Arenaviral nucleoproteins and the identification of their cellular interacting partners

Andrew Dennis Buckley BSc (Hons)

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Molecular and Cellular Biology

June 2016

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Andrew Dennis Buckley to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

<u>Acknowledgements</u>

When I started my PhD, I remember being more nervous than I can ever remember. The support of my supervisor, John, has made this project really very enjoyable, and makes those nerves back then seem so silly now. Thank you for all your support, encouragement and, above all, belief that it would come good in the end.

Working in 8.61 has been great, and everyone always made all the tough moments that much easier. Countless hilarious moments has always made coming into work less of a chore, and more like seeing some great friends everyday. If it hadn't been for all Emma's broken centrifuges, Hessa's mobile, Joe's puns or Carsten's grumbles, this would have been a much duller, and less caffeine fuelled, 3 years. Becky Surtees has to get a special mention, for without Becky, half the work done here would've seemed impossible without her advice.

Thank you most of all to my friends family, and especially Helen, who was always there to help me when I needed it most, and could always think of a solution when I'd nearly given up. And t' Christopher Paul Ward, who now owes me a pint.

Finally, being free to volunteer to help in the Ebola outbreak response was an experience I will never forget, nor wish to. I am eternally grateful to John and the graduate school in letting me go to Sierra Leone to help such a fantastic country get back to its feet. Thanks to Fabulous Team Five, especially Nicole Zitzmann as my travel buddy and her excellent leadership, and my stranded buddy Becky.

Na migo don Ebola.

Abstract

Members of the *Arenaviridae* family are able to cause severe disease in humans. The members of this family which induce haemorrhagic fever (HF) are responsible for over 300,000 reported cases per annum, making arenaviruses the largest cause of HF worldwide. Complications from arenavirus induced disease vary from cerebral and developmental complications in infants and the immunocompromised, to HF with cases ultimately fatal in up to 80% of cases. HF causing arenaviruses can induce high grade fever; diarrhoea; vomiting and general malaise, leading to loss of vascular permeability; internal haemorrhaging; disseminated intravascular coagulation and coma. Fundamental mechanisms of the viral replication strategies are currently unknown, it is therefore of paramount importance to increase knowledge in this area in order to identify potential therapeutic strategies. The nucleoprotein (NP) of arenaviruses is essential for viral replication, and is thus a good candidate for further investigation. This study presents the analysis of the host-cell interactome of the HF causing Lujo virus (LUJV) NP, an arenaviral infection associated with 80% mortality. The NP of LUJV was found to associate with members of the translation initiation and elongation complexes by immunoprecipitation and immunofluorescence microscopy. In addition, the NP of the congenital pathogen lymphocytic choriomeningitis virus was found to co-localise with translation associated proteins, with such proteins also identified within viral particles. Finally, it is demonstrated here that the translation of arenaviral-like mRNAs is enhanced in the presence of NP, through a mechanism predicted to be driven by the association of NP with the translation initiation complex eIF4F and the circularisation of mRNAs.

Table of Contents

Table of Figures

Table of Tables

Table of Abbreviations

Abbreviation	Expansion
A	Alanine
ACDP	Advisory Committee on Dangerous Pathogens
AHF	Argentinian haemorrhagic fever
BCA	Bicinchonic acid
BHF	Bolivian haemorrhagic fever
BSA	Bovine serum albumin
BSL4	Biosafety level 4
BUNV	Bunyavirus
CCHFV	Crimean Congo Haemorrhagic Fever Virus
CDC	Centre for Disease Control and Prevention
CFR	Case fatality rate
cRNA	Anti-genomic sense RNA
CTL	Cytotoxic T-lymphocyte
D	Aspartic acid
DENV	Dengue virus
DIC	Disseminated intravascular coagulation
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulphoxide
dsRNA	doulble-stranded RNA
E. coli	*Eschericia coli*
EBOV	Ebola virus
EDTA	Ethylenediaminetetraacetic acid
eEF##	Eukaryotic translation elongation factor ##
EGFP	Enhanced green flourescent protein
eIF##	Eukaryotic translation initiation factor ##
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
FBS	Foetal bovine serum
GFP	Green flourescent protein
GP	Arenavirus glycoprotein
GPC	Arenavirus glycoprotein precursor
GTOV	Guanarito virus
HCV	Hepatitis C virus
HF	Haemorrhagic fever
HSP##	Heat shock protein##
IGR	Intragenic region
IκKε	I-kappa-B-kinase ε
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropol β-D-1-thiogalactopyranoside
IRF#	Interferon regulatory factor 3
ISG	Interferon stimulated gene

IκK	IκB kinase
JUNV	Junin virus
K	Lysine
kDa	kilo Daltons
L	Arenavirus polymerase
LASV	Lassa virus
LCMV	Lymphocytic choriomeningitis virus
LUJV	Lujo virus
MACV	Machupo virus
mM	millimolar
mRNA	Messenger ribonucleic acid
MS	Mass Spectrometry
NP	Arenavirus nucleoprotein
NW	New World serogroup of prenavirus species
ORF	Open reading frame
OW	Old World serogroup of arenavirus species
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PKR	Protein kinase R
PM	Plasma membrane
PVDF	polyvinylidine flouride
RdRp	RNA dependent RNA polymerase
RFP	Red flourescent protein
RIG-I	Retinoic acid-inducible gene 1
RLP#	60S Ribosome subunit protein
RLU	Relative luciferase units
RNP	Ribonuceloprotein
RPS#	40S Ribosome subunit protein
SABV	Sáibia Virus
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SILAC	Stable isotope labelling in cell culture
SSP	Stable signal peptide
ssRNA	single-stranded RNA
TBS	Tris buffered saline
TEMED	N,N,N',N'-tetramethylethylenediamine
tRNA	Transfer ribonucleic acid
VHF	Venezuelan haemorrhagic fever
vRNA	Genomic sense RNA
WHO	World Health Organisation
WWAV	Whitewater Arroyo Virus
Z	Arenavirus matrix protein

1. <u>Introduction</u>

1.1. <u>General Introduction</u>

Arenaviruses represent an important global public health concern, with limited to no therapeutics available to combat them. As neglected tropical diseases, knowledge regarding certain fundamental aspects of their replication cycle requires greater clarity. In particular, it is becoming increasingly clear that the nucleoprotein (NP) of arenaviruses is crucial to multiple stages of the virus replication cycle, outside of the more established roles of RNA genome encapsidation. Lujo virus (LUJV), the most recently emerged human pathogen within the *Arenaviridae* family induces disease associated with haemorrhagic fever and an 80% mortality rate. Better understanding the functions of an arenaviral protein, such as NP, could allow for the identification of targets for future therapeutic attack. Utilising LUJV NP as a model NP for arenavirus NPs in general could allow for the identification of such a therapeutic target, which can then be investigated for other arenavirus NPs, with the aim to identify a key interaction with potential implications across the *Arenaviridae* family.

1.1.1. Discovery of the *Arenaviridae Family*, Lujo virus and Lymphocytic choriomeningitis virus

The *Arenaviridae* family of ambisense RNA viruses consists of over 30 members, with 7 discovered in 2015 alone. Currently 9 arenaviruses are known to cause severe disease, including haemorrhagic fever and cerebral complications in humans and are collectively the largest cause of haemorrhagic fever (HF) in humans per annum (Gryseels et al., 2015; Bisordi et al., 2015; Aqrawi et al., 2015; Lavergne et al., 2015; Li et al., 2015; Hellebuyck et al., 2015; Witkowski et al., 2015; IUMS, 2015; CDC, 2015).

The first arenavirus discovered was lymphocytic choriomeningitis virus (LCMV) after an outbreak of infectious encephalitis in St. Louis, USA in 1933 – discussed in greater detail in section 1.1.4 (Muckenfuss et al., 1934).

Following this discovery, the detection of further novel infections, combined with morphological studies, led to the establishment of a novel taxonomic group initially coined as Arenoviruses (Rowe et al., 1970). The appearance of these viruses under electron microscopy showed prominent, dense granule-like structures within particles, giving the family its name – *Arenaviridae* – 'arena' being the Latin for sand. The most prominent member of the family is Lassa virus (LASV), the causative agent of Lassa fever, which is reported to infect upwards of 300,000 individuals per year, with a high proportion of cases remaining undetected (Monath et al., 1974; Kernéis et al., 2009; Sogoba et al., 2012). The disease severity of Lassa fever varies between outbreaks, but taken together the case fatality rate (CFR) of hospitalised patients is approximately 15-30% in acute outbreaks (Yun & Walker, 2012), with the overall CFR, once asymptomatic infections are included, roughly 1-2%. Annually, LASV is the single largest cause of HF complications, far outstripping the total number of cases during the 2013-16 Ebola virus outbreak (Peterson et al., 2014; World Health Organisation, 2015). All infectious arenaviruses exhibit a very broad tissue tropism, with evidence of viral infection in nearly all organs and tissues (Bird *et al.,* 2012, Schafer *et al.,* 2014).

Species	Serogroup	Distribution
LCMV	OW	Worldwide
LASV	OW	Guinean Africa
LUJV	OW	Zambia
JUNV	NW	Northern Argentina
GTOV	NW	Venezuela
MACV	NW	Bolivia
CHHF	NW	Bolivia
WWAV	NW	Central/Southern Mid-Western USA
SABV	NW	Sau Paulo District Brazil

Table 1 Arenaviruses known to induce severe disease in humans

Each member of the *Arenaviridae* family replicates within a specific resevoir host. These hosts are almost exclusively rodents – with exceptions being the South American Tacaribe virus (TACV), which resides in Caribbean fruit bats (Cogswell-Hawkinson et al., 2012), and a small subset of recently identified arenaviruses that can cause infections within snakes (Aqrawi et al., 2015). This close and sometimes exclusive association between virus and host species often confines each arenavirus within specific geographical regions, meaning that the infections are often endemic within these locations. This geographical restriction potentially curbs their lethality due to the behavioural changes of hosts after interacting with human populations. As human populations grow and expand, the availability of food alters, enabling host rodents to come in closer contact with humans. Mass migration of human populations has been shown to directly drive LASV evolution thorough host selection pressures, discussed in greater detail in section 1.1.3 (Lalis et al., 2012). This viral-rodent co-evolution has been shown to restrict disease in rodents and could explain the far lower CFR of LASV at present, compared to when it first became prominent in the expanding regions of West Africa (Sogoba et al., 2012; Monath et al., 1974; Rowe et al., 1970; Bannister, 2010; Peterson et al., 2014; Zapata & Salvato, 2013). The emergence of new arenaviruses into naïve human populations has caused outbreaks with far higher CFR than is now present in LASV endemic regions, such as Guanarito virus (GTOV) in Venezuela, with CFRs of approximately 15% (Salas et al., 1991; Weaver et al., 2000).

Arenaviruses are certainly neglected tropical diseases, with many aspects of their replication cycle incompletely understood. With increased focus on HF causing infections following the 2013-16 EBOV outbreak, there is a need for an increase in knowledge regarding fundamental aspects of the arenaviral replication cycle. The most recently emerged HF causing arenavirus infection is Lujo virus (LUJV), which was first detected in 2008 in Zambia, following an outbreak of severe febrile illness with associated haemorrhagic shock complications (Sewlall et al., 2014; Briese et al., 2009). The resulting nosocomial outbreak in Johannesburg was associated with a CFR of 80%,

and highlighted the concerns over newly emerging arenaviruses into naïve populations.

1.1.2. Classification

The *Arenaviridae* family, along with the *Orthomyxoviridae*, *Bunyavirdae* and *Ophioviridae* families together form a collection of segmented negative-strand RNA viruses, with the *Bunyaviridae* being the most closely related family. The *Arenaviridae* family constitutes the family with the highest number of HF causing pathogens, with 7 known to cause HF, and LCMV being able to manifest clinically as a cerebral infection involving aseptic meningitis and its associated symptoms.
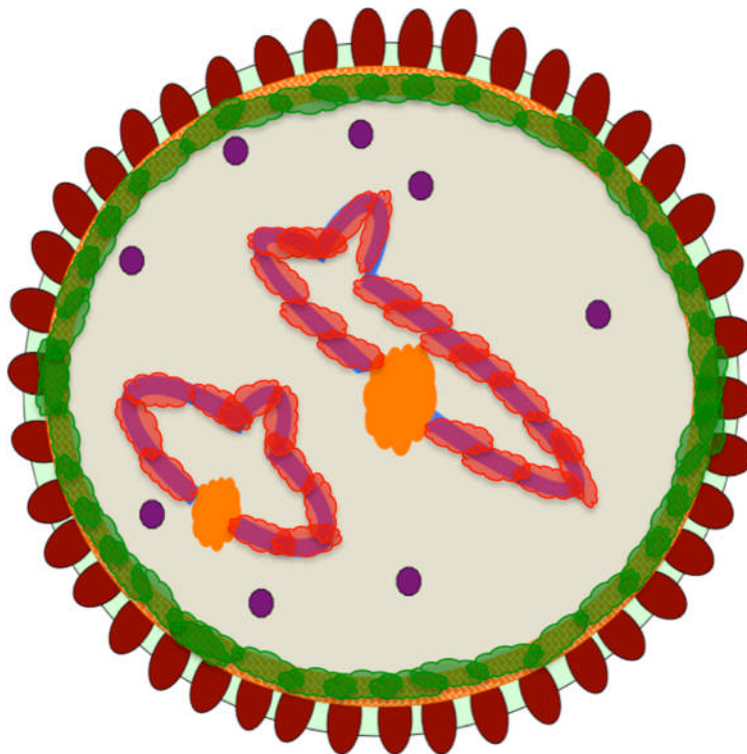
Diagram illustrating a visual representation of an arenaviral particle. Viral ss RNA genome (blue) is encapsided with NP (red) forming an RNP. At the 3' and 5' ends of this ssRNA strand is the L polymerase (orange). A host-derived lipid bilayer (light green and orange) embedded with GP spikes (maroon) bound to the matrix protein Z (green) on the matrix side of the membrane. Host-derived ribosomes have been described to populate the virion, shown in purple.

