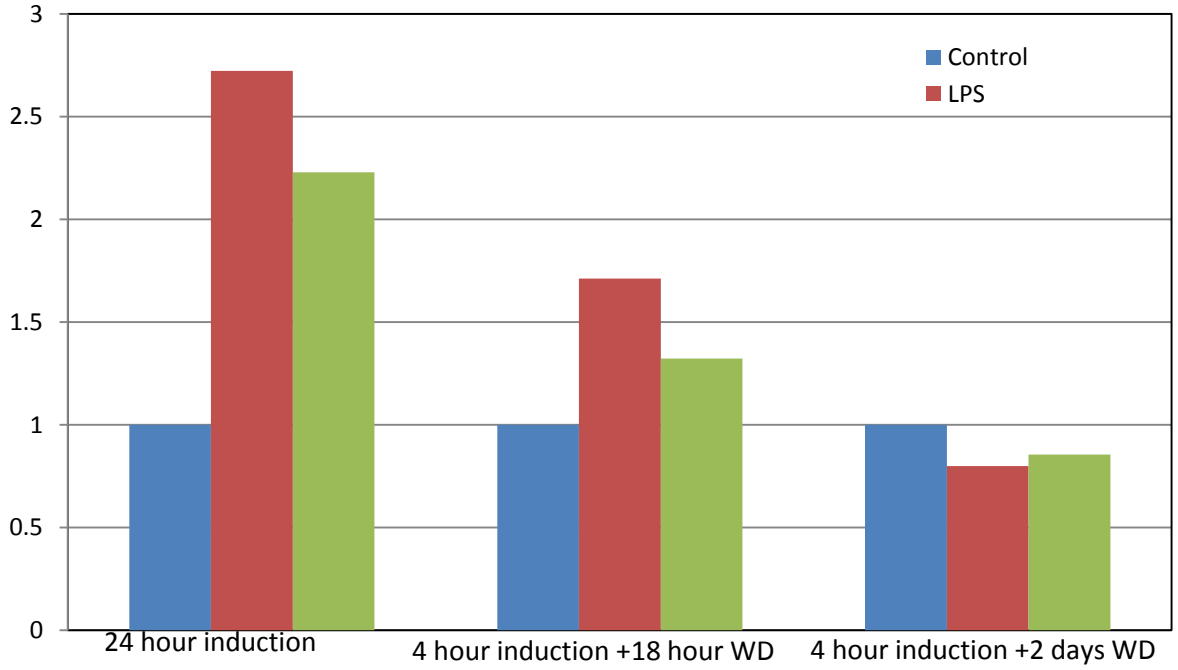



**Figure 4.2 TLR signal induction and degradation.**

LPS induced more TNFa production than poly I:C . Approximately 50% of the TNFa induction disappeared after 18 hours without agonist, and all TNFa production ceased after 2 days without TLR activation.

## 4.2 IFNbeta Reporter

To characterize the difference between Sendai and Pichende virus' abilities to induce host interferon beta production. Sendai virus is the positive control, and should induce robust interferon beta production in host macrophages. Pichende p2 and p18 were indicated to have different interferon induction profiles.

P388D.1 mouse macrophages were used for viral infection and cytokine measurement. Adherent P388D.1 cells were obtained from Norbetr Herzog (UTMB). Cells were seeded at 2 x $10^6$ / well into 6 well plates. Cells were serum starved overnight prior to infection with virus. 1:600 dilution of Sendai virus, 1:10, 1:100 dilution of Pichemde p2 and p18 were applied to the cells for 1 hour, followed by wash and reapplication of media containing serum. The supernatant were collected at 24 hours, as assayed for released IFNb using an ELISA assay for INFb (RnD systems 42400-1).

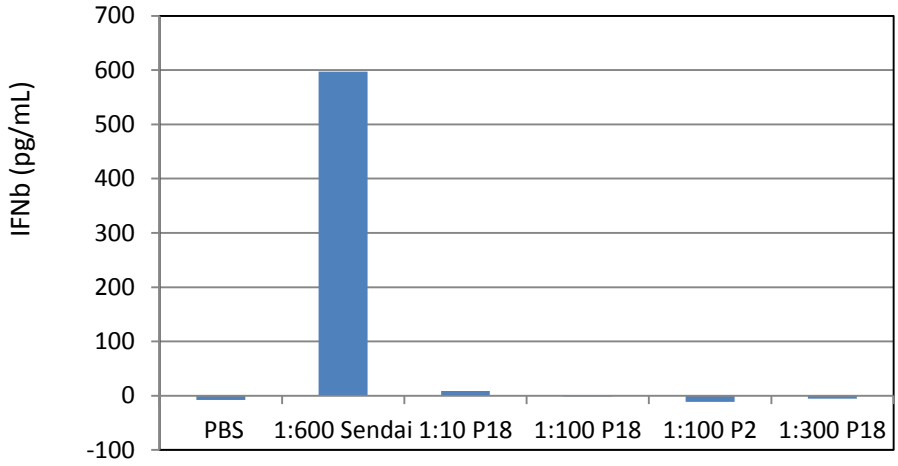



**Figure 4.3 Sendai virus versus Pichende P2, P18 IFNb production**

Robust Sendai virus induced IFNb was detected 24 hours after infection. Neither Pichende p2 nor Pichende p18 induced any IFNb production in P388D.1 cells. Pichende virus was not used for the rest of the project.

## 3.3 MICA Platform for On-chip Viral Infection and Cytokine Detection

### *4.3.1 Background*

The microimmune cell assay platform (MICA) has been used to monitor bacterial lipopolysaccharide induced macrophage activation using orthogonal fluorescent microscopy and on-chip micro flow cytometry (Srivastava, Brennan et al. 2009). The next logical step is to adapt the MICA technology for measuring host immune responses to viral infections. The advantage of using the MICA technology for viral infection is two-fold. First, the miniaturized microfluidic platform can be used to probe rare cell types such as primary cells or clinical isolates, and second, once assembled, the MICA system is completely contained, with no hands—on handling of the cells or virus. The improvements to be made to the MICA system include the capability of longer term cell culture on the chip, and the development of viral reponse pathway specific microfluidic assays.

### *4.3.2 Overnight Cell Culture on-chip*

Due to the fact that viral infections elicit slower host cell responses when compared to bacterial LPS elicited host responses, the need for longer term maintenance of live macrophages on the chip is necessary for viral infection studies. To achieve this, 5% pre-mixed $CO_2$ was plumbed into the pressure controller for generating the positive pressure used to move fluids on the chip. The pre-mixed $CO_2$ in combination with the temperature controller on the MICA platform creates similar culturing conditions as a standard mammalian cell incubator. Cell viability and morphology were examined using both microscopy and flow cytometry of p388D.1 cells grown in the MICA chip overnight (figure 4.4). The presence of daughter cell after overnight culture on the chip shows that the cells are viable. The viability was further quantified

using a fixable dead cell stain that labeled dead cells. The flow cytometric analysis showed that over 80% of the cells grown on the chip were viable by the next day.

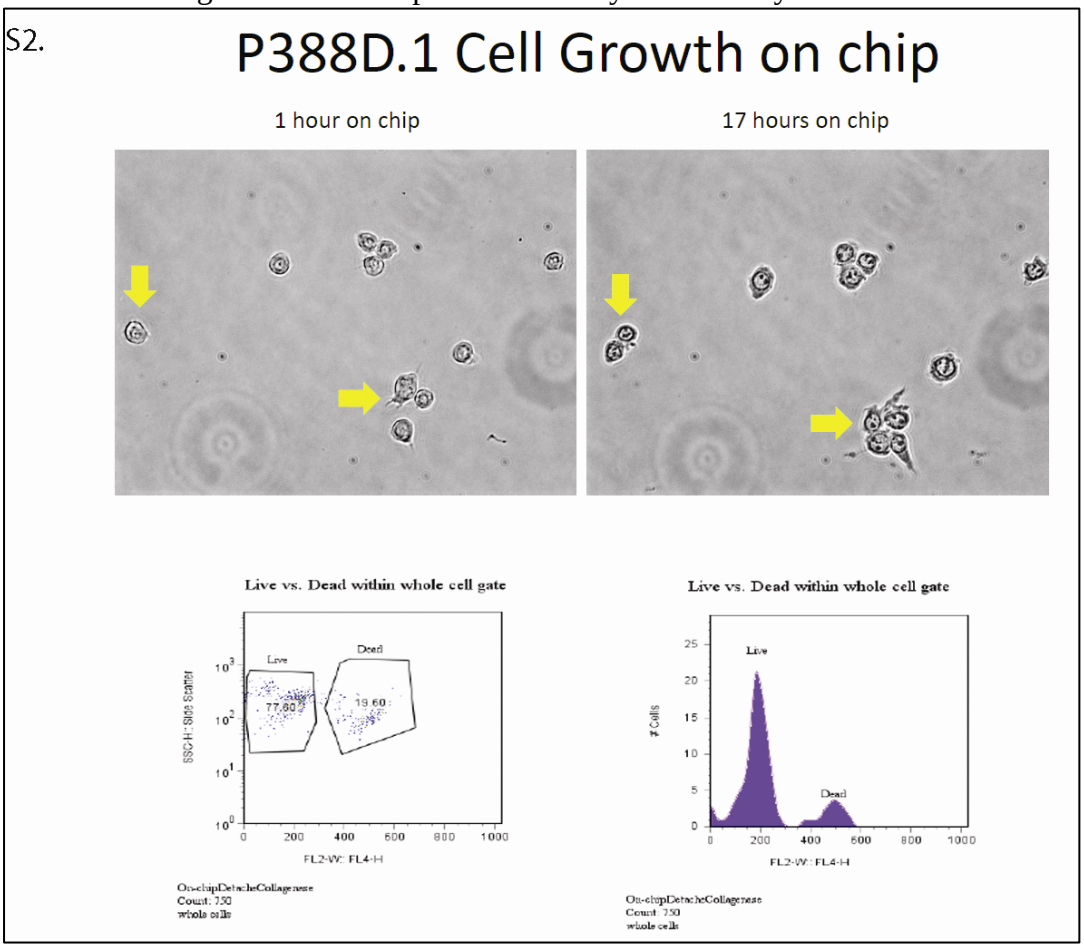



**Figure 4.4 Longer term cell culturing using microfluidic chip. Bright field microscopy reveals the presence of daughter cells (yellow arrows) after overnight culture on the MICA chip. Live/dead stain followed by flow cytometric analysis indicate that nearly all cells survived**

### *4.3.3 Monitoring of viral infections and host response using the MICA platform*

To perform viral infection studies using the MICA platform, p388D.1 mouse macrophage cells were seeded onto the chip at 1 x $10^7$ cells/ml density, and allowed to adhere for 10 minutes. To infect, Sendai virus stock was diluted 1:600 using pre-warmed RPMI media, and flown into the cell holding chambers. The flow was stopped once complete media exchange was achieved in the cell holding chambers, and the virus was incubated with the cells for 1 hour at 37°C. Following the virus incubation, the cell holding chambers were washed with pre-warmed PBS, for 5 minutes, and growth media containing 5% FBS was flown into the cell holding chambers. The flow was decreased to 0.3psi overnight and bright field images were taken the next day. In the control chambers were cells were only exposed to RPMI and washed with PBS, there was noticeable increase in number of cells 20 hours following the experiment. The morphology of the cells in the uninfected chambers were also round with no visible vacuoles or broken

membranes under 60X maginification. In contrast the Sendai infected cells showed elongated morphology, indicative of macrophage activation, as well as many dead cells with broken plasma membranes and hollow cell bodies, indicating necrosis.

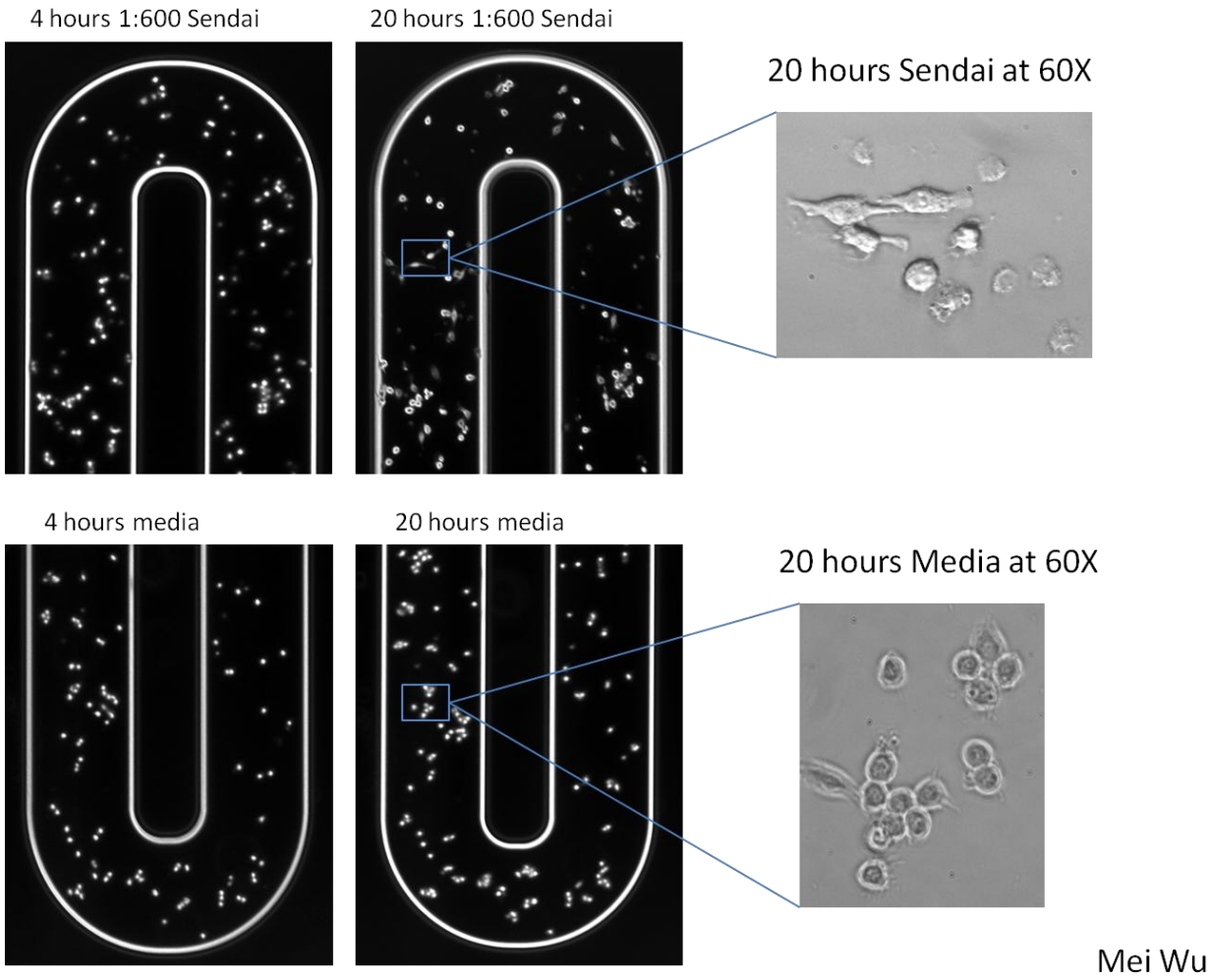



**Figure 4.5 On-chip Sendai virus infection and cytokine measurement**

### *4.3.4 On-chip cytokine detection*

To investigate the host immune response to Sendai infection, Brefeldin A, a Golgi release inhibitor was applied to grow media and flown into the chip for 4 hours, followed by fixation of cells using 1% paraformaldehyde solution. The fixed cells were permeabilized using 0.1% Triton X-100 in PBS, and washed for 5 minutes using PBS. PE labeled monoclonal antibody against TNFa (USBiological) at a dilution of 1:100 was used to measure the Sendai virus induced TNFa production. The stained cells were analyzed using flow cytometry, and data was plotted as histograms using Excel (figure 4.6).

