BIOPHYSICAL CHARACTERISATION STUDIES OF HANTAVIRAL AND HUMAN LONG NONCODING RNAS

MICHAEL D'SOUZA
Bachelor of Science, University of Calgary, 2015

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

BIOCHEMISTRY

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

BIOPHYSICAL CHARACTERISATION STUDIES OF HANTAVIRAL AND HUMAN LONG NONCODING RNAS

Michael D'Souza

Date of Defence: December 17, 2021

Dr. T. Patel Thesis Supervisor	Associate Professor	Ph.D.
Dr. U. Kothe Thesis Examination Committee Member	Professor	Ph.D.
Dr. J. Ross Thesis Examination Committee Member	Professor	Ph.D.
Dr. J. Wu External Examiner Defence Research Development Canada Suffield, Alberta	Defence Scientist	Ph.D.
Dr. N. Thakor Chair, Thesis Examination Committee	Associate Professor	Ph.D.

Abstract

Emerging zoonotic viruses are a global threat that can cause devastating outbreaks and indelible economic fallout to human systems. North American Hantavirus is a NIAID Category A Priority Pathogen, that regularly involves their Noncoding Terminal Regions to instigate viral replication and translation. Long Noncoding RNAs are integral regulatory molecules involved in a myriad of biological processes. This thesis seeks to identify human RNA-Binding Proteins involved in Hantaviral lifecycles, and to biophysically characterise human Long Noncoding RNAs that mediate cancer progression. Pull-Down Assays identified several RNA stabilising and salient viral lifecycle promoting proteins that putatively interact with Hantaviral Noncoding Terminal Regions. Small Angle X-ray Scattering and coarse-grained SimRNA computational modelling methods generated combined low/high-resolution three-dimensional structures of Long Intergenic Noncoding RNA-p21 Alu Inverted Repeats. Furthermore, we have employed orthogonal biophysical approaches involving Analytical Ultracentrifugation and Multiangle/Dynamic Light Scattering techniques to verify target *in vitro* transcribed RNA for accuracy.

Acknowledgments

Table of Contents

Chapter 1: Introduction

Chapter 2: Biodefence Implications of New-World Hantaviruses

Chapter 3: Sin Nombre Orthohantavirus S and M Segment Sense and Antisense 3' Noncoding Terminal Regions Putatively Interact with Human Host RNA-Binding Proteins

Chapter 4: Biophysical Characterisation of Human LincRNA-p21 Sense and Antisense Alu Inverted Repeats

Chapter 5: Conclusions

List of Tables

List of Figures

Abbreviations

2DSA	Two-Dimensional Spectrum Analysis
AMV	Amur Virus
ANDV	Andes Virus
ARQV	Araraquara orthohantavirus
AUC	Analytical Ultracentrifugation
BCCV	Black Creek Canal Virus
BSC	Biosafety Cabinets
BTWC	Biological and Toxin Weapon Convention
CDC	Centers for Disease Control and Prevention
CIPB	Cytoplasmic Immunoprecipitation Buffer
CLB	Cytoplasmic Lysis Buffer
CLIP	Crosslinking and Immunoprecipitation
CSD	Cold Shock Domain
D	Diffusion Coefficient

DC	Dendritic Cells
DHX9	ATP-dependent RNA Helicase A
DIG	Digoxigenin
DLS	Dynamic Light Scattering
DNA	Deoxyribonucelic Acid
DOBV	Dobrava Virus
dsRBD	Double-Stranded RNA Binding Domains
dsRNA	Double-Stranded RNAs
eEF1A	Eukaryotic Translation Elongation Factors 1 Alpha
ELAVL1	ELAV-like Protein 1
ELISA	Enzyme-Linked Immunosorbent Assay
EMSA	Electrophoretic Mobility Shift Assays
ERGIC	Endoplasmic Reticulum-Golgi Intermediate Compartment
ETAR	1-β-D-ribofuranosyl-3-ethynyl-[1,2,4] triazole
EU	European Union
FBS	Foetal Bovine Serum
FPLC	Fast Protein Liquid Chromatography
GPC	Glycoprotein Precursor
HCPS	Hantavirus Cardiopulmonary Syndrome
HFRS	Haemorrhagic Fever with Renal Syndrome
HHP	High-Hit Proteins
HIV	Human Immunodeficiency Virus
hnRNP-K	Heterogenous Nuclear Ribonuclear Protein-K
HPLC	High Performance Liquid Chromatography
HPMEC	Human Pulmonary Microvascular Endothelial Cells
HSPA8	Heat Shock Cognate 71 kDa Protein
HTNV	Haantan Virus
HuR	Human Antigen R

IAV	Influenza A Virus
IFN	Interferon
IR	Inverted Repeats
IRES	Internal Ribosome Entry Site
IRF	IFN-Regulatory Factors
JAK/STAT	Janus Kinases and Signal and Activators of Transcription Pathways
L	Large (Segment)
LC-MS/MS	Liquid Chromatography Tandem Mass Spectrometry
lincRNA	Long Intergenic Noncoding RNAs
LINES	Long Interspersed Nuclear Elements
lncRNA	Long Noncoding RNAs
LTR	Long Terminal Repeats
M	Medium (Segment)
MALS	Multiangle Light Scattering
miRNA	microRNAs
mRNA	Messenger RNA
mRNP	messenger Ribonucleoproteins Particles
MST	Microscale Thermophoresis
N	Nucleocapsid
ncRNA	Noncoding RNA
NE	Nephropathis Epidemica
NFκB	Nuclear Factor κB
NIAID	US National Institute of Allergy and Infectious Diseases
NIH	National Institute of Health
NS	Non-Structural Protein
NSD	Normalized Spatial Discrepancy
nt	Nucleotides
NTR	Noncoding Terminal Region

ORF	Open Reading Frame
PAGE	Polyacrylamide Gel Electrophoresis
PAMP	Pathogen Associated Molecular Patterns
PCBP1	Poly(rC)-binding protein 1
PCBP2	Poly(rC)-binding protein 2
PHV	Prospect Hill Virus
PKM	Pyruvate Kinase (M1/M2 Isoform)
PRKRA	Interferon-inducible Double-stranded RNA-dependent Protein Kinase Activator A
PRR	Pattern Recognition Receptors
PUUV	Puumala Virus
RBP	RNA-Binding Protein
RdRp	RNA Dependent RNA Polymerase
RIP	RNA Immunoprecipitation
RMSD	Root-Mean Square Deviation
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RRM	RNA-Recognition Motif
rRNA	Ribosomal RNA
RVFV	Rift Valley Fever Virus
S	Small (Segment)
s	Sedimentation Coefficient
SAXS	Small Angle X-Ray Crystallography
SAXS	Small Angle X-ray Scattering
SDS	Sodium Dodecyl Sulphate
SEC	Size Exclusion Chromatography
SEOV	Seoul Virus
SHAPE	Selective 2'-Hydroxyl Acylation Analyses by Primer Extension
SINES	Short Interspersed Nuclear Elements

siRNA	Small Interfering RNA
SMS	Alberta Sequence Manipulation Suite
snRNA	Small Nuclear RNAs
SNV	Sin Nombre Virus
STAU1	Double-stranded RNA-binding Protein Staufen Homolog 1
SV	Sedimentation Velocity
TE	Transposable Elements
TNF	Tumour Necrosis Factor
TP53COR1	Tumour Protein p53 pathway Corepressor 1
TPH	Total Protein Hits
tRNA	Transfer RNA
UTR	Untranslated Regions
VHFs	Hemorrhagic Fevers
VLP	Virus-like Particle
YBX1	Y-box-binding protein 1
YBX3	Y-box-binding protein 3

Chapter 1: Introduction

1.0 Overview

As the world increasingly industrialises and expands interconnectivity through travel and commercial globalisation, the risks of infection from local to international outbreaks caused by emerging pathogens are similarly increasing [1]. Emerging viruses are those that have either newly appeared, are increasing in incidence and geographic range, or will have the potential to expand their distribution in the near future [2]. Incidentally, the significant threat to human systems from emerging pathogens stems from their capacity to undermine and disrupt public health infrastructure, economies, and social security apparatuses through the paucity of knowledge regarding their pathogenesis and measures to minimise their harm. Single-stranded RNA viruses are of particular concern due to their genetic plasticity, structure, and segmented organisation [3]. This enables RNA viruses to undergo rapid mutations and exchange genome segments through reassortment events during infection, including co-infection with other RNA viruses [4-6]. This invariably leads to the creation of either highly pathogenic, or relatively maladroit viruses, that have the capacity to cause local outbreaks or global pandemics given the feasibility of their transmission.

As evidenced by the ongoing zoonotic SARS-CoV-2 pandemic, the continuing but localised 2014 Ebola epidemic, and the perennial and now ritualised Influenza infections, RNA viruses maintain their capacity to cause diseases whose effects are experienced disproportionally throughout the world [7-10]. Notably, the SARS-CoV-2 pandemic has negatively affected human mental and physical health through its unfettered spread amongst global populations and its subsequent erosion of modern capital economies [11, 12]. The indolent reaction and unpreparedness of developed nations and the international community against the initial spread were critical in SARS-CoV-2's dispersion. The proliferation of emerging RNA viruses is further exacerbated by anthropogenic activity and climate change, which together put pressures on environments through habitat destruction, increases in global average temperatures, and heightened air and water pollution [13]. These effects tend to expand the habitable range of viral vectors and induce advantageous mutations towards their greater virulence [14]. Rising vector ranges produced from warming or erratic temperatures has increased the close contact of humans with RNA viral vectors and reservoir hosts which primarily include insects (ticks and mosquitos) and rodents [15-17]. This frequent exposure has increased the transmission of RNA viruses from animals to humans, and together with industrial agricultural practices and land reclamation have minimised biodiversity whilst amplifying the rate of zoonotic spillover events [18].

1

Naturally, mitigating the risks imposed by novel infectious diseases caused by RNA viruses can be implemented through strong detection, surveillance, and monitoring systems that enable widespread reporting and global familiarisation as the disease progresses [19, 20]. This can include incorporating mathematical and computational modelling and prediction techniques to evaluate zoonotic RNA viruses as they occur worldwide [21]. An aspect of this involves analysing RNA viral genomes and their specific interactions with host elements. A key yet moderately understood property of RNA viruses are the Untranslated Regions (UTRs) within their genome, particularly their end-termini. These terminal UTRs exhibit roles that regularly regulate the lifecycles of highly infectious viruses from families including Bunyaviruses and Flaviviruses [22-24]. Identifying future treatments require foundational and holistic strategies that begins with bioinformatic and biophysical characterisation studies involving comprehensive RNA sequence and structure analysis. Naturally occurring emerging viruses will nevertheless continue to beset human systems, but by understanding their key structural and genetic properties, practical and effective vaccines, medical countermeasures, and therapeutics can be designed to diminish their effects.

1.1 Hantaviruses

Hantaviruses are a genus of segmented, single-stranded, negative-sense RNA viruses [25]. They are found worldwide and pose a serious biosecurity threat to human health, agricultural production, and global trade due to their robust dispersibility and high mortality rate [26]. Belonging to the *Bunyaviridae* family, Hantaviruses are emerging zoonotic pathogens that cause two deadly diseases: Hantavirus Cardiopulmonary Syndrome (HCPS) which is predominately found in the western hemisphere; and Haemorrhagic Fever with Renal Syndrome (HFRS) found in Asia, Europe, and Africa (**Figure 1.1**) [27]. The latter HFRS causes upwards of 150,000 to 200,000 infections annually in the eastern hemisphere or Old World, retaining a 15% mortality rate, while HCPS is a highly fatal disease with a mortality rate of up to 50% [28]. Incidentally, Hantaviruses overall are a bioweapons concern being regarded as a Category C Pathogen by the Centers for Disease Control and Prevention and a Category A Pathogen by the US National Institute of Allergy and Infectious Diseases [29-31]. The latter includes emerging pathogens that can be engineered for mass dissemination because of their availability, ease of production, and potential for high morbidity and high mortality rates, constituting the highest priority for national security. This contrasts with the former which includes high-priority agents that pose a national risk to security because they can: be easily disseminated or transmitted person-to-person; result in high mortality and retain the potential for major public health impacts; may

2

cause public panic and social disruption; and require special action for public preparedness [32]. An important pathogenic North American species of New World Hantaviruses is Sin Nombre *orthohantavirus* (SNV), which is prevalent in and distributed across Canada and the United States, and alongside Andes Virus (ANDV), pose a significant threat from exploits in bioterrorism due to their high morbidity and nature of neglected reporting.
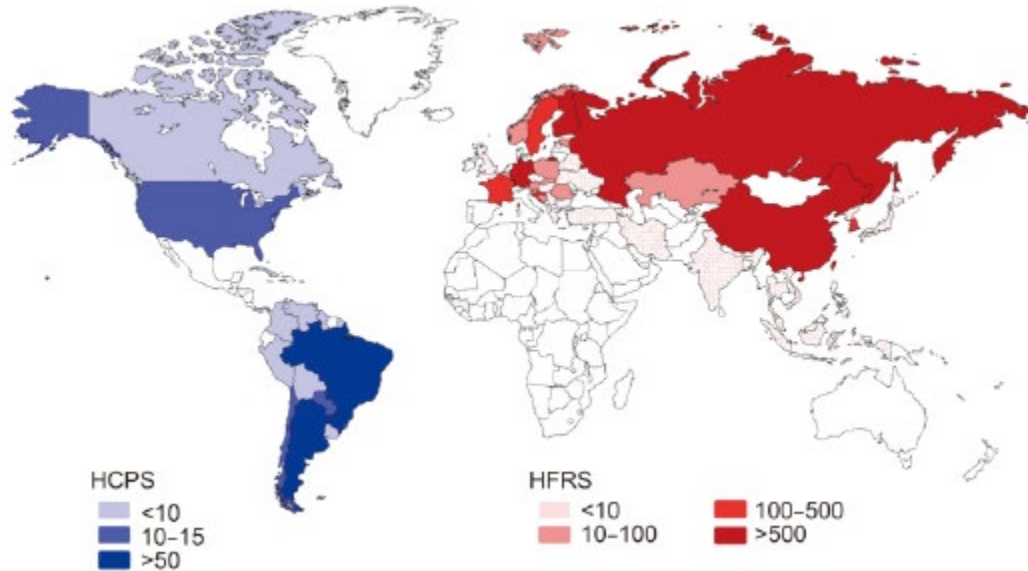
Figure 1.1: Geographical Representation of the Approximate Incidences per year of Hantavirus Cardiopulmonary Syndrome HCPS (Blue) and Haemorrhagic Fever with Renal Syndrome HFRS (Red). Projected from 2016 occurrences. Reprinted with permission from *Virologica Sinica* [33].

The origins of Hantavirus were first apparent during the Korean War (1950-1953) when 3,200 UN soldiers became ill with an unknown but HFRS-causing disease that was later identified to be Hantaan Virus (HTNV) [34-36]. Several HFRS-causing Hantaviruses are spread across Europe, Eurasia, and East-Asia, and include the Amur Virus (AMV), Seoul Virus (SEOV), Dobrava Virus (DOBV), and the Puumala Virus (PUUV). Each of these Hantaviral species causes diseases with their own severity, the latter of which causes Nephropathis Epidemica (NE) which is a milder HFRS-causing disease but still accounts for a case-fatality rate of 0.1-1% [37]. Currently, most of the cases of HFRS occur in Russia, northern and central Europe, and Mainland China, with the latter accounting for 90% of the more than 100,000 annual cases [38, 39]. As a result of being a zoonotic disease, Hantaviruses are spread through very specific rodent host reservoirs that are asymptomatic of the disease but are persistently infected, shedding the aerosolized virion particles through their skin, saliva, urine, and faeces [40, 41]. In North America, the common deer

mouse (*Peromyscus maniculatus*) is the specific vector of SNV, retaining a broad range across the Canadian prairies and American Midwest [33]. While around 1,000 HCPS cases have been reported globally, SNV cases reported in Canada primarily occurs within persons working in agriculture and operating in the military. This constitutes 70% of the more than 120 HCPS cases reported of which 29% were fatalities (**Figure 1.2**) [42]. As of 2020, more than 143 cases have been detected, with a projection of 4-5 cases occurring per year in Canada [43]. The absence of effective FDA-approved vaccines or anti-viral therapeutics makes SNV very dangerous to infected persons or those operating in close contact with its natural rodent reservoir [28]. Consequently, it is important to conduct studies investigating the role Hantavirus RNA has within viral infection and its impact on host cells.
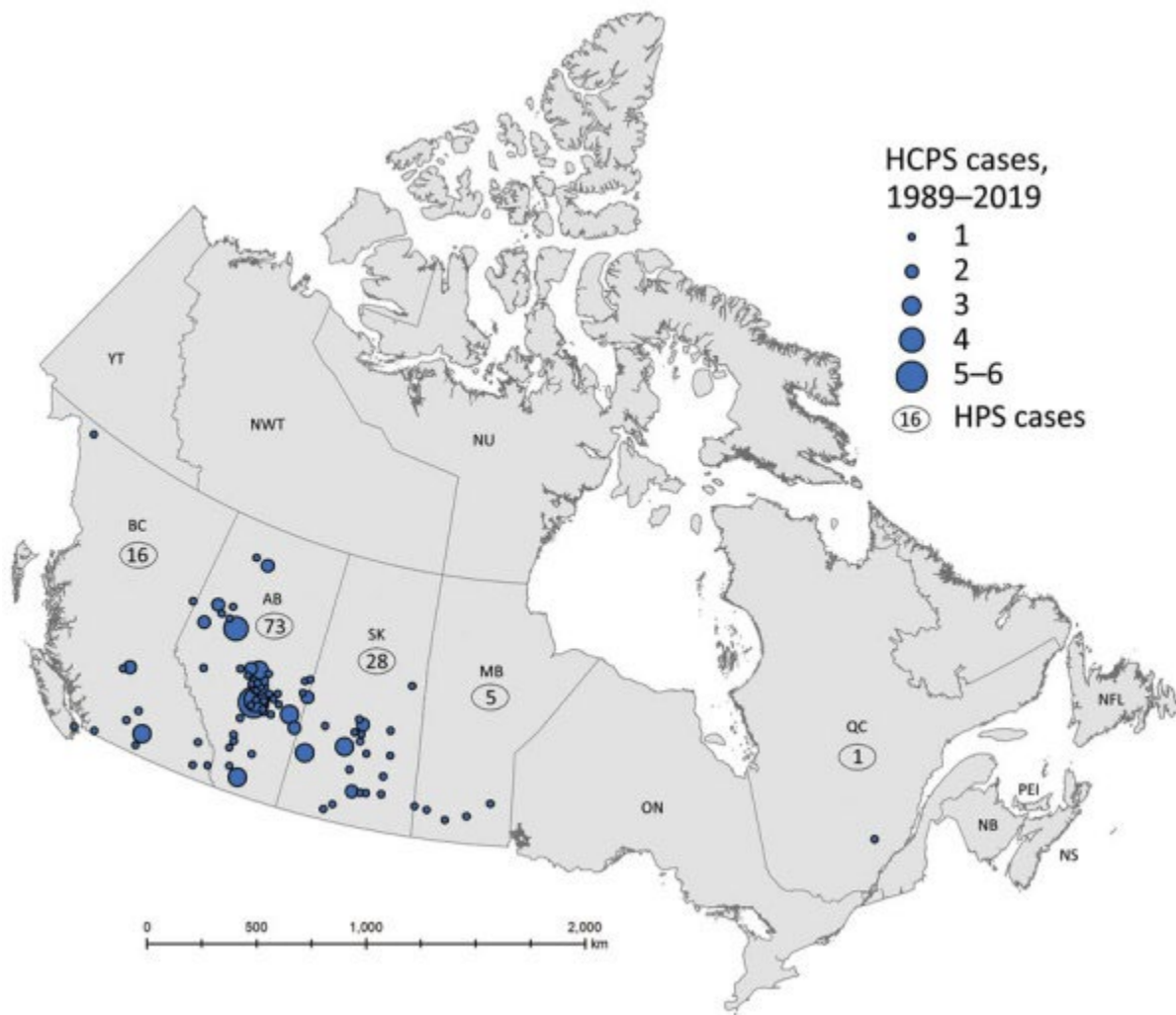
Figure 1.2: The Geographic Distribution of Confirmed Cases of Hantavirus Cardiopulmonary Syndrome (HCPS) Occurring from 1989-2019 in Canada. The map indicates locations of cases where Sin Nombre Virus infection were likely contacted. Reprinted within the Public Domain from *Emerging Infectious Diseases* [43].

1.1.1 Hantaviral Genome: Noncoding Terminal and Panhandle Regions

The Hantavirus genome consists of three segments: the Small (S) Segment composed of 2,059 nucleotides (nt); the Medium (M) Segment comprised of 3,696 nt; and the Large (L) Segment containing 6,562 nt, each with their own single open reading frame flanked by noncoding terminal regions (**Figure 1.3**) [44, 45]. The S Segment encodes the Nucleocapsid (N) Protein while the M Segment encodes the Glycoprotein Precursor (GPC) which is post-translationally cleaved to form the Gn and Gc Glycoproteins [46, 47]. The L Segment encodes the RNA-Dependent RNA Polymerase (RdRp) which possesses transcriptase, replicase, and endonuclease functions and processes both viral RNA transcription and genome replication [48, 49]. The N Protein is significant, as it is involved in the packaging and encapsidation of the viral genome through the formation of a Ribonucleoprotein (RNP) Complex [50]. Enclosed around the tripartite RNP is an envelope that is common to all Bunyaviruses [51].

Both the 3′ and 5′-Noncoding Terminal Regions (NTRs) of each negative-sense segment possess a highly conserved terminal sequence in addition to their extensive untranslated region that is present across Hantavirus species that is capable of forming conserved panhandle structures through complementary base-pairing [52]. Generally consisting of 14–18 nt, these 5′ and 3′ sequences are highly conserved throughout the Hantavirus genus, with the first 8 nt being absolutely conserved, enough that the N protein can distinguish between its secondary structure and panhandles form other genera [53]. Hantaviral panhandles were comprised of the following sequence at the 3′-termini: 3′AUCAUCAUCUGAGG-5′; and the following sequence at the 5′-termini: 5′-UAGUAGUAU(G/A)CUCC-3′ [44]. Recognised by the Trimeric N Protein Complex, these panhandle regions are involved in the process of initiation and regulation of viral transcription, replication and encapsidation [53]. Both the primary sequence and secondary structure within the terminal regions' panhandle are important for the binding of the N Protein [52]. The secondary structure of HTNV panhandle regions form dsRNA segments which are punctuated by symmetrically aligned but unpaired nucleotides at positions 9 and either 18 or 19 from their 5′ and 3′-ends. The negative-sense vRNA also has greater affinity to the Trimeric N Protein Complex compared to the positive-sense mRNA which has a 3′-truncation that prevents base pairing at the complementary 5′ and 3′-termini [52, 54]. The N protein itself is a viral analogue of eIF4F cap-binding complex and is actively involved in the initiation of viral translation through a complex formation with the RdRp [24, 55]. Additionally, after its disassembly from the Ribonucleoprotein Complex, the N Protein assist the vRNA in its transcription and subsequent translation while acting as a chaperone [56].
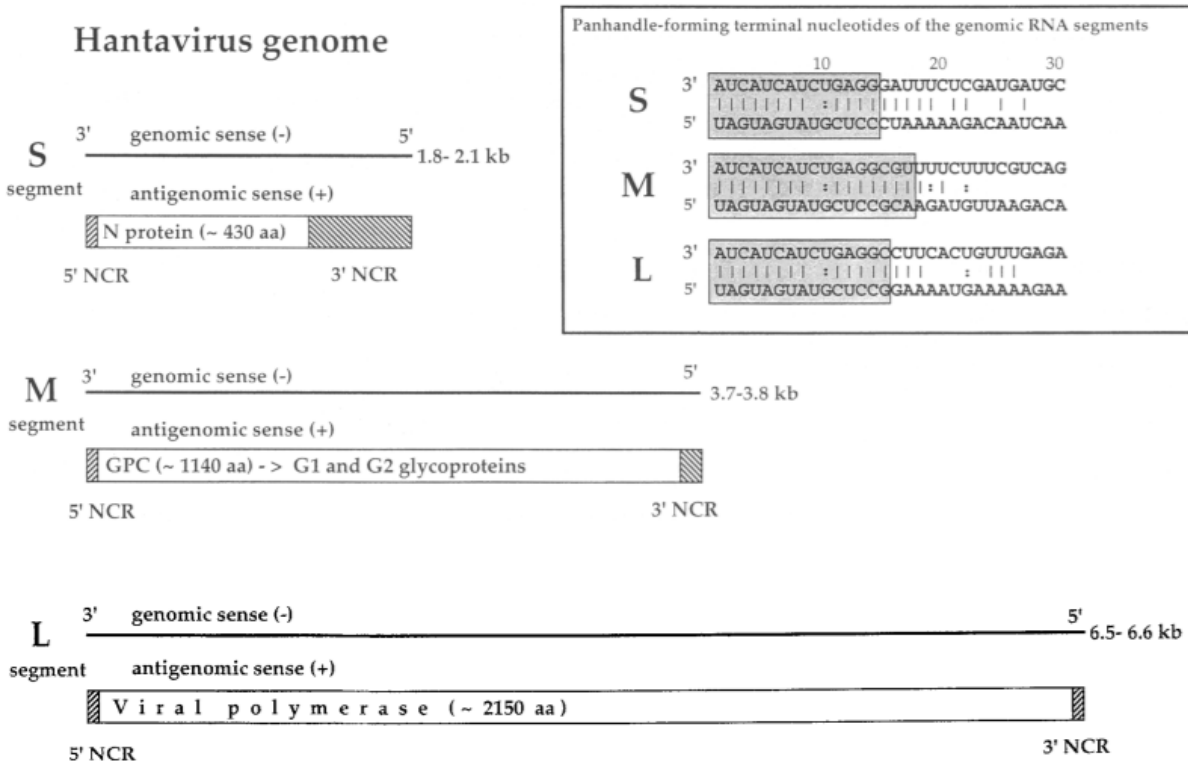
Figure 1.3: The Genomic Structure of Hantaviruses. Panhandle forming Noncoding Terminal Regions of the tripartite segments of HTNV are present in the inset box, with genus-specific terminal regions highlighted in grey. NTRs are represented by Noncoding Regions (NCR) in the figure, however, we make the distinction from general UTR/ncRNA with the term Noncoding *Terminal* Region to emphasise locality. Reprinted within Fair Dealing, under the Canadian Copyright Act, from the *Journal of General Virology* [57].

The Noncoding Terminal Regions of ssRNA viruses are integral in a variety of different lifecycle processes, including cyclisation (panhandle formation), packaging, and the initiation of replication and translation [23]. Previous studies have shown the accumulation of defective 3′-deleted S, M, and L vRNA to elicit a decline in Hantaviral titre and vRNA concentration, with their 3′-NTR deletions hypothesized to downregulate replication by competing with the standard Hantavirus for N Protein binding [58]. In other Bunyaviruses, such as the *Phlebovirus* Rift Valley Fever Virus (RVFV), RNA synthesis is regulated by the highly conserved noncoding regions of its S, M, and L segments which similarly retain panhandle structures associated with viral replication [59, 60]. Additionally, RVFV contains highly ordered secondary structures in their NTRs adjacent to these panhandle regions that facilitate its tripartite S, M, and L segment's translation and replication [61]. Hantaviral NTRs are likely necessary for viral replication with the conserved panhandle sequence deletion being observed to reduce viral replication in infected cultures [25]. Their interaction is strongly linked with the RNA chaperone functions of the N Protein which not only protects the segments but also initiates replication through the dissociation of the panhandle structure and the genomic RNA's exposure to

RdRp activity [62]. Replication of the viral genome occurs in the cytoplasm and is carried out by the RdRp which, in conjunction with the N Protein, carries out a cap-snatching process which cleaves the 5′-terminus of host mRNA to act as a primer for viral replication [63]. The N Protein assists in the protection of cap-degradation and viral mRNA persistence by binding to their 5′-terminus and enables the RdRp to carry out a prime and realign mechanisms to elongate viral mRNA for translation by cellular ribosomes [64-66].

Pathogenic Hantavirus entry is mediated by β_3-integrin receptors which are present on platelets, macrophages, and endothelial cells, with SNV's main target being pulmonary endothelial cells [67, 68]. Hantavirus infections alter the barrier properties of these cells resulting in vascular leakage and target organ failure, with HCPS infections causing pulmonary oedema or an increased vascular permeability in pulmonary endothelial cells, thrombocytopenia, and respiratory distress resulting in organ damage, dysfunction, and shock [69-71]. Humans are dead-end hosts for HCPS-causing Hantaviruses, however, rodents, and particularly common deer mice, exhibit negligible levels of pathology and can remain persistently infected for life [72, 73]. Since host reservoirs are associated with specific Hantavirus strains, there is a basis for a deep co-evolutionary history which enables their resistance and unelicited immune response to infection. This could be in the form of Hantavirus regulating the rodent T-cell response and especially the anti-inflammatory adaptive response that will limit immunopathology [74]. In humans, the Type I Interferon (IFN) response to SNV infections has been observed, resulting in the nuclear translocation of Transcription Factors, especially Nuclear Factor κB (NFκB), IFN-Regulatory Factors 1, 3, and 7 (IRF-1/3/7) [75]. The induction of the infected cell's antiviral immune response attracts stimulated effector T-cells which release pro-inflammatory cytokines including Tumour Necrosis Factor-α (TNF-α) and IFN-γ which invariably causes vasodilation, leading to endothelial leakage and severe oedema due in part to the heightened activation of $CD8^+$ T-cells [76]. SNV must antagonize the antiviral interferon (IFN) response to propagate successfully in infected cells, as doing so prevents the elicitation of the JAK/STAT pathway and its subsequent signal cascade promoting the expression of antiviral proteins [77]. The N Protein of New-World Hantaviruses (ANDV, SNV) does not inhibit NFκB [78, 79].

Hantavirus replication occurs in the cytoplasm followed by the disassembly of RNP cores from a late endosome occurring near the Endoplasmic Reticulum-Golgi Intermediate Compartment (ERGIC) [48, 63]. Transcription follows a cap-snatching, prime and realignment mechanism undertaken by the RdRp and initiated by the N Protein which is also responsible for the encapsidation and packaging of the viral genome [48, 80, 81]. Consequently, the importance of the N Protein as a chaperone in binding to the base-pairing panhandle structure of

the terminal regions with specificity invites the possibility for other RNA and host RNA-Binding Proteins (RBPs) to occur [24]. This is important in discerning how vRNA is transported to the ERGIC and free ribosomes for translation after its release from the endosome, or if host proteins are assisting the N Protein in its stability and manipulation of the genomic RNA during RNP disassembly, or encapsidation [50]. Since the N protein recognises a common structure or sequence within the NTRs' panhandle region, the presence of similarly interacting host RBPs that would either mitigate or enhance viral activity would be important for the development of future therapeutics against Hantaviruses [53].

1.1.2 Hantaviral Treatment and Biodefence Considerations

There are currently no approved post-exposure therapeutic countermeasures against Hantavirus infection for both HCPS and HFRS-causing infections, in addition to there being no USA FDA-approved vaccines and antivirals [82]. However, there are a variety of novel and experimental therapeutics that provide varying degrees of efficacy against Hantaviral infections. In the Republic of Korea, an HFRS vaccine exists in the form of Hantavax which has been accessible commercially and provided to the Republic of Korean Armed Forces due to their close contact with Hantaviral reservoir hosts. Hantavax has increasingly reduced patients hospitalised with the HFRS disease since its administration in 1991, with patients retaining 50% of the HTNV neutralizing antibodies a year after the first dose and increased immune responses after a 3-dose strategy. Despite the treatment, the HFRS-disease severity hasn't been reduced with its application [83-85]. Several disease treatment drugs are in experimentation and are administered either as a post-exposure prophylactic when viral infection occurs prior to the start of viremia or the manifestation of clinical signs, or as a therapeutic to treat patients past the stages of viremia [69]. First there is Ribavirin which affects the biological function of the Hantaviral RdRp and Favipiravir which has also been shown to reduce Bunyavirus RdRp activity and its ability to replicate and transcribe the Hantaviral genome. Favipiravir has been shown to reduce viral loads in hamster serum for lethal ANDV and SNV challenged hamster models, however, no human trials have been conducted [86, 87]. ETAR (1-β-D-ribofuranosyl-3-ethynyl-[1,2,4] triazole) has also been shown to increase immunocompromised mice survivability from HFRS-causing diseases. Both ETAR and Ribavirin act as nucleoside analogues which disrupts the replication of Hantaviral vRNA and the subsequent translation of mRNA, as well as the inhibition of the RdRp [88]. Most of these trials have been limited to mice or hamster models, with macaques being the closest non-human primate model employed.

HCPS-causing Hantaviruses are themselves potential bioweapons as elaborated by considerations from the CDC and NIAID. NIAID assesses its importance as a Category A Bioweapon threat because of HCPS-causing SNV or ANDV exerting a high mortality rate of up to 50% and maintaining a rapid disease progression that elicits serious pulmonary symptoms [89]. ANDV is also the only known Hantavirus that can be transmitted person-to-person which can include aerosolised droplets, saliva, and infection from breast milk [90-93]. When combined with other insidious properties of New-World Hantaviruses including HCPS being an emerging infectious disease with limited global surveillance and monitoring, and defined reporting in developing nations, can result in ANDV and SNV becoming potential candidates for bioweapons research and application. The spread of these viruses can greatly affect systems experiencing infrastructural and environmental decline and destruction due to the effects of climate change, civil wars and its subsequent refugee crises which will result in close human proximity to rodent hosts and an increase in the frequency of Hantaviral cases [94].

Some characteristics enable HCPS-causing Hantaviruses to be challenging agents. The Hantaviral virion is incredibly stable, and can survive for more than 18-days between -20°C and 4°C, and 10-days at room temperature enabling it to persist in the environment for an extended period of time [95]. Given the widely present availability of North and South American rodent hosts in the environment, HCPS-Causing New-World Hantaviruses can be easily acquired and proliferated given weaponization cells' laboratory expertise and handling [42, 96]. However, Hantaviruses can be inactivated and degraded by heat treatment of 60°C sustained for 30 minutes, in addition to the application of detergents, UV radiation, organic solvents, and hypochlorite solutions [27]. Concentrating Hantaviral virions for effective dispersion also remains difficult due to the requirement for repeated passaging of Hantaviral virions in immune-compromised rodent hosts although technology and new animal infection models such as non-rodent and non-human primate models can increase viral yields [97-100]. Regardless of its implementation or health risks, the deployment of Hantaviral bioweapons as an area denial weapon or to disrupt civilian activities can impact society by impeding global trade, affecting local or international economies, and increasing the cost of personnel, resources, and equipment for quarantining and sanitation [94].

1.2 Long Noncoding RNAs

Despite being from a higher order complex of organisms, the human genome retains a smaller than anticipated capacity to only express ~30,000 protein-coding genes, roughly ~2% of its entire genome [101, 102]. The

reliance on alternative splicing then becomes essential in the evolution of higher order organisms for the generation of functional proteins that contribute to necessary biological life processes [103]. These functions are additionally supplemented with the application of noncoding RNAs (ncRNA) which account for the remaining ~98% of the human genome. Several classes of ncRNAs are employed in biological processes and can include Small Nuclear RNAs (snRNA), Transfer RNAs (tRNA), Ribosomal RNA (rRNA), Double-Stranded RNAs (dsRNA), Small Interfering RNA (siRNA), and microRNAs (miRNA). Many of these RNAs function in regulatory roles, whether that is decoding the messenger RNA (mRNA) sequences into proteins during translation processes such as the tRNA or are employed in splicing events of mRNA during transcription such as snRNA [104]. Another significant fraction of the transcriptional activity present in the human genome is represented by Long Noncoding RNAs (lncRNA) which constitutes another class of regulatory ncRNAs. LncRNAs are characterised by a minimum length of 200 nt, an absence of protein-coding potential, and the presence of a 5′-cap and a poly-adenyl tail, whilst maintaining a sequence that can be spliced but contains fewer exons compared to mRNA [104, 105]. Many LncRNAs function in critical biological roles with special implications in gene regulatory processes.

The number of human LncRNAs has been estimated to be more than 60,000 with the majority of LncRNAs being localised in the nucleus where they primarily engage in the regulation of gene expression [106]. This crucial function is executed through chromatin modification and remodelling, histone modification, and changes to the nucleosome's localisation [107]. However, some LncRNAs are also present in the cytoplasm [108]. After being transcribed by RNA Polymerase II and exported from the nucleus, cytosolic LncRNAs help regulate mRNA stability, modulate translation, and can interfere with post-translational modifications [109]. These actions regulate a wide range of biological processes including cellular proliferation, cell cycle, metabolism, apoptosis, and the differentiation and maintenance of pluripotent cells [110]. Depending on their relative proximity to protein-encoding genes in the human genome, LncRNAs can be further classified into sense, antisense, and bidirectional LncRNAs as well as Intron, Intergenic, and Enhancer LncRNAs [107]. Due to their multiplicity and diverse range of action, LncRNAs are often found regulating and inducing apoptosis of cancerous cells, and contribute towards critical pathways in lung, breast, liver, and colorectal cancers [111, 112].

Since LncRNAs exhibit a broad range of functionality, identifying their integral structural elements that contribute towards LncRNA functionality – either through the identification of secondary stricture motifs and domain organisation or three-dimensional topology and shape mapping and motif identification – has become a point of

importance [113]. As RNA molecules, LncRNAs are conferred with structurally dependent chemical properties that enable them to carry out biological processes, particularly their regulatory potential as epigenetic modulators [109]. Lacking an open reading frame, which distinguishes them from protein coding transcripts, many LncRNA transcripts function intrinsically as RNA molecules which retain structures that can bind and interact with proteins and other RNAs [114]. LncRNAs then have a tendency of folding into thermodynamically stable secondary structures because of their Watson-Crick, Hoogsteen, and ribose-oriented hydrogen bonding [114]. These inherently stable secondary structures include double helices, hairpins, stem-loops, bulges, and pseudoknots that can assist in higher order, tertiary structure folding. However, because RNA are dynamic, flexible molecules that are negatively charged, they require additional stabilisers in the form of molecular crowders, and cations from salts to help them overcome electrostatic barriers and assembly misfolding [115-117]. Additionally, LncRNAs also rely on the stacking of their aromatic nucleic acid bases to form stable tertiary structures that are not just mediated in a non-Watson Crick base pairing manner [115]. This architecture can ultimately form RNA, Protein, or DNA binding domains, including conformational switches that can be employed in multifaceted operations within the cell and contribute towards cell life and regulation [118]. Incidentally, folding of LncRNAs is multiplexed, and isn't entirely dependent on primary and secondary structural arrangements. This is especially important with regards to tumour suppression and DNA Damage Repair during instances of cancer and disease progression caused by viral infections which allows an analysis of RNA secondary structure to be essential in identifying LncRNA functions.

1.2.1 LincRNA-p21

Much like LncRNAs, Long Intergenic Noncoding RNAs (LincRNA) are characterised as being autonomously transcribed RNA molecules longer than 200 nt with no protein coding potential [118]. An important and multifaceted regulatory LincRNA is LincRNA-p21 or Tumour Protein p53 pathway Corepressor 1 (TP53COR1), which is a 3100 nt RNA molecule found on chromosome 17 and located approximately 15kb upstream of the *Cdkn1a* (p21) gene [119]. Due to its relationship with p53, a transcription factor that suppresses tumour growth through a signal cascade eliciting the activation of tumour suppressor genes, LincRNA-p21 has been shown to influence the progression of cancers which regularly present mutations of the TP53 gene itself [120-122]. LincRNA-p21 transcription is induced by p53 upon DNA damage whereby LincRNA-p21 will physically associate with and recruit

11

the nuclear factor Heterogenous Nuclear Ribonuclear Protein-K (hnRNP-K) to specific promoters that mediates the p53-dependent transcriptional response [119, 123].

LincRNA-p21 is also known for modulating transcription, whereby it acts as a post-transcriptional inhibitor of translation. This is carried out by its interaction with Human Antigen R (HuR), whose presence causes LincRNA-p21 to become unstable through the recruitment of the RNA-Binding Protein complex let-7/Ago2 which subsequently promotes the translation of *CTNNB1* and *JUNB* mRNAs [124]. In HuR's absence, LincRNA-p21 accumulates which allows the Rck/p54 DEAD-box Helicase to promote its association with the *CTNNB1* and *JUNB* mRNAs thus repressing their translation. *CTNNB1* and *JUNB* encode transcriptional activators JunB and β-catenin which have been shown to regulate colorectal cancer progression [125]. Equally, in hepatocellular cancers, LincRNA-p21 is downregulated where it originally inhibited hepatocarcinoma cell proliferation and colony formation whilst inducing apoptosis of hepatocellular carcinoma cells [126]. The manifold targets of LincRNA-p21 indicates that it potentially has specific RNA-Protein Binding regions and hence ordered secondary structures that interact plainly and with high-affinity to downstream targets such as p53 and hnRNP-K. These targets are further explored in **Figure 1.4**.
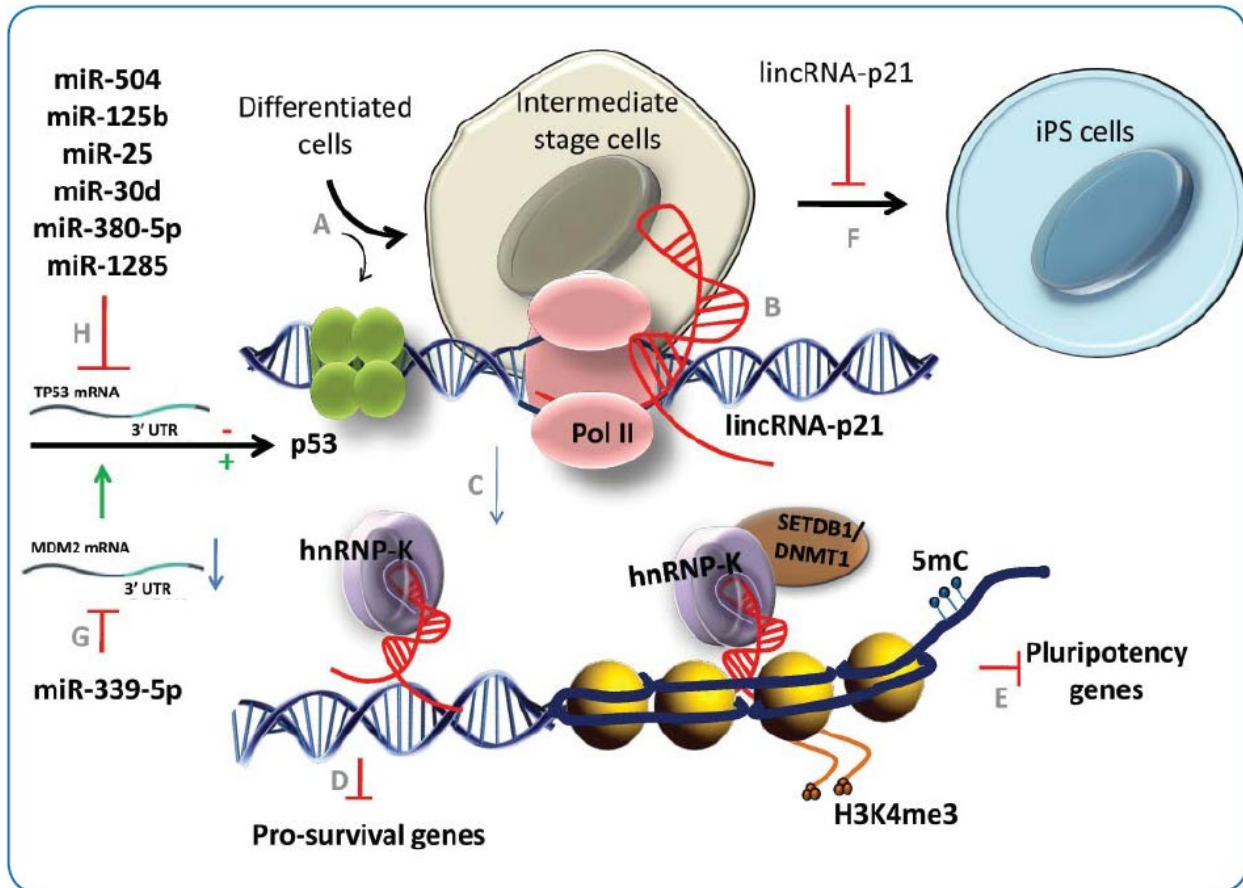
Figure 1.4: The Role LincRNA-p21 plays in Impairing Cellular Reprogramming and its Relationship with the p53 Tumour Suppression Pathway. Pathway A indicates the transformation of differentiated cells into intermediate stage cells which promote the expression of p53. Pathway B indicates p53 promoting transcription of LincRNA-p21 (Red RNA sequence). Pathway C indicates LincRNA-p21 combining with hnRNP-K which together inhibit pro-survival gene expression depicted in Pathway D leading to cellular apoptosis. Pathway E depicts the LincRNA-p21 role in inhibiting pluripotent genes after binding to hnRNP-K and SETDB1/DNMT1 complex which can be impacted by epigenetic modifications. Pathway F indicates the inhibition of induced pluripotent stem cells (iPSC). Tumour suppression by miR-339-5p activated p53 is indicated by Pathway G, where direct binding of p53′s 3′-UTR to MDM2 mRNA inhibits increasing levels of p53. Pathway H depicts the action of microRNA that interferes with p53 expression by interacting with p53′s 3′UTR with sequence-specificity. Reprinted within Fair Dealing, under the Canadian Copyright Act, from *Molecular Life* [127].

1.2.2 LincRNA-p21 Alu Inverted Repeat Elements

The human genome itself is highly dynamic and mutable, affected by genomic rearrangements that are the result of the movement of mobile endogenous DNA throughout the genome and the integration of exogenous DNA into the genome [128]. These rearrangements ultimately impact the human genome's plasticity and evolution which itself is benefitted through the insertion of repetitive sequences derived from Transposable Elements (TE). TEs are mobile genetic elements, or pieces of DNA that can move within the genome. There are two classes of TEs: Class I Retrotransposons and Class II DNA Transposons [128, 129]. Retrotransposons are additionally categorised by the

presence of Long Terminal Repeats (LTR) that contain the Transposable Elements' functions for mobility and its regulatory sequences. There are ~500,000 copies of LTR sequences in humans which constitutes ~8% of the human genome. Class I Retrotransposons are also composed of Long Interspersed Nuclear Elements (LINES) and Short Interspersed Nuclear Elements (SINES), of which the former elements are mobile while the latter are non-autonomous DNA elements that cannot encode for proteins and require LINES to assist in their own mobility and propagation.

The most common TE in humans and other primates is the Alu element which likely originated from the 7SL RNA gene that encodes for the noncoding RNA of the Signal Recognition Particle [130, 131]. Alu elements are a family of primate-specific SINEs, consisting of ~280-300 nt and are further divided into the subfamilies *AluJ* (oldest element), *AluS*, and *AluY* (youngest element) [132]. Their relative age indicates when they were evolutionarily inserted into the genome. Alu elements are embedded in LincRNAs where they directly participate in RNA-RNA and RNA-Protein binding interactions with mRNA and proteins like transcription factors [133, 134]. Their diverse range of action can be accredited to Alu sequences being present in both the sense and antisense orientation with inversely oriented Alus frequently but imperfectly base-pairing with each other and forming stem-loops that contribute towards RNA functionality within the cell [135, 136]. Incidentally, Alu elements are associated with genetic diseases, with current models proposing that their removal or alterations made to them affect the *cis* (genes on the same chromosome) regulation of adjacent genes and results in genetic disease progression [137].

LincRNA-p21 contains highly conserved sense and antisense Inverted Repeats (IR) Alu elements which have been identified through chemical probing and traditional sequencing techniques to possess secondary structures that are involved in the localisation of LincRNA-p21 into the nucleus (**Figure 1.5**) [113, 138]. Since LincRNA-p21 is an integral component in the p53-mediated stress response, the structure of these Alu IRs appear to directly impact cancer progression. The structure-function correlation of LincRNA-p21 is likely observed through LincRNA-p21's Alu elements involvement with direct RNA-Protein interactions with hnRNP-K and its subsequent localisation into the nucleus. Although there is no relation of LincRNA-p21 to viral mechanisms, analysing its biophysical properties as a Long Noncoding RNA can influence observances of viral Noncoding RNAs acting in a similarly regulatory role.
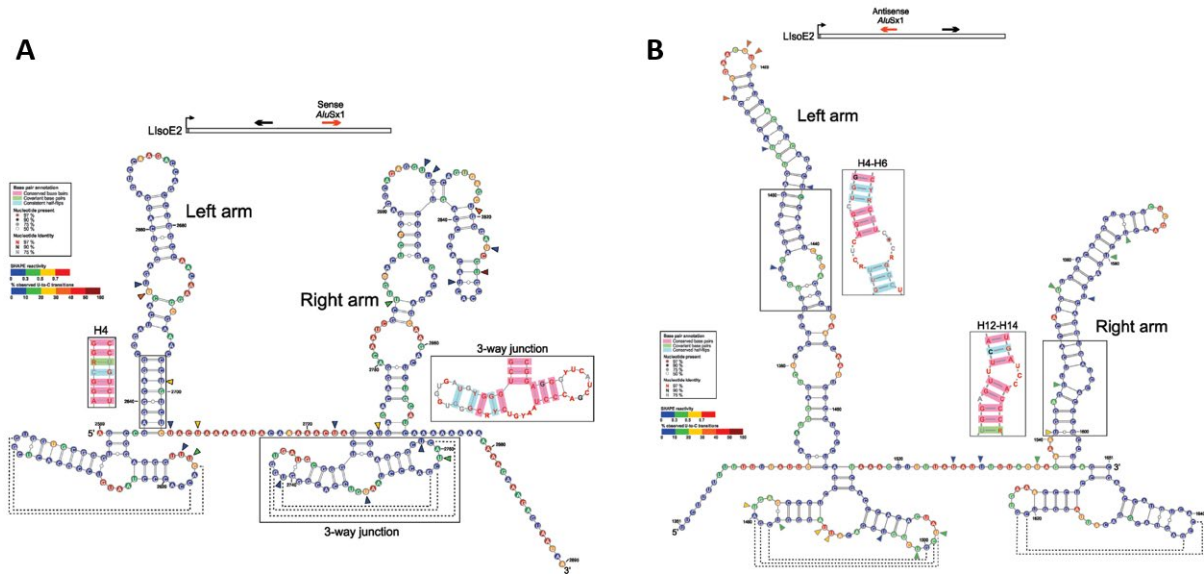
Figure 1.5: Schematic Representing the Experimentally Derived Secondary Structure of hLincRNA-p21 Sense (A) and Antisense (B) Inverted Repeat Alu Elements (IR Alu). Colours of the circle organised nucleotides indicate the relative SHAPE reactivities with dotted lines representing the putative tertiary contacts between the terminal loops and the three-way junctions. Reprinted with permission from *Nucleic Acid Research* [138].