Figure 1.   **Illustration of virion structure**

Arenaviruses exhibit a cytoplasmic replication strategy. Virions are on average approximately 100nm in diameter, although virion diameter varies

between 80 and 130nm (Murphy & Whitfield, 1975). The "sand" particles which give their name to the family are known to be host-derived ribosomes, represented in figure 1, although these packaged ribosomes are not thought to be essential for viral replication (Rowe et al., 1970; Leung & Rawls, 1977). The four encoded proteins consist of the polymerase (L); matrix protein (Z); glycoprotein – expressed as a precursor protein (GPC) and post translationally modified into a tripartite complex known as GP; and the nucleoprotein (NP).

A defining characteristic of all arenaviruses is their bi-segmented genome. While arenaviruses are classified as negative sense viruses, their coding strategy is in fact ambisense, whereby both the single stranded RNA genome segments consist of both positive and negative sense open reading frames (ORFs), as illustrated in figure 2.
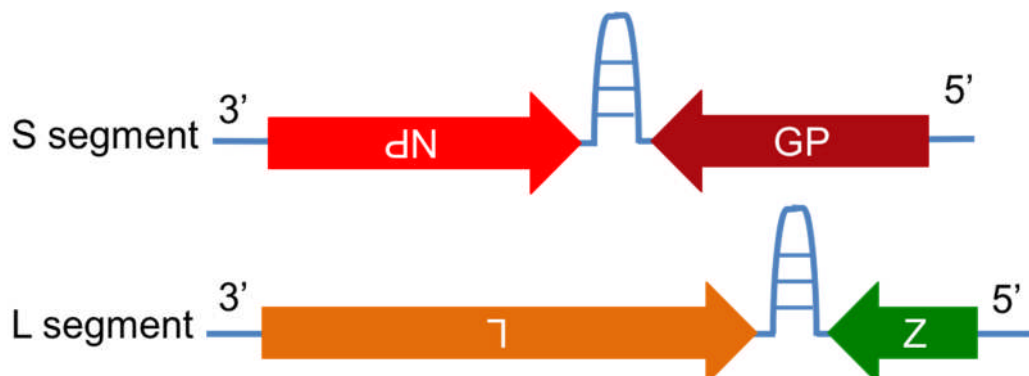
Diagram illustrating the genome structure of the *Arenaviridae*. Ambisence genome segments are shown linearised for diagrammatical purposes. The L segment encodes the polymerase protein L (orange), and the matrix protein Z (green). The S segment encodes the NP (red) and GP (maroon). Inverted letters indicate negative polarity.

Figure 2.   **Illustration of genome structure**

In total, there are over 30 arenavirus species, subdivided into two genera; *Mammarenavirus* and *Reptarenavirus*. The *Mammarenavirus genus* comprises the mammalian host-infecting viruses, whereas the recently described *Reptarenavirus* genus represents reptilian pathogens (IUMS, 2015). The mammarenaviruses are zoonoses, with several known to infect humans. In addition other mammals such as mountain gorillas and a variety of other non-human primates are known to become infected with LASV,

potentially conferring a transmission route through the bushmeat trade (Ogbu et al., 2007). Transmission from the reservoir host is through exposure to infected tissue or excreta, with virions stable in dried rodent urine and faeces for prolonged periods (CDC, 2015). Human-to-human transmission is again through coincidental ingestion or exposure of infected bodily fluids with LASV known to persist in semen for up to 6 weeks (Mary Milazzo et al., 2011; Kernéis et al., 2009).

The mammarenaviruses are subdivided into two geographically limited clades – the Old World arenaviruses (OW) originating from Africa, Europe and Asia; and the New World arenaviruses (NW), or Tacaribe Complex viruses found in the Americas. LCMV is the exception, an OW virus that exhibits a worldwide distribution through association with its natural reservoir, the common house mouse. The essentially asymptomatic, persistent infections observed in host reservoir species are evident of a long-term association with their respective hosts (Zapata & Salvato, 2013). Given the ancient geographical isolation that the OW and NW clades exhibit, these infections are generally unable to infect other rodent species, and show a degree of co-evolution with their hosts (Zapata & Salvato, 2013).

The relatively recently discovered reptarenaviruses cause a condition known as Inclusion Body Disease in snakes. The first member of this genus was discovered in 2012, with three more species isolated since (Stenglein et al., 2012). The isolation of these viruses from snakes, which are known predators of multiple rodent species, is further evidence of long-term association with hosts given that distinct viruses, some significantly diverse (Bodewes et al., 2014), have evolved to establish stable infections in predatory reptiles.

1.1.3. Rodent and human life cycle

The host tropism of arenaviruses means that they are held stably within one, or a small number, of specific reservoir host species, limited to specific ecosystems and habitats within that ecosystem. Stable host infection results in persistent, sub-clinical manifestation of infection, which presents with

rodents appearing marginally smaller in size than non-infected counterparts (Fichet-Calvet et al., 2014; Fichet-Calvet et al., 2007).

The predominant identifying features of arenavirus-infected rodents – smaller size – does not appear to inhibit the ability of those infected to thrive in a non-stressed environment. However, recent studies have identified mutations in LASV genomes after the Biafra war in Nigeria, and the Sierra Leonean and Guinean civil wars. The mass migration of human populations stressed the local rodent populations, altering the availability of food. This competition favoured non-infected rodents, creating a selection pressure to generate viral clones which would not inhibit the host's ability to compete for food and mates (Lalis et al., 2012; Shaffer et al., 2014). This competition-driven evolution could allow virally-infected rodents to be likely to interact with humans due to increased ability to compete with their non-infected counterparts. Indeed, in the aftermath of the conflicts, LASV infection rates increased noticeably. (Lalis et al., 2012).

The infectious cycle in rodents is one of persistence; permissive rodents exhibit minimal clinical signs and can possess high titres in blood and excreta (Chiller & Oldstone, 1984). Strains which cannot block interferon signalling are cleared by the immune system. The specific residues controlling acute vs chronic infection in rodents have been identified for LCMV, with a mutation in GP, F260L, and the polymerase K1079Q responsible for the persistent form (Teijaro et al., 2013; Wilson et al., 2013; Matloubian et al., 1990). This persistence contributes to the ability of rodents to thrive, and therefore infect humans through their behaviour and exploitation of human dwellings.

Virus transmission between rodent hosts is predominantly horizontal, although vertical transmission is also common, and is often virus species specific. For example, LCMV is known to able to cross the placenta in humans and cause complications in infants as a result (Zapata & Salvato, 2013; Fichet-Calvet et al., 2014; Milazzo et al., 2011; Barton et al., 2002). Arenavirus infections in humans occur following close interaction with a virus-infected rodent host resulting in exposure to infected bodily fluids, most commonly

through inhalation of virus-contaminated urine or faeces via airborne droplets (Centers for Disease & Prevention, 2006; Emond et al., 1982; Dylla et al., 2008). Ingestion and preparation of virus-contaminated food, such as bush meat, is also a possible infection route (Dylla et al., 2008). In addition, in incidents of HF outbreaks, nosocomial transmission is common, thus the exposure risk of healthcare workers must be taken into account during HF outbreaks (Paweska et al., 2009). A final transmission route that is both extremely rare and often fatal is via organ transplantation; cases of LCMV infection after transplantation have been reported, with the harvesting of organs common after trauma rather than illness. Individuals who are sub-clinically infected with LCMV are able to pass on this infection to recipients of their harvested organs. The immunosuppression medication allows the transplanted infection to take hold, with infection by this route often exhibiting a CFR of 100% (Wright & Fishman, 2014; Palacios et al., 2008; Macneil et al., 2012; Fischer et al., 2006).

There are not thought to be other stable hosts within mammals outside of the specific limited host species. Given the constitutively expressed receptor utilised by many arenaviruses of the OW, α-dystroglycan, it is conceivable that predatory mammals could become infected in the wild upon hunting of infected rodents. Macaques and other primates are utilised experimentally to study disease progression, and thus could become infected in the wild (Hensley et al., 2011).

1.1.4. Geographical distribution and epidemiology

Arenaviruses exhibit worldwide, pocketed, distribution through their association with rodents. The extent to which each disease is geographically restricted corresponds exactly to the natural habitat of its host species. The geographical isolation of the ancestral host of most rodent species and subsequent divergent evolution is evidence of co-evolution between virus and host (Zapata & Salvato, 2013). The first recognised arenavirus outbreak in humans was the St. Louis encephalitis outbreak of 1933, but as these infections are thought to enjoy a long-term association within their natural

hosts it is likely these infections have been causing disease for many centuries (Muckenfuss et al., 1934; Charrel et al., 2011).

The most widespread arenavirus is LCMV, with the distribution pattern of its host, the common house mouse (*Mus muluscus)* near-global, as illustrated in figure 3 (IUCN, 2015). The life cycle of the common house mouse is one of direct reliance on human dwellings. This association carried the rodent globally, likely through colonial-era naval trade. In certain areas, up to 60% of rodents show sero-prevalence of LCMV.

Members of OW clade, with the exception being LCMV, show geographical isolation in isolated regions of Africa and Asia; correspondingly, NW species are found throughout the Americas. This distribution pattern suggests that wherever there are local geographically isolated rodent populations, there could be an associated arenavirus species. For example, Whitewater Arroyo virus localised in central USA, Catarina virus found in Texas and the Amapari virus found in Northern Mexico were all isolated from specific areas from closely related rodent species, and are all able to induce cerebral illness in humans (Lele et al., 2003; Milazzo et al., 2010; Cajimat et al., 2007; Mary Milazzo et al., 2011). Recent identification of rodent-borne arenaviral infections in South East Asia are also evidence that arenaviruses have a far wider distribution pattern than originally thought, and absence of obvious infection in rodents is not indicative of an absence of infection (Lele et al., 2003; Van Cuong et al., 2015).
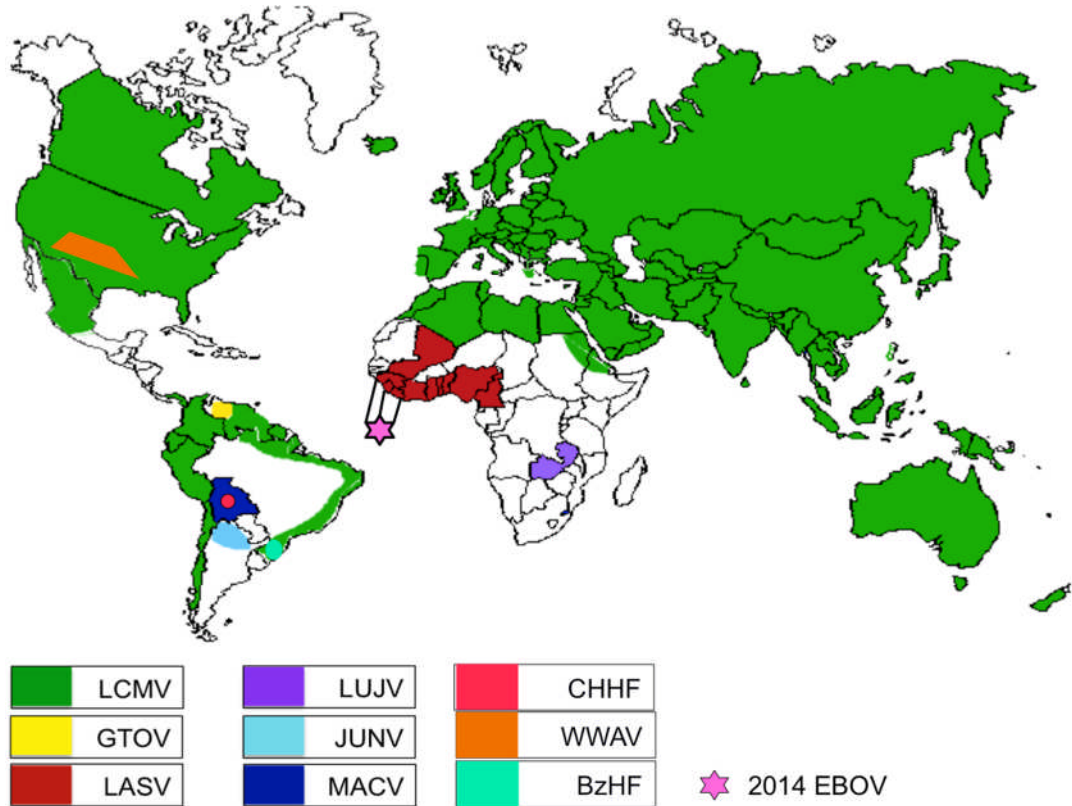
Map illustrating the relative geographical distribution of selected members of the *Arenaviridae.* Data based on clinical reporting, host species distribution and historical reports. LCMV – Lymphocytic choriomeningitis virus; GTOV – Guanarito Virus; LASV – Lassa Virus; LUJV – Lujo virus; JUNV Junin Virus; MACV – Machupo Virus; CCHF – Chapare virus; WWAV – Whitewater Arroyo Virus; BzHF – Brazillian Haemorrhagic Fever Virus; EBOV – Ebola virus, 2013-16 outbreak. (Weaver et al., 2000; Mary Louise Milazzo et al., 2011; Shao et al., 2015; Charrel & de Lamballerie, 2010; Maiztegui, 1975; Childs et al., 1992; Briese et al., 2009; Ogbu et al., 2007)