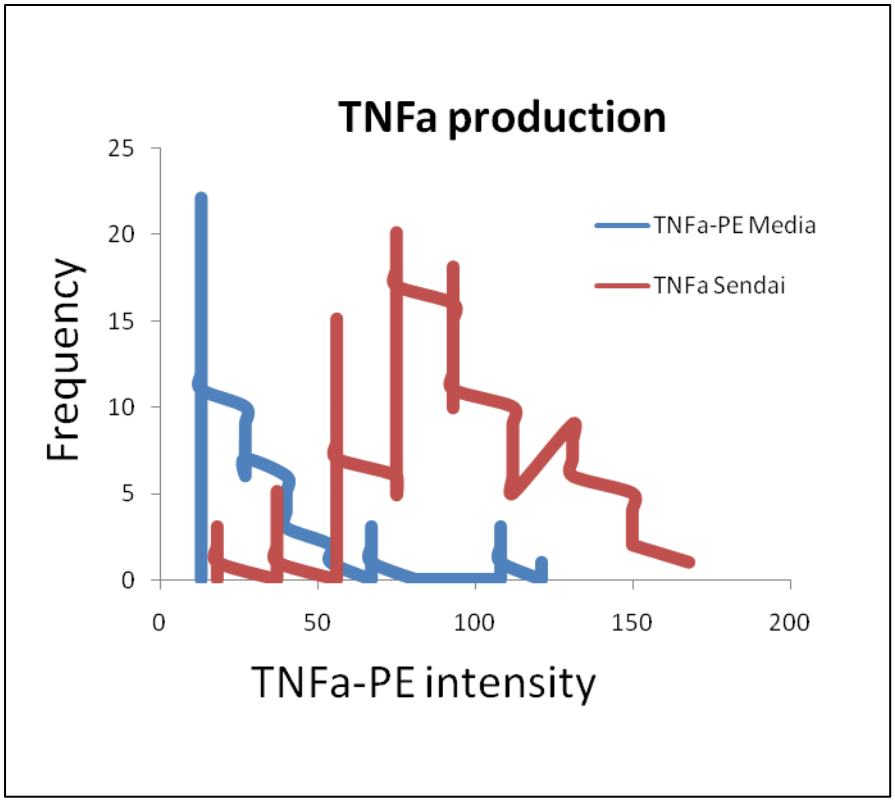



**Figure 4.6 On-chip measurement of Sendai virus induced TNFa production (intracellular cytokine staining). 20 hours after contact with Sendai virus, the P388D.1 macrophages produced significant amounts of TNFa cytokine, indicating the activation of TLR pathway.**

The same cells were also assays for the production of another cytokine, IFNb, a virus-specific cytokine. The amount of IFNb detected was much lowere compared to that of TNFa, but there is definitely detectable amounts of IFNb at 20 hours post infection. The production of IFNb indicates the activation of a slower response that is specifically mediated through interferon response factor 3 (IRF3). The detection of these two cytokines TNFa and IFNb can be multiplexed using different fluorephores that are conjugated to the monoclonal antibodies.

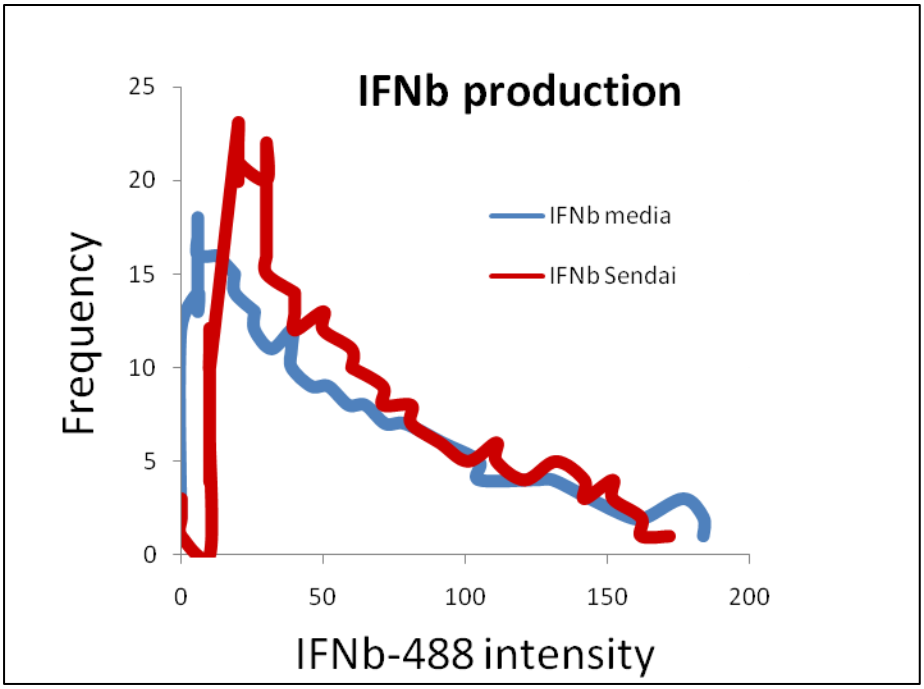



**Figure 4.7 On-chip measurement of Sendai virus induced IFNb production.**

In conclusion, The MICA platform was successfully adapted to allow longer term cell culture necessary for viral infection studies. The on-chip detection of virus induced cytokine was developed successfully, and can be multiplexed using multiple fluorephores.

### *4.3.5 Improvements to the MICA platform*

One of the objectives of this project is to improve the multiplexing capability of the MICA platform. The maximum allowable number of parameters using our existing microscope and micro flow cytometry setup is 2 fluorescence channels for microscopy and 3-color flow cytometry from the same sample. The chips designed during the MISL project only allowed two conditions to be monitored in one chip experiment, and that is not enough throughput for any realistic biological study.

New chip designs

As the microfluidic assays became more sophisticated, it became necessary to upgrade the throughput of the MICA chip to be compatible with biologically relevant problems. A series of new chip designs were generated, and tested. First, a four-chamber chip was tested and used to generate data for platform publication. Then, newer chip designs with even higher throughput was tested. The 6 chamber design proved very useful in the subsequent florescent *in situ* hydridization experiments. The 14 chamber chip, however, was a disappointment due to the fact that there were too many ports for the existing pressure control unit. All chips were designed and drafted using AutoCAD, and the photomasks were generated by Photosciences, and the chips were fabricated by Caliper.

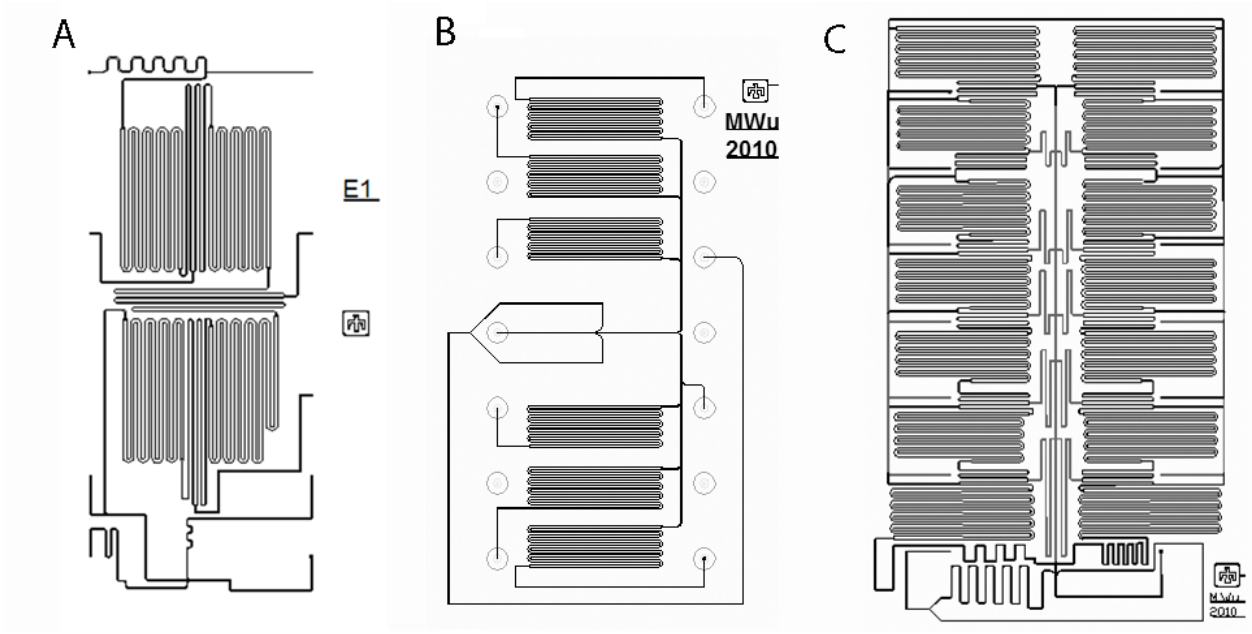


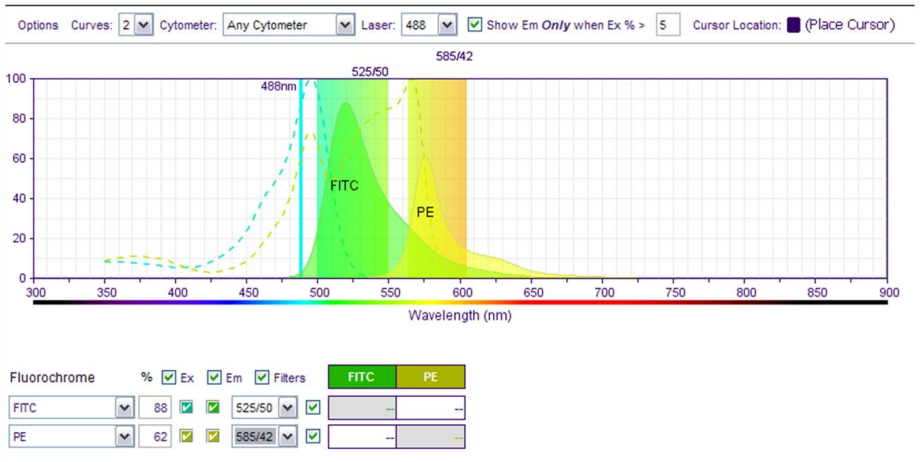

**Figure 4.8 4,6, 14 chamber chip designs.**

The MICA setup already has GFP and R-PE filters installed for simultaneous visual inspection of green and yellow fluorescence. The IRF3-GFP translocation studies can be multiplexed with a dead cell stain that emits fluorescence at 585 nm.

Fluorechromes and Filters For Microscopy:

Flourechrome: IRF3-GFP
Emission: 520 nm

Microscope Filter: GFP(Chroma)
Emission: 525/50 nm

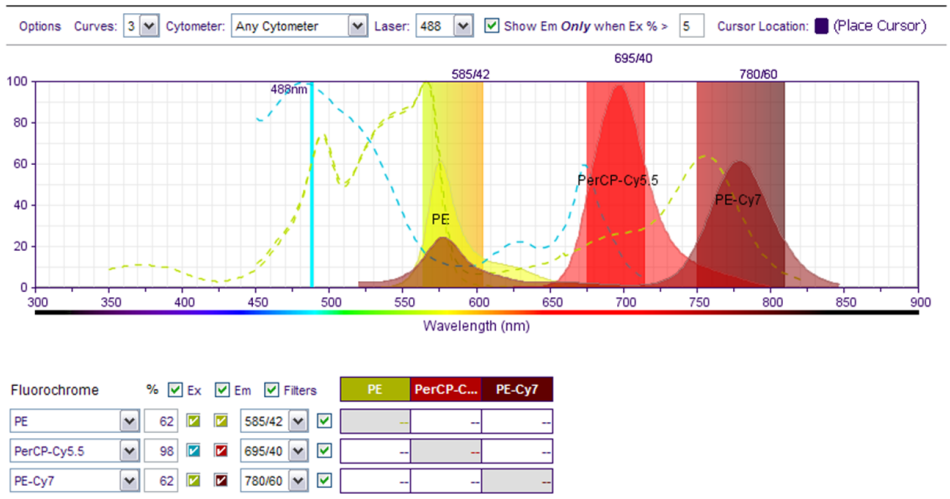



Fluorephore: Live/Dead Fixable Red Stain
Emission max: 585 nm

Microscope Filter: R&B Phycoerythrin (Chroma)
Emission: 585/40 nm

**Figure 4.9 Microscope filters that will be used to monitor IRF3-GFP translocation and cell viability.**

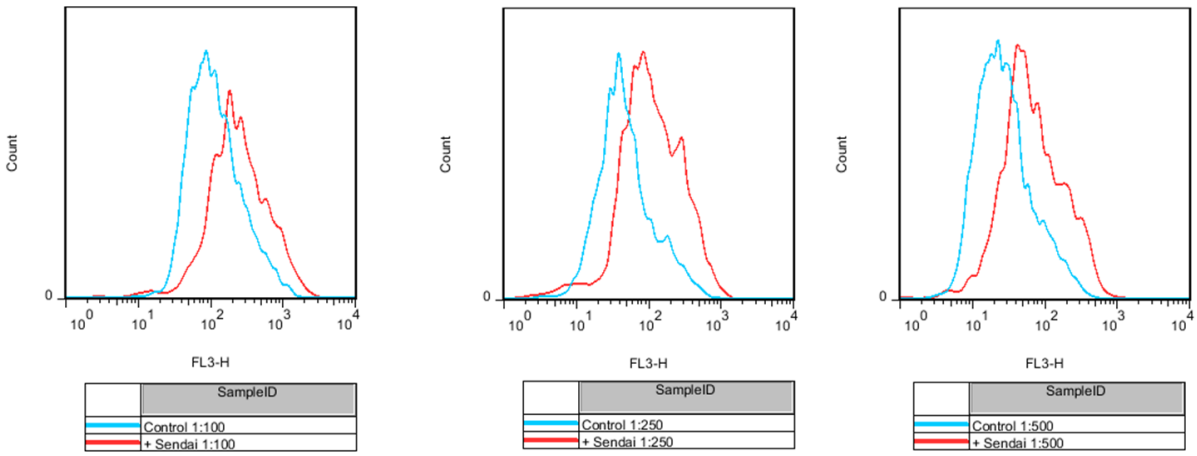
Cytokines and fluorephores for μFC




Live/dead fixable stain at 585 nm
TNFα-PerCP-Cy5.5 (1:25)
IFNβ-PE-Cy7 (conjugate myself, optimize concentration)

**Figure 4.10 The fluorephores that can be multiplexed in multi-color on-chip micro flow cytometry.**

Optimization of cytokine antibody conjugations for multiplexed on-chip measurement: In order to take advantage of the multiplex capability of the MICA platform, antibody against IFNb needs to be conjugated to PE-Cy7, and optimized for intracellular staining assay. The Lightning-link antibody conjugation kit was purchased from Innovabiosciences, and used to labeled mouse-monoclonal interferon beta antibody (USBiological). Following the conjugation, the labeled antibody was tested at various concentrations for optimal binding conditions for maximal signal detection (figure 4.10). Results in figure 4.10 shows that the best condition for IFNb detection is blocking with BD stain buffer for 30min after permeabilization and fixation, followed by incubation with 1:250 dilution of in-house labeled IFNb-PE-Cy7.