1.4 Objectives of Thesis

Interactions between human RNA-Binding Proteins and nucleic acids from viral and human sources are one of many cardinal requirements for the progression of viral infectious diseases and the development of cancers [112, 139-141]. Human RNA-Binding Proteins have also been identified to regulate and minimise viral infections and could be potential targets for medical therapeutics and drug development against pathogenic ssRNA viruses [142, 143]. Similarly, the activity of Long Noncoding RNAs exercises regulatory actions against cancers and within the innate antiviral response of the immune system, becoming an invaluable means to fight diseases [144]. Tentative antivirals are being designed to target the N Protein and vRNA relationship which has disrupted Hantavirus viral genomic RNA synthesis and N-mediated translation [145]. Given the importance of the Hantaviral Noncoding Terminal Regions and its relationship with the Hantaviral N Protein for lifecycle progression, I hypothesise that the Hantaviral terminal regions' panhandle structure, and additional secondary structural motifs are involved in RNA-Binding Protein interactions. These interactions would be exclusive to human host relationships that are outside of the regulatory roles unique to the Hantavirus N Protein that are essential for viral transport, packaging, replication, translation, and evasion from the immune system. Identifying and biophysically characterising said RNA-Binding Protein interactions can beneficially be exploited towards the development of medical countermeasures against Hantaviral infections.

My thesis work intends to explore Long Noncoding RNAs present in both human and viral systems. My initial work, outlined in **Chapter 2**, observes the classification of HCPS-causing Hantaviruses as potential bioweapons which is contradicted by the US National Institute of Allergy and Infectious Diseases and the Centers for Disease Control and Prevention. We clarified the categorisation of Hantaviruses as a potential bioweapon and highlighted the practicality of employing New-World Hantaviruses as bioweapons. My thesis work extends to identify potential human RNA-Binding Proteins that putatively interact with Hantaviral ssRNA which is demonstrated in **Chapter 3**. The intent is to produce a foundational library of RNA-Binding Proteins with Hantaviral NTRs that can be characterised with kinetic and biophysical approaches to further develop the role that Terminal Regions plays in the progression of viral infectious diseases. This first experimental objective is to identify potential RNA-Binding Proteins that interact with the 5′ and 3′ Noncoding Terminal Regions of the tripartite, negative-sense Sin Nombre *orthohantavirus* using a Pull-Down Assay involving Digoxigenin-labelled RNA. My second experimental objective is to characterise the biophysical properties of LincRNA-p21 *AluSx1* Alu Inverted Repeats in solution to define and corroborate previously modelled secondary structures. Outlined in **Chapter 4**, we employed Small Angle X-ray Scattering and computationally generated tertiary structural models to validate chemical probing techniques. Doing so generates a pipeline towards identifying critical secondary structural elements of viral and human Long Noncoding RNAs using chemical probing techniques and validation through low-resolution three-dimensional solution scattering and computational modelling to elucidate their functional roles.

1.5 References

1.   Morse, S.S., *Factors in the emergence of infectious diseases.* Emerg Infect Dis, 1995. **1**(1): p. 7-15.
2.   Holland, D.J., *Emerging viruses.* Curr Opin Pediatr, 1998. **10**(1): p. 34-40.
3.   Comas-Garcia, M., *Packaging of Genomic RNA in Positive-Sense Single-Stranded RNA Viruses: A Complex Story.* Viruses, 2019. **11**(3).
4.   McDonald, S.M., et al., *Reassortment in segmented RNA viruses: mechanisms and outcomes.* Nat Rev Microbiol, 2016. **14**(7): p. 448-60.
5.   Vijaykrishna, D., R. Mukerji, and G.J. Smith, *RNA Virus Reassortment: An Evolutionary Mechanism for Host Jumps and Immune Evasion.* PLoS Pathog, 2015. **11**(7): p. e1004902.
6.   Sanjuán, R. and P. Domingo-Calap, *Mechanisms of viral mutation.* Cell Mol Life Sci, 2016. **73**(23): p. 4433-4448.
7.   Holmes, E.C., et al., *The evolution of Ebola virus: Insights from the 2013-2016 epidemic.* Nature, 2016. **538**(7624): p. 193-200.
8.   Jacob, S.T., et al., *Ebola virus disease.* Nat Rev Dis Primers, 2020. **6**(1): p. 13.
9.   Taubenberger, J.K. and J.C. Kash, *Influenza virus evolution, host adaptation, and pandemic formation.* Cell Host Microbe, 2010. **7**(6): p. 440-51.
10.   Acuti Martellucci, C., et al., *SARS-CoV-2 pandemic: An overview.* Adv Biol Regul, 2020. **77**: p. 100736.

11. Hiscott, J., et al., *The global impact of the coronavirus pandemic.* Cytokine Growth Factor Rev, 2020. **53**: p. 1-9.

12. Nicola, M., et al., *The socio-economic implications of the coronavirus pandemic (COVID-19): A review.* Int J Surg, 2020. **78**: p. 185-193.

13. Rahman, M.T., et al., *Zoonotic Diseases: Etiology, Impact, and Control.* Microorganisms, 2020. **8**(9).

14. Dennehy, J.J., *Evolutionary ecology of virus emergence.* Ann N Y Acad Sci, 2017. **1389**(1): p. 124-146.

15. Robert, M.A., A.M. Stewart-Ibarra, and E.L. Estallo, *Climate change and viral emergence: evidence from Aedes-borne arboviruses.* Curr Opin Virol, 2020. **40**: p. 41-47.

16. Klempa, B., *Hantaviruses and climate change.* Clin Microbiol Infect, 2009. **15**(6): p. 518-23.

17. Elliott, R.M., *Bunyaviruses and climate change.* Clin Microbiol Infect, 2009. **15**(6): p. 510-7.

18. Wilkinson, D.A., et al., *Habitat fragmentation, biodiversity loss and the risk of novel infectious disease emergence.* Journal of the Royal Society, Interface, 2018. **15**(149): p. 20180403.

19. Buckeridge, D. and G. Cadieux, *Surveillance for Newly Emerging Viruses.* Perspectives in medical virology, 2006. **16**: p. 325-343.

20. Al-Tawfiq, J.A., et al., *Surveillance for emerging respiratory viruses.* Lancet Infect Dis, 2014. **14**(10): p. 992-1000.

21. Morse, S.S., et al., *Prediction and prevention of the next pandemic zoonosis.* Lancet, 2012. **380**(9857): p. 1956-65.

22. Brinton, M.A. and M. Basu, *Functions of the 3' and 5' genome RNA regions of members of the genus Flavivirus.* Virus Res, 2015. **206**: p. 108-19.

23. Liu, Y., et al., *Structures and Functions of the 3' Untranslated Regions of Positive-Sense Single-Stranded RNA Viruses Infecting Humans and Animals.* Front Cell Infect Microbiol, 2020. **10**: p. 453.

24. Mir, M.A. and A.T. Panganiban, *The triplet repeats of the Sin Nombre hantavirus 5' untranslated region are sufficient in cis for nucleocapsid-mediated translation initiation.* J Virol, 2010. **84**(17): p. 8937-44.

25. Meyer, B.J. and C.S. Schmaljohn, *Persistent hantavirus infections: characteristics and mechanisms.* Trends Microbiol, 2000. **8**(2): p. 61-7.

26. Hussein, I.T., et al., *Recent advances in hantavirus molecular biology and disease.* Adv Appl Microbiol, 2011. **74**: p. 35-75.

27. Avsic-Zupanc, T., A. Saksida, and M. Korva, *Hantavirus infections.* Clin Microbiol Infect, 2019. **21s**: p. e6-e16.

28. Llah, S.T., et al., *Hantavirus induced cardiopulmonary syndrome: A public health concern.* J Med Virol, 2018. **90**(6): p. 1003-1009.

29. Mittler, E., et al., *Hantavirus entry: Perspectives and recent advances.* Adv Virus Res, 2019. **104**: p. 185-224.

30. CDC. *Bioterrorism Agents/Diseases*. 2018 April 4, 2018; Available from: https://emergency.cdc.gov/agent/agentlist-category.asp.

31. NIAID. *US NIAID Emerging Infectious Diseases/ Pathogens*. 2018 July 26, 2018; Available from: https://www.niaid.nih.gov/research/emerging-infectious-diseases-pathogens.

32. Christian, M.D., *Biowarfare and Bioterrorism.* Critical Care Clinics, 2013. **29**(3): p. 717-756.

33. Jiang, H., et al., *Hantavirus infection: a global zoonotic challenge.* Virol Sin, 2017. **32**(1): p. 32-43.

34. Lee, H.W., P.W. Lee, and K.M. Johnson, *Isolation of the etiologic agent of Korean hemorrhagic fever. 1978.* J Infect Dis, 2004. **190**(9): p. 1711-21.

35. Schmaljohn, C. and B. Hjelle, *Hantaviruses: a global disease problem.* Emerg Infect Dis, 1997. **3**(2): p. 95-104.

36. Smadel, J.E., *Epidemic hemorrhagic fever.* Am J Public Health Nations Health, 1953. **43**(10): p. 1327-30.

37. Brummer-Korvenkontio, M., et al., *Nephropathia epidemica: detection of antigen in bank voles and serologic diagnosis of human infection.* J Infect Dis, 1980. **141**(2): p. 131-4.

38. Jiang, H., et al., *Hemorrhagic Fever with Renal Syndrome: Pathogenesis and Clinical Picture.* Front Cell Infect Microbiol, 2016. **6**: p. 1.

39. Tkachenko, E.A., et al., *Hemorrhagic Fever with Renal Syndrome, Russia.* Emerg Infect Dis, 2019. **25**(12): p. 2325-2328.

40. Forbes, K.M., T. Sironen, and A. Plyusnin, *Hantavirus maintenance and transmission in reservoir host populations.* Curr Opin Virol, 2018. **28**: p. 1-6.

41. Milholland, M.T., et al., *Global Diversity and Distribution of Hantaviruses and Their Hosts.* Ecohealth, 2018. **15**(1): p. 163-208.

42. Drebot, M.A., et al., *Hantavirus pulmonary syndrome in Canada: An overview of clinical features, diagnostics, epidemiology and prevention.* Can Commun Dis Rep, 2015. **41**(6): p. 124-131.

43. Warner, B.M., et al., *Hantavirus Cardiopulmonary Syndrome in Canada.* Emerg Infect Dis, 2020. **26**(12): p. 3020-3024.

44. Chizhikov, V.E., et al., *Complete genetic characterization and analysis of isolation of Sin Nombre virus.* J Virol, 1995. **69**(12): p. 8132-6.

45. Spiropoulou, C.F., et al., *Genome structure and variability of a virus causing hantavirus pulmonary syndrome.* Virology, 1994. **200**(2): p. 715-23.

46. Cifuentes-Munoz, N., N. Salazar-Quiroz, and N.D. Tischler, *Hantavirus Gn and Gc envelope glycoproteins: key structural units for virus cell entry and virus assembly.* Viruses, 2014. **6**(4): p. 1801-22.

47. Rissanen, I., et al., *Structural Transitions of the Conserved and Metastable Hantaviral Glycoprotein Envelope.* J Virol, 2017. **91**(21).

48. Muyangwa, M., et al., *Hantaviral Proteins: Structure, Functions, and Role in Hantavirus Infection.* Front Microbiol, 2015. **6**: p. 1326.

49. Hepojoki, J., et al., *Hantavirus structure--molecular interactions behind the scene.* J Gen Virol, 2012. **93**(Pt 8): p. 1631-44.

50. Guo, Y., et al., *Crystal Structure of the Core Region of Hantavirus Nucleocapsid Protein Reveals the Mechanism for Ribonucleoprotein Complex Formation.* J Virol, 2016. **90**(2): p. 1048-61.

51. Elliott, R.M., *Molecular biology of the Bunyaviridae.* J Gen Virol, 1990. **71 (Pt 3)**: p. 501-22.

52. Mir, M.A. and A.T. Panganiban, *The hantavirus nucleocapsid protein recognizes specific features of the viral RNA panhandle and is altered in conformation upon RNA binding.* J Virol, 2005. **79**(3): p. 1824-35.

53. Mir, M.A., et al., *Hantavirus N protein exhibits genus-specific recognition of the viral RNA panhandle.* J Virol, 2006. **80**(22): p. 11283-92.

54. Severson, W., et al., *Characterization of the Hantaan nucleocapsid protein-ribonucleic acid interaction.* J Biol Chem, 1999. **274**(47): p. 33732-9.

55. Cheng, E., Z. Wang, and M.A. Mir, *Interaction between hantavirus nucleocapsid protein (N) and RNA-dependent RNA polymerase (RdRp) mutants reveals the requirement of an N-RdRp interaction for viral RNA synthesis.* J Virol, 2014. **88**(15): p. 8706-12.

56. Mir, M.A. and A.T. Panganiban, *Characterization of the RNA chaperone activity of hantavirus nucleocapsid protein.* J Virol, 2006. **80**(13): p. 6276-85.

57. Plyusnin, A., O. Vapalahti, and A. Vaheri, *Hantaviruses: genome structure, expression and evolution.* J Gen Virol, 1996. **77 (Pt 11)**: p. 2677-87.

58. Meyer, B.J. and C. Schmaljohn, *Accumulation of terminally deleted RNAs may play a role in Seoul virus persistence.* J Virol, 2000. **74**(3): p. 1321-31.

59. Elliott, R.M., C.S. Schmaljohn, and M.S. Collett, *Bunyaviridae genome structure and gene expression.* Curr Top Microbiol Immunol, 1991. **169**: p. 91-141.

60. Bouloy, M. and F. Weber, *Molecular biology of rift valley Fever virus.* Open Virol J, 2010. **4**: p. 8-14.

61. Gauliard, N., et al., *Rift Valley fever virus noncoding regions of L, M and S segments regulate RNA synthesis.* Virology, 2006. **351**(1): p. 170-9.

62. Brown, B.A. and A.T. Panganiban, *Identification of a region of hantavirus nucleocapsid protein required for RNA chaperone activity.* RNA Biol, 2010. **7**(6): p. 830-7.

63. Jeeva, S., et al., *Hantavirus RdRp Requires a Host Cell Factor for Cap Snatching.* J Virol, 2019. **93**(5).

64. Mir, M.A., et al., *Storage of cellular 5' mRNA caps in P bodies for viral cap-snatching.* Proc Natl Acad Sci U S A, 2008. **105**(49): p. 19294-9.

65. Garcin, D., et al., *The 5' ends of Hantaan virus (Bunyaviridae) RNAs suggest a prime-and-realign mechanism for the initiation of RNA synthesis.* J Virol, 1995. **69**(9): p. 5754-62.

66. Olschewski, S., S. Cusack, and M. Rosenthal, *The Cap-Snatching Mechanism of Bunyaviruses.* Trends Microbiol, 2020. **28**(4): p. 293-303.

67. Gavrilovskaya, I.N., et al., *beta3 Integrins mediate the cellular entry of hantaviruses that cause respiratory failure.* Proc Natl Acad Sci U S A, 1998. **95**(12): p. 7074-9.

68. Macneil, A., S.T. Nichol, and C.F. Spiropoulou, *Hantavirus pulmonary syndrome.* Virus Res, 2011. **162**(1-2): p. 138-47.

69. Brocato, R.L. and J.W. Hooper, *Progress on the Prevention and Treatment of Hantavirus Disease.* Viruses, 2019. **11**(7).

70. Vergote, V., et al., *A lethal disease model for New World hantaviruses using immunosuppressed Syrian hamsters.* PLoS Negl Trop Dis, 2017. **11**(10): p. e0006042.

71.   Zaki, S.R., et al., *Hantavirus pulmonary syndrome. Pathogenesis of an emerging infectious disease.* Am J Pathol, 1995. **146**(3): p. 552-79.

72.   Botten, J., et al., *Experimental infection model for Sin Nombre hantavirus in the deer mouse (Peromyscus maniculatus).* Proc Natl Acad Sci U S A, 2000. **97**(19): p. 10578-83.

73.   Botten, J., et al., *Persistent Sin Nombre virus infection in the deer mouse (Peromyscus maniculatus) model: sites of replication and strand-specific expression.* J Virol, 2003. **77**(2): p. 1540-50.

74.   Schountz, T., et al., *Regulatory T cell-like responses in deer mice persistently infected with Sin Nombre virus.* Proc Natl Acad Sci U S A, 2007. **104**(39): p. 15496-501.

75.   Sundstrom, J.B., et al., *Hantavirus infection induces the expression of RANTES and IP-10 without causing increased permeability in human lung microvascular endothelial cells.* J Virol, 2001. **75**(13): p. 6070-85.

76.   Borges, A.A., et al., *Hantavirus cardiopulmonary syndrome: immune response and pathogenesis.* Microbes Infect, 2006. **8**(8): p. 2324-30.

77.   Gallo, G., et al., *Interactions of Viral Proteins from Pathogenic and Low or Non-Pathogenic Orthohantaviruses with Human Type I Interferon Signaling.* Viruses, 2021. **13**(1).

78.   Taylor, S.L., et al., *Hantaan virus nucleocapsid protein binds to importin alpha proteins and inhibits tumor necrosis factor alpha-induced activation of nuclear factor kappa B.* J Virol, 2009. **83**(3): p. 1271-9.

79.   Taylor, S.L., R.L. Krempel, and C.S. Schmaljohn, *Inhibition of TNF-alpha-induced activation of NF-kappaB by hantavirus nucleocapsid proteins.* Ann N Y Acad Sci, 2009. **1171 Suppl 1**: p. E86-93.

80.   Cheng, E. and M.A. Mir, *Signatures of host mRNA 5' terminus for efficient hantavirus cap snatching.* J Virol, 2012. **86**(18): p. 10173-85.

81.   Mir, M.A., *Hantaviruses.* Clin Lab Med, 2010. **30**(1): p. 67-91.

82.   Liu, R., et al., *Vaccines and Therapeutics Against Hantaviruses.* Front Microbiol, 2019. **10**: p. 2989.

83.   Cho, H.W. and C.R. Howard, *Antibody responses in humans to an inactivated hantavirus vaccine (Hantavax).* Vaccine, 1999. **17**(20-21): p. 2569-75.

84.   Yi, Y., H. Park, and J. Jung, *Effectiveness of inactivated hantavirus vaccine on the disease severity of hemorrhagic fever with renal syndrome.* Kidney Res Clin Pract, 2018. **37**(4): p. 366-372.

85.   Song, J.Y., et al., *Long-term immunogenicity and safety of inactivated Hantaan virus vaccine (Hantavax™) in healthy adults.* Vaccine, 2016. **34**(10): p. 1289-95.

86.   Safronetz, D., et al., *Antiviral efficacy of favipiravir against two prominent etiological agents of hantavirus pulmonary syndrome.* Antimicrob Agents Chemother, 2013. **57**(10): p. 4673-80.

87.   Safronetz, D., et al., *In vitro and in vivo activity of ribavirin against Andes virus infection.* PLoS One, 2011. **6**(8): p. e23560.

88.   Chung, D.H., et al., *Synthesis of 1-beta-D-ribofuranosyl-3-ethynyl-[1,2,4]triazole and its in vitro and in vivo efficacy against Hantavirus.* Antiviral Res, 2008. **79**(1): p. 19-27.

89.   Jonsson, C.B., J. Hooper, and G. Mertz, *Treatment of hantavirus pulmonary syndrome.* Antiviral Research, 2008. **78**(1): p. 162-169.

90.   Bellomo, C., et al., *A newborn infected by Andes virus suggests novel routes of hantavirus transmission: a case report.* Clin Microbiol Infect, 2020. **26**(1): p. 130-131.

91.   Alonso, D.O., et al., *Person-to-Person Transmission of Andes Virus in Hantavirus Pulmonary Syndrome, Argentina, 2014.* Emerg Infect Dis, 2020. **26**(4): p. 756-759.

92.   Godoy, P., et al., *Andes virus antigens are shed in urine of patients with acute hantavirus cardiopulmonary syndrome.* J Virol, 2009. **83**(10): p. 5046-55.

93.   Martinez, V.P., et al., *Person-to-person transmission of Andes virus.* Emerg Infect Dis, 2005. **11**(12): p. 1848-53.

94.   D'Souza, M.H. and T.R. Patel, *Biodefense Implications of New-World Hantaviruses.* Frontiers in bioengineering and biotechnology, 2020. **8**: p. 925-925.

95.   Vaheri, A., et al., *Hantavirus infections in Europe and their impact on public health.* Rev Med Virol, 2013. **23**(1): p. 35-49.

96.   Astorga, F., et al., *Distributional ecology of Andes hantavirus: a macroecological approach.* Int J Health Geogr, 2018. **17**(1): p. 22.

97.   Prescott, J., H. Feldmann, and D. Safronetz, *Amending Koch's postulates for viral disease: When "growth in pure culture" leads to a loss of virulence.* Antiviral research, 2017. **137**: p. 1-5.

98.   Safronetz, D., et al., *Pathophysiology of hantavirus pulmonary syndrome in rhesus macaques.* Proc Natl Acad Sci U S A, 2014. **111**(19): p. 7114-9.

99.   Safronetz, D., et al., *Hamster-adapted Sin Nombre virus causes disseminated infection and efficiently replicates in pulmonary endothelial cells without signs of disease.* J Virol, 2013. **87**(8): p. 4778-82.

100. Warner, B.M., et al., *Development and Characterization of a Sin Nombre Virus Transmission Model in Peromyscus maniculatus.* Viruses, 2019. **11**(2).

101. Djebali, S., et al., *Landscape of transcription in human cells.* Nature, 2012. **489**(7414): p. 101-8.

102. Mattick, J.S., *Non-coding RNAs: the architects of eukaryotic complexity.* EMBO Rep, 2001. **2**(11): p. 986-91.

103. Ule, J. and B.J. Blencowe, *Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution.* Mol Cell, 2019. **76**(2): p. 329-345.

104. Hombach, S. and M. Kretz, *Non-coding RNAs: Classification, Biology and Functioning.* Adv Exp Med Biol, 2016. **937**: p. 3-17.

105. Schmitz, S.U., P. Grote, and B.G. Herrmann, *Mechanisms of long noncoding RNA function in development and disease.* Cell Mol Life Sci, 2016. **73**(13): p. 2491-509.

106. Iyer, M.K., et al., *The landscape of long noncoding RNAs in the human transcriptome.* Nat Genet, 2015. **47**(3): p. 199-208.

107. Wang, K.C. and H.Y. Chang, *Molecular mechanisms of long noncoding RNAs.* Molecular cell, 2011. **43**(6): p. 904-914.

108. Cabili, M.N., et al., *Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution.* Genome Biol, 2015. **16**(1): p. 20.

109. Yao, R.W., Y. Wang, and L.L. Chen, *Cellular functions of long noncoding RNAs.* Nat Cell Biol, 2019. **21**(5): p. 542-551.

110. Geisler, S. and J. Coller, *RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts.* Nat Rev Mol Cell Biol, 2013. **14**(11): p. 699-712.

111. Chi, Y., et al., *Long Non-Coding RNA in the Pathogenesis of Cancers.* Cells, 2019. **8**(9).

112. Jonas, K., G.A. Calin, and M. Pichler, *RNA-Binding Proteins as Important Regulators of Long Non-Coding RNAs in Cancer.* Int J Mol Sci, 2020. **21**(8).

113. Chillón, I. and M. Marcia, *The molecular structure of long non-coding RNAs: emerging patterns and functional implications.* Crit Rev Biochem Mol Biol, 2020. **55**(6): p. 662-690.

114. Mercer, T.R. and J.S. Mattick, *Structure and function of long noncoding RNAs in epigenetic regulation.* Nat Struct Mol Biol, 2013. **20**(3): p. 300-7.

115. Butcher, S.E. and A.M. Pyle, *The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks.* Acc Chem Res, 2011. **44**(12): p. 1302-11.

116. Kilburn, D., et al., *Molecular crowding stabilizes folded RNA structure by the excluded volume effect.* J Am Chem Soc, 2010. **132**(25): p. 8690-6.

117. Tan, Z.J. and S.J. Chen, *Salt contribution to RNA tertiary structure folding stability.* Biophys J, 2011. **101**(1): p. 176-87.

118. Ransohoff, J.D., Y. Wei, and P.A. Khavari, *The functions and unique features of long intergenic non-coding RNA.* Nat Rev Mol Cell Biol, 2018. **19**(3): p. 143-157.

119. Huarte, M., et al., *A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response.* Cell, 2010. **142**(3): p. 409-19.

120. Sullivan, K.D., et al., *Mechanisms of transcriptional regulation by p53.* Cell Death & Differentiation, 2018. **25**(1): p. 133-143.

121. Vousden, K.H. and D.P. Lane, *p53 in health and disease.* Nature Reviews Molecular Cell Biology, 2007. **8**(4): p. 275-283.

122. Tang, S.S., B.Y. Zheng, and X.D. Xiong, *LincRNA-p21: Implications in Human Diseases.* Int J Mol Sci, 2015. **16**(8): p. 18732-40.

123. Chen, S., et al., *LincRNa-p21: function and mechanism in cancer.* Med Oncol, 2017. **34**(5): p. 98.

124. Yoon, J.H., et al., *LincRNA-p21 suppresses target mRNA translation.* Mol Cell, 2012. **47**(4): p. 648-55.

125. White, B.D., A.J. Chien, and D.W. Dawson, *Dysregulation of Wnt/β-catenin signaling in gastrointestinal cancers.* Gastroenterology, 2012. **142**(2): p. 219-32.

126. Yang, N., et al., *LincRNA-p21 activates endoplasmic reticulum stress and inhibits hepatocellular carcinoma.* Oncotarget, 2015. **6**(29): p. 28151-63.

127. Catana, C.-S., D. Gulei, and I. Berindan - Neagoe, *New insights into the role of non-coding RNAs as transcriptional targets of p53.* Endogenous locus-driven H-Ras G12V expression induces senescence-like phenotype in primary fibroblasts of the Costello syndrome mouse model, 2017: p. 43-49.

128. Hadjiargyrou, M. and N. Delihas, *The intertwining of transposable elements and non-coding RNAs.* Int J Mol Sci, 2013. **14**(7): p. 13307-28.

129.   Feschotte, C. and E.J. Pritham, *DNA transposons and the evolution of eukaryotic genomes.* Annu Rev Genet, 2007. **41**: p. 331-68.
130.   Daniel, C., M. Behm, and M. Öhman, *The role of Alu elements in the cis-regulation of RNA processing.* Cell Mol Life Sci, 2015. **72**(21): p. 4063-76.
131.   Ullu, E. and C. Tschudi, *Alu sequences are processed 7SL RNA genes.* Nature, 1984. **312**(5990): p. 171-2.
132.   Willard, C., H.T. Nguyen, and C.W. Schmid, *Existence of at least three distinct Alu subfamilies.* J Mol Evol, 1987. **26**(3): p. 180-6.
133.   Polak, P. and E. Domany, *Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes.* BMC Genomics, 2006. **7**: p. 133.
134.   Gong, C. and L.E. Maquat, *lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements.* Nature, 2011. **470**(7333): p. 284-8.
135.   Hanke, J.H., J.E. Hambor, and P. Kavathas, *Repetitive Alu elements form a cruciform structure that regulates the function of the human CD8 alpha T cell-specific enhancer.* J Mol Biol, 1995. **246**(1): p. 63-73.
136.   Grechishnikova, D. and M. Poptsova, *Conserved 3' UTR stem-loop structure in L1 and Alu transposons in human genome: possible role in retrotransposition.* BMC Genomics, 2016. **17**(1): p. 992.
137.   Holdt, L.M., et al., *Alu elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks.* PLoS Genet, 2013. **9**(7): p. e1003588.
138.   Chillón, I. and A.M. Pyle, *Inverted repeat Alu elements in the human lincRNA-p21 adopt a conserved secondary structure that regulates RNA function.* Nucleic Acids Res, 2016. **44**(19): p. 9462-9471.
139.   Qin, H., et al., *RNA-binding proteins in tumor progression.* J Hematol Oncol, 2020. **13**(1): p. 90.
140.   Pereira, B., M. Billaud, and R. Almeida, *RNA-Binding Proteins in Cancer: Old Players and New Actors.* Trends Cancer, 2017. **3**(7): p. 506-528.
141.   Li, Z. and P.D. Nagy, *Diverse roles of host RNA binding proteins in RNA virus replication.* RNA Biol, 2011. **8**(2): p. 305-15.
142.   Diosa-Toro, M., et al., *Role of RNA-binding proteins during the late stages of Flavivirus replication cycle.* Virol J, 2020. **17**(1): p. 60.
143.   Zhu, J., et al., *RNA-binding proteins that inhibit RNA virus infection.* Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3129-34.
144.   Wang, Y., et al., *Roles of long non-coding RNAs and emerging RNA-binding proteins in innate antiviral responses.* Theranostics, 2020. **10**(20): p. 9407-9424.
145.   Salim, N.N., et al., *Targeting a Novel RNA-Protein Interaction for Therapeutic Intervention of Hantavirus Disease.* J Biol Chem, 2016. **291**(47): p. 24702-24714.

Chapter 2: Biodefence Implications of New-World Hantaviruses[1]

2.0 Introduction

Hantaviruses are emerging zoonotic viruses that are responsible for two human diseases: Hantavirus Cardiopulmonary Syndrome (HCPS) associated with New-World Hantaviruses found in the western hemisphere; and Haemorrhagic Fever with Renal Syndrome (HFRS) associated with Old-World Hantaviruses in the eastern hemisphere [1]. Collectively, 150,000 – 200,000 cases of hantavirus disease are reported annually with the majority of HFRS cases occurring in Asia, specifically in the People's Republic of China which constitutes upwards of 90% of cases [2-4]. HCPS, comparatively, presents a stark minority in annual cases, roughly 300, with most cases being in South America and primarily Brazil [5, 6]. The annual average cases of New-World HCPS-causing hantaviruses in the western hemisphere are summarized in **Figure 2.1**. Both HFRS and HCPS exhibit drastically different mortality rates, with the former causing upwards of 12% while the latter inflicting upwards of 35-50% mortality on infected persons [7, 8]. Due to hantaviruses being emerging pathogens with HCPS-causing infections retaining a high mortality rate, there remains a possible risk of hantaviruses being engineered into novel bioweapons [9-11].

Biological systems that can potentially be used as weapons have been divided into three groups designated Category A, B, and C (**Table 2.1**) [12]. Category A agents are described as organisms or toxins that pose a national security risk because they can be easily transmitted or disseminated, can result in high mortality with a major public health impact, can cause public panic and social disruption, and require special action to ensure public health preparedness [12]. Category A agents include *Bacillus anthracis*, *Clostridium botulism* neurotoxin, and viral haemorrhagic fever viruses such as Ebola and Marburg. Category A agents are especially important because of their high mortality rate and rapid disease progression. For example, the spores of *B. anthracis* are highly resistant to adverse environmental conditions such as heat, cold, humidity, and radiation [13]. These spores are easily produced in laboratories, dried, and refined as a powder that can be released as an aerosol, which if inhaled, can result in inhalation anthrax, meningitis, and bacteraemia. If untreated, the disease is highly fatal [14].

---

[1] Reprinted with permission from D'Souza, MH., and Patel, TR (2020). Biodefense Implications of New-World Hantaviruses. *Frontiers in Bioengineering and Biotechnology*, 925(8). Differences are unique to this publication.

A

B
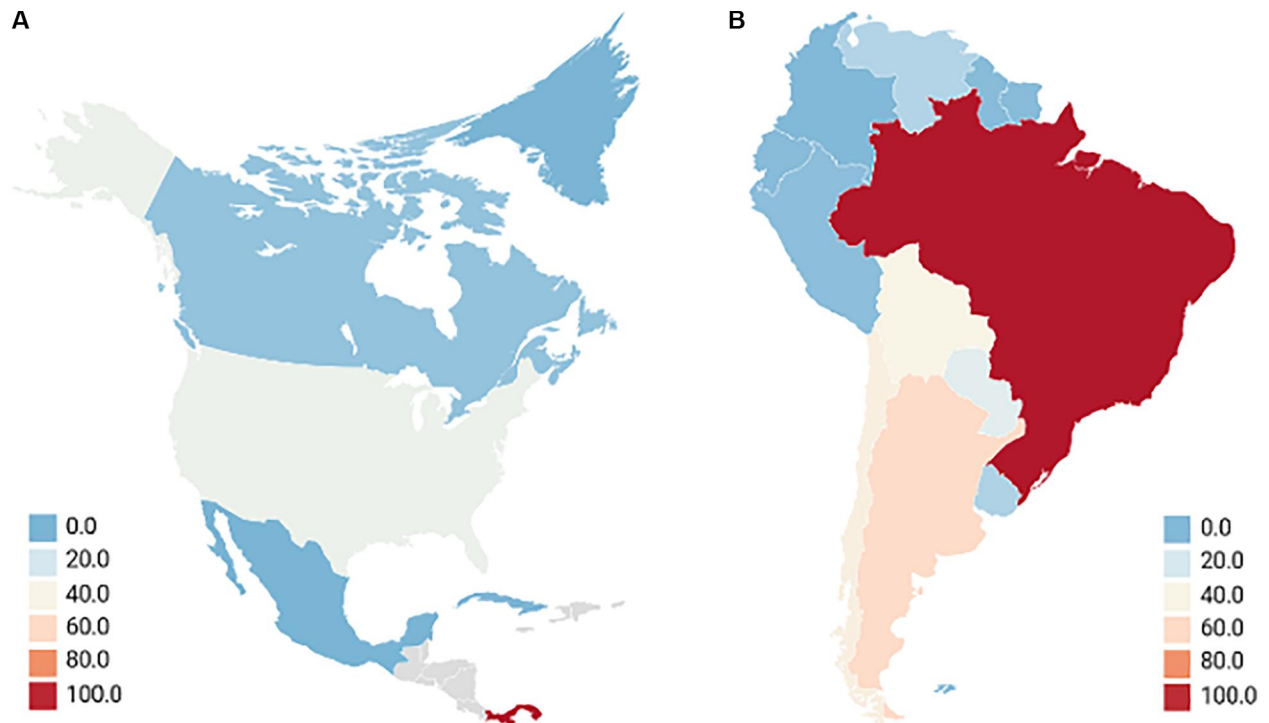
Figure 2.1: Annual Average Cases of New-World Hantaviruses in the Western Hemisphere. North America (**a**) Although Mexico has reported there being no HCPS cases, seroprevalence of hantaviruses exists in 10.15% of rodents, much of which occurs in Mexican states that border the USA where 299 cases of HCPS were reported between 1993 to 2017. Seropositive humans were identified, and the lack of reporting is attributed to the febrile disease being misconstrued with other illnesses [15]. This is very similar to other Central Latin American nations which have very limited reporting or insufficient data but show seroprevalence of hantavirus in rodents, up to 20.8% in Honduras as an example [16-19]. Grey locations indicate countries with no reporting of hantavirus. South America (**b**) Most cases occurred in rural or forested environments with farming being a major concern. Reporting is an issue as to actual annual cases for Brazil, Columbia, and Venezuela are considered to be significantly higher [18, 20-27]. Averages of HCPS-causing hantavirus cases taken from studies carried out between 2000 – 2019.

Category B agents are the second-highest priority agents and typically include agents that are responsible for moderate morbidity and low mortality rates, and are moderately dispersible [12]. They tend to include food safety threats and diseases from toxins like the Ricin toxin from Castor beans (*Ricinus communis*) that can be employed in local attacks and assassinations that have a low death rate compared to Category A agents but can still inflict significant damage to political and social systems [28]. This contrasts with the third-highest priority, Category C Pathogens which can include emerging pathogens that could be engineered for mass dissemination through their: availability, ease of production and dissemination; and their potential for high mortality resulting in a major health impact [12]. Hantavirus weaponization is speculative as there are no known major weapon development programs occurring. However, their weaponization remains attractive due to their potential to cause high mortality (up to 60% during the height of the 1993 Four Corners outbreak), and their ability to target young and healthy adults in risk occupations such as agriculture and forestry [29, 30]. Ultimately, the priority difference between the two doesn't make one any less meaningful, as

the employment of bioagents from any category of bioweapon could have a public health impact with implications to national security.

Table 2.1: Centers for Disease Control and Prevention (CDC) Biological Agent Categories [12, 31].

Category	Category Definition	Diseases	Organisms and Agents
A	High-priority agents that include organisms that pose a risk to national security because they: • Can be easily disseminated or transmitted person-to-person • Result in high mortality and have the potential to cause a major public health impact • Might cause public panic and social disruption • Require special action for public health preparedness	Anthrax Botulism Plague Smallpox Tularaemia Viral Haemorrhagic Fevers (VHFs)	*Bacillus anthracis* Clostridium botulinum toxin *Yersinia pestis* Variola major *Francisella tularensis* Filoviruses (Ebola, Marburg) Arenaviruses (Lassa, Machupo)
B	Second highest priority agents which include those that: • Are moderately easy to disseminate • Result in moderate morbidity rates and low mortality • Require specific enhancements of laboratory capacity and enhanced disease surveillance	Brucellosis Epsilon Toxin Food Safety Threats Glanders Melioidosis Psittacosis Q Fever Ricin Toxin Staphylococcal Enterotoxin B Typhus Fever Viral Encephalitis Water Safety Threats	Brucella species *Clostridium perfringens* Salmonella species *Escherichia coli* O157:H7 Shigella *Burkholderia mallei* *Burkholderia pseudomallei* *Chlamydia psittaci* *Coxiella burnetii* *Ricinus communis* (Castor beans) *Staphylococcus aureus* *Rickettsia prowazekii* Alphaviruses (Venezuelan Equine Encephalitis, Eastern Equine Encephalitis, Western Equine Encephalitis) *Vibrio cholerae* *Cryptosporidium parvum*
C	Third highest priority agents include emerging pathogens that could be engineered for mass dissemination in the future because of: • Availability • Ease of production and dissemination • Potential for high morbidity and mortality causing major health impacts	Emerging Infectious Diseases	Nipah Virus Hantaviruses Tick-Borne Hemorrhagic Fever Viruses Tick-Borne Encephalitis Viruses Yellow Fever Multidrug-Resistant Tuberculosis

Category considerations of biological agents vary between the Centers of Disease Control and Prevention (CDC) and the US National Institute of Allergy and Infectious Diseases (NIAID) based upon circumstance of the

infectious agent. The CDC assesses bioagent risks in support of US public health systems and primary healthcare providers and how, based upon the categorization, they should respond to biological agents and pathogens including those that seldom occur in the US [31]. NIAID's categories refer to documented priority pathogens A, B, and C, and emerging infectious diseases defined as those that have newly appeared in a population or have existed but are rapidly increasing in incidence or geographic range [32]. There is much categorical confusion though for hantaviruses, and specifically New World Hantaviruses such as Sin Nombre *orthohantavirus* (SNV), as to what priority of a bioagent and subsequent threat they pose [1, 33]. Hantaviruses as a whole are categorized as emerging viruses along with Nipah Virus in the CDC as Category C Pathogens; whereas NIAID places hantaviruses as part of the Category A Pathogens [31, 32]. Hantaviruses, specifically Old-World Hantaviruses causing HFRS, are listed in Category C due to their shared symptoms to other agents causing Viral Hemorrhagic Fever (VHF) that cause capillary leakage syndrome and haemorrhaging [34]. Category C retains the lowest priority of risk to national security and ultimately the lowest potential as a biowarfare agent. However, it doesn't diminish the risk that hantaviruses pose globally. With their widespread nature, being present on every continent except for Australia and Antarctica, hantaviruses continue to pose a risk to human systems and activities that closely engage with their rodent-specific reservoirs including military personnel, agricultural workers, and transport industries including warehouse and shipping staff [35]. This review paper seeks to clarify the categorization of hantaviruses as bioweapons as well as to define the practicality of employing hantaviruses, specifically HCPS-causing SNV and Andes Virus (ANDV), as novel bioagents against modern militaries and industries.

2.1 HCPS-Causing New-World Hantaviruses

The hantavirus genus forms part of the *Bunyavirus* family and is composed of well-defined serotypes that are each associated with a specific primary rodent reservoir [36, 37]. The hantavirus genome is tripartite and is composed of three segments of negative-sense, single-stranded RNA [38]. The three segments are organized by size and are designated as the Large (L), Medium (M), and Small (S) segments since the tripartite genome lengths are generally 6.6 kb, 3.7 kb, and 2.1 kb for the L, M, and S segments respectively (**Figure 2.2**) [39]. The genomic L, M, and S segments encode for the 250 kDa RNA-dependent RNA polymerase (RdRp), 125 to 127 kDa Glycoprotein Precursor (GPC) and subsequent co-translationally cleaved Gn and Gc Glycoproteins, and the 48 kDa Nucleocapsid (N) Protein respectively [38, 40, 41]. SNV, amongst other hantaviruses with the exception of Hantaan Virus (HTNV), Seoul Virus

(SEOV), and Dobrava Virus (DOBV), have an open reading frame (ORF) for a putative Non-Structural Protein (NSs ranging between 7 to 10 kDa in size) [39]. Additionally, each genomic segment is flanked by 5′ and 3′ Non-Coding Terminal Regions (NTRs) which are common to Bunyaviruses [42].
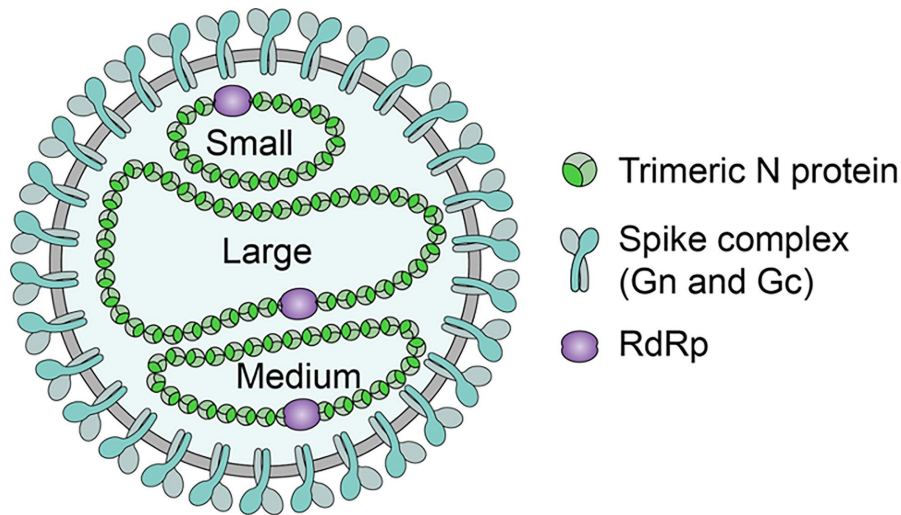
Figure 2.2: Hantavirus Structure. Hantaviruses are enveloped with a lipid bilayer containing Glycoprotein spikes assemblies comprised of Gn and Gc Glycoproteins. Contained within the envelope are the equimolar amounts of N Protein packaged S (small), M (medium), and L (large) segments vRNA, which are associated with a RdRp [43, 44]. SNV structure is generally spherical with a dense envelope [45]. N Protein forms trimers that selectively encapsidates the negative-sense vRNA into RNPs and assists in its panhandle formation for packaging [46].

The L Segment's RdRp acts as the RNA transcriptase and replicase, transcribing mRNA and replicating the genomic RNA using the positive-sense RNA as an intermediate [47] (**Figure 2.3**). Hantaviral RNA segments are each associated with the RdRp and are packaged within a ribonucleoprotein complex formed by the N Protein [43]. The RdRp is responsible for vRNA transcription and replication, additionally retaining endonuclease activity which is used to cleave the 5′-termini of host mRNA to act as a primer which initiates viral mRNA transcription in a process called cap-snatching and prime and realignment [47]. This occurs in conjunction with the N Protein which is found to form an N-RdRp complex for RNA synthesis whilst also binding to mRNA caps by recognizing a five-nucleotide sequence adjacent to the 5′cap for high-affinity binding [48, 49]. During cap-snatching, the viral RdRp binds to methylated capped 5′ends of host mRNAs and cleaves them for use as a primer for mRNA synthesis with a preference for host mRNAs that contain a Guanine prior to the cleave site [50]. The prime and realignment follow the methylated 5′cap whose aforementioned G nucleotide at the -1 position would align opposite a Cytosine nucleotide at the +3 position on the negative-sense vRNA genome. After the primer is extended up to 3 nucleotides, the nascent chain will realign to shift the original 3′ Guanine back to the -1 position ultimately generating two to four UAG repeats [51].

Pathogenic hantaviruses retain glycoproteins that target and interact with the β_3 chain of Integrins that is especially abundant as a surface receptor on endothelial cells, dendritic cells (DC), and platelets where they are critical in maintaining capillary integrity [52]. Endothelial cells are ubiquitously infected throughout the body by HFRS and HCPS-causing hantaviruses, however, pulmonary endothelial cells are the primary targets during HCPS infections with DCs and platelets being involved in the pathogenic process of vascular leakage and thrombocytopenia [53-55]. Hantaviruses inducing HCPS employ $\alpha 11\beta 3$ Integrins with hantaviruses inducing HFRS employing $\alpha v\beta 3$ Integrins for entry both of which are β_3 Integrins; $\alpha 5\beta 1$ Integrins are employed by non-pathogenic hantaviruses [52, 56]. The hantavirus virion is itself enveloped retaining a lipid bilayer whose membrane is 5 nm thick and studded with the Gn and Gc glycoprotein spike assemblies that project 10 nm from the membrane in fourfold rotational symmetry [43]. The virion's shape appears as a rounded, pleiomorphic particle ranging between 70 to 350 nm in diameter [39, 43]. The M Segment retains a five amino acid sequence (WAASA) that precedes the co-translational cleavage site for the GPC which is conserved across all hantaviruses [37].

The N Protein's role is multifaceted but is primarily involved in the encapsidation of the vRNA and protects it from host cellular nucleases by binding selectively to Hantaviral panhandle structures [49, 57]. Each SNV segment possesses conserved terminal sequences at the 5′ and 3′ NTRs that are capable of complementarily base pairing to form panhandle structures [40]. These conserved sequences, consisting of 14 to 17 nucleotides, were found to be highly conserved throughout the hantavirus genus, being comprised of the following sequence at the 3′-termini: 3′AUCAUCAUCUGAGG-5′; and the following sequence at the 5′-termini: 5′-UAGUAGUAU(G/A)CUCC-3′ [40]. Trimeric N Protein subsequently recognizes these panhandle structures with specificity and encapsidates the vRNA, with trimerization being required for high-affinity binding [46, 58]. The N Protein is also genus-specific and can bind to vRNA and cRNA of other hantavirus species, acting as an RNA chaperone [59]. An RNA chaperone is a protein involved with regulating RNA activity by enabling them to fold into functional molecules whilst improving their interactions with RNA targets, and maintaining quality control over RNP biogenesis [60]. The N Protein performs as an RNA chaperone in an ATP-independent manner, whereby it non-specifically dissociates misfolded RNA structures using Mg^{2+} ions to unwind dsRNA helices [61]. The N Protein's RNA chaperone roles are also employed when interacting with the vRNA and its panhandle structures, assisting in the dissociation of the RNA duplexes and initiating replication by the RdRp [47, 62]. This nonspecific RNA helix-unwinding or RNA chaperone activity of the N protein utilizes a single-stranded region at the 3′ end of the vRNA duplex with dissociation occurring from the 3′ to 5′ direction

in either an enzymatic catalytic activity of displacement of the RNA duplex using cooperative single-stranded RNA binding interactions that unfolds the vRNA [62]. The N Protein's RNA chaperone properties also enable it to bind to misfolded vRNA, refolding it to allow the high-ordering of panhandle structures to form and to prevent RNA structures from falling into inoperable kinetic traps that adversely affects the hantaviral lifecycle [63]. Deletions of the first 50 amino acids from the N-terminus region of the N Protein results in negated interactions with the RdRp, affecting transcription and replication [48]. Similarly, Hantaviral N Protein mutants with various N-terminal deletions, whose coiled-coil domain is necessary for RNA binding, exhibit diminished binding to the vRNA panhandle regions which consequently affects trimerization, encapsidation, and packaging because of the loss of panhandle recognition [58, 64]. This has a negative effect on the hantaviral lifecycle.
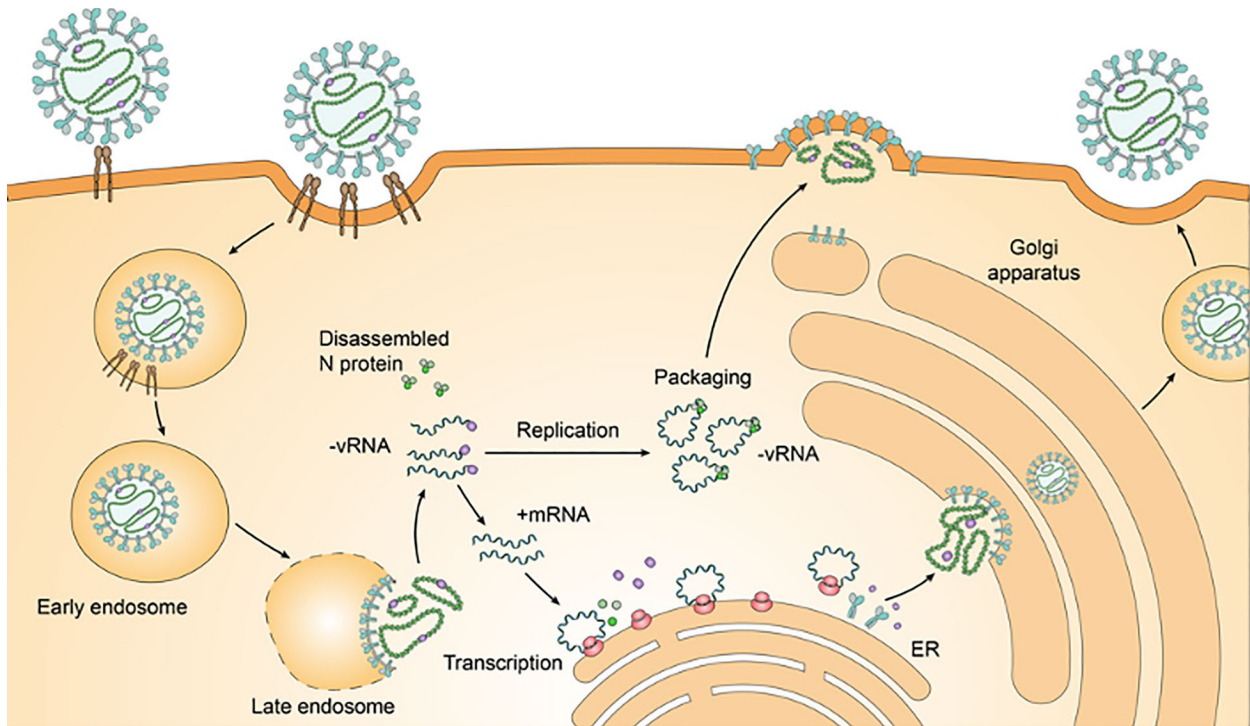
Figure 2.3: HCPS-Causing Hantavirus Life Cycle. Pathogenic HCPS-causing hantaviruses such as SNV or ANDV will first bind to β_3-Integrin receptors on endothelial cells which will mediate endocytosis. The formation of an early endosome helps traffic the virion to the Golgi Complex. Following a pH-mediated membrane fusion, the now late endosome disassembles and releases the ribonucleoproteins (RNPs) near the Endoplasmic Reticulum-Golgi Intermediate Compartment (ERGIC). The RNPs disassemble and the RdRp carries out transcription and replication in the cytoplasm, cleaving cellular mRNA to form capped primers to initiate viral mRNA transcription. Transcribed S, M, and L Segment mRNA is translated into N Protein, GPC (and then into co-translated Gn and Gc Glycoproteins), and RdRp respectively. Negative sense vRNA serves as the template for the transcription of mRNA. RdRp undergoes a transition from transcription to the replication of negative-sense vRNA which is considered to be mediated by the increase in free N Protein concentration [65]. New-World Hantaviruses will be assembled at the plasma membrane compared to Old-World Hantaviruses that are assembled at the Golgi complex [66]. Nascent virions bud from the plasma membrane.