Figure 3. **Map illustrating the geographical distribution of selected arenaviruses**

The arenavirus that is associated with the largest burden of human disease is LASV, which is hosted within the *Mastomys natalensis* species prominent across much of sub-Saharan Africa (IUCN, 2015). LASV itself is limited to the Guinea region of West Africa, extending from Guinea to Nigeria. However, recent analysis of habitats and social changes has expanded the risk areas of LASV infection further East and South, from Cameroon, Democratic Republic of Congo and south towards Angola (Fichet-Calvet & Rogers, 2009; Peterson et al., 2014). The reason for the lack of LASV within rodents in other areas of Africa is unknown. The Guinea region of West Africa, from Cameroon west towards Senegal, is where LASV is prevalent and is the most populous region

of Africa, with Nigeria predicted to become the world's third most populous country by 2050 (United Nations Populations Division, 2015). The potential for more devastating outbreaks of LASV within this region to occur following behavioural changes in the local rodent population is clear (Shaffer JG; Grant DS; Schieffelin, 2014; Olugasa & Dogba, 2015).

M. natalensis is also the rodent host species for a number of other arenaviruses in geographically distinct locations (Gryseels et al., 2015) including: the non-pathogenic Mopeia virus in Mozambique and Zimbabwe; Morogoro virus in Tanzania; Luna virus in Zambia; and finally Gairo virus in central Tanzania. This divergent distribution of separate viral infections within the same host species could indicate divergent evolution, with the virus adapting to its host differently in different geographical locations. Gairo virus itself is more closely related to an Ethiopian arenavirus not hosted within *M. natalensis* – Mobala virus – whilst still exhibiting features similar to other *M. natalensis* infecting species (Gryseels et al., 2015).

NW arenaviruses share this distribution trend, with localised pockets of human infections due to interaction with the resident rodent species in their habitats. Junin virus (JUNV) is the predominant human NW pathogen, responsible for Argentinian haemorrhagic fever. JUNV is localised to the rural, agricultural regions of Northern Argentina, with few cases as a result. In untreated and un-vaccinated individuals, the CFR can reach 15-30%, with this brought down to 1-2% when patients are treated with convalescent plasma (Yun et al., 2008; Gomez et al., 2011). JUNV is carried within a number of host rodent species, the accepted host reservoir being *Colomys musculinus,* and several small field mouse species that have similar habitats to the reservoir species. In addition, the predatory *Galictis cuja* rat is known to be infected by JUNV in the wild (Gomez et al., 2011; Yun et al., 2008; Garcia et al., 2000).

The pathogen responsible for Bolivian haemorrhagic fever (BHF) – Machupo virus (MACV) – is hosted within large cane mice of the *Calomys callosus* species in the northern rainforest region of Bolivia. They are known to invade towns and villages at times of environmental stress such as low food

availability, and thus cause urban case outbreaks of BHF with CFRs of up to 40%, with an average CFR of approximately 20% (Shao et al., 2015). Rodent control measures in Bolivia have proven effective in limiting the outbreaks of BHF; however, the ability for rodents to persist in urban environments and adapt is well documented (Barnett & Dickson, 1989), and thus control measures may not prove adequate in combatting arenaviral diseases.

Another arenavirus-mediated haemorrhagic fever is Venezuelan haemorrhagic fever (VHF), the causative agent being Guanarito virus (GTOV). It is responsible for several cases of VHF per year and is associated with a CFR of approximately 30% (Milazzo et al., 2011; Salas et al., 1991). The host species of GTOV is the cane mouse *Zygontontomys brevicuda*, which is prevalent in the far northern regions of South America, from northern Brazil to Panama. Cases of GTOV in humans are isolated to the rapidly expanding western regions of Venezuela's Orinoco oil belt. This provides further evidence that human population expansion has the potential to force otherwise endemic but isolated diseases into human populations (Talwani, 2002; Weaver et al., 2000).

Finally, the newly described Lujo virus (LUJV) of the OW clade of arenaviruses emerged in Zambia in 2008, causing an index case and a further 4 cases of nosocomial transmission under patient containment conditions (Briese et al., 2009; Sewlall et al., 2014). To date, this is the only known outbreak and while the initial outbreak was small, the ease with which it was transmitted nosocomially despite standard infection control procedures indicates that LUJV could have the potential to be a devastating human pathogen. To date, the host species of LUJV is unknown, despite efforts to trap multiple rodent species in efforts to detect the host species (Ishii et al., 2012).

1.1.5. Genetic diversity and evolution

The two distinct genera of the *Arenaviridae* family – the mammarenaviruses and the reptarenaviruses – are further subdivided into clades. The mammarenaviruses are subdivided into two serogroups with GP antibody cross-reactivity throughout the two serogroups. These two groups generally

correspond to their relative geographical distributions; the Old World (OW) group of agents originating from the African and Eurasian continents and the New World (NW) serogroup, which is further subdivided into clades A, B and C based on overall amino acid sequence analysis, of agents originating from the Americas (Charrel et al., 2008; Radoshitzky et al., 2015). The specific separation between species can be defined as a 12% overall amino acid sequence (Radoshitzky et al., 2015). Within the NW group the majority of human pathogens are found in clade B, with a few, such as Whitewater Arroyo virus, found in clade A. The amino acid and nucleotide differences between different arenavirus species within a serogroup is generally over 30%, with some exhibiting as much as a 50% difference (figure 4) (Briese et al., 2009). The geographical separation of arenavirus species results in distinct differences in viral evolution. Despite being one of the more closely related arenaviruses, LASV and Mopeia virus (MOPV), exhibit an approximately 26% amino acid difference and are both found in the same reservoir host. Despite this, MOPV is not known to infect humans, indicating a different evolutionary path perhaps due to less interaction between its host and humans in the much less densely populated Central and Southern Africa compared to West Africa (Georges et al., 1985; Briese et al., 2009).
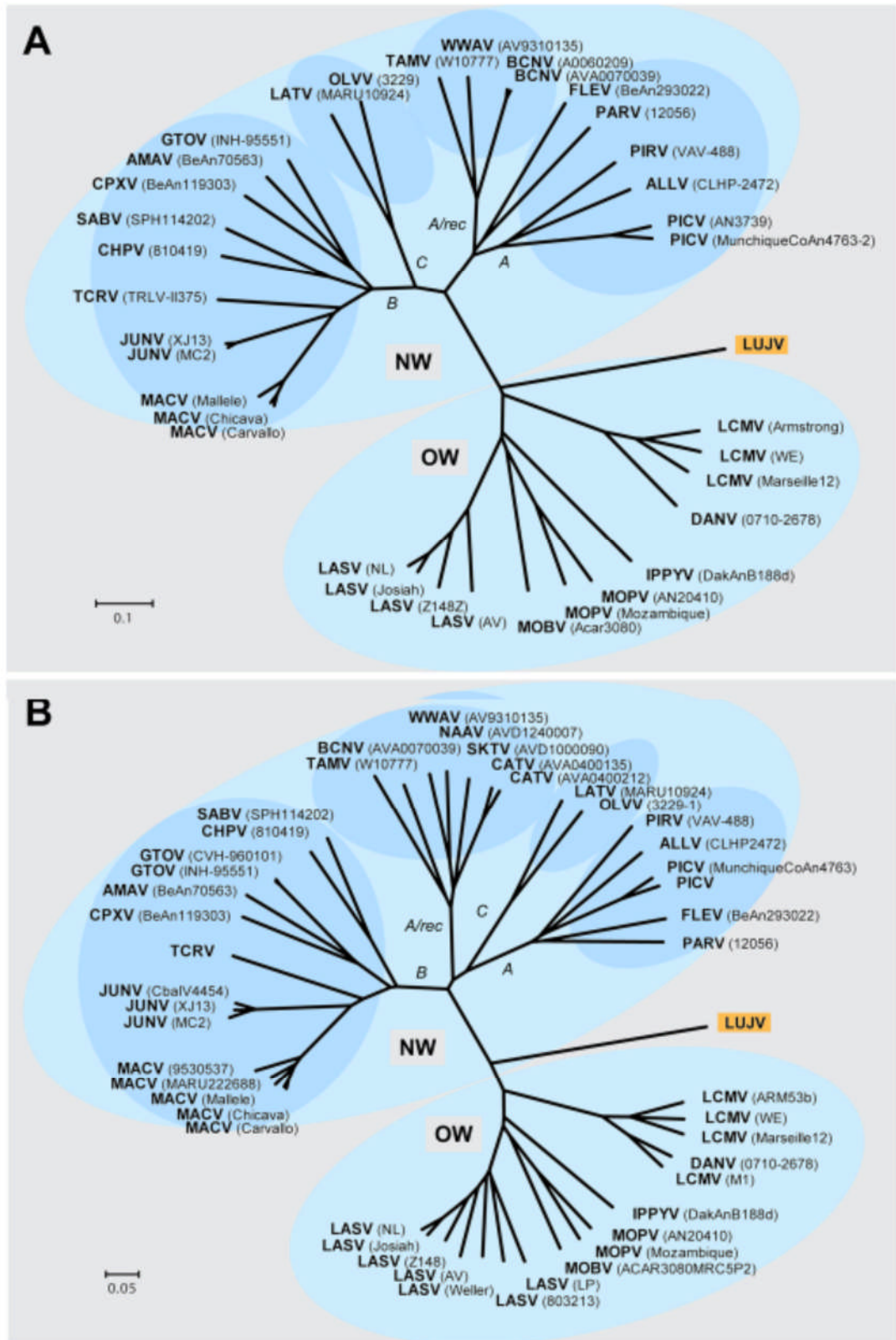
Phylogenetic relationships of LUJV inferred on full L (panel **A**) and S segment nucleotide sequence (panel **B**) Phylogenies were reconstructed by neighbour-joining analysis applying a Jukes-Cantor model; the scale bar indicates substitutions per site; robust boostrap support for the positioning of LUJV was obtained in all cases (>98% of 1000 pseudoreplicates). GenBank Accession numbers for reference sequences are provided in appendix I (Briese et al., 2009).

Figure 4. **Neighbour-joining phylogenetic tree depicting LUJV L and S segment relationships with other members of the *Arenaviridae* family members**

In 2008 LUJV emerged in Lusaka, Zambia followed by nosocomial secondary and tertiary transmission at a clinic in Johannesburg, South Africa – giving the virus its name. Sequence analysis of the virus isolated from two of the infected patients confirmed the presence of a novel arenavirus. Phylogenetic analysis showed that LUJV was significantly distinct from other OW arenaviruses (Briese et al., 2009). This analysis placed LUJV at the very root of the phylogenetic analysis; whilst being most closely related to other OW arenaviruses, the LUJV GP sequences are more closely related to its NW cousins than other OW viruses, shown in figure 4 (Briese et al., 2009). The discovery of such a diverse arenavirus is evidence that there could be several more highly divergent arenaviruses yet to have emerged into human populations.

1.1.6. Arenaviral infectious syndromes in humans and disease pathogenesis

1.1.6.1. HF complications of Arenaviral infections

Humans, and other secondary hosts – *i.e.* organisms in which infections become clinically observable – can become infected with arenaviruses. Individuals who encounter infectious material through exposure to infected urine; faeces; blood; ingestion of contaminated meat; direct human-human contact; airborne liquid droplets and sexual contact transmission can become infected. Less common routes of infection include vertical transmission and through breast feeding (Peterson et al., 2014; Dobec et al., 2006; Kernéis et al., 2009; Charrel et al., 2011; Lalis et al., 2012). LASV exhibits an incubation period varying from 3-21 days. The usual HF clinical manifestations begin with

mild febrile illness progressing to severe 'flu-like' symptoms including high-grade fever, weakness and malaise in the majority of cases. Coughing, pharyngitis and severe headaches are common, with gastrointestinal manifestations including diarrhoea, vomiting and general nausea also common complications (Yun & Walker, 2012). In HF-causing infections, such as LUJV, GTOV, JUNV and, in certain cases, LASV, it is the disturbance of vascular function and subsequent increase in vascular permeability which manifests as haemorrhagic complications, oedema and pleural and pericardial effusions (Yun & Walker, 2012). Disseminated intravascular coagulation (DIC) is also a life-threatening complication of HF infections, such as LUJV, and indicates a poor prognosis (Sewlall et al., 2014; Ogbu et al., 2007).