Conjugation successful, optimal staining condition:
block with BD stain buffer 30min RT
1:250 IFNb-PE-Cy7 (0.04ug/ml) for 30min RT

**Figure 4.11 IFNbeta-PE-Cy7 staining optimization**

### *4.3.6 Locked nucleic acid based mRNA detection using Flow-FISH on a Chip*

The detection of viral infection is vital for combating potential biowarfare agents and newly emergent diseases. The most popular detection methods are viral antibody based, quantitative RT-PCR, and microarray based detections (Cho, Hwang et al. 1996; Steininger, Kundi et al. 2002; Wang, Coscoy et al. 2002). Antibodies are very specific for viral proteins, but they must be customized for every target. Production of specific antibodies to newly emergent viruses can take anywhere from months to years to generate, and even then, some viruses can mutate faster than researchers can generate specific antibodies. The PCR and microarray based methods are highly sensitive, but have their own shortcomings as well. The sample preparation for PCR or microarray detection requires the lysis of the host cells, and therefore information regarding the viability and responses of the host cell is lost. In addition, the lysis and pooling of populations of infected cells produces an averaged measurement, obliterating the cell-to-cell variation in infection level and cellular response information. A more complete snapshot of host/virus interaction can be captured by using the cytometric-fluorescence *in situ* hybridization method, or Flow-FISH. Flow-FISH was first used for telomere length determination, where telomeric DNA was stained using fluorescently labeled oligonucleotide probes, and the fluorescence emitted was used to determine the amount of telomeric DNA and telomeric length(Baerlocher, Mak et al. 2002). FISH allows for independent observation of single cells that contain viral RNA of interest, and the multiplexing with host nucleic acid and protein analysis to provide a picture of the host/virus interaction at the single cell level. The combination of FISH with flow cytometry provides the throughput to analyze tens of thousands of cells per second, and the ability to capture rare infected cells in a large heterogeneous cell population whose signals would be lost if an average measurement was taken. The downside to traditional Flow-FISH is that the method is labor and time-intensive. Hybridization and washing

steps frequently stretch into multiple days. The availability of locked nucleic acid (LNA) RNA probes, in which several nucleotides in the RNA probe have methylene bridges between the 2'-oxygens and 4'-carbons of the ribose sugars, bind to DNA and RNA at higher affinity and specificity than any other nucleotide analog(Thomsen, Nielsen et al. 2005). Recently, Robertson et. al (Robertson, Verhoeven et al. 2010) demonstrated that LNA RNA probes can be used to directly detect viral RNA in infected cells by means of cytometric-fluorescent *in-situ* hybridization (Flow-FISH) much faster and with comparable sensitivity to RT-PCR. While this is a very promising development, LNA reagents are highly proprietary and very expensive. Large scale LNA detection based studies can easily run into thousands of dollars per experiment when done in traditional coverslip format. The miniaturization of microfluidic chip assays in combination with fine temperature control on the MICA platform provides the perfect solution. The MICA platform can reduce reagent consumption by 95% compared to benchtop methods, and the new chip designs allows for experiments with up to 6, 8, even 10 conditions using less LNA than required for even 1 LNA Flow-FISH sample.

*LNA Flow-FISH on a chip method:* Detection of β-actin mRNA using flow-FISH

As a positive control, LNA-β actin-Biotin probes were used to establish a chip-based LNA flow-FISH protocol. Non-adherent RAMOS cells were used for the pilots as the ultimate goal will be to use Flow-FISH to screen blood from potentially infected patients. Flow-FISH detection of β-actin was successfully developed and replicated, shown in figure 4.11. Future studies will use LNA Flow-FISH to detect viral as well as host mRNAs of response genes.

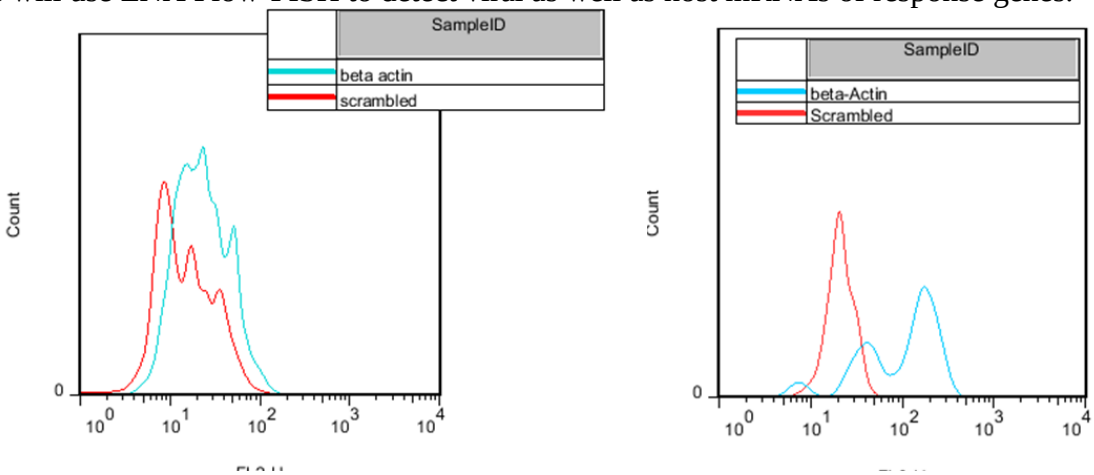



**Figure 4.12 Flow-FISH detection of β-actin by using biotin-labeled LNA probe and PE-labeled strepavidin. A biotinylated scrambled LNA probe was used as negative control. The experiment was replicated.**

*Detection of microRNA 155 using Flow-FISH*

It is known that viruses can induce production of host micro RNAs as a defense mechanism. LNA probes are the only available method for detecting small RNAs. miRNA 155 was chosen as the target for method development because it was shown to be upregulated in RAMOS cells upon CpG stimulation. Several cell lines were screened for various miRNAs that are immune related, and miR155 in Ramos cells were the most highly expressed upon CpG stimulation.

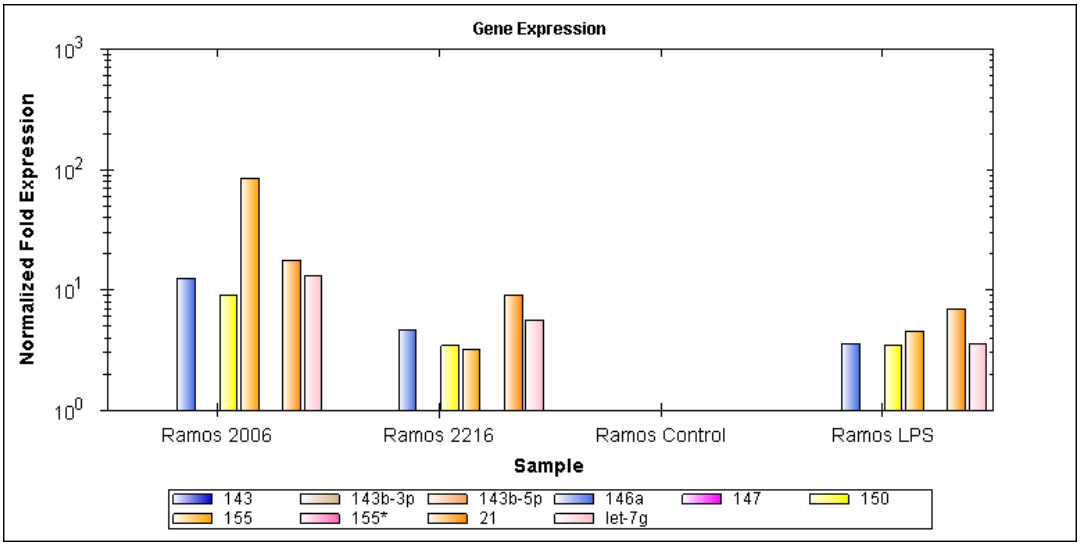
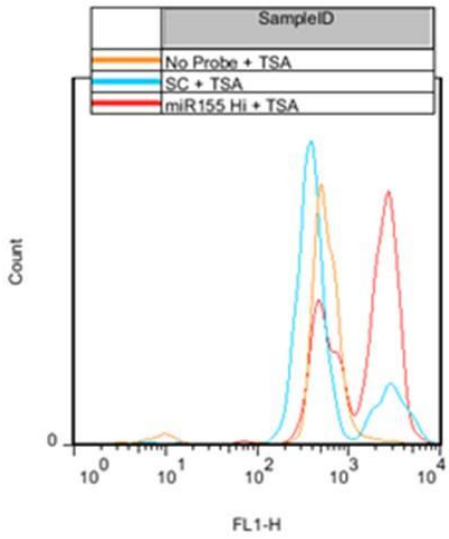



**Figure 4.13 Quantitative RT-PCR analysis of micro RNA expression in Ramos cell lines under stimulation of CpG 2006, CpG 2216, and LPS. miR155 was most highly induced when Ramos cells were stimulated with CpG 2006.**

Great effort was expended into LNA Flow-FISH detection of microRNA species. Double DIG labeled LNA probes specific for miRNA 155 was used to establish miRNA flow-FISH. After numerous failed attempts at detecting miRNA155 using fluorescently labeled anti-DIG antibody, tyramide signal amplification (TSA) was in attempt to amplify the miRNA signal for detection. miRNA155 did appear to be detectable upon stimulation of RAMOS with hypermethylated CpG.



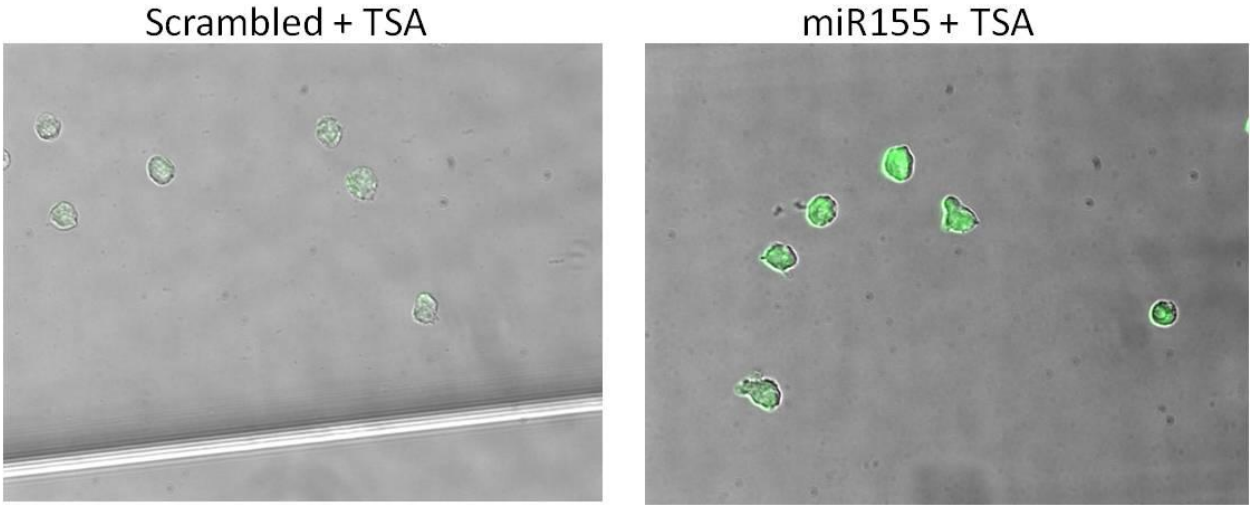




**Figure 4.14 miR155 detection using both flow cytometry and microscopy. SC – scrambled probe for negative control.**

## 4.4 References


Baerlocher, G. M., J. Mak, et al. (2002). "Telomere length measurement by fluorescence in situ hybridization and flow cytometry: tips and pitfalls." Cytometry **47**(2): 89-99.

Cho, S. W., S. G. Hwang, et al. (1996). "In situ detection of hepatitis C virus RNA in liver tissue using a digoxigenin-labeled probe created during a polymerase chain reaction." J Med Virol **48**(3): 227-233.

Robertson, K. L., A. B. Verhoeven, et al. (2010). "Monitoring viral RNA in infected cells with LNA flow-FISH." RNA **16**(8): 1679-1685.

Srivastava, N., J. S. Brennan, et al. (2009). "Fully integrated microfluidic platform enabling automated phosphoprofiling of macrophage response." Anal Chem **81**(9): 3261-3269.

Steininger, C., M. Kundi, et al. (2002). "Effectiveness of reverse transcription-PCR, virus isolation, and enzyme-linked immunosorbent assay for diagnosis of influenza A virus infection in different age groups." J Clin Microbiol **40**(6): 2051-2056.

Thomsen, R., P. S. Nielsen, et al. (2005). "Dramatically improved RNA in situ hybridization signals using LNA-modified probes." RNA **11**(11): 1745-1748.

Wang, D., L. Coscoy, et al. (2002). "Microarray-based detection and genotyping of viral pathogens." Proc Natl Acad Sci U S A **99**(24): 15687-15692.

# 5. IDENTIFICATION OF A CRITICAL DOMAIN WITHIN NUCLEOPROTEIN OF TACARIBE VIRUS IMPORTANT FOR ANTI-INTERFERON ACTIVITY

(Oscar Negrete, Carol Kozina, Dianna Maar, Brooke Harmon, Catherine Branda, Bryan Carson)

## 5.1 Overview

Pathogeneic arenaviruses can cause lethal hemorrhagic fever in humans and are categorized as biothreat agents. Investigations into the interaction between arenaviruses and the host innate immune system may identify key determinants of their pathogenicity. The arenavirus nucleoprotein (NP) contains the ability to suppress the induction of type I interferon (IFN). This anti-IFN activity is shared by all arenaviruses with the exception of Tacaribe virus (TACV). To identify the TACV NP amino acid residues that prevent its IFN-countering ability, we created a series of NP chimeras between residues of TACV NP and Pichinde virus (PICV) NP, an arenavirus NP with potent anti-IFN function. Reference sequences of PICV NP and TACV NP were codon optimized and chimeric NPs were generated by overlapping PCR. NP chimeras were expressed in the presence of an IFNβ reporter that was activated through Sendai virus infection. Chimeric NP analysis revealed that a minimal four amino acid stretch derived from PICV NP could impart efficient anti-IFN activity in TACV NP. Strikingly, the TACV NP gene cloned and sequenced from a viral strain obtained from BEI resources deviated from the reference sequence at this particular four amino acid region. Substitution of these four viral derived residues, GPPT, into the codon optimized reference sequence of TACV NP rescued IFN antagonism. Since these results suggested that the NP in BEI resources TACV strain contains the ability suppress the induction of type I interferon, the IFNβ response was further characterized during TACV infection. We determined that TACV did not stimulate a potent IFNβ response early in infection in multiple cells types and that TACV infection suppressed IFNβ activation stimulated by a secondary infection. Similar infection results were obtained with PICV. Collectively, these data indicate that certain strains of TACV contain the important NP domains necessary for anti-IFN activity observed during infection.

## 5.2 Introduction

The *Arenaviridae* family contains more than 25 virus species, which are currently subdivided into two major groups: Old World (OW) and New World (NW) arenaviruses. The OW arenaviruses constitute a single lineage, while the NW arenaviruses are divided into three phylogenetically distinct clades designated A, B and C (1). OW and NW clade B arenaviruses contain several significant human pathogens capable of causing hemorrhagic fever (HF), and as such represent a significant threat to human health worldwide. Nonpathogenic arenaviruses such as Tacaribe virus (TACV) also merit significant attention since close phylogenetic relationships with family members causing HF provide an excellent opportunity to identify determinants of pathogenesis. In addition, Pichinde virus (PICV) is a nonpathogenic arenavirus that acquires virulence in guinea pigs upon long term passaging in the animals that serves as a model for pathogenic human Lassa fever virus (LASV) infection.

Arenaviruses cause chronic infections of rodents and each arenavirus species is associated mainly with a particular rodent host species, except for the NW TACV, which has been isolated only from fruit-eating bats in Trinidad. Asymptomatically infected rodents move freely in their natural habitat and may invade human dwellings. Humans are most frequently infected through contact with infected rodent excreta, commonly via inhalation of dust or aerosolized virus-containing materials, or ingestion of contaminated foods. While human arenavirus infection, such as LASV can occur at frequencies of 100,000–500,000 infections per year, most infections are typically sub-clinical or associated with mild febrile illness. However, some South American arenavirus outbreaks have been reported to cause severe HF in humans with fatality rates of 15-35%.

Arenaviruses are enveloped viruses with a bisegmented RNA genome and a unique ambisense coding strategy. Despite this coding strategy, the *Arenaviridae* are classified as segmented single-strand, negative sense RNA viruses. The large (L) segment encodes two proteins, the viral RNA-dependent RNA polymerase (L protein) and the Z protein, a matrix protein important for virus budding. On the small (S) segment, the glycoprotein precursor (GPC) and the nucleoprotein (NP) are encoded. The NP is the most abundant viral protein in both infected cells and the virus particles. The NP binds to both the genomic and antigenomic RNA and together with the L polymerase, forms the ribonucleoprotein core necessary for virus transcription and replication. Several groups have now demonstrated that the NP is multifunctional and has the ability to inhibit the host innate immune defenses, specifically the type I interferon response (3, 8).