SNV and ANDV are the two major causative agents of HCPS, with ANDV occurring in South America which is spread by the *Sigmondontinae* subfamily and mainly by *Oligoryzomys longicaudatus* or the long-tailed pygmy rice rat reservoir [67, 68]. Before the presence of SNV, hantaviruses were not considered to be a serious public health threat as other hantaviruses like Prospect Hill Virus (PHV) and HFRS-causing SEOV have been found in several US cities [37]. The disease took significant attention when the respiratory illness of HCPS was first observed in the Four Corners region of the United States in 1993, with the outbreak causing upwards of 75% mortality in healthy adults between the ages of 20 to 40 years [38]. The deer mouse (*Peromyscus maniculatus*) was identified as the primary rodent reservoir for SNV; it is one of the most abundant small mammals in North America, prevalently found throughout the Canadian prairies and the American Midwest [37]. The emergence of SNV is likely due to environmental factors that favoured the natural reservoir of deer mice to increase, allowing for increasing opportunities for human infection [69]. The reoccurrence and increased rodent-human contact can be attributed to increased food availability from erratic weather conditions that produced higher precipitation and warmer climates [6]. Climate change and its impact of lowering biodiversity has caused a dilution effect that altered reservoir behaviour and forced population migration, and the ultimate spread of the infectious agents to human systems [6].

The reservoirs themselves can spread the virus horizontally, being nearly asymptomatic but chronically infected [70]. HFRS and HCPS are generally acquired from the inhalation of aerosolized excreta including faeces, urine, and saliva infected with hantavirus [71-74]. This can also include the direct contamination of food or household articles with rodent excreta as well as virion particles shed from rodent skin and fur [75, 76]. However, SNV transmission requires direct contact between SNV-infected rodents and humans as contaminated cages proved to be ineffectual in transmitting the disease horizontally to uninfected deer mice [74]. ANDV is unique in that it can transmit hantavirus through person-to-person contact [77]. No other hantavirus exhibits the property of person-to-person transmission, which makes ANDV a preferable candidate for weaponization, which would take advantage of the additional spread mechanism. No person-to-person hantavirus infections have been reported in North America, making SNV less effective as a bioweapon, comparably [78]. The person-to-person transmission of ANDV occurs mainly in family clusters or close activities with infected case-patients during the disease's prodrome phase, occurring during the interval of 12 to 27 days between the initial exposure and the onset of symptoms [22]. Sexual partners have a higher risk of infection compared to non-sexual partners [79].

Hantaviruses generally enter cells utilizing a clathrin-dependent pathway which follows the formation of an early endosome and subsequent low-pH initiating dissolution of the late endosome for infectious entry [80] (**Figure 2.3**). SNV and ANDV can enter endothelial cells primarily by a receptor-mediated endocytic pathway involving β₃-integrins but also a clathrin-dependent pathway [52, 81]. New-World Hantaviral replication occurs predominately in pulmonary endothelial cells which have exhibited the highest viral loads, resulting in increased vascular permeability [76]. Macrophages, follicular DCs, and DCs are also known to replicate the Hantaviral genome with the virus also being found in human tissues of the kidney, spleen, pancreas, lymph nodes, skeletal muscles, heart, intestines, adipose tissue, urinary bladder, and brain [82]. Replication occurs in the cytoplasm with the budding of the Hantaviral virion occurring in the Endoplasmic Reticulum Golgi Intermediate Complex (ERGIC); SNV uniquely, but principally, buds from the plasma membrane [45, 83] (see **Figure 2.3**).

Hantavirus infections activates the innate immune system with downstream effects that induces disease. The innate immune system recognizes pathogens through their interaction with Pattern Recognition Receptors (PRRs) which are expressed by many cell types, including endothelial and epithelial cells. Viruses present Pathogen-Associated Molecular Patterns (PAMPs) which are recognized by PRRs which activate signalling cascades and transcription factors that modulate the expression of type I Interferons and Interferon-Stimulating Genes (ISGs) involved in antiviral functions. Toll-like Receptors (TLRs) and Retinoic acid-Inducible Gene-I (RIG-I), including the RIG-I-like receptor Melanoma Associated Gene 5 (MDA5), are PRRs that are involved in the recognition of pathogenic RNA viruses by binding to vRNA [84]. Binding of PAMPs in the form of vRNA to TLRs and RIG-I receptors activates transcription factors NF-κB and IFN Regulatory Factor 3 and 7 (IRF3/7) that are translocated into the nucleus to bind to ISGs that are used to express Interferons (IFN) [85, 86]. Type I Interferons (IFN-α/β) are critical regulators of immune cell activation, development towards antiviral activity, cell growth, and apoptosis and are involved in stimulating the Janus Kinases and signal and activators of transcription pathways (JAK/STAT) [87].

New World Hantaviral proteins antagonize virus recognition by suppressing the JAK/STAT signalling pathways with ANDV utilizing its N Protein and the GPC to disrupt antiviral activity while SNV employs its GPC alone [88]. The N Protein of hantaviruses has been reported to inhibit IFN activity and NF-κB activation, with ANDV N Protein inhibiting signalling responses instigated by RIG-I and MDA5 and upstream IRF3 phosphorylation [89-91]. Both HTNV and ANDV N Proteins inhibit Tumour Necrosis Factor α induced activation from NF-κB by preventing the transcription factors translocation into the nucleus [90]. Additionally, reports identified the highly conserved

30

domains of the Glycoprotein Gn's cytoplasmic tail which also functions in early IFN responses by blocking IRF3 and NF-κB activation and subsequent downstream antiviral function of the early immune response [92, 93]. Vero cell lines are used preferably to isolate and amplify hantaviruses since they are deficient in IFN-I and IFN-II expression and will not elicit an immune response to infection although New World Hantaviruses have been shown to elicit IFN-λ activation in Vero cell lines [94-96]. Pre-treatment of IFN-λs have been shown to induce antiviral activity against HTNV infection by activating the JAK-STAT pathway in A549 cells [97]. Pathogenic hantaviruses tend to regulate the early induction of IFN to replicate successfully with pre-treated Type I IFNs only being successful shortly after infection [93].

2.2 Symptoms

Patients suffering from HCPS generally present fever, headache, muscle aches, and chills as well as leukocytosis and thrombocytopenia, which rapidly progresses to more severe respiratory diseases [69]. After 4-10 days, individuals infected with HCPS-causing hantaviruses developed influenza-like illnesses followed by rapidly progressing pulmonary oedema caused by pulmonary capillary leak syndrome, resulting in respiratory dysfunction and shock [4, 76]. HCPS is particularly important because unlike other respiratory diseases, it occurs in young, healthy adults [68]. Death occurred 2-10 days after the onset of the illness within almost 50% of patients observed [82]. Although HCPS shares some similarities with HFRS, like the febrile prodrome and capillary leakage, the kidneys are largely unaffected with capillary leakages occurring exclusively in the lungs and resulting in shock and cardiac complications despite sufficient tissue oxygenation [69].

2.3 Vaccines and Therapeutics

There are no US FDA-approved vaccines available for hantavirus infections, however, there are a variety of live-attenuated vaccines (Hantavax), DNA vaccines, subunit vaccines, and virus-like particle (VLP) vaccines that all demonstrate varying degrees of effectiveness [3, 98-101]. The Hantavax vaccine is available and is instituted in the Republic of Korea, with effectiveness against HFRS-causing hantaviruses such as HTNV and SEOV and resulting in a subsequent reduction in HFRS-related hospitalizations [102]. However, its immunogenicity is dependent on early booster vaccinations in tandem with its two-dose primary vaccination which was demonstrated to provide timely protection to high-risk groups like farmers and those in the military [103]. DNA vaccines that use recombinant

Vesicular-Stomatitis virus vectors expressing SNV and ANDV glycoproteins in Syrian hamster models were also effective at eliciting an immune response and conferred protection against lethal ANDV [104]. DNA vaccines are preferable because they can present the most immunogenic antigens to the host immune system whilst avoiding the need to propagate inactivated hantaviruses that are universally difficult to grow, isolate, and purify, with many DNA vaccines expressing Old-World Hantavirus glycoprotein genes and eliciting successful immune responses in hamster models [105].

Vaccine development for Hantavirus infections appears largely oriented around traditional techniques and could benefit from current developments utilizing mRNA vaccine technologies like that of the SARS-CoV-2 mRNA vaccines [106-108]. Many Hantaviral elements, which includes the Gn/Gc spike protein complex and the N Protein, can elicit the humoral immune response [109]. The Gn/Gc Glycoprotein spike is a composite of four Gn and four Gc protomers on the virion envelope, with exposure of N-terminally located Gn complex on the solvent side [110]. Although a likely candidate for mRNA vaccine application, the Gn experiences a higher frequency of mutations causing it to be a difficult candidate to sequence and model to account for new strains. The Gc complex, comparatively, is less exposed but maintains a greater conservation. Its inaccessibility causes it to be an unsuitable candidate. Despite this, employing antibodies to block Hantavirus entry could be an alternative strategy. Antibodies have had efficacy in targeting the $\alpha_v\beta_3$ Integrins on human umbilical vascular endothelial cells and were able to decrease the infectivity of pathogenic species like SNV or ANDV [111]. The N Protein is also highly antigenic and exhibits immunodominant antiparallel coiled-coil domains on its N-terminus whilst also possessing serotype specific epitopes on its C-terminus parallel coiled-coil structure, both of which can be targeted by antibodies [112]. Further characterization of non-neutralizing antibodies of the N Protein is needed to produce effective DNA or mRNA vaccines. However, it has been conceptualized that both recombinant and native N Protein can elicit a high immune response, attracting CD4 and CD8 T-cell activity and can be employed as a possible mRNA vaccine strategy [113].

There are currently no US FDA-approved post-exposure therapeutics against Hantaviral infections, however, there are treatment strategies present to manage HFRS and HCPS [3]. Virus-targeting antivirals including antiviral drugs, antibodies, or novel-small molecules are designed to block hantavirus entry or to reduce viral replication. Ribavirin is an effective anti-Hantaviral drug that affects the biological function of RdRp and has had some success in treating HFRS cases including protecting Syrian hamsters in lethal HCPS models [104, 114]. Ribavirin was effective at preventing lethal HCPS disease by having an inhibitory effect on ANDV replication [114]. Ribavirin also

inhibits SNV *in vitro* while the pre-treatment of deer mice followed by daily therapy of Ribavirin reduced SNV infection and viral RNA synthesis [115]. However, Ribavirin has some limitations as at high doses it is toxic to humans and animals and causes anaemia [116, 117]. It was also noted that intravenous Ribavirin was ineffective at treating HCPS-patients after the onset of the cardiopulmonary phase [116, 118]. Antivirals also function effectively only during the early infection stage and not after the start of viremia [119]. This could largely be attributed to the uncontrolled immune response which predominates the Hantaviral pathogenesis process after immediate infection [3]. The Hantaviral prodrome phase can also be difficult to differentiate from other febrile illnesses, which may benefit infection by impeding proper identification and treatment [119]. Another antiviral is Favipiravir that has shown broad-spectrum antiviral activity against RNA viruses including Bunyaviruses, being better than Ribavirin in that it is well-tolerated in humans without haemolytic anaemia related side effects [120]. Favipiravir was evaluated using *in vivo* studies for both SNV and ANDV infected hamster lethal disease models and resulted in complete survival as well as the reduction of ANDV RNA and antigens in the blood and lungs, although it was no longer effective after the onset of viremia in delayed antiviral treatment studies [119, 120].

Hantaviruses can be inactivated by heat (sustained 30 minutes at 60°C), detergents, UV radiation, organic solvents, and hypochlorite solutions [36]. Despite this, hantaviruses are fairly durable and unexpectedly stable outside of a host, being able to survive longer than 10 days at room temperature and more than 18 days between the -20°C to 4°C range [121]. For most hantaviruses, contaminated dust or aerosols can transmit the virus to other rodents for up to 15 days after being excreted with viral infectivity in the culture being lost within 5-11 days when incubated at 23°C [72, 122]. 70% Ethanol completely inactivates Bunyaviruses broadly, with HTNV being partially resistant to 30% Ethanol [122]. These are largely chemical prophylactics designed to maintain sanitation and treat hantavirus-contaminated facilities and would create risk for livestock and personnel unprotected by strong detergents or hypochlorite solutions. Consequently, the absence of any effective vaccines or therapeutics makes hantavirus infections particularly dangerous to those working or operating in risk environments including agriculture, forestry, mining, and military operations.

2.4 Biowarfare Potential of HCPS-Causing Hantaviruses

Hantaviruses have generally remained in the Category C position from the CDC and biodefense categorizations which is different from the laboratory biosafety criteria summarized in **Table 2.2**. Hantaviruses are

considered a biosafety level 3 bioagent with regards to NIH and across the European Union (EU), with the exception of HCPS-causing hantaviruses in the EU being considered a level 2 because their criteria differ with regards to an agent that causes human disease and might be a hazard to workers, but is unlikely to spread to the community, and there is usually effective prophylactic treatment available [123]. For comparison, the NIH treats Ebola virus as a level 4; *Bacillus anthracis* as a level 2; SARS-associated coronavirus (SARS-CoV) as a level 3; and Human Immunodeficiency Virus (HIV) as a level 3; regardless of the biosafety levels, both Ebola Virus and Anthrax are considered very high biothreats with hantavirus being a high threat in the EU [123]. BSL 3 laboratory requirements are intensive, especially for highly pathogenic diseases that can cause harm to materials and personnel. Incidentally, for a research or industrial laboratory to study hantaviruses they require: direct physical protection from the virus in the form of PPE including gloves, masks, gowns, respiratory protection, and positive pressure ventilation suits; Biosafety Cabinets (BSC) as primary containments to isolate the pathogen and the user; secondary containments to mitigate or prevent the pathogen's presence outside the BSC and its exit outside BSL 3 containment; and physical barriers in the form of walls, fences, or exclusion zones to prevent outside contamination [124]. The initial infrastructural costs and maintenance of BSL 3 containment protocols would be prohibitively expensive and complex for uninitiated bioterrorist organizations making its development by smaller, resource poor organizations unfeasible.

Table 2.2: Biosafety Categorization based upon the National Institute of Health (NIH) Criteria [123].

	Requirements
1	Agents that are not associated with disease in healthy adult humans
2	Agents that are associated with human disease which is rarely serious and for which preventive or therapeutic interventions are often available.
3	Agents that are associated with serious or lethal human disease for which preventive or therapeutic interventions may be available (can cause high individual risk but low community risk).
4	Agents that are likely to cause serious or lethal human disease for which preventive or therapeutic interventions are not usually available (can cause high individual risk and high community risk).

Hantaviruses are cited as being possible bioweapons that can be used against humans. When focused on specific serotypes of hantavirus, like SNV and ANDV, it becomes apparent that with their high mortality rate and rapid disease course with serious cardiopulmonary symptoms New-World Hantaviruses as opposed to HFRS-causing Old-World Hantaviruses are the more severe threat [125]. Since HCPS-causing New-World Hantaviruses exhibit a high mortality (up to 50% in older patients) but low morbidity, it would preclude them from the Category B bioweapons which specifically are classified by their moderate morbidity and low mortality rates [16]. This causes

SNV or ANDV bioagents to be assessed within Category A or C terms, although the moderate dissemination quality of Category B is reflective of hantaviruses and their limited projection by aerosols and rodents. Nevertheless, successful bioweapons though have very strict requirements listed in **Table 2.3**.

2.4.1 HCPS-Causing Reservoirs are Available and are affected by Environmental Factors

The presence of SNV-infected deer mice across the American Midwest is fairly high as seroprevalence of SNV antibodies were discovered in 38% of captured rodents in Indiana, with up to 25% of seroprevalence in the western US and 7% in the eastern US [126]. SNV-infected deer mice are somewhat discontinuous across Canada, but are located in every Canadian province as well as the Yukon territory and tend to display greater than 30% seroprevalence in large, close proximity populations [16]. Seropositivity of ANDV was prevalent across South America, particularly Patagonia in Chile and Argentina with antibodies being present at 5.9% specifically for *Oligoryzomys longicaudatus* [67, 127]. Male deer mice have a higher seroprevalence of SNV antibodies compared to female deer mice which is the same for *Sigmodontinae* species infected with ANDV [127, 128]. Consequently, acquiring HCPS-causing Hantaviruses is relatively easy and requires access to natural habitats and peridomestic environments that harbour the rodent reservoir. The relative abundance of HCPS-causing rodents will be dependent on precipitation but overall maintain high ecological densities [7].

Climate change will also have impacts to the acquisition and maintenance of Hantaviral reservoirs. Rodent population dynamics are particularly affected by a combination of unusually high rainfall followed by drought which is evidenced by the 1993 US Four Corners outbreak which was preceded by a dramatic increase in rainfall following the 1992-1993 El Niño warming phase [129]. These favourable conditions led to increases in rodent food sources and a significant increase in rodent population which took advantage of the Four Corners' environment which provided favourable habitats conducive for the growth of *P. maniculatus* [130]. This likely contributed to rising deer mouse populations which resulted in increased exposure of rodent-human contact, similar to the PUUV outbreak in Northern Europe which was also precipitated by an unusually wet spring season which affected bank vole populations beneficially [37]. The increase in North and Western European vole populations is adjusted by elevated average temperatures which improves mast production. Higher densities of rodents benefited from high seed production, itself improved by warmer summer conditions which benefited winter survival and subsequent spring breeding [131]. Incidentally, human-reservoir contact increased as the reservoir population increased.

Bioterrorist cells have the potential to take advantage of high rodent population densities. Having higher populations of asymptomatic but chronically infected rodent specimens can be utilized in either a one-target distribution model, or as a means to generate a critical concentration of passaged virions to achieve a weaponizable aerosol. Both methods would require a large-scale capture and maintenance of rodents, with the latter being more onerous in the process of passaging and isolating the hantavirus. However, the role of climate change provides access for bioterrorist groups to acquire the virus through freely available infected rodents because of their increased populations. This can change depending on the effects of human activity which is being accelerated by agricultural expansion, deforestation, land reclamation, irrigation projects, and infrastructural developments [131].

2.4.2 Difficulty in Cell Culturation Reduces Ease of Production

Hantaviruses have historically been very difficult to isolate and grow in both cell culture and animal models, which have limited their ability to be previously concentrated and weaponized [40]. The first successful passage of HTNV in a laboratory setting occurred in 1978, and the first successful passage and isolation of SNV occurred in 1994 [132]. The virus itself requires passaging by rodent-to-rodent transmission followed by cell culturing in Vero E6 cells, with the virus replicating specifically in *P. maniculatus* cells despite repeated attempts of using RT-PCR to amplify positive hantavirus from human or rodent samples [132]. Isolation from the reservoir host or from diseased human patients tends to require extensive blind passaging in cell culture to acquire adequate viral titres for characterization studies, with viral propagation being observed to elicit reduced infectivity in natural rodent reservoirs [133, 134]. SNV propagation in Vero cultures seems to cause mutations in the RdRp which potentially attenuates the virus and makes it less virulent [135]. The problem arises from attempting to adapt the viruses to new hosts through sequential passaging from animal to animals as well as amplifying the virus in large stocks of Vero cell lines which have resulted in the attenuation of the viral culture [136]. Conversely, attempts at experimentally recreating signs and symptoms of HFRS or HCPS in a nonhuman primate model demonstrated that various nonhuman species can be infected by the disease but they do not develop obvious symptoms. This trend is observed in the attenuation of the Old-World Hantavirus PUUV in cell culture due to point mutations occurring in its S Segment. PUUV's propagation in Vero E6 cells replicated with high efficiency but did not retroactively infect its natural reservoir host the bank vole (*Clethrionomys glareolus*) or cause severe disease in cynamalogous macaques [137-139]. Comparatively, SNV

36

propagated in deer mice after passaging in Vero cel lines elicited severe disease in its nonhuman primate model of rhesus macaques [135].

Table 2.3: Summary of ideal Biological Warfare Requirements adapted and modified from Meyer and Morse, and Christian [10, 12]. Table also includes the capabilities required by conventional or unconventional militaries to conduct and deliver a bioagent attack. Note that the ranking is unimportant. We also describe the risks associated with bioweapons requirements as they pertain to the feasibility of developing and deploying New-World Hantaviruses. Risk is graduated regarding Low, Medium, and High.

	Requirements	Risk	Condition
1	Availability in the Environment	Medium	Wide range of Rodent Host Reservoirs in North and south America
		Medium	Possible Attenuation through Passaging
2	Ease of Design and Production		Few Nonhuman Primate Lethal Disease Models Difficulty in Isolation and Purification of Virions New Technology improving yields and Virulence of Passaged Virions
3	Stability after Production and Persistence in the Environment	Medium	Long Durability and Persistence in Contained and Isolated Environments Sensitivity to Light and Heat
4	Effective Transmission Pattern and Routes of Entry	Medium	Inhalation of both ANDV and SNV Person-to-Person Transmission of ANDV
5	Effective Delivery Systems and Mode of Transportation	Medium	Effective Deployment Indoors Rodent Delivery is Onerous and Resource Intensive
6	Susceptible Target Population	High	Novel Emerging Infectious Disease with no known natural immunity within Human Populations
7	Absence of Specific and Effective Treatment including countermeasures that have the Ability of a Vaccine to Protect Certain Groups	High	No US FDA approved Antivirals or vaccines Some Antivirals and Vaccines for Old-World Hantaviruses with varying Degrees of Effectiveness
8	The Ability to Incapacitate or Kill Target Host	Medium	High Mortality Rate Low Morbidity Rate
9	Appropriate Particle Size for Aerosolization and Airborne Transmission	Medium	Can be Aerosolized Airborne Efficiency within Closed Environments
10	The Ability to be Disseminated in Food or Water Supplies	High	Can Contaminate Food and Water as well as commercial products
11	Logistic Requirements to Manufacture and Disperse Bioagents which include Infrastructural and Financial Support, Expertise, and Organizational Capabilities	Low	Intensive Laboratory Equipment and High Expertise Requirements High Costs Deployment of Infected Rodents Reservoirs is Demanding

The previous reporting of hantavirus being difficult to isolate have also been attributed to the low concentrations of infectious virion particles extracted from the clinical or wild-caught infected rodents, with virion replication peaking at the time of death for the HCPS-infected human patients [40]. Combined with the slow and non-cytopathic growth of hantaviruses in cell culture are considerations as to why isolation becomes onerous [40]. Passaging has been successful in non-rodent, non-human primate models involving rhesus macaques, but they had to

be previously passed in deer mice to maintain virulence and infectivity which increases the requirements for weaponization [74, 135]. Furthermore, SNV propagation in Vero cultures seems to cause mutations in the RdRp which potentially attenuates the virus and makes it less virulent [135]. Consequently, the significant absence of any strong disease models outside of macaques and Syrian hamsters poses a challenge for weaponization, as the inability to replicate a similar human disease progression in primates from passaged and isolated virions will hinder the lethality of any engneered bioweapon.

Although tough and resource demanding, concentrating hantaviruses is not impossible and may become more efficient with newer technologies and techniques as Warner *et al*. demonstrated [74]. A way to increase viral stocks is to avoid using the standard intramuscular model of infection and instead use the intraperitoneal infection of deer mice which was demonstrated to produce SNV stocks with high viral RNA copy number [74]. New immunotherapies methods leading towards lethal disease models are also helping to increase the viral load as the infection of immunocompetent Syrian hamsters with cell-cultured SNV resulted in lower levels of viral dissemination compared to immunocompromised hamsters [140, 141]. Improvements in viral isolation for biological characterization studies has been conducted with HTNV and PUUV in suckling mice and Syrian hamsters respectively because of their sensitivity to infection [94, 142]. There is still a reliance on Vero cell lines for viral propagation which has its own challenges. Vero E6 cells have been shown to produce an IFN-λ response to Hantaviral infection consequently reducing viral yields and affecting their quality [96]. The challenge of viral isolation, culturing, and modification in recent years has become relatively easy as indicated by the isolation and sequencing of the SARS-CoV-2 virions which demonstrates that synthetic biology methods are available for facilitating virion production which could include reverse engineering [143]. Hantavirus components and virions as well as pseudovirions are already produced by passaging in Vero E6 cell lines, with RT-PCR methods and Vesicular Stomatitis Virus vectors being employed for sequencing and for the detection of hantavirus infection though the presence of their neutralizing antibodies [132, 144, 145].

Additionally, given the rise in genetic engineering tools and techniques such as TALEN and CRISPR/Cas9, the ability to synthetically engineer more pathogenic bioagents is available [146]. CRISPR/Cas9 was employed to reduce HIV viral replication in infected T-cells, and could potentially be employed to increase virulence and viral replication for other pathogens including HCPS-causing hantaviruses [147]. Incidentally, the limitation of culturing hantaviruses virions now may be improved overtime with developments in gene editing and Do-It-Yourself

technologies which have made sophisticated techniques more accessible to conventional militaries and terrorist organizations [148].

2.5. Susceptible Targets

2.5.1. Conventional Warfare Settings

Militaries and hantaviruses have a deep history which is largely tied with the operation of war and the requirements that are needed to support it [149]. One important factor of militaries is their strength component, comprised of large bodies of soldiers congregating in theatres of operations for extended periods of time. This has the unfortunate consequence of consolidating resources, especially food that has the tendency to attract animals such as rodents and insects, as well as disrupting natural habitats that affect ecosystems and the reservoirs that inhabit them [150]. Warfare also extends disruptions to manmade infrastructure which generally creates barriers to illnesses, including housing and sanitation and the access to medical care facilities that could prevent the spread of diseases. A variety of these factors could be taken advantage of in warfare, whether it be a passive allowance of weakening military strength in the face of soldiers' worsening living conditions, or the intentional spread of a pathogenic biological agent by natural vectors or artificial delivery systems. Like other major wartime diseases like Influenza and Typhoid Fever, hantaviruses have been identified in a variety of different conflicts.

Puumala virus causes a milder form of HFRS called Nephropathia Epidemica (NE) and is spread by bank vole in Europe [83]. It is suspected that hantaviruses, specifically PUUV, spread across Europe during WWI in the form of Trench Nephritis which can be attributed to the congestion of soldiers and rodents in tight places, including trench lines that destroyed farmland and undermined infrastructure [4, 149, 151]. Trench diseases, including Trench Foot, Trench fever, and Trench Nephritis, constituted 25% of the British Expeditionary Force's triage bed occupancy, and when the US entered the war in 1917, 0.54% of their 370,000 military personnel were affected by NE [151]. HFRS and hantaviruses came to the attention of western medicine during the Korean War (1950-1953) which observed 3,200 United Nations troops becoming infected, with HTNV being isolated and identified as the etiological agent in 1978 [152, 153]. Similarly, in the early 1930s, Soviet troops encountered a similar disease along the Amur River that caused nephritis, bleeding, and shock while Imperial Japanese forces suffered 12,000 cases as they invaded Manchuria during the Second Sino-Japanese War (1937-1945) [4, 151]. Aside from the 1993 HCPS outbreak, major hantavirus outbreaks such as HTNV and PUUV are associated with war.

Consequently, the military is a natural target for hantavirus as a result of their activities occurring in largescale field exercises or in land-based combat which can disrupt natural habitats and cause exposure to hantaviruses by dispersing HCPS-causing rodents as was the case for HFRS-causing *Apodemus agrarius* or the striped field mouse [154]. HCPS cases continue to be reported following military personnel encounters with the rodent reservoir, especially in large-scale military exercises that overlap with the rodent reservoirs' habitats [155]. The congregation of soldiers in poorly ventilated or rarely maintained defences such as trench lines or housing complexes are at risk to the infestation of HCPS-causing rodents regardless of SNV or ANDV's weaponization. As defences and facilities decay overtime due to resource scarcity and war attrition, the ability to maintain sanitation and regular hygiene will be compromised enabling a return of Trench Nephritis and HCPS pulmonary disease. This can be further exacerbated by the influx and settling of refugees in consolidated camps which lack proper infrastructure and sanitation to prevent the spread of diseases let alone maintain barriers to hantavirus reservoir spread. This is evident by the influx of refugees generated from the conflict during the Yugoslav Wars (1991-2001) where civil unrest and internecine conflict caused massive movements of people and resulted in military and civilian exposure to hantaviruses, including the novel, HFRS-causing hantavirus DOBV which inflicted a 20% mortality rate [156, 157]. Incidentally, as infrastructure decays or is undermined by war, more people will be exposed to debilitating hantaviruses as contact with rodent reservoirs increases.

Within the range of the military, it would be appropriate to develop strategies to delay or inundate military forces by exposure to hantavirus through natural infection models. This would observe HCPS-causing hantaviruses to be deployed as area denial weapons which are employed to slow the advance or endanger target militaries. Area denial weapons tend to restrict the momentum of target forces, usually forcing them into positions of vulnerability which may include adopting additional precautions to manage and mitigate the effects of the bioweapon itself. Employing New-World Hantaviruses in this respect, whether it be the physical dispersal of HCPS-infected rodents to undermine the entrenched living conditions of soldiers, or the deployment of aerosolized virion particles would significantly affect the morale, strength, and movement of the target army. A strategy of area denial would be to harbour HCPS-causing rodents in built-up areas to prevent the appropriation of urban infrastructure by an invading force. Abandoned facilities would be especially exceptional since SNV is found highly aerosolized in small, $\leq 1\mu m$ particulate matter that is far-more easily disturbed to the breathing zone (1.5m height) from walking rather than sweeping [158]. Soldiers seizing urban areas would be most vulnerable, especially during the spring and summer months where reservoir

breeding and particulate aerosolization increases [159]. SNV may also persist in excreta for longer since sunlight and UV radiation are blocked from actively degrading the virions due to the protection vacant buildings provide [160].

2.5.2. Unconventional Warfare Settings and Civilian Targets

Targets to a country's civilian populace or economic and industrial sectors are important alternatives for bioterrorist organizations or low-parity nations that cannot compete with modern industrial militaries. As illustrated by the 2019 SARS-CoV-2 pandemic, infectious diseases have the effect of compromising the entire socio-economic systems of countries and will be a practical target for most bioterrorist organizations. A bioweapons attack will likely force the civilian populace to seek shelter or undergo rigorous quarantine measures which will affect the consumption of products from primary and secondary sectors of industry. The 2019 SARS-CoV-2 pandemic's quarantining especially reduced demand for oil and petroleum products, manufacturing, and agriculture worldwide as isolated civilians were no longer able to consume and grow the economy at previous rates resulting in a decline in overall national GDP [161]. Furthermore, impediments to the social fabric caused by the pandemic resulted in an overall abated pattern of life that observed closures of schools, increased hospitalization and pressures on the medical systems, as well as an increase in government debt and expenditure to maintain the stability of their financial sectors [161]. HCPS-causing Hantaviral bioweapons could be deployed in this way to afflict damage to a nation's industrial output or to invoke panic amongst a civilian populace which would affect a country's ability to fight conflicts abroad or domestically.

Farmers are naturally affected by the presence of hantaviruses due to their outdoor activities and cultivation of farmland which overlap rodent reservoir habits [162]. A possible biothreat scenario involving hantavirus would likely target agriculture centres by increasing the incidence of contact with HCPS-causing rodents preventing farmers from working or by forcing them to require additional and costly protective equipment which would create delays in production. This includes traditional farmers utilizing lumber as a fuel source as firewood handling could result in the close contamination from Hantaviral infected aerosols or dust [163]. The risk from storage or lumber shelters will especially affect those in the lumber and forestry industries and can thus be a target for a slow delivery in addition to an either targeted or widescale dispersal of hantavirus bioweapon which would delay or harass industrial production.

Hantaviruses, because of their global nature, have the capacity to affect infrastructure especially shipping and trade. Given a major outbreak, major ports contaminated with aerosolized hantaviruses have the capacity to create

delays in trade which will endanger the economy of a target country. HFRS-causing SEOV is found worldwide because of its ubiquitous rodent reservoir the brown rate (*Rattus norvegicus*) whose close relationship with humans and subsequent dissemination through global trade, human migrations, and settlement has enabled its transit and viability [164]. The presence of the brown rat in most major urban centres and in key transport industries such as maritime and land shipping can create a vulnerability to trade if targeted by bioterrorist organizations. Having a simple rodent infestation can threaten food stores and given hantaviruses general durability in moderate temperatures and low-UV light environments such as storage containers will allow aerosolized hantaviruses to survive up to two-weeks and create hazards for government or civilian responders. Selective pressures and challenging environments like highlands, deserts, and cities will likely prevent dispersal of natural reservoirs of hantavirus such as deer mice. However, in North America, similar species to deer mice such as *Peromyscus leucopus* or the white-footed mouse have occupied effective niches in cities on the east coast of Canada and the United States and have taken advantage of urban environment's lack of predators and natural competitors, its warmer climate for mating, and its abundance of small forest fragments for habitation [165]. Attempting to build a natural reservoir in the city would take generations of rodent colonies and would itself be unviable given time, resources, and current rodent controls and proofing.

Hantaviruses also pose a risk to food consumption as well if improperly stored [166, 167]. Hantaviruses, with PUUV and ANDV being studied, are not easily digested by stomach acids, and can survive long enough to be passed into the gastrointestinal tract. Despite the requirements for intragastric route infection being the least effective, the oral route of infection is plausible for PUUV [168]. Contaminating food and water supplies with biological weapons generally produces fewer casualties compared to an airborne release, but may be a secondary consequence resulting from a primary release [169]. This would be the effect of having a warehouse contaminated with Hantaviral aerosols which will contaminate food stocks contained in tin cans or boxed containers. The consequence is two-fold. The first involves the vast stores of merchandise and material needing to undergo rigorous decontamination or disposal to prevent subsequent human contact and illness which will affect economic output. The second, if successful, will cause an indirect aerosol route of transmission to humans which could be widespread due to the nature of modern supply-chains and distribution.

Hantaviruses are only currently pathogenic to humans as their rodent reservoirs remain chronically infected and asymptomatic of the disease, however, they continue to be highly viraemic as the natural host produces antibodies including neutralizing antibodies [53, 170]. HFRS-causing hantaviruses are found to infect a wide array of rodents

and insectivore species including bats, as well as hantavirus antibodies being found in domestic animals such as cats, dogs, rabbits, and pigs [171]. The infection of domestic animals and livestock such as cows is a concern because it produces another route of transmission between animals to humans. PUUV was demonstrated to experimentally infect bovine aortic endothelial cells, however, it is unknown whether asymptomatic persistent infections exist in domestic animals [172]. Consequently, the effect on animals is fairly low as HCPS-causing hantaviruses do not cause disease in animals which remain largely asymptomatic [173]. This includes monkeys, with the only nonhuman primate exception being SNV-infected rhesus macaques and ANDV in Syrian hamsters which experienced severe HCPS-disease because of Vero E6 propagated virions [135, 174]. Incidentally, targets for livestock and agriculture within bioterrorism attacks are very unlikely when using hantaviruses.

One of the limitations of hantaviruses is that they spread by specific rodent host species with most being spread by one or a few closely related rodents which reflect the co-evolutionary relationships hantaviruses generally have with their reservoirs [175]. Hantavirus evolution and reassortment is limited to intraspecies reassortment and inter-lineage events within the same, single rodent reservoir [176]. Like Influenza, Hantaviruses are segmented and can undergo reassortment events with the exchange of gene segments between viruses that infect the same cells. The formation of antigenic shifts through reassortment events act as new ways for segmented viruses to adapt to new animal hosts and to increase infectivity. This can result in the formation of novel progeny viruses that are genetically distinct from the parental viruses and could be employed as the method of developing more pathogenic hantaviruses by bioterrorist organizations, especially with different HCPS-causing reservoir hosts co-located in close proximity [176].

ANDV and SNV are genetically distinct hantaviruses that circulate in different regions and different rodent reservoirs. Despite ANDV not being maintained in deer mice, it can infect the SNV rodent reservoir allowing for new serotypes to occur [53]. Additionally, ANDV and SNV reassortment events produced diploid and monoploid viruses with SNV S and L Segments and ANDV M Segments, which efficiently replicated in Vero E6 cells [177]. Infectivity of these new viruses takes on the characteristics of the ANDV M Segment they have adopted and is suggested that reassortments of M Segment substitutions promote virus survival by increasing its infectivity [177]. Previous *in vitro* studies have observed the mixing of distinct strains of SNV in Vero E6 cells generating new reassorted viruses [178]. Pathogenic SNV NMR11 strains were also able to reassort with non-pathogenic Black Creek Canal Virus (BCCV), a distantly related New-World hantavirus that infects a different rodent species, the cotton rat (*Sigmodon hispidus*)

[178]. Despite the low frequency of reassortment and the lack of predominance of any specific segment over the other, the ability for the strains to reassort highlights the importance reassortment as a genetic mechanism is in the emergence of new and possibly more lethal hantaviruses [178].

Naturally occurring SNV reassortments are rather limited though and generally occur within local deer mouse populations just because of the local ecology supporting so few rodent species who rarely encounter each other and allow natural reassortment to occur. This is also undermined by the increasing genetic distance between rodent species that make reassortments less frequent [179]. However, bioterrorist organizations can artificially force these interactions by ensuring infected rodent species are grouped together with similar species like bats, voles, or shrews to enable the reassortment of pathogenic hantaviruses to form. This is similar to the case of a lethal genotype of ANDV, *Araraquara orthohantavirus* (ARQV) being documented in neotropical bats in Brazil which exposes the possibility of creating recombinant viruses with more infectious and morbid segmented negative-sense RNA pathogens such as Ebola Virus or Influenza-type viruses [180]. The threat of reassortment enables hantaviruses to develop new opportunities to host-switch. This is especially important since mixing M Segments and their expressed glycoproteins enable the virus to interact with cell membrane proteins for entry, creating new routes of entry and new cell targets from old viruses [176]. Considering that reassortments can be done *in vitro*, the opportunity for bioengineering by random reassortments of pathogenic and non-pathogenic hantaviruses is possible.

2.5.3. Transmission

Compared to other hantaviruses like PUUV which readily transmits between bank voles and persists effectively within the environment, SNV retains some limitations to its transmission both horizontally between rodents and vertically to humans [72]. SNV horizontal transmission between its rodent reservoir has been observed to occur through biting and scratching, frequently among males with indirect transmission being possible among laboratory-inoculated rodents [181]. Although transmission from contaminated excreta is possible, freshly infected deer mice were more likely to shed the virus and transmit it horizontally at the 14-day post-infection stage where SNV replication appeared the highest [74]. These recent studies of horizontal infection between deer mice discovered that SNV-infection only occurred in 24% of the uninfected deer mice caged with a same-sex SNV-infected mate for 6 weeks. Additionally, subsequent experiments accounting for long-term shedding noticed no further uninfected deer mice contracting SNV from uncleaned cages alone [74]. In contrast, ANDV transmitted more efficiently between uninfected

cage mates whilst maintaining higher persistence in the environment [128]. Additionally, reproductively active males with wounds comprised the majority of ANDV seropositive *Oligoryzomys longicaudatus* rodent members with horizontal transmission being primarily through male intersexual competition [182]. This offers a way to increase transmission horizontally to amplify hantavirus presence in the environment but is itself a difficult and resource consuming method. Ultimately, the positive pressure of SNV infection horizontally is limited to direct and aggressive interactions within the reservoir which affects how quickly a reservoir can be infected and dispersed against a military target. This also affects the cultivation of Hantaviral virions for concentration as the viral replication is impeded by the slow infection rate between deer mice.

While all hantaviruses are spread to humans via the inhalation of contaminated dust and aerosols dispersed from rodent faeces, urine, saliva, and fur, the viability of its spread is limited to peridomestic risk areas such as barns, cabins, or warehouses [160, 183]. The extent of SNV infection is thus restricted to the presence of deer mice as the main delivery system until the foundation of more effective passaging and isolation techniques arise to make artificial airborne dispersal techniques more effective.

Unlike SNV and other hantaviruses, ANDV has a distinguishable route of transmission because of its ability to spread person-to-person exemplified by several small cluster outbreaks in Southern Chile and Argentina [184, 185]. For person-to-person transmission to occur, close contact is required which increases the risk to people living in the same household as well as sexual partners. The presence of ANDV in the alveolar epithelium and salivary glands of *Sigmondontine* rodents reinforces intraspecies transmission from saliva and biting [128]. ANDV infected patients have shown the virus to be present in pneumocytes and pulmonary macrophages, with ultrastructural and immunocytochemical studies revealing viral replication occurring in the alveolar epithelial cells with virus-like particles being released into the alveolar lumen [186]. ANDV is likely secreted into human saliva and transmitted through close, intimate encounters or by exposure to respiratory droplets released through coughing or sneezing. There has also been reports of person-to-person transmission of ANDV from breast milk to new-borns, compounded by the new-borns' inadequate immune system and the presence of vRNA in the breast milk [187]. Nevertheless, ANDV person-to-person transmission appears to be limited to close contacts and not nearly similar to the transmission rate and basic reproduction numbers (R_O) of SARS-CoV-2 or Category A Pathogens such as Ebola Virus which have been assessed to be greater than 1 [188, 189]. ANDV's R_O number has been estimated to be significantly less than 1 and would likely not initiate a pandemic within the parameters of the current data [190]. Incidentally, the risk, albeit

present, is rather limited because the efficacy of ANDV being rapidly disseminated throughout a target group is dependent largely on aerosol inhalation or contact with contaminated saliva with the latter being an unpractical method to strike at large target populations.

2.5.4. Dispersal and Delivery

Since hantaviruses are transmitted to humans from rodents, a rudimentary but deliberate release of infected rodents into a target location would be a relatively easy way to threaten public health [191]. The impact would be low, but a strike against a country's key infrastructure like trade ports, warehouses, hospitals, governments centres, or public gatherings with infected rodents would create delays to productivity and the economy. Modern, industrialized countries in the west would not undergo famine or experience food insecurity as a result of a biological attack to the agriculture sector because of its robustness in diversity, high-production and heightened regulation [192]. However, disruptions caused by the presence of suspected or confirmed biological agents has the potential to inflict market speculation and contraction through bans on international food exports resulting in lost revenue, job losses, and the destruction of capital including livestock or contaminated merchandise. This is indicated by the pig and cow culling during the Foot and Mouth Disease and the Bovine Spongiform Encephalopathy outbreaks in the UK, US, and the Republic of China in 1997 and the early 2000s [193]. The intent would be to cripple infrastructure and overburden the economy and medical apparatus. If introduced into a target rural or urban area, infected rodents have the potential to cause long-term medical incidents and create public panic that will have the effect of consuming municipal or federal resources required to manage the attack [191]. Additionally, retaining and cultivating deer mice for Hantaviral preparation is cost effective and commercially available. Deer mice are easily maintained following standard laboratory mouse protocols with deer mice being no different from lab mice in terms of handling with the exceptions of their aggressive tendencies involving biting and their agility resulting in escape which prompts increased biosafety measures to be taken [194, 195]. Deer mice also suffer from not being genetically homogeneous resulting in inconsistency in experimentation due to widespread and significant genetic polymorphisms.

Rodent dispersal is discreet and innocuous and can go unnoticed compared to bioweapons deployed by artillery, missiles, or by aerial deployment by aerosols. One limitation of Hantaviral deployment by artillery or missile is due to its 60°C heat sensitivity as any incendiary or kinetic deployment system would inactivate and degrade Hantaviral virions [169]. The preferable deployment mechanism would be an aerosolization or powdered pathway

which is undetectable and can achieve rapid dispersal over a wide area [169]. Hantavirus delivery could benefit from similar dispersal methods employed to transport Anthrax or Ricin toxins such as letters and mailed packages due to their persistence in UV-free environments [28]. This becomes a problem since hantaviruses are undetectable in these delivery systems which urges for the development of new diagnostic and detection equipment. Hantavirus infections in humans are diagnosed with tedious enzyme-linked immunosorbent assay (ELISA), or IgM-capture tests to detect IgM antibodies and also RT-PCR detection of viral RNA in rodent or insectivore hosts [196]. However, the limitation of hantavirus isolation would prevent it from being assembled into an effective aerosol which would require high Hantaviral concentrations. This would require substantial laboratory resources and technical expertise to maintain sufficient viral stocks for weaponization and would prove to be the leading difficulty for uninitiated bioterrorist organizations in accomplishing. Additionally, aerosols dispersed outside during the day have the risk of being degraded by the viricidal properties of UV radiation which poses another limitation to outdoor dispersal [197]. The optimal route would be to have dispersal mechanisms deploy indoors to prevent the seizure of facilities by militaries or to create disruptions for civilian personnel employed in key industrial sectors.

There does exist substantial methods of rodent and pest controls that target deer mice through bait and trappings, structural proofing and rodenticides which have proven both economical and effective in preventing rodent entry to structures including underdeveloped residences [198-200]. By culling or isolating rodents through said techniques, rodent controls help to minimize human-rodent contact and ultimately transmission. This is exemplified by other rodent-borne *Bunyaviruses* such as the Arenavirus Lassa Virus which experienced a reduction in seroprevalence proportional to reductions in its North-West African reservoirs (*Mastomys natalensis*, *Mastomys erythroleucus*, *Hylomyscus pamfi*) through the use of rodenticides and urban proofing that targeted rodent and human food stocks and housing [201]. However, complete seroprevalence reduction of Lassa Virus relied upon an 80% reduction in rodent population densities indoors and in peridomestic environments to avoid lateral viral transmission which becomes labour and resource intensive and may not be feasible in developing countries or those affected by war [202]. Nevertheless, deployment of rodent hosts as physical carriers of a hantavirus bioweapon would be seriously hampered by a proactive application of bait poisons, fumigant poisons, or non-poisonous measures including traps. However, handling of caught rodents through traps including diseased rodents is both labour intensive and increases the risks associated with hantavirus exposure [203].

ANDV virus would be the preferred model for dispersal if conducting a specific one-target attack with collateral to personnel within an immediate vicinity. This is because ANDV can occur within household person-to-person contact and can cause up to 25-35% mortality rates [173]. An attack on a single target with the intent of causing panic and successive but limited infections within a household would be an unideal although possible diversionary method for assassination. Household contacts of ANDV are at risk of developing HCPS infections within 4 weeks with ANDV vRNA being routinely detected in blood cells for up to two weeks before symptoms or anti-hantavirus antibodies arise [204]. This enables a person-to-person model to be employed for targets that require discretion since infection can take effect weeks after the attempt has been made compared to overt assassination or sabotage attempts which risk immediate suspicion and association. This also enables the virus to be spread asymptomatically within an infected group, although an influenza-like pandemic seems unfeasible due to the strict requirements of ANDV infection relying on person-to-person contact being very close [173].

2.6. Conclusions

With the limitations present, HCPS-causing hantaviruses are generally restricted to the Category C definition largely because their spread and ability to be concentrated in the laboratory faces difficult barriers. The feasibility of developing HCPS-causing bioweapons comes from the few strengths hantaviruses possess which includes its high mortality rate. Hantaviruses can also be easily dispersed through aerosols if limited to indoor facilities or warehouses with no insolation and can effectively target personnel – especially the military – operating in close proximity to rodent reservoir habitats. HCPS-causing hantaviruses also benefit from being difficult to treat since no Old-World Hantavirus antivirals or vaccines have effective specificity against them. However, the morbidity rate of New-World Hantaviruses is very low, with ANDV being a potential but somewhat viable agent because of its person-to-person transmission pathway which would likely infect more people. Additionally, the ability to manufacture and produce ANDV or SNV into a lethal form that can be dispersed poses a problem to its weaponization due to the presence of attenuating mutations and absence of a strong disease model for nonhuman primates. Although widespread, the ability for HCPS-causing hantavirus reservoirs to adopt new urban environments is limited but may be improved by the effects of climate change and the increase in human industrial activity. As technologies improve and barriers to passaging and replicating hantavirus virions in culture become easier and more viable, the ability to mass-produce pathogenic HCPS-causing hantaviruses like SNV or ANDV may upgrade hantaviruses from a Category C to A definition.

Nevertheless, despite being a Category C pathogen, the threat of hantaviruses infections generally and the potential for it to be weaponized should not be ignored. Hantaviruses are emerging pathogens that require the attention of government and medical health research as globally they still occur frequently in developing countries with poor infrastructure or in rural, agrarian environments that have close contact with Hantaviral rodent-reservoirs. Hantavirus and HCPS continue to be a serious pathogen and disease to be considered carefully due to the environmental-associated risks of frequent rodent-human contact that expose military personnel, farmers and agriculture workers, and warehouse and shipping staff to the virus. The paucity of reporting in developing countries and the neglect that hantaviruses face allows it to slip under the radar and can be exploited by organizations that could potentially field extensive laboratory equipment and rodent reservoirs towards the development of hantavirus-based biological weapons. Infectious diseases generally can be mitigated with better reporting and surveillance, especially by monitoring the incidence of disease through extensive international health and medical networks. This can be accomplished by governments and academic agencies resolving to be proactive in testing, freely sharing clinical and experimental details, and maintaining intergovernmental transparency regarding pandemics or the occurrence of bioweapon threats [205]. Naturally, with a stronger observation and tracking of infectious diseases the easier it is to identify and manage them when they occur.

Globally, the ability to employ biological agents is prohibited by the Biological and Toxin Weapon Convention (BTWC) ratified by the United Nations and 170 of their member states which has limited international biological warfare. However, because of their lack of inspection mechanisms, rogue states and terrorist organizations could circumvent the BTWC treaty and employ biological weapons against target nations [206]. The intent may not be to singly destroy a nation or completely kill its people, with terrorist objectives being more nuanced and complex such as the case with Al-Qaeda attempting to destabilize and disrupt US power in the Middle-East [207]. Instead, the importance of disruption is key since any bioterror attack regardless of its category could inflict damage to a nation's populace, economy, and prestige which have deeper ramifications to global security. Consequently, further research into weaponization and surveillance are essential to prevent or mitigate the effects of bioweapons.

As a consideration, significant international and national cooperation must occur to safeguard global trade, public health, and international security from bioterrorism. Mitigation strategies against bioagent attacks can only be effective given the invested interests of governments and research scientists in protecting the health of their peoples. Incidentally, research into medical health science must be focused on building towards detection, identification,

mitigation and management equipment and techniques with the concentration of resources from cooperating governments to fund developments in counterterrorism and medical therapeutics. This would require an intergovernmental exchange of communication between research scientists, policy-makers, and the public to broaden transparency towards international security and scientific research **[208]**. Through cooperation, predictions of future attacks or employment of bioagents can be ascertained, preventing socio-economic collapses that could occur from industry-paralyzing infectious diseases.

