2014

Characterization of Rift Valley fever virus nucleocapsid protein-RNA binding interactions and development of a high-throughput screening assay for identification of n-RNA binding inhibitors

Mary Lee Ellenbecker

University of Montana, Missoula

CHARACTERIZATION OF RIFT VALLEY FEVER VIRUS NUCLEOCAPSID

PROTEIN-RNA BINDING INTERACTIONS AND DEVELOPMENT OF A HIGH-

THROUGHPUT SCREENING ASSAY FOR IDENTIFICATION OF N-RNA

BINDING INHIBITORS

By

Mary Lee Ellenbecker

B.S., University of Wisconsin – Stevens Point, Stevens Point, WI, 2002

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Ph.D.
in Integrative Microbiology and Biochemistry

The University of Montana
Missoula, MT

January 14th 2014

Approved by:

Sandy Ross, Dean of The Graduate School
Graduate School

Dr. J. Stephen Lodmell, Research Advisor
Division of Biological Sciences

Dr. Scott Samuels, Examination Chair
Division of Biological Sciences

Dr. Michele McGuirl
Division of Biological Sciences

Dr. Jean-Marc Lanchy
Division of Biological Sciences

Dr. Christopher Palmer
Department of Chemistry

UMI Number: 3611852

Dissertation Publishing

UMI 3611852

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Ellenbecker, Mary, Ph.D., Fall 2013   Integrative Microbiology and Biochemistry

Abstract Title: Characterization of Rift Valley fever virus nucleocapsid protein-RNA binding interactions and development of a high-throughput screening assay for identification of N-RNA binding inhibitors

Chairperson: Dr. J. Stephen Lodmell

Rift Valley fever virus (RVFV) is an emerging infectious pathogen that causes severe disease in humans and livestock and has the potential for global spread. Currently, there is no treatment for RVFV infection and there is no proven safe and effective vaccine. Inhibition of RNA binding to RVFV nucleocapsid (N) protein represents an innovative antiviral therapeutic strategy because several essential steps in the RVFV replication cycle involve N binding to viral RNA. The overall goals of our research are to better understand how RVFV N protein interacts with RNA, to develop a high-throughput drug screening assay for identification of compounds that inhibit a N-RNA binding interaction and to test the efficacy of potential antiviral drug compounds during RVFV infection. Completion of the proposed research will result in an increased understanding of the basic biology of this important human pathogen and aid in the design of more powerful and effective anti-RVFV drugs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

CHAPTER 1: GENERAL INTRODUCTION

Rift Valley fever virus (RVFV) was first discovered and characterized in the Great Rift Valley of Kenya in 1931 following the sudden death of approximately 4,700 lambs and ewes over a four week period (24). Since then, RVFV has caused numerous alarming and devastating epidemics throughout Africa, Saudi Arabia and Yemen and is currently considered an emerging infectious pathogen due to increases in the incidence and geographic range of the virus. Extensive efforts have been devoted to understanding the pathology, epidemiology and molecular biology of RVFV to learn how to better control the spread of disease during outbreaks and to develop vaccines and drug therapies for the treatment and prevention of viral infection. This chapter begins by presenting a summary of RVFV transmission, infection and therapy. Next it provides an overview of the information obtained over the past three decades regarding the molecular biology of RVFV and concludes with a discussion of the structure, function and therapeutic potential of viral nucleocapsid proteins.

I. Rift Valley fever virus

Rift Valley fever virus (RVFV) is a single-stranded RNA virus that belongs to the *Bunyaviridae* family and *Phlebovirus* genus. The *Bunyaviridae* family is composed of five genera of viruses: *Bunyavirus, Hantavirus, Nairovirus*, *Phlebovirus*, and *Tospovirus*.

A. Epidemiology

RVFV can infect a wide range of vertebrate mammals but disease is limited to domestic livestock and humans. RVFV is most commonly transmitted through the bite of an infected mosquito. However, the virus can also be acquired by exposure to the blood or body fluids of infected animals or by consuming raw, unpasteurized milk from

infected cows (6). The virus is maintained at low levels in the environment by occasional

infection and amplification of the virus in native wildlife or livestock and by transovarial

transmission in *Aedes* mosquitos. Infected mosquito eggs are extremely hardy and the

virus can lie dormant for years in the soil. A shift to epidemic RVFV activity is typically

preceded by periods of extended rainfall causing flooding and an increase in mosquito

breeding and population density (25). Infected *Aedes* mosquitos feed on livestock, which

serve as amplification hosts by quickly becoming viremic (within 12-24 hours). Other

mosquito species that emerge after flooding act as secondary vectors (30 species from six

genera have been documented thus far) and increase transmission of the virus to humans

and other animals (70). Man-made factors can also contribute to the initiation of an

epidemic cycle. For example, dam building creates reservoirs, stagnant backwater and

generates new breeding ground for mosquitos (69). Also, worldwide trade and

international travel can aid in further dissemination of the virus. Since RVFV can infect

a wide range of mosquito and animal species it could easily survive on other continents

besides Africa and the virus will become more widely distributed in the future.

B. Pathogenesis

The main sites of viral replication in a vertebrate host are the liver, spleen and brain.

Disease symptoms in livestock include excessive salivation, nasal discharge, diarrhea,

and high rates of miscarriage. Newborn lambs and kids are most susceptible and

mortality rates are between 70-100%. The fatality rate in adult cattle and sheep is lower

and may reach 30% (20, 21). Disease in humans can take on several forms. Most people

experience a mild flu-like illness, however complications for infected individuals can be

grave and can result in the development of hemorrhagic fever, hepatitis, encephalitis and

blindness (1, 2). Historically, the case fatality rate for human RVFV infection has been

low (2-5%) (68), however, reports from more recent outbreaks are much higher.

Mortality rates of 14% were reported for the first Arabian outbreak in September 2000

and rates as high as 23-45% were reported from outbreaks between November 2006 and

May 2007 in Kenya, Somalia and United Republic of Tanzania (6, 120). More research

is required to determine if this is a result of under-reporting of mild cases, the virus

encountering a naïve host as it spreads beyond the borders of its endemic region, or if the

emergence of more virulent strains is causing an increase in fatal human cases.

C. Vaccines and anti-viral therapies

There is currently no specific treatment for RVFV infection other than supportive

care. Ribavirin is a guanosine analog that is thought to function by inhibiting RNA virus

replication and was used to treat patients in Saudi Arabia suffering from the hemorrhagic

form of RVF (13). However, its use remains limited because it is a known teratogen in

some animal species and can cause other adverse side effects such as hemolytic anemia.

Furthermore, one study reported that treatment of RVFV infected laboratory animals with

ribavirin resulted in a change in disease characteristics rather than a clearing of viral

infection. A shift from sudden onset of hepatitis to delayed onset of neurological disease

was observed and is likely the result of ribavirin being too large of a molecule to cross

the blood-brain barrier (93). Favipiravir (T-705) is a pyrazinecarboxamide compound

and its metabolite inhibits viral RNA-dependent RNA polymerases without affecting host

cellular DNA or RNA synthesis (36). It acts as a purine in viral RNA replication and

displays more selectivity and specificity for its target than ribavirin. Favipiravir is

efficacious against RVFV infection in cell culture and prevents death from hamsters

infected with the closely related Punta Toro virus (41). It is also a promising broad-spectrum antiviral compound because it is active against a variety of other RNA viruses. Phase II clinical evaluations of favipiravir for influenza therapy have recently been completed in the United States (35).

Vaccines are in varying stages of development. A formalin-inactivated vaccine (TSI-GSD-200) developed in the 1970's has been administered to people at high risk of contracting RVFV but its use is limited because it is in short supply and difficult and expensive to produce (95). Inactivated vaccines are generally not considered practical for immunization of livestock because they require multiple initial inoculations followed by annual boosters. The use of a live-attenuated vaccine is the preferred method for immunization because protection against the virus requires only a single inoculation and during an outbreak it is important to be able to attain rapid immunization. Many of the live-attenuated strains of RVFV tested thus far (Smithburn and MP-12) cause miscarriage and/or birth defects in the fetuses of pregnant animals (12, 50), however, one promising live-attenuated vaccine candidate is R566. R566 is a reassortant of the L and M segments MP-12 and the S segment is from a strain of virus (Clone 13) where the NSs gene (a virulence factor) has been deleted. A vaccination trial conducted using pregnant ewes reported that no animals showed signs of illness or aborted and all animals developed antibody responses. A challenge experiment against live virus was not conducted (13).

D. Public health implications

RVFV represents a significant public health threat for several reasons. RVF epidemics quickly overwhelm the local medical care system and cause both human and veterinary medical resources to be strained. Extensive loss of livestock can have a major

economic impact on a region as well as individual farmers and concerns regarding food safety can have long lasting effects on trade. Furthermore, RVF outbreaks tend to occur in developing countries that do not have the financial means or infrastructure to deal with emergency situations.

II. Molecular biology of Rift Valley fever virus

The Rift Valley fever virion is spherically shaped and measures between 80-120 nm in diameter (48). The outermost layer of the virus is composed of a lipid bilayer (envelope) that contains heterodimers of RVFV-specific glycoproteins Gn and Gc (31). Packaged within each viral particle are the three segments of the viral RNA genome complexed with nucleocapsid protein (ribonucleocapsids) and the viral RNA-dependent RNA polymerase (Fig. 1-1).
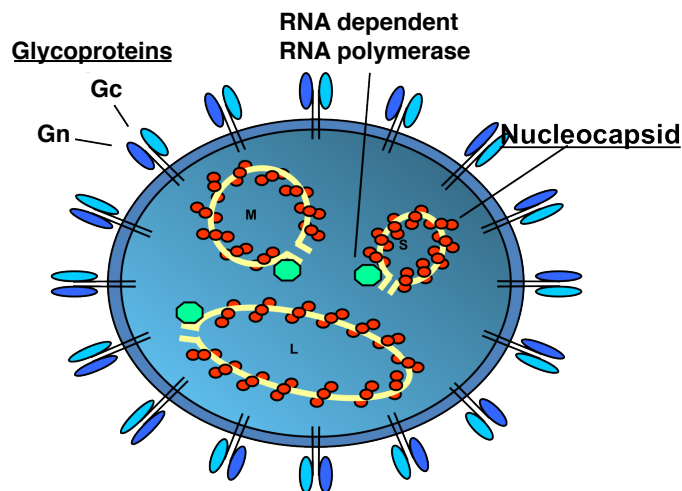
Figure 1 - 1: Schematic cross section of the Rift Valley fever virion.
Three RNA segments (S, M, and L) are complexed with nucleocapsid protein (red circles) to form ribonucleocapsid structures. The ribonucleocapsids and RNA-dependent RNA polymerase (green pentagon) are packaged within a lipid envelope that contains the viral glycoproteins Gn and Gc (blue and aqua ovals).

A. Genome coding strategy

The RVFV single-stranded RNA genome contains three segments designated large (L), medium (M) and small (S) based on their relative nucleotide length. All genome segments share identical termini (3'-UGUGUUUC-5' or 3'-GAAACACA-5') that are complementary. Base pairing of terminal nucleotides results in the formation of panhandle structures and the genomic RNAs adopt a closed circular conformation (49). The L segment encodes the RNA-dependent RNA polymerase responsible for transcription and replication the viral genome (82, 83). The M segment encodes glycoproteins (Gn and Gc) involved in attachment and entry into the host cell and non-structural proteins NSm and 78-kDa (23). The NSm protein acts as a suppressor of virus-induced apoptosis (121) and the function of the 78-kDa protein (if any) is currently unknown. The S segment utilizes an ambisense strategy to encode the nuclocapsid (N) and NSs proteins (51) whereas genes encoded by the M and L segments are oriented in the negative sense. The ambisense coding strategy of the S segment produces NSs from a subgenomic mRNA of the same polarity as the viral RNA. Time course studies conducted in the *Phlebovirus* Uukuniemi show that the mRNA for NSs is copied from complementary RNA (cRNA) after genome replication (117). Non-structural protein NSs is an important virulence factor and utilizes several strategies to prevent activation of the host anti-viral response system. NSs decreases the transcriptional activity of the host cell by sequestering TFIIH (63), a basal transcription factor, and acts as an interferon antagonist by inhibiting production of interferon-β (9, 14). It also promotes the degradation of protein kinase R (43, 52), a protein that can activate the innate immune system by detecting the presence of double stranded viral RNA in the host cell. The viral

16

N protein is an RNA-binding protein that plays an essential role during viral replication and its multiple functions will be discussed later in this chapter.

B. Replication cycle

The RVFV replication cycle begins by the virus attaching to the plasma membrane of the host cell. Attachment is mediated by interactions between the viral glycoproteins (Gn and Gc) and host proteins present on the surface of the cell (59). The virus is endocytosed and acidification of the endosome causes a conformational change in Gn and/or Gc that facilitates fusion of viral and cellular membranes and allows the viral genome and polymerase to be released into the host cell cytoplasm (31, 104). After entry, primary transcription of the negative sense viral RNA to mRNA via the RNA dependent RNA polymerase occurs (Fig. 1-2; step 1). Only ribonucleocapsids and not free RNA can serve as templates for primary transcription (64). Capped oligos are scavenged from host mRNA by the viral polymerase, which has both endonuclease and polymerase activity (45, 90). Some nucleotide preference for endonuclease cleavage has been observed amongst the Bunyaviruses and likely exists because the capped 5' oligos are used as primers to synthesize the viral mRNA (96). L and S segment mRNAs are translated by free ribosomes and translation of M segment mRNAs occurs on membrane-bound ribosomes (Fig. 1-2; step 2). After primary transcription of viral mRNA the virus replicates the RNA genome. The transition to genome replication requires a switch from mRNA synthesis to synthesis of full-length complementary RNA templates and then production of full length viral RNA (Fig. 1-2; step 3). How this switch occurs is not well understood but some viral or host factor(s) suppress the transcription termination signal that generates truncated mRNAs. Next, secondary transcription of NSs mRNA from full-

length complementary RNA occurs (117) as well as continued translation of viral protein

and RNA replication (Fig. 1-2; step 4). Gn, Gc and NSm proteins are transported through

the ER, glycosylated and accumulate in the Golgi apparatus(57). The viral genome is

packaged into virus particles formed by budding from the Golgi apparatus and

subsequent fusion with the plasma membrane results in the release of mature virions (3).

Virus budding directly from the plasma membrane has been described as well (3).
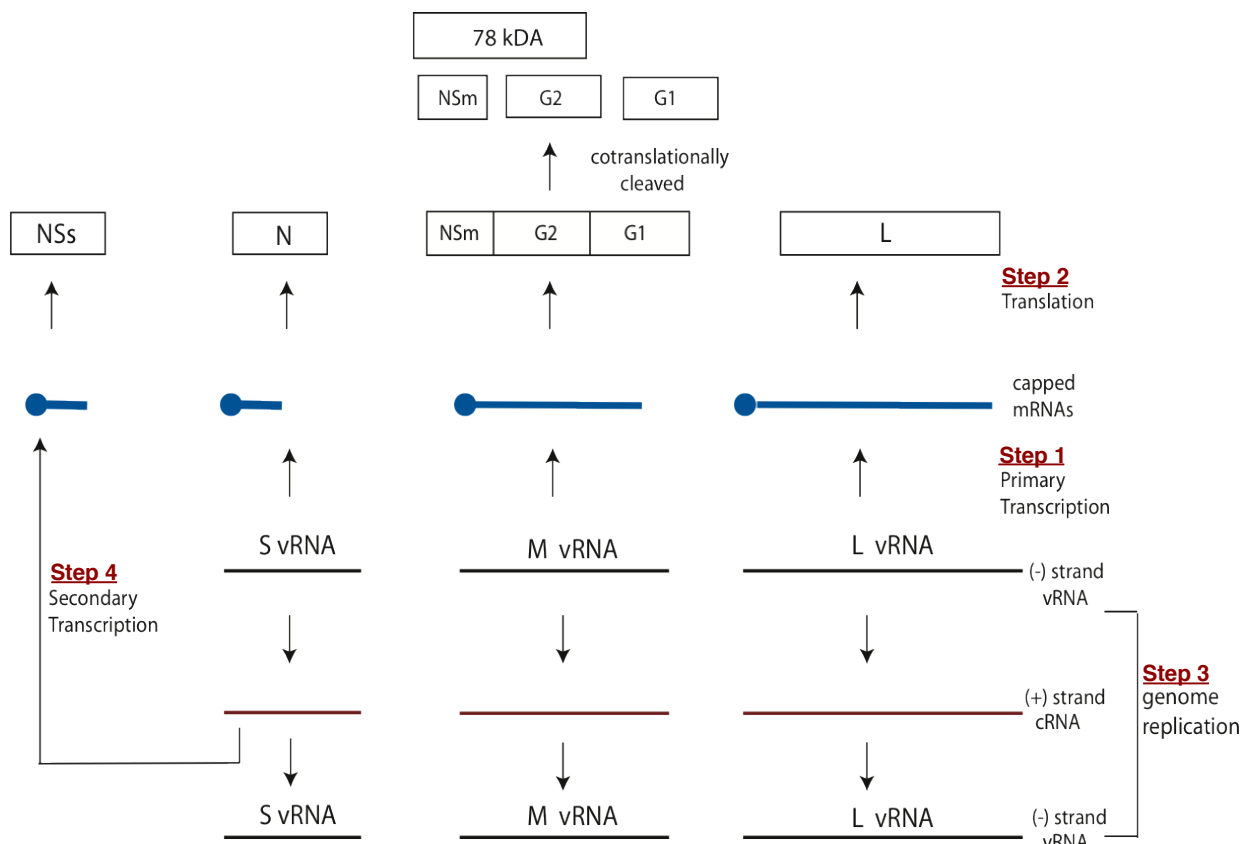
Figure 1 - 2: Principal steps during viral replication.
1.) After the viral genome has entered the cytoplasm of the host cell, primary transcription of negative-sense viral RNA (vRNA) to mRNA occurs. 2.) L and S segment capped mRNAs are translated by free ribosomes and the M segment is translated by membrane-bound ribosomes. 3.) A switch from primary mRNA synthesis to genome replication occurs. Full-length complementary RNA (cRNA) templates are synthesized followed by production of full-length vRNA. 4.) Due to the ambisense nature of the S segment RNA, secondary transcription of NSs mRNA using full-length complementary RNA (cRNA) as a template occurs.

C. Nucleocapsid protein: structure, function and specificity of RNA binding

Nucleocapsid (N) proteins are essential, multifunctional, RNA-binding and RNA chaperone proteins that play an important role during several stages of the viral replication cycle. N proteins are involved in regulation of viral transcription, translation and accumulation of N is thought to trigger the switch from mRNA synthesis to full-length cRNA synthesis during genome replication in some RNA viruses (77, 118). N interacts with viral RNA by binding to and forming polymers along the entire length of the genome (Fig. 1-3). This encapsidation process protects the viral RNA and prevents the formation of double-stranded RNA during replication, which would activate the host anti-viral response. N encapsidates viral genomic and antigenomic RNA but not other cellular RNA's (44). Only encapsidated viral genomes are packaged into virions and assembly of the RNP complexes into the budding viral particle is mediated by interactions between N and the cytoplasmic tail of viral glycoproteins Gn and Gc (88, 101, 114). Most RNA viruses, retroviruses and hepadnaviruses encode a gene for N protein. A rigid comparison of primary sequence and secondary structure suggests that these proteins are unrelated and structurally diverse. However, a broad and general comparison reveals that N proteins are functionally homologous and common structural themes do exist, including a core domain that contains an RNA-binding cleft and N and/or C- terminal extensions that allow the protein to oligomerize.
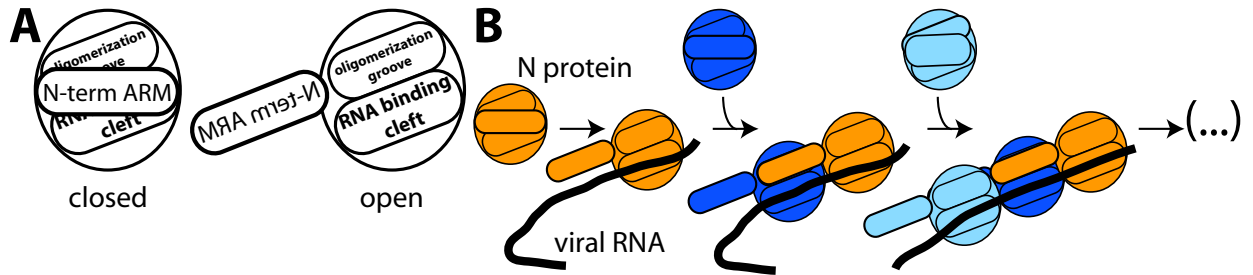
Figure 1 - 3: Schematic of N structural domains and RNA encapsidation.
A) Cartoon of N structure based on the crystal structural model (32, 100). The 33-residue N-terminal arm blocks the RNA-binding cleft in the N monomer. B) A specific RNA-binding event triggers cooperative binding between N monomers, which then switch to a non-specific mode of RNA binding that allows N to coat the viral genome.

The most recent crystal structure depicts RVFV N as a circular shaped multimer (hexa, penta or tetramer) in a complex with either DNA or RNA. Analysis of this structure shows that interactions between the flexible N-terminal domains of neighboring proteins allow multimers to form (Fig. 1-4A). An RNA-binding slot that contains several highly conserved basic amino acid residues lines the inner circle of the ring-like structure (Fig. 1-4B). The RNA sugar phosphate backbone contacts polar residues located on the rim of the binding groove and nitrogenous bases are involved in hydrophobic and base stacking interactions with N. Since this crystal structure shows a non-specific mode of N binding, no H-bond interactions between N and RNA bases are observed (100). However, since N is able to discriminate between different species of RNA in the host cell, N must possess a specific mode of binding in addition to the indiscriminant mode described above.

Research using several different bunyaviruses as model systems shows that N recognizes and binds to specific sequence and/or secondary structures on its target RNA. One previously mentioned characteristic feature of all bunyavirus RNA segments is a conserved, self-complementary sequence present at the 5' and 3' ends of the genome and anti-genome that undergoes base pairing to form panhandle-like structures. Preferential

binding of N to this panhandle structure has been demonstrated in Sin Nombre hantavirus

(72, 76, 78) and it is thought that N binding to and unwinding this RNA helix allows the

viral polymerase to bind and initiate replication (74, 75). N is also able to mediate

translation initiation by acting as a substitute for the eIF4F cap-binding complex (77).

Experiments conducted in Sin Nombre hantavirus showed that although N increased

translation of nonviral RNA, N-mediated translation of viral RNA was even greater. This

suggests that there must be some *cis*-acting signal present in the 5' UTR of viral mRNA

that allows for preferential initiation of viral mRNA. Further investigation revealed that

N binds with high affinity to a sequence present in the 5' untranslated region of viral

mRNA (GUAGUAG) and this sequence is necessary and sufficient for N-mediated

translation initiation (79). Another important function of N during replication is

encapsidation of the viral genome and antigenome. Experiments conducted in La Crosse

virus infected cells provided the first clue that the nucleation site for encapsidation might

be located at the 5' end of the genome. This was proposed because an unusual uncapped

plus-sense S RNA transcript truncated at its 3' end by ~ 100 nucleotides was found

encapsidated in La Crosse virus infected cells (98). This observation implies that the

sequence near the 3' end of the antigenome is not important for ribonucleocapsid

assembly and the highly conserved 5' end of the genome is likely where the

encapsidation signal is located. Support for this hypothesis was later obtained by binding

assays conducted in Jamestown Canyon, Hantaan and Bunyamwera viruses that reported

the first ~40 nucleotide residues at the 5' end of the viral RNA mediate the specificity of

N-RNA binding and contain the encapsidation signal (85, 87, 109). Together, these data

led to a proposed model of RNA encapsidation that involves an initial, specific

interaction occurring at the 5' end of the nascent viral RNA, followed by N-N

interactions that drive the nonspecific binding of the remaining RNA template (Fig. 3).
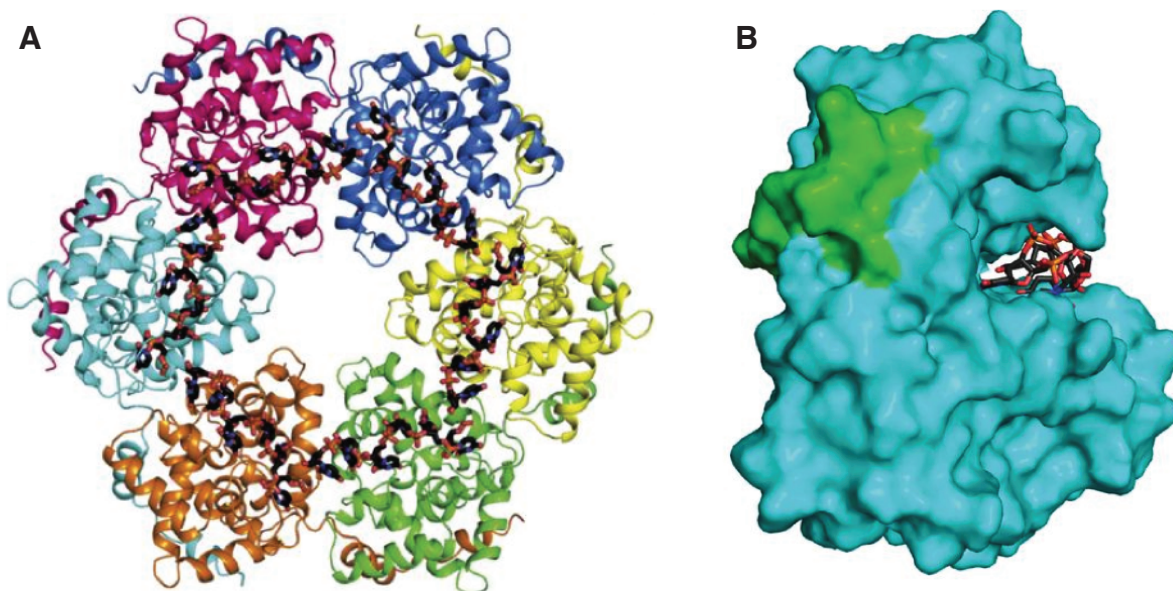
Figure 1 - 4
A) The structure of a RVFV N hexamer bound to RNA. Each individual N subunit is depicted using a different color and the N-terminal arm of each subunit wraps around the outside of the adjacent subunit. Single-stranded RNA (black) bind along the inner groove of the hexameric ring structure. B) The core domain of RVFV N (aqua) bound to RNA (sticks). The area where the helical arm of an adjacent N monomer interacts with the core domain of N is depicted in green (100).

III. Nucleocapsid protein as an anti-viral therapeutic target

Inhibition of RNA binding to Rift Valley fever virus (RVFV) nucleocapsid (N)

protein represents a potentially attractive anti-viral therapeutic strategy because several

essential steps in the RVFV replication cycle involve N binding to viral RNA.

Nucleocapsid-targeting drugs have been shown to be effective and are in various stages

of development for other RNA viruses such as HIV, respiratory syncytial virus and

influenza. Although these drugs all target N, many have been shown to utilize unique

and interesting strategies to exert their inhibitory effect. This is a result of the

multifunctional nature of N during viral infection where each function represents a

potential antiviral strategy. Examples of a variety of inhibitors that disrupt nucleocapsid protein RNA binding interactions have been described in HIV. Some inhibitors promote zinc ejection of the zinc finger RNA binding domain (28, 103) while others disrupt the ability of nucleocapsid to coordinate zinc resulting in protein aggregation and diminished protease processing (55, 71). Inhibitors that target the nucleocapsid chaperone activity for HIV RNA have been described as well (110). A small molecule that inhibits influenza virus replication by causing aberrant oligomerization of viral nucleocapsid has recently been described. The ligand, nucleozin, induces the formation of stable nucleocapsid oligomers that interfere with ribonucleocapsid transport making it unavailable for assembly into polymerase complexes and virions (38). A promising potential drug compound targeted against a respiratory syncytial virus nucleocapsid protein, RSV604, is currently in a phase II clinical trial and Arrow Therapeutics (a London-based pharmaceutical company) plans to develop and commercialize the small molecule inhibitor (18).

IV. Concluding remarks

Despite extensive research that has vastly increased our understanding of the molecular virology of Rift Valley fever virus, several key questions remain. Although the structure of RVFV N bound non-specifically to nucleic acid has been solved how nucleocapsid proteins interact with their cognate RNA is not well understood and the mechanism of the dynamic encapsidation process has not been characterized. Furthermore, since N performs a variety of functions during the virus life cycle, it is likely that N interacts specifically with internal RNA sequences and secondary structural elements of the viral RNA and possibly even cellular RNAs. The affinity of N for RNA

sequences other than the 5' and 3' termini of the viral genome and antigenome has not been explored for bunyaviruses. Finally, it has been demonstrated in several other RNA viruses that N is promising drug target; however, no FDA-approved anti-viral drugs are currently available to treat RVFV infected individuals. The research presented in this dissertation aims to address these underlying questions and issues that persist in the field of Rift Valley fever virus biology. We use biochemical methods to explore the RNA-binding properties of RVFV N and develop a high-throughput assay for identification of compounds that inhibit an N-RNA binding interaction. The efficacy of potential anti-viral drug compounds identified using the high-throughput assay during RVFV infection is also tested.

CHAPTER 2: IDENTIFICATION AND CHARACTERIZATION OF RIFT VALLEY FEVER VIRUS NUCLEOCAPSID PROTEIN-RNA BINDING MOTIF

I. Introduction

Rift Valley fever virus (RVFV) is a mosquito-borne bunyavirus (genus *Phlebovirus*) endemic to sub-Saharan Africa. It is responsible for widespread outbreaks of severe disease such as hepatitis, encephalitis and hemorrhagic fever in humans and abortion storms in livestock. In recent years the virus has spread as far north as Egypt and is now considered endemic in the Arabian Peninsula (6). The ability to cross extensive geographic barriers as well as the presence of competent mosquito vectors in Europe, Asia and the Americas signify the potential for global distribution. Currently there is no specific treatment for RVFV infection. Ribavirin, a nucleoside analog, is one of few drugs approved for treatment of viral hemorrhagic fevers. Although it has been administered to patients during past outbreaks, its use is limited due to adverse side effects (60, 67, 81). T-705 is a pyrazinecarboxamide compound that inhibits viral replication post-infection *in vitro* and efficacy has been demonstrated for several RNA viruses in animal models (40, 41). Despite these advances, T-705 is not licensed and therefore not a feasible treatment option for RVFV infection. RVF is an emerging infectious disease and a major zoonotic threat, and the characteristics described above emphasize the need for therapeutics to treat infected individuals.