The most important cytokines produced by the host to defend itself against viral infection are the type I interferons (IFN-$\alpha$ and $-\beta$, now referred to as IFN) (7). Innate immune cells such as dendritic cells and macrophages play an important role in IFN production. IFN produces an antiviral state in cells by inducing interferon stimulated genes. IFN activation occurs through the recognition of pathogen associated molecular patterns or PAMPs and on viruses they are typically nucleic acids. On the host side, pathogen recognition receptors (PRR) sense foreign nucleic acids and signal through transcription factors to produce IFN. Some PRRs are expressed on endosomal compartments as is the case for TLR receptors while others such as RIG I and MDA5 are expressed in the cytoplasm. It is believed that for arenaviruses, like many other negative strand RNA viruses that have 5' triphosphate (5'PPP) motifs on the ends of their genomic RNA (6), are mainly detected by the RIG I sensor to signal a cascade that involves the activation and nuclear translocation of the transcription factor IRF3 which in turn, induces expression of the IFN$\beta$ gene. The NP is thought to inhibit IFN production by blocking activation and translocation of IRF3, though the mechanism of inhibition remains to be fully characterized (4). The OW LASV NP was recently crystallized and it reveals that the C terminal domain contains 3'-5' exonuclease activity (2, 5). Although a discrepancy in the NP exonuclease ligand specificity exists, the general hypothesis is that the NP prevents the virus-induced IFN induction by degrading the viral PAMP RNA ligands that otherwise would trigger the viral sensors in the cells.

All the arenaviruses NP are believed to contain anti-IFN activity with the exception of TACV NP (3). It was also demonstrated that TACV infected cells fail to inhibit interferon production, a result consistent with the lack of IFN antagonism function in the NP. Understanding why the TACV NP protein lacks efficient anti-IFN activity may lead to insights into the potential pathogenic function of other areanvirus NP proteins. To address this problem, we generated NP chimeras to map the regions in TACV that prevent its IFN-countering ability.

This method of analysis uses a gain-of-function assay that should preserve the structural integrity of TACV NP, to narrow down regions important for IFN-countering function. Through NP chimeric analysis, we discovered that a four amino acid region in the reference sequence of TACV NP is preventing anti-IFN activity. However, the TACV virus obtained from BEI resources had the GPPT amino acid sequence found in other arenavirus NPs. The TACV strain with the GPPT sequence in its NP did not stimulate a potent IFNβ response early in infection in multiple cells types and TACV infection was capable of suppressing IFNβ activation induced in a co-infection assay. Based on these results, we believe the NP in the viral strain of TACV has immunosuppression function characteristic to all arenaviruses.

## 3.4 Results

### *5.3.1 Chimeric nucleoprotein analysis identifies minimal residues important for anti-interferon activity in Tacaribe NP*

In a study reported by de la Torre and colleagues (3), all arenavirus nucleoproteins, with the exception of Tacaribe virus, were found to inhibit type I interferon activation by blocking IRF-3 translocation. Since the mechanisms by which arenaviruses subvert the host's innate immune defenses may be determinants of their pathogenicity, we decided to map the regions and specific amino acid residues within NP involved in its anti-IFN activity. To do this, the NP gene Genbank sequences from PICV and TACV were codon optimized and synthesized for expression in mammalian cells. Of all other arenavirus NPs, the PICV NP is the most potent inhibitor of type I interferon and therefore chimeric NP proteins were designed to insert PICV NP residues into homologous positions in the TACV NP with the intent of gaining anti-IFN function within the TACV NP. The exact residues that were swapped between the PICV and TACV NPs are illustrated in Figure 5.1A. To assess the ability of the chimeric constructs to inhibit IFNβ activation, 293T cells were transfected with the NP constructs along with an IFNβ promoter reporter plasmid that expresses the mCherry fluorophore upon activation. Twenty four hours post transfection, the cells were infected with Sendai virus (SeV) in order to trigger the IFNβ promoter. As control for this experiment, the NP expression plasmid was removed from the transfection and replaced with an empty expression plasmid pcDNA3.1. The positive control was infected with SeV while the negative control was not infected. As seen in Figure 5.1B, transfection with PICV NP inhibits IFNβ activation by Sendai virus infection while TACV NP does not activate the IFNβ promoter. The chimeric TACV NP containing the C-terminal residues of PICV NP (TP377-561) was able to efficiently inhibit IFNβ activation by SeV infection. Upon further chimeric analysis, the minimal residues from Pichinde virus located at positions 377-388 were needed to gain IFN antagonistic function in TACV NP.

Because the level of NP expression in the cell is likely to affect the level of IFNβ inhibition, we next determined the level of NP expression during transfection by western blot analysis. Each NP construct was C-terminally tagged with an HA peptide sequence and whole cell lysates from the transiently transfected 293T cells were ran under reducing SDS-PAGE conditions and subjected to western blot using anti-HA antibodies (Figure 5.1C). Although the NP expression between the different constructs were found to vary, the NP chimeric construct

TP377-388 with one of the least amount of protein expression comparatively to the rest still had anti-IFNβ activity equivalent to the highly expressed PICV NP. Therefore, only minimal amounts of NP protein expression are needed for efficiently inhibition of the IFN response and although the PICV NP was highly expressed relatively to the TACV NP, these levels are mostly likely at saturated levels for anti-IFN activity.

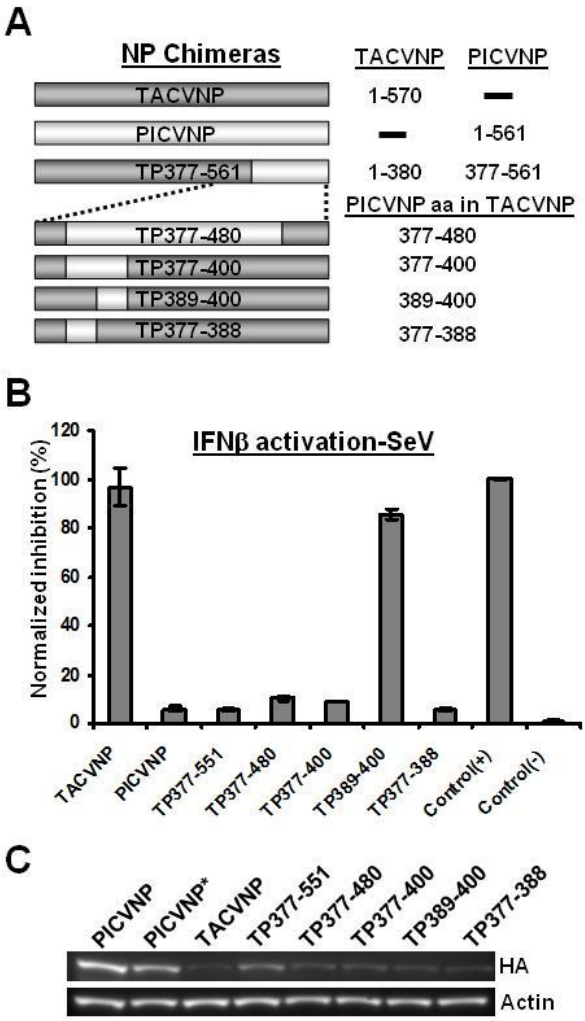



**Figure 5.1 The TACV NP chimera with residues 377-388 from PICV NP rescues anti-interferon activity. (A) Schematic representation of the TACV NP- based chimeras constructs containing C-terminal residues from PICV NP. The numbers represent the amino acid from either TACVNP or PICVNP. (B) Inhibition of an IFNβ promoter by the NP chimeras. 293T cells were seeded into a 12 well tissue culture plate and co-transfected with an IFNβ promoter plasmid expressing mCherry upon activation, a GFP expression plasmid for transfection efficiency normalization, and the NP chimeras using TranIT transfection reagents. For control experiments, the NP expression plasmid was replaced with an empty plasmid vector. All conditions, except for the negative control, were then infected with SeV 24 hours post transfection. Sixteen hours later, the cell lysates were analyzed by fluorescent plate measurements and the IFNβ activity measured by mCherry fluorescence was normalized to the GFP transfection efficiency signal. The positive control experiment was set to 100% activation and the all experiments were normalized**

**to this control. (C) Protein expression levels of the NPs and chimeras. 293T clarified cell lysates were prepared from the above transfection experiments, before SeV infection, and analyzed by western blot using anti-HA (HA) and anti-actin (Actin) antibodies.**

### *5.3.2 Complementary chimeric nucleoprotein analysis identifies amino acids involved in the lack of anti-interferon activity in Tacaribe virus NP*

Having identified that residues 377-388 from PICV NP could be substituted in the homologous TACV NP region to gain anti-IFN function, we hypothesized that replacing those same residues in PICV NP with those from TACV NP would knock out anti-IFN function. The PICV NP chimera with residues 383-396 from TACV NP termed the PT 383-396 NP chimera was constructed and analyzed for IFNβ activation potential. Strikingly, using the same activation experiment in 293T transfected cells as described previously, this chimera was found to retain anti-IFN activity. Thus, to map the region in TACV NP that would knock out anti-IFN in PICV NP, we made another series of chimeras, similar to those made in Figure 5.1A, but using PICV NP as the template to place TACVNP residues into the C-terminal domain. The PICV NP chimera constructs are detailed in Figure 5.2A. The inhibitory effect of the NP chimeras was again tested by transfecting 293T cells with the IFNβ reporter plasmid, the NP expression plasmids and a GFP expression plasmid for transfection efficiency normalization. The controls are similar to those described for Figure 5.1B. Upon SeV infection, we found when residues 383-406 from TACV NP were substituted into the analogous position in the PICV NP, the anti-IFN function was eliminated. To verify the expression of these NP chimeras, whole cell lysates were prepared from the tranfected 293T cells and analyzed by western blot using both anti-HA and anti-actin antibodies (Figure 5.2C). Although the proteins were again varied in expression levels, only the amounts equivalent to the TACV NP are necessary for function as seen previously in Figure 5.1C. The amino acid region containing the TACV NP residues 383-406 was aligned to other representative New World arenavirus NPs (Figure 5.2D). Interestingly, two domains each containing four amino acid stretches seem to deviate from the rest of the arenavirus NP residues within this region. Both of these domains are necessary to abolish anti-IFN function in PICVNP, while only the DLQL residues from TACV NP need to be replaced by PICV NP GPTN to gain antagonistic IFN activity.

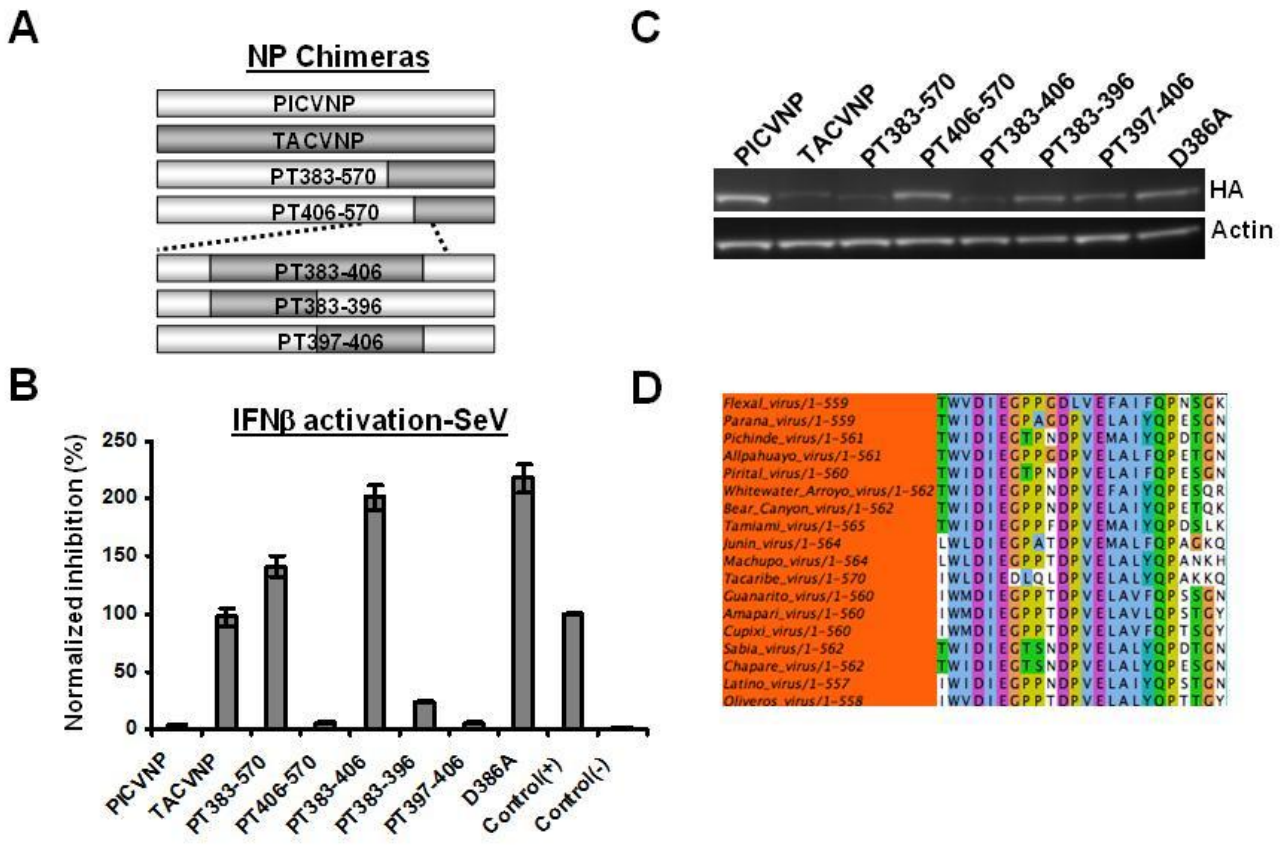



**Figure 5.2 The PICV NP chimera with residues 383-406 from TACV NP abolish anti-interferon activity. (A) Schematic representation of the PICV NP based chimeras constructs containing C-terminal residues from TACV NP. The numbers represent the amino acids from TACVNP that were placed into analogous regions of PICVNP. (B) Inhibition of an IFNβ promoter by the NP chimeras during SeV infection. 293T cells were co-transfected as previously described for Fig 1B. The positive control experiment was set to 100% activation and the all experiments were normalized to this control. (C) Protein expression levels of the NPs and chimeras. 293T clarified cell lysates were prepared, ran under reducing SDS-PAGE conditions and analyzed by western blot using anti-HA (HA) and anti-actin (Actin) antibodies.**

### *5.3.3 A TACV strain contains a GPPT domain in the NP with functional anti-IFN activity*

To check whether sequence variations occur at regions near the important DLQL domain of the TACV NP, we obtained a TACV strain (TRVL-11573) from BEI resources. According to the BEI resources reference material, both the S and L RNA segments of TCRV strain TRVL-11573 have been sequenced (GenBank: M20304 and J04340, respectively). Our TACV NP protein sequence used in Figures 5.1 and 5.2 are 100% homologous to the NP protein found in reference sequence M20304. Interestingly, when the TACV NP gene was cloned and sequenced from the new BEI resources stock, we found that one of the dominant viral species contained a TACV NP protein sequence with 98.2 % sequence similarity to the original reference sequence,

with a key substitution of residues GPPT in the DLQL domain (Figure 5.3). To test whether these GPPT residues in TACV NP have functional anti-IFN activity, the original codon optimized TACV NP, now referred to as coTACV NP, was mutagenized to contain the GPPT residues in place of the DLQL domain (coTNP-GPPT). Similar to experiments performed in Figures 5.1 and 5.2, 293T cells were transfected IFNβ-mCherry reporter plasmid, an NP construct and a GFP transfection normalization plasmid. Twenty four hours post-transfection, the IFNβ pathway was activated either by SeV infection (Figure 5.3B), a synthetic dsRNA ligand poly I:C (Figure 5.3C), or purified TACV genomic RNA (Figure 5.3D). In every method of IFN activation, the coTNP-GPPT construct was able to efficiently inhibit the IFNβ response to levels comparable to PICV NP.