2.7 References

1.   Mittler, E., et al., *Hantavirus entry: Perspectives and recent advances.* Adv Virus Res, 2019. **104**: p. 185-224.
2.   Iannetta, M., et al., *Viral Hemorrhagic Fevers Other than Ebola and Lassa.* Infectious Disease Clinics of North America, 2019. **33**(4): p. 977-1002.
3.   Liu, R., et al., *Vaccines and Therapeutics Against Hantaviruses.* Front Microbiol, 2019. **10**: p. 2989.
4.   Schmaljohn, C., *Vaccines for hantaviruses.* Vaccine, 2009. **27 Suppl 4**: p. D61-4.
5.   Duehr, J., et al., *Neutralizing Monoclonal Antibodies against the Gn and the Gc of the Andes Virus Glycoprotein Spike Complex Protect from Virus Challenge in a Preclinical Hamster Model.* mBio, 2020. **11**(2).
6.   Watson, D.C., et al., *Epidemiology of Hantavirus infections in humans: a comprehensive, global overview.* Crit Rev Microbiol, 2014. **40**(3): p. 261-72.
7.   Jonsson, C.B., L.T. Figueiredo, and O. Vapalahti, *A global perspective on hantavirus ecology, epidemiology, and disease.* Clin Microbiol Rev, 2010. **23**(2): p. 412-41.
8.   Llah, S.T., et al., *Hantavirus induced cardiopulmonary syndrome: A public health concern.* J Med Virol, 2018. **90**(6): p. 1003-1009.
9.   Kruger, D.H., G. Schonrich, and B. Klempa, *Human pathogenic hantaviruses and prevention of infection.* Hum Vaccin, 2011. **7**(6): p. 685-93.
10.   Meyer, R.F. and S.A. Morse, *Viruses and Bioterrorism*, in *Encyclopedia of Virology (Third Edition)*, B.W.J. Mahy and M.H.V. Van Regenmortel, Editors. 2008, Academic Press: Oxford. p. 406-411.
11.   Williams, M. and D.C. Sizemore, *Biologic, Chemical, and Radiation Terrorism Review*, in *StatPearls*. 2020, StatPearls Publishing
StatPearls Publishing LLC.: Treasure Island (FL).
12.   Christian, M.D., *Biowarfare and Bioterrorism.* Critical Care Clinics, 2013. **29**(3): p. 717-756.
13.   Doganay, M. and H. Demiraslan, *Human anthrax as a re-emerging disease.* Recent Pat Antiinfect Drug Discov, 2015. **10**(1): p. 10-29.
14.   Miroslav, P., *Bacillus anthracis as a biological warfare agent: infection, diagnosis and countermeasures.* Bratisl Lek Listy, 2020. **121**(3): p. 175-181.
15.   Vigueras-Galvan, A.L., et al., *Current Situation and Perspectives on Hantaviruses in Mexico.* Viruses, 2019. **11**(7).
16.   Drebot, M.A., et al., *Hantavirus pulmonary syndrome in Canada: An overview of clinical features, diagnostics, epidemiology and prevention.* Can Commun Dis Rep, 2015. **41**(6): p. 124-131.
17.   Knust, B. and P.E. Rollin, *Twenty-year summary of surveillance for human hantavirus infections, United States.* Emerg Infect Dis, 2013. **19**(12): p. 1934-7.
18.   Montoya-Ruiz, C., F.J. Diaz, and J.D. Rodas, *Recent evidence of hantavirus circulation in the American tropic.* Viruses, 2014. **6**(3): p. 1274-93.
19.   Rovida, F., et al., *Imported hantavirus cardiopulmonary syndrome in an Italian traveller returning from Cuba.* New Microbiol, 2013. **36**(1): p. 103-5.

20.	Alonso, D.O., et al., *Epidemiological description, case-fatality rate, and trends of Hantavirus Pulmonary Syndrome: 9 years of surveillance in Argentina.* J Med Virol, 2019. **91**(7): p. 1173-1181.

21.	Escalera-Antezana, J.P., et al., *Orthohantavirus pulmonary syndrome in Santa Cruz and Tarija, Bolivia, 2018.* Int J Infect Dis, 2020. **90**: p. 145-150.

22.	Figueiredo, L.T., et al., *Hantaviruses and cardiopulmonary syndrome in South America.* Virus Res, 2014. **187**: p. 43-54.

23.	Fonseca, L.X., S.V. Oliveira, and E.C. Duarte, *Magnitude and distribution of deaths due to hantavirus in Brazil, 2007-2015.* Epidemiol Serv Saude, 2018. **27**(2): p. e2017221.

24.	Goeijenbier, M., et al., *Emerging Viruses in the Republic of Suriname: Retrospective and Prospective Study into Chikungunya Circulation and Suspicion of Human Hantavirus Infections, 2008-2012 and 2014.* Vector Borne Zoonotic Dis, 2015. **15**(10): p. 611-8.

25.	Jiang, H., et al., *Hantavirus infection: a global zoonotic challenge.* Virol Sin, 2017. **32**(1): p. 32-43.

26.	Matheus, S., et al., *Hantavirus Pulmonary Syndrome Caused by Maripa Virus in French Guiana, 2008-2016.* Emerg Infect Dis, 2017. **23**(10): p. 1722-1725.

27.	Riquelme, R., et al., *Hantavirus pulmonary syndrome, Southern Chile, 1995-2012.* Emerg Infect Dis, 2015. **21**(4): p. 562-8.

28.	Bhalla, D.K. and D.B. Warheit, *Biological agents with potential for misuse: a historical perspective and defensive measures.* Toxicol Appl Pharmacol, 2004. **199**(1): p. 71-84.

29.	Hjelle, B., et al., *Emergence of hantaviral disease in the southwestern United States.* West J Med, 1994. **161**(5): p. 467-73.

30.	de St Maurice, A., et al., *Exposure Characteristics of Hantavirus Pulmonary Syndrome Patients, United States, 1993-2015.* Emerg Infect Dis, 2017. **23**(5): p. 733-739.

31.	CDC. *Bioterrorism Agents/Diseases.* 2018 April 4, 2018; Available from: https://emergency.cdc.gov/agent/agentlist-category.asp.

32.	NIAID. *US NIAID Emerging Infectious Diseases/ Pathogens.* 2018 July 26, 2018; Available from: https://www.niaid.nih.gov/research/emerging-infectious-diseases-pathogens.

33.	Mir, M.A., *Hantaviruses.* Clin Lab Med, 2010. **30**(1): p. 67-91.

34.	Clement, J.P., *Hantavirus.* Antiviral Res, 2003. **57**(1-2): p. 121-7.

35.	Forbes, K.M., T. Sironen, and A. Plyusnin, *Hantavirus maintenance and transmission in reservoir host populations.* Curr Opin Virol, 2018. **28**: p. 1-6.

36.	Avsic-Zupanc, T., A. Saksida, and M. Korva, *Hantavirus infections.* Clin Microbiol Infect, 2019. **21s**: p. e6-e16.

37.	Spiropoulou, C.F., et al., *Genome structure and variability of a virus causing hantavirus pulmonary syndrome.* Virology, 1994. **200**(2): p. 715-23.

38.	Nichol, S.T., et al., *Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness.* Science, 1993. **262**(5135): p. 914-7.

39.	Plyusnin, A., O. Vapalahti, and A. Vaheri, *Hantaviruses: genome structure, expression and evolution.* J Gen Virol, 1996. **77 (Pt 11)**: p. 2677-87.

40.	Chizhikov, V.E., et al., *Complete genetic characterization and analysis of isolation of Sin Nombre virus.* J Virol, 1995. **69**(12): p. 8132-6.

41.	Kamrud, K.I. and C.S. Schmaljohn, *Expression strategy of the M genome segment of Hantaan virus.* Virus Res, 1994. **31**(1): p. 109-21.

42.	Amroun, A., et al., *Bunyaviridae RdRps: structure, motifs, and RNA synthesis machinery.* Crit Rev Microbiol, 2017. **43**(6): p. 753-778.

43.	Hepojoki, J., et al., *Hantavirus structure--molecular interactions behind the scene.* J Gen Virol, 2012. **93**(Pt 8): p. 1631-44.

44.	Hussein, I.T., et al., *Recent advances in hantavirus molecular biology and disease.* Adv Appl Microbiol, 2011. **74**: p. 35-75.

45.	Goldsmith, C.S., et al., *Ultrastructural characteristics of Sin Nombre virus, causative agent of hantavirus pulmonary syndrome.* Arch Virol, 1995. **140**(12): p. 2107-22.

46.	Mir, M.A. and A.T. Panganiban, *Trimeric hantavirus nucleocapsid protein binds specifically to the viral RNA panhandle.* J Virol, 2004. **78**(15): p. 8281-8.

47.	Kukkonen, S.K., A. Vaheri, and A. Plyusnin, *L protein, the RNA-dependent RNA polymerase of hantaviruses.* Arch Virol, 2005. **150**(3): p. 533-56.

48.   Cheng, E., Z. Wang, and M.A. Mir, *Interaction between hantavirus nucleocapsid protein (N) and RNA-dependent RNA polymerase (RdRp) mutants reveals the requirement of an N-RdRp interaction for viral RNA synthesis.* J Virol, 2014. **88**(15): p. 8706-12.

49.   Mir, M.A., et al., *Hantavirus nucleocapsid protein has distinct m7G cap- and RNA-binding sites.* J Biol Chem, 2010. **285**(15): p. 11357-68.

50.   Garcin, D., et al., *The 5' ends of Hantaan virus (Bunyaviridae) RNAs suggest a prime-and-realign mechanism for the initiation of RNA synthesis.* J Virol, 1995. **69**(9): p. 5754-62.

51.   Hutchinson, K.L., C.J. Peters, and S.T. Nichol, *Sin Nombre virus mRNA synthesis.* Virology, 1996. **224**(1): p. 139-49.

52.   Gavrilovskaya, I.N., et al., *beta3 Integrins mediate the cellular entry of hantaviruses that cause respiratory failure.* Proc Natl Acad Sci U S A, 1998. **95**(12): p. 7074-9.

53.   Ermonval, M., F. Baychelier, and N. Tordo, *What Do We Know about How Hantaviruses Interact with Their Different Hosts?* Viruses, 2016. **8**(8).

54.   Connolly-Andersen, A.M., T. Thunberg, and C. Ahlm, *Endothelial activation and repair during hantavirus infection: association with disease outcome.* Open Forum Infect Dis, 2014. **1**(1): p. ofu027.

55.   Connolly-Andersen, A.M., et al., *Increased Thrombopoiesis and Platelet Activation in Hantavirus-Infected Patients.* J Infect Dis, 2015. **212**(7): p. 1061-9.

56.   Gavrilovskaya, I.N., et al., *Cellular entry of hantaviruses which cause hemorrhagic fever with renal syndrome is mediated by beta3 integrins.* J Virol, 1999. **73**(5): p. 3951-9.

57.   Mir, M.A. and A.T. Panganiban, *The hantavirus nucleocapsid protein recognizes specific features of the viral RNA panhandle and is altered in conformation upon RNA binding.* J Virol, 2005. **79**(3): p. 1824-35.

58.   Brown, B.A. and A.T. Panganiban, *Identification of a region of hantavirus nucleocapsid protein required for RNA chaperone activity.* RNA Biol, 2010. **7**(6): p. 830-7.

59.   Mir, M.A., et al., *Hantavirus N protein exhibits genus-specific recognition of the viral RNA panhandle.* J Virol, 2006. **80**(22): p. 11283-92.

60.   Woodson, S.A., S. Panja, and A. Santiago-Frangos, *Proteins That Chaperone RNA Regulation.* Microbiol Spectr, 2018. **6**(4).

61.   Reuter, M. and D.H. Krüger, *The nucleocapsid protein of hantaviruses: much more than a genome-wrapping protein.* Virus Genes, 2018. **54**(1): p. 5-16.

62.   Mir, M.A. and A.T. Panganiban, *Characterization of the RNA chaperone activity of hantavirus nucleocapsid protein.* J Virol, 2006. **80**(13): p. 6276-85.

63.   Mir, M.A. and A.T. Panganiban, *The bunyavirus nucleocapsid protein is an RNA chaperone: possible roles in viral RNA panhandle formation and genome replication.* Rna, 2006. **12**(2): p. 272-82.

64.   Boudko, S.P., R.J. Kuhn, and M.G. Rossmann, *The coiled-coil domain structure of the Sin Nombre virus nucleocapsid protein.* J Mol Biol, 2007. **366**(5): p. 1538-44.

65.   Jonsson, C.B. and C.S. Schmaljohn, *Replication of hantaviruses.* Curr Top Microbiol Immunol, 2001. **256**: p. 15-32.

66.   Muyangwa, M., et al., *Hantaviral Proteins: Structure, Functions, and Role in Hantavirus Infection.* Front Microbiol, 2015. **6**: p. 1326.

67.   Astorga, F., et al., *Distributional ecology of Andes hantavirus: a macroecological approach.* Int J Health Geogr, 2018. **17**(1): p. 22.

68.   Beltran-Ortiz, C.E., et al., *Expression and purification of the surface proteins from Andes virus.* Protein Expr Purif, 2017. **139**: p. 63-70.

69.   Schmaljohn, C. and B. Hjelle, *Hantaviruses: a global disease problem.* Emerg Infect Dis, 1997. **3**(2): p. 95-104.

70.   Guterres, A. and E.R.S. de Lemos, *Hantaviruses and a neglected environmental determinant.* One Health, 2018. **5**: p. 27-33.

71.   Godoy, P., et al., *Andes virus antigens are shed in urine of patients with acute hantavirus cardiopulmonary syndrome.* J Virol, 2009. **83**(10): p. 5046-55.

72.   Kallio, E.R., et al., *Prolonged survival of Puumala hantavirus outside the host: evidence for indirect transmission via the environment.* J Gen Virol, 2006. **87**(Pt 8): p. 2127-34.

73.   Kariwa, H., et al., *Urine-associated horizontal transmission of Seoul virus among rats.* Arch Virol, 1998. **143**(2): p. 365-74.

74.   Warner, B.M., et al., *Development and Characterization of a Sin Nombre Virus Transmission Model in Peromyscus maniculatus.* Viruses, 2019. **11**(2).

75.   Yanagihara, R., H.L. Amyx, and D.C. Gajdusek, *Experimental infection with Puumala virus, the etiologic agent of nephropathia epidemica, in bank voles (Clethrionomys glareolus).* J Virol, 1985. **55**(1): p. 34-8.

76.   Zaki, S.R., et al., *Hantavirus pulmonary syndrome. Pathogenesis of an emerging infectious disease.* Am J Pathol, 1995. **146**(3): p. 552-79.

77.   Alonso, D.O., et al., *Person-to-Person Transmission of Andes Virus in Hantavirus Pulmonary Syndrome, Argentina, 2014.* Emerg Infect Dis, 2020. **26**(4): p. 756-759.

78.   Hartline, J., et al., *Hantavirus infection in North America: a clinical review.* Am J Emerg Med, 2013. **31**(6): p. 978-82.

79.   Martinez-Valdebenito, C., et al., *Person-to-person household and nosocomial transmission of andes hantavirus, Southern Chile, 2011.* Emerg Infect Dis, 2014. **20**(10): p. 1629-36.

80.   Jin, M., et al., *Hantaan virus enters cells by clathrin-dependent receptor-mediated endocytosis.* Virology, 2002. **294**(1): p. 60-9.

81.   Chiang, C.F., et al., *Endocytic Pathways Used by Andes Virus to Enter Primary Human Lung Endothelial Cells.* PLoS One, 2016. **11**(10): p. e0164768.

82.   Borges, A.A., et al., *Hantavirus cardiopulmonary syndrome: immune response and pathogenesis.* Microbes Infect, 2006. **8**(8): p. 2324-30.

83.   Vapalahti, O., et al., *Hantavirus infections in Europe.* Lancet Infect Dis, 2003. **3**(10): p. 653-61.

84.   Kawai, T. and S. Akira, *TLR signaling.* Semin Immunol, 2007. **19**(1): p. 24-32.

85.   Sen, G.C. and S.N. Sarkar, *Transcriptional signaling by double-stranded RNA: role of TLR3.* Cytokine Growth Factor Rev, 2005. **16**(1): p. 1-14.

86.   Kato, H., et al., *Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses.* Nature, 2006. **441**(7089): p. 101-5.

87.   McNab, F., et al., *Type I interferons in infectious disease.* Nat Rev Immunol, 2015. **15**(2): p. 87-103.

88.   Levine, J.R., et al., *Antagonism of type I interferon responses by new world hantaviruses.* J Virol, 2010. **84**(22): p. 11790-801.

89.   Pan, W., et al., *Effects of Different Doses of Nucleocapsid Protein from Hantaan Virus A9 Strain on Regulation of Interferon Signaling.* Viral Immunol, 2015. **28**(8): p. 448-54.

90.   Taylor, S.L., et al., *Hantaan virus nucleocapsid protein binds to importin alpha proteins and inhibits tumor necrosis factor alpha-induced activation of nuclear factor kappa B.* J Virol, 2009. **83**(3): p. 1271-9.

91.   Cimica, V., et al., *An innate immunity-regulating virulence determinant is uniquely encoded by the Andes virus nucleocapsid protein.* mBio, 2014. **5**(1).

92.   Mackow, E.R., et al., *Hantavirus interferon regulation and virulence determinants.* Virus Res, 2014. **187**: p. 65-71.

93.   Matthys, V. and E.R. Mackow, *Hantavirus regulation of type I interferon responses.* Adv Virol, 2012. **2012**: p. 524024.

94.   Seto, T., et al., *An efficient in vivo method for the isolation of Puumala virus in Syrian hamsters and the characterization of the isolates from Russia.* J Virol Methods, 2011. **173**(1): p. 17-23.

95.   Emeny, J.M. and M.J. Morgan, *Regulation of the interferon system: evidence that Vero cells have a genetic defect in interferon production.* J Gen Virol, 1979. **43**(1): p. 247-52.

96.   Prescott, J., et al., *New World hantaviruses activate IFNlambda production in type I IFN-deficient vero E6 cells.* PLoS One, 2010. **5**(6): p. e11159.

97.   Li, N., et al., *IFN-λs inhibit Hantaan virus infection through the JAK-STAT pathway and expression of Mx2 protein.* Genes Immun, 2019. **20**(3): p. 234-244.

98.   Cho, H.W. and C.R. Howard, *Antibody responses in humans to an inactivated hantavirus vaccine (Hantavax).* Vaccine, 1999. **17**(20-21): p. 2569-75.

99.   Choi, Y., et al., *Inactivated Hantaan virus vaccine derived from suspension culture of Vero cells.* Vaccine, 2003. **21**(17-18): p. 1867-73.

100.   Ying, Q., et al., *Construction and immunological characterization of CD40L or GM-CSF incorporated Hantaan virus like particle.* Oncotarget, 2016. **7**(39): p. 63488-63503.

101.   Schmaljohn, C.S., S.E. Hasty, and J.M. Dalrymple, *Preparation of candidate vaccinia-vectored vaccines for haemorrhagic fever with renal syndrome.* Vaccine, 1992. **10**(1): p. 10-3.

102. Yi, Y., H. Park, and J. Jung, *Effectiveness of inactivated hantavirus vaccine on the disease severity of hemorrhagic fever with renal syndrome.* Kidney Res Clin Pract, 2018. **37**(4): p. 366-372.

103. Song, J.Y., et al., *Long-term immunogenicity and safety of inactivated Hantaan virus vaccine (Hantavax) in healthy adults.* Vaccine, 2016. **34**(10): p. 1289-95.

104. Warner, B.M., et al., *Vesicular Stomatitis Virus-Based Vaccines Provide Cross-Protection against Andes and Sin Nombre Viruses.* Viruses, 2019. **11**(7).

105. Schmaljohn, C.S., K.W. Spik, and J.W. Hooper, *DNA vaccines for HFRS: laboratory and clinical studies.* Virus Res, 2014. **187**: p. 91-6.

106. Zhang, C., et al., *Advances in mRNA Vaccines for Infectious Diseases.* Front Immunol, 2019. **10**: p. 594.

107. Li, Y.-D., et al., *Coronavirus vaccine development: from SARS and MERS to COVID-19.* Journal of biomedical science, 2020. **27**(1): p. 104-104.

108. Kowalski, P.S., et al., *Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery.* Molecular therapy : the journal of the American Society of Gene Therapy, 2019. **27**(4): p. 710-728.

109. Engdahl, T.B. and J.E. Crowe, Jr., *Humoral Immunity to Hantavirus Infection.* mSphere, 2020. **5**(4).

110. Battisti, A.J., et al., *Structural studies of Hantaan virus.* J Virol, 2011. **85**(2): p. 835-41.

111. Gavrilovskaya, I.N., E.E. Gorbunova, and E.R. Mackow, *Pathogenic hantaviruses direct the adherence of quiescent platelets to infected endothelial cells.* J Virol, 2010. **84**(9): p. 4832-9.

112. Yoshimatsu, K. and J. Arikawa, *Antigenic properties of N protein of hantavirus.* Viruses, 2014. **6**(8): p. 3097-109.

113. Bae, J.Y., et al., *The Immune Correlates of Orthohantavirus Vaccine.* Vaccines (Basel), 2021. **9**(5).

114. Safronetz, D., et al., *In vitro and in vivo activity of ribavirin against Andes virus infection.* PLoS One, 2011. **6**(8): p. e23560.

115. Medina, R.A., et al., *Ribavirin, human convalescent plasma and anti-beta3 integrin antibody inhibit infection by Sin Nombre virus in the deer mouse model.* J Gen Virol, 2007. **88**(Pt 2): p. 493-505.

116. Chapman, L.E., et al., *Intravenous ribavirin for hantavirus pulmonary syndrome: safety and tolerance during 1 year of open-label experience. Ribavirin Study Group.* Antivir Ther, 1999. **4**(4): p. 211-9.

117. McKee, K.T., Jr., et al., *Ribavirin prophylaxis and therapy for experimental argentine hemorrhagic fever.* Antimicrob Agents Chemother, 1988. **32**(9): p. 1304-9.

118. Mertz, G.J., et al., *Placebo-controlled, double-blind trial of intravenous ribavirin for the treatment of hantavirus cardiopulmonary syndrome in North America.* Clin Infect Dis, 2004. **39**(9): p. 1307-13.

119. Brocato, R.L. and J.W. Hooper, *Progress on the Prevention and Treatment of Hantavirus Disease.* Viruses, 2019. **11**(7).

120. Safronetz, D., et al., *Antiviral efficacy of favipiravir against two prominent etiological agents of hantavirus pulmonary syndrome.* Antimicrob Agents Chemother, 2013. **57**(10): p. 4673-80.

121. Vaheri, A., et al., *Hantavirus infections in Europe and their impact on public health.* Rev Med Virol, 2013. **23**(1): p. 35-49.

122. Hardestam, J., et al., *Ex vivo stability of the rodent-borne Hantaan virus in comparison to that of arthropod-borne members of the Bunyaviridae family.* Appl Environ Microbiol, 2007. **73**(8): p. 2547-51.

123. Tian, D. and T. Zheng, *Comparison and analysis of biological agent category lists based on biosafety and biodefense.* PLoS One, 2014. **9**(6): p. e101163.

124. Pastorino, B., X. de Lamballerie, and R. Charrel, *Biosafety and Biosecurity in European Containment Level 3 Laboratories: Focus on French Recent Progress and Essential Requirements.* Front Public Health, 2017. **5**: p. 121.

125. Jonsson, C.B., J. Hooper, and G. Mertz, *Treatment of hantavirus pulmonary syndrome.* Antiviral Research, 2008. **78**(1): p. 162-169.

126. Berl, J.L., et al., *INTERSPECIFIC COMPARISON OF HANTAVIRUS PREVALENCE IN PEROMYSCUS POPULATIONS FROM A FRAGMENTED AGRO-ECOSYSTEM IN INDIANA, USA.* J Wildl Dis, 2018. **54**(1): p. 147-150.

127. Medina, R.A., et al., *Ecology, genetic diversity, and phylogeographic structure of andes virus in humans and rodents in Chile.* J Virol, 2009. **83**(6): p. 2446-59.

128. Padula, P., et al., *Transmission study of Andes hantavirus infection in wild sigmodontine rodents.* J Virol, 2004. **78**(21): p. 11972-9.

129. Gubler, D.J., et al., *Climate variability and change in the United States: potential impacts on vector- and rodent-borne diseases.* Environ Health Perspect, 2001. **109 Suppl 2**: p. 223-33.

130.   Engelthaler, D.M., et al., *Climatic and environmental patterns associated with hantavirus pulmonary syndrome, Four Corners region, United States.* Emerg Infect Dis, 1999. **5**(1): p. 87-94.

131.   Klempa, B., *Hantaviruses and climate change.* Clin Microbiol Infect, 2009. **15**(6): p. 518-23.

132.   Elliott, L.H., et al., *Isolation of the causative agent of hantavirus pulmonary syndrome.* Am J Trop Med Hyg, 1994. **51**(1): p. 102-8.

133.   Galeno, H., et al., *First human isolate of Hantavirus (Andes virus) in the Americas.* Emerging infectious diseases, 2002. **8**(7): p. 657-661.

134.   Fulhorst, C.F., et al., *Isolation, characterization and geographic distribution of Caño Delgadito virus, a newly discovered South American hantavirus (family Bunyaviridae).* Virus Research, 1997. **51**(2): p. 159-171.

135.   Safronetz, D., et al., *Pathophysiology of hantavirus pulmonary syndrome in rhesus macaques.* Proc Natl Acad Sci U S A, 2014. **111**(19): p. 7114-9.

136.   Prescott, J., H. Feldmann, and D. Safronetz, *Amending Koch's postulates for viral disease: When "growth in pure culture" leads to a loss of virulence.* Antiviral research, 2017. **137**: p. 1-5.

137.   Eckerle, I., M. Lenk, and R.G. Ulrich, *More novel hantaviruses and diversifying reservoir hosts--time for development of reservoir-derived cell culture models?* Viruses, 2014. **6**(3): p. 951-967.

138.   Lundkvist, A., et al., *Cell culture adaptation of Puumala hantavirus changes the infectivity for its natural reservoir, Clethrionomys glareolus, and leads to accumulation of mutants with altered genomic RNA S segment.* Journal of virology, 1997. **71**(12): p. 9515-9523.

139.   Klingström, J., et al., *Wild-type Puumala hantavirus infection induces cytokines, C-reactive protein, creatinine, and nitric oxide in cynomolgus macaques.* J Virol, 2002. **76**(1): p. 444-9.

140.   Vergote, V., et al., *A lethal disease model for New World hantaviruses using immunosuppressed Syrian hamsters.* PLoS Negl Trop Dis, 2017. **11**(10): p. e0006042.

141.   Brocato, R.L., et al., *A lethal disease model for hantavirus pulmonary syndrome in immunosuppressed Syrian hamsters infected with Sin Nombre virus.* J Virol, 2014. **88**(2): p. 811-9.

142.   Li, J.L., et al., *An efficient method for isolation of Hantaan virus through serial passages in suckling mice.* Intervirology, 2013. **56**(3): p. 172-7.

143.   Thao, T.T.N., et al., *Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform.* Nature, 2020.

144.   Higa, M.M., et al., *Efficient production of Hantaan and Puumala pseudovirions for viral tropism and neutralization studies.* Virology, 2012. **423**(2): p. 134-42.

145.   Niskanen, S., et al., *Evaluation of Real-Time RT-PCR for Diagnostic Use in Detection of Puumala Virus.* Viruses, 2019. **11**(7).

146.   Benjamin, R., et al., *TALEN gene editing takes aim on HIV.* Hum Genet, 2016. **135**(9): p. 1059-70.

147.   Ophinni, Y., et al., *CRISPR/Cas9 system targeting regulatory genes of HIV-1 inhibits viral replication in infected T-cell cultures.* Sci Rep, 2018. **8**(1): p. 7784.

148.   Quetier, F., *The CRISPR-Cas9 technology: Closer to the ultimate toolkit for targeted genome editing.* Plant Sci, 2016. **242**: p. 65-76.

149.   Johnson, K.M., *Hantaviruses: history and overview.* Curr Top Microbiol Immunol, 2001. **256**: p. 1-14.

150.   Lawrence, M., et al., *The effects of modern war and military activities on biodiversity and the environment.* Environmental Reviews, 2015. **23**.

151.   Lameire, N., *Wars, disasters and kidneys.* Acta Clin Belg, 2014. **69**(6): p. 418-25.

152.   Gajdusek, D.C., *Virus hemorrhagic fevers. Special reference to hemorrhagic fever with renal syndrome (epidemic hemorrhagic fever).* J Pediatr, 1962. **60**: p. 841-57.

153.   Lee, H.W., P.W. Lee, and K.M. Johnson, *Isolation of the etiologic agent of Korean hemorrhagic fever. 1978.* J Infect Dis, 2004. **190**(9): p. 1711-21.

154.   Clement, J., et al., *Hantavirus outbreak during military manoeuvres in Germany.* Lancet, 1996. **347**(8997): p. 336.

155.   Parkes, L.O., et al., *A Cluster of Three Cases of Hantavirus Pulmonary Syndrome among Canadian Military Personnel.* Can J Infect Dis Med Microbiol, 2016. **2016**: p. 2757969.

156.   Bugert, J.J., et al., *Hantavirus infection--haemorrhagic fever in the Balkans--potential nephrological hazards in the Kosovo war.* Nephrol Dial Transplant, 1999. **14**(8): p. 1843-4.

157.   Markotic, A., et al., *Hantaviruses are likely threat to NATO forces in Bosnia and Herzegovina and Croatia.* Nat Med, 1996. **2**(3): p. 269-70.

158. Richardson, K.S., et al., *Human exposure to particulate matter potentially contaminated with sin nombre virus.* Ecohealth, 2013. **10**(2): p. 159-65.

159. Waltee, D., et al., *Seasonal dispersal patterns of sylvan deer mice (Peromyscus maniculatus) within Montana rangelands.* J Wildl Dis, 2009. **45**(4): p. 998-1007.

160. Douglass, R.J., et al., *Deer mouse movements in peridomestic and sylvan settings in relation to Sin Nombre virus antibody prevalence.* J Wildl Dis, 2006. **42**(4): p. 813-8.

161. Nicola, M., et al., *The Socio-Economic Implications of the Coronavirus and COVID-19 Pandemic: A Review.* Int J Surg, 2020.

162. Vapalahti, K., et al., *Puumala virus infections in Finland: increased occupational risk for farmers.* Am J Epidemiol, 1999. **149**(12): p. 1142-51.

163. Van Loock, F., et al., *A case-control study after a hantavirus infection outbreak in the south of Belgium: who is at risk?* Clin Infect Dis, 1999. **28**(4): p. 834-9.

164. Lin, X.D., et al., *Migration of Norway rats resulted in the worldwide distribution of Seoul hantavirus today.* J Virol, 2012. **86**(2): p. 972-81.

165. Munshi-South, J. and J.L. Richardson, *Peromyscus transcriptomics: Understanding adaptation and gene expression plasticity within and between species of deer mice.* Seminars in cell & developmental biology, 2017. **61**: p. 131-139.

166. Risteska-Nejashmikj, V., et al., *Facing of Family Doctor with Hantavirus Infection.* Open Access Maced J Med Sci, 2019. **7**(10): p. 1660-1664.

167. Wang, X., et al., *A case-control study on the risk factors for hemorrhagic fever with renal syndrome.* BMC Infect Dis, 2020. **20**(1): p. 103.

168. Witkowski, P.T., et al., *Gastrointestinal Tract As Entry Route for Hantavirus Infection.* Front Microbiol, 2017. **8**: p. 1721.

169. White, S.M., *Chemical and biological weapons. Implications for anaesthesia and intensive care.* Br J Anaesth, 2002. **89**(2): p. 306-24.

170. Netski, D., B.H. Thran, and S.C. St. Jeor, *Sin Nombre Virus Pathogenesis inPeromyscus maniculatus.* Journal of Virology, 1999. **73**(1): p. 585-591.

171. Zhang, Y.Z., et al., *Hantavirus infections in humans and animals, China.* Emerg Infect Dis, 2010. **16**(8): p. 1195-203.

172. Muranyi, W., et al., *Bovine aortic endothelial cells are susceptible to hantavirus infection; a new aspect in hantavirus ecology.* Virology, 2004. **318**(1): p. 112-22.

173. Krüger, D.H., G. Schönrich, and B. Klempa, *Human pathogenic hantaviruses and prevention of infection.* Hum Vaccin, 2011. **7**(6): p. 685-93.

174. Hooper, J.W., et al., *A lethal disease model for hantavirus pulmonary syndrome.* Virology, 2001. **289**(1): p. 6-14.

175. Hughes, A.L. and R. Friedman, *Evolutionary diversification of protein-coding genes of hantaviruses.* Mol Biol Evol, 2000. **17**(10): p. 1558-68.

176. Klempa, B., *Reassortment events in the evolution of hantaviruses.* Virus Genes, 2018. **54**(5): p. 638-646.

177. Rizvanov, A.A., S.F. Khaiboullina, and S. St Jeor, *Development of reassortant viruses between pathogenic hantavirus strains.* Virology, 2004. **327**(2): p. 225-32.

178. Rodriguez, L.L., et al., *Genetic reassortment among viruses causing hantavirus pulmonary syndrome.* Virology, 1998. **242**(1): p. 99-106.

179. Henderson, W.W., et al., *Naturally occurring Sin Nombre virus genetic reassortants.* Virology, 1995. **214**(2): p. 602-10.

180. Sabino-Santos, G., Jr., et al., *Natural infection of Neotropical bats with hantavirus in Brazil.* Sci Rep, 2018. **8**(1): p. 9018.

181. Bagamian, K.H., et al., *Transmission ecology of Sin Nombre hantavirus in naturally infected North American deermouse populations in outdoor enclosures.* PLoS One, 2012. **7**(10): p. e47731.

182. Juan, E., et al., *Mechanisms of Hantavirus Transmission in Oligoryzomys longicaudatus.* Ecohealth, 2019. **16**(4): p. 671-681.

183. Lonner, B.N., et al., *Seroprevalence against Sin Nombre virus in resident and dispersing deer mice.* Vector Borne Zoonotic Dis, 2008. **8**(4): p. 433-41.

184. Toro, J., et al., *An outbreak of hantavirus pulmonary syndrome, Chile, 1997.* Emerg Infect Dis, 1998. **4**(4): p. 687-94.

185.	Martinez, V.P., et al., *Person-to-person transmission of Andes virus.* Emerg Infect Dis, 2005. **11**(12): p. 1848-53.

186.	Pizarro, E., et al., *Immunocytochemical and Ultrastructural Evidence Supporting That Andes Hantavirus (ANDV) Is Transmitted Person-to-Person Through the Respiratory and/or Salivary Pathways.* Front Microbiol, 2019. **10**: p. 2992.

187.	Bellomo, C., et al., *A newborn infected by Andes virus suggests novel routes of hantavirus transmission: a case report.* Clin Microbiol Infect, 2020. **26**(1): p. 130-131.

188.	Park, S.E., *Epidemiology, virology, and clinical features of severe acute respiratory syndrome -coronavirus-2 (SARS-CoV-2; Coronavirus Disease-19).* Clin Exp Pediatr, 2020. **63**(4): p. 119-124.

189.	Althaus, C.L., *Estimating the Reproduction Number of Ebola Virus (EBOV) During the 2014 Outbreak in West Africa.* PLoS Curr, 2014. **6**.

190.	Woolhouse, M.E., et al., *Assessing the Epidemic Potential of RNA and DNA Viruses.* Emerg Infect Dis, 2016. **22**(12): p. 2037-2044.

191.	Lõhmus, M.J., I.; van de Goot, F.; van Rotterdam B.J., *Rodents as Potential Couriers for Bioterrorism Agents.* Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science, 2013. **11**(S1): p. S247-S257.

192.	Wheelis, M., R. Casagrande, and L. Madden, *Biological Attack on Agriculture: Low-Tech, High-Impact Bioterrorism.* BioScience, 2009. **52**: p. 569-576.

193.	Dudley, J.P. and M.H. Woodford, *Bioweapons, bioterrorism and biodiversity: potential impacts of biological weapons attacks on agricultural and biological diversity.* Rev Sci Tech, 2002. **21**(1): p. 125-37.

194.	Joyner, C.P., et al., *Deer Mice As Laboratory Animals.* Ilar j, 1998. **39**(4): p. 322-330.

195.	Martin, T.L., et al., *Cost and Effectiveness of Commercially Available Nesting Substrates for Deer Mice (Peromyscus maniculatus).* J Am Assoc Lab Anim Sci, 2016. **55**(4): p. 412-8.

196.	Vaheri, A., O. Vapalahti, and A. Plyusnin, *How to diagnose hantavirus infections and detect them in rodents and insectivores.* Rev Med Virol, 2008. **18**(4): p. 277-88.

197.	Kraus, A.A., et al., *Inactivation of Hantaan virus-containing samples for subsequent investigations outside biosafety level 3 facilities.* Intervirology, 2005. **48**(4): p. 255-61.

198.	Hopkins, A.S., et al., *Experimental evaluation of rodent exclusion methods to reduce hantavirus transmission to residents in a Native American community in New Mexico.* Vector Borne Zoonotic Dis, 2002. **2**(2): p. 61-8.

199.	Baldwin, R.A., et al., *Effectiveness of rodenticides for managing invasive roof rats and native deer mice in orchards.* Environ Sci Pollut Res Int, 2014. **21**(9): p. 5795-802.

200.	Glass, G.E., et al., *Experimental evaluation of rodent exclusion methods to reduce hantavirus transmission to humans in rural housing.* Am J Trop Med Hyg, 1997. **56**(4): p. 359-64.

201.	Mari Saez, A., et al., *Rodent control to fight Lassa fever: Evaluation and lessons learned from a 4-year study in Upper Guinea.* PLoS Negl Trop Dis, 2018. **12**(11): p. e0006829.

202.	Mariën, J., et al., *Evaluation of rodent control to fight Lassa fever based on field data and mathematical modelling.* Emerg Microbes Infect, 2019. **8**(1): p. 640-649.

203.	Meerburg, B.G., F.W. Brom, and A. Kijlstra, *The ethics of rodent control.* Pest Manag Sci, 2008. **64**(12): p. 1205-11.

204.	Ferres, M., et al., *Prospective evaluation of household contacts of persons with hantavirus cardiopulmonary syndrome in chile.* J Infect Dis, 2007. **195**(11): p. 1563-71.

205.	Arthur, R.R., J.W. Leduc, and J.M. Hughes, *Surveillance for Emerging Infectious Diseases and Bioterrorism Threats.* Tropical Infectious Diseases, 2006: p. 195-200.

206.	Jansen, H.J., et al., *Biological warfare, bioterrorism, and biocrime.* Clin Microbiol Infect, 2014. **20**(6): p. 488-96.

207.	Keeney, G.L. and D. Von Winterfeldt, *Identifying and structuring the objectives of terrorists.* Risk Anal, 2010. **30**(12): p. 1803-16.

208.	Zhao, J., et al., *Canadian Science Meets Parliament: Building relationships between scientists and policymakers.* Science and Public Policy, 2020.

Chapter 3: Sin Nombre *Orthohantavirus* S and M Segment Sense and Antisense 3′ Noncoding Terminal Regions Putatively Interact with Human Host RNA-Binding Proteins

3.0 Introduction

Sin Nombre orthohantavirus (SNV) is a species of Hantavirus found in the family *Bunyaviridae* [1]. The virus was first identified in 1993 when an outbreak occurred within the Four Corners Region (Utah, New Mexico, Colorado, Arizona) of the USA and caused a sudden onset of acute respiratory failure in healthy individuals [2, 3]. Patients suffering from the infection developed an influenza-like illness that was followed by a rapidly progressing pulmonary oedema, respiratory difficulties, cytokine storm and eventual shock [4-6]. The disease, Hantavirus Cardiopulmonary Syndrome (HCPS), has a case fatality rate of up to 50% [7, 8]. SNV is specifically spread by the common deer mouse (*Peromyscus maniculatus*), an asymptomatic reservoir host which is prevalent in Canada and the USA [9-11]. Both New-World and Old-World Hantaviruses across the western and eastern hemispheres respectively are transmitted by rodents which is distinct from other Bunyaviruses that are normally transmitted by arthropods [12]. Although infections are rare in North America which usually number in the hundreds, thousands of HCPS infections – with many ongoing and unreported – occur in South America which are caused by the Andes Virus (ANDV), another highly pathogenic *orthohantavirus* [13, 14]. Hantaviruses generally are transmitted to humans through the inhalation of aerosolised virions embedded most often in rodent excreta or faecal matter [15, 16]. However, ANDV uniquely can spread from person-to-person in close contact spaces [17]. As an emerging virus with a high mortality rate and transmissibility, HCPS-causing Hantaviruses pose a serious risk to Global North and developing nations, which is further exacerbated by climate change which provokes to increase the natural range of the Hantaviral reservoir hosts as well as adversely affecting infrastructure all of which can lead to increased human-rodent exposure [18-22]. Despite a few vaccine candidates, there are no current US FDA approved antivirals and vaccine therapeutics for Hantavirus infections [23].

Hantaviruses are single-stranded, negative sense, tripartite RNA viruses [24]. SNV's genome contains three segments distinguished by their overall length: Small (S, 2.1 knt), Medium (M, 3.7 knt), and Large (L, 6.6 knt) which encode for the Nucleocapsid (N) Protein, Glycoprotein Precursor (GPC), and an RNA-dependent RNA-polymerase (RdRp), respectively [25-27]. The GPC is post-translationally cleaved to form heterodimer Glycoproteins Gn and Gc which combine to assemble a tetrameric spike protein embedded in the virion's envelope [28]. Hantaviral infection

and entry into endothelial, epithelial, and dendritic cells, and macrophages, are mediated by their spike proteins' interactions with host β3-Integrins [29-31]. β3-Integrins are a class of adhesion receptors that mediate cell-cell and cell-extracellular matrix interactions which are important for the maintenance of tissue integrity, cellular migration, and the regulation of gene expression, cell survival, adhesion, and cellular differentiation events [32, 33]. The genomes of Bunyaviruses are generally conserved at the Noncoding Terminal Regions (NTRs), where the formation of a cyclised, base-pairing structure called a panhandle occurs at the 5′/3′-ends comprising generally up to 60-70 nt for their segmented genomes [34]. Hantaviruses specifically have highly conserved 14-18 nt 5′/3′-terminal sequences comprising of 5′-UAGUAGUAU(G/A)CUCC-3′ at the 5′terminus, and 3′-AUCAUCAUCUGAGG-5′ at the 3′-terminus which form complementary base-pairing elements [25].

The panhandle sequence and its secondary structures are important in relation to the Hantaviral N Protein which packages the tripartite genome into a Ribonucleoprotein Complex (RNP) whilst acting as a chaperone to initiate genome replication and translation and enable genome packaging, traffic, and disassembly [35-38]. The N Protein specifically recognises secondary structure elements on the panhandle sequence, enabling it to bind with high affinity and interact with Hantaviral components to progress its lifecycle [39-42]. Evidently, the N Protein behaves as a viral RNA-Binding Protein (RBP) which is necessary for SNV's replication and survival.

Apart from the complementary, base-pairing panhandle region at the terminal ends, SNV's genome has particularly long noncoding regions on the 3′-termini of the S and M Segments, adjacent to the panhandle sequences [25]. Much akin to eukaryotic Long Noncoding RNAs (LncRNA), these NTRs are likely involved in RNA-Protein binding interactions with human host proteins, behaving similarly in a regulatory manner. LncRNAs in eukaryotes are RNA molecules no smaller than 200 nt, maintaining no protein coding potential whilst also contributing towards regulatory activities in the cell including antiviral functions, gene expression and repression [43, 44]. LncRNAs are also transcribed by plant and animal viruses, and perform functions to help suppress host immune responses thus increasing the virus' survivability [45]. Despite not being autonomously transcribed RNA molecules independently acting from the Hantaviral genome, the NTRs of SNV have the potential to form complex secondary structures and subsequent functional domains to assist in its survival and propagation. This is exhibited by the roles of SNV and ANDV's NTRs in their lifecycles. Principally, promoter regions have been identified in these NTRs in other Hantaviruses such as orthobunyavirus Bunyamwera whose 5′/3′-terminal region base-pairing was required for RdRp binding and subsequent RNA synthesis [46]. The NTRs of SNV also have triplet repeats which specifically bind to

the N Protein and assist with translation initiation, whereas ANDV 3′-Untranslated Region (UTR) stimulates the Cap-Dependent translation initiation of viral RNA [42, 47]. Apart from the base-pairing panhandle sequence, the secondary structures of these NTRs have not been fully characterised, which have the potential to further impact the viral lifecycle through their length and complexity [48].

This investigation intends to identify RBPs that interact with the 3′-NTRs of the S and M Segments of SNV. Due to their length and potential secondary structure complexity, the 3′-NTRs likely have the capacity to interact with human host RBPs that will affect viral lifecycle and viability. Using a Digoxigenin-labelled immunoprecipitation Pull-down assay, the present study has identified several RBPs that interact with the complete NTRs of both the S and M 3′-NTRs [49, 50]. Additionally, the study has mapped their biological processes and functional roles, identifying potential co-interactions of human host RBPs likely occurring during SNV infection. This work intends to build a foundational library of host-pathogen interactions at the RBP-level to expose new, potential drug targets against Hantaviral infections.

3.1 Methods

3.1.1 Plasmid Preparation

SNV Transcripts for the positive and negative polarities of the S, M, and L Segment NTRs were derived from the SNV NMH10 strain (NC_005216.1, NC_005215.1, and NC_005217.1 Accession numbers from NCBI respectively) [25, 27]. SNV transcript sequences were computationally run under a combination of Vienna RNAfold and SFold computational software to predict tentative RNA secondary structure at the lowest free energy for downstream SAXS visualization and potential RBP binding domains [51-57]. Sequence design also included the development of Scrambled Sequences as Pull-down assay controls. Scrambled Sequences were comprised of the same nucleotides from the native NTR sequences, being rearranged and shuffled by the University of Alberta Sequence Manipulation Suite (SMS) and later analysed for sequence and structure similarity using a combination of ClustalW and the University of Alberta SMS pairwise DNA sequence alignments to observe for less than 50% similarity [58-60]. Scrambled Sequences were also visualized with RNA secondary structure software to visually verify differences between the native and scrambled sequences. Transcript design included a 3′ and 5′ flanking Selective 2'-Hydroxyl Acylation Analysed by Primer Extension binding cassettes (12 and 48 nt respectively) on the SNV S3TR(+/-) and M3TR(+/-) to conduct future secondary structure chemical probing determination [61-63]. SHAPE binding cassettes

were omitted from Scrambled NTRs and appeared visibly smaller than the NTRs with SHAPE cassettes (**Table 3.1**). All sequences were provided with a 5′-upstream T7 RNA Polymerase promoter sequence and a 3′-end XBaI restriction endonuclease cut-site sequence.

Table 3.1: Sin Nombre *orthohantavirus* Complete Genome and Non-Coding Terminal Region Lengths for the Small, Medium, and Large Segments. Long segments, S and M3TR NTRs included Selective 2'-Hydroxyl Acylation Analyses by Primer Extension (SHAPE) Cassettes (a 5′- and 3′- linker of 12 and 24nt respectively) for future secondary structure determination by chemical probing techniques [61]. The interest of this study is on S and M3TR NTRs exclusively.

Segment	Sequence Length (nt)	Native NTR Length (nt)	NTR Length SHAPE (nt)	Scrambled NTR Length (nt)
S	2,060	S5TR(+/-) – 42 S3TR(+/-) – 730	S3TR(+/-) – 787	S3TR – 730
M	3,696	M5TR(+/-) – 51 M3TR(+/-) – 222	M3TR(+/-) –279	M3TR – 222
L	6,562	L5TR(+/-) – 35 L3TR(+/-) – 65		

SNV NTR sequences were inserted into Genewiz pUC-57-KAN plasmids and transformed into NEB DH5α competent cells (including Dam-/Dcm- cells for S5TR(+), M5TR(+), and L3TR(-) sequences). A single colony from each respective LB agar plate was inoculated in 50 mL of Lysogeny Broth containing 50μg/mL Kanamycin and incubated at 37°C with 220rpm for 16 hours. SNV plasmids were purified using NEB Monarch Miniprep Kits (Canada). 10μL of purified plasmids (2μL sample, 8μL MilliQ ddH₂O, and 2μL 6x DNA Loading Dye) were loaded and run on a 1% agarose gel (Cole-Parmer E1101 Mini electrophoresis system, USA; for 100V, 25min) against a 100 bp-10 Kb Wide Range DNA marker (BioBasic, USA) for purity from RNA and cell debris contamination. Agarose gels were made with 0.4μL/50 mL 1xTAE SYBR™ Safe DNA Gel Stain (Invitrogen, USA). All agarose and PAGE gels were visualized on the Amersham Typhoon (GE Healthcare, USA) for purity and monodispersity of the plasmids and RNA samples. Cognate purified plasmids were generated from multiple minipreps to optimum plasmid concentrations (measured by BioDrop μLITE+, USA), pooled and EtOH precipitated and stored at -80°C for downstream *in vitro* transcription reactions.

3.1.2 *In vitro* Transcriptions and RNA Purification by Size Exclusion Chromatography

Concentrated plasmid samples were digested using XbaI restriction endonuclease (NEB, Canada) and were visualized on a 1% agarose gel (100V, 35min) for complete digestion. 1 mL *in vitro* transcription reactions were

conducted for 4 hours at 37ºC, using laboratory purified T7 RNA Polymerase and RiboLock RNase Inhibitor (Thermo Fisher Scientific, USA) [64]. Intricate plasmids were additionally incubated with 5% DMSO and 0.1% Triton X-100 detergent [65]. RNA Transcripts were purified using SEC on an ÄKTA pure 25 FPLC with a Superdex 200 10/300 GL column (GE Healthcare, USA). Sense and antisense S3TR RNA Transcripts were purified using a Cytiva Sephacryl S-400 High Resolution column and resin (Cytiva, USA) [66, 67]. All samples were eluted using a HEPES Buffer (50mM HEPES, 150mM NaCl, 15mM MgCl$_2$, 3% Glycerol, pH 7.4). Urea-PAGE and Native PAGE gels were performed from the selected SEC chromatogram peaks [68, 69]. 8-10% Urea-PAGE gels were ran specifically for the larger (222nt and 730nt) transcripts of S3TR(+/-) and M3TR(+/-) while 12% Urea-PAGE gels were ran for the remaining transcripts of the Hantaviral NTRs (35-65nt). A second 2% denaturing agarose gel was run in tandem (100V, 25min) to confirm for the presence and quality of purified RNA [70]. Confirmed RNA was stored in elution fractions in -80ºC followed by pooling and EtOH or LiCl concentration dependent precipitation prior to 5nt 3′Digoxigenin RNA-Labelling [71].