RVFV has a tripartite single-stranded negative-sense RNA genome that encodes seven proteins. One of these gene products, the nucleocapsid protein, N, is an RNA binding protein required for the production of viable virus because of its involvement in several stages of viral replication. N protects the viral genome from degradation and prevents the formation of double stranded RNA intermediates during replication and transcription by encapsidating viral genomic and antigenomic RNA (105). N also has

RNA chaperone activity predicted to function during initiation of viral replication by unwinding an RNA helix that sequesters the self-complementary ends of the genome and allowing the polymerase to bind (74, 75). This activity could also remove higher order RNA structures that decrease the speed and efficiency of translation of viral mRNA.

RVFV is a member of the large bunyavirus family (*Bunyaviridae*), which is composed of five genera: *Orthobunavirus, Hantavirus, Phlebovirus* (including RVFV)*, Nairovirus* and *Tospovirus*. A notable feature of all bunyavirus RNA segments is a conserved, self-complementary sequence present at the 5' and 3' ends of the genome and antigenome. This sequence undergoes base pairing to form panhandle-like structures and preferential binding of the termini of viral RNA segments by N has been demonstrated in Sin Nombre hantavirus (SNV) (72, 76, 78). Further analysis of the RNA binding properties of SNV N reveal that it binds with high affinity to UAG repeats located in the 5' untranslated region of viral mRNA. It has been proposed that this sequence is a *cis*-acting signal that enhances the initiation step of translation and subsequently increases the production of viral proteins (79). The first 40 nucleotides present at the 5' end of the viral genome and antigenome have also been shown to contain an RNA binding domain for N protein isolated from several members of the *Bunyaviridae*, such as Jamestown Canyon, Hantaan and Bunyamwera viruses (85, 87, 109). This region is thought to contain a stem-loop structure specifically recognized by N and binding to this site may initiate encapsidation of viral RNA (87, 98). Thus, all of the known essential functions performed by N during an infection involve RNA binding. Furthermore, the multifunctional and viral specific nature of the N protein underscores its potential as an antiviral therapeutic target. Because nearly all RNA viruses, retroviruses, and

hepadnaviruses encode RNA-binding nucleocapsid proteins, disrupting the N-RNA interaction represents a potentially powerful new antiviral therapeutic strategy with broad applicability. In this study, we explore the RNA binding properties of RVFV N using biochemical, biophysical, and combinatorial methods and describe construction of a fluorescent biosensor that reports on N-RNA interactions.

II. Materials and Methods

A. Overexpression and purification of Rift Valley fever virus N protein

Plasmid encoding (His_6)-tagged RVFV N protein was transformed into *E. coli* strain BL21(DE3)pLysS. Transformed cells were grown at 37°C in LB media containing 100 μg/ml ampicillin and 50 μg/ml chloramphenicol until OD_{600}=0.6-0.8. Protein expression was induced by addition of IPTG to a final concentration of 0.5mM and cultures were grown overnight at room temperature (24°C). Cells were harvested by centrifugation and stored at -80°C.

Cells were lysed by resuspension of thawed pellets in BugBuster (Novagen) in the presence of protease inhibitors. Benzonase (Novagen) was added to the lysis buffer at a concentration of 12 units/mL to degrade DNA and RNA. Insoluble cell debris was removed from the lysate by centrifugation at 12,000 RPM for 45 minutes at 4°C. The (His_6)-tagged N was batch purified from the cell lysate by using Ni-charged IMAC resin (Bio-Rad). The resin was washed 5 times with 35mM imidazole/0.5M NaCl/1.5M urea in 50mM Tris-HCl buffer, pH 8.0 and the protein eluted with 300mM imidazole/0.5M NaCl in 50mM Tris-HCl buffer, pH 8.0. The eluate was concentrated and the buffer exchanged to 0.5M NaCl in 50mM Tris-HCl pH 8.0 using Amicon centrifugal filters (10K MWCO). After flash freezing with liquid nitrogen the N protein was stored at -

80°C. The purity of the protein was checked by SDS-PAGE and its concentration was determined by the absorbance at 280 nm using extinction coefficient obtained from the ExPASy website ($33,920$ $M^{-1}cm^{-1}$).

B. Preparation of pool 0 DNA

In order to construct a pool of DNA template with a 30 nucleotide long randomized region flanked by constant regions, MMLV reverse transcriptase (Invitrogen) was used to perform an extension reaction on two chemically synthesized oligos that had complementary 3' ends. The sense primer contained an *Eco*RI site and a T7 promoter near the 5' end followed by a 13 nucleotide long constant region (Fig. 2-1A, EcoT7SX5'). The antisense primer contained a *Xba*I site near the 5' end and a 30 nucleotide randomized cassette flanked by constant regions (Fig. 2-1A, XBbSx30). After extension the DNA was digested with *Eco*RI and *Xba*I to obtain templates of uniform size and purified by phenol/chloroform extraction and ethanol precipitation.

C. Preparation of pool 0 RNA

Pool 0 RNA was generated using pool 0 DNA as a template and a MEGAshortscript Kit (Ambion) (Fig. 2-1B). The transcription reaction was DNase treated and purified by phenol/chloroform extraction and ethanol precipitation. The pool RNA was further purified by denaturing gel electrophoresis. The band corresponding to the RNA was visualized by UV shadowing, excised, eluted into 0.5M ammonium acetate, 1.0mM EDTA pH 8.0 and 0.2% SDS, and ethanol precipitated.

D. Selection of RNA aptamers to Rift Valley fever virus N protein

Pool RNAs were incubated at 30°C with Ni-charged IMAC resin (Bio-Rad) in binding buffer (10mM HEPES pH 7.5, 150mM NaCl and 5mM MgCl$_2$) using a rotisserie in a hybridization oven. After incubation the reaction was placed on ice, the resin pelleted by gravity, and the unbound RNA in the supernatant collected and used for subsequent selections. This "pre-selection" step was necessary to prevent amplification of RNAs that bind with high affinity to Ni-charged resin. Next preselected RNAs were incubated with purified (His$_6$)-tagged N protein that was pre-bound to Ni-charged resin in binding buffer with constant rotation at 30°C. The binding and nonbinding RNAs were partitioned by allowing the resin to settle to the bottom of the microfuge tube on ice. The unbound RNAs in the supernatant were removed and discarded. The Ni resin pellet containing the target RNA-N complexes was washed multiple times with binding buffer to remove weakly bound RNAs and then phenol/chloroform extracted and ethanol precipitated to purify RNAs that bound to N with high affinity.

The selected RNAs were then subjected to RT-PCR to obtain DNA templates for subsequent rounds of selection. The reverse transcription step was performed in a final volume of 20 μL using AMV reverse transcriptase (Promega) according to the manufacturer's protocol. Approximately 200 ng XbaBbs primer (5'-GCT CTA GAA GAC GC-3') and 400 ng RNA template (a 1:2 primer/template ratio was always used) were incubated at 70°C for 3 minutes and snap cooled on ice. Next, AMV RT buffer, dNTP mix (0.6mM of each NTP) and 0.3 μL AMV RT enzyme were added and the reaction incubated at 42°C for 1 hour. The PCR step was performed in a final volume of 50 μL containing 5 μL of the cDNA template, 0.2μM of EcoT7SX5' and XbaBbs

primers, dNTP mix (0.2mM of each dNTP), $MgCl_2$ (1.5mM final), 1X PCR buffer, and

1.25 units of GoTaq Hot Start DNA polymerase. Twenty five cycles of amplification

were used and the annealing temperature varied and was empirically determined for each

round of selection by performing a small scale gradient PCR. The RT- PCR products

were digested with *Eco*RI and *Xba*I and used as templates for T7 transcription to generate

RNA for subsequent rounds of selection.

Fifteen rounds of selection for RNA aptamers to RVFV N were conducted using the

protocol described above (Fig. 2-1C). We gradually increased the stringency of the

selection each round by decreasing the amount of time allotted for RNA-protein

complexes to form. Pool 0 RNAs and N were incubated together for 5 hours whereas

pool 14 and 15 RNAs were incubated with N for less than 1 minute. The final (16th)

round of selection was done by incubating N that had the His-tag removed by tobacco

etch virus (TEV) protease digestion with a mixture of radioactive (200,000 cpm) and

unlabeled (300 ng) pool 15 RNA for less than one minute at room temperature (24°C).

The binding reactions were loaded with 2μL of glycerol loading dye onto a pre-chilled

6% acrylamide/TBE gel and samples migrated at 10V/cm on the gel in the cold room for

75 minutes. The gel was visualized using phosphorimagery (Fuji) and the bands

corresponding to the N-RNA complex were excised from the gel. To obtain round 16

RNA, slices of the excised bands were subjected to RT-PCR followed by T7 transcription

using essentially the same protocol described for previous selections. The only alteration

involved performing the reverse transcription step in a final volume of 50 μL to

synthesize cDNA from a 10 μL equivalent slice of gel. For comparative analysis the

digested RT-PCR products of rounds 1, 8, 15 and 16 were ligated into the *Xba*I and

31

*Eco*RI sites of the pUC18 plasmid and the DNA from individual clones isolated and sequenced.

E. Synthesis and purification of selected pool and clonal RNAs

RNA was synthesized using either clones digested with *Xba*I or pool DNA obtained from digested RT-PCR products as templates for T7 RNA polymerase. MEGAshortscript kits (Ambion) were used to synthesize non-radioactive RNA and MAXIscript kits (Ambion) were used to synthesize RNA in the presence of radioactive α-^{32}P-UTP (800Ci/mmol; PerkinElmer). The transcription reactions were then DNase treated and purified by phenol/chloroform extraction and denaturing gel electrophoresis as described for the synthesis of pool 0 RNA.

F. Nitrocellulose filter binding assays

N was serial diluted to varying concentrations in a binding buffer (10mM HEPES pH 7.3, 150mM NaCl, 20mM KCl, and 5mM $MgCl_2$ final concentration). Nine μL of various dilutions of N were added to each tube along with 1 μL of RNA that had been subjected to denaturation at 90°C for two minutes followed by snap cooling on ice. In each experiment 25,000 cpm/μL of reference RNA (wild type) was added. The cpm/μL of the other RNAs was adjusted based on their uridine content to ensure an equal number of molecules were used. After incubation at 30°C for one hour, reactions were diluted to 100 μL with ice-cold binding buffer and filtered through pre-soaked nitrocellulose filters (Millipore HAWP). The filters were washed twice with 500 μL ice-cold binding buffer, dried, and the radioactivity retained on filters was measured by scintillation counting.

G. Electrophoretic mobility shift assays

Samples for EMSAs were prepared as described for the filter binding assays. After incubation the reactions were placed on ice and 2 μL of glycerol loading dye was added to each tube. Samples were loaded onto pre-chilled 6% acrylamide/1X TBE gels that were run in the cold room at 10V/cm for 75 minutes. Gels were then dried, exposed to a phosphorimager screen, and visualized on a Fuji FLA3000G Image Analyzer.

H. Synthesis of mutant and truncated aptamer RNAs

DNA template that contained the desired mutations to the aptamer sequence was generated by performing a PCR extension reaction on two chemically synthesized oligos that had complementary 3' ends. The PCR step was performed in a final volume of 50 μL containing 0.3μM of sense and antisense oligo, dNTP mix (0.2mM of each dNTP), $MgCl_2$ (1.5mM final), 1XPCR buffer, and 1.25 units of GoTaq Hot Start DNA polymerase. Thirty cycles of amplification with an annealing temperature of 55°C were used. The PCR products were digested with *Eco*RI and *Xba*I and used as templates T7 transcription to generate mutant aptamer RNA. A MAXIscript kit (Ambion) was used to synthesize RNA in the presence of radioactive α-^{32}P-UTP (800Ci/mmol; PerkinElmer) using the same protocol described for synthesis and purification of clonal RNAs.

I. RNA solution structure probing

RNAs were dephosphorylated using Antarctic phosphatase (New England BioLabs). The 5' ends were labeled with γ-^{32}P-ATP (6000Ci/mmol; PerkinElmer) using T4 polynucleotide kinase (Invitrogen) and the RNA purified by denaturing gel electrophoresis. Labeled RNAs were heated in water for two minutes at 90°C followed by snap cooling on ice. Binding buffer (10mM HEPES pH 8, 150mM NaCl, 20mM KCl

and 5mM $MgCl_2$) was added and samples incubated at 37°C for 10 minutes. The folded

RNA (60,000 cpm/rxn) was incubated in binding buffer with 2 μg tRNA and either

RNase T1 (0.25 units/rxn) or RNase U2 (10 units/rxn at 37°C for 10 minutes. A T1

ladder was prepared by incubating the RNA at 50°C in a denaturing buffer (6mM sodium

citrate pH 5, 7M urea and 1 mM EDTA) prior to digestion with RNase T1. An alkaline

ladder was prepared by incubating the RNA at 90°C for 4 minutes in 100mM sodium

carbonate buffer pH 9.0. All samples were ethanol precipitated, resuspended in

formamide loading dye and analyzed by denaturing gel electrophoresis.

J. Fluorescence polarization measurements

A 35 nucleotide long RNA was chemically synthesized, 3' end labeled with FAM

and HPLC purified by TriLink BioTechnologies. The fluorescent label was attached to

the RNA using a six-carbon linker. N was serial diluted to varying concentrations in a

binding buffer (10mM HEPES pH 8, 150mM NaCl, 20mM KCl and 5mM $MgCl_2$). To

determine the equilibrium dissociation constant of N/FAM-RNA complex, various

dilutions of N were added to 10 nM FAM-labeled aptamer RNA that had been subjected

to denaturation at 90°C for two minutes followed by snap cooling on ice. Reactions were

incubated at 30°C for 1 hour. Samples for competitive binding experiments were

prepared by incubating varying concentrations unlabeled competitor RNA (450 pM-5.0

μM) with 4.5 μM N for 40 minutes at 30°C. Next, 10 nM FAM-labeled RNA was added

and the reaction incubated for 40 minutes at 30°C. Fluorescence polarization values were

measured using a VICTOR *X* multilabel plate reader. Binding profiles were plotted and

apparent K_d and K_i values were calculated using GraphPad Prism software.

K. Cell culture studies

Human 293 cells were maintained in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), penicillin and streptomycin. Vero cells were maintained in minimum essential medium (MEM) alpha medium supplemented with 10% FBS, penicillin and streptomycin. The Rift Valley fever virus (RVFV) vaccine strain, MP-12, was provided by Brian Gowen (Utah State University, Logan, UT). A Rift Valley fever reporter virus containing the gene for *Renilla* luciferase was provided by Richard Elliott (University of St. Andrews, St. Andrews, United Kingdom). The *Renilla* luciferase open reading frame was inserted in place of the NSs gene encoded by the S segment of Rift Valley fever virus (RVFV) and is known to not be required for viral infection. During infection, human 293 cells and virus were maintained in DMEM supplemented with 2% FBS, penicillin and streptomycin. Incubations were carried out at 37°C and 5% CO_2 unless otherwise stated.

L. Cell culture and transfection

Confluent monolayers of human 293 cells grown in 24-well plate were transiently cotransfected with 15 pmole RNA/well and 0.5 µg plasmid DNA/well using the Lipofectamine 2000 kit (Life Technologies). The next morning transfected cells were harvested, counted and reseeded in 96-well plate format.

M. Luciferase assays

Confluent monolayers of transfected human 293 cells grown in 96-well plate format were infected with RVF-luc virus using a MOI of three. Cells were harvested at eight hours post infection by removing the growth media, adding 20 µL of lysis buffer and rocking the cells for 15 minutes at room temperature. The cell lysate substrate was

diluted (5-20x) in reagent buffer and 20 µL was transferred to a white 96-well plate. A

Biotek Synergy 2 plate reader was used to autoinject each sample well with 100 µL of

Renilla luciferase assay reagent (Promega) and measure the luminescence.

N. Plaque Assays

Confluent monolayers of transfected human 293 cells grown in 96-well plate format

were infected with RVF MP-12 virus using a MOI of 0.1, centrifuged at 2500 RPM for

100 minutes and incubated for 1 hour. After incubation the media was removed and

replaced with DMEM supplemented with 2% FBS, penicillin and streptomycin.

Supernatants of virus infected cells were harvested at 3 days post infection. The

supernatants were serially diluted in MEM and subsequently used to infect confluent

Vero cells grown in 6-well plate format. After a 2 hour incubation the cells were overlaid

with MEM containing 1% agarose and incubated for 7 days. Cells were fixed with

9.25% formaldehyde and plaques revealed using crystal violet stain.

III. Results

A. In vitro evolution of RNA aptamers to Rift Valley fever virus N protein

An 84 bp segment of DNA that contained a 30 bp random region flanked by constant

sequences was used as a template for T7 RNA polymerase transcription. The template

contained a 5' *Eco*RI site, a T7 promoter, a 13 nucleotide-long constant region followed

by a 30 nucleotide-long randomized region and then another constant region with a 3'

*Xba*I site (Fig. 2-1A). The presence of all 4 nucleotides in the randomized region was

verified by batch sequencing of pool 0 DNA (data not shown) and by sequencing of

individual clones from pool 1 DNA. The recombinant nucleocapsid protein (N) from the

M12 strain, a clinical isolate of RVFV (39), was overexpressed in *E.* coli and purified

using Ni-charged IMAC resin. Next, a selection/amplification technique (SELEX) was used to isolate several families of RNA aptamers that bind with high affinity and specificity to the N protein from a starting pool of approximately 10^{14} small RNAs (Fig. 2-1B). The basic methodology for the *in vitro* selection is outlined in Figure 2-1C and the technique described in detail in Materials and Methods.

The iterative selection step consisted of partitioning the unbound and bound RNA species by immobilizing the N protein onto Ni-charged IMAC resin. This allowed the bound RNAs to be pelleted with the resin beads and the unbound RNAs in the supernatant to be discarded. The selected RNAs were then purified from both N and Ni-resin by phenol/chloroform extraction and amplified using RT-PCR and primers specific to the constant sequences. Sixteen rounds of selection were conducted and negative selections were periodically performed to eliminate Ni-resin binding species and ensure isolation and amplification of N-specific aptamers. During the first rounds of selection N was allowed extended incubation with pool RNAs in order to capture and amplify a small number of functional molecules. In later rounds, the stringency of selection was increased using a greater RNA to protein ratio (100:1) and shorter incubation periods. The final (16th) round of selection was performed using non his-tagged N and the bound and unbound species of RNA were separated by electrophoretic mobility shift. The band on the gel corresponding to high molecular weight RNA/protein complexes was excised, and the RNA eluted from the gel slice and subjected to RT-PCR followed by T7 transcription.

Figure 2 - 1: Construction of random sequence nucleic acid pools and *in vitro* selection methodology.

A. Oligos used to construct pool 0 DNA. The sense oligo (EcoT7SX5') contained an *Eco*RI restriction site, a T7 promoter and a 3' constant sequence that was complementary to the 3' constant sequence on the antisense oligo (XBbsSx30). The antisense oligo contained the 30 nucleotide long randomized region flanked by constant regions and a 5' *Xba*1 restriction site. Shaded boxes highlight constant regions present in both the oligo DNA and RNA. B. Randomized RNA sequence used for the *in vitro* selection. C. Schematic diagram of the technique used to select and amplify RNAs that bound with high affinity to RVFV N. Non-binding and binding RNAs were separated using N immobilized to Ni-charged resin as a target. Therefore, the resin also immobilized RNAs that bound to the N, allowing unbound RNAs to be washed away and selected RNAs were purified and amplified. RT-PCR was used to create more DNA template for transcription, generating more RNA for subsequent rounds of selection. This cycle was repeated 15 times. D. Evolution of RNAs with high affinity to N after repeated rounds of selection. The pools RNA binding affinity for N was analyzed by EMSA. The bands that migrate faster in the gel are unbound RNA species and the slower migrating bands represent N/RNA complexes. RNAs were internally labeled with ^{32}P-UTP and the gels visualized by phosphorimaging.

B. Binding affinity of pool RNAs to N increases with repeated rounds of selection

Binding assays were performed at several intermediate rounds of the selection to assess the affinity of pool RNA sequences for N. Pool RNAs were transcribed from RT-PCR products and internally labeled with ^{32}P-UTP. A constant amount of labeled RNAs (25,000 cpm) was incubated in binding buffer with either 1.2μM or 12μM RVFV N for 1 hour at 30°C and then subjected to non-denaturing gel electrophoresis (Fig. 2-1D). The range of N protein concentrations used in our study (0.6-35μM) was determined empirically. The bands that migrate faster on the gel correspond to the free RNA species and the higher migrating bands represent high molecular weight RNA/protein complexes. Round 0 RNAs exhibited low overall binding affinity to N whereas the binding affinity of round 8 RNAs had increased dramatically. When incubated with the highest concentration of N, most of round 8 and round 11 RNAs were present in the slowest migrating band, indicating that they were able to bind stably to N and that the affinity for N was increasing with subsequent rounds of selection. Although the binding affinity of round 11 RNAs for N was high, several further rounds of selection were carried out to decrease the population complexity.

C. Selected aptamer RNAs have a distinct primary sequence signature

The RT-PCR products from rounds 1, 8, 15 and 16 of the selection were cloned into pUC18 plasmid. Individual clones from rounds 1 and 8 and a total of 100 individual clones from rounds 15 and 16 were sequenced. Analysis of round 1 sequences showed a distribution of 27% A, 19% G, 32 % C, and 23 % T (Table 2-1). Interestingly, when these percentages were compared to the nucleotide distribution of sequences isolated from rounds 8, 15 and 16 of the selection, a progressive enrichment for both guanosine

and thymidine residues was evident. While the enrichment observed for guanosine residues stabilized by round 8, the percentage of thymidine residues continued to increase from 31% at round 8 to 40% in rounds 15 and 16. Also, the percentage of adenosine and cytidine residues was reduced in sequences isolated from later rounds of selection compared to round 1 clones. Analysis of individual sequences from rounds 15 and 16 of the selection showed that 49 out of the 100 sequences occurred with multiple frequencies or were at least 90% similar (Table 2-2). These results suggest that, while identifiable sequence trends were evident throughout the population, several sequence families were beginning to dominate the pool.

Table 2 - 1: Distribution of nucleotides within the RVFV genome and the randomized region of clonal RNAs[a].

Sequence Analyzed	% A	% G	% C	% T
Round 1 RNA	27 ± 9	19 ± 6	32 ± 11	23 ± 9
Round 8 RNA	17 ± 8	29 ± 9	19 ± 5	31 ± 7
Rounds 15&16 RNA	19 ± 6	28 ± 4	15 ± 5	40 ± 6
RVFV S segment	23.6	23.1	25.9	27.4
RVFV M segment	27.3	25.4	20.1	27.3
RVFV L segment	29.2	23.8	20.3	26.7

[a]Sequences corresponding to the viral genome were obtained from the NCBI nucleotide database using the following accession numbers: NC_014395.1, NC_014396.1, and NC_014397.1 that correspond to the S, M, and L segments respectively. Eight sequences from round 1 and round 8 and a total of 100 sequences from rounds 15 and 16 were analyzed. ± deviation from the mean.

D. Characterization of binding affinity between selected aptamer RNAs and N protein

Several clones from rounds 1, 8, and 15 of the selection were chosen to assess the evolution of aptamer affinity (Table 2-2). The filter binding assay was used to compare the binding of individual clonal RNAs (Fig. 2-2A). As expected, the binding affinity of the round 1 RNA (MBE1.3) was very low while the binding affinity of round 15 RNA (MBE15.3) to N was high. Moreover, the persistence of RNAs with low binding affinity to N at round 8 underscored the need for further rounds of selection to cull lower-affinity

binders from the pool. To confirm the high affinity of our aptamers cloned from the last

selection rounds, we tested them in a competitive binding assay against a specific natural

target, the terminal panhandle structure of RVFV genomic RNA.

Table 2 - 2: Alignment of the randomized region (5'-3') of RNA aptamers isolated from the different rounds of selection.

RNA[a]	sequence[b]	Frequency (100 clones)
MBE16.49	-------AGUAAUGCUAUGAUGAUAAUGUGUCCCGGGG--	1
MBE16.41	-------AGUAACGCUAUGAUGAUAAUGUAUCCCGGGG--	1
MBE15.12	----CUGUUUACUGAACUAUGAUACUUU---CUGGGG---	5
MBE16.64	-----AUGUUACUGACUUCUUCUGAUUU---CUAGUGG--	2
MBE15.2	----UGUCGUACUGAU---UGAUGAUUUACUCUGGGG---	5
MBE15.8	----UGUCGUACUGAU---UGAUGAUUUACUCGGGG---	1
MBE16.87	AGCCAGUAUUACUGAU---UGAUGAUU----CUGUGG---	2
MBE16.59	-UGAUGCAGAACUGAU---UGAUUACUU---CUGGGG---	2
MBE15.21	--GCUACU--UCUGAU---UGAUUAUUU---CUCUGGUGG	2
MBE16.50	---AUGUCAUACUGAUUACUGAUGACUU----UGGGG---	2
MBE16.74	---AUGUCAUACUGAUUACUGAUGAAUU----UAGGGG--	1
MBE8.1	--------CUUCUGAUUACUGAUUUCUUUCUCUAGUGG--	ND[c]
MBE16.113	-------AAUGUCGAUUACUGGUUGCUGACUCUGGGG---	1
MBE16.114	-------AAUGUCGAUUACUGAUUGCUGAUUCUGGGG---	2
MBE15.7	-----UGCAUGCUGAUUACUG---AUUUUAUCCUGGGG--	1
MBE15.20	-----UGCAUGUUGAUUACUG---AUUU-AUCCUGGGG--	1
MBE15.14	-----UGCAGACUGAUUACUG---ACUUUCUCUAGUGG--	3
MBE16.40	---AAUGCGGACUGAUUACUG---AUACUCUGGGG-----	5
MBE8.6	--GUGGACGUCCUGCGUUGUG--GGCUUG---UGUGG---	ND
MBE15.3	-------CGUCCCGUAGUGUCGGUACUGAUUGAUGUG---	9
MBE16.57	-------CGUCCCGUAGUGUCGGUACUGAUUAAUGUG---	1
MBE8.3	-------CGCCCCGUAGUGCAUAGUCUGAUUGAUGUG---	ND
MBE16.112	----UGCGAUACUGAAU-AUGAAAGCUUUAGGUGG-----	2
MBE1.4	--ACGAGUGCCAUGCACUACAUUGUCUGUUGA--------	ND
MBE1.3	----CACCGUAACAAAGUUGUAAGACACCAACCA------	ND
Others	Orphan sequence	51

[a]Aptamer RNAs are labeled using two numbers. The first number corresponds to the round the aptamer was first isolated from and the second number is a unique identifier. [b]The constant regions are not shown. Sequences from rounds 15 and 16 that occurred with multiple frequencies or were <90% similar are shown. The GAUU motif is in bold font and the pyrimidine/guanine motif is underlined. [c]Not determined
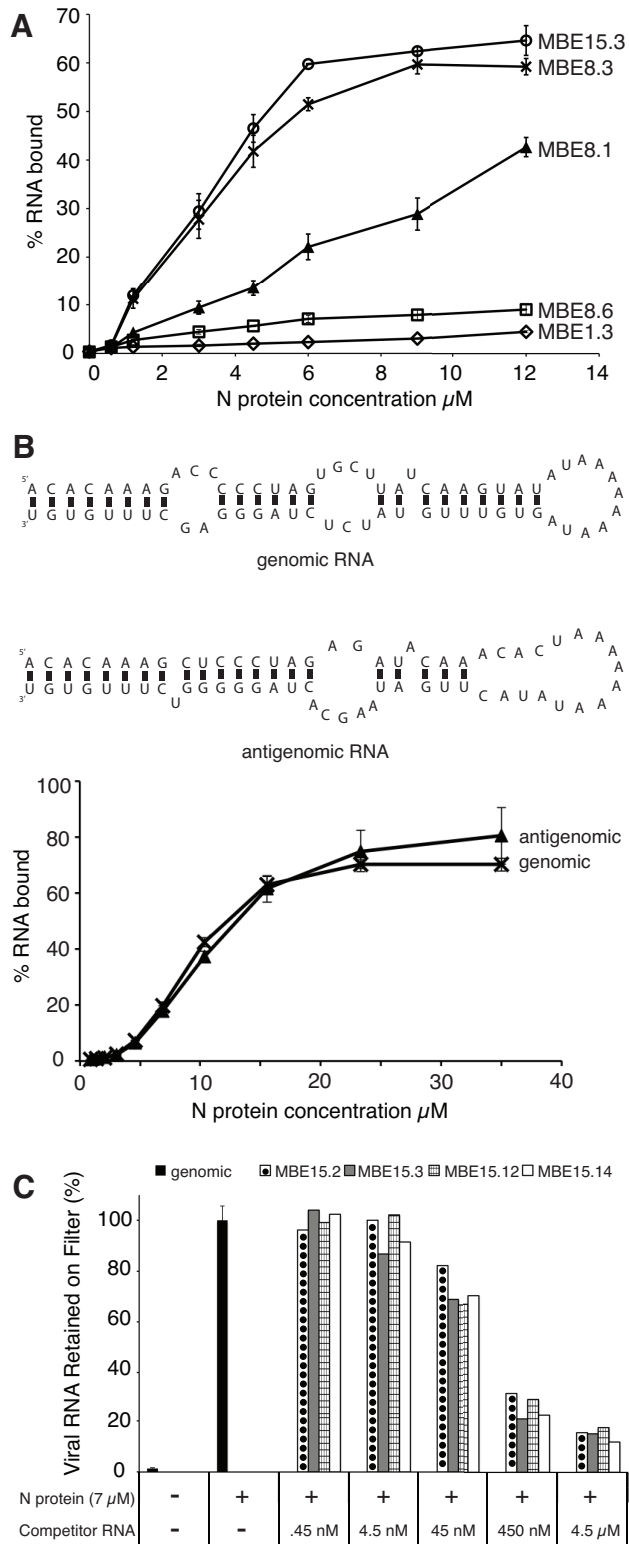
Figure 2 - 2: Characterization of binding interactions between selected RNAs and N.