In patients who are unlikely to recover, deterioration manifests as severe oedema, respiratory distress, seizures, encephalopathy and can result in coma, DIC and fatal haemorrhagic shock (Yun & Walker, 2012). Primary bleeding in most cases of LASV is not sufficient to induce life-threatening shock. However other infections such as those caused by LUJV may be severe enough to induce life-threatening haemorrhagic shock,as bleeding starts at an earlier stage of disease (Bird et al., 2012; Sewlall et al., 2014). After symptom onset, recovery will usually begin between 8 and 10 days, and in cases of severe disease, deterioration will occur between days 6 and 10 (Yun & Walker, 2012). Sensorineural deafness is common amongst those who recover from arenaviral disease, a complication also commonly observed in EBOV survivors (Okokhere et al., 2009).

In the majority of locations where human disease-causing arenaviral infections are endemic, other febrile illnesses are common. This can cause a high proportion of mis-diagnoses as large proportions of populations are sub-clinically infected with malaria, or other more common tropical diseases such as Dengue fever. Mis-diagnosis can lead to increased risk of nosocomial transmission and the prolonging of outbreaks as a result.

Viraemia is a good clinical indicator for prognosis. LASV infected patients found to have a lower viral titre, corresponding to approximately a $TCID_{50}$/mL of $<10^3$, being indicative of a better prognosis (Faye et al., 2015; Asogun et al., 2012; Yun & Walker, 2012). Viraemia peaks between 4 and 9 days post symptom onset. Patients will generally clear virus from their blood within three weeks post symptom relief. However, it has been noted to persist in mammary glands and testes for months, conferring sexual transmission via semen for several months after recovery (Asogun et al., 2012; Bausch et al., 2000; Mate et al., 2015).

The precise cellular pathogenesis that gives rise to the symptoms of arenavirus infection is poorly understood. In particular the cellular processes which induce or facilitate the systemic symptoms and terminal syndromes. The lack of T-cell immune responses to LASV is indicative of poor prognosis, but the underlying mechanisms are unknown (Yun & Walker, 2012). The lack of an adaptive immune response means that neutralising antibodies are not present in the blood and can lead to fatal disease.

The clinical manifestations of LUJV mimic those of LASV but to a more severe extent clinically, rivalling EBOV in terms of severity. The 2008 LUJV outbreak was associated with rapid onset of fever and rapid deterioration to HF symptoms (Sewlall et al., 2014). Out of the 5 patients, only one survived long enough to receive high dose broad spectrum anti-viral ribavirin, in addition to clotting factor VIIa to combat DIC, with this patient ultimately surviving. Mild immunosuppressants were also administered in order to prevent induction of a cytokine storm (Sewlall et al., 2014). The availability of many of these agents in rural outbreaks would be low, and any larger-scale outbreaks of this newly emerged pathogen would require swift responses from aid agencies and governments (Sewlall et al., 2014).

Histological observations and changes in human tissue infected with LUJV is unknown, with analysis of representative the animal model – LUJV-infected 13/N Guinea pigs – indicating histological changes within all organs, especially kidneys, liver, spleen and lungs (Bird et al., 2012). Clinical

observations in humans matched this systemic distribution. This results in rapid deterioration from the relatively benign initial onset followed by rapid and severe multi-organ failure. The risk of nosocomial infection is increased if primary differential diagnoses do not include haemorrhagic fever, due to lack of protective measures (Sewlall et al., 2014).

Healthcare professionals should be aware of this disease progression when dealing with severe febrile illnesses in patients from rural locations in developing nations, and arenaviral HF should be considered as an important differential diagnosis.

1.1.6.2. Cerebral complications of LCMV infection

LCMV is generally asymptomatic and rapidly cleared in otherwise healthy adults (Bonthius, 2012). However, LCMV is still a major public health concern due to its high CFR in immunocompromised patients. LCMV is, as discussed above (1.1.3), able to cross the placenta, conferring severe developmental abnormalities in infected infants or prenates. Additionally, LCMV is an agent known to be transmitted through tissue and solid organ transplantation; as it is generally asymptomatic the infection in the organ donor is missed (Schafer et al., 2014; Palacios et al., 2008; Macneil et al., 2012). The majority of transplants occur after either brain or cardiac death, where a clear cause of death has been confirmed (Manara et al., 2012; Souter & Van Norman, 2010). The required use of immunosuppressant medication after receiving donated organs allows LCMV to thrive unchecked in an immunosuppressed patient, which can cause CFRs of up to 100% through transplantation transmission (Palacios et al., 2008; Fischer et al., 2006).

LCMV is the most widely researched arenavirus, and thus more is known about its immuno-regulatory mechanisms. In mice, LCMV is able to establish persistent infections through modulating the host interferon response. Interfering with these processes prevent the activation of cytotoxic T-lymphocytes (CTLs), thus allowing infected cells to persist (Chiller & Oldstone, 1984; Teijaro et al., 2013). In humans, LCMV is unable to overcome the intracellular type I interferon signalling, thus allowing the immune system to

clear the virus. In immunosuppressed individuals the activation of CTLs is impaired, thus allowing the infection to continue unperturbed. In mice, it is the NP which is responsible for interfering with interferon signalling by preventing the activation of IκB Kinase (IKK) and interferon regulatory factor-3 (IRF3) (Pythoud et al., 2012; Rodrigo et al., 2012). While cells infected with LCMV are able to signal appropriately in humans, this signal is not acted upon in immunocompromised patients.

1.1.7. Small animal models of LUJV and LCMV

The reservoir species of LUJV is unknown despite efforts to capture infected rodents in the region of Zambia from where the outbreak occurred (Ishii et al., 2012). The disease progression of LUJV within its reservoir species is therefore unknown, and in the absence of an effective infectious model, details of its disease progression is limited to the case reports of the patients from the 2008 outbreak (Sewlall et al., 2014; Paweska et al., 2009).

LUJV was found to infect clone 13/N Guinea pigs, presenting with similar clinical features as patients infected in the 2008 outbreak (Bird et al., 2012). The disease, as is common for all arenaviral infections, was able to infect multiple host tissues, indicating the use of a ubiquitous entry receptor. This may be α-dystroglycan, as is utilised by LASV and LCMV and other OW and NW clade C arenaviral infections (Moraz et al., 2013; Shah et al., 2006).

Disease onset was approximately 5 days post exposure – as opposed to between 7 and 13 days in humans – with fever, loss in weight, ocular discharge, dehydration, haematuria, genitourinary haemorrhage and eventual moribundity and death (Bird et al., 2012; Sewlall et al., 2014). All infected animals had either died or had been euthanised within 17 days of infection. Autopsy of deceased animals identified significant internal frank haemorrhage in multiple organs; most notably within the large and small bowel, liver, bladder and in the surrounding lymphatic tissue. Evidence of DIC was found in the heart and liver, with secondary necrosis of the liver and cardiomyopathy (from ~12 days post infection (PI).) also observable (Bird et al., 2012).

13/N guinea pigs provide an effective clinical model of LUJV infection. However, the severe limitations presented by working with a Bio-safety level 4 (BSL4) pathogen within experimental animals restricts their use significantly. The identification of an effective BSL2 model for LUJV infection would greatly assist in evaluating therapeutic procedures or testing novel therapeutics prior to assessment in LUJV infected animals.

LCMV has for some time been used in studies involving experimental animals, most notably mice, for research regarding general immune responses to viral infections (Teijaro et al., 2013; Wilson et al., 2013). The clinical presentation of LCMV is also well established, and divergent from the majority of arenaviral infections, exhibiting cerebral tropism, causing encephalitis rather than any haemorrhagic complications. Utilising LCMV as a model for HF causing infections must take into account effects that are known to be common between the LCMV and the HF infection of interest. Finally, in cases of infant and pre-/post-natal exposure LCMV has been indicated in severe developmental abnormalities and any animal model must thus produce similar developmental impairments in infected animal infants. In this instance, LCMV itself is of research interest in the prevention of developmental abnormalities in children, and infections of the immunocompromised.

1.1.8. Diagnosis

Diagnosis of viral HF is notoriously challenging, with many alternative and more common causes of acute and severe febrile illness endemic in most areas where pathogenic arenaviruses are present. In addition, malaria can be endemic, with persistent and asymptomatic infections in swathes of the population in Africa, with a positive test resulting in improper treatment. For LASV, dedicated facilities in Nigeria utilise PCR and ELISA-based techniques (Asogun et al., 2012). After the 2013-16 EBOV outbreak new laboratories were established in Sierra Leone, and other West African nations, in order to diagnose acute febrile illnesses more quickly as a direct result of the international effort to combat the outbreak (Asogun et al., 2012; Crowe et al., 2015; Personal Experience, 2015).

1.1.9. Treatments

The number of effective therapies for arenaviral infections is limited, with high-dose intravenous (IV) ribavirin indicated in treatment of LASV, and convalescent serum useful in treatment of NW infections (Hadi et al., 2010; Ruggiero et al., 1986). Ribavirin is often used as post-exposure prophylaxis, given to close contacts and healthcare workers after a case of LASV is confirmed. However, post-exposure prophylaxis has its limitations, for example, the efficacy of ribavirin is only observed if administered early, and does not appear to be capable of reversing the course of a well-established infection (Hadi et al., 2010). Commonly, disease in the index case will have progressed too far for successful treatment after diagnosis is confirmed, meaning the infection could have spread to other individuals necessitating rapid post-exposure therapy.

The side-effects of ribavirin, including severe 'flu-like' symptoms and suicidal depression, make it a less than ideal choice as a post-exposure prophylactic, and thus new therapies are required. Ribavirin is contraindicated in a number of circumstances, most prominently in patients with anaemia, sickle cell anaemia (SCA), pregnancy or bleeding disorders making its use as a therapeutic in HF patients more complicated. With SCA also highly prevalent in areas endemic for malaria, ribavirin is an undesirable therapeutic agent in tropical regions. In addition, ribavirin is severely contraindicated in pregnancy, causing severe congenital defects. Its use is not recommended in expectant mothers, those wishing to become pregnant and males whose partner is planning to become pregnant, and its use prohibits in either male or female recipients pregnancy attempts for 6 months (US National Library of Medicine, 2015).

These defects and side-effects are due to ribavirin's proposed mechanism of anti-viral action. As a nucleoside antimetabolite, ribavirin is incorporated into the viral genome during RNA replication, but without the canonical Watson-Crick base pairing between purines and pyrimidines. Subsequent rounds of template-directed RNA synthesis allow further mis-incorporation events to

occur leading to exponential increase in the number of viral polymerase mistakes. Eventually, the error rate reaches a point where the accumulated viral genomes are no longer viable, known as 'error catastrophe'. As ribavirin can also be incorporated into host DNA and RNA, this lack of specificity confers the observed side-effects and contraindication in humans, particularly during pregnancy (McEvoy, 2005; MeSH, 2015; US National Library of Medicine, 2015). Ribavirin is supplied as a prodrug, and it is the ribavirin-5'-monophosphate, -diphosphate and -triphosphate forms metabolised and phosphorylated in the liver, that are likely the agents behind its proposed activity (McEvoy, 2005).

The use of convalescent serum has been shown to be efficacious in treating those infected with JUNV, and could therefore be used in other cases of arenaviral infectious outbreaks. Again, this comes with severe limitations; in novel outbreaks, the availability of convalescent serum will be essentially non-existent, with a reliance on stable healthcare networks to provide a continuous cold storage chain to maintain efficacy. In the case of acute, novel outbreaks both ribavirin and convalescent serum are less than ideal therapies. Novel therapeutic approaches are therefore warranted, and thus identifying mechanisms of viral replication strategies are essential.

In addition to limited therapeutics, there are no widely available vaccines for arenaviral infections. The Candid #1 strain of JUNV has been licenced for use in Argentina for decades, but is restricted to use only in Argentina. It is, however, given to front-line healthcare workers and armed forces personnel in the United States in the event of bioterrorism threats (Ambrosio et al., 2011; Bausch et al., 2010; Centers for Disease et al., 2012; Murphy, 2008).

The vaccine was developed in 1979 by an international consortium of Argentinian, US, Pan-American and UN public health organisations. It was isolated after serial passages of the XJ44 strain of JUNV in FRhL-2 cells, generating a protective and safe immunogenic response in Guinea pigs, mice and Rhesus monkeys prior to human clinical trials (Ambrosio et al., 2011). As is the case with many neglected, yet significant, human pathogens, the

vaccine was not commercially viable to produce. The Argentinian government financed the manufacture of the vaccine, with the final product providing 95% efficacy (Ambrosio et al., 2011).