3.1.3 Culturing Epithelial Adenocarcinoma A549 Cell line for Pull-down assay

Epithelial Adenocarcinoma A549 cells were used to produce the RBP enriched cell lysate because of SNV's infection of endothelial pulmonary cells and epithelial cells [72, 73]. A549 cells were obtained from the University of Calgary (Christopher Mody, Department of Microbiology, Immunology and Infectious Diseases, Canada) and were grown in T-75 culture flasks in an atmosphere of 5% CO$_2$ at 37ºC [74]. A549 cells were maintained in a proliferation medium (F12K) which contained 10% Foetal Bovine Serum (FBS), and 1% penicillin-streptomycin. Cells were seeded at Day 0 in culture plates at a density of ~2,000,000 cells/flask with proliferation media. Cells attached over 24 hours with proliferation media replaced daily. Cells reached 80% confluency within three days and passaged to produce one flask per NTR sequence replicate for the Pull-down assays (150mm dish plate). A549 cells were freshly harvested using scraping from the 150mm dish plates into cold PBS and pelleted by centrifugation at 1500 RPM for 15 minutes. Cells were then resuspended in cold PBS and transferred to 2 mL microcentrifuge tubes to be re-pelleted at 2500 RPM for 3 minutes and homogenised to be used directly in the Pull-down assay.

3.1.4 RNA Labelling with 5nt 3′-Digoxigenin RNA Linker and Crosslinking with Anti-Digoxin Antibodies to A/G Magnetic Beads, Pull-down assay, and Mass Spectrometry

10 µL of a minimum of ~30 µM Purified RNA was labelled with 5nt 3′-Digoxigenin (DIG) (IDT, USA) in a thermocycler (PTC-100 Programmable Thermocycler, BioRad, USA) at 12ºC for 12 hours using T4 RNA Ligase 1 (NEB, Canada) [49, 75, 76]. Undergoing end-labelling reactions, the Digoxigenin-11-dUTP is ligated to the NTRs' 3′-end to minimise interrupting RNA structure formation and subsequent nonspecific RBP binding [49, 50, 77]. A total of four replicates of sense and antisense S3TR and M3TR as well as S3TR and M3TR Scrambled Dig-Labelled RNA was produced. A total of two replicates for antisense M3TR Dig-Labelled RNA was produced. DIG-Labelled RNA was then purified using an EZ-10 Spin Column RNA Clean-up and Concentration Kit (BioBasic, USA) and eluted in nuclease-free water to achieve a minimum ~3µM DIG-Labelled RNA. The DIG-Labelled RNA binding efficiency was assessed using 6% Native Page EMSAs run at 4ºC in PBS running buffer (50mM potassium phosphate buffer pH 7, 2% glycerol, 6% 29:1 Acrylamide:Bisacrylamide). Binding Efficiency gels included a cross-comparison of RNA-DIG labelled RNA against RNA-DIG-Antibody complexes to observe a visible shift in band migration. Quick 2% agarose gels (100V, 30 min) were also run to corroborate the results.

A/G Magnetic beads (Pierce Protein A/G Magnetic Beads, Thermo Scientific, USA) were washed and prepared using PBS and a magnetic bar. Anti-Digoxin antibody (IgG Fraction Monoclonal Mouse Antibody, Jackson ImmunoResearch, USA) was incubated with A/G Magnetic beads at 4ºC for 1 hour. Beads were washed with 0.2mM Triethanolamine, pH 8.2 and resuspended with buffer containing 25mM Dimethyl pimelimidate and 0.2mM Triethanolamine, pH 8.2. Crosslinking of the Anti-Digoxin antibodies and A/G Magnetic beads followed incubation on a tube roller for 45 minutes with 15-minute interval fresh washes of 25mM Dimethyl pimelimidate. Additional washes included 0.1M Ethanolamine, pH 8.2 and incubation on a tube roller at room temperature for 30 minutes. Bead were washed again in PBS and once with 0.1M glycine, pH 2.5. Samples were finally washed using PBS with 0.1% Tween and 0.02% Sodium azide for storage overnight. Final Cross-linked Antibody Magnetic Bead Complex was stored in PBS-Tween and Sodium azide overnight.

DIG-Labelled RNA and Anti-Digoxin antibody-labelled A/G Magnetic beads were incubated together at 4ºC for 1 hour. DIG-Labelled RNA and A/G Magnetic bead complex was equilibrated using a Cytoplasmic Lysis Buffer (CLB: 25mM HEPES, pH 7.9, 5mM KCl, 0.5mM $MgCl_2$, 0.5% NP-40, Ribolock RNase Inhibitor, 1 HALT phosphatase and protease inhibitor cocktail (Thermo Scientific, USA), and DNase. Pelleted A549 cells were

resuspended in CLB and incubated on a tube roller at 4ºC for 5 minutes before being pelleted at 5000 rpm for 5 minutes to remove insoluble material. The cytoplasmic fraction of the cell lysate was combined with Cytoplasmic Immunoprecipitation Buffer (CIPB: 25mM HEPES, pH 7.9, 5mM KCl, 0.5mM $MgCl_2$, 200mM NaCl, Ribolock RNase Inhibitor, and 1x HALT phosphatase and protease inhibitor cocktail), with an aliquot set aside as the Pre-Immunoprecipitation control for the SDS-PAGE gels. An aliquot of cell lysate was added to the DIG-labelled RNA A/G Magnetic bead complex and mixed on a tube roller for 4 hours at 4ºC. Beads were pelleted using the magnetic bar and an aliquot of supernatant was set aside as the Post-Immunoprecipitation control. Beads were washed several times with CIPB and resuspended. Aliquots were taken from each sample and ran on an SDS-PAGE gel (6% Stacking, 12% Resolving 29:1 Acrylamide:Bisacrylamide) at 120-150V for 30 minutes or 1 cm into the gel. Gels were fixed using the modified "Blue-Silver" Coomassie Staining Neuhoff procedure developed by Richard Fahlman (University of Alberta, Department of Biochemistry, Canada), in preparation for Mass Spectrometry [78]. Immunoprecipitated bands were excised from the gel and subsequently submitted to the Alberta Proteomics and Mass Spectrometry Facility (Jack Moore, University of Alberta, Canada) for in-gel digestion and protein identification by Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) using a Q Extractive Mass Spectrophotometers [79-81].

3.1.5 Bioinformatic Analysis of Human Proteins that can Potentially Interact with Hantaviral NTRs

RBPs immunoprecipitated during the Pull-down assay and identified by MS underwent identity and property classifications to distinguish potentially viable proteins from loosely binding, to non-specific binding proteins. Total Protein Hits (TPH) were analysed for cytosolic activity and underwent bioinformatic review. Nuclear proteins and cytoskeletal proteins were excluded with proteins being analysed by STRING to correlate known protein-protein interactions and relationships [82-87]. STRING analysis observed interaction sources derived from experiments, databases, co-expression, and gene fusion with a minimum required interaction score of 0.700 (high confidence). Identified proteins binding to the Scrambled RNA sequences were also excluded producing remaining High-Hit Proteins (HHP). A PANTHER analysis was conducted to identify biological pathways, and evolutionary and function dependent relationships with the remaining HHPs [88, 89]. HHPs exhibiting a high MS score and a potential RNA-Binding role were included in the proteins expressed for future directed biophysical interaction studies.

3.2 Results

3.2.1 Purification of Sense and Antisense S3TR and M3TR NTRs using Size Exclusion Chromatography

The Digoxigenin-Labelling reactions required a minimum of 30μM of each NTR to generate roughly 3μM of Digoxigenin-Labelled RNA in a 100μL reaction. Incidentally, RNA transcription and purification underwent multiple iterations to generate suitable, unaggregated sample concentrations. **Figures 3.1A/B** show the purification of sense and antisense and scrambled S3TR and M3TR RNAs by SEC using the 16/60 Sephacryl-400 and Superdex 200 GL 10/300 columns. Peak fractions between 12 and 16 mL from the S-400 column were pooled and concentrated for sense and antisense S3TR. S3TR Scrambled NTRs were pooled between 14 to 16 mL, being comparatively smaller in size as indicated in **Figure 3.1C**. M3TR RNAs were pooled and concentrated between 10 to 12 mL; M3TR Scrambled RNA was pooled from 11 to 13 mL, being comparatively smaller as indicated in **Figure 3.1D**. S3TR NTRs demonstrated pure, monodisperse RNA depicted in the 8% Urea gel and migrating to ~800 bp (**Figure 3.1C**). M3TR sense and antisense RNA generally migrated in single bands around ~300 bp with mild contamination. M3TR Scrambled migrated around ~200 bp with mild degradation (**Figure 3.1D**).
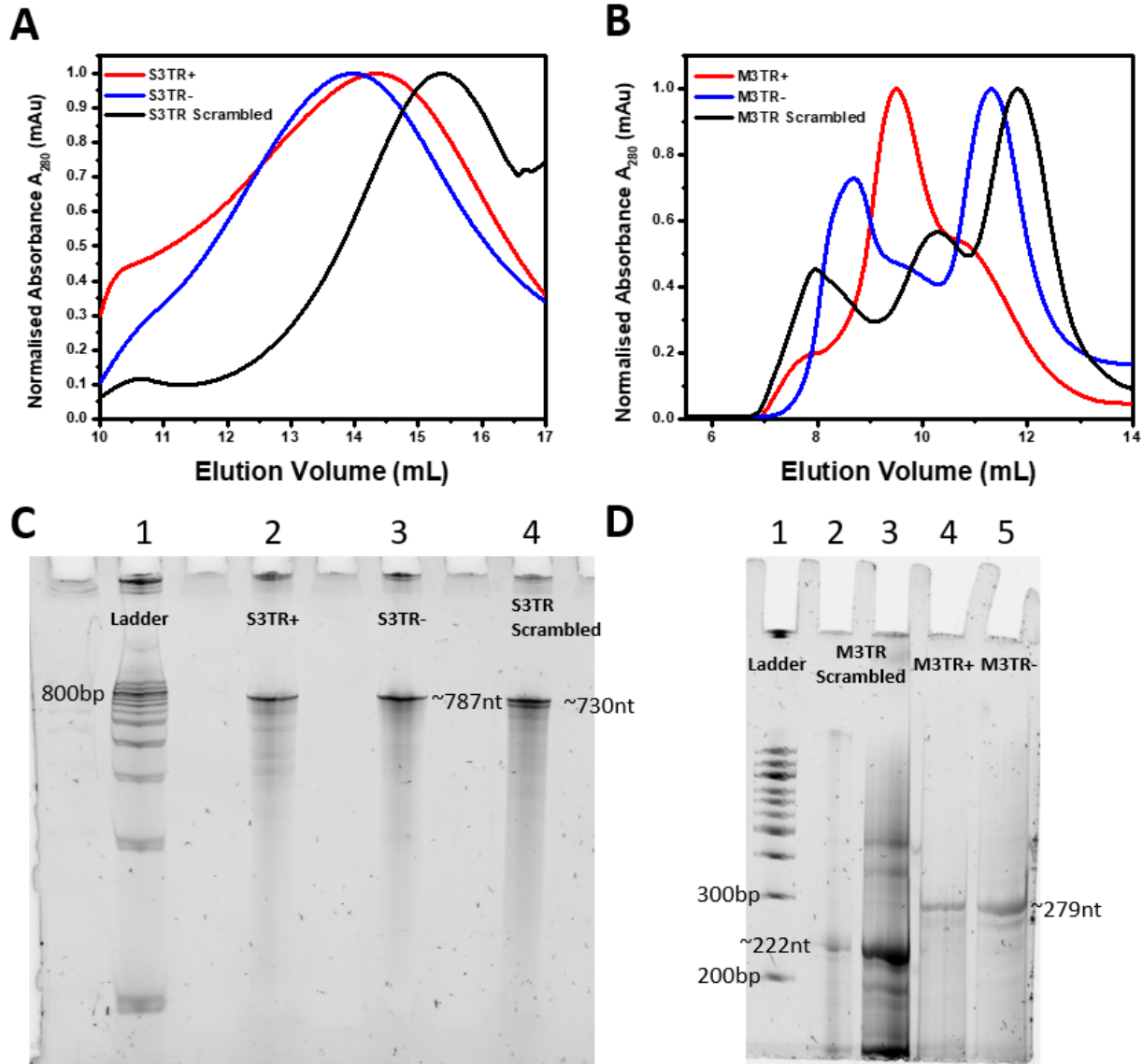
Figure 3.1: The Purification of Sense and Antisense S3TR and M3TR NTRs. (**A**) and (**B**) depict the Size Exclusion Chromatogram of the sense and antisense S3TR and M3TR NTRs respectively, using the 16/60 Sephacryl-400 HR and Superdex 200 GL 10/300 columns. (**C**) depicts the 8% Urea denaturing gel for S3TR NTRs, with Lanes 2, 3, and 4 representing S3TR+, S3TR-, and S3TR Scrambled RNA, respectively. (**D**) depicts the 10% Urea denaturing gels M3TR NTRs, with Lane 2 and 3 representing the M3TR Scrambled RNAs (diluted and undiluted), 4 and 5 representing M3TR+ and M3TR- respectively. Quick-Load® Purple 100 bp DNA Ladder (NEB, Canada) was used for the Urea denaturing gels indicated in Lanes 1 for each gel. RNAs loaded onto gels were highly concentrated from the consolidated SEC fractions which were then supplied to Digoxigenin-Labelling reactions and downstream Pull-down assay Immunoprecipation experiments.

66

3.2.2 Putative RNA-Binding Proteins from S and M3TR Sense and Antisense RNA Pull-down assays form Multiple STRING Networks

Following MS, total proteins were grouped based upon their overall scores and were selected or excluded based upon their nuclear or cytosolic localisation and overlap with scrambled proteins. Special inclusions observed proteins that occurred at least three out of the four hits in each replicate Pull-down assay. Proteins which numbered fewer than three hits were excluded. Some proteins that experienced a single hit in Scrambled RNA Pull-Downs, but a majority hit on native RNA Pull-down assays, were included. Nuclear proteins such as Histones, Heterogeneous Ribonucleoproteins K, A/B, and C/D, as well as Nucleosome Assembly proteins were excluded. Nuclear proteins would likely not interact with the Hantaviral vRNA or mRNA, because Hantaviral assembly, replication, and translation occurs exclusively in the cytoplasm [48]. Mitochondrial lumen and Endoplasmic Reticulum lumen proteins, such as the 60 kDa Heat Shock Protein (HSP) or Endoplasmic Reticulum Chaperone BiP, were also excluded. Proteins having functional or regulatory roles present in both the nucleus and cytoplasm were included. Cytoskeletal proteins such as Actin, Myosin, Plectin, Prelamin-A/C, and Vimentin were excluded. Fibrous intermediate filament proteins were considered for potential proteins involved in vRNA trafficking but were not included in the final Hight-Hit RBP list. Keratin, a major contaminant from human skin and hair as well as organic clothing, was highly present in all the MS runs. Keratin detection itself can be minimised using laminar flow hoods and Keratin-free clean rooms but can be acquired through unexpected sources with detection being somewhat inevitable [90]. Trypsin-1 and RNase Inhibitors as part of the A549 cell processing and RNA maintenance steps were excluded. However, harvested cells were conducted using a scraping tool to avoid the introduction of Trypsin which would have likely contaminated and degraded RBPs isolated during the Pull-Down. The intent was to minimise protein degradation. Final High-Hit Proteins (HHP) have been organised in **Table 3.2**.

The Total Protein Hits (TPH) were analysed using STRING to identify interactions between potential RBPs immunoprecipitated using Digoxigenin-labelled SNV NTRs as to determine their pathway, function, and network relationships. STRING analyses two nonidentical proteins that originate from different protein-coding gene loci and generates an association that does not differentiate between splicing variants or post-translationally modified protein isoforms that could be encoded from the same gene locus [87]. This is instead consolidated and represented as a single protein with proteins relationships being scaled between zero and one on whether they're biologically meaningful as dependent from the literature. **Figure 3.2A** depicts the six major clusters (**L**) of TPH likely involved across the RBPs

67

immunoprecipitated by sense S3TR. **L1** depicts several immunoprecipitated proteins that are independently arranged with no known interactions. Included is an HHP, Y-Box-Binding Protein 3 (YBX3). **L2** groups translational and ribosomal interacting proteins, including several amino acid-tRNA ligases and complex forming proteins such as Aminoacyl tRNA Synthase Complex-Interacting Multifunctional Protein 2. **L2** is connected to **L3** and **L4** through association with several Transferases, Helicases, and Transcription Factors. **L3** itself is comprised of pre-mRNA-Binding Proteins including rRNA 2'-O-Ethyltransferase Fibrillarin, RNA Cytidine Acetyltransferase, and including Nucleolar GTP-binding Protein 1. Nonincluded in **L3** is another HHP, Y-Box-Binding Protein 1 (YBX1), which is highly associated with Pre-mRNA-Splicing Factor ATP-dependent RNA Helicase DHX15. **L4** is highly connected with **L3′**s Probable ATP-dependent RNA Helicase DDX5, and **L2**'s Bifunctional Glutamate/proline-tRNA Ligase with Interleukin Enhancer-Binding Factor (ILF) 2. **L4** contains proteins involved in transcription, including two HHPs: Poly(rC)-Binding Protein 1 and 3 (PCBP1/3). **L3′**s DHX15 is associated with **L5** which is comprised of Mitochondrial and Ribosomal HSP 70 Family of Molecular Chaperones. Finally, **L6** is composed of glycosyltransferase proteins associated to 78 kDa Glucose-regulated Protein Heat Shock 70 protein.

Figure 3.2A: STRING network of Protein-Protein interactions from the Immunoprecipitation Pull-down assays of sense S3TR RNA [86]. Each protein is represented by a coloured node and is filled with a known or predicted three-dimensional structure; unknown structures are filled by an empty node. Edges represent Protein-Protein associations which are specific and meaningful. Known interactions are indicated by lines: (–) from curated databases; and (--) from experimentally determined sources. Predicted interactions are indicated by the lines: (--) from gene neighbourhoods; (--) gene fusions; and (--) gene co-occurrences. Other interactions are indicated by (--) text mining; (--) co-expression; and (--) protein homology. Immunoprecipitated proteins include top scorers from MS and have not been excluded from overlapping proteins from the scrambled sequences. Clusters are grouped by colour: **L1** (Cyan) groups independently arranged proteins with no known database interactions; **L2** (Light Yellow) groups proteins involved in Translation and Ribosomal activity; **L3** (Light Green) groups proteins involved in RNA stability and regulation; **L4** (Light Red) groups proteins involved in Transcription; **L5** (Light Blue) groups proteins involved in Molecular Chaperone activities; and **L6** (Light Grey) groups proteins involved in Glycosyltransferase activities. Sense S3TR HHPs from MS are indicated by a Red Circle.

Figure 3.2B: STRING network of Protein-Protein interactions from the Immunoprecipitation Pull-down assays of antisense S3TR RNA [86]. Each protein is represented by a coloured node and is filled with a known or predicted three-dimensional structure; unknown structures are filled by an empty node. Edges represent Protein-Protein associations which are specific and meaningful. Known interactions are indicated by lines: (—) from curated databases; and (--) from experimentally determined sources. Predicted interactions are indicated by the lines: (--) from gene neighbourhoods; (--) gene fusions; and (--) gene co-occurrences. Other interactions are indicated by (—) text mining; (--) co-expression; and (—) protein homology. Immunoprecipitated proteins include top scorers from MS and have not been excluded from overlapping proteins from the scrambled sequences. Clusters are grouped by colour: **L1** (Cyan) groups independently arranged proteins with no known database interactions; **L2** (Light Yellow) groups proteins involved in Translation and Ribosomal activity; **L3** (Light Brown) groups proteins involved in Helicase and Transferase Ribosomal Activity; **L4** (Light Red) groups proteins involved in Transcription; and **L5** (Light Blue) groups proteins involved in Molecular Chaperone activities. Antisense S3TR HHPs from MS are indicated by a Red Circle.

Similarly, **Figure 3.2B** depicts the STRING analysis of antisense S3TR TPH which shows four major clusters reticulated between three trajectories. Again, independently arranged proteins with no known database associations are grouped in **L1**, which also contains HHP YBX3. **L2** depicts a cluster of aa-tRNA Ligases with similar associations to **L3** Helicase and Transferase Ribosomal cluster proteins and **L4** Transcription proteins through RNA Cytidine Acetyltransferase and ILF2 and the ATP-dependent RNA Helicase A (DHX9) respectively. IFL2/3 and regulatory mRNA packaging and splicing proteins like Heterogeneous Nuclear Ribonucleoprotein A1 contribute to **L4** which includes properties of inflammation and immunity. The HHP YBX1 is located within **L4**. **L3** primarily involves proteins associated to Ribosomal RNA (rRNA) regulation. A similar Molecular Chaperone cluster compared to sense S3TR TPH **L5** in **Figure 3.2A** is comprised of Heat Shock 70 Protein Family proteins are present in **L5**. Several of these associations overlap between gene fusions, neighbourhoods, and co-occurrences for a major translation-oriented cluster. The RNA Transferases and DNA Helicases which include DDX27 and DDX18 form a major cluster of experimentally determined relationships oriented around ribosomal regulation. A significant node, DHX15 interacts with a primary HSP-oriented cluster that retained experimentally determined interactions which included mitochondrial regulatory proteins like ADP/ATP Translocase 1/2/3.
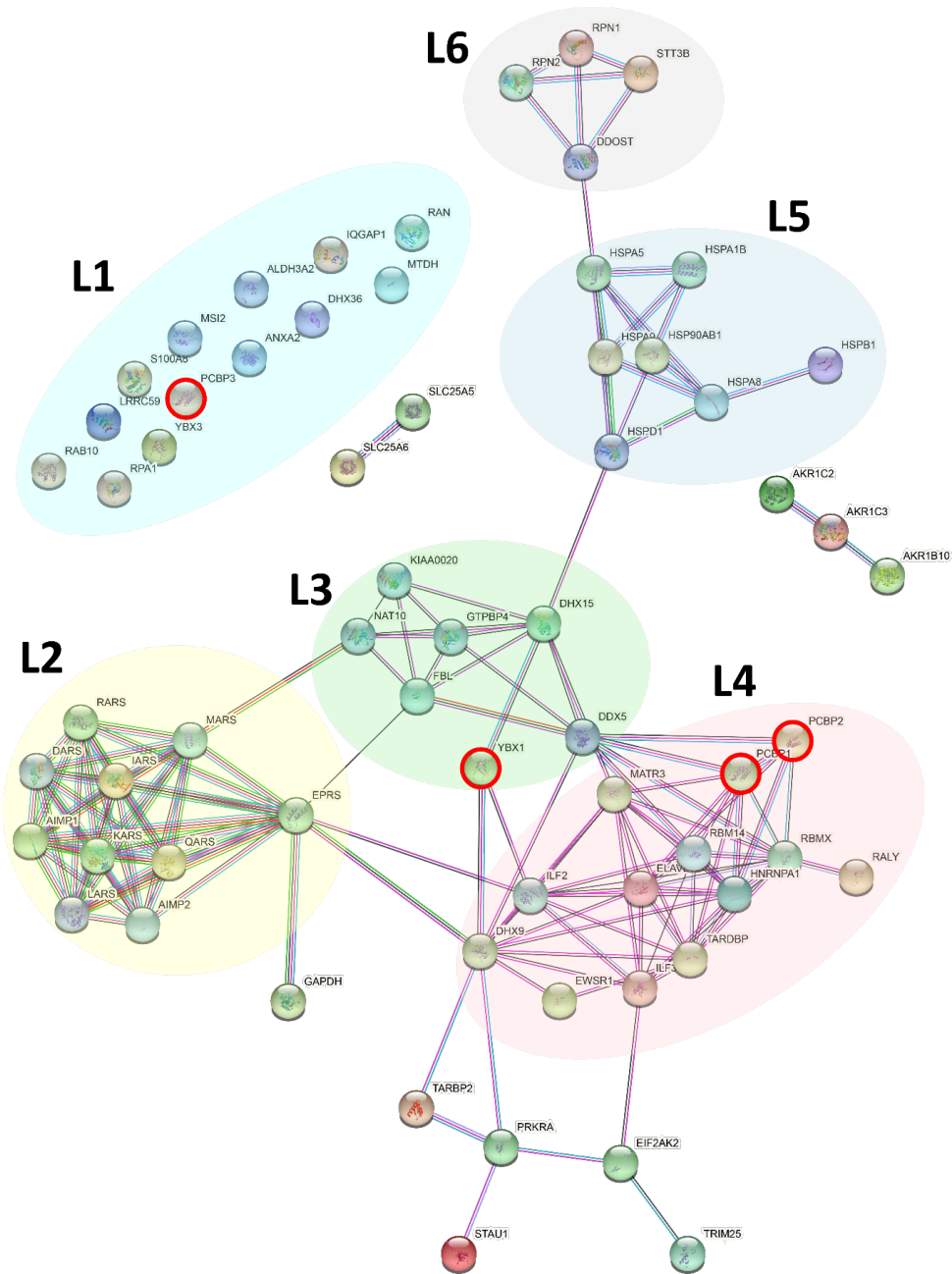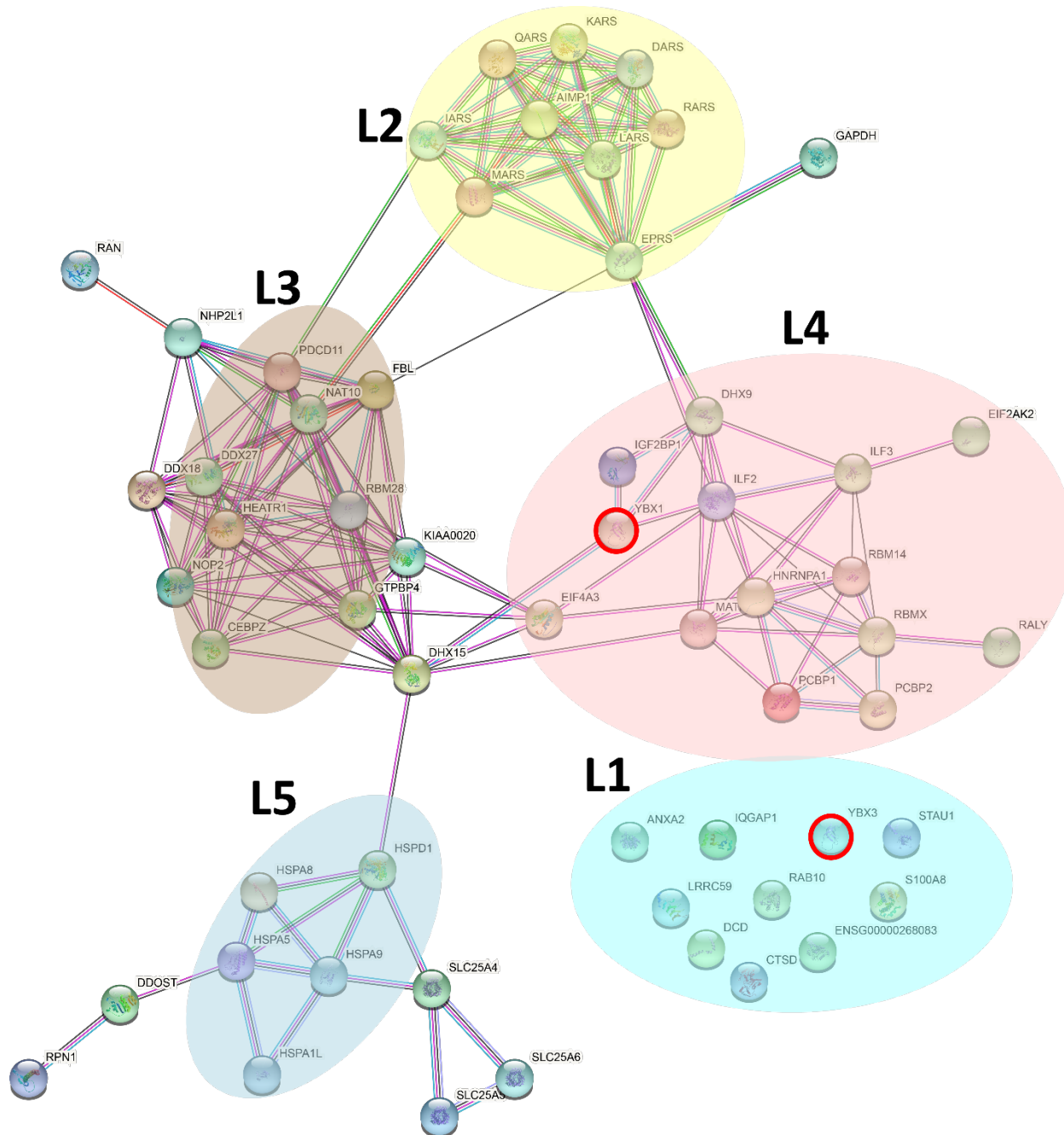
Figure 3.2C: STRING network of Protein-Protein interactions from the Immunoprecipitation Pull-down assays of sense M3TR RNA [86]. Each protein is represented by a coloured node and is filled with a known or predicted three-dimensional structure; unknown structures are filled by an empty node. Edges represent Protein-Protein associations which are specific and meaningful. Known interactions are indicated by lines: (--) from curated databases; and (--) from experimentally determined sources. Predicted interactions are indicated by the lines: (--) from gene neighbourhoods; (--) gene fusions; and (--) gene co-occurrences. Other interactions are indicated by (--) text mining; (--) co-expression; and (--) protein homology. Immunoprecipitated proteins include top scorers from MS and have not been excluded from overlapping proteins from the scrambled sequences. Clusters are grouped by colour: **L1** (Cyan) groups independently arranged proteins with no known database interactions; **L2** (Light Red) groups proteins involved in Transcription; **L3** (Light Purple) groups proteins involved in pre-RNA Splicing; and **L4** (Light Blue) groups proteins involved in Molecular Chaperone activities. Sense M3TR HHPs from MS are indicated by a Red Circle.

STRING analysis of sense M3TR immunoprecipitated TPH reveals a single networked organisation linking

three major clusters and a smaller independent organisation with several independently arranged proteins. **Figure**

3.2C depicts the arrangement of a single, smaller cluster organised around Glycolysis, including Pyruvate kinase PKM and Glyceraldehyde-3-phosphate dehydrogenase. **L1** presents once more independently arranged proteins. Inflammation and immune response proteins and those involved with transcription such as IFL2/3, ELAVL1, with nuclear regulatory proteins such as RNA-binding Motif Protein X, again are observed (**L2**). The **L3** cluster depicts a smaller arrangement of pre-RNA regulatory and Helicase proteins that connect **L2** and **L4** through the associative network. Again, mitochondrial regulatory proteins form their own cluster (**L4**) around the HSP 70 Family of proteins specifically, with 60 kDa HSP interacting with DHX15. Included is the HHP Heat Shock Cognate 71 kDa Protein (HSPA8). Double-stranded RNA-binding Protein Staufen Homolog 1 (STAU1), an RBP is associated by experimental determination to Interferon-inducible Double-stranded RNA-dependent Protein Kinase Activator A as part of the inflammation and immune response cluster. STAU1 is a HHP associated through **L2**'s Transcription associated proteins, by way of ATP-dependent RNA Helicase A.

Figure 3.2D depicts the STRING analysis of TPH for antisense M3TR RNA, which comprise three distinctly unconnected clusters. **L1** again shows the independently arranged proteins with no known database associations. **L2** includes Transcriptional and immune response proteins which also contains two HHPs: ELAV-like Protein 1 (ELAVL1) and Interferon-inducible Double-stranded RNA-dependent Protein Kinase Activator A (PRKRA). **L3** contains the Molecular Chaperone cluster of mainly HSP 70 Family proteins. Heat shock Cognate 71 kDa Protein (HSPA8) is a HHP present within. Pyruvate Kinase, a glycolysis protein and HHP, forms an independent organisation with other glycolysis proteins Glyceraldehyde-3-phosphate dehydrogenase and Glyceraldehyde-3-phosphate dehydrogenase.
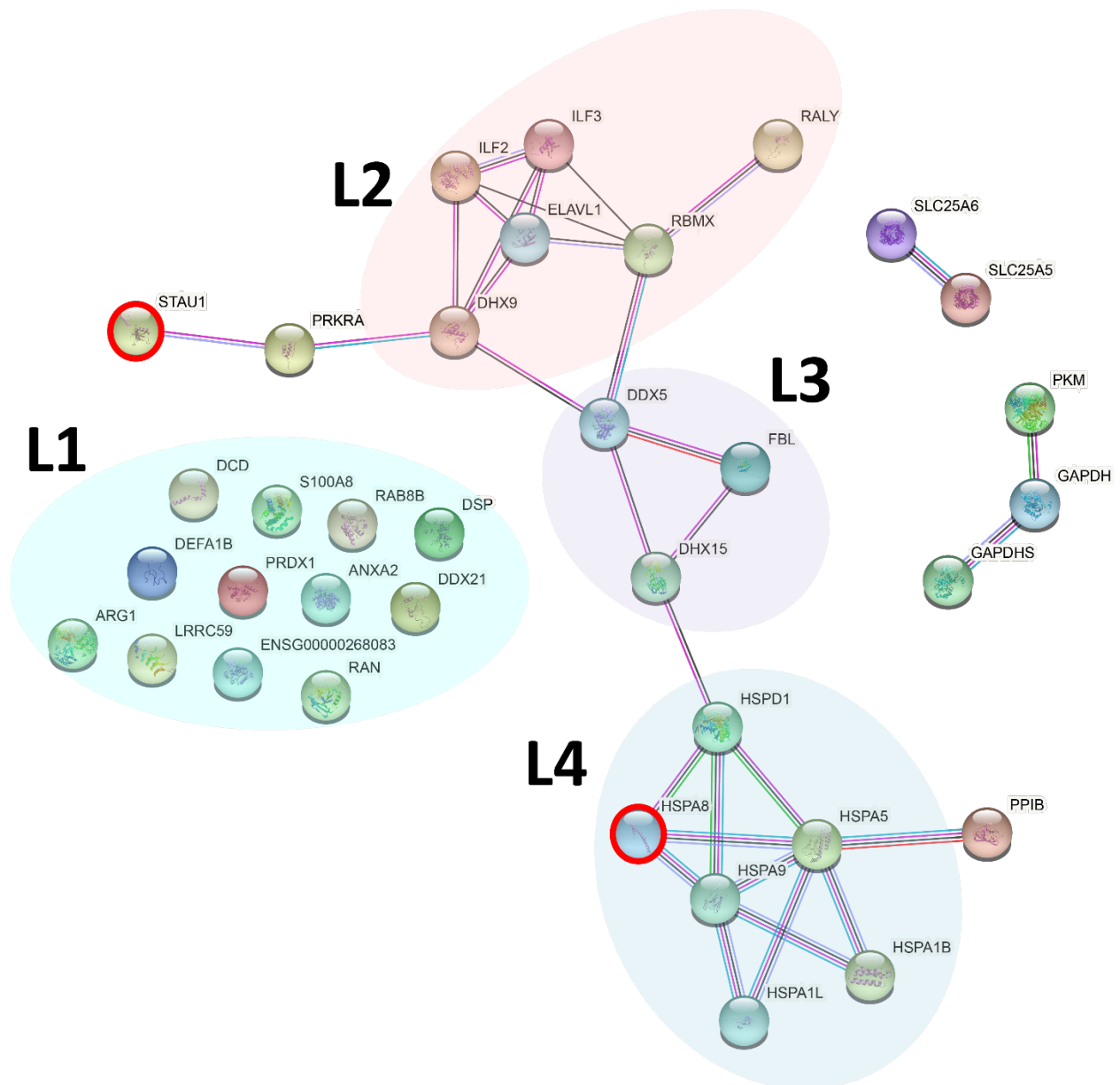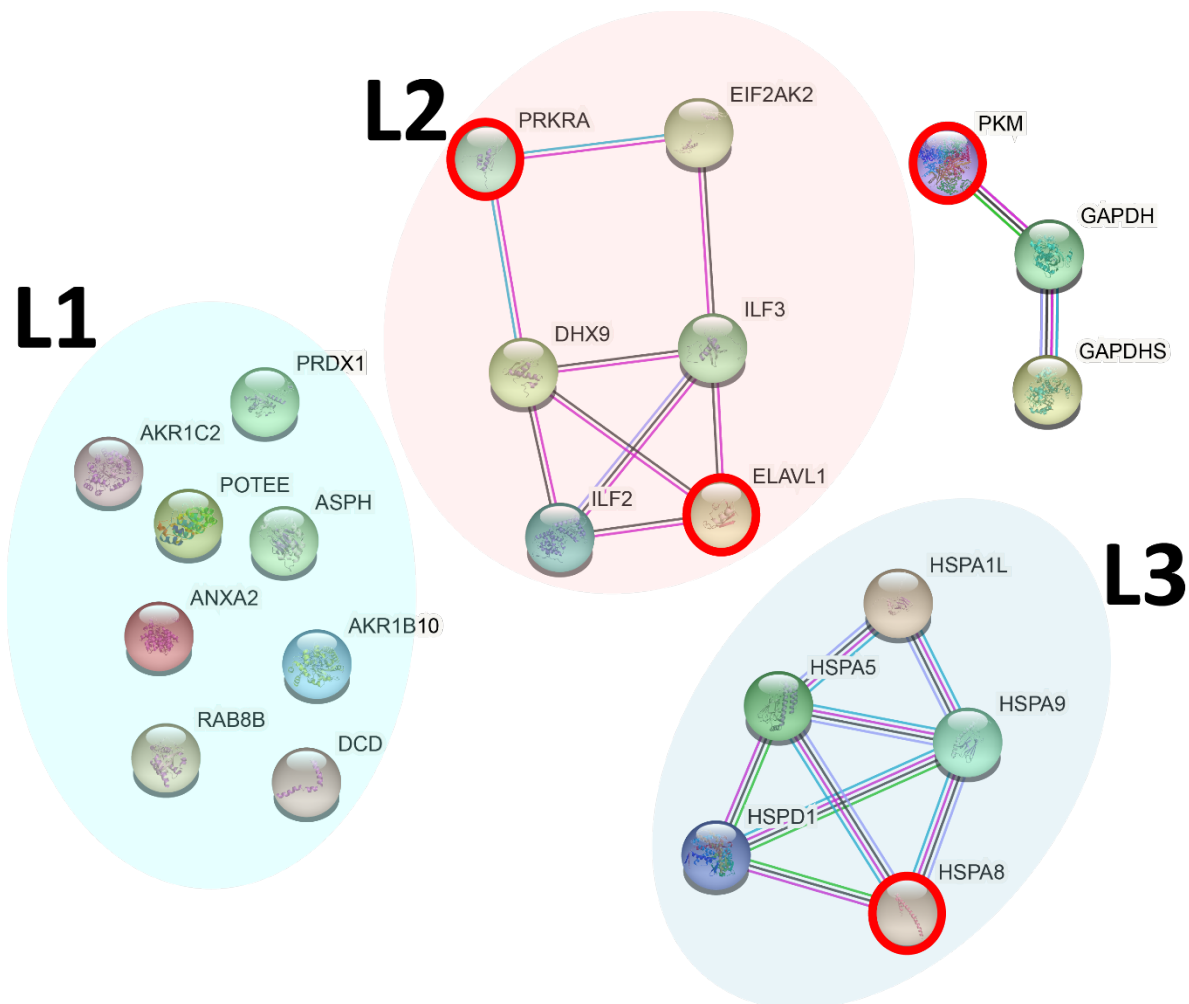
Figure 3.2D: STRING network of Protein-Protein interactions from the Immunoprecipitation pull-down assays of antisense M3TR RNA [86]. Each protein is represented by a coloured node and is filled with a known or predicted three-dimensional structure; unknown structures are filled by an empty node. Edges represent Protein-Protein associations which are specific and meaningful. Known interactions are indicated by lines: (—) from curated databases; and (--) from experimentally determined sources. Predicted interactions are indicated by the lines: (--) from gene neighbourhoods; (--) gene fusions; and (--) gene co-occurrences. Other interactions are indicated by (--) text mining; (--) co-expression; and (--) protein homology. Immunoprecipitated proteins include top scorers from MS and have not been excluded from overlapping proteins from the scrambled sequences. Clusters are grouped by colour: **L1** (Cyan) groups independently arranged proteins with no known database interactions; **L2** (Light Red) groups proteins involved in Transcription; and **L3** (Light Blue) groups proteins involved in Molecular Chaperone activities. Antisense M3TR HHPs from MS are indicated by a Red Circle.

3.2.3 PANTHER Classification of Refined Putative RNA-Binding Proteins from S and M3TR Sense and Antisense RNA Pull-down assays demonstrate Molecular and Biological Importance

Table 3.2: Total High-Hit, Scramble Excluded Immunoprecipitated Proteins. HHPs are organised by their respective NTR with corresponding Gene and Accession Number. All proteins are *Homo sapiens* specific as isolated during the MS process. MS Scores were averaged and organised from highest to lowest, with most prevalent HHP at the top.

	Protein Name	Gene	Accession Number	Averaged MS Score
Sense S3TR	1. **Poly(rC)-binding protein 1**	**PCBP1**	Q15365	**13.64**
	2. **Poly(rC)-binding protein 2**	**PCBP2**	Q15366	**13.39**
	3. **Y-box-binding protein 1**	**YBX1**	P67809	**9.65**
	4. **Y-box-binding protein 3**	**YBX3**	P16989	**8.21**
Antisense S3TR	1. **Y-box-binding protein 3**	**YBX3**	P16989	**7.49**
	2. **Y-box-binding protein 1**	**YBX1**	P67809	**6.47**
Sense M3TR	1. **Heat Shock Cognate 71 kDa Protein**	**HSPA8**	P11142	**10.44**
	2. **Double-stranded RNA-binding Protein Staufen Homolog 1**	**STAU1**	O95793	**5.58**
Antisense M3TR	1. **Heat Shock Cognate 71 kDa Protein**	**HSPA8**	P11142	**11.09**
	2. **ELAV-like Protein 1**	**ELAVL1**	Q15717	**6.54**
	3. **Pyruvate Kinase**	**PKM**	P14618	**6.27**
	4. **Interferon-inducible Double-stranded RNA-dependent Protein Kinase Activator A**	**PRKRA**	O75569	**2.94**

The PANTHER (Protein ANalysis THrough Evolutionary Relationships) classification system was employed to classify immunoprecipitated proteins and their genomes according to family and subfamily, molecular function, biological processes, and pathways which includes relationships between interacting molecules [91]. PANTHER itself is an integrated knowledgebase of both evolutionary and functional relationships between proteins and their genomes. The classification occurs over two axes: evolutionary groupings such as protein class, family, and subfamily; and functional groupings that includes gene ontology and pathways. Evolutionary groupings correspond to a natural classification based on their evolutionary histories while functional groupings classify individual proteins according to their specific functions.
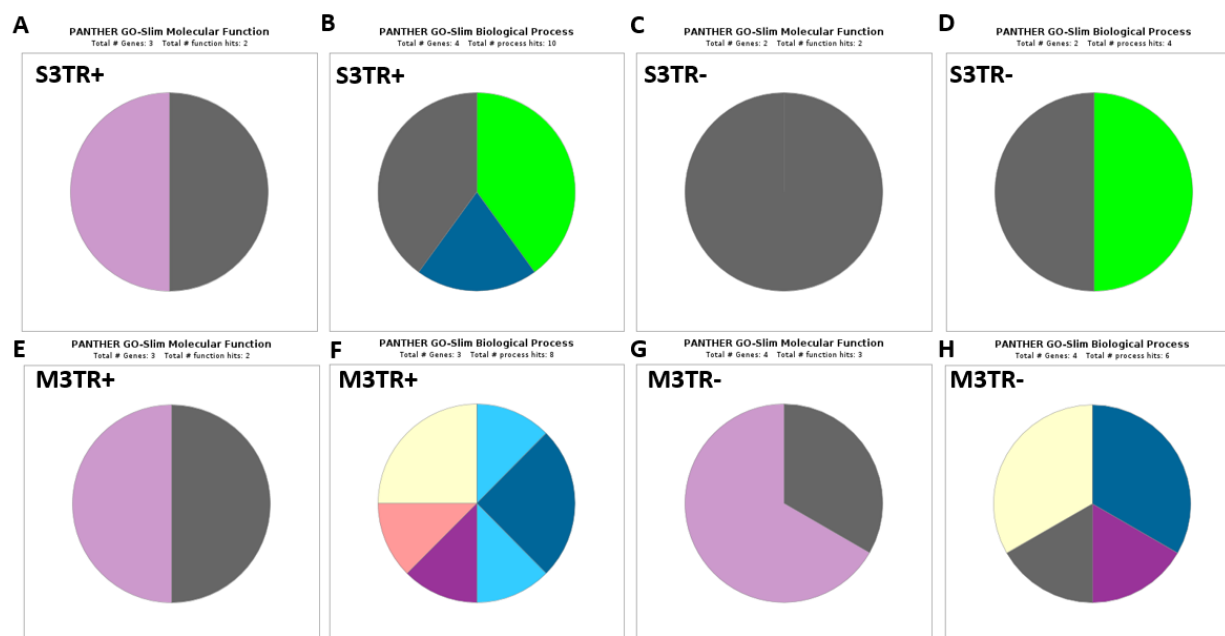
Figure 3.3: PANTHER Classification System of High-Hit Scrambled Excluded Immunoprecipitated Proteins [91]. Proteins are divided based on Molecular Function (MF) and Biological Processes (BP). Protein categories for S3TR+ is represented by (A) and (B) for MF and BP respectively. Protein categories for S3TR- is represented by (C) and (D) for MF and BP respectively. Protein categories for M3TR+ is represented by (E) and (F) for MF and BP respectively. Protein categories for M3TR- is represented by (G) and (H) for MF and BP respectively. MF is further divided based on Catalytic Activity (**Light Purple**) and Binding (**Dark Grey**). BP is further divided based on Biological Regulation (**Green**), Metabolic Process (**Dark Grey**), Cellular Process (**Navy Blue**), Biological Adhesion (**Light Blue**), Developmental Process (**Teal**), Localisation (**Dark Purple**), Multicellular Organismal Process (**Pink**), Response to Stimulus (**Sand**), and Signalling (**Dark Blue**).

The HHPs for sense S3TR all were involved in Binding interactions, with Poly(RC)-Binding Protein 1 and 2 both being involved in RNA metabolism activities which is defined as a protein that binds RNA and is involved in RNA processing or metabolism **(Figure 3.3A).** Both Poly(RC)-Binding Proteins 1 and 2, and Y-Box-Binding Proteins 1 and 3 are involved in biological regulation activities **(Figure 3.3B)**. The HHPs for antisense S3TR are comprise solely of Y-Box-Binding Protein 1 and 3, both of which are involved in regulating gene expression **(Figure 3.3C/D)**. HHPs for sense M3TR were involved in Binding and Catalytic Activity interactions **(Figure 3.3E)**. Heat Shock Cognate 71 kDa Protein was involved in localisation and response to stimuli processes **(Figure 3.3F)**. HHPs for antisense M3TR were involved in Binding and catalytic Activity interactions. Pyruvate Kinase is a kinase involved in catalysing the transfer of a phosphate from ATP to a second substrate and serves a function within metabolism **(Figure 3.3G/F)**. Heat Shock Cognate 71 kDa Protein's role is extended from sense M3TR. Between the two, Interferon-inducible Double-stranded RNA-dependent Protein Kinase Activator A and Double-stranded RNA-binding Protein Staufen Homolog 1 are not reflected in the Biological Process organisation.

3.3 Discussion

Noncoding RNAs (ncRNAs) are those RNA sequences that do not encode a protein but perform a variety of essential roles – that are generally highly pronounced in eukaryotic cells but are also present in viral genomes – and perform the regulation of transcription, translation, metabolic and enzymatic processes of other proteins, or provide infrastructural support for cellular activities [92]. The structures of these ncRNAs, which are predominantly localised at the terminal regions of ssRNA viruses, are essential in this regulation. For instance, the ssRNA viruses Enterovirus retains key secondary structures on its 5′-NTR which are important in the initiation of transcription and translation of viral elements [93]. More so, RNA replication for the highly pathogenic alphavirus, Venezuelan Equine Encephalitis Virus, is dictated by the cooperation of two structural RNA elements (a nonstable secondary structural promoter and a short RNA stem) in its 5′-untranslated terminal region that which when mutated, result in deleterious effects to its genomic replication [94, 95]. Many RBPs interact with the NTRs of ssRNA viruses to facilitate the initiation of viral replication or translation which is the case for Eukaryotic Translation Elongation Factors 1 Alpha (eEF1A) which interacts with three sites that include multiple Stem-loops on the conserved 3′-NTR of the *Flavivirus* family of viruses [96, 97]. As indicated, RNA-binding events through viral ncRNAs and host RBPs are important since these interactions may retain high specificity and are likely critical in viral lifecycles. Congruently, RBPs retain secondary and tertiary structural arrangements that enable them to bind to ncRNAs tightly as to effectuate biological processes. Many cellular processes rely on proteins assembled with modularised architecture, which enables RBPs to bind to RNA with increased specificity and affinity compared to individual domains acting alone [98]. Several motifs include the highly characterised and prevalent RNA-Recognition Motif (RRM), Zinc Fingers, or Double-Stranded RNA Binding Domains (dsRBD). Evidently, with the relationships that RNA viruses have with host and viral RBPs, cataloguing them is critical in mapping their interactions and establishing a foundation for future therapeutics. Since the secondary structure of Hantaviral NTRs has not been experimentally determined, identifying potential RBPs is a first step towards characterising likely proteins of importance, of which likely have the capacity to affect viral pathogenesis.

Identifying Host-Pathogen interactions itself is important in determining how viral systems infect and proliferate throughout human and animal hosts thereby creating a framework of infection in its totality. To study these interactions between the viral genome and the host RBPs, an RNA Immunoprecipitation (RIP) technique is employed and can also include Crosslinking and Immunoprecipitation (CLIP) to identify RNA substrates of RBPs [99]. The

RNA-RBP Pull-down assay is a RIP technique that utilises a target RNA sequence that is labelled with a linker molecule that can be bound to a crosslinked antibody and bead complex. The benefits of conducting RIP studies allows users to perform immunoprecipitations under physiologic conditions thus preserving the native complexes formed between RNAs and RBPs, and additionally requires little specialised equipment and reagents. Conducting RIPs in conjunction with CLIP studies enables the application of extensive washes to properly distinguish against strong and weak interactions. There are some limitations, however, RIP experiments require highly specialised antibodies for tight label binding or the use of tagged RBPs. They also need additional controls to distinguish against true interactions as opposed to non-specific or loosely associating binding. Finally, RIPs cannot identify the exact location of the RBPs binding sites to their respective RNAs which may require excess RNA truncations to specifically narrow the site of interaction [99].

Typical RIP Pull-down assays use high affinity epitope labels such as biotin which are bound to the target RNA and incubated with RBPs. Streptavidin agarose beads immobilise the RNA-Biotin complex and pull-down associating RBPs which can be subsequently identified using MS [100, 101]. Within this study, we utilise Digoxigenin (DIG, 0.390kDa) as a 3′-NTR nonradioactive label for purified RNAs which will be pulled-down using crosslinked anti-digoxigenin digoxin magnetic beads as opposed to biotin and streptavidin [49]. Here, the target NTRs are labelled at their 3′-terminus with DIG-labelled dideoxy-UTP and are catalysed using a transferase such as T4 RNA Ligase I [77]. The 3′-end labelling of RNA sequences is intended to promote appropriate RNA secondary structure formation and folding as alternative pull-down strategies of internal labelling can produce misfolding due to steric hindrance and subsequently result in synthetic or false RNA-RBP binding that would not normally occur in nature [102]. Once crosslinked with an immobilising magnetic bead matrix, the DIG-labelled NTRs can interact and bind to RBPs derived from the cancerous cells' lysate.