A. A nitrocellulose filter binding assay was used to compare N binding to clonal RNAs from round 1(MBE1.3, open diamonds), round 8 (MBE8.1, filled triangles; MBE8.3, "X"; MBE8.6, open squares) and

round 15 (MBE15.3, open circles). RNAs were internally labeled with ^{32}P-UTP and radioactivity retained on filters was quantified by liquid scintillation counting. B. Top panel: mfold predicted secondary structure of viral genomic and antigenomic small RNA constructs. Bottom panel: a filter binding assay was used to test the binding ability of genomic ("X") and antigenomic (filled triangles) RNA constructs to the N protein. C. A competitive filter binding assay was used to test the ability of RNAs from the final (15th and 16th) rounds of selection to inhibit binding of viral genomic RNA. The amount of radiolabeled (^{32}P-UTP) viral RNA retained on the nitrocellulose filters in either the absence (black fill) or presence of unlabeled competitor was measured by liquid scintillation counting. The following aptamer RNAs were tested: MBE15.2, circle fill; MBE15.3, gray fill; MBE15.12, grid fill; MBE15.14; white fill.

E. Aptamer RNAs compete with viral RNA sequences to bind N

First, a binding experiment was performed to test the affinity of N for small RNA molecules that mimic the panhandle structures formed by the genomic or antigenomic RVFV RNAs (S segment). The stem-loop structures are formed by the base-pairing interaction of the 5' and 3' ends of one RNA molecule. For practical reasons, the long intervening sequence of the S segment was replaced with a stretch of adenine residues (Fig. 2-2B). The filter binding results showed that N binds well to both small RNA constructs (Fig. 2-2B).

Next, we performed competitive binding experiments using the internally radiolabeled genomic panhandle RNA construct and unlabeled aptamer RNAs. A constant amount of N (7μM) was incubated with varying amounts of unlabeled aptamer RNA (0.45 nM – 4.5 μM for 1 hour at 30°C. Labeled genomic RNA (100,000 cpm/rxn) was added and the binding reaction incubated for 5 more hours at 30°C. Binding of labeled RNA was then assessed by filter binding assay. The results from the competition filter binding assay showed that when increasing amounts of unlabeled aptamer RNA were present during the binding reaction, less genomic panhandle RNA was retained on the filter (Fig. 2-2C). Clearly, the presence of aptamer RNA inhibited the formation of genomic RNA-N complexes, suggesting the aptamer RNA binds to the same region on N as the viral panhandle.

F. Two conserved motifs are found in late-round aptamers

We then checked for the presence of conserved motifs among the high-affinity aptamers of rounds 15 and 16 using primary sequence alignment and visual inspection. The results are summarized in Table 2-2. A GAUU-containing motif was identified in 82% of the 49 sequences isolated with high frequency. This motif was variable in the number of GAUU repeats per RNA and the exact sequence observed. For example, MBE16.74 had a GAUUACUGAUGAAUUU motif while sequencing of another isolate (MBE16.87) revealed a GAUUGAUGAUU motif (Table 2-2). Despite these minor differences, the prevalence of variations of the GAUU motif in a high percentage of the RNAs isolated indicates that it may be an important recognition sequence for N. Also, a pyrimidine/guanine motif we call the U/G-rich motif, located near the 3' end of the random region, was present in many of the sequences isolated from rounds 15 and 16 (Table 2-2). Interestingly, alignment of aptamer sequences with the RVFV genome and antigenome using BLAST revealed that the U/G-rich motif UUUCUGGGGGCG isolated 11 times in the selected RNAs was also present in the RVFV L antigenome (Table 2-3). The evolution of these two motifs could be responsible for the observed enrichment of guanosine and thymidine residues in sequences isolated from rounds 8, 15 and 16. There was a strong correlation between the emergence of guanosine and thymidine rich primary sequences during the selection and the evolution of RNAs that bind with high affinity to N. Together, these data suggest that N may preferentially bind to GAUU and/or U/G-rich sequences.

44

G. Directed mutagenesis of GAUU and U/G-rich motifs in aptamer RNAs

Mutant aptamer RNA constructs were generated to test the hypothesis that N specifically recognizes and binds to GAUU and U/G-rich motifs on selected RNAs. Two variants of the GAUUGAUGAUU motif, either obtained by scrambling the sequence or by introducing complementary nucleotides in place of the U/G-rich motif, were tested in the context of the MBE15.8 aptamer (Fig. 2-3A). The binding affinity of MBE15.12 RNA with similar disruptions to its U/G-rich motif (UUUCUGGGG) (Fig. 2-3B) was also tested.

Filter binding assays were used to compare the affinity of mutant and wild type aptamer RNAs to N. Scrambling one GAUU repeat in the MBE15.8 sequence caused a one-third decrease in the level of RNA bound at 12 μM N. Complementary mutation of the first 7 nucleotides of the GAUU motif caused an even more dramatic decrease in binding. The binding affinity of the complementary MBE15.8 mutant (15.8 COMP mutant) to N was even lower than the affinity observed for a round 1 RNA (Fig. 2-3A). Changing the U/G-rich motif from MBE15.12 RNA from UUUCUGGGG to UGUAAUCGU caused a slight decrease in the binding affinity to N compared to the affinity of the original sequence. Unexpectedly, an RNA that contained a complementary mutation of the U/G-rich motif (15.12 COMP mutant) exhibited better binding than the original MBE15.12 aptamer (Fig. 2-3B). Interestingly, closer examination of this mutant revealed that it contained an 11 nucleotide sequence identical to the 5' end of the S segment of viral genomic RNA (Table 2-3).

These data show that the GAUU motif and, to a lesser extent, the U/G motif, are important for N binding to aptamer RNA and they likely facilitate interactions between N

and other RNA molecules as well. The location of the motifs and the binding affinities of

the RNA aptamers were then used to design truncated aptamers of diminished size.

Table 2 - 3: Regions of sequence similarity between RNA aptamers and the RVFV genome and antigenome identified using BLAST.

Subject	Sequence identified by BLAST		Aptamer	Frequency in aptamer pool (100 clones)

S genome	`Aptamer` `ACAAAGACCCC`
	`            |||||||||||`
	`Sbjct    3 ACAAAGACCCC 13`
			MBE12COMP mutant	1

S anti-genome	`Aptamer` `AUGAUUGUUCUGG`
	`            |||||||||||||`
	`Sbjct  207 AUGAUUGUUCUGG 219`
			MBE35	1

M genome	`Aptamer` `UCCAAAUGACUA`
	`            ||||||||||||`
	`Sbjct  770 UCCAAAUGACUA 781`
			MBE101	1

M genome	`Aptamer` `CAGAACUGAUU`
	`            |||||||||||`
	`Sbjct 2095 CAGAACUGAUU 2105`
			2xMBE59	2

L genome	`Aptamer` `GUUACUGACUUU`
	`            ||||||||||||`
	`Sbjct 1789 GUUACUGACUUU 1800`
			3xMBE14, 2xMBE54, MBE81	6

L genome	`Aptamer` `UUCUGAAUUUU`
	`            |||||||||||`
	`Sbjct 3585 UUCUGAAUUUU 3595`
			MBE1	1

L genome	`Aptamer` `CAGAACUGAUU`
	`            |||||||||||`
	`Sbjct 3932 CAGAACUGAUU 3942`
			2xMBE59	2

L anti-genome	`Aptamer` `UUCUGGGGCGUCUUC`
	`            ||||||| ||||||||`
	`Sbjct 1894 UUCUGGAGCGUCUUC 1908`
			MBE35	1

L anti-genome	`Aptamer` `UUUCUGGGGGCG`
	`            ||||||||||||`
	`Sbjct 4536 UUUCUGGGGGCG 4547`
			5xMBE12, 2xMBE29, 2xMBE107, MBE10, MBE45	11

H. Identification of truncated RNA aptamers that bind with high affinity to RVFV N

Starting with the sequences of several aptamers described above, a series of truncated

RNA aptamers was generated in order to find one or more small RNAs that could

eventually be chemically synthesized and fluorescently labeled. We chose to truncate

RNAs that were either highly represented in the final round of selection, that contained

GAUU or U/G-rich sequence motifs, or contained sequences that were also present in

viral RNA (Fig. 2-4). During the truncation process we generally sought to preserve

important sequence motifs and predicted secondary structure. A combination of RNase T1 and U2 structure probing and mfold were used to determine the conformation of several aptamer RNAs. Overall, results obtained from the structure probing experiments provide support for the mfold predicted structures. However, one common exception was that the mfold algorithm often predicted base pairing interactions at the apical stems of the aptamer RNAs. The RNase T1 and U2 structure probing data showed that guanosine and adenosine nucleotides at the apical stems were reactive (single-stranded) (Fig. 2-4A, B and C). All three RNAs presented here exhibit a strikingly similar stem loop structure that may serve as a binding site for N. Our results indicate that preservation of this structure is necessary for truncated RNAs to retain their ability to bind N (Fig. 2-4D, E and F).

MBE15.8 and MBE16.40 contained a GAUU motif that was shown above to be a requirement for N to bind MBE15.8 RNA. When these aptamers were truncated and the stem loop structure containing the GAUU motif was preserved, both retained their ability to bind N (Fig. 2-4D and F). Conversely, a truncation of MBE16.41 that disrupted the pyrimidine portion of U/G-rich motif caused the RNA to completely lose its affinity for N (Table 2-2; 7% input RNA bound at 12.5 μM). These results further support our hypothesis that the GAUU and U/G-rich motifs are important components of N-RNA recognition.

A truncated version of MBE15.14 that contained an 11 nucleotide-long sequence of RNA present in the L segment of viral genomic RNA bound N with high affinity (62% of input RNA bound at 12.5 μM). This U-rich sequence was isolated a total of 6 times from the final pool of RNA and was also found in two aptamers unrelated to MBE15.14 (Table

2-3). The 15.12 COMP mutant aptamer bound N better than the wild type aptamer

MBE15.12 (Fig 2-3B). MBE15.12 RNA was of interest because it contained a U/G-rich

motif also found in the L segment of viral antigenomic RNA (Table 2-3). Interestingly,

the complement of the U/G-rich motif is a sequence found at the 5' end of the S segment

of viral genomic RNA. The truncated version of 15.12 COMP mutant (15.12 TRNK

COMP mutant) retained this viral sequence, bound N with high affinity and was

predicted by RNA structure probing experiments and mfold to be a stem loop structure

(Fig 2-4E). Mutation of the large 15 nucleotide long loop in 15.12 TRNK COMP mutant

to create RNAs with smaller loops (TRNK loop A and TRNK loop B) did not affect

binding affinity to N (data not shown). TRNK loop A and TRNK loop B RNAs retained

all and most of the viral RNA sequence, respectively, which suggests that this sequence

may be important for N recognizing and binding to both viral and aptamer RNA.
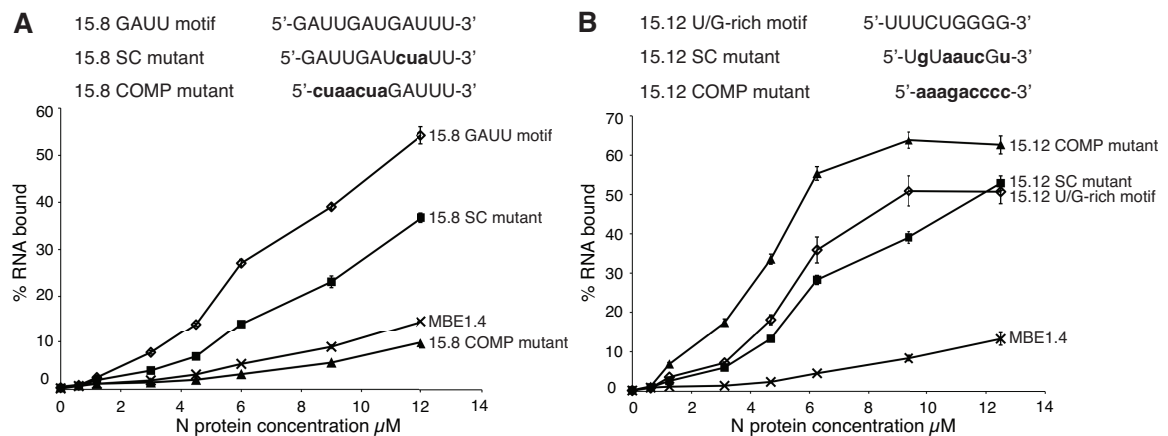
Figure 2 - 3: Characterization of binding affinity of mutant aptamer RNAs to N.
A nitrocellulose filter binding assay was used to compare N binding to wild type and mutant aptamer RNAs where either the GAUU or U/G-rich primary sequence motifs were disrupted. RNAs were internally labeled with ^{32}P-UTP and radioactivity retained on filters was quantified by liquid scintillation counting. A. Top panel: MBE15.8 aptamer RNA GAUU motif and mutant constructs. Changes to sequence motif are in bold and lowercase font. Bottom panel: Filter binding assay of wild type MBE15.8 aptamer, mutant constructs and round 1 MBE1.4 RNA. MBE15.8 wild type RNA; open diamonds, MBE15.8 SC mutant; filled squares, MBE15.8 COMP mutant; filled triangles, and round 1 MBE1.4 RNA; "X". B. Top panel: MBE15.12 aptamer RNA U/G rich motif and mutant constructs. Changes to sequence motif are in bold and lowercase font. Bottom panel: Filter binding assay of wild type MBE15.12 aptamer, mutant constructs

and round 1 MBE1.4 RNA. MBE15.12 wild type RNA; open diamonds, MBE15.12 SC mutant; filled squares, MBE15.12 COMP mutant; filled triangles, round 1 MBE1.4 RNA; "X".

A

ladder
Ø OH- T1 T1 U2

G46-
G41

A33
G32
A30
G29
A26
G25
A22

B

ladder
T1 T1 OH- Ø

G48
G46
G40

G32

G25

G18

ladder
Ø OH- T1 U2

A35
A33
A30
A27

A22

C

ladder
T1 T1 OH- Ø

G47
G45-
G41

G33

G26

G22
G21
G19

ladder
Ø OH- T1 U2

A36
A34

A30
A27

A23

D

MBE15.8 (+)
TRNK 15.8 (++)

E

15.12 COMP mutant (+++)
15.12 TRNK COMP mutant (+++)

F
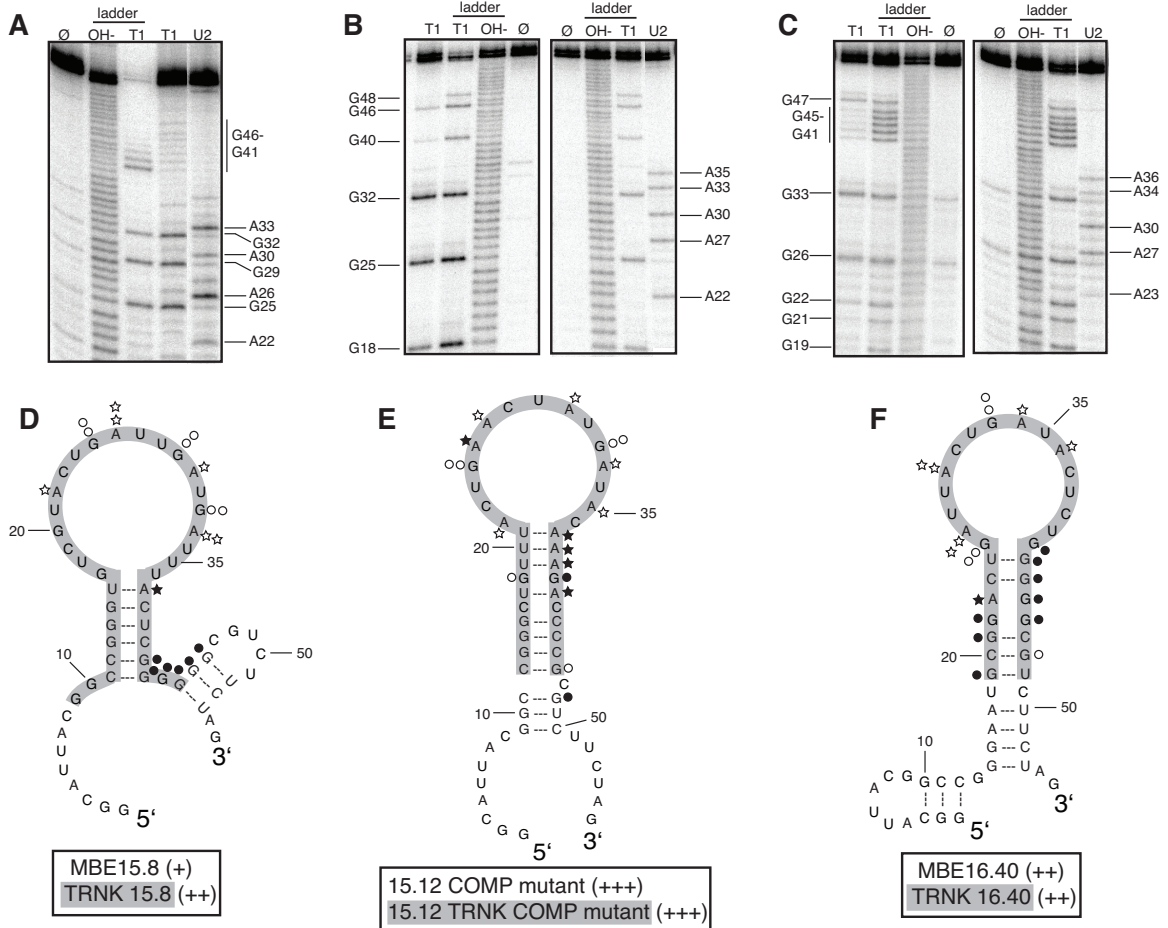
MBE16.40 (++)
TRNK 16.40 (++)

Figure 2 - 4: Secondary structure model and binding affinity of full length and truncated RNA aptamers to RVFV N.

A. to C. Structure probing of full length RNA aptamers MBE15.8 (A), 15.12 COMP mutant (B) and MBE16.40 (C) was conducted to determine the reactivity of guanosine (lane T1) and adenosine (lane U2) residues to enzymes RNase T1 and RNase U2. Lane Ø represents undigested RNA. T1 and OH⁻ ladders are used as RNA sequencing lanes. D. to F. Summary of nucleotide reactivity superimposed on secondary structure models of MBE 15.8 (D) 15.12COMP mutant (E) MBE16.40 (F). Circles represent results obtained from experiments conducted with the single-strand guanosine specific T1 nuclease and stars are representative of the single-strand adenosine specific U2 nuclease. The greater the number of symbols associated with a specific nucleotide, the stronger the reactivity. Open symbols indicate a lack of reactivity. Truncated aptamers are highlighted in gray. Aptamers were rated based on percent of input RNA bound at 12.5 µM N: 40-49% = (+), 50-59% = (++), and 60-69% = (+++).
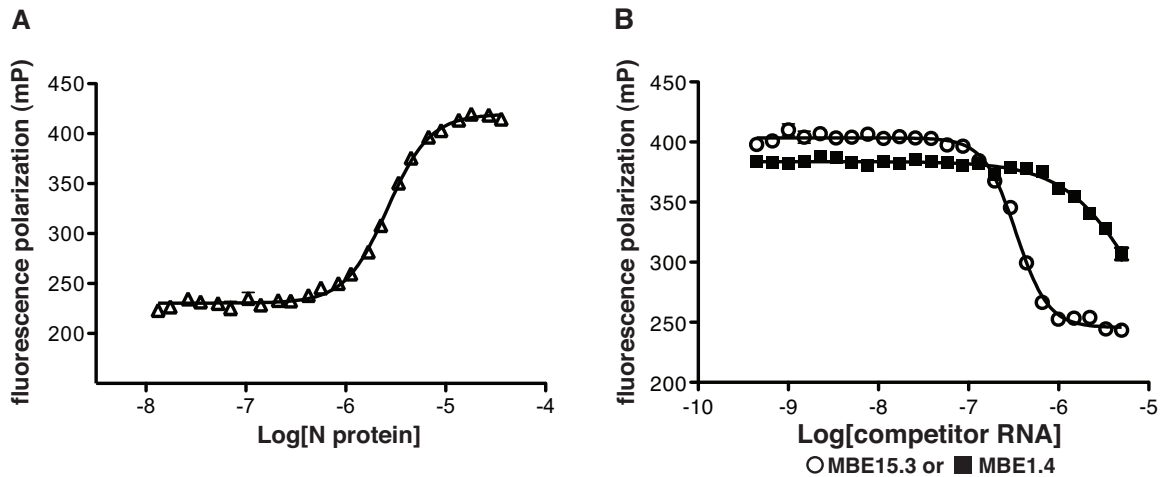
Figure 2 - 5: Analysis of aptamer RNA-N binding interactions using fluorescence polarization (FP).
A. Binding profile for the association of N with fluorescently labeled aptamer RNA (open triangles). B. A competition assay where a fixed concentration of N was incubated with varying concentrations of unlabeled competitor RNAs (round 1 MBE1.4 RNA; filled squares, MBE15.3 aptamer RNA; open circles). Fluorescently labeled aptamer RNA was added and the FP signal is plotted versus competitor RNA concentration.

I.   Fluorescence polarization detection of N-aptamer binding

A truncated RNA aptamer (15.12 TRNK COMP mutant, Fig. 2-4E) was chemically

synthesized and 3' end-labeled with fluorescein (TriLink BioTechnologies). Fluorescent

RNA was incubated in binding buffer with varying concentrations of RVFV N at 30°C

for 1 hour and fluorescence polarization (FP) measured. The results showed that the FP

value was low (220 mP) when the label was attached to RNA alone and increased to 420

mP when the RNA was incubated with increasing concentrations of N (Fig. 2-5A).

Titration of N with the 3'-FAM-labeled RNA aptamer gave an apparent K_d of 2.6 μM.

Next we performed competitive binding experiments by incubating a constant amount of

N (4.5 μM) with varying concentrations (450 pM-5.0 μM) unlabeled competitor RNA at

30°C for 40 minutes. Labeled aptamer RNA was added (10 nM) and the reaction

incubated for 40 more minutes at 30°C. Fluorescence polarization measurements were

taken and the results showed unlabeled aptamer RNA competed for N binding with

labeled aptamer RNA. Competition experiments were conducted with four different aptamers and the apparent K_i values were all in the ~200 nM range. A nonbinding RNA (MBE1.4 RNA) was unable to effectively inhibit the labeled RNA from binding N (Fig 2-5B). These data demonstrate that fluorescence polarization can be used to report binding of fluorescently labeled aptamer RNA to N as a function of competitor binding and that this approach could find utility as a biosensor or drug-screening tool.

J. Selected RNAs inhibit viral replication and production of viral protein in human cell culture

To determine if RNAs selected to bind with high affinity to RVFV N protein could inhibit viral replication and/or production of viral protein during an infection, we cotransfected cells with a green fluorescent protein (GFP) reporter plasmid and either non-binding RNAs (pool 0), RNAs that have undergone one round of the *in vitro* selection process (pool 1 and Round 1-4) or an aptamer RNA that has undergone 16 rounds of selection and binds N with high affinity *in vitro* (MBE87). The GFP reporter plasmid was used to verify that equal and high levels of transfection efficiency were achieved. After transfection in 24-well plate format cells were harvested, counted and reseeded in two 96-well plates. To test the ability of selected RNAs to inhibit viral replication over the course of several days, a 96 well plate of transfected cells was infected with the MP-12 strain of RVFV using a low MOI (0.1) and the number of infectious viral particles produced at three days post infection was measured using a plaque assay. The results show that RNAs that experienced either one or 15 rounds of selection to bind N were able to significantly inhibit virus replication during a long term (three day) infection (Fig. 2-6A).

51

In parallel, to determine if selected RNAs could inhibit production of viral protein during the early stages of an infection, the second 96-well plate of transfected cells was infected with a RVF reporter virus that contains the gene for *Renilla* luciferase (RVF-luc) using a MOI of 3. Cells were harvested at 8 hours post infection and the amount of protein produced by RVF-luc virus was quantified by adding a substrate to each sample that when oxidized by *Renilla* luciferase produces light. By measuring the luminescence of each sample using a Biotek Synergy 2 plate reader the amount of protein produced by RVF-luc virus in the presence or absence of a selected RNA can be measured. The results show that transfection of non-binding RNA into cells causes a slight (approximately 2 fold) decrease in the production of viral protein when compared to cells that were transfected with only plasmid DNA. However, transfection of RNA that has been selected once or multiple times to bind N causes an order of magnitude decrease in viral protein production (Fig 2-6B).

To verify that selected RNAs specifically inhibit production of viral (and not cellular) protein, cells were cotransfected with a plasmid containing the gene for *Renilla* luciferase (phRL-CMV; Promega) and either GFP, non-binding or selected RNA. After transfection in 24-well plate format, cells were harvested, reseeded in 96-well plate format and infected with the MP-12 strain of RVFV using a MOI of 3. At 8 hours post infection, the amount of protein translated by the host cell was quantified by adding substrate to lysed cells and measuring luminescence (as described above). The results show that the presence of non-binding RNA and RNA that has experienced one round of selection have little to no effect on host cell translation. MBE 87 RNA that has undergone 15 rounds of selection and binds N with high affinity causes an approximately

2-fold decrease in host cell protein synthesis. However, since the decrease in translation

of viral protein is significantly greater than the decrease host cell translation, these data

suggest that the selected RNAs are specifically targeted to and inhibiting RVFV (Fig. 2-

6C).
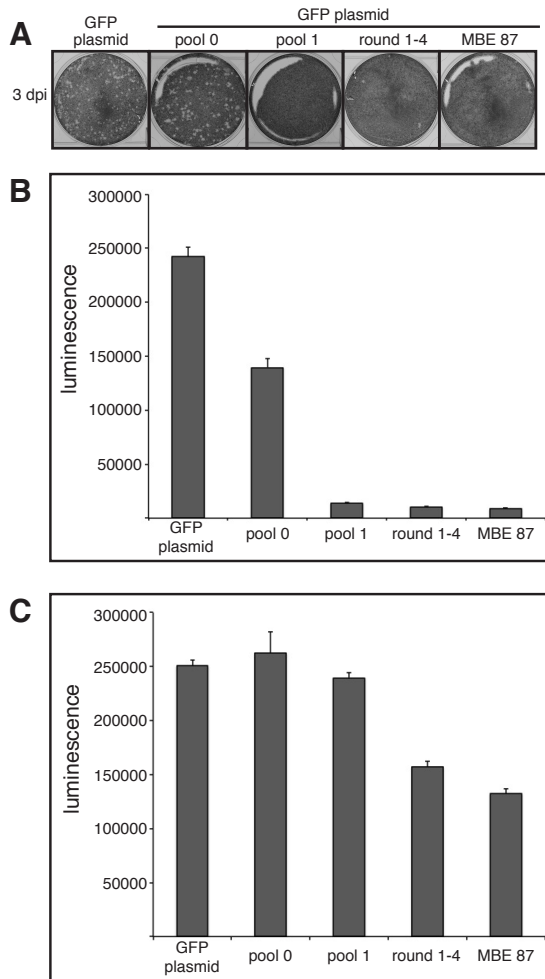
Figure 2 - 6: Analysis of the ability of selected RNA to inhibit RVFV protein production and replication in cell culture.
A) Plaque assay to determine the number of infectious viral particles produced from cells transfected with either plasmid DNA alone (GFP plasmid), non-binding RNA (pool 0), RNA that has undergone one round of selection (pool 1 and round 1-4) or aptamer RNA (MBE87). B) Luciferase assay showing the effects of selected RNAs on the production of viral protein. Cells were transfected with either plasmid DNA alone (GFP plasmid), non-binding RNA (pool 0), RNA that has undergone one round of selection (pool 1 and round 1-4) or aptamer RNA (MBE87). Transfected cells were infected with a luciferase expressing RVFV at a MOI of 3. At 8 hours post infection cells were harvested, lysed and protein quantitated by measureing luminescence using a microplate reader. C) Luciferase assay showing the effects of selected RNA on the ability of the host cell to produce protein. Cells were cotransfected with a plasmid containing the gene for *Renilla* luciferase (phRL-CMV; Promega) and either plasmid DNA alone (GFP plasmid), non-binding

RNA (pool 0), RNA that has undergone one round of selection (pool 1 and round 1-4) or aptamer RNA (MBE87). Transfected cells were infected with the MP-12 strain of RVFV at a MOI of 3. At 8 hours post infection cells were harvested, lysed and protein quantitated by measuring luminescence using a microplate reader.

IV. Discussion

All known functions for N during viral replication involve binding RNA in either a non-specific or specific manner. On the one hand, N possesses sequence-independent binding properties that allow it to form N-RNA complexes along the entire length of viral RNA. On the other hand, specific recognition of cognate RNA by N has been observed in a number of viruses related to Rift Valley fever virus. In the present study, we characterize specific N-RNA interactions for Rift Valley fever virus.

RNA constructs that resemble the panhandle of RVFV viral RNA S segment bind with medium to high affinity to N. These results corroborate studies in Sin Nombre virus (SNV), a member of the *Hantavirus* genus, that show N specifically binds to a panhandle structure formed by the self-complementary 5' and 3' ends of the viral genome (72, 76, 78). The triplet repeat UAGUAGUAG is required for N to bind SNV mRNA and this sequence is also present at the 5' end of all three segments of the SNV genome and antigenome (79). Together these data show that the binding specificity of N is influenced by both sequence and structural features of RNA. N from several other species of bunyavirus are known to preferentially bind to the 5' end of the viral genome and antigenome (85, 87, 109). This binding event is thought to be an important signal that causes N to switch from a specific to a nonspecific mode of binding and initiates nucleocapsid assembly (98). Binding outside this region was also observed in Jamestown Canyon virus (*Orthobunyavirus* genus) but the sequences or secondary structures involved in this N-RNA interaction were not specifically identified (85).