1.2. **Arenaviral structures, genome and proteins**

1.2.1. Structure and genome

Arenavirus particles are lipid soluble and are approximately 100nm in diameter. They consist of a host-derived lipid membrane, forming an envelope with embedded glycoprotein spikes. Particles are pleomorphic but generally spherical, and are considerably larger than non-enveloped viruses such as poliovirus, and comparable to certain herpesviruses (Chevaliez & Pawlotsky, 2006; Mocarski Jr, 2007; Murphy & Whitfield, 1975). Arenaviruses are composed of four virus-encoded proteins. The Glycoproteins (GP) GP1, GP2 and SSP are predominantly responsible for cell entry, membrane fusion and budding (Geisbert et al., 2005). On the inside of the viral envelope is the matrix (Z) protein. Z is a multifunctional protein, with roles in viral assembly, viral and cellular protein trafficking – such as ESCRT complex proteins and GP, NP and RNP complexes – and immune-modulation (Fehling et al., 2012; Fan et al., 2010). Within the interior of the virus particle the two circularised ambisense genome segments, circularised through direct RNA-RNA interaction (figure 5) are encapsidated with the viral nucleoprotein (NP), forming a ribonucleoprotein (RNP) when associated with the polymerase (L) (Pedersen & Konigshofer, 1976). Finally, the large polymerase (L) is thought to interact with the 3' and 5' ends of the ambisense genomes, associated with the RNP, in a similar manner to La Crosse Virus (LACV) of the closely related *Bunyaviridae* (Pedersen & Konigshofer, 1976; Gerlach et al., 2015).

Diagram depicting the 3' and 5' ends of the LCMV S segment accession #NC_004294, forming the proposed panhandle structure where L is thought to bind.

Figure 5. **Diragram depicting proposed panhandle structure of LCMV S segment**
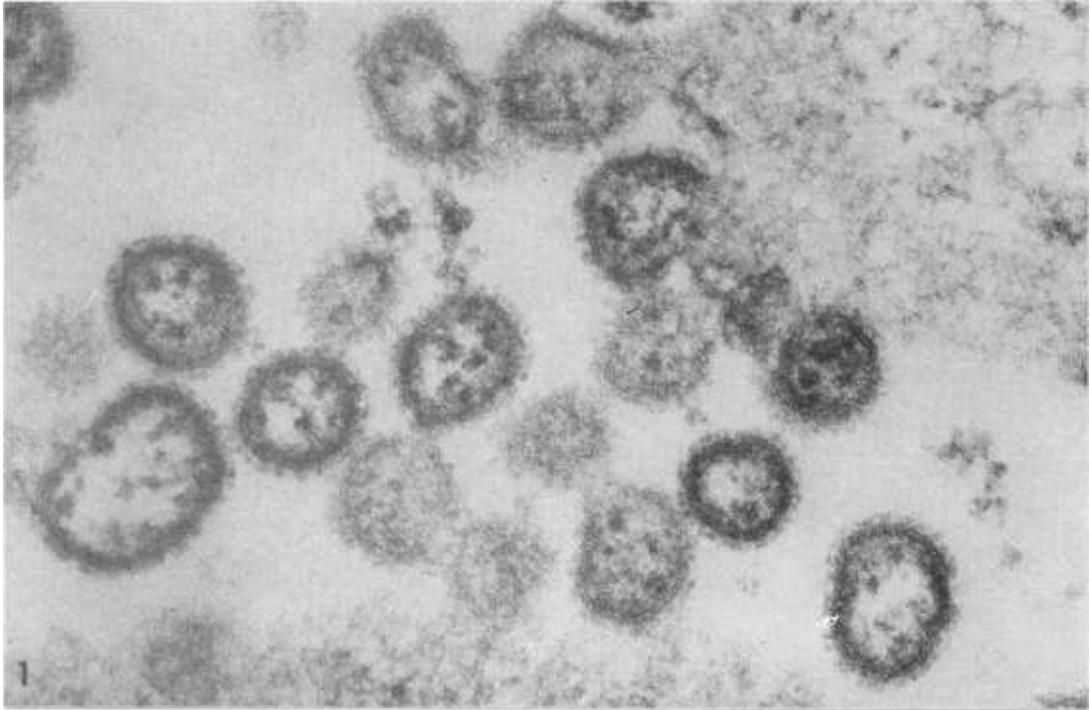
Electron micrograph of LCMV particles isolated from infected Vero cells (Murphy & Whitfield, 1975)

Figure 6. **Electron micrograph showing LCMV morphology**

As briefly discussed above, an early observation made from electron micrograph images is that purified arenavirus particles contain large dense structures within the virion interior, that give virus particles a 'sandy' appearance. These electron-dense sites are widely considered to represent host cell ribosomes (Rowe et al., 1970), and their sandy appearance was chosen to derive the family name – 'arena' being Latin for sand (shown in figure 6 (Murphy & Whitfield, 1975)). Subsequent investigations have further confirmed the presence of ribosomal subunits in virions (Pedersen & Konigshofer, 1976; Leung & Rawls, 1977).

The presence of ribosomes within viral particles is an intriguing element of the arenavirus life cycle. Their presence could indicate that other cellular components are packaged within viral particles. The presence of such factors – while not necessarily essential – could play a role in establishing an infection more efficiently than without.

1.2.2. Ambisense strategy of gene expression

Arenaviruses employ an ambisense genome replication strategy. Each virion-incorporated genomic RNA segment (vRNA) has the potential to encode two open reading frames (ORF) (Zapata & Salvato, 2013); one from the vRNA strand and another from its antigenomic, complementary copy (cRNA). The sequences that encode these ORFs are separated by an intragenic region (IGR), which has the potential to form a large hairpin secondary structure, and is believed to mediate transcription termination during mRNA synthesis. During transcription of the vRNA template, the virus-encoded RNA-dependent RNA-polymerase (RdRp) complete with a 5' capped oligonucleotide snatched from cellular mRNAs, binds to the 3' end of the template, thus priming mRNA synthesis. The RdRp moves towards the template 5' end, but upon reaching the IGR, the polymerase is thought to respond to cis-acting signals and is forced to terminate mRNA synthesis (Pinschewer et al., 2003; Morin et al., 2010; Tortorici et al., 2001). A similar transcription reaction can also occur following replication of this vRNA genome to yield a cRNA antigenome. As described above, the viral RdRp complete with snatched capped RNA oligonucleotide can also bind to the cRNA 3' end, and similarly extend a nascent RNA. When the IGR hairpin is reached, the RdRp again terminates transcription and releases the nascent mRNA. In the case of the S segment, the mRNA transcribed from the vRNA possesses an ORF that encodes NP, whereas the mRNA transcribed from the cRNA encodes GPC (figure 7). In the case of the L segment, the ORF of the polymerase protein L is transcribed from the vRNA strand, with the ORF of the matrix protein Z transcribed into mRNA from cRNA. GPC and Z can then be synthesised, with the switch between primary mRNA and antigenome formation controlled by a feedback system, proposed to be driven by NP abundance (Iapalucci et al., 1991; Burri et al., 2012). Extension past the IGR will allow for the generation of GP and Z and the following formation of virions.
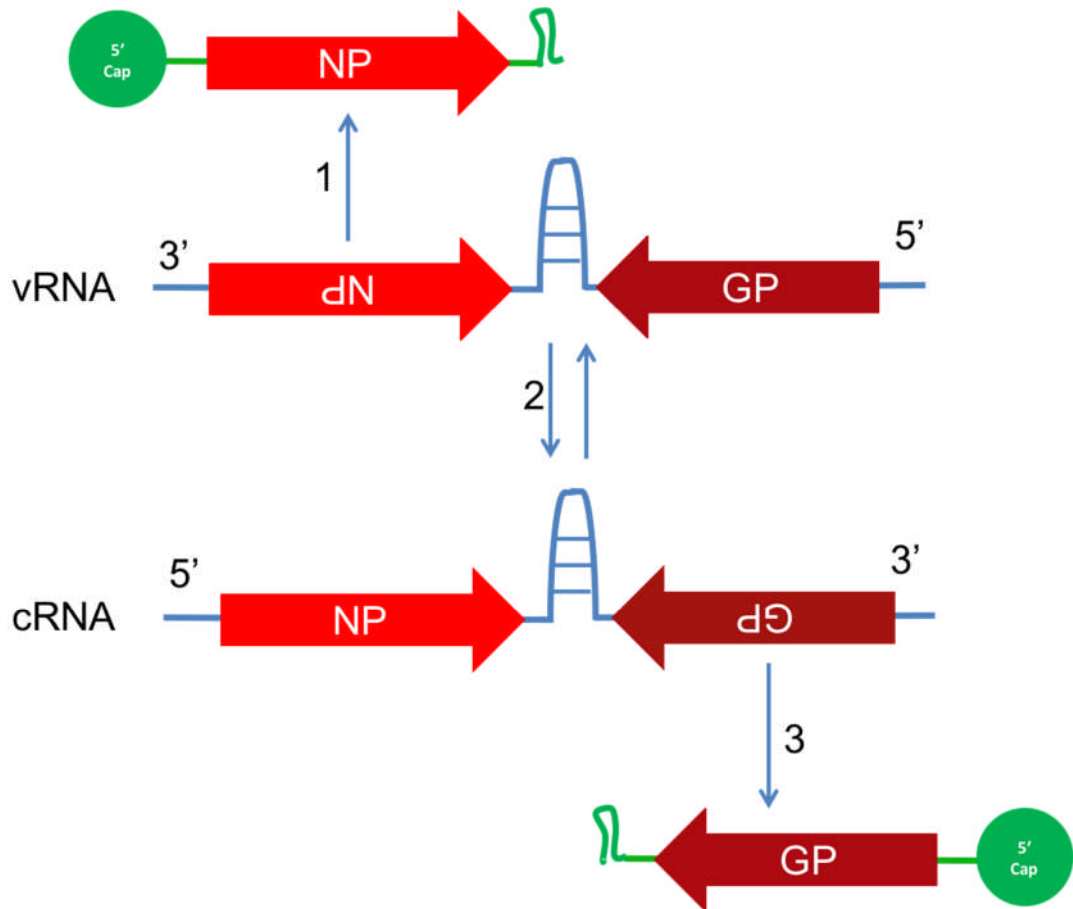
Diagram illustrating the ambisense strategy of arenaviruses, demonstrated by the S segment here. **1)** Transcriptional copying of the vRNA genome by L results in the generation of NP mRNA. **2)** Generation of complementary RNA (cRNA) to from an anti-genome**. 3)** Generation of GPC mRNA from the anti-genome cRNA template. The same applies between L/Z in the L segment of RNA. Sequences in the 'negative' sense within the ambisense vRNA and cRNA are shown with upside-down annotation.

Figure 7.   Arenavirus S segment transcription events

Present at the 3' and 5' ends of the vRNA and cRNA are conserved and complementary sequences, (figure 5) which are proposed to form a panhandle structure within virions, where the polymerase sits, in a similar manner to bunyavirus polymerases (Gerlach et al., 2015). Within the 5' and 3' ends of both the S and L segment are the promoter sequences for the arenavirus polymerase L (Perez & de La Torre, 2003). Alteration of these promoter sequences results in the loss of reporter expression, suggesting that initiation of RNA synthesis by L is sequence dependent and is sensitive to mutation (Perez & de La Torre, 2003).

None of the arenavirus mRNAs possess 3' poly(A) tails, but instead are thought to possess an extended hairpin secondary structure, copied from the conserved untranslated region (UTR) within the vRNA and cRNA templates (Meyer & Southern, 1994). How the poly(A) tail-less arenavirus mRNAs are efficiently translated is unknown. The established role of the cellular mRNA 3' poly(A) tail in translation is to both protect the mRNA 3' end from endo- and exonuclease attack (Weill et al., 2012) and to bind poly(A) binding protein (PABP). PABP subsequently associates with components of the eIF4F cap binding complex, leading to mRNA circularisation. Circularisation is essential to confer efficient primary scanning of ribosome complexes to find the initiator codon (Jackson et al., 2010).

One possibility is that the predicted arenaviral 3' secondary structure encoded by the IGR is able to functionally replace the conventional poly(A) tail. This may be through limiting RNA endonuclease degradation, and possibly interacting with other cellular or virus-encoded proteins to mediate eIF4F interactions. However, these possibilities remain to be tested. Other viruses, such as vaccinia virus and human respiratory syncytial virus (hRSV), are known to generate poly(A) tails through a virus encoded activity in order to ensure that their mRNAs are translated efficiently (Mohamed et al., 2001; Grosfeld et al., 1995), and thus it is likely that arenaviruses will have evolved a strategy to ensure efficient translation.

This proposed 3' hairpin structure would be predicted to trigger a significant immune response due to double stranded (ds) RNA being a pathogen-associated molecular pattern (PAMP) detected by multiple internal surveillance proteins, such as PKR, RIG-I and TLR3 (Saito & Gale Jr., 2008; K M Hastie et al., 2012; Akira et al., 2006). Therefore, it seems unlikely that such a significant immune trigger would be so well-conserved amongst all arenaviruses, suggestive of an important conserved role. Identifying a function for this predicted structure could help establish a novel therapeutic target.