Next, to understand whether the new virus-derived TACV NP sequence (vTACVNP) could inhibit IFNβ activation, the vTACVNP, coTNP-GPPT and the PICVNP constructs were transfected in increasing concentrations into 293Ts and tested using the standard SeV activation protocol used previously (Figure 5.3E). The protein expression levels of these constructs were also tested by western blot (Figure 5.3F). We found that the coTNP-GPPT was able to effectively inhibit the IFNβ response induced by SeV infection to levels similar to PICVNP even though the PICVNP was expressed at greater levels than coTNP-GPPT. Interestingly, the vTACVNP that contains a natural GPPT domain was only able to moderately inhibit the IFNβ response. This reduced level of IFNβ inhibition was mostly likely due to the reduce amount of protein expression seen during western blot analysis of the NP proteins (Figure 5.3F).

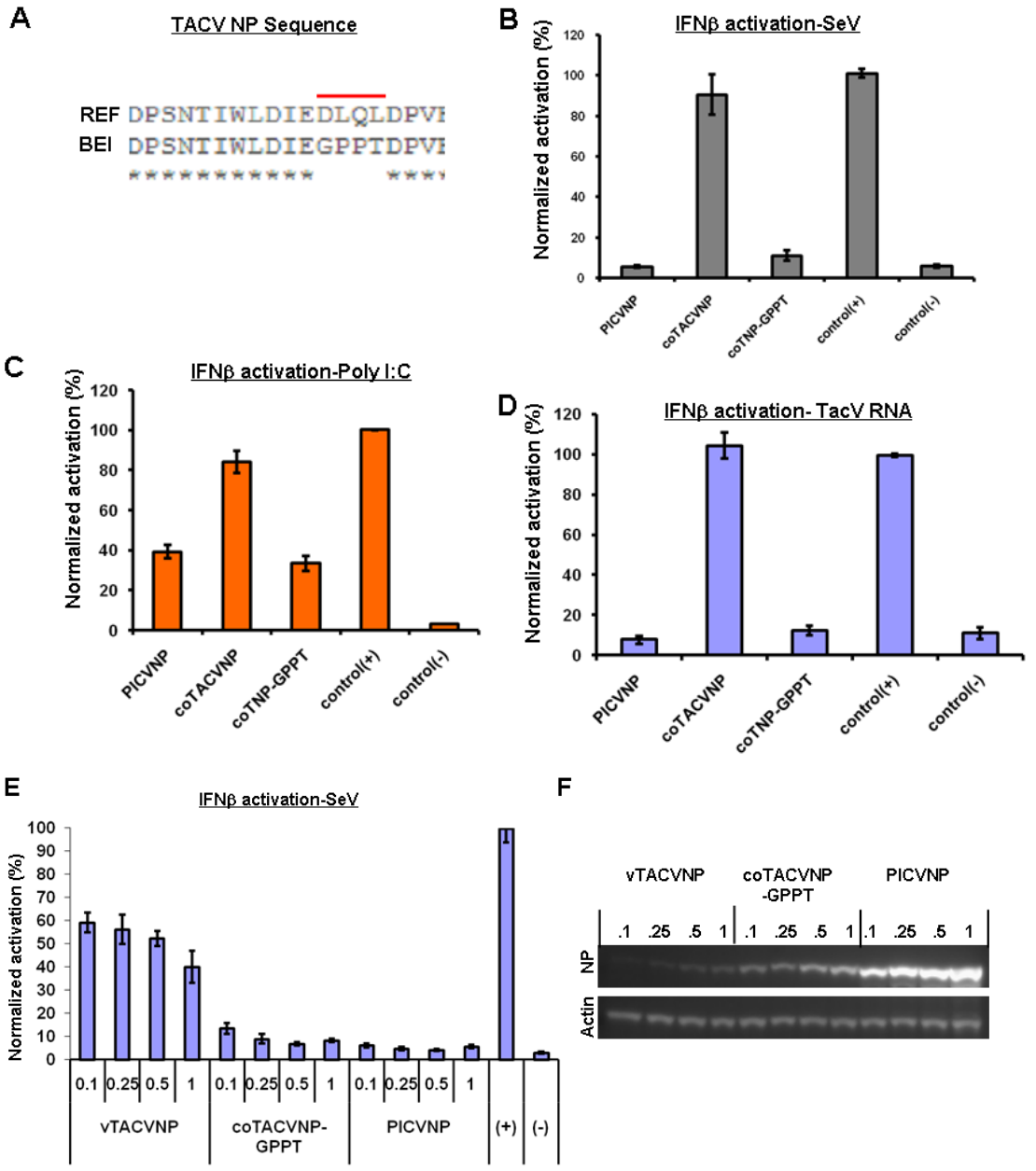



**Figure 5.3 Analysis of a TACV NP with a GPPT domain important for anti-IFN activity. (A) Alignment of the protein sequence near the DLQL domain of the TACV NP reference sequence (REF) to the newly sequenced BEI resources TACV NP strain 11573 (BEI). (B) IFNβ inhibition by the NP containing GPPT during SeV following methods described in previous figures. (C) IFNβ inhibition by the coTNP-GPPT upon transfection with 10□g/ml of low molecular weight poly I:C. (D) IFNβ inhibition by the coTNP-GPPT upon transfection with 250ng of purified TACV RNA per condition. (E) Varying amounts of the vTACVNP, coTACVNP-GPPT, and PICVNP ranging from 0.1-1 µg of plasmid DNA were analyzed by the SeV infection IFN activation assay. The postitive control was set 100% activation and the rest of the conditions were normalized to this value. (F) The levels of NP expression were analyzed 24 hours post transfection by western blot using anti-HA (NP) and anti-actin (Actin) antibodies.**

### *5.3.4 Analysis of virus replication and production in NP overexpression studies*

Two recent studies indicate that the arenavirus NP C-terminal domain functions as a 3'-5' exonuclease. Both of the reports agree that dsRNA can serve as a substrate for the NP exonuclease domain. SeV, like other negative strand RNA viruses, can produce dsRNA species during virus replication. Therefore, it's possible that NP overexpression in the SeV-infected IFN reporter assay may actually limit IFN activation by reducing virus replication and production through the degradation of genomic dsRNA substrates. To test this hypothesis, we used a reporter virus that could easily be monitored for RNA genome integrity after virus entry by monitoring transcription of a reporter gene, and in analysis of virus production. The reporter virus is a recombinant Rift Valley Fever virus expressing GFP in place of the NSs gene (RVFV-GFP). The NSs gene in RVFV inhibits IFN activation and therefore RVFV-GFP infection leads to IFNβ production (9). Using the IFN promoter assay as described in previous experiments, 293T cells were transfected with the PICVNP, coTACVNP and coTNP-GPPT and the IFN promoter activation was monitored after infection with RVFV-GFP. Images taken from these experiments are shown in Figure 5.4A to show the qualitative differences in IFNB activation and RVFV-GFP infection levels. The results of IFN activation were quantified in Figure 5.4B and showed that PICVNP and the TACV NP containing the GPPT residues (coTNP-GPPT) efficiently inhibit IFNβ activation, while the coTACNP construct containing the DLQL residues could not. Analysis of the GFP fluorescence from RVFV-GFP infection indicated that NP overexpression did not affect virus entry or transcription since the levels of GFP in the infected cells were similar in all NP transfected condition regardless of IFN promoter inhibition function (Figure 5.4C). When virus was collected 8, 24 and 48 hours post infection and the virus titers were measured, NP overexpression in the IFN reporter assay did not alter virus production levels in the conditions with anti-IFN activity (PICVNP, coTNP-GPPT) versus those without (coTACVNP and controls) (Figure 5.4D). Collectively, the data indicates that NP overexpression do not interfere with the normal virus lifecycle of RVFV-GFP in the IFN reporter assay.

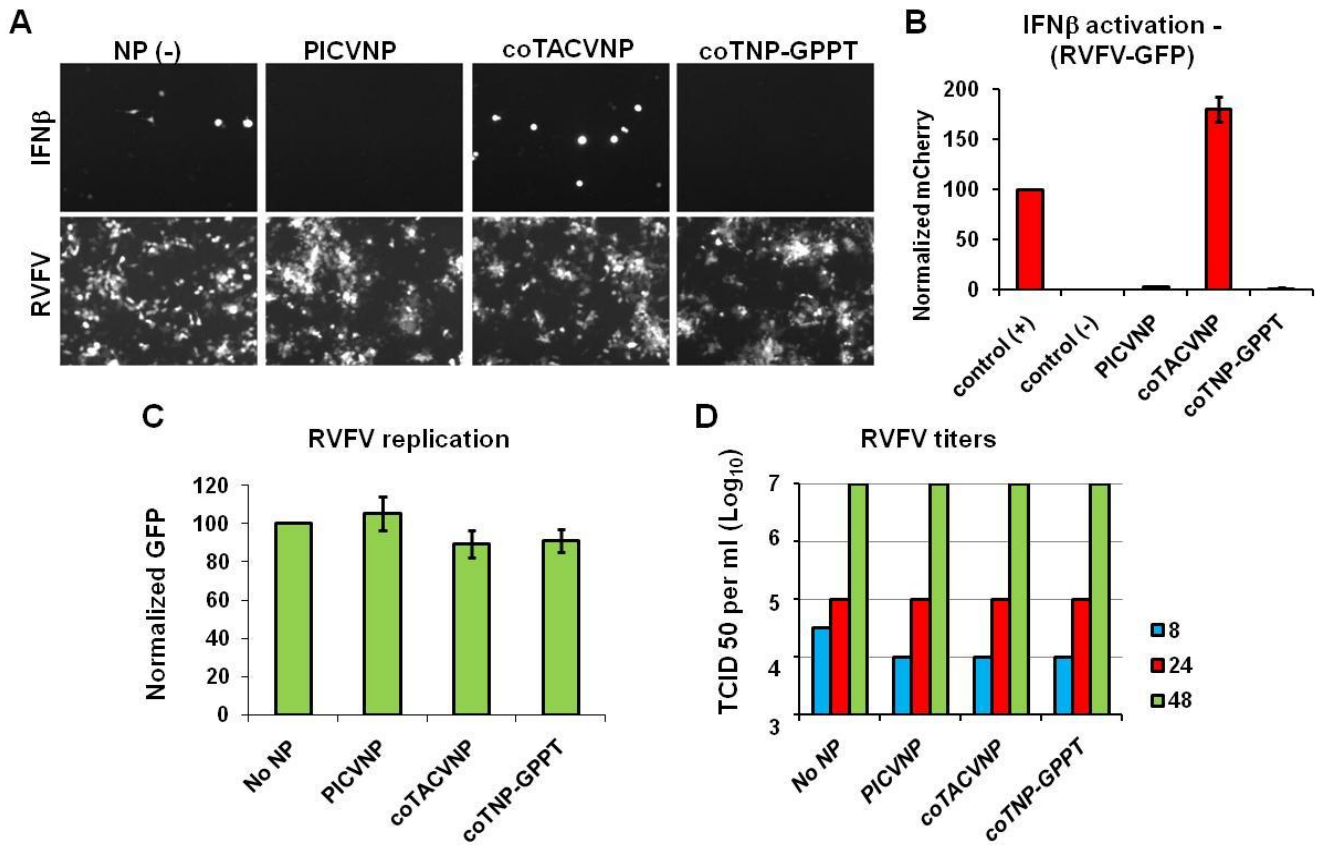



**Figure 5.4 NP overexpression in the IFN reporter assay does not interfere with virus replication or production. (A) Fluorescent images were taken at 20X resolution in both the mCherry channel to measure the IFNβ activation (IFNb) and the GFP channel to measure RVFV-GFP infection (GFP). The images were taken 24 hours post infection. The experimental conditions from the (A) were quantified by cell lysate preparations and fluorescent measurements by a Tecan M1000 plate reader. (B) The mCherry relative fluorescent units in the positive control (+) experiment were set to 100% activation and all other values were normalized to this condition. (C) Similarly, the GFP units in the No NP control were set to 100% and all other values were normalized to this value. (D) RVFV-GFP produced 8, 24 and 48 hours post infection were collected and the virus titers were measured by tissue culture infectivity 50 (TCID50) via serial dilutions into a 96 well microtiter plate of Vero cells.**

### *5.3.5 Minimal IFNβ activation during TACV or PICV infection in A549 or P388D1 cells*

Since our TACV strain contains the important GPPT domain residues in the NP gene critical of IFN inhibition, we wanted to determine whether the TACV strain could prevent IFN activation during virus infection. A549 cells are a human lung epithelial cell line that can produce an endogenous IFNβ response, unlike 293T cells. Therefore, A549 cells were chosen for IFN activation level measurements by real time reverse transcription PCR (RT-PCR). We first determined whether these cells were permissive to infection by TACV. By comparison, the A549 cells were also infected with PICV since the PICV NP has such potent anti-IFN activity and

PICV infection therefore should serve as a control for IFN inhibition. Starting with a range from $10^4$-$10^5$ TCID50 with PICV or $10^4$-$10^6$ for TACV, virus production in A549 cells could be measured as early as 12 hours post infection during PICV or TACV infection (Figure 5.5A) and the viral titers increased during the first 48-72 hours post infection. Total RNA from the infected cells was isolated at 12, 24, 48 and 72 hours post infection. As a control for IFN activation, the A549 cells were infected with SeV and RVFV-GFP. The total RNA samples were subjected to analysis by real time PCR for IFNβ mRNA expression the levels of IFN activation were compared to the no virus infection control (Figure 5.5B). SeV produced a strong IFN response early during the course of infection (12hr), while RVFV-GFP tended to peak around 24hr post infection. Interestingly, both TACV and PICV did not stimulate a potent IFNβ response through 72 hours post infection, not even with high titers of initial virus infection.

IFNβ activation by TACV and PICV infection was also measured in a P388D1 mouse macrophage cell line with an IFNβ reporter stably integrated (Figure 5.5). A range of virus from $10^3$-$10^5$ TCID50 with PICV or $10^4$-$10^6$ for TACV was used to infect the P388D1 cells. IFNβ activation was monitoring through expression of a fluorescent reporter via the IFNβ promoter. For a period of 24-72 hours, the cells did not activate the IFNβ promoter above the background of a no virus infection control. The SeV and RVFV-GFP infection both induced IFN promoter activation with a signal that was elevated through the 72 hours measurement period. The virus titer from these infections were measured and although the PICV infection was productive in all initial virus titer tested, only the highest levels of initial TACV infection produced moderate amount of virus in P388D1 (Figure 5.6C). In all, both TACV and PICV infection of A549 cells or P388D1 cells cause minimal-to-no IFNβ activation, consistent with NP induced IFN antagonistic function.