To further characterize N-RNA binding motifs, we used an unbiased *in vitro*

evolution approach to search for RNA ligands of N without any *a priori* assumptions of

sequence or structural constraints. A library of randomized RNA molecules was

incubated with N and the RNA species that bound N with high affinity were isolated,

amplified and used as substrates for the next round of selection. Sequence analysis of

100 individual RNAs isolated from the final rounds of the *in vitro* selection showed

enrichment of guanosine and thymidine residues compared to sequences analyzed from

earlier rounds (Table 2-1). Identification of regions of similarity between aptamer and

viral RNA sequences using BLAST revealed the presence of motifs located in different

regions of all three viral RNA segments. The most frequent similarity is represented by

an U/G-rich motif in the RVFV L antigenome, found in 11 out of the 100 sequenced

aptamers (Table 2-3). Surprisingly, an RNA that contained a complementary mutation of

the U/G-rich motif bound with even higher affinity to N than the wild type aptamer (Fig

2-3B). These data suggest that N is able to recognize both a specific sequence and its

complement. The same phenomenon has been observed in Jamestown Canyon virus,

where it was shown that N binds with high affinity a short segment of ssRNA that

corresponds to the 5' end of the viral genome, as well as its complement (85).

Several GAUU rich aptamer sequences aligned with sequences present on the M

and L segments of the RVFV genome and the S antigenome (Table 2-3). Mutation of the

GAUUGAUGAUU motif in the MBE15.8 aptamer showed that preservation of this

sequence is a requirement for N binding (Fig. 2-3A). Although this sequence has not yet

been shown to be important for N-RNA recognition in other bunyaviruses, the bacterial

chaperone protein Hfq preferentially binds AU rich RNA (27, 108). It is possible that

like SNV, RVFV N has chaperone activity. If so, binding and unwinding GAUU rich stem loop structures in viral mRNA could increase the rate of production of viral proteins. Indeed, chaperone proteins have been shown to be essential for some RNA viruses, such as phage Qβ (33, 34, 80).

Interestingly, the novel N-RNA binding motifs that we identified in aptamer sequences map to coding regions of the different viral RNA segments (Tables 2-2 and 2-3). Further analysis of these data could potentially shed light on N-RNA interactions that are important for N to function during other phases of the viral replication cycle. For example, interaction with host translational machinery, specifically ribosomal protein S19, suggests a role for N in translation of viral mRNA (19, 46). Also, accumulation of N is required to trigger the switch from transcription to full-length cRNA synthesis (106). It is possible that during these processes N interacts with non-terminal viral RNA sequences. Future work will involve determining if the N binding motifs identified in this study are important for regulation of viral replication.

Information obtained from sequence analysis of the RNA aptamers was also used to make a biosensor that reports on N-RNA binding. We chose the 15.12 TRNK COMP aptamer as the RNA sensor for two reasons. First, it was one of the RNA constructs that retained a high affinity for the N protein upon truncation (Fig 2-4E). Second, analysis of sequence similarities revealed that its 3' end contained an 11-nucleotide motif present at the 5' terminus of the S segment of RVFV RNA (Table 2-3). In fact, the first 6 nucleotides of this motif are highly conserved, as they are present at the 5' end of all segments of the viral RNA genome and antigenome. Significant changes in fluorescence polarization (FP) of a labeled 15.12 TRNK COMP RNA were detected as a result of N

binding and this change in FP could be inhibited by addition of an unlabeled RNA that also binds N with high affinity. Aptamer RNAs could thus be used as a sensitive fluorescence-based sensor of RVFV N binding with potential applications for drug screening and imaging methodologies.

Several lines of evidence suggest that targeting nucleocapsid protein is a viable therapeutic strategy. Experiments conducted in La Crosse virus (genus *Orthobunyavirus*) show that interferon-induced cellular MxA protein inhibits viral genome production by binding and depleting the infected cell of N. This binding interaction causes N to be sequestered in cytoplasmic inclusion bodies and inaccessible during viral replication (61). Furthermore, RNA aptamers against HIV-1 Gag polyprotein have been shown to cause inhibition of viral replication in cell culture. Some of these aptamers function by targeting the nucleocapsid (NC) component of the polyprotein and competing with viral genomic RNA for binding (99). Finally, the characterization of several novel antiviral agents that target HIV-1 and respiratory syncytial virus nucleocapsid proteins (18, 102) underscore the potential utility of understanding and exploiting N-RNA interactions for antiviral strategies.

CHAPTER 3: IDENTIFICATION OF RIFT VALLEY FEVER VIRUS NUCLEOCAPSID PROTEIN-RNA BINDING INHIBITORS USING A HIGH-THROUGHPUT SCREENING ASSAY

I. Introduction

Rift Valley fever virus (RVFV) is a mosquito-transmitted bunyavirus that causes disease ranging from a mild influenza-like illness to hemorrhagic fever in humans. Domestic livestock such as sheep, goats and cattle are also susceptible to RVFV infection. The mortality rate for young and newborn animals is close to 100% and the virus causes abortion in pregnant livestock (10). Although historically confined to sub-Saharan Africa, RVFV has spread north to Egypt and was discovered outside of Africa for the first time in 2000. Disease symptoms can be particularly severe when the virus encounters a naïve host and outbreaks along the Nile River in Egypt and the Arabian Peninsula resulted in several devastating epidemics (6, 15, 69, 119). Global climate change, international trade and the ability of RVFV to infect a broad range of mosquito species demonstrate the possibility that RVFV will continue to disseminate geographically. Currently, treatment options for RVFV infected individuals and livestock are extremely limited. Ribavirin has been used to treat patients during past outbreaks but its use is limited due to undesirable side effects and concern related to the potential to cause birth defects (11, 111). T-705 (favipiravir), a promising new broad-spectrum inhibitor of viral hemorrhagic fever, is able to inhibit bunyavirus replication in experimentally infected hamsters when administered 24 hours post infection (41). Studies with influenza virus showed that T-705 metabolite inhibits the viral RNA-dependent RNA polymerase (36). However, the efficacy of the drug in RVFV infected humans or livestock has not yet been determined. The paucity of drugs currently licensed and available to treat viral hemorrhagic fever infection coupled with the ability of RNA

59

viruses to mutate and develop resistance to drugs emphasize the need for continued identification of antiviral compounds.

RVFV has a tripartite single-stranded negative-sense RNA genome that encodes seven proteins. One of these gene products, the nucleocapsid (N) protein, is a 26kDa RNA binding protein that plays an important role during several stages of the viral replication cycle. N helps to package the RNA genome into the budding virus particle and regulates transcription and translation of viral mRNA and protein (79, 89, 115). To discriminate between different species of RNA in the host cell, N recognizes and binds to specific sequences and secondary structures of viral RNA (76, 78). After an initial specific binding event, N switches to a non-specific mode of binding and forms polymers along the entire length of the viral genome or antigenome. This encapsidation process protects the viral RNA and prevents the formation of double stranded RNA during replication, which would activate the host anti-viral response. Inhibition of RNA binding to the essential viral nucleocapsid protein represents an attractive anti-viral therapeutic strategy because all of the functions performed by N during an infection involve RNA binding. Furthermore, significant progress has been made toward exploiting nucleocapsid protein as a drug target in HIV, respiratory syncytial virus and influenza (18, 38, 99).

To target RNA binding activity of N, we developed a fluorescence polarization-based high-throughput drug screening (HTS) assay to test chemical compounds for their ability to disrupt an N-RNA complex *in vitro*. From libraries of FDA approved drugs, drug-like molecules, and natural product extracts, 26,424 compounds were screened and we identified 79 compounds that are inhibitors of the N-RNA interaction. We then

retested the 79 hit compounds to eliminate false positives and used fluorescence

polarization and filter binding assays to determine the IC_{50} of lead compounds.

II. Materials and Methods

A. Compound libraries

The chemical libraries screened are a subset of either commercially available small

molecule or natural product extract libraries maintained at the NERCE National

Screening Laboratory (NSRB), located in the ICCB Longwood screening facility

(Harvard Medical School, Boston MA). Compounds are stored desiccated at -20˚C in

384-well plates at comparable concentration of either 3.3 to 10 mM or 2 to 15 mg/ml in

DMSO. Since the 23rd and 24th columns are empty for all compound plates we tested,

these columns were used for screening controls.

B. Fluorescent RNA aptamer to RVFV nucleocapsid N protein

The sequence of the fluorescent RNA aptamer used in the screening was obtained in

our laboratory by *in vitro* selection (30). Briefly, a pool of randomized RNA molecules

was positively selected for efficient binding to RVFV N protein using a SELEX

procedure. After 16 rounds of selection, sequencing of individual clones revealed several

families of RNA aptamers. Representative clones of each family were then either

terminally truncated or mutated by substitution of internal motifs. The best and smallest

N aptamer resulting from these studies was the 15.12 TRNK COMP clone (30). For that

reason, this aptamer was chosen to be the one used during HTS. A 35 nucleotide long

15.12 TRNK COMP RNA was chemically synthesized, 3' end labeled with FAM and

HPLC purified (TriLink Biotechnologies, San Diego, CA). The fluorescent label was

attached to the RNA using a six-carbon linker. The fluorescent RNA was stored at -80°C at a concentration of 86 µM in water.

C. Overexpression and purification of RVFV nucleocapsid N protein

The nucleocapsid N protein sequence used in this study is derived from the experimental live attenuated RVF virus vaccine strain (MP-12; (16)), cloned into the N-terminal (His-Asn)$_6$-tag pEcoli plasmid (Clontech, Mountain View, CA). The N protein expression plasmid was transformed into *E. coli* strain BL21 (DE3) pLysS. Transformed cells were grown at 37°C in LB media containing 100 µg/ml ampicillin and 50 µg/ml chloramphenicol until OD$_{600}$=0.6-0.8. Protein expression was induced by addition of IPTG to a final concentration of 0.5 mM and cultures were grown overnight at room temperature (24°C). Cells were harvested by centrifugation and stored at -80°C until processed. Eight to twelve liters of culture were processed weekly for three months. Because of the modest level of soluble protein expression, a total amount of purified N protein from 130 liters of culture was required for the screening.

Cells were lysed by resuspension of thawed pellets in Solulyse (Genlantis, San Diego, CA) lysis buffer containing a cocktail of protease inhibitors (Thermo Fisher Scientific, Waltham MA). Benzonase (EMD Chemicals, Rockland, MA) was added to the lysis buffer at a concentration of 12 units/mL to degrade nucleic acids. Insoluble cell debris was removed from the lysate by centrifugation at 10,000 RPM for 30 minutes at 4°C. The (His-Asn)$_6$-tagged N was batch purified from the cell lysate by using Co^{2+}-charged Talon resin (Clontech). The resin was washed with 15 volumes of wash buffer (35 mM imidazole, 500 mM NaCl, 1 M urea, 50 mM Tris-HCl pH 8.0), and the protein eluted with 4 volumes of elution buffer (300 mM imidazole, 500 mM NaCl, 50 mM Tris-

HCl pH 8.0). The eluate was dialyzed overnight against the storage buffer (500 mM

NaCl, 10% v/v glycerol, 50 mM Tris-HCl pH 8.0, two buffer changes) and concentrated

using 10K MWCO centrifugal filters (Amicon Millipore, Billerica, MA). After flash

freezing with liquid nitrogen, the aliquots were stored at -80˚C. The purity of the protein

was checked by SDS-PAGE (90%; evaluated by Coomassie staining) and the protein

concentration was determined by the absorbance at 280 nm using extinction coefficient

obtained from the ExPASy website (33,920 $M^{-1}.cm^{-1}$)
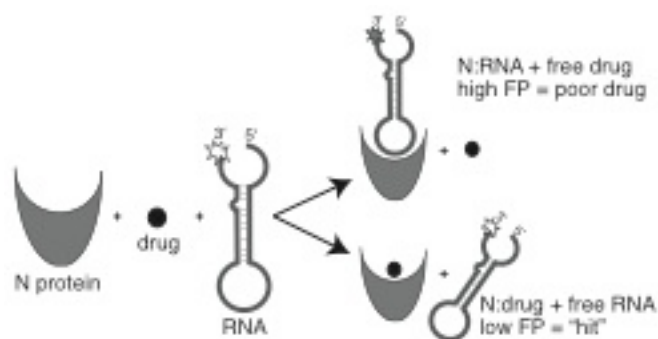
Figure 3 - 1: High throughput screening methodology.
Schematic diagram of the technique used to identify compounds that inhibit RNA binding to N protein *in vitro*. N was loaded into 384 well microplates, drug compound was added and then fluorescently labeled RNA. After a one hour incubation, fluorescence polarization (FP) measurements were taken. A high FP signal indicates that the compound was unable to prevent the RNA from binding N protein and would not be considered a potential antiviral drug. A low FP signal is indicative of a compound that inhibits the N-RNA interaction. Experimental wells that exhibit a low FP signal are considered "hits" and the compounds in those wells could be good drug candidates. Hits identified during the initial screen are subjected to further testing and verified in follow up assays.

D. High-throughput screening

The binding of the fluorescent RNA aptamer to the RVFV N protein was monitored

by fluorescence polarization. The high-throughput assay tested the ability of compounds

to prevent the binding of the RNA aptamer to N (Fig. 3-1). Screening was performed in

384-well low volume flat bottom black microplates (Corning #3820, Corning, NY). A

pilot screening (June 2011) was used to test and streamline the standard operating

procedure. The frozen stocks of 10-fold binding buffer, N protein, and fluorescent RNA

were sent to the screening facility ahead of time (NSRB/ICCB Longwood, Harvard Medical School, Boston MA). The optimum N protein and RNA aptamer concentrations were previously determined to be 12 μM and 10 nM, respectively because that resulted in a greater than 90% fractional saturation of the N-RNA complex (30). The assay was further checked for HTS robustness during the pilot screening.

The main screening (August 2011) started by the thawing and pooling of all N protein aliquots. The concentration of the resulting stock was 130 μM (*i.e.* 11-fold concentrated) and was kept at 4˚C. Each morning the amount necessary to screen the plates for the day was removed and diluted from 11- to 1.2-fold with binding buffer. The binding buffer (final concentration) is composed of 10 mM Hepes-NaOH pH 8, 150 mM NaCl, 20 mM KCl, and 5 mM $MgCl_2$. The standard operating procedure consists of several sequential steps. First, ten to twenty six assay plates were consecutively loaded with 10 μl of N protein (1.2-fold in binding buffer) with a Matrix WellMate dispenser (Thermo Scientific) in all but the last (24th) column, which received 10μl of binding buffer. Second, 100 nl of each compound from the compound plates were transferred into the corresponding wells of the 384-well assay plates with the Epson Compound Transfer Robot. Third, two μl of fluorescent RNA was added to all wells of the assay plates with a Multidrop Combi nL (Thermo Fisher Scientific). After a quick centrifugation, the plates were incubated in a closed drawer at room temperature (24˚C) for 30 min. Finally, the fluorescence polarization (FP) was measured with the EnVision plate reader (Perkin Elmer, Waltham, MA) and the FITC FP 480 (excitation) and 535 (emission) filters. Results were stored in Microsoft (Redmond, WA) Excel files.

E. Data analysis and hit scoring

The results were transferred into a single Microsoft Excel 2011 file and analyzed using the FP, parallel and perpendicular channels values. The fluorescence polarization is defined as (92):

$$FP = (I_s-I_p)/(I_s+I_p) \qquad \text{Eq. (1)}$$

where FP is polarization, I_s and I_p are parallel and perpendicular polarization emission light, respectively. Practically, the FP signal is calculated as follow:

$$FP=1000*(S-G*P)/(S+G*P) \qquad \text{Eq. (2)}$$

where FP is polarization in milliP (mP) units, S and P are the values of the parallel and perpendicular channels, respectively. The G-factor used during our screening was preset to 0.63.

To evaluate the robustness of the assay, we quantified the Z' parameter (122) for each plate using the FP values of columns 23 and 24. A Z' value above 0.5 describes a suitable HTS assay (122). The FP values of column 24 (RNA alone) are called positive controls, since they give the same signal than one would expect with an ideal compound, *i.e.* inhibiting 100% of protein/RNA interactions. Accordingly, FP values in column 23 (N protein and RNA) represent negative controls.

Several parameters were first calculated: the sum of the fluorescence of the parallel and perpendicular channels, the average and standard deviation of positive and negative controls, the plate Z', and the amplitude of FP signal for the plate (called amplitudeFP, *i.e.* average of negative controls minus average of positive controls). Experimental wells were excluded from scoring if the sum of the parallel and perpendicular channels was lower than 10 million or greater than 20 million fluorescence units. These data allowed us to eliminate compounds that interfere with the assay

detection either through quenching or compound fluorescence. The next calculation expressed the percentile of FP for an experimental well (X) related to the plate FP amplitude, provided that the well was not excluded:

$$\%FP^X = 100*(FP^X\text{-average}^{POS})/\text{amplitudeFP} \qquad \text{Eq. (3)}$$

where FP^X is the FP value for the well and the other parameters are described in previous paragraphs. The $\%FP^X$ values were then filtered-in if they are below a certain percentile or filtered-out if they are above. Finally the $\%FP^X$ for duplicates was averaged and the resulting hits were visually inspected.

F. Cherry picking and final compound assessment

The protocol used during HTS was modified as followed. The final volume of the reaction in binding buffer 1x was increased from 12.1 to 24.6 µl that includes 20 µl of N protein at a concentration of 14.8 µM, 4 µl of fluorescent aptamer RNA at 61.5 nM, and 0.6 µl of diluted compound in DMSO 50% v/v. The cherry-picked compounds (1.2 µl each in DMSO) were received in a 384-well plate format. A volume of 2.4 µl DMSO 25% v/v was added to each 1.2 µl of compound, so that 0.6 µl of diluted compound (in DMSO 50% v/v) could be pipetted per reaction. Except for DMSO, all other molecules, including N protein, fluorescent aptamer RNA, and compound, were tested at the same concentrations than the ones used during HTS. However, we checked that the slightly higher DMSO concentration used (1.2% vs. 0.8%) did not affect fluorescence polarization (data not shown). The compounds were tested in triplicate and the results were analyzed as described above for the main HTS experiment.

G. Fluorescence polarization competition assay

Fluorescence polarization competition assays were performed in 30 μL volumes and 384-well microplate format. Samples were prepared by incubating varying concentrations of compound (8.9 nM–500 μM final) or natural product extract ($2.37E^{-5}$–0.266 μg/μL final) for 20 min. at room temperature (~24°C) with 10 μM N in binding buffer (10 mM Hepes-NaOH pH 8, 150 mM NaCl, 20 mM KCl, and 5 mM $MgCl_2$). Next FAM-labeled aptamer RNA (10 nM final) was added and the reaction incubated for 1 hour at room temperature. Fluorescence polarization values were measured using a Biotek Synergy 2 plate reader. Excitation and emission wavelengths were 485 nm and 528 nm, respectively and the G-factor was 0.87. Binding profiles were plotted and IC_{50} values were calculated using Graph-Pad Prism software.

H. Filter binding competition assay

Samples were prepared by incubating varying concentrations of compound (8.9 nM–500 μM final) or natural product extract ($2.37E^{-5}$–0.266 μg/μL final) for 20 min. at room temperature with 10 μM N in binding buffer (10 mM Hepes-NaOH pH 8, 150 mM NaCl, 20 mM KCl, and 5 mM $MgCl_2$). Next, 1 μL of radiolabeled RNA (^{32}P-UTP, 25,000 cpm/μL) was added to the reaction and the final volume was 12 μL. After incubation at room temperature for one hour, reactions were diluted to 100 μL with ice-cold binding buffer and filtered through pre-soaked nitrocellulose filters (Millipore HAWP). The filters were washed twice with 500 μL ice-cold binding buffer, dried, and the radioactivity retained on filters was measured by scintillation counting. Binding profiles were plotted and IC_{50} values were calculated using Graph-Pad Prism software.

I. Plaque assays

Confluent monolayers of human 293 cells grown in 96-well plate format were treated with one microliter of natural product extract, unique compound, or DMSO. Next, cells were infected with RVF MP-12 virus using a MOI of 0.1. Supernatants of virus infected and uninfected cells were harvested at 2 and 3 days post infection. The supernatants were serially diluted in MEM and subsequently used to infect confluent Vero cells grown in 6-well plate format. After a 2 hour incubation the cells were overlaid with MEM containing 1% agarose and incubated for 7 days. Cells were fixed with 9.25% formaldehyde and plaques revealed using crystal violet stain. After 2 or 3 days the cytotoxicity of natural product extracts or compounds was measured using resazurin fluorescence. Resazurin (0.08 mM final) was added to wells containing cells treated with natural product extract, unique compound, or DMSO and the fluorescence measured at 490 nM using a Biotek Synergy 2 plate reader.

III. Results and Discussion

A. HTS assay validation

To identify compounds that could be used to develop prophylactic molecules against Rift Valley Fever Virus infection, we developed a screening assay to find inhibitors of the RVFV nucleocapsid (N) protein RNA-binding activity. Since viral RNA molecules are too large to chemically synthesize and fluorescently label we used an *in vitro* selection technique to isolate a small RNA aptamer that bound to N protein with high affinity. Competitive binding experiments using a RNA molecule that mimics the panhandle structure formed by the S segment of RVFV genomic RNA showed that the aptamer RNA binds to the same region on N as the viral panhandle (30). We used

fluorescence polarization (FP) to measure binding between N and a fluorescent RNA aptamer. Compounds that inhibit the protein-RNA binding decrease the FP value because the anisotropy of the free fluorescent RNA is less than that of the N-RNA complex. For instance, an ideal compound inhibiting 100% would have the same FP value than the positive control (RNA alone). During the HTS screen, the average value for the positive controls was 158.3±8.3 mP and the negative control average was 282.7±10.8 mP, resulting in an amplitude of 124.4 mP (Fig. 1).

We validated our binding assay for HTS by testing in a low volume 384-well plate format and calculating the HTS-related Z' score (122). Initial experiments conducted in our own and Dr. Brian Geiss' laboratory (Colorado State University, Fort Collins, CO) showed that Z' values at or above 0.5 could be obtained (data not shown). When the experiments were done in a complete HTS environment (NSRB/ICCB Longwood, Harvard Medical School, Boston MA), the value of the pilot Z' plate was 0.51, while the plates Z' of the main screen ranged from 0.437 to 0.801. The overall Z' average of 0.73±0.052 and the 95% confidence interval for the mean of 0.73±0.01 underscored the validity of the assay for HTS.

B. Compound screening results

The main screen was conducted over two days with 4708 and 21076 compounds tested during the first and second day, respectively. The first three series of experiments were done in duplicate. The stable and reproducible nature of the controls, as evidenced by individual plate Z' values, the reasonable number of positive hits and the rare occurrence of positive/negative pairs among duplicates (99.4% of duplicate values were 15% or less apart), as well as the limited amount of N available to work with prompted us

to screen the last 52 compound plates only once in the initial screen. Compounds showing activity in this broad initial screen would be verified in follow-up assays.

Hit analysis was done for all compound well values using Microsoft Excel (see Materials & Methods) and the results are reported in Fig. 3-2 and Table 3-1. We chose a cut-off of no less than 42% inhibition (= a FP amplitude decrease of more than 58%) that led to a total of 79 hits (hit rate at 0.3%, 79/26424). The average inhibition for the hits was 52±13% with a 48% median. Five of the 79 hits (6%) may be either false positive or negative: one of the two duplicate values would not have registered the compound as a hit on its own. Fifty-four of the hits belong to natural products extracts libraries (Table 3-1).

To validate the hits, we tested the 79 compounds and extracts in 384-well plates in our laboratory. We doubled the reaction volume and 3-fold diluted the stock of compounds we received, so a small volume of compound could be pipetted effectively. We checked that the small increase of DMSO in the final reaction had no effect on the binding yield (data not shown). The total confirmation rate was 63%, 40% for unique compounds and 74% for natural products extracts (Table 3-1).

Structural analysis of the 79 hit compounds revealed that four were from Life Chemicals F1860 series of aryl-benzofuran compounds. Two other molecules, Life Chemicals F0862-0529 and Chembridge 5406174, contain a quinone with an attached derivatized benzene ring (Table 3-2). We have yet to pursue medicinal chemistry of

these compounds to determine critical features of the inhibitory mechanism.
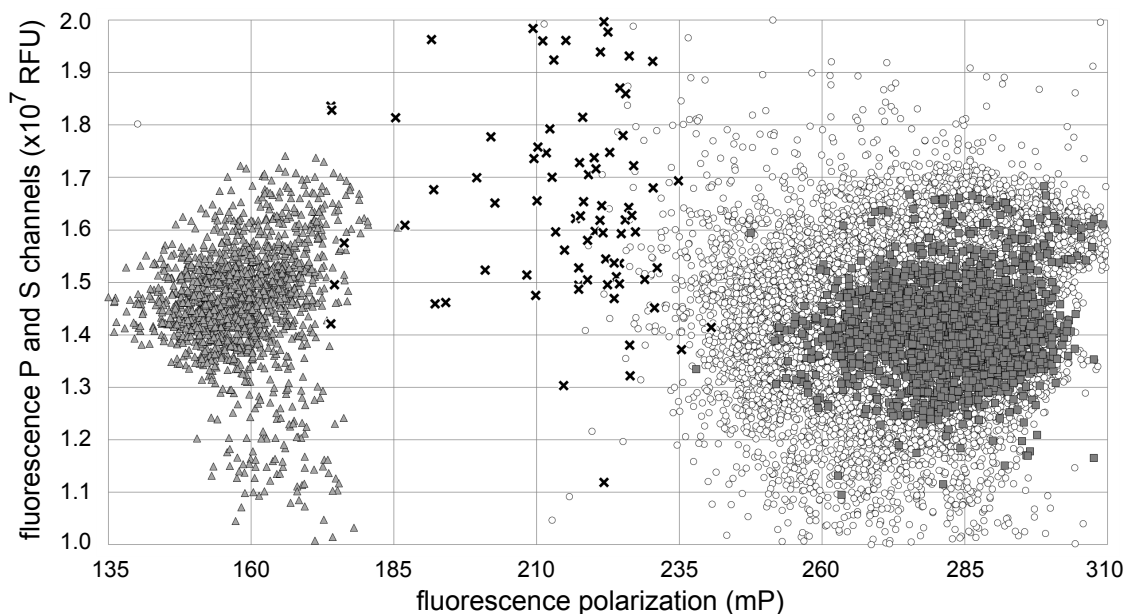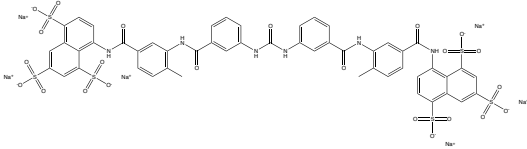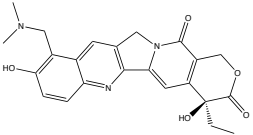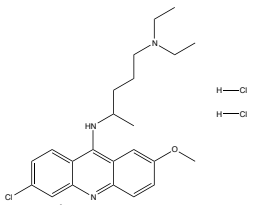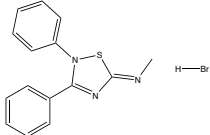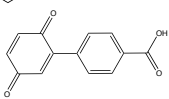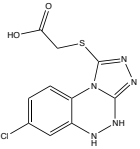
Figure 3 - 2: Scatter plot of HTS results.
The X-axis represents the fluorescence polarization in mP units for each well. The Y-axis represents the sum of the parallel and perpendicular fluorescence channels in relative fluorescence units. The limits of the Y-axis correspond to the values used to filter in experimental values during hit analysis (between 1-2 x 10^7 RFU). Positive (RNA alone) and negative (RNA and protein, no compound or extract) controls are indicated by dark triangles and dark squares, respectively. The values for the 79 hits and 26345 non-hits compounds wells are indicated by black crosses and open circles, respectively.

Table 3 - 1: Summary of statistical data from the HTS

	Total	Unique compounds	Natural product extracts
Number of compounds screened	26424	21144	5280
Total hits	79	25	54
Overall hit rate	0.3%	0.12%	1%
Percent repeat	63%	40%	74%
Number of compounds repeated	50	10	40
Number of compounds tested	8	6	2
Average Z'	0.727±0.052	0.734±0.047	0.707±0.063

Table 3 - 2: Lead compounds and natural product extracts

Structure	Company/Library & Compound ID	Percent Inhibition HTS	IC$_{50}$ [a]	Avg. Hill slope
	Sigma Lopac suramin sodium salt	93.3	16.3±7.3 (μM)	-2.76
	BIOMOL topotecan	85.9	ND[b]	ND[b]
	BIOMOL 4 quinacrine dihydrochloride	65.7	162±26 (μM)	-2.99
	Sigma Lopac S 4063	50.3	6.7±2.9 (μM)	-1.64
	Life Chemicals 5406174	47.9	5.1±0.6 (μM)	-0.97
	Maybridge BTB 11461	44.1	17.2±3.9 (μM)	-1.27
Not applicable	ICGB Extracts 18 600 6990	72.8	0.015 ± 0.0006 (ng/μL)	-3.24
Not applicable	ICGB Extracts 18 600 7051	45.7	0.030 ± 0.0004 (ng/μL)	-0.99

[a]Competitive binding assays were conducted using fluorescence polarization with the exception of quinacrine which was done using a filter binding assay technique previously described in Ellenbecker et al., 2012. ND, not determined. [b] Topotecan caused protein precipitation; none of the other compounds caused obvious precipitation.

C. Dose-response studies

Cherry picked compounds that were confirmed as giving a positive result in our

laboratory were purchased and the IC$_{50}$ values for five commercially available

compounds was determined using fluorescence polarization (FP). Competitive filter

binding assays were also conducted to ensure that the results obtained using FP were accurate (Fig. 3-3). Chemical structures, results from the HTS and results from the dose-response assays for five unique compounds are summarized in Table 3-2. Nearly all of the compounds tested had an IC_{50} value in the low micromolar range (5-17.2 µM). Suramin, the best inhibitor of N binding RNA in the HTS assay, is a drug that has been used to treat trypanosomiasis. Suramin also exhibits antiretroviral activity *in vivo* by inhibiting reverse transcriptase enzymes and has been shown to have an affinity for purinergic binding sites (26, 58, 116). These data suggest that in our system suramin prevents formation of the N-RNA complex by interacting with the RNA binding cleft of N. Suramin is amenable to medicinal chemistry, as evidenced by the several derivatives that have been described.