1.2.2.1. Replication

At a certain time point post infection, the overall activity of the viral polymerase shifts from transcription to replication, and the switch for this fundamental change in polymerase activity is poorly understood (Burri et al., 2012; Iapalucci et al., 1991). During replication, which requires the synthesis of a complementary full-length copy of the vRNA or cRNA templates, with the presence of multiple copies of NP believed to be responsible for allowing L to pass over the IGR.

As depicted in figure 7, after the abundance of NP has reached sufficient levels, L is proposed to be able to pass over the IGR, generating a complete cRNA sequence or, in the reverse, new vRNA copies. The promoter regions held within the 3' and 5' vRNA & cRNAs allow for L to initiate this replication, in conjunction with NP conferring genome replication as opposed to transcription of mRNAs, during the extension both the template and nascent genomic/anti-genomic strands are encapsidated to form either genomic or anti-genomic, uncapped RNPs. Comparatively, the replication stage of bunyavirus RNAs has been proposed whereby the template strand is unwound from the RNP, and threaded through the polymerase, before re-associating with NP to re-form the original RNP (Gerlach et al., 2015). Simultaneously, the nascent strand would then be encapsidated by free NP, possibly by a second polymerase (Gerlach et al., 2015). This proposal has yet to be tested extensively, and how the process occurs during arenaviral replication is not fully understood. Nascent RNPs can subsequently act as a template for further genomic, antigenomic or messenger RNA strands. Each step in the replication cycle is explained in greater detail in 1.2.3.

1.2.3. Viral Proteins

The four arenaviral proteins, NP, GP, L and Z contribute to the viral life cycle in four specific areas: L is responsible for the transcription and replication of the viral genome (Salvato et al., 1989); GP is responsible for membrane attachment and subsequent fusion within endosomes (York et al., 2008); Z has roles in virion organisation and intracellular transport, in addition to

intracellular reorganisation and immune ablation (Fehling et al., 2012); with NP finally responsible for the formation of RNPs – and therefore RNA stability – and immune regulation (Martinez-Sobrido et al., 2006; Pedersen & Konigshofer, 1976). In addition to these defined roles, there is significant multifunctionality and overlap of function, where viral proteins work in combination. This includes the association of Z and GP while embedded within the viral envelope (Fehling et al., 2012) or NP and L working in conjunction to ensure efficient RNA replication and translation. In addition, there is a degree of overlap in functions of NP and Z, with both proteins implicated in innate immune regulation (Fan et al., 2010; Martinez-Sobrido et al., 2006); and regulation of L and, in turn, transcription and replication (Kranzusch & Whelan, 2011; Tortorici et al., 2001; Pinschewer et al., 2003). Overall, each arenaviral protein can be considered multifunctional, and important in many stages of the replication cycle within an infected cell.

1.2.3.1. Glycoprotein

The glycoprotein of arenaviruses is responsible for cell entry, membrane fusion and budding, along with conferring the assembly and budding sites of newly formed virions. GP is initially translated as a glycoprotein precursor protein (GPC) of approximately 75 kDa. This is then cleaved within the endoplasmic reticulum (ER), initially by signal pepsidases to remove the 58 residue stable signal peptide (SSP), with GPC cleaved at conserved sites between the GP1 and GP2 domains, corresponding to RX(hydrophobicX)X by pepsidases and subtisin kexin isosyme-1 (SKI-1) and Site-1 protease (S1P) to form three constituent subunits of mature GP; GP1, GP2 and SSP (Burri et al., 2012). In the case of LASV and LCMV respectively, these sites are RRLL and RRLA. GP1, GP2 and SSP are then processed, with GP1 & GP2 glycosylated within the endoplasmic reticulum (ER), enabling the formation of mature GP trimers.

Through GP, arenaviruses utilise ubiquitous receptors for cell entry, with the exact mechanisms regarding co- and secondary receptors unknown. The primary receptor usage for the OW viruses LASV and LCMV was first

identified in 1998 as α-dystroglycan, and later, in 2007 transferrin-1 (TfR1) was established as the primary receptor for NW human infections (Cao et al., 1998; Radoshitzky et al., 2007). In both instances, neutralising antibodies against the surface receptor were utilised to identify the receptor responsible for viral entry. In the case of NW clade C arenaviruses – non-pathogenic in humans – these utilise α-dystroglycan similar to their closely related OW counterparts.

The interaction between the receptor binding domain of GP (GP1) and the primary receptor for NW infections is known to be distinct from the usual transferrin binding domain, at the very apical tip of the receptor (Nunberg & York, 2012; Abraham et al., 2010). As there is not a major conformational change in the structure of the receptor it appears that – while crucial – the primary receptor binding is passive, rather than actively stimulating endocytosis (Abraham et al., 2010).

The evolution of viral sequences is evident in those species that can infect humans. The predominant clade B pathogens – JUNV, GTOV and MACV – all utilise TfR1, with others, such as TACV also using TfR1 but are not able to infect humans. Single nucleotide changes in GP1 are sufficient to allow human infection by non-pathogenic arenaviruses and it is possible that this mutation occurs naturally after increased exposure to humans, conferring infection (Martin et al., 2010; Abraham et al., 2010).

Upon binding, the virion is endocytosed, with GP-mediated fusion of viral and endosome membranes occurring as the endosome matures, and the pH drops (York et al., 2008). TfR1 is constantly recycled via clathrin-mediated endocytosis, and thus it is likely that TfR1-mediated endocytosis will confer efficient entry. Indeed, the passive nature of the initial binding with TfR1 could mean that the cell internalisation mechanisms are not alerted of virions that are bound to the receptor, with the internalisation of TfR1 simply part of the routine recycling of surface receptors. The body of work related to arenaviral entry has encompassed NW infections, with the broad basis of internalisation mechanisms theorised to be similar between NW and OW infections.

SSP is responsible for targeting GPC polypeptides into the endoplasmic reticulum (ER) for processing and glycosylation. SSP contains two hydrophobic membrane spanning regions, with a constitutively conserved lysine (K33) present on the extracellular/lumen side of membranes and known to be essential for membrane fusion (York & Nunberg, 2006). Both C and N termini of SSP are present on the cytosolic/matrix side (Nunberg & York, 2012; York et al., 2004). SSP is also essential in GP translocation from the ER, masking the retention motif of GP2 (Burri et al., 2012). Upon initial ER processing, N- glycosylation is the next essential step in GP maturation and function, with both GP1 and GP2 containing sites of significant glycosylation (Burri et al., 2012; Eichler et al., 2006). The stable and correct glycosylation of GP1 and GP2 is essential for transport to the membrane and external functions such as subsequent membrane fusion in newly infected cells.

In assembly, the matrix protein Z is the major modulator of the transport machinery (Fehling et al., 2012), and it is especially important in transporting GP to the plasma membrane (PM). Z interacts with the mature GP complex of GP1, GP2 and SSP in a manner dependent on Z myristolisation and SSP (Burri et al., 2012; Perez et al., 2004; Capul et al., 2007). Interaction between Z and the GP complex is required for egress and, in addition, this GP/Z complex must also interact with NP/RNP and L to form mature progeny particles. Interfering with the process of this binding is, therefore, an important target for therapeutic development.

1.2.3.2. L

As its name suggests, the arenavirus polymerase protein L is large, being the largest arenaviral protein, with a molecular weight of approximately 250kDa. It is responsible for several activities that all result in viral RNA synthesis; these include recognition and binding of promoter RNA sequences, cellular cap binding, endonuclease cleavage, nucleotide polymerization as well as likely mediating NP dissociation and re-association of the RNP template (Morin et al., 2010; Tortorici et al., 2001).

After membrane fusion, the process of viral replication can begin with L transcribing the vRNA genome into mRNA (Wilda et al., 2008). L recognises promoter sequences that by analogy with other negative-stranded RNA viruses likely include the 3' end of genomic and anti-genomic RNAs, and possibly the 5' end also. During both mRNA transcription and RNA replication, the L protein moves along the RNP template and synthesises the nascent RNA strand in a template-dependant manner. Also, by analogy with other RdRps, the arenavirus L protein is thought to mediate the transient dissociation of individual NP monomers from the RNA strand, followed by their subsequent re-addition, and reformation of the RNP. The naked RNA that is exposed during NP removal enters the RdRp active site and is copied into nascent RNA. This emerges from L, presumably through a dedicated exit channel. In the case of mRNA transcription, the nascent RNA is capped through the NL1 domain of the polymerase, where the removal of the cap structure of host capped mRNA takes place (Morin et al., 2010). The nascent mRNA is then extended until the IGR, at which point the RdRp terminates transcription and releases a mRNA, dependent on the abundance of NP (Burri et al., 2012; Iwasaki et al., 2015). In the case of RNA replication, the nascent RNA strand emerges from the exit channel and must be encapsidated to generate a new RNP. The exact mechanism is not fully understood, but evidence is emerging from the bunyavirus and orthomyxovirus fields that this involves a second RdRp, first reported by Gerlach *et al.* in 2015 through the structure of LaCrosse virus polymerase. MACV L EM images (figure 8) indicated a similar overall structure to the LaCrosse virus polymerase, with recent findings implicating that arenavirus L proteins could act in a similar manner to those of the *Bunyaviridae* family members (Gerlach et al., 2015; Kranzusch et al., 2010; Lehmann et al., 2014).
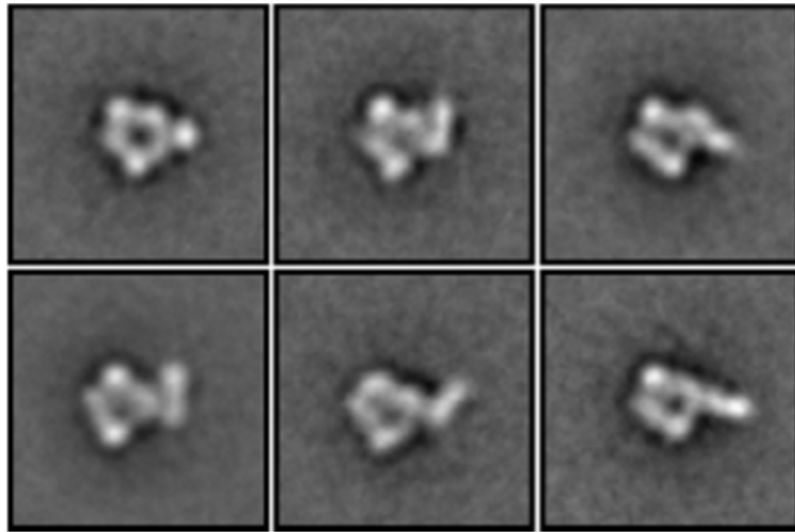
EM negative stain images of the MACV polymerase. Class averages of L after visualisation, observed by negative stain with each containing ≈100 particles. (Kranzusch et al., 2010).

Figure 8.   **EM observation of MACV L**

1.2.3.3. Z

The smallest arenavirus protein is the matrix protein known as Z. It is known to have multiple roles in the viral life cycle, with its most well established functions being internal transport of GP and RNP/L complexes and the facilitation of assembly and budding (Perez et al., 2004; Fehling et al., 2012). Beyond these functions, Z has a role in immune modulation and internal modulation of the cellular environment.

Z is between 90 and 103 residues in length for all the mammarenaviruses, making it all the more remarkable that it performs such a plethora of functions (Fehling et al., 2012). Z sequences across the *Arenaviridae* family share similar features: a myristoylation signal; a really interesting new gene (RING) zinc binding domain and tetra-peptide late domains (Fehling et al., 2012). These structural similarities underpin the basis of their multifunctional nature. The myristoylation site aids in the trafficking and membrane anchoring of mature GP complexes, as well as the incorporation of RNP and L into the budding virion (Perez et al., 2004). This anchoring can also facilitate Z in its other functions and self-assembly. By allowing the internal transport of proteins and interactions with other cellular and viral proteins through

myristoylation, arenaviral replication and release is enhanced (Strecker et al., 2006; Loureiro et al., 2011).

The Z protein RING domain is thought to act as a platform for interactions with cellular and viral proteins (Volpon et al., 2010). The interaction between LASV Z and eIF4E identified a novel RING domain sequence, which provided evidence that Z inhibits the cap-binding function of eIF4E. This interaction has been implicated in the inhibition of arenaviral translation. This function is thought to coincide with growing abundance of Z within a cell, whereby a feedback system stimulates – not dissimilar to NP's stimulation of the switch between transcription and replication – termination of translation and confers the recruitment of viral proteins and RNPs to viral assembly areas. The RING domain of Z is also essential for Z self-assembly, in conjunction with the myristoylation of the N terminus of the polypeptide (Fehling et al., 2012).