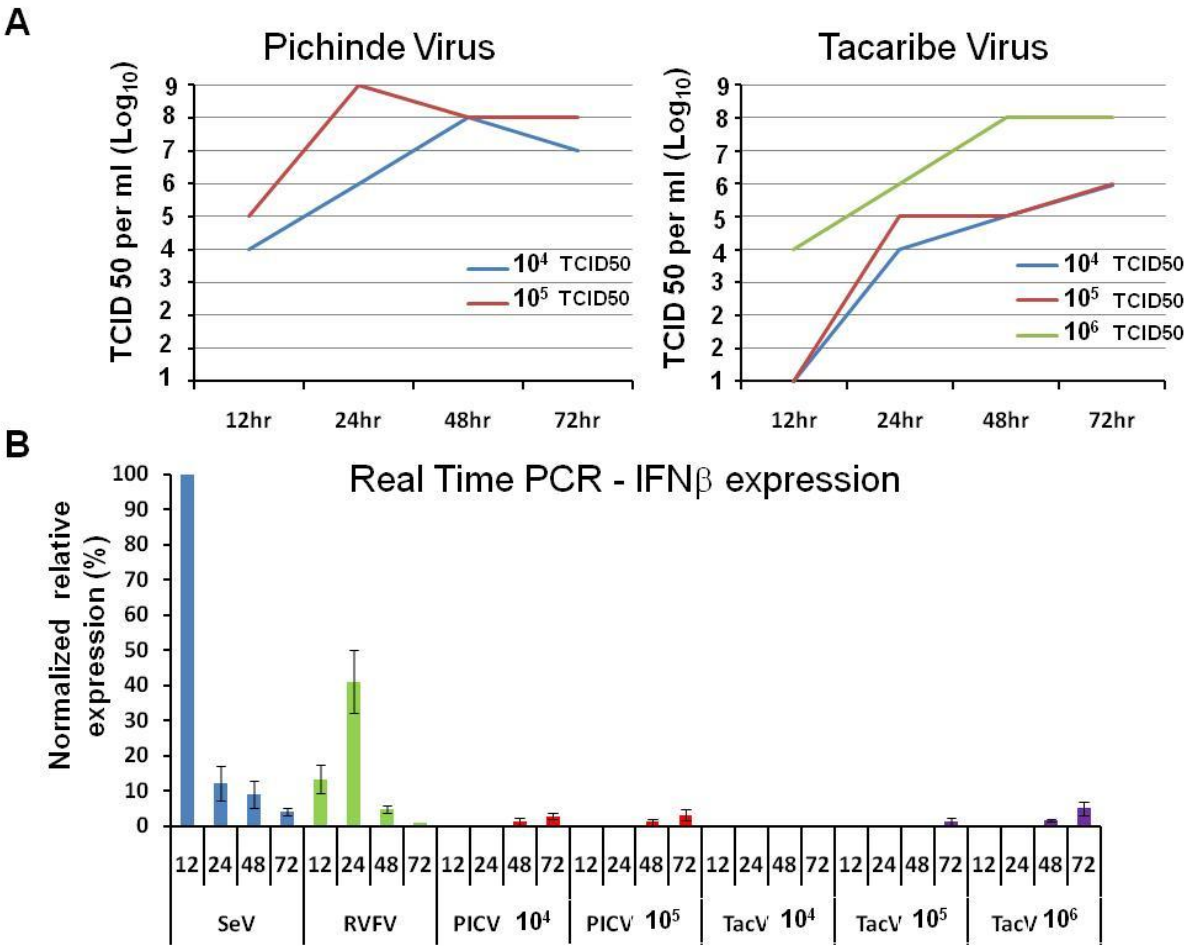



**Figure 5.5 Minimal IFNβ activation during TACV or PICV infection in A549 cells (A) Virus titers for TACV and PICV were measured using TCID50 measurements in Vero cells. Viral supernantants were collected at 12, 24, 48 and 72 hours post infection and clarified before serial dilutions were made. 7-8 days after infection in Veros, the cells were visualized for cytopathic effects (CPE). The TCID50 value was calculated as the last dilution that caused CPE. The PICV was the virulent P18 strain and a gift from Norbert Herzog at the University of Texas Medical Branch. (B) IFNβ expression in A549 cells during virus infection. Total RNA from the TACV, PICV, SeV, RVFV-GFP or no virus mock infections in A549 cells were extracted using a Trizol reagent and protocol. The total RNA was further purified using a Noragen kit and manufactures directions. The real time RT-PCR primer and probe for human IFNβ were purchased from Applied Biosystems the IFNβ expression measurement were analyzed on a Biorad CFX384. The IFNβ levels were calculated as the difference from the no virus infection control. The SeV time point at 12 hours was set to 100% activation and all the values were normalized as a percent of this reference value.**

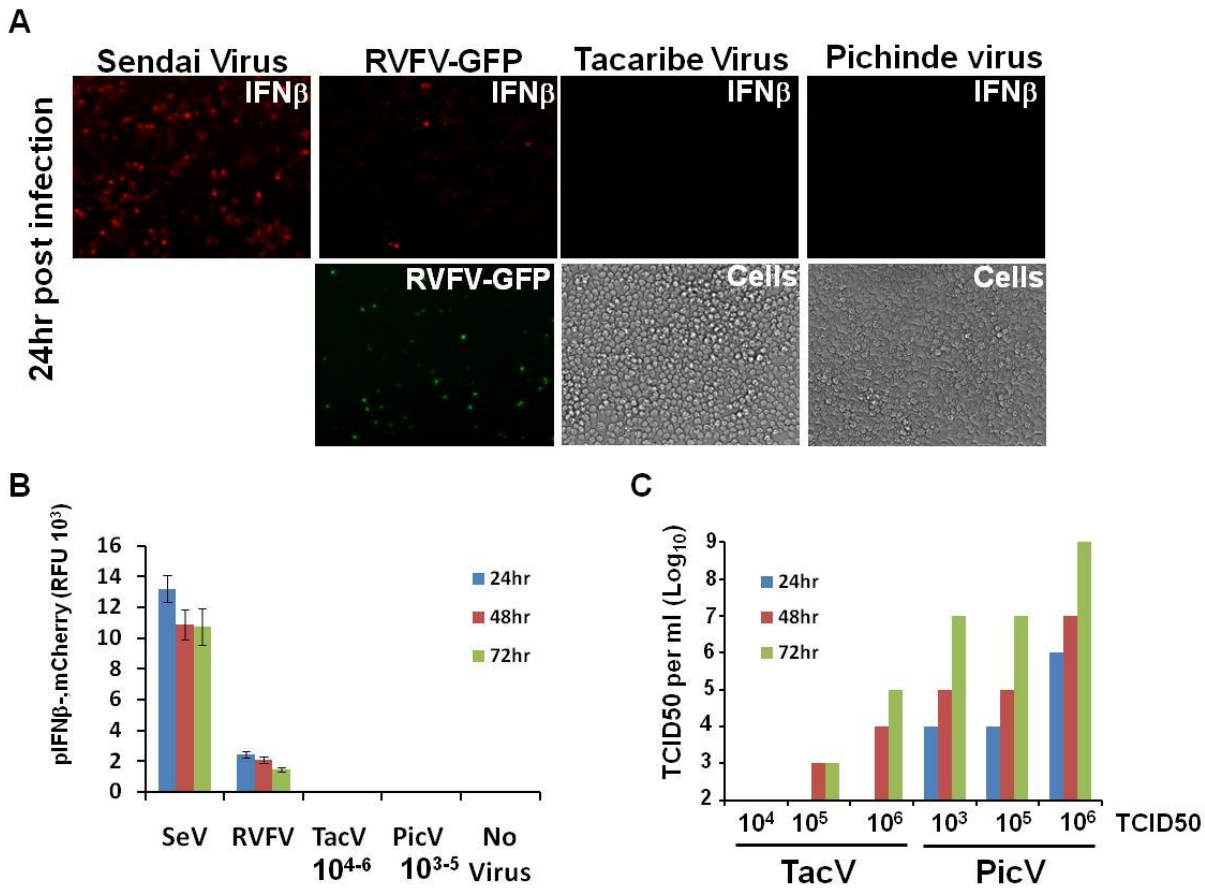



**Figure 5.6 TACV and PICV infection does not activate the IFNβ promoter during virus infection of a mouse macrophage (P388D1) stable cell line. (A) Images of the P388D1 IFN reporter cells were taken at 24 hours post infection with SeV, RVFV-GFP, TACV, or PICV. The IFNβ activity is shown by the reporter mCherry in red (IFNβ). GFP fluorescence indicating that RVFV-GFP infection occurred is shown in green (RVFV-GFP). Bright field images of the cells in the TACV and PICV infection are also shown (Cells). (B) The mCherry reporter fluorescence was measured at indicated time points. These infections were performed in a black-clear bottom 96 well plate and fluorescence was measured using a Tecan M1000 plate reader. (C) Virus titers for TACV and PICV were measured using TCID50 measurements in Vero cells as described in Figure 5.5A.**

### *5.3.6 TACV infection can suppress IFNβ activation when stimulated by a secondary infection*

While we found that TACV did not stimulate a strong IFNβ response during infection in IFN producing cells, it was still unclear whether TACV infection involves activate suppression of the innate immune response or evasion of the cytoplamic sensors. Both mechanisms have been cited as reasons of immune suppression by other arenaviruses. Some viruses replicate in particular compartments in the cell not available to the host PRRs as a method of evading the host nucleic acid sensors. It has been prosposed that the complementary ends of the single

stranded arenavirus genomes anneal in a manner to leave a one base overhang that is unable to trigger RIG I. This would also be a method of evasion. Arenavirus NPs have also been shown to co-immunoprecipate with RIG I and MDA5 sensor as a potential means of active immune suppression. To test whether TACV can complement anti-IFN activity during a secondary virus infection, we subjected 293Ts to transfection with an IFNβ reporter plasmid, followed by primary infecton with TACV or PICV 24 hours post transfection. Virus titers ranging from $10^3$-$10^5$ TCID50 with PICV or $10^4$-$10^6$ for TACV were used to infect the transfected 293Ts. To determine whether TACV and/or PICV could produce production infection in these 293Ts, viral titers were measured at 24, 48 and 72 hours post infection. Both TACV and PICV could be detected from the supernatants of the 293Ts as early as 24 hours and production of virus increase through the 72 hours time point tested (Figure 5.7A). A secondary infection was performed with SeV 24 hours after initial infection with TACV or PICV. Remarkably, the IFN reporter was not activated in condition where the primary infection consisted of PICV at $10^5$ TCID50 or TACV at $10^6$ TCID50 (Figure 5.7B). In this case however, we could not determine whether SeV infection occurred in those conditions. Alternatively, we used RVFV-GFP to stimulate the IFN response in TACV or PICV infected cells. The results showed that both TACV and PICV could reduce the IFN response induced by RVFV-GFP (Figures 5.7C). We also verified that RVFV-GFP entered the infected cells by measuring the GFP fluorescence in the cells (Figures 5.7D). In all, TACV and PICV could actively suppress IFNβ activation induced by co-infection.

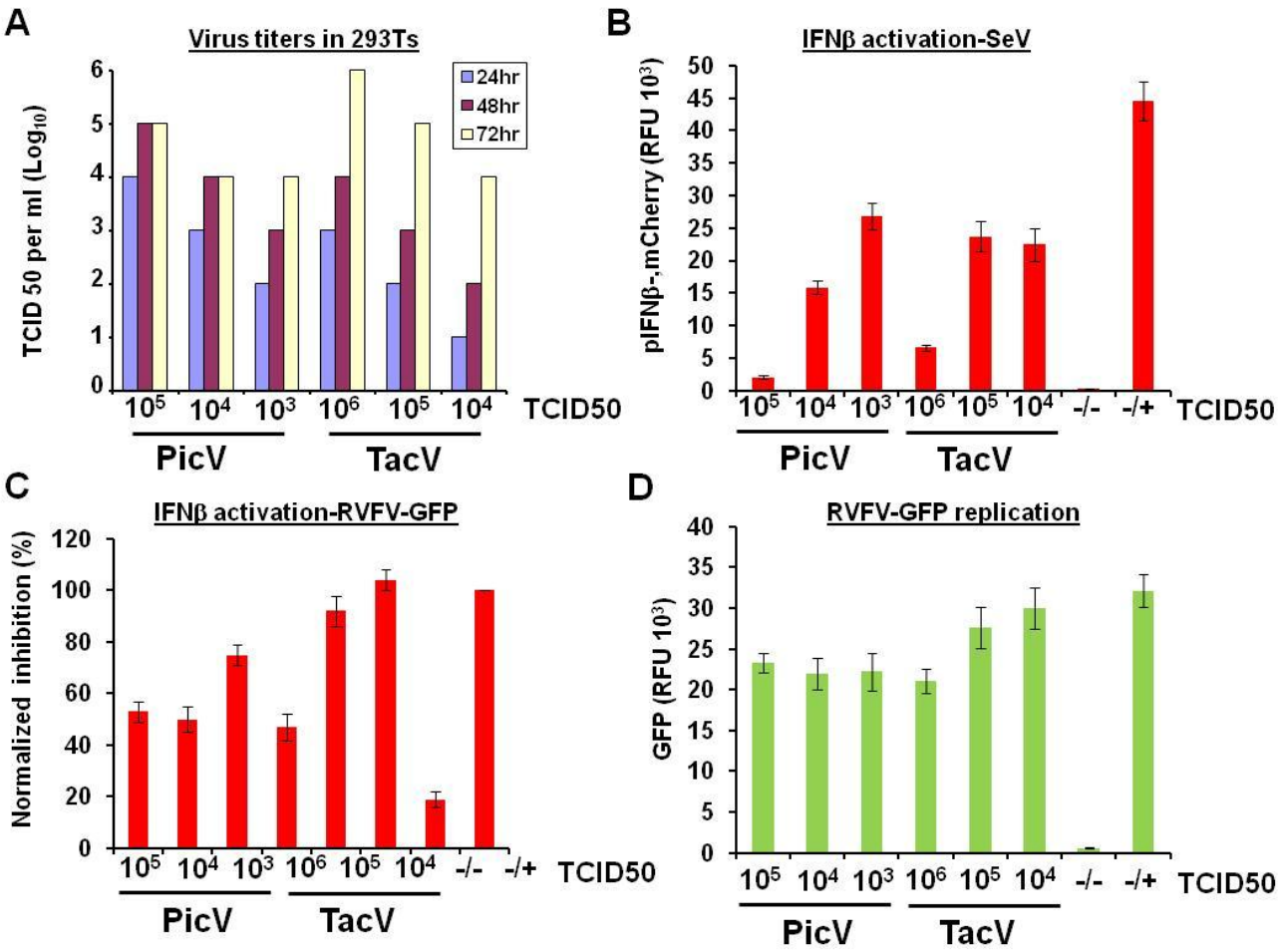



**Figure 5.7 Infection with TACV reduces IFNβ promoter activation after a secondary infection. 293Ts were tranfected with the IFNβ reporter plasmid and 24 hours later infected with virus titers ranging from 103-105 TCID50 with PICV or 104-106 for TACV. (A) Viral supernantants were collected at 24, 48 and 72 hours post infection and TCID50 titers were determined as described in Figure 5.5A. (B) A secondary infection was performed with SeV 24 hours after initial infection with TACV or PICV. Control conditions consisted of no primary or secondary infection (-/-) or only SeV infection (-/+). IFN promoter activity was measured through reporter mCherry fluorescence in relative fluorescent units (RLU). (C) RVFV-GFP was used in the secondary infection assay seen in (B). IFN promoter activity by reporter fluorescence was collected RLUs and normalized to the (-/+) condition that was set to 100% activation. (D) Similarly, the GFP RLUs from RVFV-GFP infection were collected using the same samples in (C).**

## 5.4 Conclusion

Here, we document for the first time that the nucleoprotein of TACV can suppress the type I interferon response. Consistent with this anti-IFN activity in the TACV NP, Tacaribe virus infection does not potently activate IFNβ production in host cells as compared to infection with

Sendai virus or Rift Valley Fever virus deleted in the NSs protein. Additionally, TACV infection can suppress IFNβ activation stimulated through SeV or RVFV-GFP infection. Since the immunosuppression induced by the NPs appears to be a conserved feature of all arenaviruses, both pathogenic and nonpathogenic, this anti-IFN function may serve more as a basic survival mechanism of arenavirus infection than a determinant of pathogenicity.