The IC_{50} value of quinacrine dihydrochloride was higher (162 µM) than the other compounds tested. Since quinacrine is known to intercalate into double-stranded DNA and structured RNA molecules, it is likely that this compound is targeting RNA itself, rather than the interaction with N protein directly (84, 97, 113). We hypothesize that quinacrine binds to the aptamer RNA, altering its structure and thus decreasing its affinity for N. Studies conducted in picornaviruses showed that quinacrine inhibited translation of viral proteins by binding to internal ribosome entry sites in 5' untranslated region of viral mRNA (37).

Two natural product extracts selected for further testing based on their performance in the HTS and follow-up assays were provided by Jon Clardy (Harvard Medical School). The IC50 values of extracts 6990 and 7051 were 0.015 ng/µL and 0.030 ng/µL respectively. It is difficult to directly compare the commercially available compounds to

the natural products extracts because we do not know the size or number of compounds

present in an extract. As a loose approximation for discussion purposes, if we assume

that the extracts contained ~20 compounds averaging ~300 Da then the IC50 for the

active component(s) would be in the low micromolar range (2-5 µM) and would be

performing as well, or even better than the unique compounds (Table 3-2). We are

currently seeking to identify and characterize the active component(s) in the natural

product extracts that show good activity in the molecular target assay.
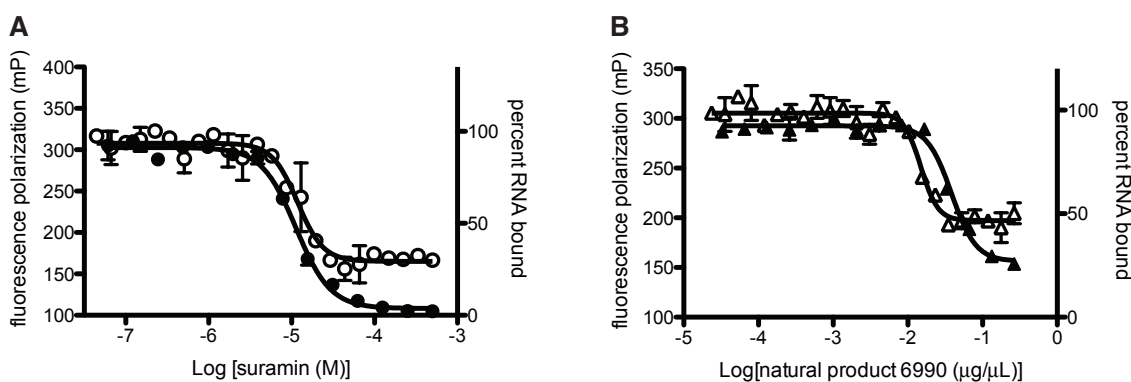
Figure 3 - 3: Dose-response curves of suramin and natural product extract 6990.
Nitrocellulose filter binding assays (closed symbols) and fluorescence polarization (FP) (open symbols)
were used to measure the ability of either suramin (A; circles) or natural product extract 6990 (B; triangles)
to inhibit binding of aptamer RNA to N. The competition assays involved incubating a fixed concentration
of N with varying concentrations of compound and either fluorescent or radiolabeled RNA. For the filter
binding assay the amount of radiolabeled (^{32}P-UTP) aptamer RNA retained on filters was measured by
liquid scintillation counting and the percent RNA bound is plotted versus compound or natural product
extract concentration. The fluorescence polarization experiments used fluorescently labeled aptamer RNA
and the FP signal (mP) is plotted versus compound or natural product extract concentration. The IC_{50} value
calculated for suramin using a filter binding assay (A; closed circles) was 11.3 µM and using FP (A; open
circles) was 16.3 µM. The IC50 value calculated for natural product extract 6990 using a filter binding
assay (B; closed triangles) was .038 ng/µL and using FP (B; open triangles) was .015 ng/µL. Similar
experiments were conducted to calculate IC_{50} values for the other unique compounds and natural product
extract (Table 2). The results are the average of at least two independent experiments.

D. Suramin and 5406174 inhibit RVFV replication in cell culture

Plaque assays were conducted to measure the ability of the unique compounds

identified in the HTS assay to inhibit the production of viable virus over the course of

several days. Human 293 cells were infected at a low multiplicity of infection (MOI 0.1)

in the presence or absence of compound and the number of infectious viral particles

produced was quantitated using a plaque assay. Of the eight unique compounds tested

thus far, two have exhibited antiviral activity in cell culture. Treatment of cells with

either 5406174 or suramin decreased the number of plaque forming units per mL by

approximately one log at two and three days post infection. Conversely, no decrease in

virus production was observed when cells were treated with compound 0250 (Fig. 3-4).

The cytotoxicity of the compounds was measured using resazurin fluorescence. The

metabolic activity of living cells converts resazurin into a fluorescently active compound

that can be detected at 490 nm using a microplate reader. The amount of compound used

to treat cells was previously determined to be non-cytotoxic (data not shown).
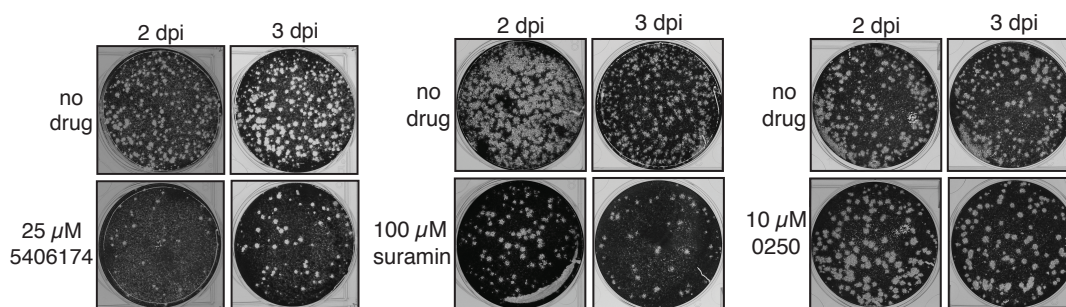
Figure 3 - 4: Analysis of the effect of unique compounds on virus yield in cell culture.
Plaque assay to determine the number of infectious viral particles produced from cells treated with either DMSO (no drug), 5406174, suramin or 0250 at two and three days post infection (dpi).

E. A subset of natural product extracts identified in the HTS assay inhibit RVFV replication in cell culture

Plaque assays were conducted to measure the ability of the natural product extracts to

inhibit production of viable virus over the course of several days. The cytotoxicity of the

extracts was also measured using resazurin fluorescence and only extract N5 exhibited

cytotoxicity in cell culture (Fig. 3-6B). Unexpectedly, some natural products stimulated

virus production. For example, at both two and three days post infection natural product

extract 6990 increased virus yield (Fig. 3-5). Six of extracts tested inhibited RVFV

replication by 50% or more and were not cytotoxicity to cells (Fig 3-6A and B).
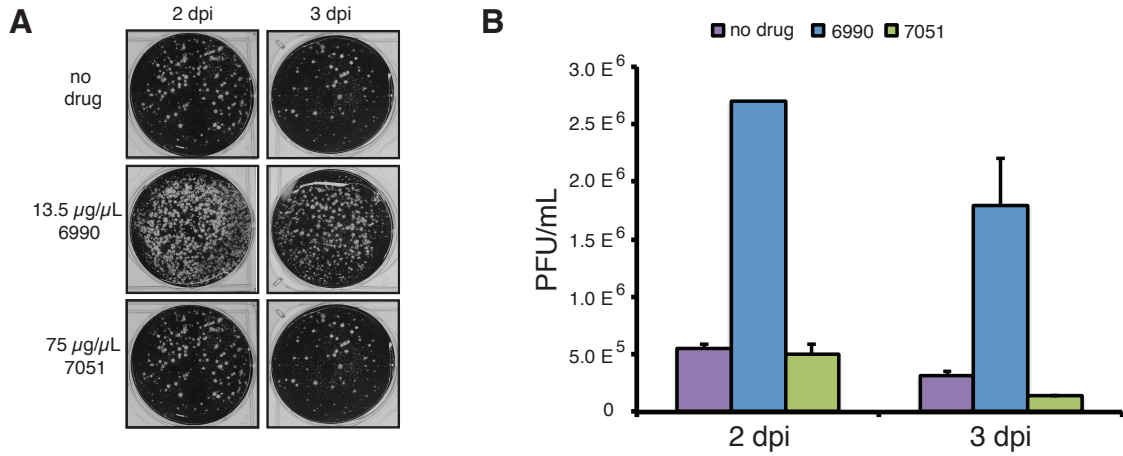
Figure 3 - 5: Analysis of the effect of natural product extracts on virus yield in cell culture.
A) Plaque assay to determine the number of infectious viral particles produced from cells treated with either DMSO (no drug), extract 6990 or extract 7051 at two and three days post infection. B) Quantification of the plaque assay in Fig. 3-2A. Cells were treated with either DMSO (purple fill), extract 6990 (blue fill), or 7051 (green fill).
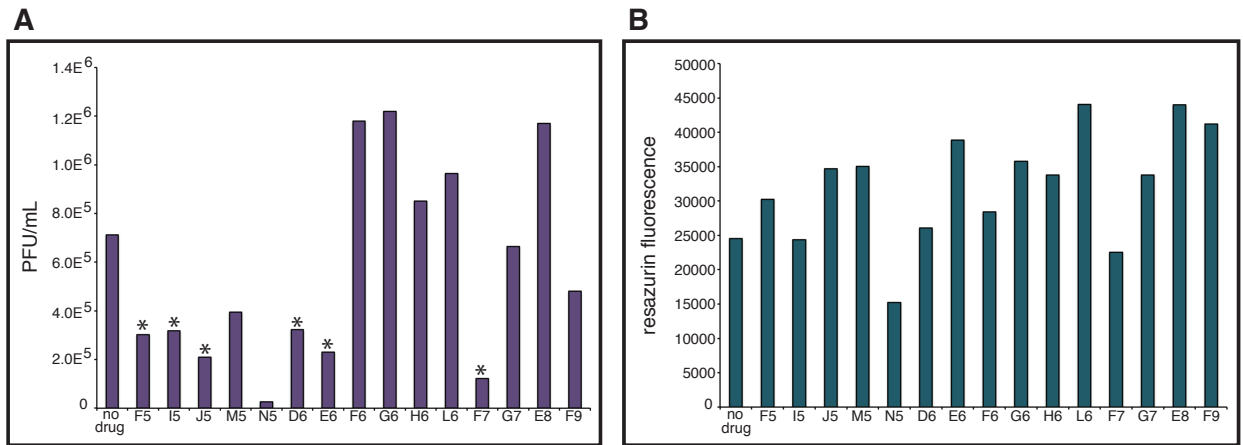
Figure 3 - 6: Analysis of the effects of natural product extracts on virus yield and cell viability in human cell culture.
A) Quantification of plaque assays to determine the number of infectious viral particles produced from cells treated with either DMSO (no drug) or natural product extract (F5-F9). Asterisks indicate natural product extracts that decrease virus production by 50% or more and are not cytotoxic to cells. B) Cytotoxicity assay to measure the effects of natural product extracts on cell viability. Resazurin was added to wells treated with either DMSO or natural product extract (F5-F9) and the fluorescence measured at 490 nm

In conclusion, we describe a high-throughput molecular target assay that was used

to identify FDA approved drugs, unique compounds and natural product extracts that

have the ability to disrupt the RVFV N-RNA complex *in vitro*. Since N is indispensible

for bunyavirus replication, we anticipate that a subset of the lead compounds or derivatives designed using medicinal chemistry will be efficacious in cell culture (64). The ability of the compounds and natural product extracts to inhibit viral replication over the course of several days was determined using plaque assays to quantitate production of viable virus. These secondary assays allowed us identify the compounds and extracts that have the potential to become antiviral drugs and provide us with promising candidates for future *in vivo* studies. The work presented here represents a promising start toward identifying new inhibitors of RVFV infection.

.

CHAPTER 4: CHARACTERIZATION OF RIFT VALLEY FEVER VIRUS NUCLEOCAPSID PROTEIN-RNA COMPLEX FORMATION AND DISSOCIATION BY THE INHIBITOR SURAMIN

I. Introduction

Rift Valley fever virus (RVFV) is a mosquito-transmitted bunyavirus (genus *Phlebovirus*) that causes severe disease in humans and ruminant livestock. In humans disease symptoms can range from a mild flu-like illness to hemorrhagic fever, encephalitis, neurological disorders and blindness (65). Pregnant livestock are at risk for miscarriage and high mortality rates among newborn animals have been reported (20-22). RVFV can be transmitted by a diverse set of mosquito species and outbreaks have been more severe outside of the historically defined endemic areas of Sub-Saharan Africa. The potential for RVFV to cause devastating epidemics worldwide is evidenced by its classification as a Category A high priority disease agent by the National Institute for Allergy and Infectious Diseases (NIAID) (10, 112). There are currently no proven safe, effective treatment options for RVFV infected individuals or livestock. Increasing our understanding of the basic molecular virology of this important pathogen represents an essential step toward identification of new drug targets and the development of more efficacious antiviral therapeutic compounds.

The single-stranded RNA genome of RVFV is composed of three segments (L, M and S) that appear circular by electron microscopy due to complementarity between the 5' and 3' terminal regions (49). The L segment encodes the RNA dependent RNA polymerase responsible for transcription and replication of the viral genome. The M segment encodes glycoproteins (Gn and Gc) involved in entry into the host cell and non-structural proteins Nsm and 78-kDa (62). The S segment utilizes an ambisense strategy to encode the nucleocapsid (N) and NSs proteins while the genes of the M and L segments are in the negative sense (39). Among all viral gene products, the importance

of N is underscored by its involvement during many stages of the RVFV replication cycle. N is an RNA binding protein that plays important roles during transcription, translation, replication and packaging of the viral genome into virions (74, 79, 88, 89, 101, 114). During replication N monomers cooperatively bind along the entire length of the viral genome and antigenome. This encapsidation event protects the viral RNA from degradation and prevents the formation of double stranded RNA intermediates that could activate the host innate immune response (105). To discriminate between full-length viral RNA and the plethora of other abundant species of RNAs present in the host cell, it is thought that N must initially recognize some distinct sequence and/or secondary structural element specific to viral RNA. Subsequently, additional N molecules bind the entire length of the viral RNA. Studies conducted in other related viruses such as Bunyamwera, Sin Nombre, Hantaan and La Crosse viruses identified single or double stranded structures within the terminal non-coding regions of the RNA genome as containing important sequence recognition elements for N (72, 76, 78, 87, 98, 109). Our laboratory used an *in vitro* selection technique (SELEX) to identify and amplify RNAs called *aptamers* that bind to RVFV N with high affinity. In the *in vitro* selection scheme we observed a recurring GAUU motif that was found to be important for N recognition and binding to many of these aptamer RNAs (30).

A high-resolution crystallographic structure of N bound to RNA was recently published (100). Analysis of this structure revealed an RNA binding groove containing several highly conserved basic amino acid residues located within the C-terminal core domain, but the structure did not lend insight into how N may recognize RNAs in a sequence or structure-dependent fashion. The N monomer also contained a flexible N-

terminal arm that interacts with the neighboring N monomer and this protein-protein interaction likely facilitates cooperative N binding (32, 100). However, the details of how RVFV N and other viral nucleocapsid proteins interact with their cognate RNAs are not well understood. We hypothesize that N binds RNA in a biphasic manner. After an initial specific binding event, subsequent N monomers bind in a non-specific mode and coat the entire length of the viral genome or antigenome. Furthermore, we propose that the inhibition of RNA binding to N represents an attractive antiviral therapeutic strategy because several essential steps in the RVFV replication cycle involve N binding to viral RNA.

In this study we characterize the cooperative assembly of N monomers onto an RNA molecule using mutational analysis, biochemical binding assays and RNA structure probing techniques. We elucidate the mechanism utilized by suramin, a small molecule identified by our laboratory in a high-throughput screen as an N-RNA binding inhibitor, to disrupt both the initial specific binding event, as well as subsequent non-specific binding events important for ribonucleoprotein (RNP) complex formation (29). Finally, we show that suramin inhibits RVFV replication in human cell culture. Using time of addition analysis, we determine that suramin exerts its inhibitory effect both by interfering with N-RNA binding and by blocking virus uptake into cells and/or other step(s) post entry that have not yet been described.

II. Materials and Methods

A. Overexpression and purification of RVFV wild type and mutant N proteins

The WT N protein sequence used in this study corresponds to the experimental live attenuated RVFV MP-12 vaccine strain (17). A mutated variant of N lacking the 33

amino acid residues that comprise the N-terminal oligomerization domain (Δ33 mutant) and a mutant that is unable to bind RNA (RNA binding (-) mutant) were also used. Three conserved basic residues in the RNA binding cleft of N were changed to aspartic acid (R64D, K67D and K74D) to abolish the RNA binding activity of N (32). Sequences for the WT N and mutant constructs were cloned into the N-terminal (His-Asn)$_6$-tag pEcoli plasmid (Clontech, Mountain View, CA). Plasmids encoding WT, Δ33 or RNA binding (-) mutant RVFV N protein were transformed into *E. coli* strain BL21-CodonPlus. Transformed cells were grown at 37°C in LB media containing 100 μg/ml ampicillin and 50 μg/ml chloramphenicol until the optical density at 600 nm reached 0.6-0.8. Protein expression was induced by addition of IPTG to a final concentration of 0.5 mM and cultures were grown overnight at room temperature (24°C). Cells were harvested by centrifugation and stored at -80°C.

The thawed cell pellets were lysed by resuspension in lysis buffer (50 mM sodium phosphate pH 8.0, 50 mM NaCl, 10 mM imidazole, 1 M urea, 1% v/v Triton X-100) in the presence of protease inhibitors. Benzonase (Novagen) was added to the lysis buffer at a concentration of 12 units/mL to degrade DNA and free RNA. Insoluble cell debris was removed from the lysate by centrifugation at 10,000 RPM for 30 minutes at 4°C. The (His-Asn)$_6$-tagged proteins were batch purified from the cell lysate using Co^{2+}-charged Talon resin (Clontech). The resin was washed 4 times with 50 mM sodium phosphate pH 8.0, 500 mM NaCl,10 mM imidazole, 1M urea and the protein eluted with 25 mM HEPES pH 8.0, 500 mM NaCl, 300 mM imidazole. The eluate was dialyzed overnight against the storage buffer (500 mM NaCl, 10% v/v glycerol, 50 mM Tris-HCl pH 8.0, two buffer changes) and concentrated using Amicon centrifugal filters (10K

MWCO). After flash freezing with liquid nitrogen, the aliquots were stored at -80˚C.

The purity of the protein was checked by SDS-PAGE and the protein concentration was

determined by the absorbance at 280 nm using extinction coefficients obtained from the

ExPASy website (33,000 $M^{-1}cm^{-1}$ for WT and RNA binding (-) mutant and 24,750 $M^{-1}cm^{-1}$ for the Δ33 mutant).

B. Synthesis of aptamer RNA constructs

Two PCR reaction and ligation steps were conducted to build an RNA construct that

contains an aptamer RNA sequence (previously shown to bind N with high affinity)

attached to the 3' end of the human 5S rRNA sequence. First, a sense primer containing

a *Sal*I site (5'- CCA TAC GAG TCG ACG GCA TTA CGG CCG GG-3') and an anti-

sense primer containing a *Xba*I site (5'-GCT CTA GAA GAC GC-3') were used to

amplify a 55-nt-long sequence that corresponds to MBE59 aptamer RNA (30). The PCR

fragment was agarose gel-purified, digested with *Sal*I and *Xba*I and ligated into a pAV

plasmid using the Quick ligation kit (New England Biolabs). The pAV plasmid

containing the 5S rRNA gene followed by *Sal*I and *Xba*I restriction sites for insertion of

sequences encoding test RNAs was a kind gift from David Engelke (University of

Michigan, Ann Arbor, MI) (91). Second, a sense primer containing an *Eco*RI site with a

T7 RNA polymerase promoter (5'-CGG AAT TCT AAT ACG ACT CAC TAT AGG

GTC TAC GGC CAT ACC ACC-3') and an anti-sense primer containing a *Bam*HI site

(5'-ACA GGA TCC GCT CTT CCA AAA GCG GAC CGA AG-3') were used to

amplify the 200-nt-long segment corresponding to the human 5S rRNA gene plus the

MBE59 aptamer RNA sequence. The PCR product was agarose gel-purified, digested

with *Eco*RI and *Bam*HI and ligated into the pUC18 plasmid. The MBE59/5S plasmid DNA construct was verified by DNA sequencing.

RNA was synthesized *in vitro* using the MBE59/5S construct digested with *Xba*I as a template for T7 RNA polymerase. MEGAshortscript kits (Ambion) were used to synthesize non-radioactive RNA and MAXIscript kits (Ambion) were used to synthesize RNA in the presence of radioactive α-^{32}P-UTP (800Ci/mmol; PerkinElmer). The transcription reaction was DNase treated and purified by phenol/chloroform extraction and ethanol precipitation. The RNA was further purified by denaturing gel electrophoresis. The band corresponding to the RNA was visualized by UV shadowing, excised, eluted into 0.5M ammonium acetate, 1.0mM EDTA pH 8.0 and 0.2% SDS, and ethanol precipitated.

C. Electrophoretic mobility shift assays

For electrophoretic mobility shift assays, a 35nt long fluorescently labeled aptamer-based RNA was used (sequence: 5'- CGG GCU GUU UAC UGA ACU AUG AUA CAA AGA CCC CG-fluorescein-3'; previously described in reference 9). N protein (50 μM final) was incubated with FAM-RNA (1.0 μM final) in a binding buffer (10 mM HEPES pH 7.3, 150 mM NaCl, 20 mM KCl, and 5 mM MgCl$_2$ final concentration) at 10°C for varying amounts of time. Samples for competitive binding experiments were incubated at room temperature (~24°C) for varying amounts of time before competitor RNA (25 μM final) was added and the reaction incubated for 1 more hour. After incubation the reactions were placed on ice and glycerol loading dye was added to each tube. Samples were loaded onto pre-chilled 6% acrylamide/1X TB gels that were run in the cold room at

120 volts for 75 minutes. Gels were visualized and quantified on a Fuji FLA3000G Image Analyzer.

D. Fluorescence polarization measurements

Aliquots of WT and $\Delta 33$ N were serially diluted to varying concentrations in a binding buffer (10mM HEPES pH 8, 150mM NaCl, 20mM KCl, 5mM $MgCl_2$ and 10% v/v glycerol). To determine the equilibrium dissociation constant of N/FAM-RNA complex, various dilutions of N were added to 10 nM FAM-labeled aptamer RNA that had been subjected to denaturation at 90°C for two minutes followed by snap cooling on ice. Reactions were incubated at room temperature (24°C) for 1 hour. Samples for competitive binding experiments were prepared by incubating either WT (15 μM) or $\Delta 33$ mutant (50 μM) with 10 nM FAM-labeled RNA in binding buffer at room temperature for 1 hour. Next, varying concentrations of suramin were added (524 pM to 588 μM final concentration) and the reaction incubated for 1 hour at room temperature. Fluorescence polarization values were measured using a Synergy 2 plate reader (BioTek). Binding profiles were plotted and apparent K_d and EC_{50} values were calculated using GraphPad Prism software.

E. Gel filtration chromatography

Purification of recombinant WT N from bacteria results in a heterogeneous population of monomers and multimers in a complex with *E. coli* RNA. For some experiments isolation of the lowest molecular weight species (monomeric population) was required and gel filtration chromatography was used to fractionate N. A Superose 12 10/300 GL column (GE Healthcare Life Sciences) was equilibrated in FPLC running

buffer (10 mM HEPES pH 8.0, 500 mM NaCl, 20 mM KCl and 5 mM $MgCl_2$) using an

ÄKTA FPLC system (Amersham Biosciences). After elution from the cobalt resin, N

was concentrated to approximately 10 mg/ml and 200 μL was loaded onto the column.

Protein was eluted at a flow rate of 0.2 ml/min and monitored by absorbance at 280 nm.

The following standard proteins were used to calibrate the column: carbonic anhydrase

(29 kDa), alcohol dehydrogenase (ADH) monomer (44 kDa), albumin (66 kDa), ADH

dimer (74 kDa) and tetramer (150 kDa). Blue dextran (2000 kDa) was used to determine

the void volume.

The Superose 12 column was also used analytically to compare the multimeric state

of N in the presence and absence of suramin. First, the column was equilibrated in FPLC

running buffer. Purified WT or RNA binding (-) N was adjusted to a concentration of 5

mg/mL in 10 mM HEPES pH 8.0/500 mM NaCl/20 mM KCl/5 mM $MgCl_2$ and

incubated at room temperature for 1 hour in either the presence or absence of varying

concentrations of suramin. N was loaded onto the column (600 μg; 120 μL), eluted at a

flow rate of 0.2 ml/min and absorbance at 280 nm measured. Fractions (0.5 ml) were

collected and analyzed for RNA content by phenol/chloroform extraction, ethanol

precipitation and denaturing gel electrophoresis. Gels were stained with ethidium

bromide, scanned using a Fuji FLA-3000G phosphorimager and RNA content quantified

by measuring band intensities using Image Gauge software.

F. Time-dependent RNA structure probing

First, MBE59/5S was subjected to denaturation at 90°C for two minutes followed by

snap cooling on ice. Next, WT monomeric N (30 μM) and RNA (0.5 μM) were

incubated in a binding buffer (10mM HEPES pH 8, 150mM NaCl, 20mM KCl and 5mM

MgCl$_2$) for 5 minutes or 5 hours at 10°C to allow ribonucleoprotein (RNP) complexes to form. The MBE59/5S alone and RNP complexes were probed using either RNase T1 or RNase A. T1 RNase reactions were each incubated with 0.18 units of enzyme per tube for 4 minutes at 10°C and RNase A reactions were each incubated with 6.25 ng of enzyme per tube for 2 minutes at 10°C. Reactions were stopped by phenol/chloroform extraction followed by ethanol precipitation. Samples were resuspended in Nanopure water and stored at -20°C.

G. Analysis of probing experiments

The probed RNA samples were reverse transcribed using three different γ-^{32}P-labeled DNA oligonucleotides located 10-40 nucleotides upstream of the region of interest on the RNA and AMV-RT (Promega). The sequences of the primers used are: 5'-TCT AGA AGA CGC CCC-3', 5'-ACG CCT ACA GCA CCC-3' and 5'-AGG CCC GAC CCT GC-3'. After a 30 minute primer extension at 42°C NaOH was added at a 0.1N final concentration to each sample and incubated at 65°C for 10 minutes. Samples were ethanol precipitated, dried, resuspended in formamide loading dye and subjected to denaturing gel electrophoresis. Gels were exposed using a Fuji FLA- 3000G phosphorimager and quantification of band intensity (which corresponds to the reactivity of RNA bases) was performed using Image Gauge software. Band intensities were adjusted to eliminate small variation in lane loading by dividing all bases affected by N by an internal control value. In other words, the intensity of each experimental band was divided by the intensity of a control band present in each lane. For control bands, we used naturally occurring RT pause/stop sites referred to as K-bands. These bands were identified by their appearance in the MBE59/5S control lanes (K-lanes) where samples

were not treated with RNases. Since the intensities of the K-bands are consistent for each sample, these bands were used to normalize experimental bands and compensate for any differences in band intensity that could be attributed to experimental error. After normalizing all of the experimental bands, the intensity of bands in samples incubated with N for 5 minutes or 5 hours were compared to the band intensities of the RNA alone control.

H. Cell culture studies

Human 293 cells were maintained in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), penicillin and streptomycin. Vero cells were maintained in minimum essential medium (MEM) alpha medium supplemented with 10% FBS, penicillin and streptomycin. The Rift Valley fever virus (RVFV) vaccine strain, MP-12, was kindly provided by Brian Gowen (Utah State University, Logan, UT). During infection, human 293 cells and virus were maintained in DMEM supplemented with 2% FBS, penicillin and streptomycin. Incubations were carried out at 37°C and 5% CO_2 unless otherwise stated.

I. Virus yield reduction assays

Growth media from confluent monolayers of human 293 cells in 96-well plate format was removed and replaced with DMEM supplemented with 2% FBS. Cells were treated with one microliter of suramin (20 mM stock; 100 μM final concentration) or DMSO. Next, cells were infected with RVFV using a MOI of 0.1 and incubated for 2 hours. After incubation media was removed, cells were washed with PBS and fresh media as well as one microliter of suramin or DMSO were added. Supernatants of virus infected

and uninfected cells were harvested at two and three days post infection. The amount of virus present in the supernatant was quantitated using plaque assays.

J. Time of addition virus yield reduction assays

To determine if suramin blocks an early or late step in the virus lifecycle we modified the protocol described in the previous paragraph. Growth media from confluent monolayers of human 293 cells grown in 96-well plate format was removed and cells were washed with cold DMEM. Next, 100 μL of cold DMEM supplemented with 2% FBS was added to each well. Either ammonium chloride (an inhibitor of viral entry; 12 mM final concentration) or suramin (50 or 100 μM final concentration) was added to cells that received drug before infection. Cells were infected with RVFV using a MOI of 1.0 and incubated for 1 hour at 4°C to allow virus to attach but not enter cells. After incubation media was removed, cells were washed with cold PBS to remove unbound virus and 200 μL of prewarmed media was added to each well. Ammonium chloride and suramin were added back to wells that received drug before infection and the plate was incubated for one hour to allow the virus to enter the cells. After 1 hour either ammonium chloride or suramin was added to experimental wells that received drug after the virus entered cells. Supernatants of virus infected cells were harvested the next morning.

K. Plaque assays

Supernatants from virus infected and uninfected cells were serially diluted in MEM and subsequently used to infect confluent monolayers of Vero cells grown in 6-well plate format. After a 2 hour incubation the cells were overlaid with minimum essential medium (MEM) containing 1% agarose and incubated for 7 days. Cells were fixed with

9.25% formaldehyde and plaques revealed using crystal violet stain. After 2 and 3 days the cytotoxicity of suramin was measured using resazurin fluorescence. Resazurin (0.08 mM final) was added to wells containing cells treated with either suramin or DMSO and the fluorescence measured at 490 nM using a Biotek Synergy 2 plate reader.