The recently described structure, shown in figure 10, of LASV Z in a dodecomeric orientation, was found to be comprised from 6 dimers (Hastie et al., 2016). The monomeric form of Z is known to be responsible for the interaction of LASV Z with eIF4E, however LCMV Z is known to interact with eIF4E through a multimeric form of Z (Volpon et al., 2010; Kentsis et al., 2002).
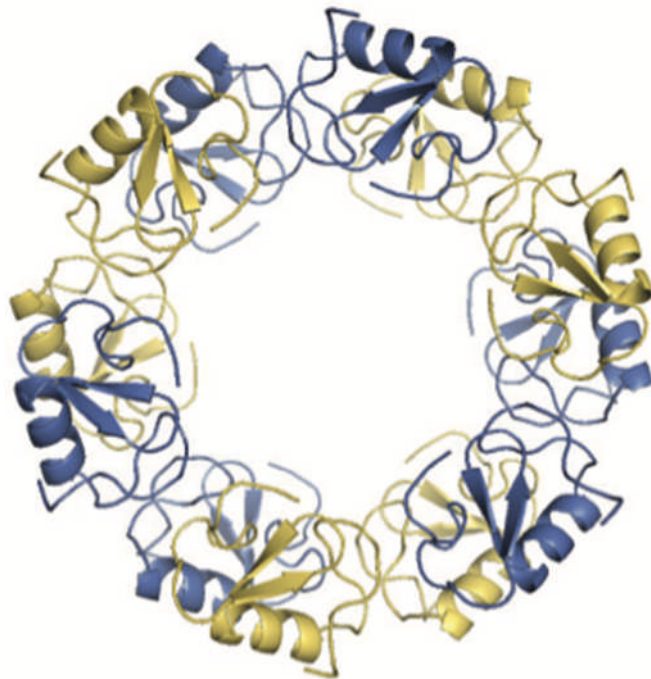
X-Ray crystal structure of the dodecomeric structure of LASV Z at a resolution of 2.9Å (Hastie et al., 2016)

Figure 9. **Crystal structure of LASV Z**

Z is also known to regulate transcription and replication during infection, dependent on the relative abundance of L and NP within the cell, in combination with the relative abundance of Z (Garcin et al., 1993; Jácamo et al., 2003). Indeed, this regulatory mechanism is so strong, that cells transiently expressing Z exhibit a degree of protection against LCMV infection through polymerase inhibition (Cornu & de La Torre, 2001). The mechanism behind this is not fully understood, but could indicate that Z acts in a feedback cycle; when the abundance is reaching high levels, Z would be able to direct components for assembly and budding, rather than continuing with strand elongation. Two Z binding sites on L indicate that Z directly binds L through the RING finger domains, inhibiting the polymerase's function (Kranzusch & Whelan, 2011; Wilda et al., 2008).

In assembly, Z is responsible for the anchoring and trafficking of proteins to the PM prior to release. Its association with the mature GP complex stabilises it in the membrane, with the association not dependent on other viral proteins, and the interaction appears to be in the embedded SSP of the mature GP

complex (Capul et al., 2007). Z appears to be able to generate 'assembly areas' in cells, forming vesicle-like structures of GP-Z near the nuclei – most likely derived from the ER where GP is processed (Schlie et al., 2010; Burri et al., 2013). This indicates an internal trafficking mechanism, again most likely facilitated by Z.

The assembly of arenaviruses occurs at the PM of infected cells, with viral egress from polarised lung epithelial cells generally limited to the basolateral face of cells, rather than the apical face (Dylla et al., 2008). The utilisation of the basolateral face of lung tissue indicates that the virus is not expelled from the lungs in high volumes, and thus not a predominantly aerosolised virus – supporting the established transmission mechanisms (Kernéis et al., 2009). While infection through the lung is known to occur, and aerosolised transmission possible (Stephenson et al., 1984; Sewlall et al., 2014), it is not the accepted human-human transmission route. A more common transmission route is the transfer of contaminated liquid droplets between patient and carer (Paweska et al., 2009; Sewlall et al., 2014).

Given that arenaviruses are theorised to replicate within replication-transcription complexes (RTCs) distributed throughout the cytoplasm, RNPs and proteins need to be trafficked to the PM. Z is known to traffic GP to the PM, and association between NP and Z has been shown within infectious particles, indicating that Z anchors RNPs into the membrane through association with NP. Indeed, several arenaviruses, most notably LASV, JUNV, LCMV and TCRV all exhibit this conserved interaction, through the C terminus of NP (Fehling et al., 2012; Schlie et al., 2010; Shtanko et al., 2010). It is theorised that Z aids in trafficking RNPs to assembly sites at the PM.

1.2.3.4. NP

The arenaviral NP is a highly abundant structural protein with a role in encapsidation of the viral genome to form RNPs, which is responsible for protecting the genome segments and providing structural integrity. It is also known to act as an anti-terminator during genome replication, which is proposed to allow for the synthesis of full-length vRNA and cRNAs

(Pinschewer et al., 2003; Tortorici et al., 2001; Pedersen & Konigshofer, 1976). In addition, NP is known to have roles that relate to inhibiting the induction of innate immune responses within the infected host cell. The arenaviral NPs are all of approximately 60-65 kDa, with their functions predominantly conserved between NW and OW clades.

The role of NP in RNP formation has been known since the 1970s, with LCMV NP subsequently identified as capable of modulating cellular interferon responses. This is shown through the ability of certain strains of LCMV to prevent the activation of CTLs to LCMV-infected cells (Matloubian et al., 1990). Since then, NP has been shown to modulate specific pathways in interferon signalling, most notably inhibiting the passage of signalling after detection through blocking the phosphorylation of IKK and IRF3/7 (Russier et al., 2012; Rodrigo et al., 2012; Pythoud et al., 2012), shown in figure 10. This blockade of signalling indicates that arenaviral NPs are able to counteract direct activation of interferon stimulated genes (ISGs) by preventing the activation signals reaching the nucleus. By doing so, detection of any PAMP within the cytosol will not result in the generation of an antiviral state within infected cells, allowing infection to persist (Saito & Gale Jr., 2008; Levy & Garcia-Sastre, 2001). An outline of some of the known innate immune regulatory functions of NP is shown in figure 10.
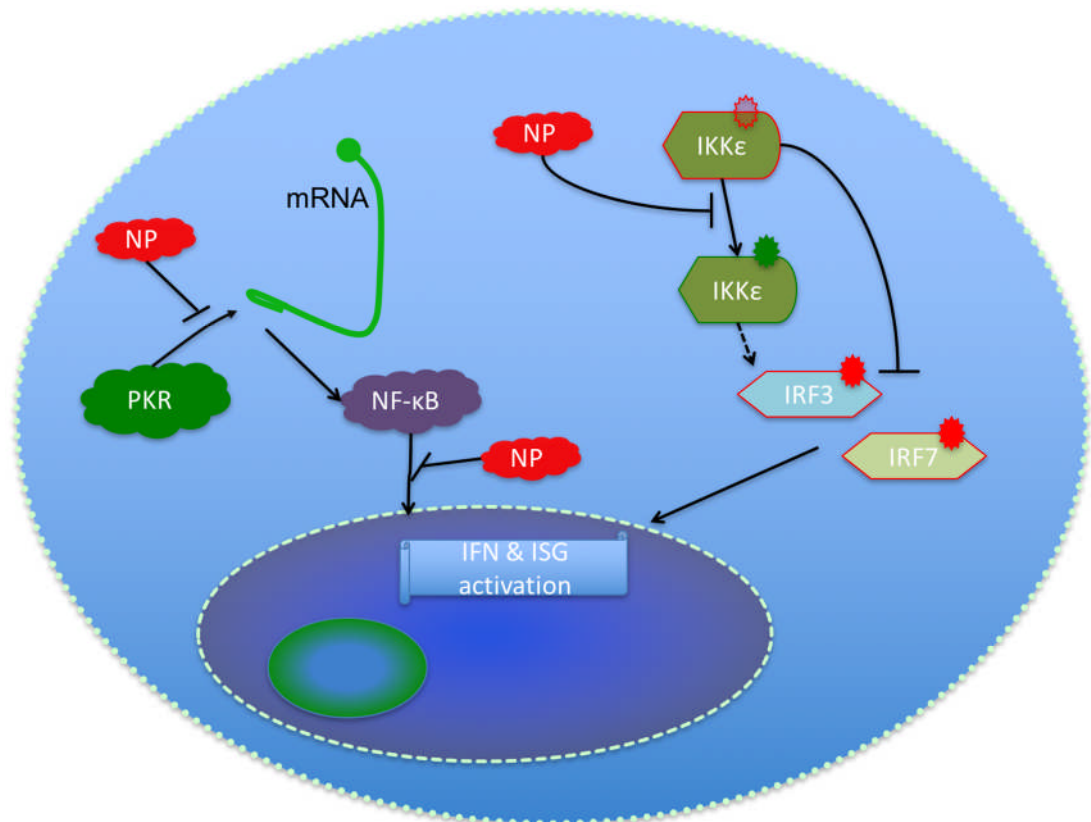
Diagram illustrating the ability of NP to disrupt interferon signalling. NP is known to inhibit RIG-I functionality and signalling to NF-κB, blocking its nuclear translocation. NP is also shown to inhibit IKK phosphorylation (red to green star), which in turn blocks IRF3 and IRF7 phosphorylation and their subsequent nuclear translocation.

Figure 10.   Diagram previously described roles of NP in the regulation cellular immune signalling cascades

NP is required to generate RNPs by encapsidating all nascent vRNAs and cRNAs within an infected cell. The association between NP and L is thought to be important to allow the appropriate removal of NP from the template and simultaneous addition of NP to nascent RNA. In the case of LACV, NP may remain in association with neighbouring NP molecules preserving the NP chain as an extended NP multimer (Gerlach et al., 2015). The RNA is thought to be transiently disassociated from the NP multimers and this exposes RNA bases to the active site of the polymerase for copying. Once replicated, the template emerges from the polymerase and is simultaneously threaded within the intact NP multimeric chain. Whether the process of arenavirus encapsidation is similar is unknown, but the LACV model is currently the most

plausible explanation proposed given the current knowledge of arenavirus RNA synthesis.

This simultaneous addition of NP to nascent RNAs will allow for the immediate generation of RNPs and thus maintain stability through disguising RNAs from the immune system (Hastie et al., 2011; Fehling et al., 2012; Iwasaki et al., 2015). Corresponding with this mechanism, the ability of NP to coat RNA means that as the genome/antigenome pool increases, the abundance of NP must increase exponentially to match it. The interaction with L is also mediated through the RNP, and thus the anti-termination effects of NP would not be present without a functional NP (Tortorici et al., 2001; Iwasaki et al., 2015). This method implies that the majority of the initial internal recruitment following viral infection of the cell will be performed by NP, as the abundance of Z and GPC will not increase significantly until the infection has been established within a cell. The elongation of nascent RNAs is dependent on NP abundance, allowing for the formation of cRNA templates for GPC and Z mRNA synthesis (Fehling et al., 2012; Tortorici et al., 2001).
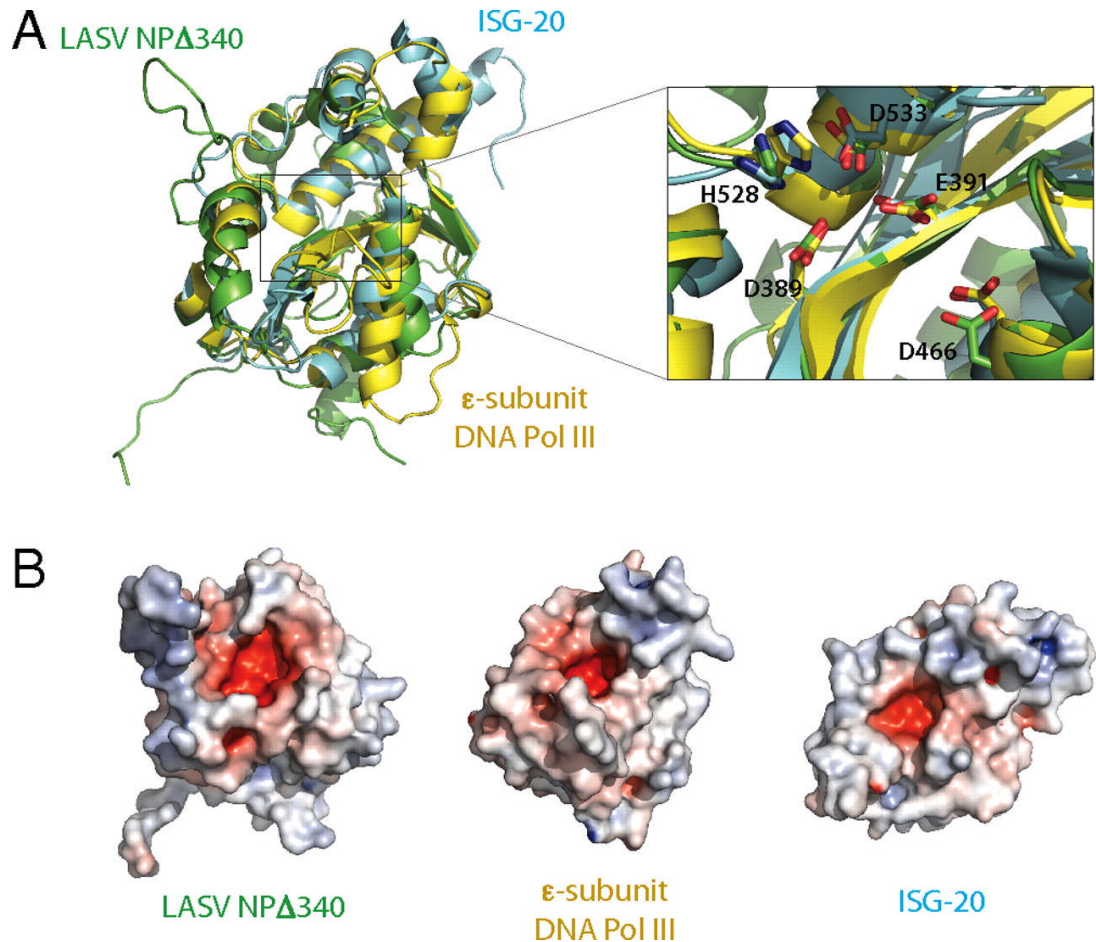
Superimposition of LASV NP and known DEDD exonucleases. (*A*) Structural comparison of NPΔ340 and two known DEDDh exonucleases. NPΔ340 is coloured green, ISG-20 (PDB ID 1WLJ) is coloured cyan, and the *E. coli* DNA pol IIIε (PDB ID 2GUI) is coloured yellow. *Inset* shows a close-up view of the superimposed DEDDh residues of the active site. Numbered residues reflect those of LASV NP. (*B*) Electrostatic surface potential calculated with APBS (the Adaptive Poisson-Boltzmann Solver) software shows that each exonuclease has an acidic active site and highlights the basic arm of LASV NP. Positive surface is coloured blue; negative surface is coloured red with limits ± 10 kT/e. (Hastie et al., 2012)