## 5.5 References


1. **Buchmeier, M. J., C. J. Peters, and J. C. de la Torre. .** 2007. Arenaviridae: the viruses and their replication, p. 1792–1827. *In* D. M. K. a. P. M. Holey (ed.), Fields virology, vol. 2. Lippincott Williams & Wilkins, Philadelphia.
2. **Hastie, K. M., C. R. Kimberlin, M. A. Zandonatti, I. J. MacRae, and E. O. Saphire.** 2011. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. Proceedings of the National Academy of Sciences **108:**2396-2401.
3. **Martinez-Sobrido, L., P. Giannakas, B. Cubitt, A. Garcia-Sastre, and J. C. de la Torre.** 2007. Differential Inhibition of Type I Interferon Induction by Arenavirus Nucleoproteins. Journal of Virology **81:**12696-12703.
4. **Martinez-Sobrido, L., E. I. Zuniga, D. Rosario, A. Garcia-Sastre, and J. C. de la Torre.** 2006. Inhibition of the Type I Interferon Response by the Nucleoprotein of the Prototypic Arenavirus Lymphocytic Choriomeningitis Virus. Journal of Virology **80:**9192-9199.
5. **Qi, X., S. Lan, W. Wang, L. M. Schelde, H. Dong, G. D. Wallat, H. Ly, Y. Liang, and C. Dong.** 2010. Cap binding and immune evasion revealed by Lassa nucleoprotein structure. Nature **468:**779-783.
6. **Rehwinkel, J., C. P. Tan, D. Goubau, O. Schulz, A. Pichlmair, K. Bier, N. Robb, F. Vreede, W. Barclay, E. Fodor, and C. Reis e Sousa.** 2010. RIG-I Detects Viral Genomic RNA during Negative-Strand RNA Virus Infection. Cell **140:**397-408.
7. **Versteeg, G. A., and A. García-Sastre.** 2010. Viral tricks to grid-lock the type I interferon system. Current Opinion in Microbiology **13:**508-516.
8. **Zhou, S., A. M. Cerny, A. Zacharia, K. A. Fitzgerald, E. A. Kurt-Jones, and R. W. Finberg.** 2010. Induction and Inhibition of Type I Interferon Responses by Distinct Components of Lymphocytic Choriomeningitis Virus. Journal of Virology **84:**9452-9462.
9. **Ikegami, T., S. Won, C. J. Peters, and S. Makino.** 2006. Rescue of Infectious Rift Valley Fever Virus Entirely from cDNA, Analysis of Virus Lacking the NSs Gene, and Expression of a Foreign Gene. Journal of Virology **80:**2933-2940.

# 6. COMPUTATIONAL PREDICTION OF HOST-VIRAL PROTEIN-PROTEIN INTERACTIONS FOR ARENAVIRUS PROTEINS

## 6.1 Background

A wide variety of biological processes, such as signaling cascades, metabolic cycles, and DNA transcription and replication, require an understanding of how proteins interact with each other. Proteins may form stable or temporary complexes of functional importance. Yet laboratory experiments to find interactions between two proteins are currently labor, cost, and time intensive. Some of these identification methods include yeast two-hybrid screening, protein affinity chromatography, or *in vitro* enrichment of protein complexes that were tagged *in vivo*. Yeast two-hybrid system can readily detect interacting protein pairs but has a high false positive rate and often needs confirmation from other methods. Tandem affinity purification can also detect real interactions but the tagging process may introduce structural changes in the tagged protein and the method may purify additional contaminating proteins. At times there may be reasonable doubt about the degree of the interaction *in vivo* since the interacting proteins are isolated as complexes and hence it may be difficult to tell direct interactions apart from indirect ones. Such limitations of biochemical methods have led to the development of computational approaches for predicting protein-protein interactions (PPI).

The methods that have been developed for PPI prediction may use structural information [1,2], previously identified domains [3,4], sequence similarity [5,6], or primary structure [7,8] of proteins. Algorithms have been designed to search for sequence conservation, occurrence of fused genes, coevolution of interacting proteins, and similarity of phylogenetic trees (reviewed in [9]). Some methods combine experimental information with previously known biological phenomena [7,10]. Another was developed at SNL and uses the *signature* descriptor to capture physicochemical information from the primary sequence in a support vector machine (SVM) classifier [8]. This method was tested on yeast, human, mouse, *H. pylori*, and *E. coli* and achieved 70-80% cross-validated accuracy. An interesting method that also uses primary sequences as input hypothesized that some PPI may be dependent on a finite set of short polypeptides that may occur across multiple species or organisms [11]. This method achieved an overall accuracy of 75% for its intended system of proteins in the yeast *S. cerevisiae*.

In this chapter, we report our experience with using the primary sequence based methods of Martin et al [8] and Pitre et al [11] in the prediction of host-viral PPI for arenavirus proteins. Arenaviruses are thought to inhibit host production of cytokines by ultimately inhibiting the NF-kB stimulatory pathway. However, specific host-pathogen PPI are not known. Project level interest in a) validating novel experimental techniques for coimmunoprecipitation of specific host-viral proteins and b) modeling the possible disruption by arenaviruses of the network pathway responsible for balancing the concentration of NF-kB dimers, provided the motivation for predicting host-arenavirus PPI. Prior knowledge through predicted PPI would enable focused efforts in both of the above areas.

## 6.2 Methods

### *6.2.1 SVM classifier based upon signature product*

This method is based on using *signature* products for developing a kernel for an SVM [8]. The kernel is used to train a classification model based on a training set of known positive interactions and random negative interactions. The model is then used to classify interacting/non-interacting predictions of unknown PPI of specific interest. This technique has previously been used with some success in predicting host-bacterial PPI at SNL, in particular for predicting interaction of Yersinia outer proteins with host TLR network proteins [12].

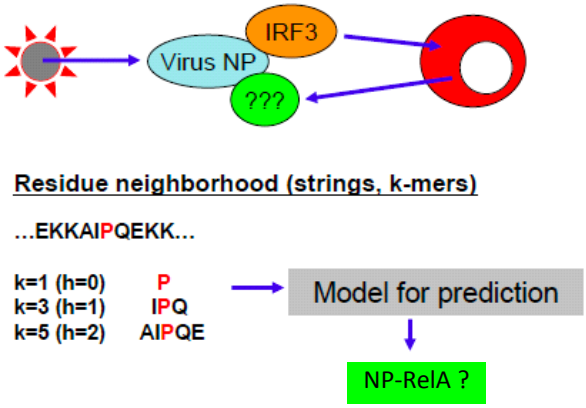



**Figure 6.1 Signature kernel used in an SVM classifier for host-arenavirus PPI**

a) *Signature* and *signature* product

In order to use a machine-learning algorithm (such as an SVM) for predicting PPI, it becomes necessary to represent the amino acid sequences as vectors containing the information that may distinguish between binding and non-binding protein pairs. In this project, this was achieved by using the technique developed earlier at SNL, namely the *signature* descriptor [8,13].

Essentially, a *signature* (of "*height*" = 1) is an amino acid and its neighbors in the primary sequence (hence it is a trimer, see k=3 in Fig. 1). The *signature* space is the set of all possible *signatures* for a given sequence. *Signature* has been formulated as a function $s$ : {variable length amino acid sequences} $\rightarrow F$ defined by $s(A) = \sum_i \sigma_i \mathbf{z}_i$, where A is an amino acid sequence, $\mathbf{z}_i$ is a basis vector in the *signature* space $F \cong \mathbb{R}^N$ and $\sigma_i$ is the number of occurrences of $\mathbf{z}_i$ in $A$. Thus, for the five-letter amino acid sequence MLLMK, there are three trimers MLL, LLM, and LMK. The middle letter of each *signature* is designated as the root and its two neighbors are listed in alphabetical order. The *signatures* for this sequence are L(LM), L(LM), and M(KL) and $s$(MLLMK) = 2 L(LM) + M(KL). Note that longer subsequences can also be handled by using *height* > 1 (for *height* = 2 the *signature* is a 5-mer and for *height* = 3 it is a 7-mer). This formulation of the *signature* descriptor has worked well for HIV protease-1 peptide prediction [14] and for inverse design of LFA-1/ICAM-1 peptides [15]. In order to

classify interacting or non-interacting *pairs* of proteins, *signature* products have been defined using the idea of a tensor product of vectors. This enables the specification of a suitable SVM kernel for use in PPI prediction.

b) Support Vector Machines

An SVM is used for classification and regression analysis [16]. It is a supervised learning method that, for classification, acts on a set of input data and assigns each data point to one of two classes. The idea is to find the largest separation or maximum margin between the two classes in some hyperspace such that the data points can be best classified. In biology, SVM's have been used to, for example, classify gene expression data [17], detect homology [18], and predict PPI [7].

c) Training set

Ideally, the SVM training set should contain instances of known host-pathogen PPI. However, this information is generally unavailable, especially for the small number of proteins present in a virus. To work around this situation, a) known host-host PPI, b) known viral-viral PPI, and c) a small number of known host-viral PPI were included in the training set, which contained a total of 297 PPI mostly from the TLR pathway. The viral-viral PPI included N-G2, N-Z, N-L, N-N, L-Z, L-L, Z-GPC, G1-G1, G2-G2, and G1-G2 for Pichinde virus. The host-viral PPI included PML-Z protein, Ribosomal P proteins-Z protein, TSG101-Z protein, and DAG1-Glycoprotein complex for Pichinde virus.

d) Program/code used

The code (sig_prod_v2.1) was obtained from it's authors: Shawn Martin and Jean-Loup Faulon. It was run on a UNIX machine with Intel Xeon W5580 3.20 GHz processors (16 threads in total). It implements version 5.00 of the SVM-light algorithm [19]. The LOO.pl wrapper script is used to call the other programs in the suite and modified for path and parameter information.

### *6.2.2 Modified "PIPE"*

It should be noted that the training set used for the SVM predictions was based on PPI's belonging to primarily TLR pathway proteins and had a preponderance of relA interactions. Consequently, it became important to evaluate possible biases in the resulting SVM model, which may help explain the constantly high predictions of PPI involving relA. Two ways to address this are: 1) training set expansion such that PPI are not restricted to the TLR pathway proteins and 2) predicting PPI using another method. Both these approaches were combined by using another method for PPI prediction. This method, called PIPE (protein-protein interaction prediction engine) also uses only primary structure information but relies on a scoring method based on PAM120 scoring matrix and not on SVM classification [11]. It was originally developed for predicting yeast PPI and was modified in this project for human-arenavirus PPI predictions.

Briefly, the method assumes that a large set of PPI will contain common interacting subsequences. The givens include two query protein sequences A and B, and a large data set of known PPI. If the query proteins contain subsequences that are also found in many of the known PPI, then it may be hypothesized that A and B are likely to interact. It is a three step process: a) initialize list (L1) of known PPI; b) create a second list (L2) of all proteins similar to the query protein A; and c) identify in L2 sequences that are similar to query protein B. The list L1 is transformed into a graph where a node is a protein and an edge is an interaction. L2 contains sequences similar to fragments of A: the fragments are obtained by using a sliding window (here of size 20 mer), while similarity is calculated using a score based on PAM120 substitution matrix. A sliding window is used again to obtain fragments of B (of size 20 mer again). The result matrix of can be visualized as a 3D surface where rows and columns represent fragments of A and B while the vertical axis represents the score (strong peak indicates good interaction). The method had a sensitivity of 61% for detecting any yeast protein interaction with 89% specificity and an overall accuracy of 75%. It was able to detect novel yeast PPI that was confirmed by TAP tag [11].

There are three tunable parameters in this method: a) the size of the sliding window (set to 20 amino acids here); b) the threshold PAM score that signifies a match between two subsequences relative to the PAM120 matrix (set to 35 here); and c) the number of matches observed from two subsequences, above which PIPE reports an interaction (set to 10 here).

The method was generalized in this project for use with databases of *all* known host-host and host-viral PPI. Thus, the training set of known PPI was enlarged so as to include all available interactions in the Database of Interacting Proteins (DIP) [20] and in VirusMINT [21]. The DIP contains known human host-host PPI while VirusMINT contains known viral-host PPI. The viral-host interactions were for papillomavirus, Epstein-Barr virus, Hepatitis B and C viruses, Human adenovirus, Human herpesvirus, HIV-1 virus, Influenza A virus, Vaccinia virus, and Simian virus 40. The size of the training set was thus expanded from 297 (primarily host-host interactions that belonged mainly to the TLR pathway) to 24,390 interactions (22,208 host-host + 2,182 viral-host).

## 6.3 Results and Discussion

### *6.3.1 SVM model*

SVM learning using the 297 interactions training set resulted in a model with 85% leave-one-out (LOO) cross-validated accuracy. Preliminary predictions for the N protein of Pichinde virus suggest that it may interact with relA and a few other host proteins (p50/p105, IκBα, IκBβ, MEKK3, and TRAF6) but not with IRF3 or TRIF (see Table 1).

**Table 6.1 SVM Classifier Rankings for Positive Host-Pichinde Virus Interactions**

| NP | | LP | | GPC | | ZP | |
|---|---|---|---|---|---|---|---|
| rela | 1.03 | rela | 1.88 | rela | 0.86 | rela | 0.68 |
| nf-κb1 | 0.86 | nf-κb1 | 1.54 | relb | 0.66 | nf-κb1 | 0.52 |
| ikbα | 0.81 | ikbα | 1.44 | nf-κb1 | 0.59 | nf-κb2 | 0.50 |
| ikbβ | 0.78 | traf6 | 1.43 | ikbβ | 0.57 | relb | 0.47 |
| mekk3 | 0.76 | ikbβ | 1.42 | ikbα | 0.55 | c-rel | 0.46 |
| traf6 | 0.75 | relb | 1.32 | c-rel | 0.53 | ikbβ | 0.43 |
| c-rel | 0.71 | ikkγ | 1.30 | traf6 | 0.48 | ikkγ | 0.42 |
| relb | 0.67 | c-rel | 1.30 | mekk3 | 0.47 | ikkα | 0.35 |
| ikkγ | 0.53 | mekk3 | 1.18 | ikbε | 0.47 | mTOR | 0.34 |
| ikkα | 0.51 | nf-κb2 | 1.14 | nf-κb2 | 0.47 | mekk3 | 0.33 |

The result points to a clear bias toward PPI involving relA. The reason for this is that 20% of the training set itself comprises of relA interactions. The 297 PPI were selected from the TLR4 pathway and were originally intended for use in bacterial PPI prediction for that pathway. That training set worked well for some gram-negative bacteria predictions (such as those for Yersinia outer proteins) but did not contain many host-viral PPI that could help train the intended host-viral model/classification. A recent finding has also questioned the use of simple sequence-based kernels for PPI prediction [22] and has implications for training set dependency.

### *6.3.2 Modified PIPE*

The best predictions turned out to be for interaction with Selenoprotein N (see Fig. 2). This protein has been hypothesized to be involved in hemorrhagic fever manifestation (see PAT 6303295) but further evidence is lacking. It was not pursued further since the projects goals were limited to pathways that may directly be affected by the virus. Predictions for interaction with proteins in the TLR pathway were inconclusive.

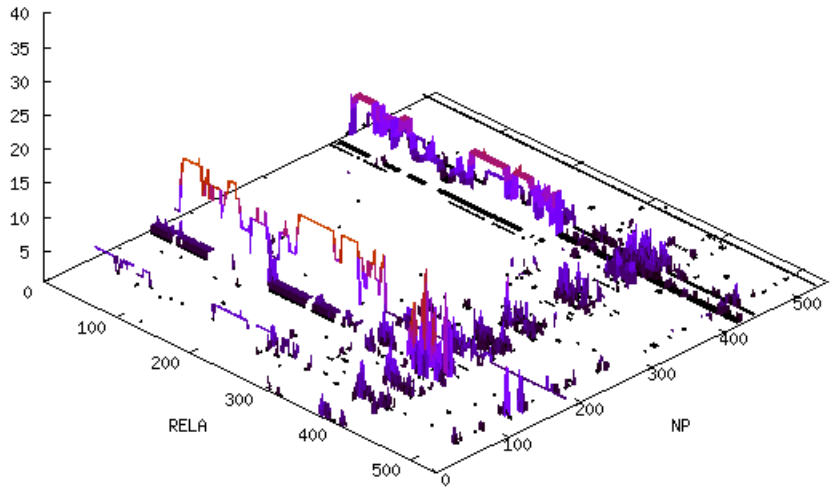



**Figure 6.2 Modified PIPE visualization for relA-NP interaction. Each grid point represents an interacting location for the two proteins on the horizontal axes. The length of the interacting sequences was set to 20 amino acids. A peak represents the number of known interacting protein pairs with sequences similar to the two query proteins. Thus, in this figure for relA and NP, the region around 400 (relA) and 150 (NP) has similar regions in about 20 known interacting protein pairs (present in the training set).**

### *6.3.3 Attempts to Confirm Interactions Empirically*

We attempted to confirm several of the initial SVM method predictions by co-immunoprecipitation. Because we were unsuccessful, what follows is a brief summary of what was attempted. Both adherent and suspension cell lines (RAW264, P388D1, and THP1) were transfected with Pichindi virus (PICV) or Tacaribe virus (TACV) nucleoprotein constructs (NP-IRES-mCherry). We included TACV as a potential negative control since published work claimed that TACV NP failed to inhibit IFN-I production. In testing for expression and translated hemagglutinin tag (HA-NP) and transcriptionally-linked fluorophore, cells transfected via FuGENE HD transfection reagent (Roche, 04 709 705 001) proved to have greater transfection efficiency than electroporation methods using Amaxa Nucleofector Kit V (Lonza,VCA-1003). Additionally we discovered PICV HA-NP protein was consistently expressed at much higher levels than TACV NP in both adherent and suspension cells (Fig.6.3). NP was effectively immunoprecipitated (IP) and protocols were optimized to maximize yield of HA-NP tag construct, as seen in silver stain (Figure 6.3) and subsequent Western blots (Figure 6.4). Immunoprecipitated product probed with with rabbit anti-HA antibody (Santa Cruz sc-805) was visible on Western blots providing evidence of technique/method success. However, rabbit anti-p105 and/or RelA did not demonstrate either method or transfection efficiency were successful. There was no visible band intensity difference in either silver staining or on Western blot. Probing with other antibodies from the prediction list above (Table 6.1) gave similar results.