III. Results

A. Competitor RNA is unable to displace prebound RNA in mature Rift Valley fever virus ribonucloprotein complexes

An electrophoretic mobility shift assay was conducted to test the ability of a fluorescently (3'-FAM) labeled RNA to bind Rift Valley fever virus (RVFV) nucleocapsid (N) protein and visualize the ribonucleoprotein (RNP) complexes of varying size that form during an N-RNA binding event. A constant amount of FAM-RNA and N were incubated at 10°C for varying amounts of time (1 minute – 5 hours). The binding reaction was incubated at 10°C to decrease the speed of RNP formation because previous experiments conducted in our laboratory indicated that N binds RNA rapidly at room temperature (data not shown). After incubation, the samples were subjected to non-denaturing gel electrophoresis at 4°C. The bands that migrate faster on the gel correspond to free RNA species and the higher migrating bands represent higher molecular weight RNA/protein complexes. The results showed that when the binding reaction was incubated for 1 minute – 2 hours two distinct N-RNA complexes were formed. However, after a 5-hour incubation period the bands that corresponded to the lower molecular weight N-RNA complex and free RNA species disappeared suggesting that over time the RNP complex becomes larger (Fig. 4-1A), suggesting that additional molecules of N bind along the length of the RNA molecule and eventually incorporate all of the FAM-RNA into the highest molecular weight N-RNA complexes.

90
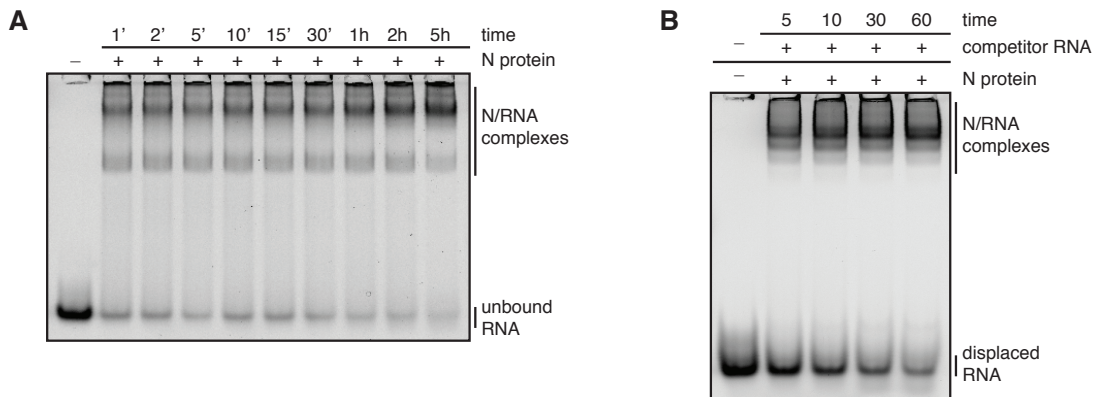
Figure 4 - 1: Characterization of N-RNA binding interactions.
A. EMSA analysis of fluorescently (3'-FAM) labeled RNA binding affinity for N after incubation for varying amounts of time. The bands that migrate faster on the gel are unbound RNA species and the slower migrating bands represent N/RNA complexes. B. A competitive EMSA to test the ability of an unlabeled competitor RNA to displace FAM-RNA that was preincubated with N for varying amounts of time.

Since N bound RNA efficiently, we next chose to study the stability of N-RNA complexes. A competitive binding experiment was performed to test the ability of an unlabeled RNA to bind N and displace a fluorescently (FAM) labeled RNA. A constant amount of FAM-RNA was incubated at room temperature with N for varying amounts of time (5-60 minutes) and then competitor RNA was added. The reactions were incubated at room temperature for an additional hour and subjected to non-denaturing gel electrophoresis. The results showed that competitor RNA was able to bind N and displace some FAM-RNA when added 5 or 10 minutes into the incubation period. However, when FAM-RNA and N were incubated for 30 to 60 minutes prior to addition of competitor, significantly less FAM-RNA was displaced (Fig. 4-1B). These data suggest that after an initial N/RNA binding event the RNP complex undergoes a conformational change and a more robust and impenetrable complex is formed. It is also possible that competitor RNA is unable to displace prebound RNA after prolonged incubation because the dissociation rate (k_{off}) of the mature complex is very slow. To further study the formation of high molecular weight N-RNA complexes we used RNA

structure probing to visualize changes in the reactivity of RNA bases that occur over time as a result of N binding.

B. A specific RNA binding event triggers cooperative binding between N monomers

Previously, we used an *in vitro* selection technique (SELEX) to generate high affinity 55 nt-long RNA ligands, or aptamers, to RVFV N. Mutational analysis of aptamer RNAs identified a GAUU motif found in many of the aptamers as an important recognition sequence for N (30). To create a longer (~200 nt) RNA molecule with additional non-specific N-binding sites and additional primer binding sites for structure probing studies, we added the human 5S rRNA sequence to the 5' end of an aptamer RNA with a GAUU motif (MBE59/5S). The 5S rRNA molecule was chosen because the structure is well characterized, robust, and folds readily and stably into its native structure. Therefore, addition of new sequence to its terminus does not alter the structure of the attached aptamer RNA.

The MBE59/5S RNA was incubated with N in a binding buffer for 5 minutes or 5 hours at 10°C to allow RNP complexes to form. The complexes were then probed using either RNase T1 (cleaves after single-stranded guanosine residues) or RNase A (cleaves after single-stranded uridine and cytidine residues). The probed RNA samples were extended using a γ-^{32}P-labeled DNA primer and AMV reverse transcriptase and the reactivity of the individual bases to the RNase enzymes in the presence and absence of N was analyzed. By comparing cDNA band patterns produced by primer extension of probed RNA alone and probed RNA complexed with N, nucleotides exhibiting either reduced or enhanced reactivates in the N-RNA complex were identified. This

information was then used to deduce the RNA sequences specifically contacted by N, as well as secondary structural rearrangements that occurred when N bound RNA.

When N was incubated with MBE59/5S and then probed with RNase T1 there was a strong and very rapid protection of two G residues located within the GAUU motif (G156 and G152), suggesting that this sequence is immediately recognized and bound by N (Fig. 4-2C; Fig. 4-3). However, G residues located away from the GAUU motif (G75, G48 and G47) exhibited little to no protection at the 5 minutes time point but after an extended incubation showed moderate (G75) and significant decreases in reactivity (G48 and G47) (Fig. 4-2A and B; Fig. 4-3). These data are consistent with the idea that an initial specific binding event triggers subsequent binding of N monomers and/or that the highest affinity binding sites are occupied first followed by binding to lower affinity sites.

A pattern similar to the one described above was observed when RNP complexes were probed with RNase A. For example, a U residue present in the GAUU motif (U159) was strongly protected at both the 5 minute and 5 hour time points whereas several C and U residues located closer to the 5' end of the RNA construct and farther away from the GAUU motif exhibit weak protections at 5 minute time point but became more protected over time (Fig. 4-2D-F; Fig. 4-3). Interestingly, residues U130 and U76 located approximately in the middle of the RNA construct initially exhibited enhanced reactivity and then became protected after 5 hours (Fig. 4-2E and F; Fig. 4-3). It is possible that binding of the first molecule of N stabilizes or alters an adjacent binding site for additional N molecules. These enhancements could be caused by a structural rearrangement in the RNA that occurs when N initially recognizes and binds the GAUU

motif. Then later, as N oligomers coat the entire RNA molecule protection of all reactive

G, C and U residues was observed (Fig. 4-2D-F; Fig. 4-3).
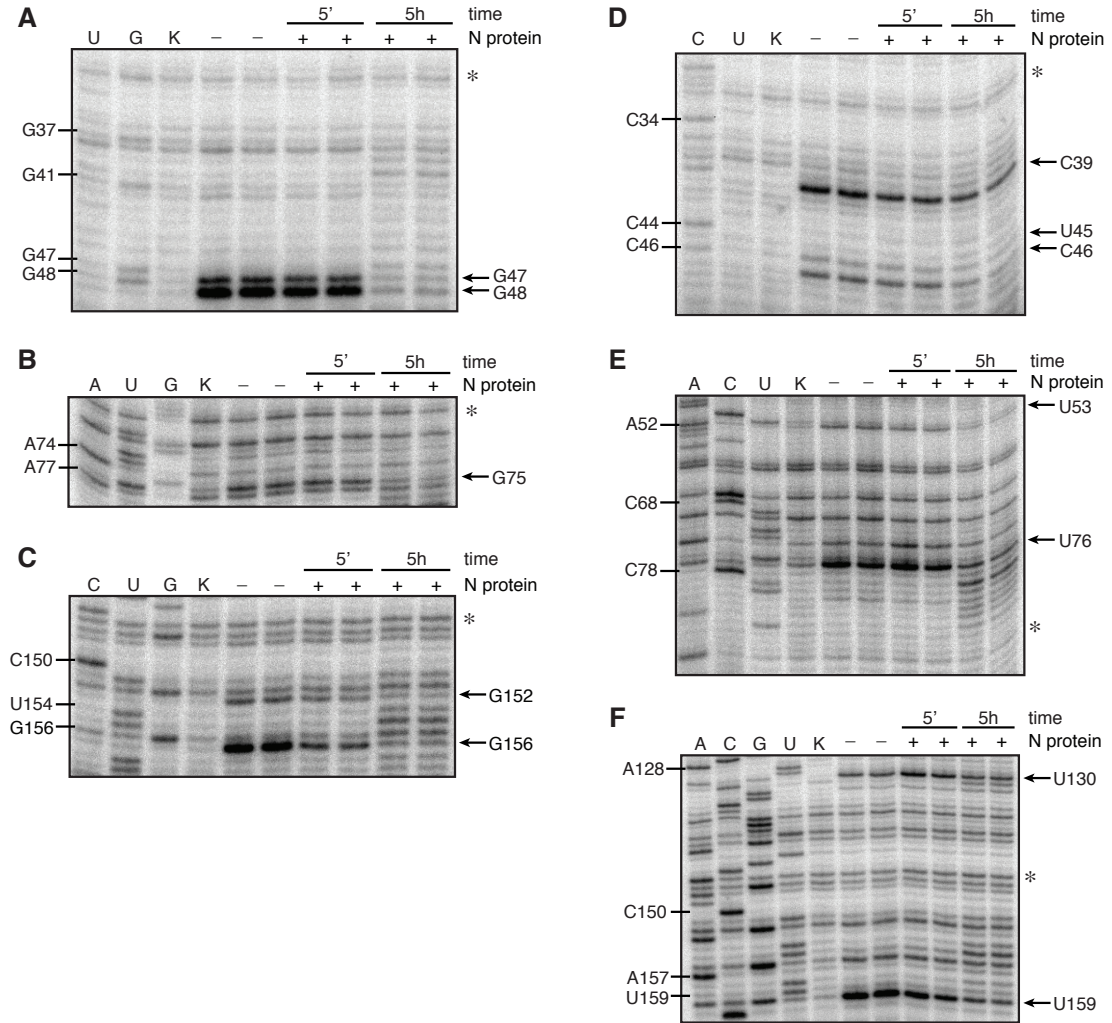
Figure 4 - 2: RNase T1 and RNase A structure probing analysis of an N-RNA binding interaction.
Lanes designated A, C, G and U are sequencing lanes. K lanes are primer extensions of MBE59/5S untreated with RNases. Lanes 5' and 5h indicate the length of time that N and RNA were incubated together before probing. Arrows indicate reactive nucleotides and asterisks (*) indicate bands used to normalize band intensities for quantification purposes as described in materials and methods. A-C. Primer extension analysis of MBE59/5S probed with RNase T1 in duplicate. D-F. Primer extension analysis of MBE59/5S probed with RNase A in duplicate.

After gaining insight into the molecular details of an N-RNA binding event, we

were interested in studying how the small molecule, suramin, identified in a high-

throughput drug-screening assay conducted by our laboratory, disrupts an N-RNA

binding interaction (29).
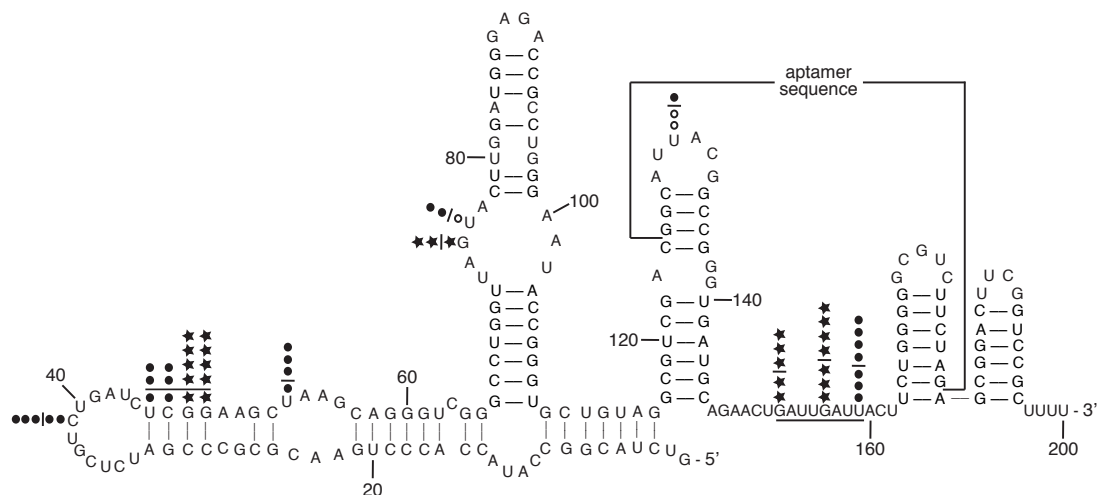
**Figure 4 - 3: Summary of the nucleotide reactivities superimposed on an Mfold
predicted secondary structure model of MBE59/5S.**
MBE59/5S is an RNA construct containing an aptamer RNA sequence with a GAUU motif attached to the
3' end of the human 5S rRNA sequence. The GAUU motif located within the 55-nucleotide long aptamer
RNA sequence is underlined. Circles represent RNase A probing of cytidine and uridine residues and stars
represent RNase T1 probing of guanosine residues. Open symbols represent enhanced reactivity and closed
symbols represent a decrease in reactivity. The number of symbols next to a residue indicate the relative
level of reactivity of the residue after incubation with N for either 5 minutes (symbols before bar) or 5
hours (symbols after bar) when compared to the RNA alone control. Band intensities were quantitated
using GraphPad Prism software and the number of symbols used is based on the percent change in band
intensity compared to the RNA alone control: 0-25% = 1 symbol, 26-50% = 2 symbols, 50-75% = 3
symbols, 76-100% = 4 symbols.

C. Suramin inhibits RNA binding to wild-type and mutant N protein

A fluorescence polarization experiment was performed to compare the ability of wild

type (WT) and a mutated version of RVFV N that lacks the amino-terminal 33 amino

acid residues important for N oligomerization (Δ33) to bind RNA. A fluorescently

labeled (3'-FAM) RNA previously shown to bind wild type RVFV N was incubated in

binding buffer with varying concentrations of WT and Δ 33 N at 30°C for 1 hour and

fluorescence polarization (FP) measured. Titration of WT N with FAM-RNA gave an

apparent K_d of 1.2 μM and a Hill coefficient of 2.2. Titration of the Δ33 mutant version

95

of N gave an apparent K_d of 9.0 μM and a Hill coefficient of 1.4 (Fig. 4-4A). The 7.5

fold decrease in the apparent K_d of WT N when compared to the Δ33 mutant shows that

removing the N-terminal arm decreases the ability of N to bind RNA. The Hill

coefficients for both WT and mutant N proteins were greater than 1 indicating that N

binds RNA in a positively cooperative manner. The N-terminal arm may play an

important role in the ability of N to oligomerize because the WT N protein exhibited a

steeper binding curve and larger Hill coefficient than the Δ33 mutant. Next we

performed competitive binding experiments by incubating a constant amount of N (15

μM of WT; 50 μM of Δ33) with 10 nM RNA at room temperature (RT) for 1 hour.

Increasing concentrations of suramin (524 pM-588 μM) were added and the reaction

incubated for one more hour at RT. FP measurements were taken and the results showed

suramin competed with aptamer RNA for N binding. The EC_{50} value of suramin for WT

N was approximately 3 fold higher than for the Δ33 mutant protein (22.3 μM for WT vs.

7.1 μM for Δ32) (Fig. 4-4B). These data show that the absence of the N-terminal arm

increases the ability of suramin to disrupt an N-RNA binding interaction and suggest that

the N-terminal arm may play an important role in stabilizing the RNP complex.
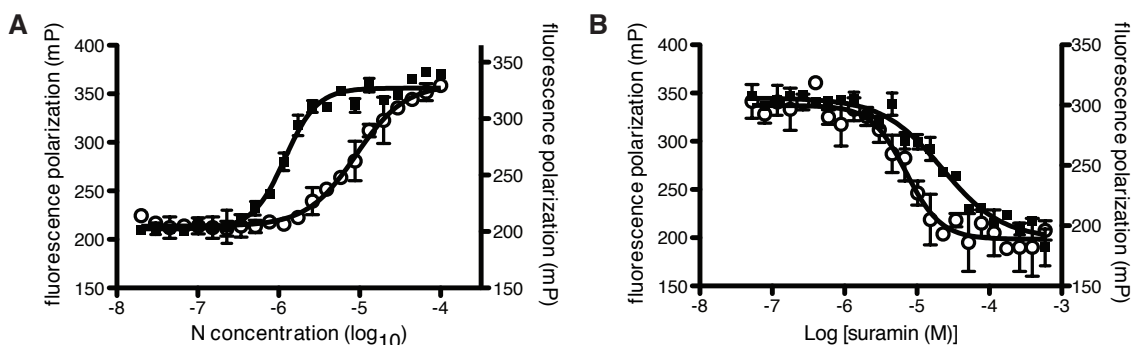
Figure 4 - 4: Analysis of the effects of suramin on WT and Δ33 N-RNA binding interactions using fluorescence polarization.
A. Binding profile for the association of WT (filled squares) and Δ33 mutant (open circles) N with fluorescently (FAM) labeled RNA. B. A competition assay where a fixed concentration of WT (filled

squares) and Δ33 mutant (open circles) N was incubated with FAM-RNA. Varying concentrations of suramin was added and the FP signal was plotted versus suramin concentration. The results are the average of two independent experiments.

D. Suramin inhibits a specific N-RNA binding interaction

Our previous results have shown that N specifically recognizes and binds to GAUU

motifs on selected aptamer RNAs. RNA structure probing experiments were conducted

(as described above) in the presence of varying concentrations of suramin (0.8 – 800 µM)

to determine if suramin is able to inhibit this specific binding interaction. Consistent with

the results obtained previously, G156 and G152 in the GAUU motif were reactive in the

absence of N and became strongly protected when N bound the MBE59/5S. When

suramin was added to the binding reaction the protective effect of N was abolished and

G156 and G152 regained their susceptibility to cleavage by RNase T1 (Fig. 4-5). These

data suggest that suramin disrupts the ability of N to specifically recognize and bind its

cognate RNA. Next we sought to determine if suramin was able to disrupt non-specific

N-RNA interactions that are important for the formation of high molecular weight RNP

complexes.
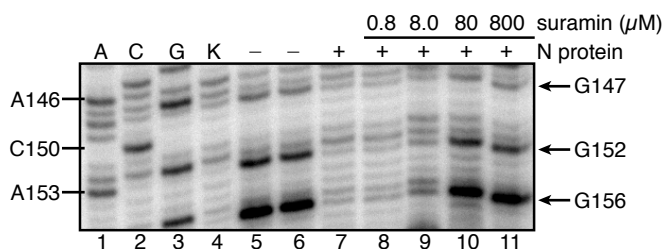
Figure 4 - 5: RNase T1 probing analysis of the effects of suramin on an N-RNA binding interaction.
RNase T1 structure probing followed by primer extension was used to visualize the effects of increasing concentrations of suramin (0.8 – 800 µM) on the ability of N to bind the GAUU motif of MBE59/5S. Lanes A, C and G are sequencing lanes, the K lane is a primer extension of untreated MBE59/5S and arrows indicate reactive guanosine nucleotides.

97

E. Suramin disrupts formation of high molecular weight N-RNA complexes

Purification of recombinant WT N protein (31 kDa) from bacteria results in a heterogeneous population of N dimers and multimers in a complex with *E. coli* RNA. Gel filtration chromatography yielded two protein/RNA fractions of ~153 and 90 kDa apparent molecular weight based on elution volumes from a Superose 12 (S-12) column that was calibrated with protein standards (Fig. 4-6A). The calibrated S-12 column was used to compare the multimeric state of wild type N in the presence and absence of suramin. Our results showed that when N was incubated with increasing concentrations of suramin higher molecular weight ribonucleoprotein complexes were disrupted, resulting in the formation of lower molecular weight complexes (Fig. 4-6A). The peak at 10.7 mL in the absence of suramin corresponded to a molecular weight of approximately 90 kDa and is likely two molecules of N bound to *E.coli* RNA. When 10 mM suramin was added this peak decreased in size significantly and a peak at 14.8 mL appeared that corresponded to a molecular weight of ~37 kDa (Fig. 4-6A). To compare the amount of RNA in higher molecular weight N-RNA complexes vs. lower molecular weight complexes formed by incubation with 10 mM suramin, fractions of N were collected, phenol/chloroform extracted and the amount of RNA in each fraction determined by denaturing gel electrophoresis followed by densitometry. Next, the protein concentration was plotted against the elution volume of the gel filtration column, and the relative amounts of co-purifying RNA was observed at each point. Large amounts of RNA were present in the fractions that corresponded to the 90 kDa peak, however, the 37 kDa peak formed during incubation with 10 mM suramin was devoid of RNA (Fig. 4-6B and C). These data show that suramin is able to displace the RNA from the higher molecular

weight N-RNA complexes resulting in the formation of a lower molecular weight species

that could be an N monomer bound to suramin.
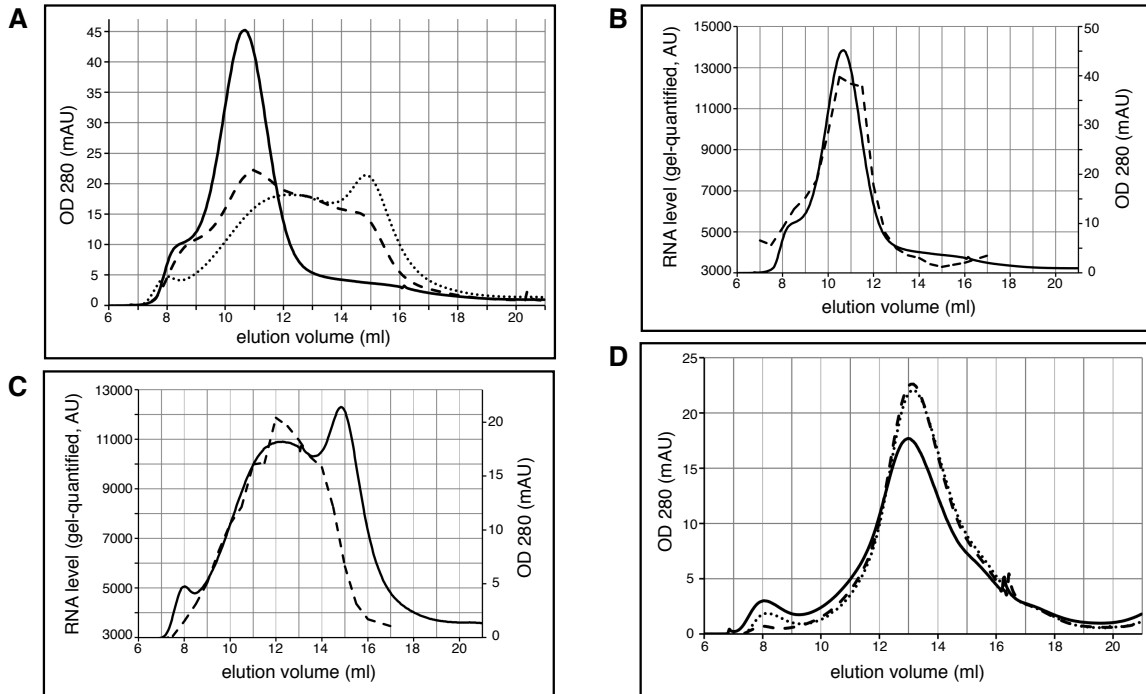
Figure 4 - 6: Characterization of the effects of suramin on the multimeric state of N using gel filtration chromatography.
A. Gel filtration profile of wild type (WT) N under the following conditions: no suramin, solid line; 1 mM suramin, dashed line; 10 mM suramin, dotted line. B. Analysis of the RNA content of fractionated WT N in the absence of suramin. RNA level (dashed line) and protein concentration (solid line) are plotted against the elution volume of the gel filtration column. C. Analysis of the RNA content of fractionated WT N after incubation with 10 mM suramin. To compare the amount of RNA in high molecular weight N-RNA complexes versus lower molecular weight N-RNA complexes formed by incubation with 10 mM suramin, RNA level (dashed line) and protein concentration (solid line) are plotted against the elution volume of the gel filtration column. D. Gel filtration profile of RNA binding (-) mutant N under the following conditions: no suramin, solid line; 1 mM suramin, dashed line; 10 mM suramin, dotted line.

To determine whether the observed shift to a lower molecular weight species is

caused by suramin blocking the RNA binding site of RVFV N protein or inhibiting N-N

interactions required for multimerization, we tested the ability of suramin to disrupt a

dimer structure formed by a mutated N that is unable to bind RNA. The predominant

peak for the RNA binding (-) mutant corresponded to a molecular weight of 54 kDa and

likely represents an N dimer without RNA (Fig. 4-6D). The gel filtration profile of the

mutated N was identical in both the presence and absence of suramin, suggesting that suramin was unable to disrupt the protein–protein interactions important for the formation and stabilization of the N dimer structure (Fig. 4-6D). Together, these data suggest that suramin disrupts formation of high molecular weight N-RNA complexes by binding directly to and obstructing the RNA binding cleft. These results illustrate the ability of suramin to block N-RNA binding interactions *in vitro* and prompted us to next test the efficacy of suramin as an antiviral agent in cell culture.

F. Suramin inhibits RVFV replication in cell culture

Plaque assays were conducted to measure the ability of suramin to inhibit the production of viable virus over the course of several days. Human 293 cells were infected at a low multiplicity of infection (MOI 0.1) in the presence or absence of suramin and the number of infectious viral particles produced was quantitated using a plaque assay. The cytotoxicity of suramin was measured using resazurin fluorescence. The metabolic activity of living cells converts resazurin into a fluorescently active compound that can be detected at 490 nm using a microplate reader. Suramin (100 μM) decreased the number of plaque forming units per mL by approximately 1 log at 2 and 3 days post infection (dpi) and was not cytotoxic to human 293 cells (Fig. 4-7A-C).
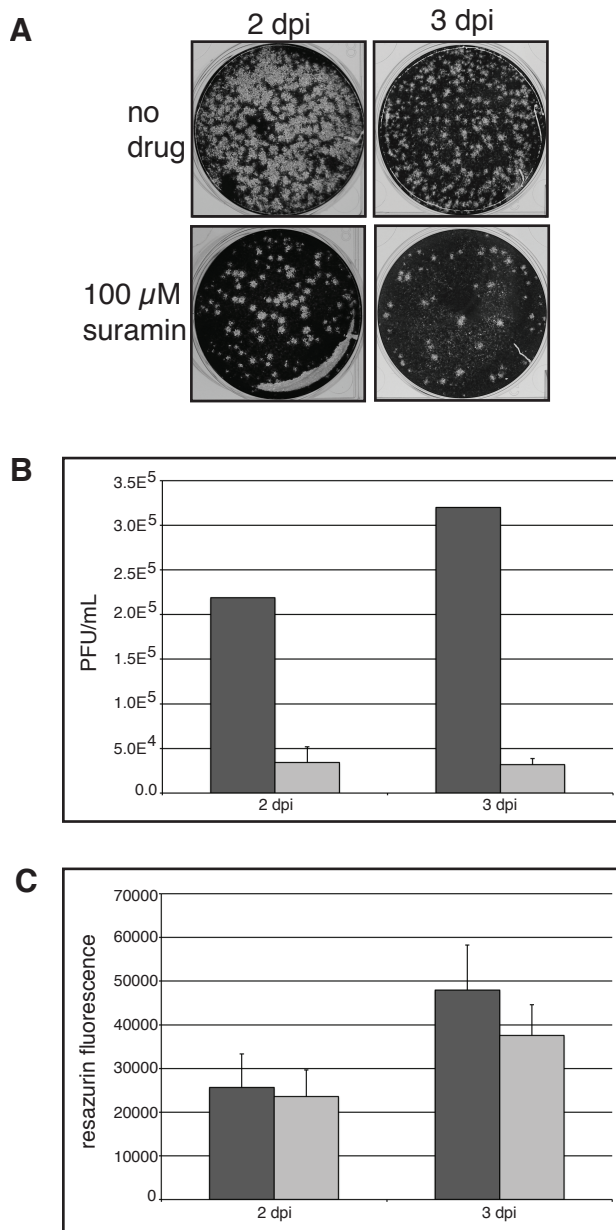
Figure 4 - 7: Analysis of the effects of suramin on virus yield and cell viability in human cell culture.

A. Plaque assay to determine the number of infectious vial particles produced from cells treated with either DMSO (no drug) or 100 μM suramin at two and three days post infection. B. Quantification of the plaque assay in Fig. 4-7A. Cells were treated with either DMSO (dark grey fill) or 100 μM suramin (light grey fill). C. Cytotoxicity assay to measure the effects of suramin on cell viability. Resazurain was added to wells treated with either DMSO (dark grey fill) or 100 μM suramin (light grey fill) and the fluorescence measured at 490 nm.