HCPS-causing Hantaviruses generally infect human pulmonary endothelial cells, Dendritic cells, and pulmonary epithelial cells [103, 104]. Incidentally, the application of either A549 epithelial cells or Human Pulmonary Microvascular Endothelial Cells (HPMEC) can be used to derive RBPs for the pull-down assays [105]. Following stringent washes, RNA-DIG-RBP complexes can be eluted and separated on SDS-PAGE gels with bands to be identified using Mass Spectrometry. Once RBPs identified using Mass Spectrometry, those with relatively high hits can be selected, expressed, and purified and subsequently validated using RNA interactions studies such as Microscale Thermophoresis (MST) and Electrophoretic Mobility Shift Assays (EMSA). The significance of Scrambled RNA as

a control is intended to provide randomLy ordered, nonspecific binding to RBPs that interact with oligonucleotides sequences generally and noncompetitively. Consequently, any length or sequence is acceptable, however, we adjusted to the specific length and randomLy arranged sequences of the sense and antisense S and M3TR NTRs. Scrambled oligomers compared to their Wildtype RNAs will fold generally, producing diverse ensembles whereas WT RNAs generate fewer ensembles whose structures are evolutionarily driven [106]. Scrambled RNAs are regularly employed as negative controls in binding studies, especially for siRNA and gene expression by Splicing Factors, where a highly structured secondary structure is necessary for specific interactions [107, 108].

The Hantaviral lifecycle is replete with several viral RBPs that engage with host machinery for its survival. Briefly, Hantavirus entry is mediated by the Gn/Gc spike forming Glycoproteins, which interacts with human Integrin receptors, with the *αVβ3*-Integrin being specific for SNV [29, 30, 109]. New-World Hantaviruses, also utilise Protocadherin-1 to enter cells [110]. Old-World Hantaviruses generally enter by a clathrin-dependent endocytosis, whereas New-World Hantaviruses like ANDV enter through clathrin-independent endocytosis, both of which rely on sodium proton exchangers and actin for virion internalisation [73, 111, 112]. Once virions enter the cell, they form an early to late endosome which migrates to the Endoplasmic Reticulum–Golgi Intermediate Compartment (ERGIC) where a low, internal pH triggers endosome disassembly, and subsequent viral RNP release in the cytoplasm near the ER [113]. Viral RNPs disassemble with the RdRp initiating replication and transcription through a cap snatching technique, that primes and realigns viral mRNA with a cleaved host 10-14 nt 5′-mRNA cap [38, 114]. N Protein subsequently binds to the 5′-caps and protect the nascent RNA from degradation from host machinery. The N Protein will also substitute as a Eukaryotic Initiation Factor 4F complex which forces host machinery to prefer the translation of Hantaviral mRNA over host mRNA [115, 116]. Hantaviral vRNA accumulates and will be packaged by the N Protein reforming the RNP which will eventually being coordinated by an RdRp at the panhandle regions [117, 118]. After replication and translation, virion components are trafficked from the ER to the plasma membrane, where SNV virions particularly bud [119]. With the N Protein and RdRp thus being the principle Hantaviral RBPs, the search for similar host RBPs that likely antagonise or promote vRNA's stability and operation is this study's primary intent.

Sense S3TR NTRs bound to four proteins: PCBP1, PCBP2, YBX1, and YBX3, four of which resulted in moderate MS scores, but were selected for their uniqueness from the Scrambled S3TR RNA and cytosolic localisation. Poly(RC)-Binding Proteins are part of the KH (hnRNP-k Homology) domain superfamily of nucleic acid binding proteins which are principally involved in mRNA stability processes [120]. The KH domain was identified in

Heterogeneous Nuclear Ribonucleoprotein K (hnRNP-K) which is a ubiquitous, domain that can bind to both ssRNA and ssDNA [121]. There are two types of PCBPs: hnRNP-K/J and αCP (α-complex proteins) proteins, both of which are present in the cytoplasm and nucleus. Viral mRNAs are also stabilised by αCPs, as both PCBP1 and 2 can interact with the 5′-NTR independent structures of Poliovirus: a cloverleaf and a more centrally located stem-loop [122, 123]. The 5′-NTR of Poliovirus RNA contains an Internal Ribosome Entry Site (IRES) for the cap-independent initiation of translation and is essential for virus survivability. Binding to PCBP1/2 increases Poliovirus mRNA to increase in stability whilst also coordinating the transition from translation to replication [124]. Similarly, PCBP2 interacts with several other 5′-NTR IRES elements in ssRNA viruses such as Hepatitis A and C, as well as Enterovirus, Coxsackievirus B3, with interactions targeting the cloverleaf structure [125-128]. Mutations made against PCBP2 affected Hepatitis Virus A translation, while removal of the 5′-NTR IRES element liberated Hepatitis A Virus' dependence on PCBP2 [126]. Consequently, having PCBP1/2 bind to sense S3TR NTRs likely indicates a role in its transcribed mRNA stability and preservation. Interestingly, the sense S3TR is the tail end of the S Segment, with both PCBP1/2 likely interacting with the 5′-end of the RNA sequence. Alternatively, capped mRNA regularly has specific binding sites limited to their 3′-UTRs which would corroborate potential interactions between PCBP1/2 and sense S3TR [120]. Being positive-sense viral mRNA, PCBP1/2's role as an mRNA stabiliser appears to apply, however, is distinctly absent in the other sense and antisense segment pull-downs. Nevertheless, the specificity of the two PCBPs likely indicates the presence of complex secondary structures at the S3TR NTR organised around four stem loops that include a cloverleaf.

Like sense S3TR, antisense S3TR also bound to YBX1 and YBX3, again with modest MS scores. Y-Box Binding Proteins are DNA/RNA-binding proteins that belong to a large family of proteins associated with the Cold Shock Domain (CSD) [129]. CSDs structurally contain five antiparallel β-strands which form a compact β-barrel, with β2/3-strands containing the RNA-Binding Motifs RNP1 and 2 [130]. YBX1 is the most studied RNA/DNA-binding Y-Box Protein, and is involved in cell differentiation, embryonical development, and stress response activities in both the cytoplasm and the nucleus. They are also involved in RNA-dependent processes that include mRNA packaging into messenger Ribonucleoproteins Particles (mRNP), mRNA translation, and mRNA stabilisation. YBX1 mRNA-binding is preferential, as RNA sequences with high GC contents are or high interaction [131]. Although replication for HIV occurs in the nucleus, YBX1 has been shown to be integral for its early and late-stage replication with the RNA chaperone stabilising newly synthesised vRNA and enhancing viral production through the binding of HIV's

Stem-loop 2 [132-134]. This is also consistent with HCV vRNA replication impairment because of YBX1 depletion. Although, this was in conjunction with HCV's Non-structural Proteins S3 and 4A which regulated the equilibrium between HCV RNA replication and viral particle production [135]. Conversely, YBX1 has been shown to repress Dengue Virus (DENV) translation through the association of YBX1 with its 3′-NTR which contains a highly conserved Stem Loop which is itself important for *Flavivirus* translation and replication [136]. Additionally, much like Hantaviruses, Influenza A Virus (IAV) is a negative sense ssRNA virus which is encapsidated by an RNP complex [137]. YBX3 has been shown to bind to the vRNP of IAV where it negatively regulates viral replication during the early stages of infection, although the mechanism is not yet known [138]. From these interactions, it can be postulated that Y-Box Binding Proteins present for both sense and antisense S3TR RNA are involved in antiviral processes, however, the role of mRNA stabilisation of specific viral mRNA and host mRNA is likely dependent on localisation and subsequent coordination from protein complexes as indicated by the protein-protein interactions throughout the **L4s** in **Figures 3.2A/B**. YBX1 appears to strongly interact with multiple transcription regulatory proteins while YBX3 is uniquely dissociative of other protein-protein interactions.

Sense M3TR NTRs broadly bound to different HHPs, including HSPA8, and STAU1 with moderate MS scores. STAU1 is a ubiquitously expressed protein whose role involves RNA localisation, splicing, stability, translation, and decay and thus contributes to cell proliferation, differentiation, migration, apoptosis, autophagy, and the cellular stress response [139]. STAU1 contains multiple dsRBDs, as well as a microtubule-binding domain and a STAU1-swapping motif of which contributes to RNA metabolism including transcription and degradation. Within the framework of viral infections, STAU1 has been shown to promote Enterovirus 71 (EV-A71) viral replication and translation, by having its RBD2-3 specifically bind to EV-A71's 5′-NTR which enables STAU1 to recruit more vRNA towards the ribosomal complexes which facilitates translation whilst also increasing viral mRNA stability [140]. Again, the M3TR is a noncoding RNA with a 3′-polarity, however, STAU1 is likely to interact with its 5′-end using its RBD2-3 *in vitro* as demonstrated. HCV replication and translation is similarly promoted by STAU1 which inactivates Protein Kinase R, an antiviral host defence protein [141]. Specifically, STAU1 interacts with the variable stem-Loop located on HCV's 3′-NTR including the 5′NTR's domain IIId which is part of HCV's IRES element. It is important to emphasise that many proteins appear to interact with cap-independent translation systems such as the IRES elements. This is not present in Hantaviruses because of the RdRp-mediated cap-snatching, prime and

realignment mechanism integral in Hantaviral translation which works in tandem with the N Protein to preferentially load ribosomes onto capped viral mRNA [142].

Antisense M3TR NTRs broadly bound to different HHPs, including HSPA8, and ELAVL1, PKM, and PRKRA with moderate to low MS scores. The HHP HSPA8 bound to both the sense and antisense M3TR NTR, which is a mostly cytosolic molecular chaperone involved in the maintenance of cellular protein regulation. Heat Shock Family proteins have two conserved structural domains: an N-terminal nucleotide-binding domain that hydrolyses ATP into ADP and a C-terminal substrate binding domain [143]. HSPA8 appears to be involved in viral entry and intracellular trafficking [144]. However, its associations with DENV internalisation are the result of the formation of protein complexes and not the interaction with NTRs [145]. Additionally, its role in trafficking the endosome is its interaction with clathrin and its subsequent disassembly [146]. The likely presence of the variety of HSPs is likely due to the induced cellular stressed caused by culturing and manipulation and subsequent harvesting through the scraping method. However, their consistent presence in the M3TR pull-downs as opposed to S3TR likely indicates a novel interaction that has yet to be substantiated.

ELAVL1 or Human Antigen R (HuR) is an RBP that is primarily located in the nucleus but can translocate to the cytoplasm after the exposure of a combination of intrinsic or extrinsic cellular stresses, where it then engages in mRNA stabilisation and translation [147]. Being mostly an RBP involved in cellular stress and cancer regulation, ELAVL1 has also been shown to interact with ssRNA viruses like HCV and HIV [147]. HCV's 3′-NTR is the site of assembly of its viral replication complex and has been shown to bind to ELAVL1 which helps HCV's replication and translation. This process is executed by ELAVL1 influencing the formation of a complex with other RBPs (La and Polypyrimidine Tract Binding Protein) that helps with HCV's circularisation and initiation of replication [148]. ELAVL1 does bind to the 3′-UTRs of mRNAs that exhibit rich AU elements, which are themselves targets of RBPs involved in mRNA decay [149]. This process helps to stabilise mRNAs and prevent their degradation. This is consistent since the antisense M3TR has a high ~73% AU-content.

Finally, there are PKM and PRKRA which are both cytosolic acting kinases. However, there are very few literature connections to PKM-NTR binding as it pertains to viral progression and lifecycle. PKM is a glycolysis protein with four isoforms involved in regulating cellular metabolism, specifically in that it catalyses the conversion of phosphoenolpyruvate and ADP to pyruvate and ATP [150]. PRKRA, conversely, is involved in antiviral and antiproliferative effects of Interferon during viral infections. PRKRA is activated by dsRNA through its dsRBD

located near its N-terminus, which results in the inhibition of viral and cellular protein synthesis [151]. This is elicited through binding of viral dsRNA that likely forms from dsRNA secondary structure from the NTR sequences. Confidence with PRKRA is dependent on its consistency in binding the two antisense M3TR replicates, although it retains the lowest score of the HHP. Additional pull-down must be carried out to fully confirm RNA-binding potential of all RBPs acquired during the antisense M3TR immunoprecipitation reactions.

3.4 Conclusions and Future Directions

The study only indicates potential RBP interactions with the Hantaviral S and M3TR NTRs and requires further validation studies to quantify and confirm the RNA-Protein interactions identified herein. These can include the Electrophoretic Mobility Shift Assays and subsequent Microscale Thermophoresis techniques to confirm RNA-Protein interaction, and to quantify the kinetics of the interactions and how strongly the NTRs and RBPs bind to each other. EMSAs will be necessary to visualise the binding of the identified RBPs with their viral NTRs under native conditions to assess the interaction qualitatively [69]. MST comparatively can determine quantitatively the binding affinities of the RNA-RBP interactions by observing the migration of RNA-RBP complexes along a temperature gradient [152]. After successfully validating RNA-RBP interactions, observing knockdown studies of these key RBPs in live virus conditions in A549 or HPMEC cell lines will be necessary in observing whether the absence of the RBPs promote or negate the Hantaviral lifecycle.

Other techniques would include Filter Binding to confirm the formation of RNA-Protein complexes, or Fluorescence Anisotropy which can measure the kinetics of RNA-Protein complexes based on their size and molecular weight changes that are affected by how the bound fluorophores tumble in solution [153]. RNA-Protein complexes can also be evaluated based on secondary structural properties of the RNA which contribute to binding interactions. Identifying structures through low- and high-resolution three-dimensional modelling like Small-Angle X-ray Scattering or Cryo-Electron Microscopy, can help evaluate potential binding domains that would play a significant role in SNV's relationship with regulatory or functional host proteins [154]. This can be combined with SHAPE studies on the SNV NTRs to determine their secondary structures in two-dimensional space, which can be modelled and overlapped using high-resolution, three-dimensional modelling packages such as SimRNA [57, 155, 156]. Auxiliary immunoprecipitation reactions would include conducting MANGO Aptamer Pull-down assays that minimise crosslinking and RNA-labelling steps and can acquire RNA-Protein interactions as they occur *in vivo* [50,

157-159]. The MANGO aptamer is an RNA tag that has nanomolar affinity to thiazole orange derivatives such as thiazole orange 1-Biotin and can increase their fluorescence ~1000-fold. Functioning as both an aptamer and fluorophore system, the MANGO aptamer can be added to the target RNA sequence of interest and can transfected and transcribed *in vivo* [158]. Transcribed RNA can interact with RBPs in A549 cells which can be immunoprecipitated directly after the cells are lysed using magnetic streptavidin beads like the Digoxigenin Pull-down assay presented in this study. Proteins identified through MS with this technique can supplement previous pull-down assay data and provide increased confidence using orthogonal techniques.

3.5 References

1. Mittler, E., et al., *Hantavirus entry: Perspectives and recent advances.* Adv Virus Res, 2019. **104**: p. 185-224.
2. Morrison, Y.Y. and R.C. Rathbun, *Hantavirus pulmonary syndrome: the Four Corners disease.* Ann Pharmacother, 1995. **29**(1): p. 57-65.
3. Nichol, S.T., et al., *Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness.* Science, 1993. **262**(5135): p. 914-7.
4. Zaki, S.R., et al., *Hantavirus pulmonary syndrome. Pathogenesis of an emerging infectious disease.* Am J Pathol, 1995. **146**(3): p. 552-79.
5. Borges, A.A., et al., *Hantavirus cardiopulmonary syndrome: immune response and pathogenesis.* Microbes Infect, 2006. **8**(8): p. 2324-30.
6. Khaiboullina, S.F., et al., *Serum Cytokine Profiles Differentiating Hemorrhagic Fever with Renal Syndrome and Hantavirus Pulmonary Syndrome.* Front Immunol, 2017. **8**: p. 567.
7. Watson, D.C., et al., *Epidemiology of Hantavirus infections in humans: a comprehensive, global overview.* Crit Rev Microbiol, 2014. **40**(3): p. 261-72.
8. Llah, S.T., et al., *Hantavirus induced cardiopulmonary syndrome: A public health concern.* J Med Virol, 2018. **90**(6): p. 1003-1009.
9. Netski, D., B.H. Thran, and S.C. St Jeor, *Sin Nombre virus pathogenesis in Peromyscus maniculatus.* J Virol, 1999. **73**(1): p. 585-91.
10. Warner, B.M., et al., *Development and Characterization of a Sin Nombre Virus Transmission Model in Peromyscus maniculatus.* Viruses, 2019. **11**(2).
11. Botten, J., et al., *Persistent Sin Nombre virus infection in the deer mouse (Peromyscus maniculatus) model: sites of replication and strand-specific expression.* J Virol, 2003. **77**(2): p. 1540-50.
12. Horne, K.M. and D.L. Vanlandingham, *Bunyavirus-vector interactions.* Viruses, 2014. **6**(11): p. 4373-97.
13. Bellomo, C., et al., *A newborn infected by Andes virus suggests novel routes of hantavirus transmission: a case report.* Clin Microbiol Infect, 2020. **26**(1): p. 130-131.
14. Martinez, V.P., et al., *Person-to-person transmission of Andes virus.* Emerg Infect Dis, 2005. **11**(12): p. 1848-53.
15. Forbes, K.M., T. Sironen, and A. Plyusnin, *Hantavirus maintenance and transmission in reservoir host populations.* Curr Opin Virol, 2018. **28**: p. 1-6.
16. Bagamian, K.H., et al., *Transmission ecology of Sin Nombre hantavirus in naturally infected North American deermouse populations in outdoor enclosures.* PLoS One, 2012. **7**(10): p. e47731.
17. Alonso, D.O., et al., *Person-to-Person Transmission of Andes Virus in Hantavirus Pulmonary Syndrome, Argentina, 2014.* Emerg Infect Dis, 2020. **26**(4): p. 756-759.
18. Elliott, R.M., *Bunyaviruses and climate change.* Clin Microbiol Infect, 2009. **15**(6): p. 510-7.
19. Gubler, D.J., et al., *Climate variability and change in the United States: potential impacts on vector- and rodent-borne diseases.* Environ Health Perspect, 2001. **109 Suppl 2**: p. 223-33.
20. Klempa, B., *Hantaviruses and climate change.* Clin Microbiol Infect, 2009. **15**(6): p. 518-23.

21.   Hansen, A., et al., *Transmission of haemorrhagic fever with renal syndrome in china and the role of climate factors: a review.* Int J Infect Dis, 2015. **33**: p. 212-8.
22.   Guterres, A. and E.R.S. de Lemos, *Hantaviruses and a neglected environmental determinant.* One Health, 2018. **5**: p. 27-33.
23.   Liu, R., et al., *Vaccines and Therapeutics Against Hantaviruses.* Front Microbiol, 2019. **10**: p. 2989.
24.   Khaiboullina, S.F., S.P. Morzunov, and S.C. St Jeor, *Hantaviruses: molecular biology, evolution and pathogenesis.* Curr Mol Med, 2005. **5**(8): p. 773-90.
25.   Chizhikov, V.E., et al., *Complete genetic characterization and analysis of isolation of Sin Nombre virus.* J Virol, 1995. **69**(12): p. 8132-6.
26.   Elliott, L.H., et al., *Isolation of the causative agent of hantavirus pulmonary syndrome.* Am J Trop Med Hyg, 1994. **51**(1): p. 102-8.
27.   Spiropoulou, C.F., et al., *Genome structure and variability of a virus causing hantavirus pulmonary syndrome.* Virology, 1994. **200**(2): p. 715-23.
28.   Serris, A., et al., *The Hantavirus Surface Glycoprotein Lattice and Its Fusion Control Mechanism.* Cell, 2020. **183**(2): p. 442-456.e16.
29.   Gavrilovskaya, I.N., et al., *beta3 Integrins mediate the cellular entry of hantaviruses that cause respiratory failure.* Proc Natl Acad Sci U S A, 1998. **95**(12): p. 7074-9.
30.   Gavrilovskaya, I.N., et al., *Cellular entry of hantaviruses which cause hemorrhagic fever with renal syndrome is mediated by beta3 integrins.* J Virol, 1999. **73**(5): p. 3951-9.
31.   Cifuentes-Munoz, N., N. Salazar-Quiroz, and N.D. Tischler, *Hantavirus Gn and Gc envelope glycoproteins: key structural units for virus cell entry and virus assembly.* Viruses, 2014. **6**(4): p. 1801-22.
32.   Kadry, Y.A. and D.A. Calderwood, *Chapter 22: Structural and signaling functions of integrins.* Biochim Biophys Acta Biomembr, 2020. **1862**(5): p. 183206.
33.   Switala-Jelen, K., et al., *The biological functions of beta3 integrins.* Folia Biol (Praha), 2004. **50**(5): p. 143-52.
34.   Wichgers Schreur, P.J., R. Kormelink, and J. Kortekaas, *Genome packaging of the Bunyavirales.* Curr Opin Virol, 2018. **33**: p. 151-155.
35.   Mir, M.A. and A.T. Panganiban, *The bunyavirus nucleocapsid protein is an RNA chaperone: possible roles in viral RNA panhandle formation and genome replication.* Rna, 2006. **12**(2): p. 272-82.
36.   Mir, M.A. and A.T. Panganiban, *Characterization of the RNA chaperone activity of hantavirus nucleocapsid protein.* J Virol, 2006. **80**(13): p. 6276-85.
37.   Brown, B.A. and A.T. Panganiban, *Identification of a region of hantavirus nucleocapsid protein required for RNA chaperone activity.* RNA Biol, 2010. **7**(6): p. 830-7.
38.   Cheng, E., Z. Wang, and M.A. Mir, *Interaction between hantavirus nucleocapsid protein (N) and RNA-dependent RNA polymerase (RdRp) mutants reveals the requirement of an N-RdRp interaction for viral RNA synthesis.* J Virol, 2014. **88**(15): p. 8706-12.
39.   Mir, M.A., et al., *Hantavirus N protein exhibits genus-specific recognition of the viral RNA panhandle.* J Virol, 2006. **80**(22): p. 11283-92.
40.   Mir, M.A. and A.T. Panganiban, *The hantavirus nucleocapsid protein recognizes specific features of the viral RNA panhandle and is altered in conformation upon RNA binding.* J Virol, 2005. **79**(3): p. 1824-35.
41.   Mir, M.A. and A.T. Panganiban, *Trimeric hantavirus nucleocapsid protein binds specifically to the viral RNA panhandle.* J Virol, 2004. **78**(15): p. 8281-8.
42.   Mir, M.A. and A.T. Panganiban, *The triplet repeats of the Sin Nombre hantavirus 5' untranslated region are sufficient in cis for nucleocapsid-mediated translation initiation.* J Virol, 2010. **84**(17): p. 8937-44.
43.   Kesheh, M.M., S. Mahmoudvand, and S. Shokri, *Long noncoding RNAs in respiratory viruses: A review.* Rev Med Virol, 2021: p. e2275.
44.   Yi, K., et al., *Long noncoding RNA and its role in virus infection and pathogenesis.* Front Biosci (Landmark Ed), 2019. **24**: p. 777-789.
45.   Shrestha, N. and J.J. Bujarski, *Long Noncoding RNAs in Plant Viroids and Viruses: A Review.* Pathogens, 2020. **9**(9).
46.   Barr, J.N. and G.W. Wertz, *Bunyamwera bunyavirus RNA synthesis requires cooperation of 3'- and 5'-terminal sequences.* J Virol, 2004. **78**(3): p. 1129-38.
47.   Vera-Otarola, J., et al., *The 3' untranslated region of the Andes hantavirus small mRNA functionally replaces the poly(A) tail and stimulates cap-dependent translation initiation from the viral mRNA.* Journal of virology, 2010. **84**(19): p. 10420-10424.

48.   Meier, K., et al., *Hantavirus Replication Cycle-An Updated Structural Virology Perspective.* Viruses, 2021. **13**(8).

49.   McCreery, T., *Digoxigenin labeling.* Mol Biotechnol, 1997. **7**(2): p. 121-4.

50.   Gemmill, D., et al., *Current approaches for RNA-labelling to identify RNA-binding proteins.* Biochem Cell Biol, 2020. **98**(1): p. 31-41.

51.   Rennie, W., et al., *Sfold Tools for MicroRNA Target Prediction.* Methods Mol Biol, 2019. **1970**: p. 31-42.

52.   Ding, Y., C.Y. Chan, and C.E. Lawrence, *RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble.* Rna, 2005. **11**(8): p. 1157-66.

53.   Ding, Y., C.Y. Chan, and C.E. Lawrence, *Sfold web server for statistical folding and rational design of nucleic acids.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W135-41.

54.   Ding, Y. and C.E. Lawrence, *A statistical sampling algorithm for RNA secondary structure prediction.* Nucleic Acids Res, 2003. **31**(24): p. 7280-301.

55.   Ding, Y. and C.E. Lawrence, *Statistical prediction of single-stranded regions in RNA secondary structure and application to predicting effective antisense target sites and beyond.* Nucleic Acids Res, 2001. **29**(5): p. 1034-46.

56.   Hofacker, I.L., et al., *Fast folding and comparison of RNA secondary structures.* Monatshefte für Chemie / Chemical Monthly, 1994. **125**(2): p. 167-188.

57.   Lorenz, R., et al., *ViennaRNA Package 2.0.* Algorithms for molecular biology : AMB, 2011. **6**: p. 26-26.

58.   Stothard, P., *The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences.* Biotechniques 2000. **28**: p. 1102-1104.

59.   Sievers, F. and D.G. Higgins, *Clustal Omega for making accurate alignments of many protein sequences.* Protein Science, 2018. **27**(1): p. 135-145.

60.   Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Molecular Systems Biology, 2011. **7**(1): p. 539.

61.   Wilkinson, K.A., E.J. Merino, and K.M. Weeks, *Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution.* Nat Protoc, 2006. **1**(3): p. 1610-6.

62.   Lusvarghi, S., et al., *RNA secondary structure prediction using high-throughput SHAPE.* J Vis Exp, 2013(75): p. e50243.

63.   Smola, M.J., et al., *Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis.* Nat Protoc, 2015. **10**(11): p. 1643-69.

64.   Beckert, B. and B. Masquida, *Synthesis of RNA by in vitro transcription.* Methods Mol Biol, 2011. **703**: p. 29-41.

65.   Chen, Z. and Y. Zhang, *Dimethyl sulfoxide targets phage RNA polymerases to promote transcription.* Biochem Biophys Res Commun, 2005. **333**(3): p. 664-70.

66.   McKenna, S.A., et al., *Purification and characterization of transcribed RNAs using gel filtration chromatography.* Nat Protoc, 2007. **2**(12): p. 3270-7.

67.   Kim, I., et al., *Rapid purification of RNAs using fast performance liquid chromatography (FPLC).* RNA (New York, N.Y.), 2007. **13**(2): p. 289-294.

68.   Ream, J.A., L.K. Lewis, and K.A. Lewis, *Rapid agarose gel electrophoretic mobility shift assay for quantitating protein: RNA interactions.* Anal Biochem, 2016. **511**: p. 36-41.

69.   Hellman, L.M. and M.G. Fried, *Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions.* Nat Protoc, 2007. **2**(8): p. 1849-61.

70.   Fisher, J.S., et al., *Preparation of a denaturing agarose gel for the analysis of RNA samples.* CSH Protoc, 2006. **2006**(1).

71.   Walker, S.E. and J. Lorsch, *RNA purification--precipitation methods.* Methods Enzymol, 2013. **530**: p. 337-43.

72.   Ermonval, M., F. Baychelier, and N. Tordo, *What Do We Know about How Hantaviruses Interact with Their Different Hosts?* Viruses, 2016. **8**(8).

73.   Torriani, G., et al., *Macropinocytosis contributes to hantavirus entry into human airway epithelial cells.* Virology, 2019. **531**: p. 57-68.

74.   Ren, H., N.P. Birch, and V. Suresh, *An Optimised Human Cell Culture Model for Alveolar Epithelial Transport.* PLOS ONE, 2016. **11**(10): p. e0165225.

75.   Kawahara, I., et al., *Site-specific isotope labeling of long RNA for structural and mechanistic studies.* Nucleic Acids Res, 2012. **40**(1): p. e7.

76.   Liu, Y., R. Sousa, and Y.X. Wang, *Specific labeling: An effective tool to explore the RNA world.* Bioessays, 2016. **38**(2): p. 192-200.

77.   Tomassi, A.H., et al., *Nonradioactive Detection of Small RNAs Using Digoxigenin-Labeled Probes.* Methods Mol Biol, 2017. **1640**: p. 199-210.

78.   Candiano, G., et al., *Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis.* Electrophoresis, 2004. **25**(9): p. 1327-33.

79.   Adaway, J.E., B.G. Keevil, and L.J. Owen, *Liquid chromatography tandem mass spectrometry in the clinical laboratory.* Ann Clin Biochem, 2015. **52**(Pt 1): p. 18-38.

80.   Brymora, A., V.A. Valova, and P.J. Robinson, *Protein-protein interactions identified by pull-down experiments and mass spectrometry.* Curr Protoc Cell Biol, 2004. **Chapter 17**: p. Unit 17.5.

81.   Michalski, A., et al., *Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer.* Molecular & cellular proteomics : MCP, 2011. **10**(9): p. M111.011015-M111.011015.

82.   Snel, B., et al., *STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene.* Nucleic Acids Res, 2000. **28**(18): p. 3442-4.

83.   von Mering, C., et al., *STRING: a database of predicted functional associations between proteins.* Nucleic Acids Res, 2003. **31**(1): p. 258-61.

84.   von Mering, C., et al., *STRING: known and predicted protein-protein associations, integrated and transferred across organisms.* Nucleic Acids Res, 2005. **33**(Database issue): p. D433-7.

85.   Franceschini, A., et al., *SVD-phy: improved prediction of protein functional associations through singular value decomposition of phylogenetic profiles.* Bioinformatics, 2016. **32**(7): p. 1085-7.

86.   Szklarczyk, D., et al., *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.* Nucleic Acids Res, 2019. **47**(D1): p. D607-d613.

87.   Szklarczyk, D., et al., *The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets.* Nucleic Acids Res, 2021. **49**(D1): p. D605-d612.

88.   Thomas, P.D., et al., *PANTHER: a library of protein families and subfamilies indexed by function.* Genome Res, 2003. **13**(9): p. 2129-41.

89.   Mi, H., A. Muruganujan, and P.D. Thomas, *PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees.* Nucleic Acids Res, 2013. **41**(Database issue): p. D377-86.

90.   Hodge, K., et al., *Cleaning up the masses: exclusion lists to reduce contamination with HPLC-MS/MS.* Journal of proteomics, 2013. **88**: p. 92-103.

91.   Mi, H., et al., *PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API.* Nucleic Acids Res, 2021. **49**(D1): p. D394-d403.

92.   Mattick, J.S. and I.V. Makunin, *Non-coding RNA.* Hum Mol Genet, 2006. **15 Spec No 1**: p. R17-29.

93.   Glenet, M., et al., *Structures and Functions of Viral 5' Non-Coding Genomic RNA Domain-I in Group-B Enterovirus Infections.* Viruses, 2020. **12**(9).

94.   Hyde, J.L., et al., *The 5' and 3' ends of alphavirus RNAs--Non-coding is not non-functional.* Virus Res, 2015. **206**: p. 99-107.

95.   Kulasegaran-Shylini, R., et al., *The 5'UTR-specific mutation in VEEV TC-83 genome has a strong effect on RNA replication and subgenomic RNA synthesis, but not on translation of the encoded proteins.* Virology, 2009. **387**(1): p. 211-21.

96.   Brinton, M.A. and M. Basu, *Functions of the 3' and 5' genome RNA regions of members of the genus Flavivirus.* Virus Res, 2015. **206**: p. 108-19.

97.   Bidet, K. and M.A. Garcia-Blanco, *Flaviviral RNAs: weapons and targets in the war between virus and host.* Biochem J, 2014. **462**(2): p. 215-30.

98.   Lunde, B.M., C. Moore, and G. Varani, *RNA-binding proteins: modular design for efficient function.* Nat Rev Mol Cell Biol, 2007. **8**(6): p. 479-90.

99.   Barra, J. and E. Leucci, *Probing Long Non-coding RNA-Protein Interactions.* Front Mol Biosci, 2017. **4**: p. 45.

100.   Popova, V.V., M.M. Kurshakova, and D.V. Kopytova, *Methods to study the RNA-protein interactions.* Mol Biol (Mosk), 2015. **49**(3): p. 472-81.

101.   Rinaldi, A.J., K.C. Suddala, and N.G. Walter, *Native purification and labeling of RNA for single molecule fluorescence studies.* Methods Mol Biol, 2015. **1240**: p. 63-95.

102.   Jazurek, M., et al., *Identifying proteins that bind to specific RNAs - focus on simple repeat expansion diseases.* Nucleic Acids Res, 2016. **44**(19): p. 9050-9070.

103.   Raftery, M.J., et al., *Hantavirus infection of dendritic cells.* Journal of virology, 2002. **76**(21): p. 10724-10733.

104.   Wahl-Jensen, V., et al., *Temporal Analysis of Andes Virus and Sin Nombre Virus Infections of Syrian Hamsters.* Journal of Virology, 2007. **81**(14): p. 7449.

105.   Krump-Konvalinkova, V., et al., *Generation of human pulmonary microvascular endothelial cell lines.* Lab Invest, 2001. **81**(12): p. 1717-27.

106.   Moss, W.N., *The ensemble diversity of non-coding RNA structure is lower than random sequence.* Noncoding RNA Res, 2018. **3**(3): p. 100-107.

107.   Li, H., et al., *Identification of mRNA binding proteins that regulate the stability of LDL receptor mRNA through AU-rich elements.* J Lipid Res, 2009. **50**(5): p. 820-31.

108.   Denichenko, P., et al., *Specific inhibition of splicing factor activity by decoy RNA oligonucleotides.* Nature Communications, 2019. **10**(1): p. 1590.

109.   Guardado-Calvo, P., et al., *Mechanistic Insight into Bunyavirus-Induced Membrane Fusion from Structure-Function Analyses of the Hantavirus Envelope Glycoprotein Gc.* PLoS Pathog, 2016. **12**(10): p. e1005813.

110.   Jangra, R.K., et al., *Protocadherin-1 is essential for cell entry by New World hantaviruses.* Nature, 2018. **563**(7732): p. 559-563.

111.   Chiang, C.F., et al., *Endocytic Pathways Used by Andes Virus to Enter Primary Human Lung Endothelial Cells.* PLoS One, 2016. **11**(10): p. e0164768.

112.   Jin, M., et al., *Hantaan virus enters cells by clathrin-dependent receptor-mediated endocytosis.* Virology, 2002. **294**(1): p. 60-9.

113.   Muyangwa, M., et al., *Hantaviral Proteins: Structure, Functions, and Role in Hantavirus Infection.* Front Microbiol, 2015. **6**: p. 1326.

114.   Jeeva, S., et al., *Hantavirus RdRp Requires a Host Cell Factor for Cap Snatching.* J Virol, 2019. **93**(5).

115.   Mir, M.A. and A.T. Panganiban, *A protein that replaces the entire cellular eIF4F complex.* Embo j, 2008. **27**(23): p. 3129-39.

116.   Hussein, I.T. and M.A. Mir, *How hantaviruses modulate cellular pathways for efficient replication?* Front Biosci (Elite Ed), 2013. **5**: p. 154-66.

117.   Arragain, B., et al., *High resolution cryo-EM structure of the helical RNA-bound Hantaan virus nucleocapsid reveals its assembly mechanisms.* Elife, 2019. **8**.

118.   Reuter, M. and D.H. Krüger, *The nucleocapsid protein of hantaviruses: much more than a genome-wrapping protein.* Virus Genes, 2018. **54**(1): p. 5-16.

119.   Goldsmith, C.S., et al., *Ultrastructural characteristics of Sin Nombre virus, causative agent of hantavirus pulmonary syndrome.* Arch Virol, 1995. **140**(12): p. 2107-22.

120.   Makeyev, A.V. and S.A. Liebhaber, *The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms.* Rna, 2002. **8**(3): p. 265-78.

121.   Backe, P.H., et al., *X-ray crystallographic and NMR studies of the third KH domain of hnRNP K in complex with single-stranded nucleic acids.* Structure, 2005. **13**(7): p. 1055-67.

122.   Blyn, L.B., et al., *Poly(rC) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5' noncoding region: identification by automated liquid chromatography-tandem mass spectrometry.* Proc Natl Acad Sci U S A, 1996. **93**(20): p. 11115-20.

123.   Murray, K.E., A.W. Roberts, and D.J. Barton, *Poly(rC) binding proteins mediate poliovirus mRNA stability.* Rna, 2001. **7**(8): p. 1126-41.

124.   Kempf, B.J. and D.J. Barton, *Poly(rC) binding proteins and the 5' cloverleaf of uncapped poliovirus mRNA function during de novo assembly of polysomes.* J Virol, 2008. **82**(12): p. 5835-46.

125.   Fukushi, S., et al., *Interaction of poly(rC)-binding protein 2 with the 5'-terminal stem loop of the hepatitis C-virus genome.* Virus Res, 2001. **73**(1): p. 67-79.

126.   Graff, J., et al., *Interaction of poly(rC) binding protein 2 with the 5' noncoding region of hepatitis A virus RNA and its effects on translation.* J Virol, 1998. **72**(12): p. 9668-75.

127.   Zell, R., et al., *Poly(rC)-binding protein 2 interacts with the oligo(rC) tract of coxsackievirus B3.* Biochem Biophys Res Commun, 2008. **366**(4): p. 917-21.

128.   Zell, R., et al., *Interaction of poly(rC)-binding protein 2 domains KH1 and KH3 with coxsackievirus RNA.* Biochem Biophys Res Commun, 2008. **377**(2): p. 500-503.

129.   Mordovkina, D., et al., *Y-Box Binding Proteins in mRNP Assembly, Translation, and Stability Control.* Biomolecules, 2020. **10**(4).

130. Budkina, K.S., et al., *Cold Shock Domain Proteins: Structure and Interaction with Nucleic Acids.* Biochemistry (Mosc), 2020. **85**(Suppl 1): p. S1-s19.

131. Dong, J., et al., *RNA-binding specificity of Y-box protein 1.* RNA Biol, 2009. **6**(1): p. 59-64.

132. Weydert, C., et al., *Correction: Y-box-binding protein 1 supports the early and late steps of HIV replication.* PLoS One, 2020. **15**(10): p. e0241893.

133. Weydert, C., et al., *Y-box-binding protein 1 supports the early and late steps of HIV replication.* PLoS One, 2018. **13**(7): p. e0200080.

134. Mu, X., et al., *YB-1 stabilizes HIV-1 genomic RNA and enhances viral production.* Protein Cell, 2013. **4**(8): p. 591-7.

135. Chatel-Chaix, L., et al., *Y-box-binding protein 1 interacts with hepatitis C virus NS3/4A and influences the equilibrium between viral RNA replication and infectious particle production.* J Virol, 2011. **85**(21): p. 11022-37.

136. Paranjape, S.M. and E. Harris, *Y box-binding protein-1 binds to the dengue virus 3'-untranslated region and mediates antiviral effects.* J Biol Chem, 2007. **282**(42): p. 30497-508.

137. Eisfeld, A.J., G. Neumann, and Y. Kawaoka, *At the centre: influenza A virus ribonucleoproteins.* Nat Rev Microbiol, 2015. **13**(1): p. 28-41.

138. Qin, Z., et al., *Y-Box-Binding Protein 3 (YBX3) Restricts Influenza A Virus by Interacting with Viral Ribonucleoprotein Complex and Imparing its Function.* J Gen Virol, 2020. **101**(4): p. 385-398.

139. Almasi, S. and B.J. Jasmin, *The multifunctional RNA-binding protein Staufen1: an emerging regulator of oncogenesis through its various roles in key cellular events.* Cell Mol Life Sci, 2021.

140. Chen, Y.-M., et al., *Staufen1 Protein Participates Positively in the Viral RNA Replication of Enterovirus 71.* Viruses, 2019. **11**(2): p. 142.

141. Dixit, U., et al., *Staufen1 promotes HCV replication by inhibiting protein kinase R and transporting viral RNA to the site of translation and replication in the cells.* Nucleic Acids Res, 2016. **44**(11): p. 5271-87.

142. Hussein, I.T., et al., *Recent advances in hantavirus molecular biology and disease.* Adv Appl Microbiol, 2011. **74**: p. 35-75.

143. Silva, N.S.M., et al., *Structural, thermodynamic and functional studies of human 71 kDa heat shock cognate protein (HSPA8/hHsc70).* Biochim Biophys Acta Proteins Proteom, 2021. **1869**(12): p. 140719.

144. Wang, Z., et al., *Mechanism and Complex Roles of HSC70 in Viral Infections.* Front Microbiol, 2020. **11**: p. 1577.

145. Vega-Almeida, T.O., et al., *Surface proteins of C6/36 cells involved in dengue virus 4 binding and entry.* Arch Virol, 2013. **158**(6): p. 1189-207.

146. Newmyer, S.L. and S.L. Schmid, *Dominant-interfering Hsc70 mutants disrupt multiple stages of the clathrin-coated vesicle cycle in vivo.* J Cell Biol, 2001. **152**(3): p. 607-20.

147. Schultz, C.W., et al., *Understanding and targeting the disease-related RNA binding protein human antigen R (HuR).* Wiley Interdiscip Rev RNA, 2020. **11**(3): p. e1581.

148. Shwetha, S., et al., *HuR Displaces Polypyrimidine Tract Binding Protein To Facilitate La Binding to the 3' Untranslated Region and Enhances Hepatitis C Virus Replication.* J Virol, 2015. **89**(22): p. 11356-71.

149. Jehung, J.P., et al., *Adenovirus infection induces HuR relocalization to facilitate virus replication.* Biochem Biophys Res Commun, 2018. **495**(2): p. 1795-1800.

150. Israelsen, W.J. and M.G. Vander Heiden, *Pyruvate kinase: Function, regulation and role in cancer.* Semin Cell Dev Biol, 2015. **43**: p. 43-51.

151. Patel, R.C. and G.C. Sen, *PACT, a protein activator of the interferon-induced protein kinase, PKR.* Embo j, 1998. **17**(15): p. 4379-90.

152. Moon, M.H., et al., *Measuring RNA-Ligand Interactions with Microscale Thermophoresis.* Biochemistry, 2018. **57**(31): p. 4638-4643.

153. Licatalosi, D.D., X. Ye, and E. Jankowsky, *Approaches for measuring the dynamics of RNA-protein interactions.* Wiley interdisciplinary reviews. RNA, 2020. **11**(1): p. e1565-e1565.

154. Théobald-Dietrich, A., et al., *Structural Analysis of RNA by Small-Angle X-ray Scattering.* Methods Mol Biol, 2020. **2113**: p. 189-215.

155. Boniecki, M.J., et al., *SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction.* Nucleic Acids Res, 2016. **44**(7): p. e63.

156. Lorenz, R., et al., *Predicting RNA secondary structures from sequence and probing data.* Methods, 2016. **103**: p. 86-98.

157. Panchapakesan, S.S., S.C. Jeng, and P.J. Unrau, *RNA complex purification using high-affinity fluorescent RNA aptamer tags.* Ann N Y Acad Sci, 2015. **1341**: p. 149-55.

158. Panchapakesan, S.S.S., et al., *Ribonucleoprotein purification and characterization using RNA Mango.* Rna, 2017. **23**(10): p. 1592-1599.

159. Dolgosheina, E.V., et al., *RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking.* ACS Chem Biol, 2014. **9**(10): p. 2412-20.

Chapter 4: Biophysical Characterisation of Human LincRNA-p21 Sense and Antisense Alu Inverted Repeats[1]

4.0 Introduction

The tumour suppressor protein p53 is an important transcription factor that regulates a variety of cellular processes, including cell-cycle control, DNA repair, apoptosis, senescence, and cellular stress responses through the activation and repression of target genes [1, 2]. Despite playing a critical role in the DNA damage response, p53's genome is frequently mutated in cancer cells, exposing a vulnerability in cell cycle regulation [3, 4]. Nevertheless, when DNA damage occurs, p53 upregulates the expression of genes involved in the cell cycle arrest and DNA repair processes, which leads to cell survival, but also facilitates the initiation of apoptosis for cancerous cells [5]. Regulation of p53 is generally achieved through post-transcriptional modification, which can include its phosphorylation, acetylation, methylation, ubiquitination, or SUMOylation with the result of different types of DNA damage affecting p53's activation [6, 7]. The p53 pathway itself is composed of a network of genes, regulatory proteins, and their transcriptional products which can help respond to intrinsic and extrinsic stress signals [8]. This can include MDM2 which interacts strongly with p53 and helps regulate its activity through post-transcriptional modification. MDM2 degrades p53 through the ubiquitin proteasome pathway and prevents the transcription of tumour suppressor genes leading to apoptosis [9-11]. Other proteins like the acetyltransferase P300 can also enhance the activity of p53 [12]. These networks enable the regulation of p53 and can be additionally modulated by long noncoding RNAs (LncRNAs) which have been shown to act in a regulatory role within the p53 pathway. Often, the transcription of LncRNA genes are the targets of p53 itself [13-16].

LncRNAs are noncoding RNA molecules devoid of an open reading frame and are generally around 200-100,000 nucleotides (nts). They also do not retain any significant protein-coding capabilities and are therefore generally not expressed [17-19]. LncRNAs were previously thought to have no biological function but have been identified to regulate biological processes by altering gene expression and signal pathways [17]. Consequently, LncRNAs play a role in the regulation of gene expression and appear poised to affect the progression of cancers. Long intergenic noncoding RNA-p21 (LincRNA-p21) is found to be a transcriptional repressor in the p53 pathway, playing a role in triggering cellular apoptosis [20]. LincRNAs are also capped, spliced, and polyadenylated due to being RNA

---

[1] Currently under review in Nucleic Acid Research. Differences are unique to this publication.

polymerase II transcripts [18]. Under stress conditions including DNA damage, p53 activates the transcription of LincRNA-p21 which accumulates in the nucleus and associates with the heterogeneous nuclear ribonuclear protein K (hnRNP-K) [21]. The hnRNP-K contains RNA recognition motifs Arg-Gly-Gly repeats or hnRNP-K homology (KH) domains and whose role is important for nucleic acid metabolism and transcription [22, 23]. The hnRNP-K is integral in the induction of apoptosis since it will combine with the p53 promoted and transcribed LincRNA-p21 which will then act to repress p53 target genes resulting in apoptosis [24]. LincRNA-p21 is thus required to help direct hnRNP-K to bind to the promoters of the target repressed genes [23]. Additionally, hnRNP-K was observed to be a transcriptional coactivator of p53, enabling gene expression in response to DNA damage [22].

An important element identified in the LincRNA-p21 gene is the presence of two isoforms that contain Alu repeats, which influence the function of the RNA [21]. Alu elements are particularly important because they are highly conserved among primates and fold to produce independent domains. These repeated DNA sequences comprise upwards of 60% of the human genome and can be divided into several classes including micro-satellites (repeat sequences greater than 7 bp), mini satellites (basic repeats of 7 bp or less), or telomeres. These interspersed repeated DNA sequences are further divided into two classes: Short interspersed elements (SINES), and long interspersed elements (LINES) [25]. Alu SINES themselves are repetitive elements present in multiple copies of the genomes they reside in and are named because the family of repeats contains a recognition site for the restriction enzyme *AluI* [26, 27]. Full-length Alu elements are roughly 300 bp long and are frequently located in the 3′-untranslated regions of genes and their intergenic genomic regions and continue to be the most abundant mobile or transposable element in the entirety of the human genome. Alu elements are also important because they maintain an impact on the human genome's mutability [26]. Alu elements can influence insertion mutations, recombination between elements, gene conversion, and gene expression, and can ultimately cause a variety of human diseases including neurofibromatosis, haemophilia, familial hypercholesterolaemia, breast cancer, insulin-resistant diabetes type II, and Ewing sarcoma [28-30]. Determining the structural-dependent role of LincRNA-p21 Alu elements will have an impact on elucidating their overall function and responsibilities within the cell.

Many studies using molecular and computational structural biology seek to identify LncRNA secondary and tertiary structures, and whether said structures have an impact on their function [31, 32]. This also includes the application of RNA secondary structure prediction techniques such as selective 2′-hydroxyl acylation analysed by primer extension (SHAPE) [33]. Doing so is important because it conceptualises structures present on LincRNA-p21

and can elucidate potential specific interactions within the p53 and hnRNP-K pathways. Previous studies have investigated the secondary structure of LincRNA-p21 Alu Inverted Repeats (IRs) and identifying important functional regions that are involved in LincRNA-p21 nuclear localisation and its subsequent transcriptional factor interactions [21]. They identified that the two isoforms of LincRNA-p21 Alu IRs retained integral secondary structures that can fold into independent domains. These structures were suggested to be conserved in primates and contribute towards the regulation of cellular localisation of LincRNA-p21 during the cellular stress response.

Incidentally, the intent of this study is to investigate the overall structure of the sense and antisense LincRNA-p21 *AluSx1* Inverted Repeats by employing small angle X-ray scattering (SAXS) and computational modelling to develop their three-dimensional structures [34]. We have employed analytical ultracentrifugation (AUC) and size exclusion chromatography coupled to multiangle light scattering (SEC-MALS) instruments to biophysically characterise transcribed and purified *AluSx1* RNAs. AUC experiments revealed that *AluSx1* RNAs were present as monomeric, full-length transcripts under denaturing conditions, while SEC-MALS characterised their Molecular Weight (M_W). By combining chemically probed secondary structure information proposed by Chillón and Pyle, 2016, and SimRNA computational modelling, several three-dimensional, high-resolution models can be calculated and fitted to SAXS determined structures. We determined that LincRNA-p21 *AluSx1* RNA adopts an asymmetrical, and extended structure in solution. We describe a workflow utilising SAXS and SimRNA computational modelling to produce three-dimensional, high-resolution models devised from two-dimensional structures determined *via* SHAPE and other secondary structure probing techniques.