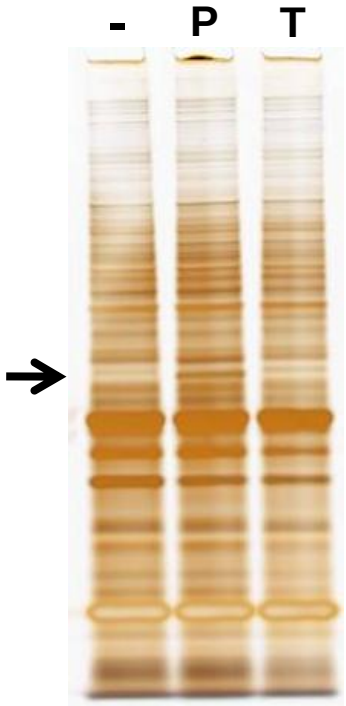



**Figure 6.3 Silver stain, HEK293 transfection with Pichinde or Tacharibe virus nucleoproteins. Arrow indicates position of NP-HA product.**

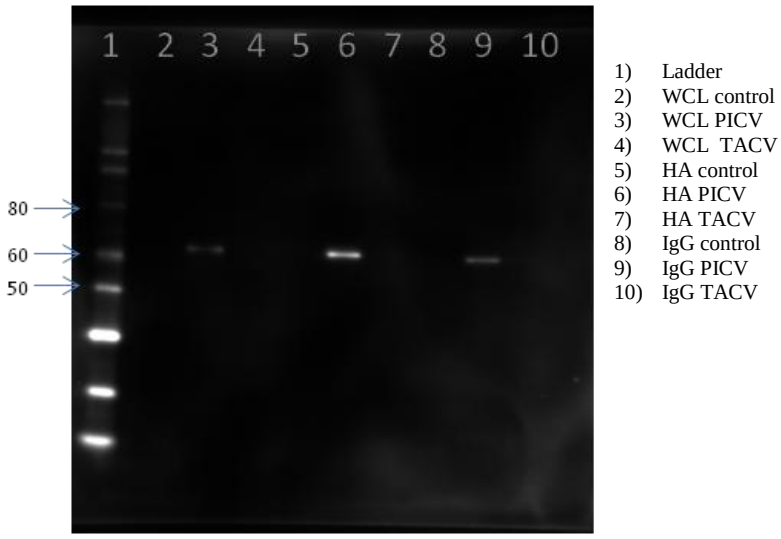



**Figure 6.4 HEK293 transfection Western blot showing successful enrichment of NP-HA.**

Comparison of FuGENE HD and Amaxa Nucleofactor ultimately proved FuGENE HD produced the highest transformation efficiency in both cell lines; Amaxa transformation efficiency was considerably lower and cell death rate unacceptably high (> %70). The ratio of reagent solution to plasmid DNA(7:2), the reagent variation of lysis/wash buffers, and the immunoprecipication (IP) method developed yield an intense band of nuclear protein in addition

to minimizing background protein bands. Increasing the stringency of the wash buffer from 0.1%Triton 100x in PBS to TRIS, 1.0% Triton in PBS, the number of washes from 3 to 5 X, and preclearing lysates (with IgG agarose conjugate) before coIP HA-NP agarose beads provided the clearest picture. However, subsequent experiments withanti-RelA (Fig. 6.5 & 6.6) and anti-p105/p50 (Fig. 8) did not show co-IP with NP. Thus we were unable to verify any of the predicted interactions. There was insufficient time remaining in the project to test the predictions from the PIPE method.

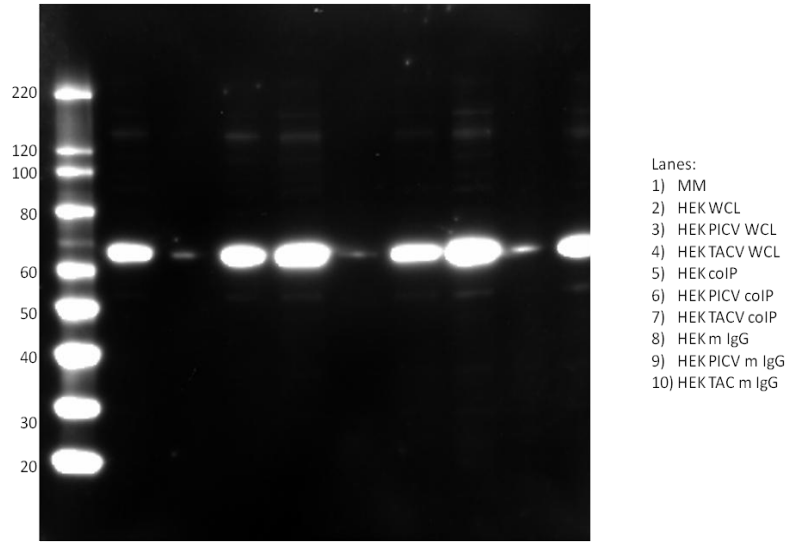



**Figure 6.5 RelA Western Blot: HEK293 whole cell lysate, IgG agrose conjugate (sc-2343), and HA probe agarose conjugate(sc-7392AC); Antibody p65 (Santa Cruz sc-372).**

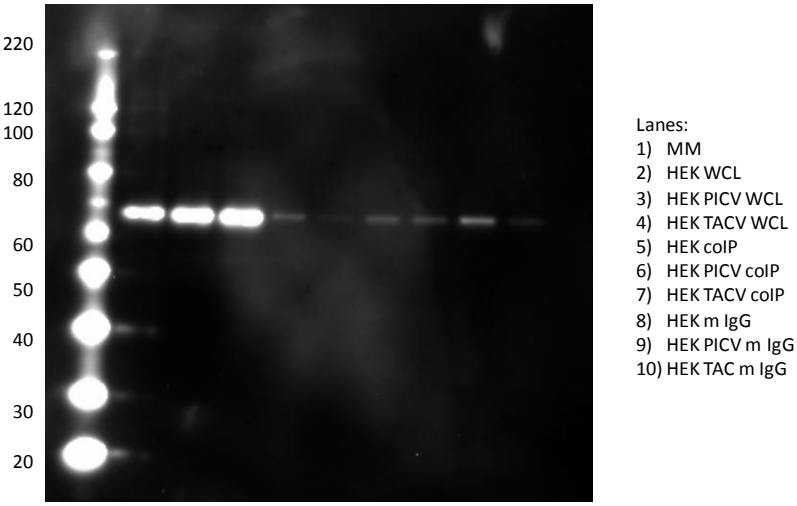



**Figure 6.6 HEK PicV & TacV coIP conjugates; Antibody NF-kappaB1 p105/p50 (CST #3035)**

## 6.4 Recommendations

*a) Build better training set using updated databases*
Although the predictions for host-viral protein-protein interactions were found to be inconclusive, it should be noted that this was an ambitious goal to begin with. The methods used for the predictions depend upon the quality of the training set of known interactions. In case of host-viral interactions, and specifically in case of human-arenavirus interactions, there is sparse documented evidence. Most of the known host-viral interactions in the training set were for viruses (noted above for the VirusMINT database) that are very different from arenaviruses. Updated databases of host-viral interactions may become available soon. These should be used to make newer training sets that are likely to be more representative of the PPI that were sought to be modeled/predicted in this project.

*b) Work with authors of the SVM code to optimize it for large data sets*
The nearly 7,000-fold increase in the pair-wise interaction matrix motivated installation of the SVM code and scripts on a faster computer. However, it was found that the code was not equipped to handle such large data sets even on a fast computer. It was estimated that it would take 6-10 months to do similar calculations on the expanded data set as on the data set of only 297 PPI's. Consequently, an attempt was made to modify the code to permit faster calculations. At a meeting with Shawn Martin (Org 1415), a key contributor to the classification code, it was suggested that the leave-one-out (LOO) method of cross-validation should be changed to leave-group-out (LGO). This should have significantly reduced the time needed to complete calculations but the idea was abandoned primarily because parts of the code were outdated and/or no longer supported by their authors. Crucially, the SVM-light version 5.00 was not compatible with newer operating systems and newer versions of the program failed to work with the wrapping code. It was also difficult to track the original authors for clarifications of their code. The earlier LOO predictions (with the 297 PPI data set) should also be re-run using the LGO cross-validation in order to directly compare them with the predictions using the expanded data set.

*c) Develop an analytical method for PIPE analysis*
PIPE was originally designed for yeast PPI prediction. It was modified in this project to work on a much larger scale and was parallelized and optimized for speed. We now have the capacity to use this method on other similarly interesting problems. Given a good training set of known interactions, and this method, it is possible to get quick predictions and to analyze them visually. Further useful modifications should be explored such that the predictions can be assessed analytically, foregoing the need for visual inspection of thousands of predictions. In this project this aspect was briefly explored by roughly fitting several functions for optimizing the parameters required for PIPE analysis. In future this should be looked at in some detail.

## 6.5 References


1. Ogmen et al; PRISM: protein interactions by structural matching. *Nucleic Acids Res* 2005, **33:**W331-336.
2. Aloy and Russell; InterPreTS: protein interaction prediction through tertiary structure. *Bioinformatics* 2003, **19:**161-162.
3. Han et al; PreSPI: a domain combination based prediction system for protein-protein interaction. *Nucleic Acids Res* 2004, **32:**6312-6320.
4. Kim et al; Large scale statistical prediction of protein-protein interaction by potentially interacting domain (PID) pair. *Genome Inform Ser Workshop Genome Inform* 2002, **13:**42-50.
5. Huang et al; POINT: a database for the prediction of protein-protein interactions based on the orthologous interactome. *Bioinformatics* 2004, **20:**3273-3276.
6. Espadaler et al; Prediction of protein-protein interactions using distant conservation of sequence patterns and structure relationships. *Bioinformatics* 2005, **21:**3360-3368.
7. Bock and Gough; Predicting protein-protein interactions from primary structure. *Bioinformatics* 2001, **17:**455-460.
8. Martin et al; Predicting protein-protein interactions using signature products. *Bioinformatics* 2005, **21:**218-226.
9. Valencia and Pazos; Computational methods for the prediction of protein interactions. *Curr Opin Struct Biol* 2002, **12:**368-373.
10. Sprinzak and Margalit; Correlated sequence signatures as markers of protein–protein interaction. *J Mol Biol* 2001, **311**:681-692.
11. Pitre et al; PIPE: a protein-protein interaction prediction engine based on the re-occurring short polypeptide sequences between known interacting protein pairs. *BMC Bioinformatics* 2006, **7:**365.
12. Falulon, J-L; personal communication.
13. Visco et al; Developing a methodology for an inverse quantitative structure–activity relationship using the signature molecular descriptor. *J Mol Graph Model* 2002, **20:**429–438.
14. Faulon et al; The signature molecular descriptor. 2. Enumerating molecules from their extended valence sequences. *J Chem Inf Comput Sci* 2003, **43:**721–734.
15. Churchwell et al; The signature molecular descriptor. 3. Inverse quantitative structure–activity relationship of ICAM-1 inhibitory peptides. *J Mol Graph Model* 2004, **22:**263-273.
16. Vapnik; Statistical Learning Theory 1998, New York Wiley Interscience.
17. Furey et al; Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 2000, **16:**906–914.
18. Leslie et al; Mismatch string kernels for SVM protein classification. In Becker et al (Eds). Advances in Neural Information Processing Systems MIT Press 2003, Vol. 15, 1441–1448.
19. Joachims, T; Making large-scale SVM learning practical. In Schölkopf et al (eds), Advances in Kernel Methods–Support Vector Learning. MIT Press, Cambridge, MA, 1999, 169–184.
20. DIP was accessed at http://dip.doe-mbi.ucla.edu
21. VirusMINT was accessed at http://mint.bio.uniroma2.it/virusmint
22. Yu et al; Simple sequence-based kernels do not predict protein-protein interactions. *Bioinformatics* 2010, **26**:2610-2614.

# 7. CONCLUSIONS

Over the past three years, this project underwent numerous changes and challenges. Initially, the team was large (~8) and composed almost entirely of postdoctoral fellows. The team had a mandate to continue building on the work in innate immunity begun during the Microscale Immune Studies Lab (MISL) grand challenge LDRD project, and we did. Despite three team members leaving Sandia, major labor cost increases during the project, and technical difficulties too numerous to count, we succeeded in gaining new expertise in Virology, building and testing many novel tools and reagents, and further strengthening Sandia's capabilities in single cell analysis.

Among the difficulties we encountered were problems with hardware. We initially thought that the equipment that was developed near the end of MISL was fully functional and ready to apply. However, we learned there was still a great deal of work to be done and we made numerous improvements to that equipment. It now stands in a far more usable state and we are already preparing to apply it to new projects, including an NIH-funded collaboration with the University of New Mexico.

We also learned that Arenaviruses are much more subtle in their effects at the individual cell level than we had thought based on the literature and on conversations with our collaborators. In fact, we observed hardly any effect on our reporter cells upon Pichinde virus infection save inhibition of IFN-I production, despite verification of live virus by plaque assay. This subtlety was confirmed in later discussions with our colleagues at UTMB. Future studies involving Arenavirus effects on individual cells will require more sensitive and detailed analysis techniques such as digital transcriptomics and siRNA studies. Perhaps the very subtlety of all the effects we attempted to measure at the single cell level point to a much different mechanism of pathogenesis, most likely involving inappropriate hyperactivation of the adaptive arm of the immune system or perhaps a particular non-macrophage cell of the innate immune system such as plasmacytoid dendritic cells.

In summary, the molecular and microfluidic tools developed during this project are already being applied to diverse problems in biodefense-related research, and we will continue to refine them as resources become available. Although we fell short of our hardest goal, identifying a difference in cellular response to infection when cells are isolated, we learned a lot and will soon submit a paper to the Journal of Virology and resubmit another to Lab on a Chip. The work we did to predict virus-host protein interactions ultimately showed us that it is a substantially harder problem than we initially thought. However, that work will continue in the form of a proposal in the very near future.

This was a very ambitious project, and I am deeply impressed with the efforts of all those who contributed to its successes.

# Distribution

| | | | |
|---|---|---|---|
| 1 | MS0359 | D. Chavez, LDRD Office | 1911 |
| 2 | MS0892 | Ronald Manginell | 1716 |
| 3 | MS0892 | Matthew Moorman | 1716 |
| 4 | MS0899 | Technical Library | 9536 (electronic copy) |
| 5 | MS1322 | Milind Misra | 1425 |
| 6 | MS1413 | Bryan Carson | 8622 |
| 7 | MS1413 | J. Bryce Ricken | 8622 |
| 8 | MS1425 | Conrad James | 1714 |
| 9 | MS9291 | Catherine Branda | 8623 |
| 10 | MS9292 | Oscar Negrete | 8621 |
| 11 | MS9671 | Carol Kozina | 8634 |
| 12 | MS9671 | Meiye Wu | 8621 |


Sandia National Laboratories