G. Suramin blocks both entry and post-entry steps in the RVFV lifecycle

Time of addition experiments were performed to determine if suramin blocks an early

and/or late step in the virus lifecycle. First, human 293 cells were treated with

ammonium chloride (12 mM), a known potent blocker of viral entry, either before or at

varying time points after infection to determine the amount of time required for RVFV

MP-12 to enter cells. We found that at one-hour post infection ammonium chloride was

no longer able to block infection, suggesting that entry is complete within one hour (data

not shown). Next, cells were treated with ammonium chloride or suramin either before or

after the virus (MOI 1.0) entered cells. Supernatants of infected cells were harvested the

next morning (15 hours post infection) and the number of infectious viral particles

produced was quantitated using a plaque assay. The results showed that 50 µM suramin

added before infection effectively blocked viral replication, but when added one-hour

post infection, the inhibitory effect was significantly diminished (Fig. 4-8A and B).

However, if a higher concentration of suramin was used (100 µM), the drug was able to

effectively inhibit replication when added either before or after virus entry into cells (Fig.

4-8A and B). The concentrations of ammonium chloride and suramin used in these

experiments had no effect on cell viability (Fig. 4-8C). These data suggest that suramin

is able to prevent RVFV replication at multiple stages of the virus life cycle. The ability

of the virus to enter cells is clearly inhibited by suramin, as well as one or more steps

post-entry that have yet to be determined.
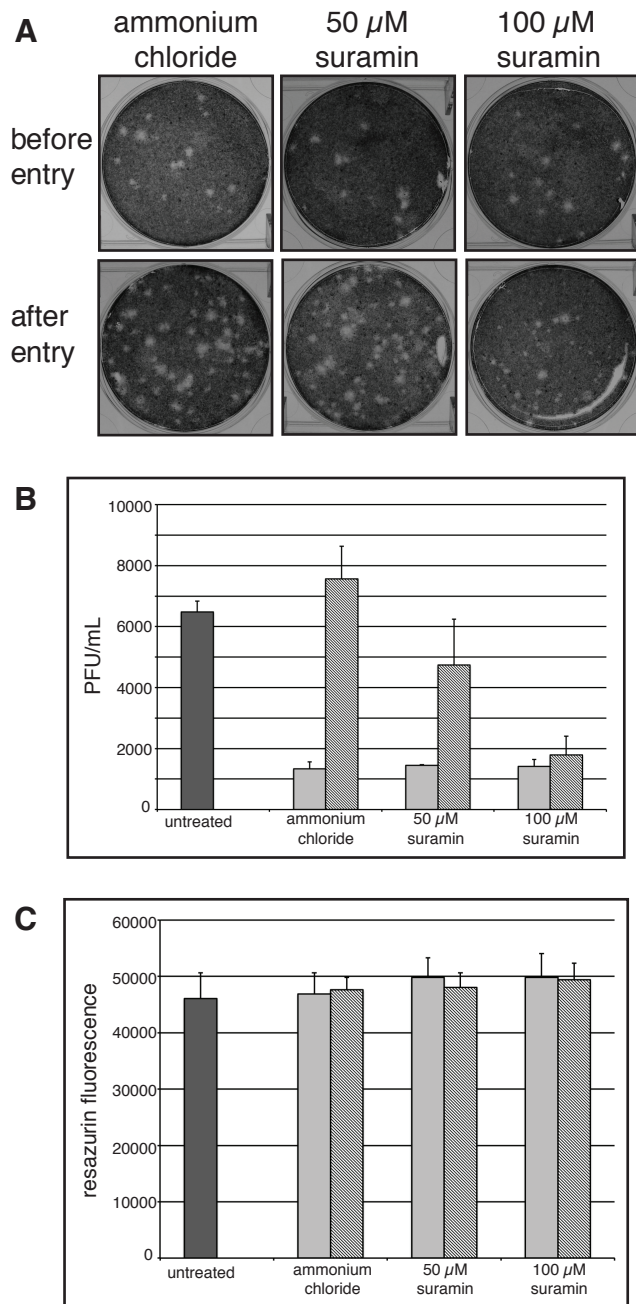
Figure 4 - 8: Time-of-addition assays to determine the stage(s) of the virus life cycle inhibited by suramin.
A. Plaque assay showing the effects of drug concentration and time of addition on the ability of suramin to inhibit viral replication. Cells were either pretreated (before entry) with suramin or ammonium chloride or treated one hour post infection (after entry) with suramin or ammonium chloride. B. Quantification of the plaque assay in Fig. 4-8A. Cells were treated with either DMSO ("untreated"; dark grey fill), pretreated with suramin or ammonium chloride (light grey fill) or treated one hour post infection with suramin or ammonium chloride (diagonal fill). C. Cytotoxicity assay to measure the effects of suramin and ammonium chloride on cell viability. Resazurain was added to wells treated with either DMSO ("untreated"; dark grey fill), pretreated with suramin or ammonium chloride (light grey fill) or treated one

hour post infection with suramin or ammonium chloride (diagonal fill) and the fluorescence measured at 490 nm.

IV. Discussion

In order for N to preferentially encapsidate full-length viral RNA during replication, N must interact with RNA in both a specific and non-specific manner. Interfering with either of these N-RNA binding modes could significantly abrogate viral replication. In the present study, we explore the nature of the RVFV N-RNA binding interaction, beginning with the initial recognition event followed by subsequent binding of N monomers in a sequence independent manner, resulting in the formation large, stable ribonucleoprotein (RNP) complexes. We demonstrate that suramin inhibits the initial specific N-RNA binding event and is capable of dissociating the high molecular weight N-RNA complexes. Furthermore, we demonstrate the therapeutic potential of suramin by showing that it functions well as an inhibitor of RVFV *in vitro* and inhibits viral replication in human cell culture.

In support of a biphasic RNA binding model, our experiments showed that N exhibited positive cooperativity when binding to an RNA molecule, resulting in the formation of high molecular weight N-RNA complexes (Figs. 4-1A and 4-4A). Over time, these complexes matured and prebound RNA became increasingly resistant to displacement by competitor RNA (Fig. 4-1B). RNP complexes are also known to be resistant to degradation by ribonucleases and high salt treatment (94). A high resolution structure of N multimers bound non-specifically to RNA shows that the RNA bases are sequestered in a deep binding groove that is contiguous along adjacent N monomers; this arrangement could explain the inability of competitor RNA to displace prebound RNA (100). Removal of 33 residues from the amino-terminal domain of N (Δ33 mutant)

decreased the cooperativity of N binding as well as the affinity of N for RNA by approximately 7.5 fold compared to WT N (Fig. 4-4A). These results are consistent with high-resolution structures that suggest that the N-terminus plays an important role in N oligomerization. N forms ring shaped multimers in which the flexible N-terminal arm wraps around the solvent side of an adjacent subunit, where it binds to a hydrophobic groove (32, 100).

We propose that RNA encapsidation by N is triggered by an initial specific binding event and previously reported that N specifically recognizes and binds appropriately-positioned GAUU sequences on aptamer RNA molecules (30). Time-resolved RNA structure probing experiments allowed us to observe sequential N binding to RNA and showed that N first bound and protected G and U residues located within a GAUU motif from cleavage by RNases. At the same time point, bases adjacent to the initial binding site exhibited a slight increase in reactivity (Fig. 4-3). These enhancements could be a result of the binding of the first N monomer causing a secondary structural rearrangement that prepares the adjacent binding site for another molecule of N to bind. Sequential binding of N monomers onto our RNA construct was also observed. RNA bases at the N-RNA recognition site near the 3' end of the RNA molecule were protected first and bases toward the 5' end were protected later (Fig. 4-3). As the complex matured over time, global protection of all reactive RNA bases was observed (Fig. 4-3). These results support a model of RNP formation that involves a specific RNA binding event initiating cooperative binding between N monomers, which then switch to a non-specific mode of RNA binding that allows N to coat the viral genome.

Interestingly, during replication encapsidation of viral RNA by N has been proposed to occur in the 5' to 3' direction (87, 109). We observed sequential binding in the opposite (3' to 5') direction, suggesting either that N is able to oligomerize in both directions or that binding occurred first at a high affinity site and subsequently at sites further upstream. That N exhibits disparate binding behaviors on different RNAs has been demonstrated in La Crosse virus infected cells (44). During the later stages of a La Crosse viral infection (24-48 h.p.i.) as the concentration of N in the host cell increases, viral mRNAs become encapsidated (44). Viral mRNAs are considered secondary targets because at low concentrations of N, only genomes and antigenomes are complexed with N. At higher concentrations, viral but not cellular mRNAs are associated with N protein (44). This suggests that the viral mRNA has a lower-affinity N recognition motif(s), but at a critical concentration of N, the mRNA can be bound first specifically and then cooperatively by N. It is thought that encapsidation of viral mRNA is a mechanism to regulate viral protein synthesis because encapsidated mRNAs cannot be translated by the ribosome.

Suramin was identified in a high-throughput screening assay for compounds that inhibit a specific RVFV N-RNA binding interaction (29). Suramin is a symmetric, polyanionic compound that contains two naphthalene-1,3,5-trisulfonic acid head moieties and has been used successfully in the past to treat parasitic infections, specifically trypanosomiasis and human onchocerciasis (4, 47). Here we observed that although competitor RNA was unable to disrupt mature N-RNA complexes, suramin was able to penetrate and disrupt both WT and Δ33 mutant N-RNA complexes (Figs. 4-1B; 4-4B; 4-5; 4-6A). Interestingly, a significantly lower concentration of suramin was required to

disrupt the Δ33 mutant complex (Fig. 4-4B). This could be because the N-terminal arm not only facilitates the formation of multimers onto an RNA molecule, but helps stabilize and strengthen the RNP complex as well. We reasoned that suramin could exert this inhibitory effect by either occluding the RNA binding site directly or inhibiting protein-protein interactions that are required for assembly of the viral nucleocapsid. Because we found that N purified by gel filtration chromatography in the presence of suramin was diminished in RNA content, and because suramin was unable to disrupt a dimer structure formed by a mutated N that is unable to bind RNA, we conclude that suramin binds directly to and obscures the RNA binding cleft.

Suramin, or derivatives thereof, may be useful in the development of broad spectrum antiviral agents. A recent crystal structure of N from a related phlebovirus (severe fever with thrombocytopenia syndrome virus, SFTSV) with a bound suramin showed that the head of the suramin molecule penetrated deeper into the RNA binding slot than an RNA molecule (56). The ability of suramin to penetrate further into the RNA binding groove could explain why we found that suramin is able to eject prebound RNA. Suramin inhibits Norovirus RNA dependent RNA polymerase by its negatively charged naphthalenetrisulfonic acid groups binding to a positively charged cleft near the active site and likely preventing incoming NTPs to access to the active site (66). It is also thought that suramin binds retroviral reverse transcriptase at or near the polymerase active site through interactions between its negatively charged head groups and basic amino acid residues (86).

Suramin was not only effective at disrupting RVFV N-RNA binding interactions *in vitro*, it also successfully inhibited viral replication in human cell culture (Fig. 4-7).

Suramin blocked virus entry into cells (or another early step) as well as a subsequent step(s) in the viral replication cycle (Fig. 4-8). Polyanionic compounds are known to interfere with the virus adsorption process. It is also possible that suramin in the media could penetrate the viral particle and begin to disrupt the RNP complex and/or the RNA dependent RNA polymerase before it enters the cell. Similar to the results obtained here, previous studies show that in addition to the anti-RT activity mentioned above, suramin prevents HIV-1 from binding to cells possibly by interacting with the viral glycoprotein gp120 (26, 107).

This study demonstrates that suramin is not only a useful tool to study RVFV N-RNA binding interactions, it may form the basis for future broad spectrum antiviral therapeutic strategies, due to its ability to inhibit RVFV replication at multiple steps in cell culture, and its ability to inhibit replication of several diverse viruses. Furthermore, continuing to increase our understanding of the molecular details of viral RNA encapsidation could ultimately aid in the design of more effective drugs targeted to viral nucleocapsid proteins.

CHAPTER 5: GENERAL DISCUSSION AND FUTURE

DIRECTIONS

The overall goals of our research were to better understand how Rift Valley fever virus (RVFV) nucleocapsid (N) protein interacts with RNA, to develop a high-throughput screening (HTS) assay for identification of compounds that inhibit a N-RNA binding interaction and to test the efficacy of potential anti-viral drug compounds during RVFV infection. We have identified interesting new sequences and secondary structural elements of RNA that are specifically recognized and bound by N. The biological relevance and function of these binding determinants during RVFV infection are currently unknown but warrant further discussion and study. Implementation of a HTS assay developed in our laboratory allowed us to identify small molecules and natural product extracts that inhibit N-RNA binding interactions *in vitro* and RVFV replication in cell culture. The compounds and extracts identified in this study represent potential antiviral therapeutic strategies as well as useful as tools to study RVFV N-RNA binding interactions. Increasing our understanding of how RVFV N recognizes and binds to its cognate RNA will aid in the design of more potent and less cytotoxic compounds targeted to viral nucleocapsid proteins.

I. Rift Valley fever virus nucleocapsid protein-RNA binding interactions

The majority of previous research aimed at elucidating requisite RNA features for a specific N-binding event has focused on the 5' and 3' non-coding ends of the viral RNA. The terminal nucleotide sequences are highly conserved among bunyaviruses of the same genus and base-pairing of the terminal nucleotides results in the formation of stable panhandle structures and allows the genome and antigenome to circularize. It is thought that the sequences and/or panhandle structures at the ends of the genome allow viral nucleocapsid protein to recognize and preferentially bind to its own RNA (72, 76, 78).

This specific binding may be important for initiation of viral transcription, replication and packaging. However, N is involved in other steps during an infection that require it to interact with non-terminal sequences and/or secondary structural elements of the viral genome and antigenome, as well as viral mRNA. For example, hantavirus N preferentially binds RNA containing a 5' m7G cap when compared with uncapped RNA or the cap analog, 2'-O-methyl G (73). This cap-binding activity plays an important role in allowing N to augment production of viral protein in *trans* by acting as a translation initiation factor. In fact, N is able to replace the eIF4F initiation complex by mimicking the cap binding activity of eIF4E, directly recruiting the 43S pre-initiation complex to the mRNA (a function normally performed by eIF4G) and the RNA chaperone activity of N obviates the need for the RNA helicase eIF4A (77, 89). Also, during an RNA virus infection a correlation between an increased concentration of N in the host cell and the switch from transcription of viral mRNA to replication of the viral genome has been observed. The mechanism underlying this switch is poorly understood but it has been proposed that N acts as an anti-terminator during mRNA synthesis and promotes full-length complementary RNA synthesis (118). The events described above likely involve binding interactions between N and viral RNA that have not yet been characterized.

To gain a better understanding of the RNA sequence and secondary structural elements specifically recognized by RVFV N during an RNA-binding event, we used an *in vitro* selection technique (SELEX) to evolve RNAs that bound with high affinity (aptamer RNAs) to N. Analysis of these aptamer RNAs allowed us to identify RNA features required for N binding without any biased assumptions about what sequences or structures constitute favorable binding targets. Analysis of the nucleotide distribution of

RNA sequences from the last two rounds of the selection showed enrichment for both guanosine and uridine residues when compared to sequences from round one. Consistent with this observation, GAUU and pyrimidine/guanine motifs were found in many of our aptamer RNAs. Mutational analysis verified that the GAUU motif and, to a lesser extent the pyrimidine/guanine motif, are important for N binding to aptamer RNA. We conducted RNA structure probing experiments that allowed us to observe N binding to a GAUU motif in real time. When N first bound the RNA it protected G and U residues located within the motif from cleavage by RNases. At the same time point, bases adjacent to the initial binding site and located within stem loop structures exhibited an increase in reactivity. These enhancements in reactivity could be a result of N disrupting adjacent stem loop structures as it binds to the GAUU motif. This type of activity could be important for relieving secondary structural elements on viral RNA that could decrease the processivity of a polymerase and/or impede the progress of the ribosome. For example, N binding to and unwinding mRNA secondary structure during translation could increase the speed and efficiency of viral protein synthesis. BLAST analysis revealed multiple regions of sequence similarity between RNA aptamers and the coding region of the viral genome and antigenome, which supports the idea that N interacts with non-terminal viral RNA sequences during a replication cycle.

As previously mentioned, a pyrimidine/guanine motif was present in many of our aptamer RNAs. The pyrimidine portion of this motif exhibited some sequence variability but it always ended with a stretch of five G residues. For example, MBE15.12 had a UUUCUGGGGG motif while sequencing of MBE16.40 revealed a CUCUGGGGG motif. Interestingly, a UG-rich RNA element is known to be involved in regulating

transcription termination in RVFV (53). This UG-rich motif (UGGGGUGGUGGGGU) is located directly upstream of a conserved pentanucleotide (CGUCG) sequence and is present on the S and M segments of viral mRNA (53). Although the mechanism of transcription termination is not well understood, it likely involves a viral or host cellular factor interacting with these RNA elements. It has also been proposed that annealing between the nascent mRNA and template during transcription is involved in causing the RNA-dependent RNA polymerase to terminate. In bunyaviruses, drugs that inhibit translation (such as cycloheximide) also inhibit mRNA synthesis. Specifically, in La Crosse viruses addition of cycloheximide did not affect the rate of mRNA initiation, but caused all S segment mRNA transcripts to terminate prematurely at nucleotide 175. However, the polymerase could be made to read through nucleotide 175 to the end of the RNA molecule in the absence of protein synthesis by replacing guanosines residues with inosine (7, 8). These data provide support for the idea that premature termination is due to base-pairing of the nascent mRNA chain with the template and that concurrent translation of viral mRNA in the cytoplasm of the host cell prevents the formation of ds-RNA and premature termination. Experiments conducted in Sendai virus (a non-segmented negative-sense RNA virus) show that as the concentration of N increases in the host cell during an infection, a switch from transcription to full-length cRNA synthesis occurs because at a certain critical concentration, N begins to encapsidate viral RNA as it is being synthesized by the RNA-dependent RNA polymerase (118). We hypothesize that in our system, RVFV N could function as an anti-terminator during transcription and facilitate the switch to full-length genome replication by specifically recognizing and binding to the UG-rich motif on the M and S segments of viral mRNA.

This would prevent base-pairing between the nascent mRNA chain and the template and promote full length complementary RNA synthesis. If this were true, one would expect to see a full-length antigenomic RNA with a 5' 7mG cap present in RVFV infected cells. Interestingly, small amounts of an S RNA transcript that starts like an mRNA but ends like the antigenome have been detected in La Crosse virus infected cells (98) and could represent an intermediate RNA molecule formed during the switch from transcription to genome replication.

When RVFV N binds RNA it is unclear whether RNA sequence, secondary structure, or both are important for determining the specificity of binding. Experiments conducted using Sin Nombre hantavirus N suggest that formation of a panhandle structure allows selective encapsidation of cognate RNA molecules during an infection (76). However hantavirus N also binds the 5' end of single-stranded RNA with higher affinity than the 3' end suggesting that sequence as well as secondary structure play a role in the ability of nucleocapsid protein to specifically recognize viral RNA (75). We propose that the GAUU sequence is recognized and bound by N and therefore it is important that the motif be located in a single-stranded region of an RNA molecule. However, this hypothesis has not been rigorously tested and it is likely that secondary structural elements are important binding determinants for other aptamer RNAs identified during our *in vitro* selection. Recently, a collaborator (Doug Raiford) developed an algorithm to mutate the sequence but not the mfold-predicted secondary structure of aptamer RNAs. We are currently in the process of synthesizing RNAs that will be used to compare the binding affinity of N for WT and mutant aptamer RNA molecules. These

studies will ultimately allow us to measure the individual contribution of a sequence or secondary structure to a RVFV N-RNA-binding event.

A recent crystal structure of single-stranded RNA encapsidated by RVFV N shows the ribonucleoprotein complex existing as a hexamer, pentamer or tetramer. Analysis of this structure revealed an RNA-binding groove that contains several highly conserved basic amino acid side chains and the N-terminal domain of an N monomer facilitates oligomerization by binding to a hydrophobic pocket on an adjacent subunit. The sugar phosphate backbone of the RNA is oriented toward the solvent exposed side of the RNA binding groove and the RNA bases, which are inserted into the slot, are involved in hydrophobic and base-stacking interactions with N (100). The mode of binding represented by this crystal structure is considered non-specific because N is only involved in H-bond interactions with the sugar phosphate backbone of RNA and is not recognizing a specific H-bond pattern formed by the nitrogenous bases. Structures of N from other negative-sense RNA viruses, including influenza and vesicular stomatitis virus (VSV) have been solved. Similar to the RVFV structure, the RNP oligomers form high-order ring structures and oligomerize by N and/or C terminal arm-like extensions interacting with neighboring monomers. However, the orientation of the RNA bases within the RNA-binding cleft of influenza and VSV RNP complexes is different. Influenza virus N coats the RNA sugar phosphate backbone and the bases are oriented toward the solvent (123). The crystal structure of VSV N shows that the RNA bases adopt a unique conformation where some of the bases are facing toward and others away from the solvent (42). The nitrogenous bases of VSV N encapsidated RNA exhibit varying

degrees of reactivity when chemically probed and these data are consistent with the high-resolution structure model that shows some of the bases are solvent exposed (54).

RNA structure probing experiments conducted in our laboratory yielded some thought provoking results. As discussed previously, when RVFV N was incubated with aptamer RNA and then probed with RNases, initially N protected G and U residues present in a GAUU motif from cleavage. After the RNA was incubated with N for an extended period of time, all reactive nucleotides were protected. We observed sequential binding in the 3' to 5' direction, suggesting either that N is either able to oligomerize in both (5' to 3' and 3' to 5') directions or that binding occurred first at a high affinity site and subsequently at lower affinity sites further upstream. Unexpectedly, when RNA was incubated with N and then probed with DMS, enhancements in the reactivity of RNA bases were observed. Since the crystal structure of RVFV N shows the nitrogenous bases buried deep within the RNA-binding groove we expected to see a decrease in reactivity of the bases when N binds. However, our results suggest that when N binds RNA it positions the Watson-Crick face of some A and C residues in such a way that they become more available for DMS modification. When N encapsidates RNA it removes secondary structures and the RNA becomes single-stranded. Perhaps disruption of H-bond interactions that stabilize stem loop structures could increase the availability of some RNA bases for chemical modification. Also, not all RVFV N-encapsidated RNA bases are sequestered within the RNA binding groove. Each N subunit binds four RNA nucleotides in the RNA binding slot and 2 to 3 additional nucleotides at the subunit interface. It is possible that DMS is able to modify the bases in between the N monomers. If this were true we would expect to see a pattern emerge where every 5th, 6th

and 7th A or C residue is reactive; however, our probing results did not show this type of periodicity.

During replication as the viral RNA exits the RNA-dependent RNA polymerase it is sequentially bound by N in the 5' to 3' direction. However, in our experimental system we were only able to observe binding in the 3' to 5' direction. To evaluate whether sequential binding can occur in both directions, it would be useful to create an RNA molecule that contains a high affinity RVFV N-binding site flanked by 100 or so nucleotides of low affinity non-aptamer-like RNA sequence. To determine the directionality of N binding, a quench flow apparatus that incubates a binding reaction for a specified amount of time ranging from milliseconds to hours could be used to perform time-resolved RNA structure probing experiments. First, we would expect to see changes in reactivity in RNA bases that correspond to the central portion of the RNA molecule that contains the high affinity aptamer RNA sequence because this is where N should bind first. Then later we would expect RNA sequences upstream and/or downstream of the aptamer RNA sequence to exhibit changes in reactivity as N subsequently binds the lower affinity sites. Our previous results suggest that the initial N-RNA binding events occur rapidly and then over the course of several minutes to hours the ribonucleoprotein complex matures. Since the quench flow apparatus can take multiple snapshots of an N-RNA binding event that has been incubating for milliseconds or seconds, we could potentially observe N monomers binding one at a time to an RNA molecule. These data would allow us to observe RNA encapsidation in real time and deduce the directionality of N oligomerization.

DMS probing results suggest that as N binds, RNA bases become more available for chemical modification. Perhaps an increase in the reactivity of RNA bases is observed when small, hydrophobic chemical probes are used because they can penetrate into the RNA-binding groove and modify bases involved in hydrophobic interactions with N. It would be interesting to try RNA structure probing using kethoxal, a small molecule that modifies N1 and N2 of guanosine to see if the results are similar to those obtained using DMS. We could also directly compare the reactivity of G residues to a chemical probe vs. RNase T1 using kethoxal. Hydroxyl radical structure probing experiments could be conducted to see how well protected by N the RNA sugar phosphate backbone becomes over time. Hydroxyl radicals react with and degrade the ribose ring of the RNA and protein bound to the RNA backbone lowers the reactivity of the ribose rings. Completion of the proposed RNA structure probing experiments would allow us to further elucidate the solvent accessibility of various components of an RNA molecule during an N-RNA binding event. This information could be used to create a detailed model of a RVFV N-RNA encapsidation event.

II. High-throughput drug screening assay

We developed a fluorescence-based high-throughput drug screening (HTS) assay to test 26,424 chemical compounds for their ability to disrupt an N-RNA complex, and potentially inhibit viral replication. We found 10 unique compounds and 40 natural product extracts that were capable of disrupting the N-RNA complex *in vitro* and a subset displayed efficacy in cell culture. Transfection of RNA aptamers directed against N inhibited viral translation and replication in cell culture providing support for our

hypothesis that inhibition of N binding to RNA represents a viable antiviral therapeutic strategy.

Two unique compounds (Life Chemicals, 5406174 and suramin) were able to decrease RVFV replication in cell culture by approximately 1 log. Compound 5406174 contains a quinone, which is a reactive moiety and can modify cysteine residues. Some scientists label compounds that contain quinones as "PAINS" (Pan Assay Interference compounds) because they appear as hits in many high-throughput screens and question the specificity of compounds that are promiscuous hitters. However, quinone-containing compounds have been successfully developed into drugs and are common in anti-cancer therapy (5). Also, only one of the two quinones that were confirmed hits from our HTS screen decreased viral replication in cell culture, suggesting that compound 5406174 exhibits specificity for its target. Therefore, we believe that compound 5406174 should not be dismissed as a hit and intend to continue to study its mechanism of inhibition both *in vitro* and in human cell culture.

Suramin is a negatively charged compound that contains two naphthalene-1,3,5-trisulfonic acid head groups. It effectively inhibited RVFV replication in cell culture and was an easy to use tool to study N-RNA binding interactions biochemically because it is colorless, stable and water soluble. Analysis of the *in vitro* inhibitory mechanism of suramin showed that it was able to disrupt N-RNA interactions by binding directly to and obscuring the RNA binding cleft. We also discovered that suramin inhibits RVFV replication in cell culture by preventing the virus from entering cells and by blocking other subsequent steps of the virus life cycle. Suramin could be useful in the development of broad-spectrum antiviral therapies due to its ability to inhibit replication

of RVFV and several other viruses such as retroviruses. Although it is difficult to put an exact figure on the cost of creating a new drug, it is extremely expensive and estimates range from hundreds of millions to billions of dollars. However, since suramin has been used successfully to treat parasitic diseases, its relative safety in humans has already been proven and this could significantly decrease length of time and cost required to bring the drug to market.

We used a variety of biochemical assays to determine how suramin disrupts an N-RNA binding event and a similar strategy could be utilized to study the inhibitory mechanism(s) of the other nine unique compounds. We anticipate that compounds identified in our HTS assay inhibit N-RNA binding by either occluding the RNA binding site or inhibiting protein-protein interactions that are required for assembly of the viral nucleocapsid. To distinguish between these potential mechanisms, we could use fluorescence polarization to report on N-RNA binding in the presence or absence of drug compound, and compare WT vs. a mutated version of N that is unable to multimerize. We could also perform gel filtration on purified N protein in the presence and absence of compound to determine how addition of compound affects formation of N multimers. RNA structure probing could be used to determine whether a compound is capable of inhibiting a specific and/or non-specific N-RNA binding interaction. It would be extremely fortuitous to discover a compound that inhibits one mode of binding but not the other.

Lead natural product extracts were tested for cytotoxicity and their ability to inhibit viral replication in cell culture. The results were promising and showed that six extracts inhibited RVFV replication by 50% or more and were not cytotoxic to cells. Since

natural product extracts are a complex mixture of compounds, the active component(s) of the extract may be dilute. Therefore, once the functional component is isolated, the natural product could be an even more potent inhibitor of viral replication. Characterization of the active components(s) in the natural product extracts could potentially result in the discovery of novel anti-viral compounds invented by Mother Nature. Interestingly, at two and three days post infection natural product extract 6990 stimulated or increased virus yield. Although it would not be useful to pursue this extract as a potential anti-viral drug, it could be useful as a tool to study the mechanism of RVFV replication. Identification of the molecular target(s) of extract 6990 could help us determine specific virus and/or host factors that play an important role during the viral replication cycle. Microarray-based experiments could be conducted to compare the gene expression profiles of cells infected with RFVF in the presence and absence of extract 6990. Identification of genes that are up-regulated in the presence of extract 6990 could help determine specific genes or cellular pathways utilized by RVFV during an infection.

This research has increased our understanding of how viral nucleocapsid proteins recognize and bind their cognate RNA and rearrange RNA structure. We have also developed a high throughput drug-screening tool and established RVFV N as a promising drug target. Natural product extracts that could contain novel anti-viral compounds have been identified and we show that suramin and Life Chemicals 5406174 possess drug development potential.