4.1 Methods

1. RNA Preparation and Biophysical Studies

2. Low-resolution Structural Studies using SAXS

3. High-resolution Structure Calculation using Computational tools and Integration of SAXS data

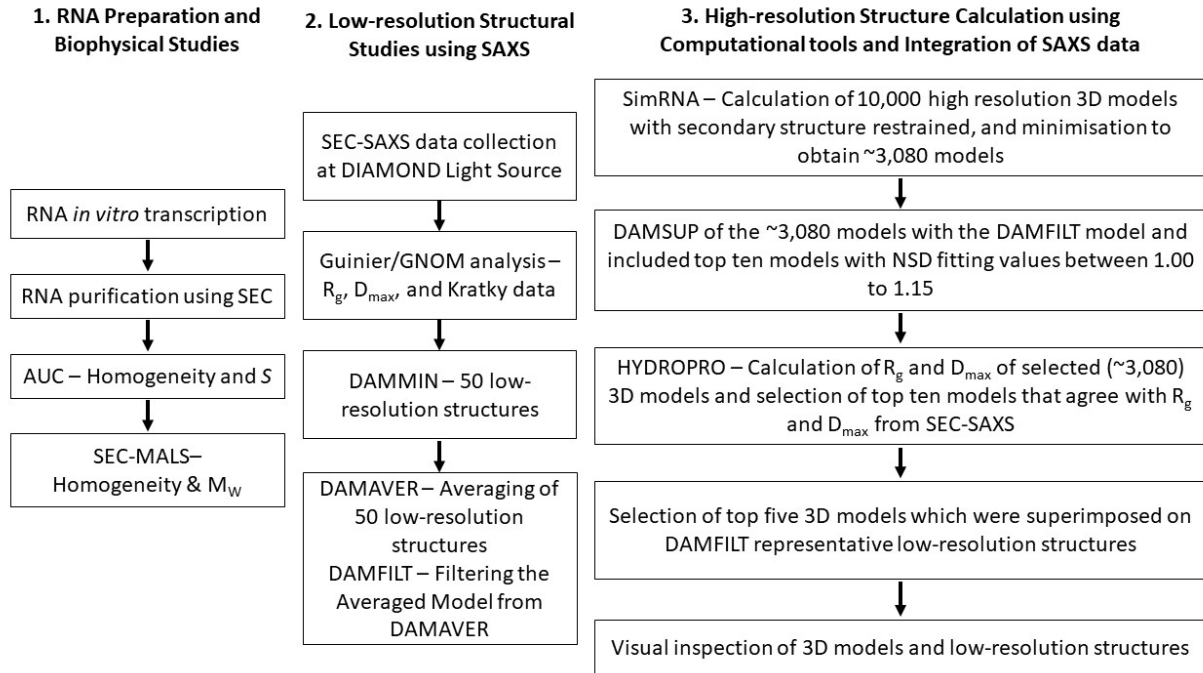|		
RNA *in vitro* transcription	SEC-SAXS data collection at DIAMOND Light Source	SimRNA – Calculation of 10,000 high resolution 3D models with secondary structure restrained, and minimisation to obtain ~3,080 models
RNA purification using SEC	Guinier/GNOM analysis– R_g, D_{max}, and Kratky data	DAMSUP of the ~3,080 models with the DAMFILT model and included top ten models with NSD fitting values between 1.00 to 1.15
AUC – Homogeneity and *S*	DAMMIN – 50 low-resolution structures	HYDROPRO – Calculation of R_g and D_{max} of selected (~3,080) 3D models and selection of top ten models that agree with R_g and D_{max} from SEC-SAXS
SEC-MALS– Homogeneity & M_W	DAMAVER – Averaging of 50 low-resolution structures DAMFILT – Filtering the Averaged Model from DAMAVER	Selection of top five 3D models which were superimposed on DAMFILT representative low-resolution structures
		Visual inspection of 3D models and low-resolution structures

Figure 4.1: Organisational Flowchart for the Purification and Characterisation of Sense and Antisense LincRNA-p21 *AluSx1* RNA. The determination of LincRNA three-dimensional, low-resolution structures overlaid by high-resolution, atomistic models was conducted in three phases: RNA preparation and biophysical studies to determine sample homogeneity and sample properties; low-resolution structure determination by SAXS; and high-resolution modelling using SimRNA, with constraints imposed by HYDROPRO. All methods are further described below.

4.1.1 Sense and Antisense LincRNA-p21 *Alu*Sx1 Plasmid Preparation

A flowchart of the procedure is outlined in **Figure 4.1**. LincRNA-p21 *AluSx1* transcripts for the sense and antisense (taken from the TP53COR1 gene located on Chr6:36,663,392-36,667,296 (GRCh38/hg38), Chr6:36,631,169-36,635,073 (GRCh37/hg19)) were designed from the sequences presented in from Chillón and Pyle [20, 21]. RNA constructs used in this experiment are represented below:

>LincRNA-p21 *AluSx1* Sense RNA Sequence (307nt) | TP53COR1_LIsoE2_*AluSx1*_P
5′-
AGCUGGGCGUGGUGGCUCACGCCUGUAAUCCCACCACUUUGGGAGGCCGAGGCAGGCGGAUCACUU
GAGGUCAGGAGUCCAAGACCAGCCUGGCCAACAAGGCGAAACCCUGUCUCUACUAAAAAUACAAAA
ACUAGCUGGGCGUAGUGGUGGGCACCUGUAAUCCCAGCUACUCGGGAGGCUGAGACAGGACAAUCG
CUUGGACUCCGGAGGCAGAGGUUGCAGUGAGCUGGGAUCGUGCCACUACACUCCAGUCUGGGCGAC
AGAGCAAGACUCUGCAUCAAAAAAAAAAAAAGAAAGAGUAAUAA-3′

>LincRNA-p21 *AluSx1* Antisense RNA Sequence (280nt) | TP53COR1_LIsoE2_*AluSx1*_P
5′-
GCAGAGGAGGAAUGGAAUCAUUCUUUUUUUUUUUAUUGGAGACGGAGUCUCACUCUGUUGCUCAG
GCUGGAGUGUAGUGGUGCGAACUUGGCUCACUGCAGCCUCCACCUCCCAGGCUCAAGCAAUUCUCC
UGCCUCAGCCUCCCGAGUAGCUGGGAUUACAGGUGUCUGCUAUCACACCCAGCUAAAGUUUUUAUA

94

UUUUUAGUAGAAAUGGAGUUUCACCAUGUUGGACAGGCUGGUCUCGAACUCCUGACCUCAGGUGA
UCCACCCGCCUCAGCCUC-3′

LincRNA-p21 *AluSx1* RNAs will be referred to as sense and antisense *AluSx1* RNAs throughout.

The plasmids were synthesised commercially, each sequence was flanked by a T7 RNA polymerase promoter sequence at the 5′-end and an *XbaI* restriction endonuclease cut-site sequence at the 3′-end. To increase RNA yield, two additional Gs were added to the 3′-end of the T7 promoter region which is reflected in the theoretical M_W (**Table 4.1**) [35]. LincRNA-p21 sequences were inserted into Genewiz pUC-57-KAN plasmids (Azenta Life Sciences, USA). Plasmids were transformed and cultured in *E. coli* NEBα (NEB, Canada) competent cells and were purified using NEB Monarch Miniprep Kits (NEB, Canada) as per manufacturer's protocol.

4.1.2 *In vitro* Transcriptions of LincRNA-p21 Sense and Antisense *AluSx1* Inverted Repeats and RNA Purification

RNA transcripts were prepared using run-off *in vitro* transcriptions (IVT) as prepared previously [36, 37]. Briefly, concentrated plasmid samples were digested by *XbaI* restriction endonuclease (NEB, Canada). 1 mL *in vitro* transcription reactions were performed using laboratory purified in-house T7 RNA polymerase and commercial RiboLock RNase inhibitor (Thermo Fisher Scientific, USA) [38]. Linearised plasmids were additionally incubated with 10% DMSO and 0.1% Triton X-100 to increase RNA transcript yields [39]. Sense and antisense *AluSx1* RNA were purified by SEC using a Superdex 200 Increase GL 10/300 (Global Life Science Solutions USA LLC, Marlborough, MA, USA) and purification buffer (PB) (10 mM Bis-tris pH 5.0, 100 mM NaCl, 15 mM KCl 15 mM MgCl2, 10% glycerol) with an ÄKTA pure FPLC (Global Life Science Solutions USA LLC, Marlborough, MA, USA) [40]. SEC peak fractions were assessed for purity by urea-polyacrylamide gel electrophoresis (Urea-PAGE) and sedimentation velocity analytical ultracentrifugation (SV-AUC) in 6M urea. Urea-PAGE (10%) was run at room temperature, 300V, for 40 min in 1x TBE (Tris-Borate-EDTA) buffer, followed by staining with Sybr™ Safe (Thermofisher Scientific, Saint-Laurant, QC, Canada) and visualisation. Pure fractions were pooled and concentrated by ethanol precipitation, with resuspension in HEPES Folding Buffer (HFB) (50mM HEPES, 150 mM NaCl, 15 mM $MgCl_2$, 3% Glycerol, pH 7.4) for SAXS submission.

4.1.3 Multiangle Light Scattering (MALS), and Analytical Ultracentrifugation (AUC) Studies of LincRNA-p21 *AluSx1* Sense and Antisense Inverted Repeats

SEC purified LincRNA-p21 *AluSx1* RNAs were subjected to an additional SEC purification by a Superdex 200 Increase 10/300 GL column in HFB and analysed directly by an in tandem DAWN Multiangle Light Scatterer (MALS) with Optilab Refractive Index System (Wyatt Technology, USA) to determine the M_W as per Wyatt Technologies guidelines [41]. Samples were eluted at a 0.5 mL/min flowrate and measured using 18 multiangle detectors, including a UV A_{260} and A_{280}, and a refractive index (RI) detector. MALS measurements were taken using a helium-neon red laser (632.8 nm) at 25ºC. For data analysis, the refractive index increment (dn/dc) was adjusted to 0.1721 mL/g for sense and antisense *AluSx1* RNA samples [42-44]. Data were analysed using the ASTRA v9 software package (Wyatt Technology, USA) [45, 46].

Purified sense and antisense *AluSx1* RNA were measured by SV-AUC under denaturing conditions in 6M urea to ascertain purity and composition of the transcript. Both samples were measured in two-channel centrepieces and spun at 25,000 rpm for 6 hours at 20ºC. Denaturing 6M urea buffer density (1.0899 g/mL) and viscosity (1.3896 cP) were estimated with Ultrascan and used to convert observed sedimentation and diffusion coefficients to standard conditions (water at 20ºC). Data were collected in intensity mode at 260 nm and processed using the UltraScan III Software [47]. SV-AUC data were processed as described in [48]. Briefly, systematic noise contributions and boundary conditions (meniscus and bottom of the cell position) samples were processed with the two-dimensional spectrum analysis [49]. Data was further refined by genetic algorithm analysis to achieve parsimonious regularisation [50]. The final step included a genetic algorithm-Monte Carlo (GA-MC) analysis, that was performed with 50 iterations to obtain 95% confidence intervals for the determined parameters (**Table 4.1**) [51]. AUC data analysis was performed on the XSEDE high-performance computing infrastructure using Expanse and Bridges-2 at the San Diego and Pittsburgh supercomputing centres, respectively. The final model produced very low residual mean square deviations (RMSD) of 0.00139 at 0.438 OD_{260} for sense and 0.00177 at 0.71 OD_{260} for antisense LincRNA-p21 *AluSx1*. All fits produced random residuals, which, together with the low RMSD is evidence for excellent convergence.

4.1.4 Small Angle X-Ray Scattering (SAXS) Analysis of LincRNA-p21 *AluSx1* Sense and Antisense

SAXS data for sense and antisense *AluSx1* RNA samples were collected at 2.5 mg/mL. Samples were run at Diamond Light Source Ltd. synchrotron (Didcot, Oxfordshire, UK) on the B21 SAXS beamline, with a high-performance liquid chromatography (HPLC) system attached upstream to ensure sample monodispersity [52]. A specialised flow cell was employed in conjunction with an inline Agilent 1200 HPLC system (Agilent Technologies, Stockport, UK); sense and antisense *AluSx1* RNA samples were injected onto a Shodex KW403-4F (Showa Denko America Inc., New York, NY, USA) size exclusion column pre-equilibrated with HFB. The flow rate of the column was maintained at 0.160 mL/minute with eluted samples being exposed to X-rays with 3 second exposure time and 600 frames.

Analysis of scattering data was carried out using the ATSAS suite [53]. Using Chromixs, the buffer contribution was subtracted from the sample peak [54]. A Guinier analyses (q^2 vs. $\ln(I(q))$) was performed on each data set to obtain the radius of gyration (R_g) and to determine the sample's quality [55]. A dimensionless Kratky plot (qR_g vs $qR_g^2*I(q)/I(0)$) was generated to evaluate folding of RNA molecules [56]. A paired-distance distribution function (P(r)) analysis was performed using GNOM to obtain real-space R_g and the maximum particle dimension (D_{max}) of the sample [57, 58]. Employing the information derived from the P(r) plot, a total of fifty sense and antisense *AluSx1* RNA models were generated using DAMMIN [59]. These models were then averaged using DAMAVER and then filtered using DAMFILT to produce a single representative model of each of the RNAs [59, 60].

4.1.5 Sense and Antisense LincRNAp-21 *AluSx1* RNA Tertiary Structure Determination

Using the secondary structure information from Chillón and Pyle, 2016, sense and antisense *AluSx1* tertiary structures were calculated using SimRNA v3.20 [34]. SimRNA v3.20 is a Monte Carlo sampler that operates on a coarse-grained model of RNA structure. SimRNA employs a five-bead system per nucleotide, as well as an empirically derived knowledge-based potential. A total of 20 million SimRNA iterations in replica exchange mode were performed for both sense and antisense *AluSx1* RNAs. SimRNA clustering was then performed within one percent of all trajectories with the lowest energy. A RMSD cut-off of five was applied to filter 3080 clusters of similar structures for both sense and antisense *AluSx1* RNAs.

4.1.6 High-Resolution Structural Modelling of LincRNA-p21 *Alu*Sx1 Sense and Antisense

The representative cluster models containing 3080 computationally generated high-resolution models for both the sense and antisense *AluSx1* were separately assessed by HYDROPRO to generate hydrodynamic properties for each model [61]. Running conditions for HYDROPRO involved buffer properties for HFB as determined by UltraScan III: a viscosity of 1.10068 cP; and buffer density of 1.014 g/cm^3 [47]. The theoretical M$_W$ of 99.418 kDa and 89.543 kDa were applied to the HYDROPRO parameters of sense and antisense *AluSx1* RNA, respectively. Models were superimposed onto the SAXS DAMFILT structures and fitted using DAMSUP. Models exhibiting an NSD (normalized spatial discrepancy) value of 1.00 to 1.15 which indicates close fitting were further selected to represent the high-resolution, atomistic RNA model [60]. Models exhibiting similar HYDROPRO determined R$_g$ and D$_{max}$ were further selected for and formed the top ten models of interest. The top ten models were energy minimised using an additional step involving QRNAS, which employed a subset of the AMBER force field to achieve energy minimisation of the structures generated from coarse-grained three-dimensional modelling [62]. 20,000 QRNAS MD iterations were performed from the original SimRNA full-atom reconstructed high-resolution models that best-fit the averaged, filtered low-resolution, three-dimensional structure obtained from DAMFILT. Subsequently, five best fit models were superimposed on SAXS structures and represented using PyMOL [60, 63].

4.2 Results

4.2.1 Purification of LincRNA-p21 Sense and Antisense AluSx1 Inverted Repeats

Both sense and antisense *AluSx1* RNAs were purified using SEC, eluting principally around ~10.0 – 12.5 mL at a flowrate of 0.5 mL/min on the Superdex 200 Increase GL 10/300 (**Figure 4.2A**). The left peak indicates plasmid excluded from subsequent analysis while the right peak represents the RNA of interest. The left plasmid peak for sense and antisense RNA was previously identified using denaturing Urea PAGE gels not depicted here. Differences between plasmid peak intensity for sense and antisense *in vitro* transcriptions is the result of a higher concentration of purified and digested plasmid being provided for antisense LincRNA-p21 IVTs compared to sense IVTs. **Figure 4.2B** depicts the 10% Urea PAGE gel for the RNA fractions indicating that both RNAs migrated closely with similar length and, around the ~300bp marker with respect to their elution volumes (Red Inset). Sense LincRNA-p21 *AluSx1*

generally produced closely eluting bands (around ~300bp). SV-AUC experiments were conducted using the single fractions collected at 11.0 mL and 11.5 mL for sense and antisense *AluSx1* RNA respectively (indicated by the right shoulder, blue inset, **Figure 4.2A**).
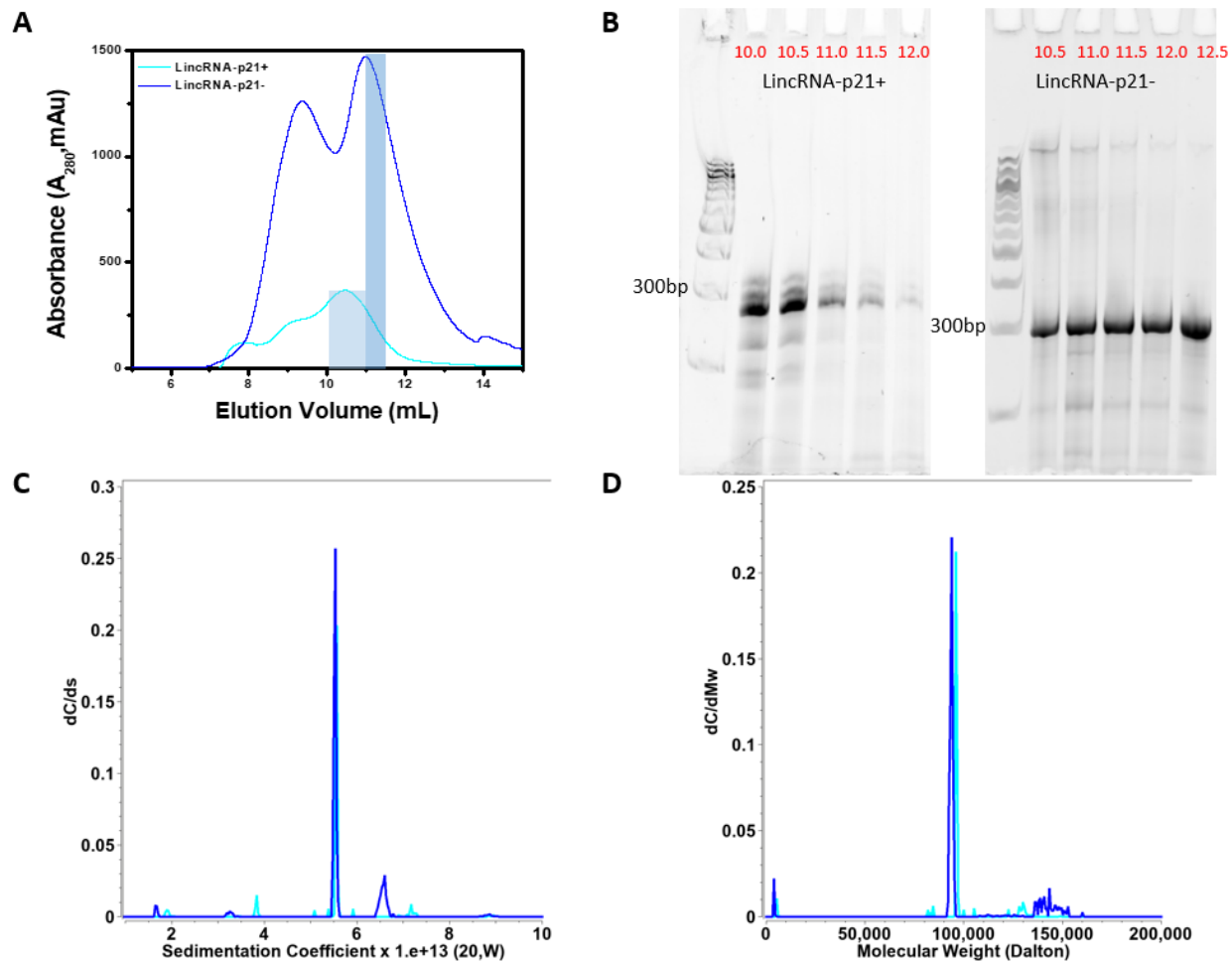

Figure 4.2: Purification of Sense and Antisense *in vitro* Transcribed LincRNA-p21 *AluSx1* RNA. (**A**) depicts the size exclusion chromatogram of the sense and antisense *AluSx1* RNA elution profile using the Superdex 200 Increase GL 10/300 column. SEC-MALS and SV-AUC experiments were performed with the fractions highlighted in light blue (sense) and dark blue (antisense). (**B**) shows the 10% urea PAGE gel used to ascertain the sense and antisense LincRNA-p21 RNA purity extracted using 0.5 mL fractions (volumes in red) using an ÄKTA Pure FPLC through a Superdex 200 Increase GL 10/300 SEC column. Fractions collected at 11.0 mL and 11.5 mL for sense and antisense *AluSx1* purifications were consolidated and used for SAXS and SV-AUC experiments. A Quick-Load® Purple 100 bp DNA Ladder (NEB, Canada) was used for the 10% urea PAGE gels in lanes 1 and 7 of each gel. (**C**) dC/ds sedimentation coefficient distributions for sense (light blue) and anti-sense (dark blue) under 6M urea denaturing conditions. (**D**) same as (**C**), except transformed to molar mass distributions assuming a partial specific volume of 0.516 mL/g.

4.2.2 Biophysical Characterisation of LincRNA-p21 Sense and Antisense AluSx1 Inverted Repeats

Sedimentation and diffusion coefficients and their 95% confidence intervals resulting from the GA-MC analyses are listed in **Table 4.1**. Together with sequence based molar masses, SV-AUC results can be used to derive

partial specific volumes and anisotropies for the RNA measurements. Since both RNA molecules were measured in the same urea buffer, it is reasonable to assume that the partial specific volume is similar for both molecules. Sedimentation experiments were performed in 6M urea to denature the molecule and disrupt hydrogen bonding within double-stranded RNA regions of the molecule. Results shown in **Figure 4.2C** and **Figure 4.2D** indicate that both samples contain one major species with similar sedimentation coefficients, and molar masses in agreement with molar masses predicted from sequence when using a partial specific volume of 0.516 mL/g. This result is consistent with a monomeric and homogeneous full-length transcript of sense (84% of total concentration) and antisense (75% of total concentration) LincRNA-p21 *AluSx1* RNA. Frictional ratios and hydrodynamic radii derived for both molecules indicate a high anisotropy for both molecules, consistent with an unfolded and extended molecule.

SEC-MALS analysis was conducted to determine the M_W of the sense and antisense *AluSx1* RNA SAXS. LincRNA-p21 *AluSx1* RNAs were purified again on a Superdex 200 Increase GL 10/300 SEC column which produced peaks eluting between 11 – 12.5 mL (**Figure 4.3A**). M_W values of sense and antisense *AluSx1* reported from SEC-MALS are slightly higher than the molar masses calculated from their sequences except for sense LincRNA-p21 *AluSx1* which exhibited less than 0.2% difference from the theoretical M_W at 99.24 ± 0.01 kDa (**Table 4.1**). M_W uniformity throughout **Figure 4.3B** and **Figure 4.3C** indicates that both RNAs are monomeric. SEC-MALS results further confirm the homogeneous composition of sense and antisense *AluSx1* RNA determined in SV-AUC. RNA degradation or shorter transcripts can be excluded since no significant smaller fragments were detected. Hydrodynamic parameters derived from SEC-MALS, and SV-AUC are summarised in **Table 4.1**.
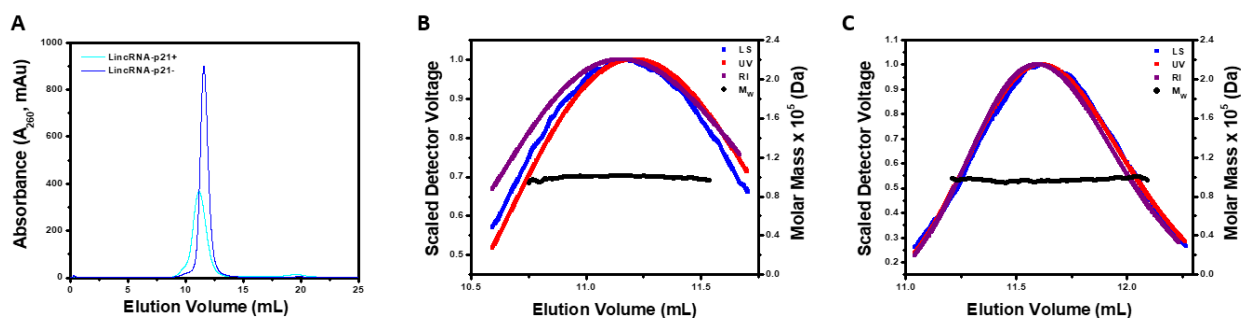
Figure 4.3: Molecular Weight Determination of Sense and Antisense LincRNA-p21 *AluSx1* RNA using SEC-MALS. (A) Portrays the elution curve from the Superdex 200 Increase GL 10/300 SEC of sense and antisense *AluSx1* RNAs. **(B)** Demonstrates the absolute molecular weight distribution across the elution peak of sense LincRNA-p21 *AluSx1* RNA's elution profile, and light scattering (blue), UV (red), and RI (purple) scattering. **(C)** Portrays the absolute molecular weight distribution across the elution peak the results fitting of antisense LincRNA-p21 *AluSx1* RNA's elution profile, and light scattering (blue), UV (red), and RI (purple) scattering.

Table 4.1: Solution Properties of Sense and Antisense LincRNA-p21 *AluSx1*.

Sample	Sense LincRNA-p21 *AluSx1*	Antisense LincRNA-p21 *AluSx1*
M_w Theoretical (kDa)	99.418	89.543
M_w AUC (kDa) [+]	94.770	92.561
M_w SEC-MALS (kDa)[∇]	99.24 ± 0.01	94.52 ± 3.71
Sedimentation Coefficient, $s_{20, w}$ (10^{-13} s) [+]	5.56 ± 0.25	5.53 ± 0.05
Diffusion Coefficient, $D_{20, w}$ (10^{-7} cm^2/s) [+]	2.95 ± 0.59	3.00 ± 0.03
Frictional Ratio, f/f_0 [+]	2.71	2.68
R_h (Å)[+]	72.7	71.3
$q.R_g$ range [#]	0.43-1.29	0.42-1.25
R_g (Å) [#]	60.87 ± 0.85	59.07 ± 0.15
$I(0)$ [Δ]	0.01 ± 9.90 × 10^{-5}	0.07 ± 8.14 × 10^{-5}
R_g (Å) [Δ]	61.71 ± 0.31	58.37 ± 0.07
D_{max} (Å)[Δ]	185.0	180.7
X^2 *	~1.148	~1.084
NSD *	1.080 ± 0.024	1.005 ± 0.022

The M_W of the sense and antisense LincRNA-p21 Alu Repeat RNA were calculated using the nucleotide sequences provided by Chillón and Pyle, 2016. [+] Molar masses and frictional ratios determined by AUC assume a partial specific volume of 0.516 mL/g and refer to conditions where the RNA is denatured by 6M urea. [+] are within 95% confidence intervals. Data points [∇] were determined from SEC-MALS experiments. Data points [#] were derived from the Guinier analysis. Data points [Δ] were determined using P(r) analysis using the GNOM program. Data points * were derived from DAMMIN and DAMAVER analysis. Terms: Hydrodynamic Radius (R_h); Radius of Gyration (R_g); Maximum Particle Dimension (D_{max}); Normalised Spatial Discrepancy (NSD).

4.2.3 Low-Resolution Structural Studies of LincRNA-p21 Sense and Antisense AluSx1 Inverted Repeats

SAXS is a powerful method that can represent the overall solution shape of biomolecules under physiologically relevant conditions. Using SEC-SAXS, which can separate different species according to their size before being applied to the SAXS measuring cell, provides confidence in the monodispersity of purified samples [64-67]. The resulting datasets were merged and presented in **Figure 4.4A** depicting the scattering intensity relative to angle for sense and antisense *AluSx1* RNA. A Guinier analysis (l(q)) vs. (q^2)) represented by **Figure 4.4B** displays the LincRNA-p21 *AluSx1* RNA samples' purity [55]. The Guinier analysis determined the Guinier R_g from the low-q region as being 60.87 ± 0.87 Å for sense LincRNA-p21 *AluSx1* and 59.07 ± 0.15 Å for antisense LincRNA-p21 *AluSx1*. Intensity data from **Figure 4.4A** was transformed to produce a dimensionless Kratky plot (**Figure 4.4C**) to determine the LincRNA-p21 *AluSx1* RNAs' conformations in solution [68]. The dimensionless Kratky plot for both

the sense and antisense AluSx1 shows a levelled-plateau which suggests them as being folded and extended in solution [69].

Figure 4.4D represents the (P(r)) plot which was derived from indirect Fourier transformations to convert the reciprocal-space information of the intensity data in **Figure 4.4A** to real-space electron pair distance distribution data [70]. Using the P(r) plot, sense LincRNA-p21 *AluSx1* presented a real-space R_g of 61.71 ± 0.31 Å and a D_{max} of 185.0 Å, while the antisense LincRNA-p21 *AluSx1* presented a real-space R_g of 58.37 ± 0.07 Å and D_{max} of 180.7 Å.

DAMMIN was performed to obtain low-resolution structures for the sense and antisense *AluSx1* RNA. Fifty models were calculated for each sense and antisense *AluSx1* RNAs which demonstrated favourable agreement as indicated by the X^2 values (~1.148 for sense LincRNA-p21 *AluSx1* and ~1.084 for antisense LincRNA-p21 *AluSx1*). DAMFILT and DAMAVER were performed to filter and averaged the models. The NSDs were estimated to be 1.080 ± 0.024 and 1.005 ± 0.022 for sense and antisense respectively (**Table 4.1**) [63].

We identified two, single representative SAXS envelopes illustrated by **Figure 4.5**. The averaged, single-representative SAXS envelope of sense LincRNA-p21 *AluSx1* is generally extended, adopting a non-spherical, nonglobular surface model (**Figure 4.5A**). The SAXS envelope is additionally asymmetrical in its rotation along its x- and y-axes, exhibiting two prominent bulges that are primarily located on its ends. **Figure 4.5B** shows the antisense LincRNA-p21 *AluSx1* SAXS envelope which is similarly elongated and asymmetrical. Antisense LincRNA-p21 *AluSx1* though has three prominent bulges, two located centrally, while the third distally protrudes outwards from the centre.
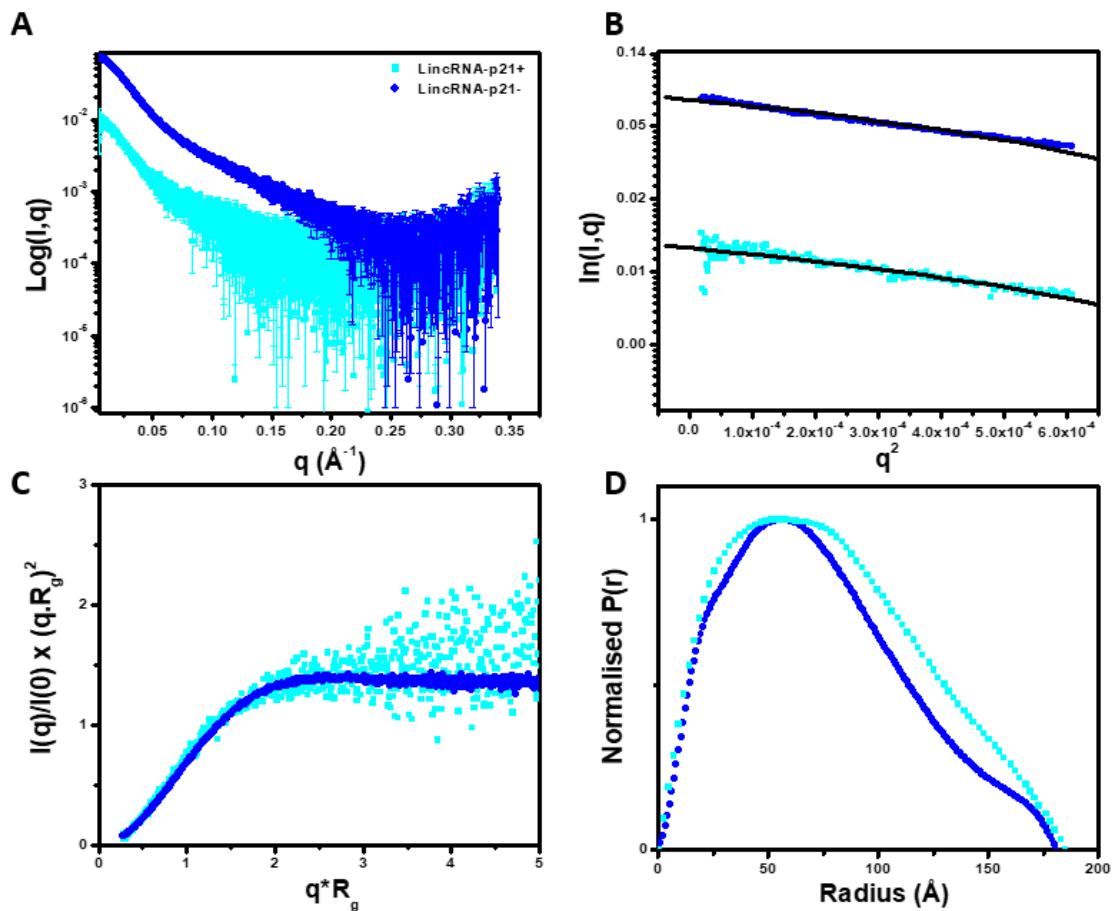
Figure 4.4: Small Angle X-Ray Scattering (SAXS) Characterisation of Sense and Antisense LincRNA-p21 ***AluSx1*** **RNA.** (**A**) merged scattering data of sense and antisense *AluSx1* RNA depicting the scattering intensity (log I(q)) vs. scattering angle ($q = 4\pi\sin\theta/\lambda$). (**B**) Guinier plots allowing for the determination of R_g from the low-angle region data and representing the homogeneity of samples. (**C**) Dimensionless Kratky plots ($I(q)/I(0)*(q*R_g)2$ vs. $q*R_g$) of sense and antisense *AluSx1* RNA depicting the elongated, tube-like structures because of the non-Gaussian, levelled-plateau shape of the curve. (**D**) Normalised pair distance distribution plots for sense and antisense *AluSx1* RNA which permits the determination of R_g derived from the SAXS dataset and including each molecule's D_{max}.

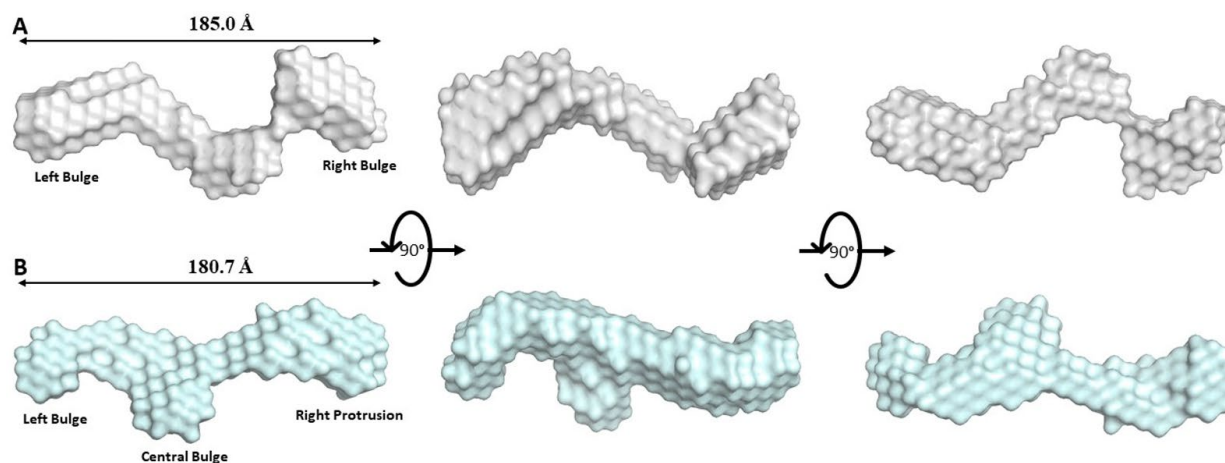**Figure 4.5: Low-Resolution Structures of Sense (A, Grey) and Antisense (B, Pale Cyan) LincRNA-p21 *AluSx1* Inverted Repeats Determined using SAXS. (A)** The averaged DAMAVER SAXS low-resolution structure of sense LincRNA-p21 *AluSx1* RNA, taking on an elongated, asymmetrical, and extended structure with maximum length of 185.0 Å. Key features include a left and right Bulge. **(B)** The averaged DAMAVER SAXS low-resolution structure of antisense LincRNA-p21 *AluSx1* RNA, adopting an elongated, asymmetrical, and extended structure with maximum length of 180.7 Å. Key features include a left bulge, central bulge, and a right protrusion. Dimensions are represented by the D_{max} obtained from the P(r) analysis. Models are rotated along their x-axis by 90º as represented by the inset.

4.2.4 High-Resolution Atomistic Models of LincRNA-p21 AluSx1 Sense and Antisense Inverted Repeats

Using SimRNA v3.20, and the secondary structure constraints for both RNAs based on previous studies, we calculated 10,000 clusters of high-resolution, atomistic models for each RNA [21]. These models were further refined through energy minimisation steps to remove models that did not satisfy constraints such as defined atom distances, bond lengths and angles. Subsequently, we obtained 3080 high-resolution models that can be superimposed on the DAMFILT SAXS envelopes using DAMSUP. Sense LincRNA-p21 *AluSx1* models when superimposed produced ten models that had an NSD range from 1.053 to 1.094, while for the antisense RNA, the top ten models retained an NSD range of 1.113 to 1.175. We further applied a selection process using the real-space R_g and D_{max} values determined from SAXS for each molecule. We employed the program HYDROPRO to calculate biophysical properties such as R_g and D_{max} from the 3080 high-resolution models, as performed previously [71]. Top ten models were further reduced to five using the HYDROPRO properties to achieve models that were in close approximations to SAXS determined R_g and D_{max}.

Both the top five high-resolution, high-fidelity sense and antisense *AluSx1* models are represented by **Figures 4.6** and 4.7 respectively. **Figure 4.6** depicts the sense LincRNA-p21 *AluSx1* RNAs that closely fit with the SAXS envelopes generated in **Figure 4.5A**. Previous chemically probed secondary structure predictions identified three

major secondary structures: a left and right arm, and a 3′-three-way junction which have been modelled using SimRNA and represented in **Figure 4.6**. High-fidelity sense LincRNA-p21 *AluSx1* characteristically exhibits the right arm (Magenta) that curls into the central RNA body while the left arm (Blue) extends outwards. Multiple models depict variance in the right arm's position, appearing to adopt multiple conformations that curl, but rarely extend outwards, towards the central RNA body. The right arm's stem-loop on its head is additionally compacted and has either bridged with the main RNA body against the 3′-three-way junction or curls outwards. We have identified a consistent 3′-adenyl tail (Cyan) that consistently wraps around the right arm's base or the connection with the 3′-three-way junctions (Yellow). A 5′-junction (Green) is also a present feature identified but not named by the previous study, but consistently appears to project outwards, perpendicularly from the RNA's x-axis (**Figure 4.6**). A flexible, and generally unnamed region – the single-stranded linker (Orange), is presented centrally between the two arms, adopting no specific structure. An animated representation of **Figure 4.6B** is attached in **Supplementary Information** as **SM 1**.

Figure 4.7 similarly presents the high-fidelity, high-resolution structures of antisense LincRNAp21 *AluSx1* which were modelled using previous chemically probed secondary structure predictions using SimRNA v3.20 [21]. They exhibit the left and right arms, and the 5′-three-way junctions. Both the left (Blue) and right (Magenta) arms of antisense LincRNA-p21 *AluSx1* project laterally in line with the x-axis. The identified 5′-uridyl tail (Cyan) consistently wraps around the left arm. Both the 3′-three-way junctions, and the identified 5′-junction are compacted centrally, with regions that project perpendicularly from the RNA body's x-axis. The right arm of antisense LincRNA-p21 *AluSx1* does not retain the characteristic stem-loop head that sense LincRNA-p21 *AluSx1* has. An animated representation of **Figure 4.7C** is attached in **Supplementary Information** as **SM 2**.

Figure 4.6: The SimRNA High-Resolution, High-Fidelity Models of Sense LincRNA-p21 *AluSx1* RNA. Figure 4.6 presents the high-resolution, high-fidelity sense models that have good fitting with their SAXS envelope as demonstrated by low NSD values. **A** represents model 514; **B** represents model 1036; **C** represents model 1476; **D** represents model 1677; and **E** represents model 1794 which exhibit chemically probed secondary structures: left arm (Blue); 5′-junction (Green); three-way junction (Yellow), right arm (Magenta), and the 3′-adenyl tail (Cyan). Terminal nucleotides are displayed as: 5′ nt (Red, Sphere Modelled) and 3′ nt (Lime Green, Sphere Modelled). Models are rotated along their x-axis by 90º as indicated by the inset. A flexible, single-stranded linker sequence is represented centrally (Orange).
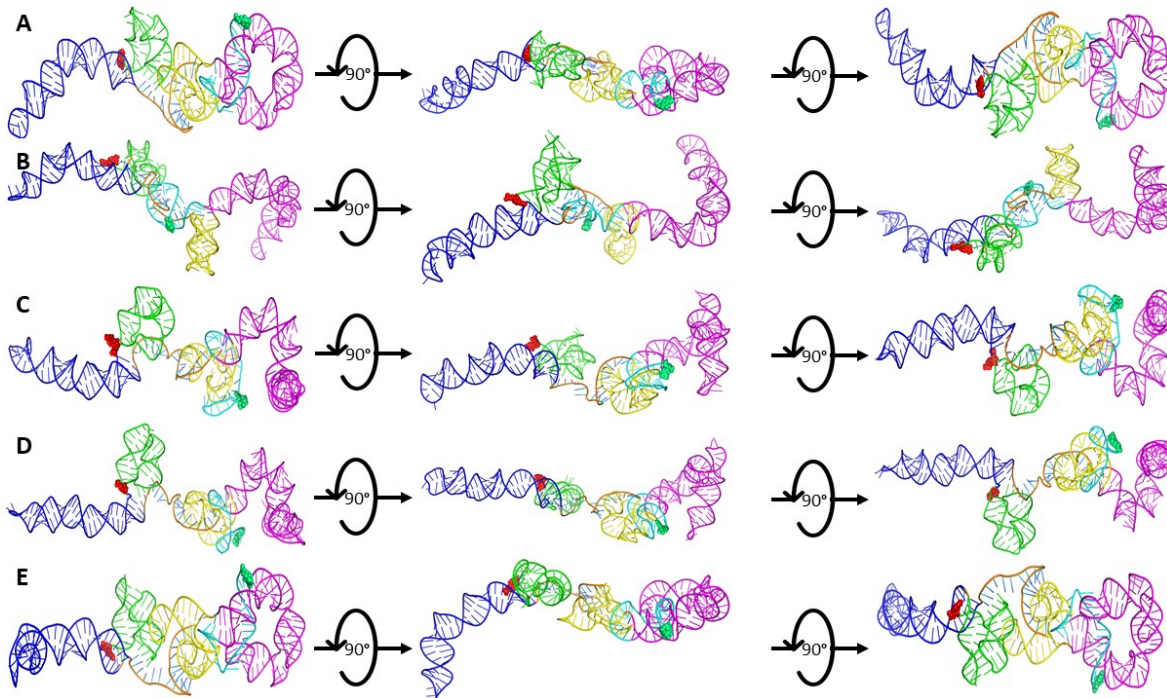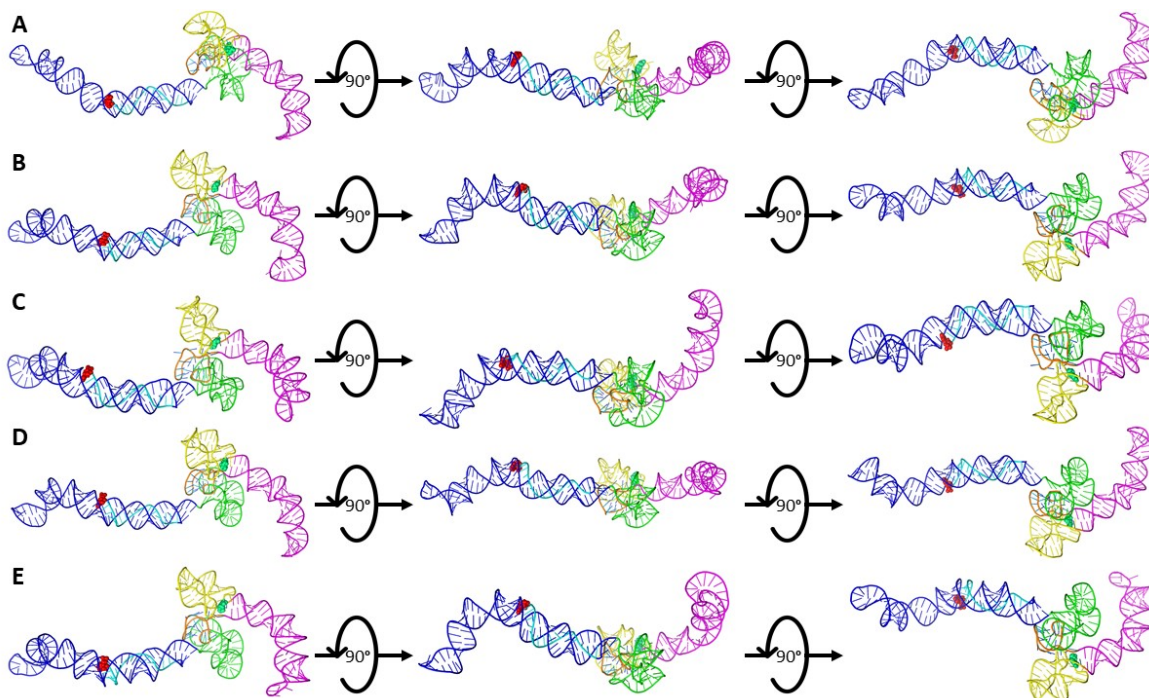
Figure 4.7: The SimRNA High-Resolution, High-Fidelity Models of Antisense LincRNA-p21 *AluSx1* RNA. **Figure 4.7** presents the high-resolution, high-fidelity sense models that have good fitting with their SAXS envelope as demonstrated by low NSD values. **A** represents model 66; **B** represents model 974; **C** represents model 1013; **D** represents model 1074; and **E** represents model 1417 which exhibit chemically probed secondary structures: left arm (Blue); 5′-junction (Green); three-way junction (Yellow), right arm (Magenta), and the 5′-uridyl tail (Cyan). Terminal nucleotides are displayed as: 5′ nt (Red, Sphere Modelled) and 3′ nt (Lime Green, Sphere Modelled). Models are rotated along their x-axis by 90º as indicated by the inset. A flexible, single-stranded linker sequence is represented centrally (Orange).

After performing DAMSUP, the high-fidelity, high-resolution models were visually inspected in terms of their alignment with the low-resolution SAXS envelope from DAMFILT and were represented in **Figures 4.8** and 4.**9** for sense and antisense *AluSx1* RNA respectively. **Figure 4.8** displays a general agreement with the sense SAXS envelope overlaid by their high-fidelity, high resolution SimRNA models. Both the left and right arms when superimposed fit neatly within the protruding two bulges of the SAXS envelope. Minor disagreement occurs for the right arm's stem-loop head which appears to be excluded from the SAXS structure, which is similar for the tip of the Left arm. Both the 5′-junction and 3′-three-way junction exhibits considerable overlap with the SAXS low-resolution structure when superimposed. **Figure 4.9** shows a general agreement with the antisense SimRNA models with their respective SAXS low-resolution structures. The left arm, 5′-junction, and 3′-three-way junctions are secondary structures that exhibit the highest relative overlap with the SAXS envelope and occur within the central bulge and right-most protrusion. The right arm (Magenta) depicts relatively lower agreement, with its tip and core regions exposed and externalised from the SAXS envelope. This is primarily confined to the left bulge, however, which

107

indicates a lack of adequate fitting. Two animated representations of one of the sense (**Figure 4.8B**) and antisense (**Figure 4.9C**) models are attached in **Supplementary Information** as **SM 3** and **SM 4** respectively.

Figure 4.8: Superimposed Overlays of Sense SAXS Envelopes with their High-Resolution, High-Fidelity SimRNA Models. **Figure 4.8** represents the combined overlays of the sense LincRNA-p21 *AluSx1* RNA envelope with their top high-resolution, SimRNA models: (**A**) 514; (**B**) 1036; (**C**) 1476; (**D**) 1677; and (**E**) 1794. Overall, SAXS envelopes indicate a general agreement with computationally generated structures, showing high overlap of the extended molecule with the major secondary structures identified using chemical probing techniques.
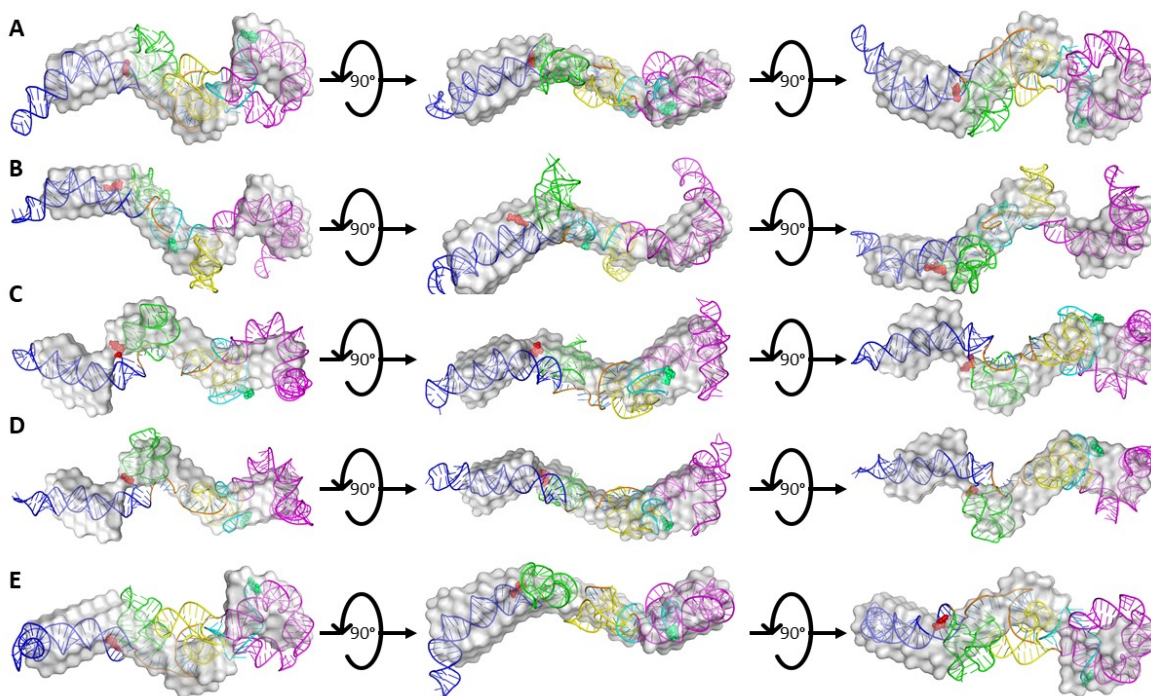
Figure 4.9: Superimposed Overlays of Antisense SAXS Envelopes with their High-Resolution, High-Fidelity SimRNA Models. Figure 4.9 represents the combined overlays of the antisense LincRNA-p21 *AluSx1* RNA envelope with their top high-resolution, SimRNA models: (**A**) 66; (**B**) 974; (**C**) 1013; (**D**) 1074; and (**E**) 1417. Overall, SAXS envelopes indicate a general agreement with computationally generated structures, showing high overlap of the extended molecule with the major seconda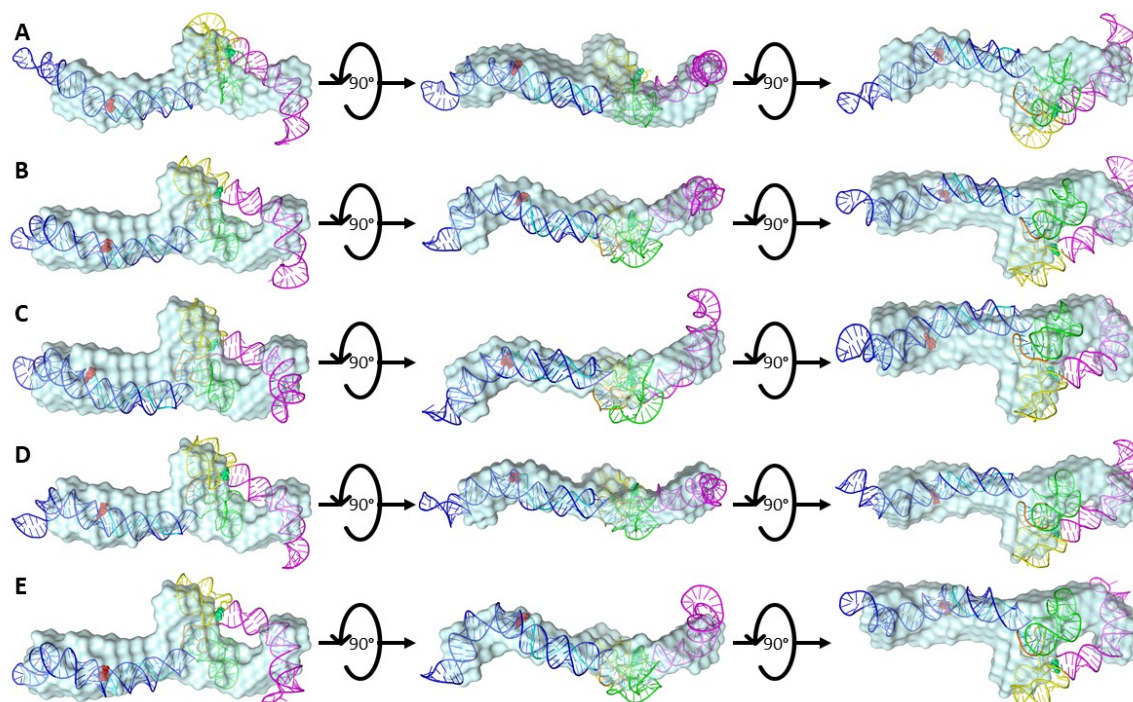ry structures identified using chemical probing techniques. However, there is a slight overhang present with the right arm (Magenta) which has an area excluded from overlapping with the SAXS envelope.