CHAPTER 6: REFERENCES

1. **Al-Hazmi, M., E. A. Ayoola, M. Abdurahman, S. Banzal, J. Ashraf, A. El-Bushra, A. Hazmi, M. Abdullah, H. Abbo, A. Elamin, T. Al-Sammani el, M. Gadour, C. Menon, M. Hamza, I. Rahim, M. Hafez, M. Jambavalikar, H. Arishi, and A. Aqeel.** 2003. Epidemic Rift Valley fever in Saudi Arabia: a clinical study of severe illness in humans. Clin Infect Dis **36:**245-252.
2. **Alrajhi, A. A., A. Al-Semari, and J. Al-Watban.** 2004. Rift Valley fever encephalitis. Emerg Infect Dis **10:**554-555.
3. **Anderson, G. W., Jr., and J. F. Smith.** 1987. Immunoelectron microscopy of Rift Valley fever viral morphogenesis in primary rat hepatocytes. Virology **161:**91-100.
4. **Anderson, J., and H. Fuglsang.** 1978. Further studies on the treatment of ocular onchocerciasis with diethylcarbamazine and suramin. Br J Ophthalmol **62:**450-457.
5. **Baell, J. B., and G. A. Holloway.** 2010. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem **53:**2719-2740.
6. **Balkhy, H. H., and Z. A. Memish.** 2003. Rift Valley fever: an uninvited zoonosis in the Arabian peninsula. Int J Antimicrob Agents **21:**153-157.
7. **Bellocq, C., and D. Kolakofsky.** 1987. Translational requirement for La Crosse virus S-mRNA synthesis: a possible mechanism. J Virol **61:**3960-3967.
8. **Bellocq, C., R. Raju, J. Patterson, and D. Kolakofsky.** 1987. Translational requirement of La Crosse virus S-mRNA synthesis: in vitro studies. J Virol **61:**87-95.
9. **Billecocq, A., M. Spiegel, P. Vialat, A. Kohl, F. Weber, M. Bouloy, and O. Haller.** 2004. NSs protein of Rift Valley fever virus blocks interferon production by inhibiting host gene transcription. J Virol **78:**9798-9806.
10. **Bird, B. H., T. G. Ksiazek, S. T. Nichol, and N. J. Maclachlan.** 2009. Rift Valley fever virus. J Am Vet Med Assoc **234:**883-893.
11. **Borio, L., T. Inglesby, C. J. Peters, A. L. Schmaljohn, J. M. Hughes, P. B. Jahrling, T. Ksiazek, K. M. Johnson, A. Meyerhoff, T. O'Toole, M. S. Ascher, J. Bartlett, J. G. Breman, E. M. Eitzen, Jr., M. Hamburg, J. Hauer, D. A. Henderson, R. T. Johnson, G. Kwik, M. Layton, S. Lillibridge, G. J. Nabel, M. T. Osterholm, T. M. Perl, P. Russell, and K. Tonat.** 2002. Hemorrhagic fever viruses as biological weapons: medical and public health management. Jama **287:**2391-2405.
12. **Botros, B., A. Omar, K. Elian, G. Mohamed, A. Soliman, A. Salib, D. Salman, M. Saad, and K. Earhart.** 2006. Adverse response of non-indigenous cattle of European breeds to live attenuated Smithburn Rift Valley fever vaccine. J Med Virol **78:**787-791.
13. **Bouloy, M., and R. Flick.** 2009. Reverse genetics technology for Rift Valley fever virus: current and future applications for the development of therapeutics and vaccines. Antiviral Res **84:**101-118.

14. **Bouloy, M., C. Janzen, P. Vialat, H. Khun, J. Pavlovic, M. Huerre, and O. Haller.** 2001. Genetic evidence for an interferon-antagonistic function of rift valley fever virus nonstructural protein NSs. J Virol **75:**1371-1377.

15. **C.D.C.** 2000. Outbreak of Rift Valley fever-Saudia Arabia, August-October, 2000. MMWR Morb Mortal Wkly Rep **49:**905-908.

16. **Caplen, H., C. J. Peters, and D. H. Bishop.** 1985. Mutagen-directed attenuation of Rift Valley fever virus as a method for vaccine development. J Gen Virol **66 (Pt 10):**2271-2277.

17. **Caplen, H., C. J. Peters, and D. H. Bishop.** 1985. Mutagen-directed attenuation of Rift Valley fever virus as a method for vaccine development. J Gen Virol **66 (Pt 10):**2271-2277.

18. **Chapman, J., E. Abbott, D. G. Alber, R. C. Baxter, S. K. Bithell, E. A. Henderson, M. C. Carter, P. Chambers, A. Chubb, G. S. Cockerill, P. L. Collins, V. C. Dowdell, S. J. Keegan, R. D. Kelsey, M. J. Lockyer, C. Luongo, P. Najarro, R. J. Pickles, M. Simmonds, D. Taylor, S. Tyms, L. J. Wilson, and K. L. Powell.** 2007. RSV604, a novel inhibitor of respiratory syncytial virus replication. Antimicrob Agents Chemother **51:**3346-3353.

19. **Cheng, E., A. Haque, M. A. Rimmer, I. T. Hussein, S. Sheema, A. Little, and M. A. Mir.** 2011. Characterization of the Interaction between hantavirus nucleocapsid protein (N) and ribosomal protein S19 (RPS19). J Biol Chem **286:**11814-11824.

20. **Coetzer, J. A.** 1977. The pathology of Rift Valley fever. I. Lesions occurring in natural cases in new-born lambs. Onderstepoort J Vet Res **44:**205-211.

21. **Coetzer, J. A.** 1982. The pathology of Rift Valley fever. II. Lesions occurring in field cases in adult cattle, calves and aborted foetuses. Onderstepoort J Vet Res **49:**11-17.

22. **Coetzer, J. A., and K. G. Ishak.** 1982. Sequential development of the liver lesions in new-born lambs infected with Rift Valley fever virus. I. Macroscopic and microscopic pathology. Onderstepoort J Vet Res **49:**103-108.

23. **Collett, M. S., A. F. Purchio, K. Keegan, S. Frazier, W. Hays, D. K. Anderson, M. D. Parker, C. Schmaljohn, J. Schmidt, and J. M. Dalrymple.** 1985. Complete nucleotide sequence of the M RNA segment of Rift Valley fever virus. Virology **144:**228-245.

24. **Daubney, R. J., Hudson, J.R. and Garnham, P.C.** 1931. Journal of Pathology and Bacteriology **34:**545-579.

25. **Davies, F. G., K. J. Linthicum, and A. D. James.** 1985. Rainfall and epizootic Rift Valley fever. Bull World Health Organ **63:**941-943.

26. **De Clercq, E.** 1979. Suramin: a potent inhibitor of the reverse transcriptase of RNA tumor viruses. Cancer Lett **8:**9-22.

27. **de Haseth, P. L., and O. C. Uhlenbeck.** 1980. Interaction of Escherichia coli host factor protein with Q beta ribonucleic acid. Biochemistry **19:**6146-6151.

28. **de Rocquigny, H., V. Shvadchak, S. Avilov, C. Z. Dong, U. Dietrich, J. L. Darlix, and Y. Mely.** 2008. Targeting the viral nucleocapsid protein in anti-HIV-1 therapy. Mini Rev Med Chem **8:**24-35.

29. **Ellenbecker, M., J. M. Lanchy, and J. S. Lodmell.** 2012. Identification of Rift Valley fever virus nucleocapsid protein-RNA binding inhibitors using a high-throughput screening assay. J Biomol Screen **17:**1062-1070.

30. **Ellenbecker, M., L. Sears, P. Li, J. M. Lanchy, and J. Stephen Lodmell.** 2012. Characterization of RNA aptamers directed against the nucleocapsid protein of Rift Valley fever virus. Antiviral Res **93:**330-339.

31. **Ellis, D. S., P. V. Shirodaria, E. Fleming, and D. I. Simpson.** 1988. Morphology and development of Rift Valley fever virus in Vero cell cultures. J Med Virol **24:**161-174.

32. **Ferron, F., Z. Li, E. I. Danek, D. Luo, Y. Wong, B. Coutard, V. Lantez, R. Charrel, B. Canard, T. Walz, and J. Lescar.** 2011. The hexamer structure of the rift valley Fever virus nucleoprotein suggests a mechanism for its assembly into ribonucleoprotein complexes. PLoS Pathog **7:**e1002030.

33. **Franze de Fernandez, M. T., L. Eoyang, and J. T. August.** 1968. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. Nature **219:**588-590.

34. **Franze de Fernandez, M. T., W. S. Hayward, and J. T. August.** 1972. Bacterial proteins required for replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein. J Biol Chem **247:**824-831.

35. **Furuta, Y., B. B. Gowen, K. Takahashi, K. Shiraki, D. F. Smee, and D. L. Barnard.** 2013. Favipiravir (T-705), a novel viral RNA polymerase inhibitor. Antiviral Res **100:**446-454.

36. **Furuta, Y., K. Takahashi, M. Kuno-Maekawa, H. Sangawa, S. Uehara, K. Kozaki, N. Nomura, H. Egawa, and K. Shiraki.** 2005. Mechanism of action of T-705 against influenza virus. Antimicrob Agents Chemother **49:**981-986.

37. **Gasparian, A. V., N. Neznanov, S. Jha, O. Galkin, J. J. Moran, A. V. Gudkov, K. V. Gurova, and A. A. Komar.** 2010. Inhibition of encephalomyocarditis virus and poliovirus replication by quinacrine: implications for the design and discovery of novel antiviral drugs. J Virol **84:**9390-9397.

38. **Gerritz, S. W., C. Cianci, S. Kim, B. C. Pearce, C. Deminie, L. Discotto, B. McAuliffe, B. F. Minassian, S. Shi, S. Zhu, W. Zhai, A. Pendri, G. Li, M. A. Poss, S. Edavettal, P. A. McDonnell, H. A. Lewis, K. Maskos, M. Mortl, R. Kiefersauer, S. Steinbacher, E. T. Baldwin, W. Metzler, J. Bryson, M. D. Healy, T. Philip, M. Zoeckler, R. Schartman, M. Sinz, V. H. Leyva-Grado, H. H. Hoffmann, D. R. Langley, N. A. Meanwell, and M. Krystal.** 2011. Inhibition of influenza virus replication via small molecules that induce the formation of higher-order nucleoprotein oligomers. Proc Natl Acad Sci U S A **108:**15366-15371.

39. **Giorgi, C., L. Accardi, L. Nicoletti, M. C. Gro, K. Takehara, C. Hilditch, S. Morikawa, and D. H. Bishop.** 1991. Sequences and coding strategies of the S RNAs of Toscana and Rift Valley fever viruses compared to those of Punta Toro, Sicilian Sandfly fever, and Uukuniemi viruses. Virology **180:**738-753.

40. **Gowen, B. B., D. F. Smee, M. H. Wong, J. O. Hall, K. H. Jung, K. W. Bailey, J. R. Stevens, Y. Furuta, and J. D. Morrey.** 2008. Treatment of late stage

disease in a model of arenaviral hemorrhagic fever: T-705 efficacy and reduced toxicity suggests an alternative to ribavirin. PLoS One **3:**e3725.

41.   **Gowen, B. B., M. H. Wong, K. H. Jung, A. B. Sanders, M. Mendenhall, K. W. Bailey, Y. Furuta, and R. W. Sidwell.** 2007. In vitro and in vivo activities of T-705 against arenavirus and bunyavirus infections. Antimicrob Agents Chemother **51:**3168-3176.

42.   **Green, T. J., X. Zhang, G. W. Wertz, and M. Luo.** 2006. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. Science **313:**357-360.

43.   **Habjan, M., A. Pichlmair, R. M. Elliott, A. K. Overby, T. Glatter, M. Gstaiger, G. Superti-Furga, H. Unger, and F. Weber.** 2009. NSs protein of rift valley fever virus induces the specific degradation of the double-stranded RNA-dependent protein kinase. J Virol **83:**4365-4375.

44.   **Hacker, D., R. Raju, and D. Kolakofsky.** 1989. La Crosse virus nucleocapsid protein controls its own synthesis in mosquito cells by encapsidating its mRNA. J Virol **63:**5166-5174.

45.   **Hacker, D., S. Rochat, and D. Kolakofsky.** 1990. Anti-mRNAs in La Crosse bunyavirus-infected cells. J Virol **64:**5051-5057.

46.   **Haque, A., and M. A. Mir.** 2010. Interaction of hantavirus nucleocapsid protein with ribosomal protein S19. J Virol **84:**12450-12453.

47.   **Hawking, F.** 1940. Concentration of Bayer 205 (Germanin) in human blood and cerebrospinal fluid after treatment. Trans R Soc Trop Med Hyg**:**37-52.

48.   **Hewlett, M. J., and W. Chiu.** 1991. Virion structure. Curr Top Microbiol Immunol **169:**79-90.

49.   **Hewlett, M. J., R. F. Pettersson, and D. Baltimore.** 1977. Circular forms of Uukuniemi virion RNA: an electron microscopic study. J Virol **21:**1085-1093.

50.   **Hunter, P., B. J. Erasmus, and J. H. Vorster.** 2002. Teratogenicity of a mutagenised Rift Valley fever virus (MVP 12) in sheep. Onderstepoort J Vet Res **69:**95-98.

51.   **Ihara, T., H. Akashi, and D. H. Bishop.** 1984. Novel coding strategy (ambisense genomic RNA) revealed by sequence analyses of Punta Toro Phlebovirus S RNA. Virology **136:**293-306.

52.   **Ikegami, T., K. Narayanan, S. Won, W. Kamitani, C. J. Peters, and S. Makino.** 2009. Rift Valley fever virus NSs protein promotes post-transcriptional downregulation of protein kinase PKR and inhibits eIF2alpha phosphorylation. PLoS Pathog **5:**e1000287.

53.   **Ikegami, T., S. Won, C. J. Peters, and S. Makino.** 2007. Characterization of Rift Valley fever virus transcriptional terminations. J Virol **81:**8421-8438.

54.   **Iseni, F., F. Baudin, D. Blondel, and R. W. Ruigrok.** 2000. Structure of the RNA inside the vesicular stomatitis virus nucleocapsid. Rna **6:**270-281.

55.   **Jenkins, L. M., J. C. Byrd, T. Hara, P. Srivastava, S. J. Mazur, S. J. Stahl, J. K. Inman, E. Appella, J. G. Omichinski, and P. Legault.** 2005. Studies on the mechanism of inactivation of the HIV-1 nucleocapsid protein NCp7 with 2-mercaptobenzamide thioesters. J Med Chem **48:**2847-2858.

56.   **Jiao, L., S. Ouyang, M. Liang, F. Niu, N. Shaw, W. Wu, W. Ding, C. Jin, Y. Peng, Y. Zhu, F. Zhang, T. Wang, C. Li, X. Zuo, C. H. Luan, D. Li, and Z. J. Liu.** 2013. Structure of severe fever with thrombocytopenia syndrome virus

nucleocapsid protein in complex with suramin reveals therapeutic potential. J Virol **87:**6829-6839.

57. **Kakach, L. T., J. A. Suzich, and M. S. Collett.** 1989. Rift Valley fever virus M segment: phlebovirus expression strategy and protein glycosylation. Virology **170:**505-510.

58. **Kassack, M. U., K. Braun, M. Ganso, H. Ullmann, P. Nickel, B. Boing, G. Muller, and G. Lambrecht.** 2004. Structure-activity relationships of analogues of NF449 confirm NF449 as the most potent and selective known P2X1 receptor antagonist. Eur J Med Chem **39:**345-357.

59. **Keegan, K., and M. S. Collett.** 1986. Use of bacterial expression cloning to define the amino acid sequences of antigenic determinants on the G2 glycoprotein of Rift Valley fever virus. J Virol **58:**263-270.

60. **Kilgore, P. E., T. G. Ksiazek, P. E. Rollin, J. N. Mills, M. R. Villagra, M. J. Montenegro, M. A. Costales, L. C. Paredes, and C. J. Peters.** 1997. Treatment of Bolivian hemorrhagic fever with intravenous ribavirin. Clinical Infectious Diseases **24:**718-722.

61. **Kochs, G., C. Janzen, H. Hohenberg, and O. Haller.** 2002. Antivirally active MxA protein sequesters La Crosse virus nucleocapsid protein into perinuclear complexes. Proc Natl Acad Sci U S A **99:**3153-3158.

62. **Kolakofsky, D., and D. Hacker.** 1991. Bunyavirus RNA synthesis: genome transcription and replication. Curr Top Microbiol Immunol **169:**143-159.

63. **Le May, N., S. Dubaele, L. Proietti De Santis, A. Billecocq, M. Bouloy, and J. M. Egly.** 2004. TFIIH transcription factor, a target for the Rift Valley hemorrhagic fever virus. Cell **116:**541-550.

64. **Lopez, N., R. Muller, C. Prehaud, and M. Bouloy.** 1995. The L protein of Rift Valley fever virus can rescue viral ribonucleoproteins and transcribe synthetic genome-like RNA molecules. J Virol **69:**3972-3979.

65. **Madani, T. A., Y. Y. Al-Mazrou, M. H. Al-Jeffri, A. A. Mishkhas, A. M. Al-Rabeah, A. M. Turkistani, M. O. Al-Sayed, A. A. Abodahish, A. S. Khan, T. G. Ksiazek, and O. Shobokshi.** 2003. Rift Valley fever epidemic in Saudi Arabia: epidemiological, clinical, and laboratory characteristics. Clin Infect Dis **37:**1084-1092.

66. **Mastrangelo, E., M. Pezzullo, D. Tarantino, R. Petazzi, F. Germani, D. Kramer, I. Robel, J. Rohayem, M. Bolognesi, and M. Milani.** 2012. Structure-based inhibition of Norovirus RNA-dependent RNA polymerases. J Mol Biol **419:**198-210.

67. **McCormick, J. B., I. J. King, P. A. Webb, C. L. Scribner, R. B. Craven, K. M. Johnson, L. H. Elliott, and R. Belmont-Williams.** 1986. Lassa fever. Effective therapy with ribavirin. N Engl J Med **314:**20-26.

68. **McIntosh, B. M., D. Russell, I. dos Santos, and J. H. Gear.** 1980. Rift Valley fever in humans in South Africa. S Afr Med J **58:**803-806.

69. **Meegan, J. M.** 1979. The Rift Valley fever epizootic in Egypt 1977-78. 1. Description of the epizzotic and virological studies. Trans R Soc Trop Med Hyg **73:**618-623.

70. **Meegan JM, B. C., Monath TP.** 1989. Rift Valley fever., p. 51-76. *In* M. D. Thomas P. Monath (ed.), The arboviruses: epidemiology and ecology, vol. 4. CRC Press Inc, Boca Raton, Florida.

71. **Miller Jenkins, L. M., D. E. Ott, R. Hayashi, L. V. Coren, D. Wang, Q. Xu, M. L. Schito, J. K. Inman, D. H. Appella, and E. Appella.** 2010. Small-molecule inactivation of HIV-1 NCp7 by repetitive intracellular acyl transfer. Nat Chem Biol **6:**887-889.

72. **Mir, M. A., B. Brown, B. Hjelle, W. A. Duran, and A. T. Panganiban.** 2006. Hantavirus N protein exhibits genus-specific recognition of the viral RNA panhandle. J Virol **80:**11283-11292.

73. **Mir, M. A., W. A. Duran, B. L. Hjelle, C. Ye, and A. T. Panganiban.** 2008. Storage of cellular 5' mRNA caps in P bodies for viral cap-snatching. Proc Natl Acad Sci U S A **105:**19294-19299.

74. **Mir, M. A., and A. T. Panganiban.** 2006. The bunyavirus nucleocapsid protein is an RNA chaperone: possible roles in viral RNA panhandle formation and genome replication. Rna **12:**272-282.

75. **Mir, M. A., and A. T. Panganiban.** 2006. Characterization of the RNA chaperone activity of hantavirus nucleocapsid protein. J Virol **80:**6276-6285.

76. **Mir, M. A., and A. T. Panganiban.** 2005. The hantavirus nucleocapsid protein recognizes specific features of the viral RNA panhandle and is altered in conformation upon RNA binding. J Virol **79:**1824-1835.

77. **Mir, M. A., and A. T. Panganiban.** 2008. A protein that replaces the entire cellular eIF4F complex. Embo J **27:**3129-3139.

78. **Mir, M. A., and A. T. Panganiban.** 2004. Trimeric hantavirus nucleocapsid protein binds specifically to the viral RNA panhandle. J Virol **78:**8281-8288.

79. **Mir, M. A., and A. T. Panganiban.** 2010. The triplet repeats of the Sin Nombre hantavirus 5' untranslated region are sufficient in cis for nucleocapsid-mediated translation initiation. J Virol **84:**8937-8944.

80. **Miranda, G., D. Schuppli, I. Barrera, C. Hausherr, J. M. Sogo, and H. Weber.** 1997. Recognition of bacteriophage Qbeta plus strand RNA as a template by Qbeta replicase: role of RNA interactions mediated by ribosomal proteins S1 and host factor. J Mol Biol **267:**1089-1103.

81. **Monath, T. P.** 2008. Treatment of yellow fever. Antiviral Res **78:**116-124.

82. **Muller, R., C. Argentini, M. Bouloy, C. Prehaud, and D. H. Bishop.** 1991. Completion of the genome sequence of Rift Valley fever phlebovirus indicates that the L RNA is negative sense and codes for a putative transcriptase-replicase [corrected]. Nucleic Acids Res **19:**5433.

83. **Muller, R., O. Poch, M. Delarue, D. H. Bishop, and M. Bouloy.** 1994. Rift Valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. J Gen Virol **75 (Pt 6):**1345-1352.

84. **O'Brien, R. L., J. G. Olenick, and F. E. Hahn.** 1966. Reactions of quinine, chloroquine, and quinacrine with DNA and their effects on the DNA and RNA polymerase reactions. Proc Natl Acad Sci U S A **55:**1511-1517.

85. **Ogg, M. M., and J. L. Patterson.** 2007. RNA binding domain of Jamestown Canyon virus S segment RNAs. J Virol **81:**13754-13760.

86.	**Ono, K., H. Nakane, and M. Fukushima.** 1985. Inhibition of the activities of DNA primase-polymerase alpha complex from KB cells by hexasodium sym-bis(m-aminobenzoyl-m-amino-p-methylbenzoyl-1-naphthylamino-4 ,6,8-trisulfonate)carbamide. Nucleic Acids Symp Ser**:**249-252.

87.	**Osborne, J. C., and R. M. Elliott.** 2000. RNA binding properties of bunyamwera virus nucleocapsid protein and selective binding to an element in the 5' terminus of the negative-sense S segment. J Virol **74:**9946-9952.

88.	**Overby, A. K., R. F. Pettersson, and E. P. Neve.** 2007. The glycoprotein cytoplasmic tail of Uukuniemi virus (Bunyaviridae) interacts with ribonucleoproteins and is critical for genome packaging. J Virol **81:**3198-3205.

89.	**Panganiban, A. T., and M. A. Mir.** 2009. Bunyavirus N: eIF4F surrogate and cap-guardian. Cell Cycle **8:**1332-1337.

90.	**Patterson, J. L., B. Holloway, and D. Kolakofsky.** 1984. La Crosse virions contain a primer-stimulated RNA polymerase and a methylated cap-dependent endonuclease. J Virol **52:**215-222.

91.	**Paul, C. P., P. D. Good, S. X. Li, A. Kleihauer, J. J. Rossi, and D. R. Engelke.** 2003. Localized expression of small RNA inhibitors in human cells. Mol Ther **7:**237-247.

92.	**Perrin, F.** 1926. Polarisation de la lumière de fluorescence. Vie moyenne des molécules dans l'état excité. J. Phys. Radium. **7:**390-401.

93.	**Peters, C. J. a. A., G.W. Jr.** 1981. Pathogenesis of Rift Valley fever. Contributions to Epidemiology and Biostatistics **3:**21-41.

94.	**Pettersson, R., L. Kaariainen, C. H. von Bonsdorff, and N. Oker-Blom.** 1971. Structural components of Uukuniemi virus, a noncubical tick-borne arbovirus. Virology **46:**721-729.

95.	**Pittman, P. R., C. T. Liu, T. L. Cannon, R. S. Makuch, J. A. Mangiafico, P. H. Gibbs, and C. J. Peters.** 1999. Immunogenicity of an inactivated Rift Valley fever vaccine in humans: a 12-year experience. Vaccine **18:**181-189.

96.	**Prehaud, C., N. Lopez, M. J. Blok, V. Obry, and M. Bouloy.** 1997. Analysis of the 3' terminal sequence recognized by the Rift Valley fever virus transcription complex in its ambisense S segment. Virology **227:**189-197.

97.	**Pritchard, N. J., A. Blake, and A. R. Peacocke.** 1966. Modified intercalation model for the interaction of amino acridines and DNA. Nature **212:**1360-1361.

98.	**Raju, R., and D. Kolakofsky.** 1987. Unusual transcripts in La Crosse virus-infected cells and the site for nucleocapsid assembly. J Virol **61:**667-672.

99.	**Ramalingam, D., S. Duclair, S. A. Datta, A. Ellington, A. Rein, and V. R. Prasad.** 2011. RNA aptamers directed to human immunodeficiency virus type 1 Gag polyprotein bind to the matrix and nucleocapsid domains and inhibit virus production. J Virol **85:**305-314.

100.	**Raymond, D. D., M. E. Piper, S. R. Gerrard, G. Skiniotis, and J. L. Smith.** 2012. Phleboviruses encapsidate their genomes by sequestering RNA bases. Proc Natl Acad Sci U S A **109:**19208-19213.

101.	**Ribeiro, D., J. W. Borst, R. Goldbach, and R. Kormelink.** 2009. Tomato spotted wilt virus nucleocapsid protein interacts with both viral glycoproteins Gn and Gc in planta. Virology **383:**121-130.

102. **Rice, W. G., C. A. Schaeffer, B. Harten, F. Villinger, T. L. South, M. F. Summers, L. E. Henderson, J. W. Bess, Jr., L. O. Arthur, J. S. McDougal, and et al.** 1993. Inhibition of HIV-1 infectivity by zinc-ejecting aromatic C-nitroso compounds. Nature **361:**473-475.

103. **Rice, W. G., J. G. Supko, L. Malspeis, R. W. Buckheit, Jr., D. Clanton, M. Bu, L. Graham, C. A. Schaeffer, J. A. Turpin, J. Domagala, R. Gogliotti, J. P. Bader, S. M. Halliday, L. Coren, R. C. Sowder, 2nd, L. O. Arthur, and L. E. Henderson.** 1995. Inhibitors of HIV nucleocapsid protein zinc fingers as candidates for the treatment of AIDS. Science **270:**1194-1197.

104. **Ronka, H., P. Hilden, C. H. Von Bonsdorff, and E. Kuismanen.** 1995. Homodimeric association of the spike glycoproteins G1 and G2 of Uukuniemi virus. Virology **211:**241-250.

105. **Ruigrok, R. W., T. Crepin, and D. Kolakofsky.** 2011. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. Curr Opin Microbiol **14:**504-510.

106. **Schmaljohn, C. S.** 1996. p. 1447-1471. *In* B. N. Fields (ed.), Virology, Raven, New York.

107. **Schols, D., R. Pauwels, J. Desmyter, and E. De Clercq.** 1990. Dextran sulfate and other polyanionic anti-HIV compounds specifically interact with the viral gp120 glycoprotein expressed by T-cells persistently infected with HIV-1. Virology **175:**556-561.

108. **Senear, A. W., and J. A. Steitz.** 1976. Site-specific interaction of Qbeta host factor and ribosomal protein S1 with Qbeta and R17 bacteriophage RNAs. J Biol Chem **251:**1902-1912.

109. **Severson, W. E., X. Xu, and C. B. Jonsson.** 2001. cis-Acting signals in encapsidation of Hantaan virus S-segment viral genomic RNA by its N protein. J Virol **75:**2646-2652.

110. **Shvadchak, V., S. Sanglier, S. Rocle, P. Villa, J. Haiech, M. Hibert, A. Van Dorsselaer, Y. Mely, and H. de Rocquigny.** 2009. Identification by high throughput screening of small compounds inhibiting the nucleic acid destabilization activity of the HIV-1 nucleocapsid protein. Biochimie **91:**916-923.

111. **Sidwell, R. W., J. H. Huffman, B. B. Barnett, and D. Y. Pifat.** 1988. In vitro and in vivo Phlebovirus inhibition by ribavirin. Antimicrob Agents Chemother **32:**331-336.

112. **Sidwell, R. W., and D. F. Smee.** 2003. Viruses of the Bunya- and Togaviridae families: potential as bioterrorism agents and means of control. Antiviral Res **57:**101-111.

113. **Sinha, R., M. Hossain, and G. S. Kumar.** 2007. RNA targeting by DNA binding drugs: structural, conformational and energetic aspects of the binding of quinacrine and DAPI to A-form and H(L)-form of poly(rC).poly(rG). Biochim Biophys Acta **1770:**1636-1650.

114. **Snippe, M., J. Willem Borst, R. Goldbach, and R. Kormelink.** 2007. Tomato spotted wilt virus Gc and N proteins interact in vivo. Virology **357:**115-123.

115. **Strandin, T., J. Hepojoki, H. Wang, A. Vaheri, and H. Lankinen.** 2011. The cytoplasmic tail of hantavirus Gn glycoprotein interacts with RNA. Virology **418:**12-20.

116.   **Ullmann, H., S. Meis, D. Hongwiset, C. Marzian, M. Wiese, P. Nickel, D. Communi, J. M. Boeynaems, C. Wolf, R. Hausmann, G. Schmalzing, and M. U. Kassack.** 2005. Synthesis and structure-activity relationships of suramin-derived P2Y11 receptor antagonists with nanomolar potency. J Med Chem **48:**7040-7048.

117.   **Ulmanen, I., P. Seppala, and R. F. Pettersson.** 1981. In vitro translation of Uukuniemi virus-specific RNAs: identification of a nonstructural protein and a precursor to the membrane glycoproteins. J Virol **37:**72-79.

118.   **Vidal, S., and D. Kolakofsky.** 1989. Modified model for the switch from Sendai virus transcription to replication. J Virol **63:**1951-1958.

119.   **W.H.O.** 2000. Rift Valley fever, Saudi Arabia (update). Wkly Epidemiol Rec **75:**321.

120.   **WHO.** 2007. Outbreaks of Rift Valley fever in Kenya, Somalia and United Republic of Tanzania. Weekly Epidemiological Record **82:**169-178.

121.   **Won, S., T. Ikegami, C. J. Peters, and S. Makino.** 2007. NSm protein of Rift Valley fever virus suppresses virus-induced apoptosis. J Virol **81:**13335-13345.

122.   **Zhang, J. H., T. D. Chung, and K. R. Oldenburg.** 1999. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen **4:**67-73.

123.   **Zheng, W., and Y. J. Tao.** 2013. Structure and assembly of the influenza A virus ribonucleoprotein complex. FEBS Lett **587:**1206-1214.