4.3 Discussion

RNA secondary structure is especially important in defining an RNA molecule's roles and functions [72, 73]. These structures can be studied using a variety of techniques including secondary structure probing methods such as SHAPE, NMR, comparative sequence alignment and analysis, or three-dimensional structure predictions software for conceptualising higher order complexes [74-77]. Identifying functional structural elements of RNA is especially important when observing the noncoding elements of the genome, whose significant contribution of noncoding RNA (ncRNA) play important regulatory roles in complex organisms. LincRNA-p21 is one such important regulatory lncRNA that is directly targeted by p53 in response to DNA damage. Being a transcriptional repressor in the p53 pathway, understanding LincRNA-p21's secondary structure is a first step evaluation towards how it binds to and modulates hnRNP-K localisation which is a process that ultimately triggers apoptosis in DNA damaged or cancerous cells.

The primary intent of our work is to obtain the high-resolution details of LincRNA-p21 *AluSx1* RNA by combining the previously determined secondary structure of the Inverted Repeat transposable elements determined by Chillón and Pyle with SAXS structure determination and computational modelling [21]. **Figure 4.1** outlines our multifaceted process for biophysically characterising molecules using SEC-SAXS, SEC-MALS, SV-AUC, and SimRNA from chemically probed secondary structure determinations. SEC-MALS and AUC act as orthogonal quality control and validation techniques when combined with PAGE [78].

We initially started by transcribing sense and antisense *AluSx1* RNAs, which were purified to homogeneity as indicated by **Figure 4.2A**. Both RNA displays multi-modal peaks with a primary peak representative of the monomeric RNA species. Fractions along the primary peak were taken and analysed by 6M urea PAGE, which showed relatively pure and pronounced RNA bands around 300 bp indicating that most of the species is the primary RNA of interest (**Figure 4.2B**). Subsequent urea PAGEs were performed, and each resulted in an inconclusive answer to RNA homogeneity. We therefore turned to SV-AUC to determine if the RNA samples were heterogenous as indicated by urea-PAGE, or homogenous enough for further characterisation. As indicated in **Figure 4.2C/D**, both sense and antisense *AluSx1* RNA adopted a single, major species in solution when denatured by 6M urea. Any minor species had negligible concentrations and reflected noise contributions. The SV-AUC experiments showed no evidence for RNA degradation or aggregation. The M_W values of sense and antisense *AluSx1* RNAs determined by AUC were less than 5% different (**Table 4.1**). AUC experiments were carried out in the presence of 6M urea as a denaturant, so we needed to assess each RNA in non-denaturing conditions. We therefore utilized SEC-MALS to determine absolute molecular weight in solution [41, 79]. Each RNA eluted as a singular, tight, and gaussian distribution elution profile evident in **Figure 4.3A**. Comparatively to AUC results, the RNA also exists monomerically in HFB as indicated by their uniform M_W demonstrated in **Figures 4.3B/C**. The sense RNA molecular weight resulted in 99.24 ± 0.01, virtually identical to the theoretical molecular weight of 99.42 kDa. Antisense RNA also resulted in a very similar molecular weight of 94.52 ± 3.71 to the theoretical value of 89.54 kDa. These results suggest the RNA was acceptable to undergo SEC-SAXS and three-dimensional structure determination since SAXS requires highly homogenous samples.

SAXS is an ideal technique to determine low-resolution three-dimensional structures of molecules in their native state and was used to investigate the solution conformations of both RNAs [59, 80, 81]. Overall, scattering intensities were acceptable even in low-angle regions as shown in **Figure 4.4A**. The linear regression of the sense and

antisense *AluSx1* samples in **Figure 4.4B** portrays the intensities within the defined low-q^2 range as being linear with the absence of upward curves, illustrating monodispersity and the absence of attractive or repulsive interactions between scatterers [69, 82, 83]. Guiner R_g approximation resulted in 60.87 ± 0.85 and 59.07 ± 0.15 for sense and antisense *AluSx1* RNAs respectively. These Rg approximations are consistent with an elongated RNA molecule [84, 85]. Furthermore, the relative foldedness of sense and antisense RNAs can be deduced from the dimensionless Kratky plot in **Figure 4.4C.** The Kratky plot depicts both sense and antisense RNAs as being elongated but relatively folded, like other ncRNA [36, 86]. Additionally, the Guinier R_g and the real-space R_g are close with less than 1.5% difference.

Figure 4.4D represents both RNA's distance distribution functions and are non-Gaussian, further consistent with extended molecules. Globular molecules will generate a Gaussian-like P(r) distributions which is not demonstrated in **Figure 4.4D** further justifying its extended shape [70]. Using the P(r) distribution, molecules can also be described on their overall shape and symmetry to confirm solution folding acquired from the dimensionless Kratky plot [70]. Generally, globular molecules will display a bell-shaped curve with a maximum at approximately $D_{max}/2$ while elongated molecules retain non-Gaussian, asymmetrical distributions with a maximum at smaller distances which appear as shoulders. For elongated molecules, this distribution will correspond to the radius of the cross section which will generally be illustrated by a tailing of the profile at larger distances [36, 87]. This is present for both sense and antisense *AluSx1* RNA whereby their respective $D_{max}/2$ do not demonstrably produce an even bell-curved but rather lead into right-leaning tails reinforcing the overall elongated shape established by **Figure 4.4C**. Sense and antisense adopt maximum distance towards 185.0 Å and 180.7 Å respectively. Both RNAs result in similar D_{max} measurements which is expected of two RNA of similar lengths, with the sense RNA being larger than the antisense RNA. These size difference can be attributed both to the RNAs' folding properties, and the 307nt vs 280nt lengths of the sense and antisense sequences as derived from Chillón and Pyle [21].

SAXS analysis of both the sense and antisense RNAs provided structures with noticeably consistent features, including a left and right bulge for the sense RNA and a left and central bulge, and right protrusion for the antisense RNA (**Figure 4.5**). Use of the standard HFB was important since human LincRNA-p21 folds at near physiological concentrations at around 5mM $MgCl_2$ with maximum compaction at 15 mM $MgCl_2$ [21]. RNA stability is dependent on non-electrostatic (hydrophobic interactions, van der Waals interactions, translational, rotational, vibrational, and configuration entropies) and electrostatic contributions, with the latter having an impact on nucleic acid charges in solution [88]. Incidentally, Mg^{2+} ions are necessary to strongly stabilise the tertiary structure of RNA molecules

through their divalent ion binding with the RNA phosphate backbone which subsequently neutralises it [89, 90]. As observed in the previous studies, LincRNA-p21 folds with greater stability between $5 - 15$ mM $MgCl_2$ concentrations which was selected for the current investigation to reproduce the secondary structure elements determined from the SHAPE studies for SAXS [21, 37]. The aforementioned SAXS features are observed to overlap with predicted features seen in the secondary structure analysis by Chillion and Pyle [21]. These features, while they can be seen in the SAXS structures, were arbitrarily based on structure orientation because directionality cannot be determined from these models. Therefore, we needed to not only computationally derive higher resolution models for clarity, but also directionality.

When analysing the high-resolution computational models derived from the secondary structure determinations, a straitened selection process was needed to screen for models that match the experimentally determined low-resolution structures. Therefore, HYDROPRO was employed to compute the hydrodynamic properties of sense and antisense *AluSx1* rigid macromolecules from their atomic-level structure [91-94]. HYDROPRO's calculation comprise the basic hydrodynamic properties including the translational diffusion coefficient, sedimentation coefficient, intrinsic viscosity, and relaxation times, and can additionally provide the radius of gyration [61]. Incidentally, HYDROPRO can be used as another orthogonal selection process against computationally generated models that don't fit the solutions scattering data. It was employed to minimise the ~3080 models which could subsequently be fit using DAMSUP based on the solution scattering range of D_{max} and R_g. Consequently, observing HYDROPRO merely as a fitting tool narrowed the selection of potential models to five.

As detailed in **Figures 4.6 and 4.7**, sense and antisense *AluSx1* RNA folds into a double-stranded RNA molecule that appear to be consistent with Chillón and Pyle's secondary structure determinations. Human LincRNA-p21 is itself a linear, single exon lncRNA which contains IR Alu repeats such as *AluSx1* [21]. Both the sense and antisense *AluSx1* RNAs were determined to comprise a left and right Arm with both arms connected by a single-stranded region. The 5′-domain of each arm is characterised by a central three-way junction while the 3′-domain is characterised by a long stem-loop. These subsequently form independent structural domains which contribute to a variety of core functions such as human LincRNA-p21's nuclear localisation following cell stress and DNA damage events [20]. They suggested that the 5′-end of LincRNA-p21 interacts with hnRNP-K which could also regulate its nuclear localisation in conjunction with its *AluSx1* inverted repeat elements. Our low-resolution, three-dimensional structure of sense LincRNA-p21 *AluSx1* RNA (**Figure 4.8**) exhibits what appears to be two dsRNA arms that fold

helically into the otherwise elongated, but radially compact model. When overlaid with the SAXS envelope, considerable overlap is present between the high-fidelity, high-resolution models. Both the 5′-junction and 3′-three-way junctions are tightly localised in the centre of the RNA body, while the left and right arms branch outwards. This likely forms a binding pocket or coordination site for hnRNP-K's one of three KH RNA/DNA binding domains which is involved in eliciting transcription regulation in the nucleus [95-97]. Antisense LincRNA-p21 *AluSx1* RNA is again represented as a dsRNA molecule with both the 5′-junction and 3′-three-way junctions being compressed around the central bulge of the SAXS envelope (**Figure 4.9**). The RNA molecule itself is mostly extended with the left and right arms projecting outward which likely contributes towards interacting with hnRNP-K RNA binding domains.

Agreement of solution structure and computational models for sense LincRNA-p21 *AluSx1* as evident by high overlaps and tight fitting. However, tangible exclusions are present for antisense LincRNA-p21 *AluSx1* RNA caused by a lack of overlap with the right arm and the tip of the left arm. This can potentially be attributed to the difference in length between the solution scattering sequence (280 nts) and the secondary structure prediction sequence (301 nts). The difference can be explained by construct synthesis whose sequence lacked 41 nts at the 3′-end but encompassed the remaining 80% of the secondary structure sequence predicted which still comprises the core and majority of the SAXS structure (**Supplementary Information**). Nevertheless, both structures detail the presence of the secondary structures identified previously, and their orientations. Applying both solution scattering techniques and coarse-grained computational modelling reveals that LincRNA-p21 Alu Inverted Repeats do not adopt completely stable, single-representative structures due in part to the conformational flexibility present in their respective DAMMIN and SimRNA models. Both the sense and antisense *AluSx1* RNAs adopt multiple conformations, however, they closely approximate into a generally similar, single-representative structure with mild conformational and structural differences when averaged. The main shape – one that is elongated, asymmetrical, with regions that encompass the main left and right arms, and three-way junction – is uniform and preserved throughout solution scattering and computational fitting techniques. Consequently, applying a combination of SAXS and computationally generated tertiary structure models concertedly determined appropriate representations of the LincRNA-p21 *AluSx1* Inverted Repeats.

4.4 Conclusion

We have demonstrated that chemical probing techniques involved in RNA secondary structure predictions such as SHAPE can be combined using a multifaceted biophysical approach involving SAXS, AUC, SEC-MALS and computational RNA modelling. Overall, sense and antisense LincRNA-p21 *AluSx1* Inverted Repeats adopt organised tertiary structures with regularly occurring secondary structure motifs that contribute towards LincRNA-p21's role within the p53 tumour suppression pathway, especially regarding hnRNP-K's nuclear localisation, as demonstrated in the previous study's knockdown experiments [21]. Nevertheless, these three-dimensional structures are important in confirming secondary structure motifs that are predicted through probing techniques. By expanding the structure to include three-dimensional native state folding, essential regulatory RNAs involved in apoptosis and tumour suppression in cancer cells can be effectively visualised to identify functional domains and potential RNA-Protein binding regions.

4.5 References

1. Rufini, A., et al., *Senescence and aging: the critical roles of p53.* Oncogene, 2013. **32**(43): p. 5129-43.
2. Amaral, J.D., et al., *The role of p53 in apoptosis.* Discov Med, 2010. **9**(45): p. 145-52.
3. Treviño, V., E. Martínez-Ledesma, and J. Tamez-Peña, *Identification of outcome-related driver mutations in cancer using conditional co-occurrence distributions.* Sci Rep, 2017. **7**: p. 43350.
4. Catana, C.-S., D. Gulei, and I. Berindan - Neagoe, *New insights into the role of non-coding RNAs as transcriptional targets of p53.* Endogenous locus-driven H-Ras G12V expression induces senescence-like phenotype in primary fibroblasts of the Costello syndrome mouse model, 2017: p. 43-49.
5. Hafner, A., et al., *The multiple mechanisms that regulate p53 activity and cell fate.* Nat Rev Mol Cell Biol, 2019. **20**(4): p. 199-210.
6. Harris, S.L. and A.J. Levine, *The p53 pathway: positive and negative feedback loops.* Oncogene, 2005. **24**(17): p. 2899-908.
7. Appella, E. and C.W. Anderson, *Post-translational modifications and activation of p53 by genotoxic stresses.* Eur J Biochem, 2001. **268**(10): p. 2764-72.
8. Levine, A.J., *p53: 800 million years of evolution and 40 years of discovery.* Nat Rev Cancer, 2020. **20**(8): p. 471-480.
9. Oliner, J.D., et al., *Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53.* Nature, 1993. **362**(6423): p. 857-60.
10. Haupt, Y., et al., *Mdm2 promotes the rapid degradation of p53.* Nature, 1997. **387**(6630): p. 296-9.
11. Yin, Y., et al., *p53 Stability and activity is regulated by Mdm2-mediated induction of alternative p53 translation products.* Nat Cell Biol, 2002. **4**(6): p. 462-7.
12. Gu, W. and R.G. Roeder, *Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain.* Cell, 1997. **90**(4): p. 595-606.
13. Guttman, M., et al., *Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals.* Nature, 2009. **458**(7235): p. 223-7.
14. Jin, S., et al., *p53-targeted lincRNA-p21 acts as a tumor suppressor by inhibiting JAK2/STAT3 signaling pathways in head and neck squamous cell carcinoma.* Mol Cancer, 2019. **18**(1): p. 38.
15. Wang, X., et al., *Long intragenic non-coding RNA lincRNA-p21 suppresses development of human prostate cancer.* Cell Prolif, 2017. **50**(2).

16.   Tran, U.M., et al., *LincRNA-p21 acts as a mediator of ING1b-induced apoptosis.* Cell Death Dis, 2015. **6**(3): p. e1668.
17.   Chen, S., et al., *LincRNa-p21: function and mechanism in cancer.* Med Oncol, 2017. **34**(5): p. 98.
18.   Tang, S.S., B.Y. Zheng, and X.D. Xiong, *LincRNA-p21: Implications in Human Diseases.* Int J Mol Sci, 2015. **16**(8): p. 18732-40.
19.   Kesheh, M.M., S. Mahmoudvand, and S. Shokri, *Long noncoding RNAs in respiratory viruses: A review.* Rev Med Virol, 2021: p. e2275.
20.   Huarte, M., et al., *A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response.* Cell, 2010. **142**(3): p. 409-19.
21.   Chillón, I. and A.M. Pyle, *Inverted repeat Alu elements in the human lincRNA-p21 adopt a conserved secondary structure that regulates RNA function.* Nucleic Acids Res, 2016. **44**(19): p. 9462-9471.
22.   Moumen, A., et al., *hnRNP K: an HDM2 target and transcriptional coactivator of p53 in response to DNA damage.* Cell, 2005. **123**(6): p. 1065-78.
23.   Sun, X., M.S.S. Haider Ali, and M. Moran, *The role of interactions of long non-coding RNAs and heterogeneous nuclear ribonucleoproteins in regulating cellular functions.* The Biochemical journal, 2017. **474**(17): p. 2925-2935.
24.   Baldassarre, A. and A. Masotti, *Long non-coding RNAs and p53 regulation.* Int J Mol Sci, 2012. **13**(12): p. 16708-17.
25.   Arcot, S.S., et al., *Alu repeats: a source for the genesis of primate microsatellites.* Genomics, 1995. **29**(1): p. 136-44.
26.   Batzer, M.A. and P.L. Deininger, *Alu repeats and human genomic diversity.* Nat Rev Genet, 2002. **3**(5): p. 370-9.
27.   Tajaddod, M., et al., *Transcriptome-wide effects of inverted SINEs on gene expression and their impact on RNA polymerase II activity.* Genome Biol, 2016. **17**(1): p. 220.
28.   Deininger, P.L. and M.A. Batzer, *Alu repeats and human disease.* Mol Genet Metab, 1999. **67**(3): p. 183-93.
29.   Payer, L.M., et al., *Structural variants caused by Alu insertions are associated with risks for many human diseases.* Proc Natl Acad Sci U S A, 2017. **114**(20): p. E3984-e3992.
30.   Polak, P. and E. Domany, *Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes.* BMC Genomics, 2006. **7**: p. 133.
31.   Novikova, I.V., S.P. Hennelly, and K.Y. Sanbonmatsu, *Sizing up long non-coding RNAs: do lncRNAs have secondary and tertiary structure?* Bioarchitecture, 2012. **2**(6): p. 189-199.
32.   Chillón, I. and M. Marcia, *The molecular structure of long non-coding RNAs: emerging patterns and functional implications.* Crit Rev Biochem Mol Biol, 2020. **55**(6): p. 662-690.
33.   Lusvarghi, S., et al., *RNA secondary structure prediction using high-throughput SHAPE.* J Vis Exp, 2013(75): p. e50243.
34.   Boniecki, M.J., et al., *SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction.* Nucleic Acids Res, 2016. **44**(7): p. e63.
35.   Conrad, T., et al., *Maximizing transcription of nucleic acids with efficient T7 promoters.* Commun Biol, 2020. **3**(1): p. 439.
36.   Mrozowich, T., et al., *Nanoscale Structure Determination of Murray Valley Encephalitis and Powassan Virus Non-Coding RNAs.* Viruses, 2020. **12**(2).
37.   Chillón, I., et al., *Native Purification and Analysis of Long RNAs.* Methods Enzymol, 2015. **558**: p. 3-37.
38.   Beckert, B. and B. Masquida, *Synthesis of RNA by in vitro transcription.* Methods Mol Biol, 2011. **703**: p. 29-41.
39.   Chen, Z. and Y. Zhang, *Dimethyl sulfoxide targets phage RNA polymerases to promote transcription.* Biochem Biophys Res Commun, 2005. **333**(3): p. 664-70.
40.   McKenna, S.A., et al., *Purification and characterization of transcribed RNAs using gel filtration chromatography.* Nat Protoc, 2007. **2**(12): p. 3270-7.
41.   Some, D., et al., *Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS).* J Vis Exp, 2019(148).
42.   Pam Wang, R.A., Michelle Chen, and Kristine Legaspi, *AN1616: SEC-MALS Method for Characterizing mRNA Biophysical Attributes.* Wyatt Technologies Moderna Therapeutics 2020.
43.   Pam Wang, R.A., Michelle Chen, and Kristine Legaspi, *SEC-MALS Method for Characterizing mRNA Biophysical Attributes.* LCGC 2020.

44.   Patel, T.R., et al., *Structural studies of RNA-protein complexes: A hybrid approach involving hydrodynamics, scattering, and computational methods.* Methods, 2017. **118-119**: p. 146-162.

45.   Wyatt, P.J., *Measurement of special nanoparticle structures by light scattering.* Anal Chem, 2014. **86**(15): p. 7171-83.

46.   Wyatt, P.J., *Measuring nanoparticles in the size range to 2000 nm.* Journal of nanoparticle research : an interdisciplinary forum for nanoscale science and technology, 2018. **20**(12): p. 322-322.

47.   Demeler, B. and G.E. Gorbet, *Analytical Ultracentrifugation Data Analysis with UltraScan-III*, in *Analytical Ultracentrifugation: Instrumentation, Software, and Applications*, S. Uchiyama, et al., Editors. 2016, Springer Japan: Tokyo. p. 119-143.

48.   Demeler, B., *Methods for the design and analysis of sedimentation velocity and sedimentation equilibrium experiments with proteins.* Curr Protoc Protein Sci, 2010. **Chapter 7**: p. Unit 7.13.

49.   Brookes, E., W. Cao, and B. Demeler, *A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape.* Eur Biophys J, 2010. **39**(3): p. 405-14.

50.   Brookes, E.H. and B. Demeler, *Parsimonious regularization using genetic algorithms applied to the analysis of analytical ultracentrifugation experiments*, in *Proceedings of the 9th annual conference on Genetic and evolutionary computation*. 2007, Association for Computing Machinery: London, England. p. 361–368.

51.   Demeler, B. and E. Brookes, *Monte Carlo analysis of sedimentation experiments.* Colloid and Polymer Science, 2008. **286**(2): p. 129-137.

52.   Cowieson, N.P., et al., *Beamline B21: high-throughput small-angle X-ray scattering at Diamond Light Source.* J Synchrotron Radiat, 2020. **27**(Pt 5): p. 1438-1446.

53.   Manalastas-Cantos, K., et al., *ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis.* J Appl Crystallogr, 2021. **54**(Pt 1): p. 343-355.

54.   Panjkovich, A. and D.I. Svergun, *CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data.* Bioinformatics, 2018. **34**(11): p. 1944-1946.

55.   Putnam, C.D., *Guinier peak analysis for visual and automated inspection of small-angle X-ray scattering data.* J Appl Crystallogr, 2016. **49**(Pt 5): p. 1412-1419.

56.   Burke, J.E. and S.E. Butcher, *Nucleic acid structure characterization by small angle X-ray scattering (SAXS).* Current protocols in nucleic acid chemistry, 2012. **Chapter 7**: p. Unit7.18-Unit7.18.

57.   Semenyuk, A.V. and D.I. Svergun, *GNOM– a program package for small-angle scattering data processing.* Journal of Applied Crystallography, 1991. **24**(5): p. 537-540.

58.   Svergun, D.I., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria.* Journal of Applied Crystallography, 1992. **25**(4): p. 495-503.

59.   Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophys J, 1999. **76**(6): p. 2879-86.

60.   Volkov, V.V. and D.I. Svergun, *Uniqueness of ab initio shape determination in small-angle scattering.* Journal of Applied Crystallography, 2003. **36**(3-1): p. 860-864.

61.   Ortega, A., D. Amorós, and J. García de la Torre, *Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models.* Biophys J, 2011. **101**(4): p. 892-8.

62.   Stasiewicz, J., et al., *QRNAS: software tool for refinement of nucleic acid structures.* BMC Structural Biology, 2019. **19**(1): p. 5.

63.   Kozin, M.B. and D.I. Svergun, *Automated matching of high- and low-resolution structural models.* Journal of Applied Crystallography, 2001. **34**(1): p. 33-41.

64.   Brosey, C.A. and J.A. Tainer, *Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology.* Curr Opin Struct Biol, 2019. **58**: p. 197-213.

65.   Pérez, J. and P. Vachette, *A Successful Combination: Coupling SE-HPLC with SAXS.* Adv Exp Med Biol, 2017. **1009**: p. 183-199.

66.   O'Brien, D.P., et al., *SEC-SAXS and HDX-MS: A powerful combination. The case of the calcium-binding domain of a bacterial toxin.* Biotechnol Appl Biochem, 2018. **65**(1): p. 62-68.

67.   Gräwert, M., et al., *Adding Size Exclusion Chromatography (SEC) and Light Scattering (LS) Devices to Obtain High-Quality Small Angle X-Ray Scattering (SAXS) Data.* Crystals, 2020. **10**: p. 975.

68.   Rambo, R.P. and J.A. Tainer, *Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law.* Biopolymers, 2011. **95**(8): p. 559-571.

69.   Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution.* Q Rev Biophys, 2007. **40**(3): p. 191-285.

70.   Kikhney, A.G. and D.I. Svergun, *A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins.* FEBS Lett, 2015. **589**(19 Pt A): p. 2570-7.

71.   Vadlamani, G., et al., *The β-lactamase gene regulator AmpR is a tetramer that recognizes and binds the D-Ala-D-Ala motif of its repressor UDP-N-acetylmuramic acid (MurNAc)-pentapeptide.* J Biol Chem, 2015. **290**(5): p. 2630-43.

72.   Vandivier, L.E., et al., *The Conservation and Function of RNA Secondary Structure in Plants.* Annual review of plant biology, 2016. **67**: p. 463-488.

73.   Chełkowska-Pauszek, A., et al., *The Role of RNA Secondary Structure in Regulation of Gene Expression in Bacteria.* International journal of molecular sciences, 2021. **22**(15): p. 7845.

74.   Lorenz, R., et al., *Predicting RNA secondary structures from sequence and probing data.* Methods, 2016. **103**: p. 86-98.

75.   Chen, J.L., S. Bellaousov, and D.H. Turner, *RNA Secondary Structure Determination by NMR.* Methods Mol Biol, 2016. **1490**: p. 177-86.

76.   Kenyon, J., L. Prestwood, and A. Lever, *Current perspectives on RNA secondary structure probing.* Biochem Soc Trans, 2014. **42**(4): p. 1251-5.

77.   Tahi, F., T.T.V. Du, and A. Boucheham, *In Silico Prediction of RNA Secondary Structure.* Methods Mol Biol, 2017. **1543**: p. 145-168.

78.   Patel, B.A., et al., *Multi-angle light scattering as a process analytical technology measuring real-time molecular weight for downstream process control.* mAbs, 2018. **10**(7): p. 945-950.

79.   Stetefeld, J., S.A. McKenna, and T.R. Patel, *Dynamic light scattering: a practical guide and applications in biomedical sciences.* Biophys Rev, 2016. **8**(4): p. 409-427.

80.   Gräwert, T.W. and D.I. Svergun, *Structural Modeling Using Solution Small-Angle X-ray Scattering (SAXS).* J Mol Biol, 2020. **432**(9): p. 3078-3092.

81.   Svergun, D.I., *Small-angle X-ray and neutron scattering as a tool for structural systems biology.* Biol Chem, 2010. **391**(7): p. 737-43.

82.   Grant, T.D., et al., *The accurate assessment of small-angle X-ray scattering data.* Acta Crystallogr D Biol Crystallogr, 2015. **71**(Pt 1): p. 45-56.

83.   Cantara, W.A., E.D. Olson, and K. Musier-Forsyth, *Analysis of RNA structure using small-angle X-ray scattering.* Methods, 2017. **113**: p. 46-55.

84.   Zhang, Y., et al., *Long non-coding subgenomic flavivirus RNAs have extended 3D structures and are flexible in solution.* EMBO Rep, 2019. **20**(11): p. e47016.

85.   Cantero-Camacho, Á., et al., *Three-dimensional structure of the 3'X-tail of hepatitis C virus RNA in monomeric and dimeric states.* Rna, 2017. **23**(9): p. 1465-1476.

86.   Nelson, C.R., et al., *Human DDX17 Unwinds Rift Valley Fever Virus Non-Coding RNAs.* Int J Mol Sci, 2020. **22**(1).

87.   Choi, K.H. and M. Morais, *Use of small-angle X-ray scattering to investigate the structure and function of dengue virus NS3 and NS5.* Methods in molecular biology (Clifton, N.J.), 2014. **1138**: p. 241-252.

88.   Misra, V.K. and D.E. Draper, *On the role of magnesium ions in RNA stability.* Biopolymers, 1998. **48**(2-3): p. 113-35.

89.   Misra, V.K. and D.E. Draper, *The linkage between magnesium binding and RNA folding.* J Mol Biol, 2002. **317**(4): p. 507-21.

90.   Yamagami, R., J.P. Sieg, and P.C. Bevilacqua, *Functional Roles of Chelated Magnesium Ions in RNA Folding and Function.* Biochemistry, 2021. **60**(31): p. 2374-2386.

91.   Carrasco, B. and J. García de la Torre, *Hydrodynamic properties of rigid particles: comparison of different modeling and computational procedures.* Biophys J, 1999. **76**(6): p. 3044-57.

92.   Fernandes, M.X., et al., *Calculation of hydrodynamic properties of small nucleic acids from their atomic structure.* Nucleic Acids Res, 2002. **30**(8): p. 1782-8.

93.   García De La Torre, J., M.L. Huertas, and B. Carrasco, *Calculation of hydrodynamic properties of globular proteins from their atomic-level structure.* Biophys J, 2000. **78**(2): p. 719-30.

94.   Garcia de la Torre, J.G. and V.A. Bloomfield, *Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications.* Q Rev Biophys, 1981. **14**(1): p. 81-139.

95.   Bomsztyk, K., O. Denisenko, and J. Ostrowski, *hnRNP K: one protein multiple processes.* Bioessays, 2004. **26**(6): p. 629-38.

96.   Xu, Y., et al., *New Insights into the Interplay between Non-Coding RNAs and RNA-Binding Protein HnRNPK in Regulating Cellular Functions.* Cells, 2019. **8**(1): p. 62.

97.   Makeyev, A.V. and S.A. Liebhaber, *The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms.* Rna, 2002. **8**(3): p. 265-78.

Chapter 5: Conclusions

5.1 Overview

By characterising the host-pathogen relationships of viral Noncoding RNA and the biophysical properties of human Noncoding RNA, we're able to define and explicate their functions that contribute towards infectious disease and cancer progression. The work presented in this thesis observes the host-pathogen interactions between human RNA-Binding Proteins and Hantaviral Noncoding Terminal Regions, and the biophysical characterisation of previously identified secondary structures of human Long Noncoding RNA [1]. Previous studies have identified the importance of ssRNA Noncoding Terminal Regions and the vital roles they perform to maintain viral lifecycles, including transcription and replication, translation initiation, packaging, and trafficking within the cytoplasm [2, 3]. In Hantaviruses, especially other viruses within the *Bunyaviridae* order, the Noncoding Terminal Regions are critical for circularisation and the formation of panhandle structures as is the case with Bunyamwera Virus, Rift Valley Fever Virus, and Arenaviruses [4-6]. These panhandle regions and their extended Noncoding Terminal Regions appear to retain specialised promoters and requisite secondary structures for specific interactions with viral RNA-dependent RNA-polymerases and the Nucleocapsid Protein [7].

Emerging viruses are continuously increasing and their impact on human health and infrastructure is a global concern. As Hantaviruses are considered emerging viruses, new taxa and species are becoming known regularly, and can introduce a myriad of health risks to humans, including animals such as livestock as is the case with other Bunaviruses [8, 9]. In Chapter 2, we outlined the characteristics of New-World Hantaviruses, particularly Sin Nombre *orthohantavirus* and Andes Virus, as potential bioweapons [10]. This results from these pathogenic New-World Hantaviruses causing Hantavirus Cardiopulmonary Syndrome, a severe infectious disease that retains a high mortality rate and severe disease progression [11]. Despite occurring infrequently in North America, primarily due to developed infrastructure and the rigid separation of human dwellings from rodent reservoir hosts, Hantavirus Cardiopulmonary Syndrome causing Hantaviruses are an ongoing risk to developing nations in Central and South America [12-15]. We also showed that its high mortality rate, transmissibility from rodents to humans (and Andes Virus human-to-human transmission), durability, prevalence in nature, and aerosolised properties make it an attractive agent for biowarfare [16-18]. Compounded with climate change and the effects of war and climate migrations on populations and infrastructure, Hantaviral agents can be deployed to create economic and civil disruptions, requiring heightened states

of security and increased sanitation measures [19-21]. However, with culturing and passaging limitations, New-World Hantavirus agents remain at the behest of improved technologies and dispersion systems [22].

We continued the Hantaviral narrative by investigating potential RNA-Binding Proteins that likely impact the Hantaviral lifecycle. In Chapter 3, we observed a variety of RNA-Binding Proteins that presumably interacts with the 3'-Noncoding Terminal Regions of the S and M Segments of Sin Nombre *orthohantavirus*. We identified Poly(rC)-binding proteins 1 and 2 and Y-box-binding proteins 1 and 3 involved in sense and antisense S3TR interactions. These proteins are widely used for mRNA stabilisation and the initiation of viral replication, noted in interactions of other ssRNA viruses and their Noncoding Terminal Regions such as Poliovirus and Dengue Virus [23-25]. Other proteins identified for the sense and antisense M3TR were Heat Shock Cognate 71 kDa Protein and Double-stranded RNA-binding Protein Staufen Homolog 1 which are involved in RNA molecular chaperone activities and viral RNA replication. Staufen Homolog 1 particularly is involved in binding interactions with similar negative-sense ssRNA viruses, like Ebola Virus and their 3'-/5'-terminal regions resulting in an increase in viral replication [26]. It is likely then that these RNA-Binding Proteins increase the stability of viral RNAs enabling them to undergo translation and replication more effectively, especially when coordinated with their own viral proteins. Hantavirus terminal regions are also highly conserved, which additionally extends to other Bunyaviruses whose Noncoding Terminal regions are highly conserved within each genus [27]. This suggests that RNA-Binding Proteins identified for Sin Nombre orthohantavirus Noncoding Terminal Regions will likely interact similarly with other pathogenic Hantaviruses such as Andes Virus, Hantaan Virus, and Puumala Virus. This can have profound impacts on developing universal and comprehensive therapeutics against infectious viruses within the same genus, as drug targets against one will likely elicit the same response in other *orthohantaviruses*.

To further understand the role structure plays in noncoding RNAs, we analysed the Long Intergenic Noncoding RNA-p21 *AluSx1* Inverted Repeat elements which have been identified to regulate cancer progression and apoptosis through its binding of Heterogeneous nuclear ribonucleoprotein K [1]. Our work involves several biophysical characterisation techniques that evaluated and optimised purified RNA using an integrative approach with Multiangle Light Scattering-Dynamic Light Scattering and Analytical Ultracentrifugation techniques to confidently assess the homogeneity and fidelity of *in vitro* transcribed RNA [28-30]. Employing Small Angle X-ray Scattering, we generated space-filling low-resolution, three-dimensional models of sense and antisense *AluSx1* Inverted Repeat RNA which has been identified to regulate Heterogeneous nuclear ribonucleoprotein K nuclear localisation and

subsequent repression of p53 genes [1, 31, 32]. In collaboration with Dr. Micheal Wolfinger, we performed atomistic RNA modelling that attained a high degree of conformity with experimentally determined solution scattering [33, 34]. These models additionally match with secondary structure models previously determined by chemical probing techniques making this pipeline applicable to the study of other viral and human Long Noncoding RNA structures.

To conclude, my thesis work presents potential human RNA-Binding Proteins that likely interact with Sin Nombre *orthohantavirus* S and M Segment's 3'-Noncoding Terminal Regions. Additionally, we have biophysically characterised LincRNA-p21 sense and antisense *AluSx1* Inverted Repeat RNAs, which have been shown to regulate cellular apoptosis and cancer progression. Studying the structure and biophysical properties of Noncoding RNAs are essential in building a foundation towards understanding holistically the mechanisms of their interactions which have significant weight in biological processes. Finally, identifying potential RNA-Binding Proteins that interact meaningfully within the Hantaviral lifecycle can introduce new avenues towards combatting Hantavirus infections which are highly pronounced and fatal, especially in developing countries which accounts for thousands of infections every year [35-37].

5.2 Future Perspectives

As emphasised, identifying RNA-Binding Proteins is not sufficient in building a framework of Host-Pathogen interactions within the study of infectious diseases. Emerging viruses, like HCPS-causing North American Hantaviruses, should be investigated promptly and defined thoroughly using multidisciplinary approaches. Consequently, conducting intensive biophysical characterisation studies are needed to explain in detail complex mechanisms and requirements viral systems undergo to maintain themselves. Advancing on current immunoprecipitation reactions, *in vivo* experiments such as utilising transfected MANGO aptamers or alternate *in vitro* Biotin-Streptavidin pull-down assays can supplement previous pull-down assays towards developing a concrete and foundational list of Hantaviral RNA-Binding Proteins [38-43]. Furthermore, visiting alternate pull-down strategies can provide increased confidence whilst supplementing and minimising currently known RBP lists to highly specific or novel proteins and should include the other Noncoding Terminal Regions of the S, M, and L Segments.

Biophysical characterisation studies of SNV Hantaviral RNA can be improved with additional methods to develop orthogonal and overlapping approaches to confirm the interactions of RNA-Binding Proteins once identified. Observing the functions and specificity of Hantaviral RNA-Protein domains can be investigated using chemical

probing techniques to determine secondary structure essential in the Hantaviruses' lifecycles. Selective 2′-Hydroxyl Acylation Analysed by Primer Extension (SHAPE) will develop the Noncoding Terminal Regions of Hantaviruses beyond their highly conserved panhandle structures, revealing and identifying additional binding domains that could be involved in activities present in the functions of Long Noncoding RNAs [44-47]. This especially applies to biological processes that advance or mitigate disease progression during Hantaviral infections and is necessary since no low- to high-resolution structural information exists for their Noncoding RNA. Analysing the experimentally predicted secondary structures can follow the pipeline we have established in Chapter 4. The application of low- and high-resolution solution scattering techniques such as Small Angle X-ray Scattering, Cryo-Electron Microscopy, and Nuclear Magnetic Resonance, with the addition of computational folding software can produce structural models that can verify chemical probing techniques and depict native RNA structure as it occurs in solution [48-51].

As mentioned, important validation steps utilising Microscale Thermophoresis and Electrophoretic Mobility Shift Assays will be necessary to quantifying specific RNA-Binding Protein interactions [52-56]. Protein expression, purification, and RNA-Protein Binding studies will be carried out by Scott Tersteeg and Mitchell Geeraert during their MSc programs to validate the immunoprecipitated proteins identified herein for S and M3TR. Future studies will also include the visualisation of SHAPE dependent secondary structure determinations to render high-resolution, atomistic models using SimRNA techniques applied herein for Hantaviral NTRs in collaboration with Dr. Wolfinger and the University of Vienna. Together, these approaches can build a picture of surface level interactions that will be foundational towards the understanding and structural behaviour of Hantaviral terminal regions. By comprehending these Hantaviral terminal regions and their structure-functional role in the Hantaviral lifecycle thought their RBP interactions, we can target structures of opportunity and RBP relationships towards the development of future antivirals, medical countermeasures, and vaccines against Hantavirus infections.

5.3 References

1. Chillón, I. and A.M. Pyle, *Inverted repeat Alu elements in the human lincRNA-p21 adopt a conserved secondary structure that regulates RNA function.* Nucleic Acids Res, 2016. **44**(19): p. 9462-9471.
2. Liu, Y., et al., *Structures and Functions of the 3' Untranslated Regions of Positive-Sense Single-Stranded RNA Viruses Infecting Humans and Animals.* Frontiers in Cellular and Infection Microbiology, 2020. **10**(453).
3. Brinton, M.A. and M. Basu, *Functions of the 3' and 5' genome RNA regions of members of the genus Flavivirus.* Virus Res, 2015. **206**: p. 108-19.
4. Barr, J.N. and G.W. Wertz, *Role of the conserved nucleotide mismatch within 3'- and 5'-terminal regions of Bunyamwera virus in signaling transcription.* Journal of virology, 2005. **79**(6): p. 3586-3594.

5.   Gauliard, N., et al., *Rift Valley fever virus noncoding regions of L, M and S segments regulate RNA synthesis.* Virology, 2006. **351**(1): p. 170-179.

6.   Brisse, M.E. and H. Ly, *Hemorrhagic Fever-Causing Arenaviruses: Lethal Pathogens and Potent Immune Suppressors.* Frontiers in Immunology, 2019. **10**(372).

7.   Sun, Y., et al., *Bunyavirales ribonucleoproteins: the viral replication and transcription machinery.* Critical Reviews in Microbiology, 2018. **44**(5): p. 522-540.

8.   Abudurexiti, A., et al., *Taxonomy of the order Bunyavirales: update 2019.* Arch Virol, 2019. **164**(7): p. 1949-1965.

9.   Dutuze, M.F., et al., *A Review of Bunyamwera, Batai, and Ngari Viruses: Understudied Orthobunyaviruses With Potential One Health Implications.* Front Vet Sci, 2018. **5**: p. 69.

10.   D'Souza, M.H. and T.R. Patel, *Biodefense Implications of New-World Hantaviruses.* Frontiers in bioengineering and biotechnology, 2020. **8**: p. 925-925.

11.   Borges, A.A., et al., *Hantavirus cardiopulmonary syndrome: immune response and pathogenesis.* Microbes Infect, 2006. **8**(8): p. 2324-30.

12.   Figueiredo, L.T., et al., *Hantaviruses and cardiopulmonary syndrome in South America.* Virus Res, 2014. **187**: p. 43-54.

13.   Hjelle, B. and F. Torres-Perez, *Hantaviruses in the americas and their role as emerging pathogens.* Viruses, 2010. **2**(12): p. 2559-86.

14.   Firth, C., et al., *Diversity and distribution of hantaviruses in South America.* J Virol, 2012. **86**(24): p. 13756-66.

15.   Vigueras-Galvan, A.L., et al., *Current Situation and Perspectives on Hantaviruses in Mexico.* Viruses, 2019. **11**(7).

16.   Alonso, D.O., et al., *Person-to-Person Transmission of Andes Virus in Hantavirus Pulmonary Syndrome, Argentina, 2014.* Emerg Infect Dis, 2020. **26**(4): p. 756-759.

17.   Bellomo, C., et al., *A newborn infected by Andes virus suggests novel routes of hantavirus transmission: a case report.* Clin Microbiol Infect, 2020. **26**(1): p. 130-131.

18.   Pizarro, E., et al., *Immunocytochemical and Ultrastructural Evidence Supporting That Andes Hantavirus (ANDV) Is Transmitted Person-to-Person Through the Respiratory and/or Salivary Pathways.* Front Microbiol, 2019. **10**: p. 2992.

19.   Guterres, A. and E.R.S. de Lemos, *Hantaviruses and a neglected environmental determinant.* One Health, 2018. **5**: p. 27-33.

20.   Klempa, B., *Hantaviruses and climate change.* Clin Microbiol Infect, 2009. **15**(6): p. 518-23.

21.   Hansen, A., et al., *Transmission of haemorrhagic fever with renal syndrome in china and the role of climate factors: a review.* Int J Infect Dis, 2015. **33**: p. 212-8.

22.   Clement, J.P., *Hantavirus.* Antiviral Res, 2003. **57**(1-2): p. 121-7.

23.   Walter, B.L., et al., *Distinct poly(rC) binding protein KH domain determinants for poliovirus translation initiation and viral RNA replication.* J Virol, 2002. **76**(23): p. 12008-22.

24.   Chase, A.J., S. Daijogo, and B.L. Semler, *Inhibition of poliovirus-induced cleavage of cellular protein PCBP2 reduces the levels of viral RNA replication.* J Virol, 2014. **88**(6): p. 3192-201.

25.   Paranjape, S.M. and E. Harris, *Y box-binding protein-1 binds to the dengue virus 3'-untranslated region and mediates antiviral effects.* J Biol Chem, 2007. **282**(42): p. 30497-508.

26.   Fang, J., et al., *Staufen1 Interacts with Multiple Components of the Ebola Virus Ribonucleoprotein and Enhances Viral RNA Synthesis.* mBio, 2018. **9**(5).

27.   Wichgers Schreur, P.J., R. Kormelink, and J. Kortekaas, *Genome packaging of the Bunyavirales.* Curr Opin Virol, 2018. **33**: p. 151-155.

28.   Patel, T.R., D.J. Winzor, and D.J. Scott, *Analytical ultracentrifugation: A versatile tool for the characterisation of macromolecular complexes in solution.* Methods, 2016. **95**: p. 55-61.

29.   Stetefeld, J., S.A. McKenna, and T.R. Patel, *Dynamic light scattering: a practical guide and applications in biomedical sciences.* Biophys Rev, 2016. **8**(4): p. 409-427.

30.   Patel, T.R., et al., *Structural studies of RNA-protein complexes: A hybrid approach involving hydrodynamics, scattering, and computational methods.* Methods, 2017. **118-119**: p. 146-162.

31.   Huarte, M., et al., *A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response.* Cell, 2010. **142**(3): p. 409-19.

32.   Moumen, A., et al., *hnRNP K: an HDM2 target and transcriptional coactivator of p53 in response to DNA damage.* Cell, 2005. **123**(6): p. 1065-78.

33. Lorenz, R., et al., *Predicting RNA secondary structures from sequence and probing data.* Methods, 2016. **103**: p. 86-98.

34. Boniecki, M.J., et al., *SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction.* Nucleic Acids Res, 2016. **44**(7): p. e63.

35. Jiang, H., et al., *Hemorrhagic Fever with Renal Syndrome: Pathogenesis and Clinical Picture.* Front Cell Infect Microbiol, 2016. **6**: p. 1.

36. Noh, J.Y., J. Jung, and J.W. Song, *Hemorrhagic Fever with Renal Syndrome.* Infect Chemother, 2019. **51**(4): p. 405-413.

37. Llah, S.T., et al., *Hantavirus induced cardiopulmonary syndrome: A public health concern.* J Med Virol, 2018. **90**(6): p. 1003-1009.

38. Autour, A., et al., *Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells.* Nat Commun, 2018. **9**(1): p. 656.

39. Dolgosheina, E.V., et al., *RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking.* ACS Chem Biol, 2014. **9**(10): p. 2412-20.

40. Panchapakesan, S.S., S.C. Jeng, and P.J. Unrau, *RNA complex purification using high-affinity fluorescent RNA aptamer tags.* Ann N Y Acad Sci, 2015. **1341**: p. 149-55.

41. Panchapakesan, S.S.S., et al., *Ribonucleoprotein purification and characterization using RNA Mango.* Rna, 2017. **23**(10): p. 1592-1599.

42. Gemmill, D., et al., *Current approaches for RNA-labelling to identify RNA-binding proteins.* Biochem Cell Biol, 2020. **98**(1): p. 31-41.

43. McQuaid, S., J. McMahon, and G.M. Allan, *A Comparison of Digoxigenin and Biotin Labelled DNA and RNA Probes for in Situ Hybridization.* Biotechnic & Histochemistry, 1995. **70**(3): p. 147-154.

44. Lusvarghi, S., et al., *RNA secondary structure prediction using high-throughput SHAPE.* J Vis Exp, 2013(75): p. e50243.

45. Smola, M.J., et al., *Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis.* Nat Protoc, 2015. **10**(11): p. 1643-69.

46. Watters, K.E., et al., *Characterizing RNA structures in vitro and in vivo with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq).* Methods, 2016. **103**: p. 34-48.

47. Wilkinson, K.A., E.J. Merino, and K.M. Weeks, *Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution.* Nat Protoc, 2006. **1**(3): p. 1610-6.

48. Chen, Y. and L. Pollack, *SAXS studies of RNA: structures, dynamics, and interactions with partners.* Wiley Interdiscip Rev RNA, 2016. **7**(4): p. 512-26.

49. Jeffries, C.M., et al., *Preparing monodisperse macromolecular samples for successful biological small-angle X-ray and neutron-scattering experiments.* Nat Protoc, 2016. **11**(11): p. 2122-2153.

50. Mrozowich, T., et al., *Structural Studies of Macromolecules in Solution using Small Angle X-Ray Scattering.* J Vis Exp, 2018(141).

51. Théobald-Dietrich, A., et al., *Structural Analysis of RNA by Small-Angle X-ray Scattering.* Methods Mol Biol, 2020. **2113**: p. 189-215.

52. Mrozowich, T., V. MeierStephenson, and T.R. Patel, *Microscale thermophoresis: warming up to a new biomolecular interaction technique.* The Biochemist, 2019. **41**(2): p. 8-12.

53. Jacob, D., et al., *Absolute Quantification of Noncoding RNA by Microscale Thermophoresis.* Angew Chem Int Ed Engl, 2019. **58**(28): p. 9565-9569.

54. Moon, M.H., et al., *Measuring RNA-Ligand Interactions with Microscale Thermophoresis.* Biochemistry, 2018. **57**(31): p. 4638-4643.

55. Ream, J.A., L.K. Lewis, and K.A. Lewis, *Rapid agarose gel electrophoretic mobility shift assay for quantitating protein: RNA interactions.* Anal Biochem, 2016. **511**: p. 36-41.

56. Rio, D.C., *Electrophoretic mobility shift assays for RNA-protein complexes.* Cold Spring Harb Protoc, 2014. **2014**(4): p. 435-40.