Figure 11. **Crystal structure of LASV NP and comparison with other DEDD exonucleases**

The crystal structure of LASV NP, shown in figure 11, was first solved in 2011. It revealed a dsRNA specific exonuclease domain, proposed to remove dsRNAs that could otherwise serve as an immune trigger (Hastie et al., 2011; Jiang et al., 2013; Hastie et al., 2012). The LASV NP exonuclease domain showed remarkable structural homology with other known DEDD exonuclease proteins ISG-20 and DNA Pol III. The presence of dsRNA regions within arenaviral mRNAs is well established, with LCMV NP and GP mRNA known to contain conserved 3' mRNA hairpins (Meyer & Southern, 1994). These

dsRNA sequences, if detected by the innate immune system, would ultimately lead to the stimulation of an interferon response and the activation of ISGs. In order to combat the initial detection of dsRNA, it was proposed that NP digested the dsRNA intermediary, preventing surveillance molecules ever encountering their dsRNA PAMP, illustrated in figure 12 (Hastie et al., 2012; Reynard et al., 2014). The possibility that these 3' hairpin structures would only be present for NP to then digest in order to prevent detection is puzzling, and raises the possibility that the 3' mRNA hairpins may serve yet undetermined roles, possibly in viral translation.
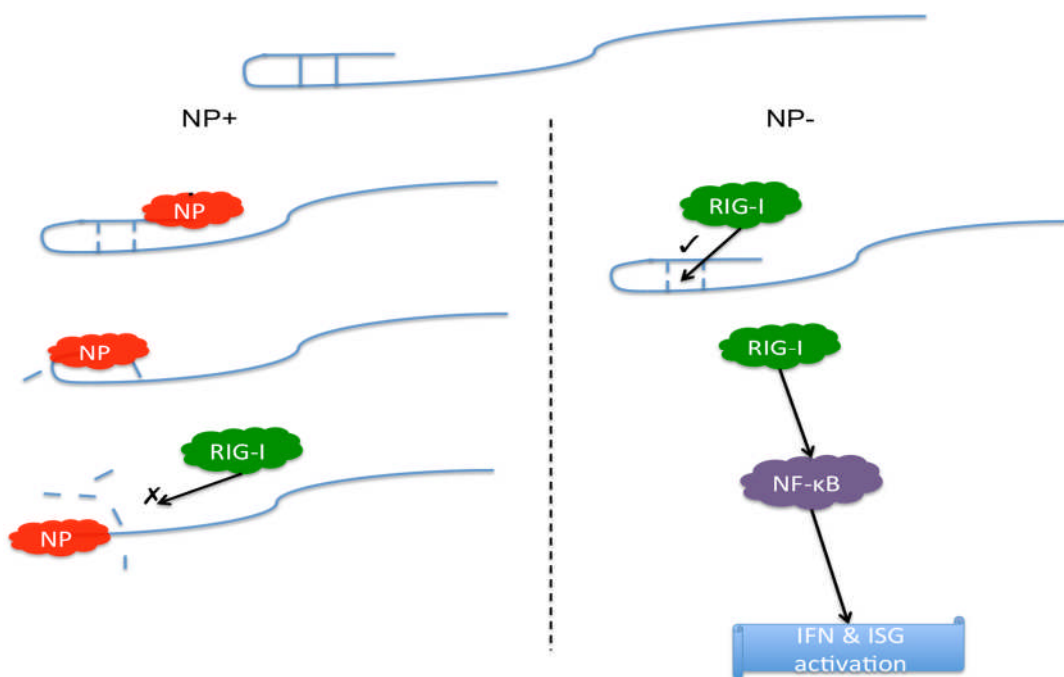
Diagram illustrating the proposed role of NP in digestion of dsRNA from 3' mRNA. In NP+, NP is shown to digest the dsRNA PAMP, preventing RIG-I (or similar surveillance proteins) detection and activation of ISG signaling. NP- shows RIG-I's normal pathway, through which dsRNA would induce ISG expression.

Figure 12.   **Proposed function of NP in dsRNA digestion**

Another intriguing aspect of the participation of NP in the viral life cycle is its involvement within so-called RTCs as diagrammatically shown in figure 13. JUNV NP was shown to co-localise with translation associated proteins inside distinct cytoplasmic puncta, present throughout the cytoplasm (Baird et al., 2012). The re-distribution of these factors was not attributed directly to NP, although cytoplasmic puncta were observed during NP expression, shown

schematically in figure 13. NP may play a role in establishing such structures or may simply be present due to high concentrations of viral RNAs. Sequestration of cellular factors – and the 'hiding' of viral replication – is not without precedent. Both HCV and DENV stimulate the formation of new sub-cellular structures as virus factories for the synthesis of viral components and RNA replication in order to isolate PAMPs from immune surveillance proteins such as PKR and RIG-I (Welsch et al., 2009; Gosert et al., 2003). RNA replication, transcription and translation are theorised to take place within these structures, yet their precise sub-cellular origin is still unknown.

NP has been shown to interact with a sequence of translation initiation factors, most notably the eukaryotic initiation factors (eIF) 4A and eIF4G, with a proposed model whereby NP substitutes the function of eIF4E in translation initiation through 5' cap binding (Linero et al., 2013). This association provides evidence that NP could have an active role in facilitating the translation of arenaviral mRNAs at the expense of cellular mRNAs, although this is yet to be assessed.
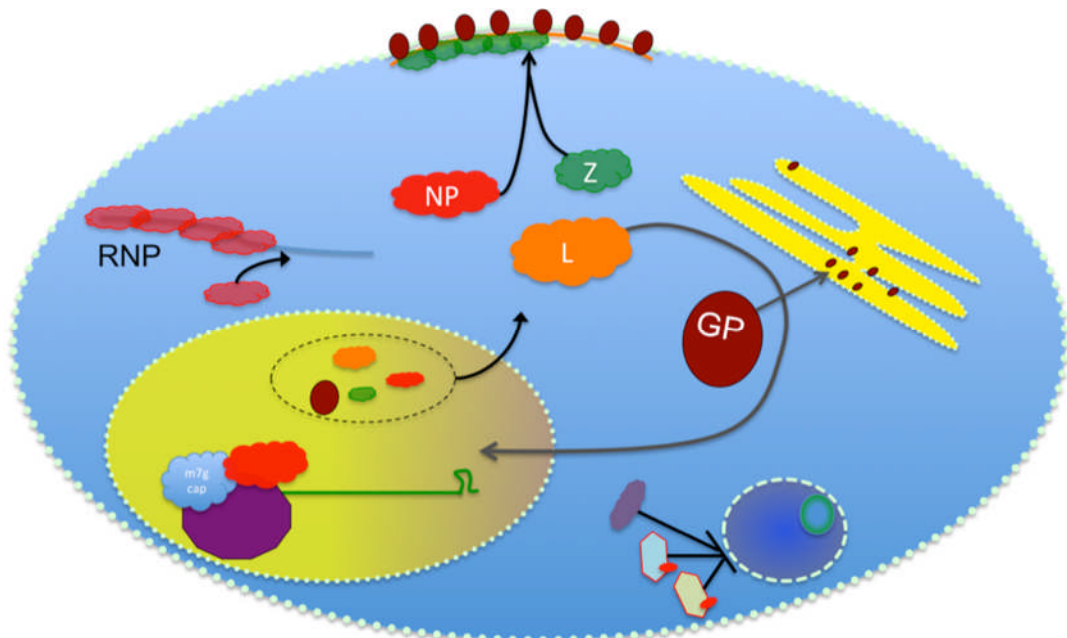
Diagram illustrating the proposed use of RTCs by arenaviruses – notably JUNV – shown as the yellow oval in the lower left. These structures were shown to contain ribosomal subunits – with a ribosome depicted by the purple dodecahedron within the RTC. Other elements of life cycle such as viral assembly (top), GP maturation (right), and NF-κB/IKK inhibition (lower right) are also shown.

Figure 13. **Diagram illustrating proposed formation of cytosolic puncta for arenaviral replication**

Potential involvement in translation would indicate a novel role for NP in the direct propagation of infection. The 3' hairpin structures on mRNA could play a role in translation. The role of this structure is unknown, but, given the lack of poly(A) tails present on viral mRNAs, the structure could facilitate the circularisation of mRNA through interaction with PABP, eIF4G or another host or viral protein. One possibility is that NP substitutes for one of these eIF4F complex components. A model proposing a similar substitution of the entire eIF4F complex has been proposed for other negative-stranded RNA viruses in the Hantavirus genus of the *Bunyaviridae* family (Mir & Panganiban, 2008)*. The NSP3 of rotavirus is known to substitute the function of PABP in 3' mRNA and eIF4G binding. The active sequestration of PABP into the nucleus during bunyamwera virus (BUNV) infection by NSs is similarly intriguing, with cellular mRNAs left un-circularised, however, no cellular or viral protein has yet been identified to circularise BUNV mRNAs, but it is possible that a cellular protein – such as stem loop binding protein (SLBP) – is responsible for BUNV mRNA circularisation. From a cellular mRNA perspective, poly(A)-less histone mRNAs possess a stable stem-loop structure, bound by SLBP, with SLBP binding to eIF4G to confer circularisation.
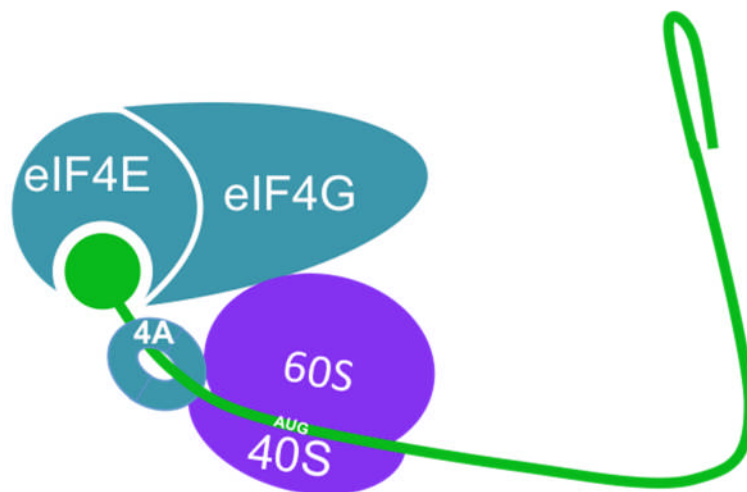
Diagram showing proposed arenaviral mRNA-protein interactions in the absence of circularisation. Circularisation could be achieved via PABP, direct interaction between eIF4G, or an intermediary such as NP binding with the proposed 3' stem loop.

Figure 14. **Diagram illustrating an arenaviral mRNA in the absence of circularisation**

Without circularisation, translational efficiency would be reduced, impacting on the virus' ability to propagate. It is therefore possible that arenaviruses possess the ability to circularise their mRNAs, in order to maximise translational output and limit mRNA degradation, with such an organisation depicted in figure 14 (Kahvejian et al., 2012; Wells et al., 2012; Jackson et al., 2010). In addition, many viruses initiate a host cell translation shut-off to increase viral translation output, such as for influenza and herpes viruses (Covarrubias et al., 2009; Yanguez et al., 2011). Maximising the availability of translation factors through the sequestration of host translation factors by a viral protein would also lower cellular translational efficiency while simultaneously increase viral translational output.

NP, given its wide role in immune regulation, is an excellent target for antiviral therapy, as perturbing its ability to shut-down the immune response could allow the host to clear the virus, rather than relying on virocidal therapies. Most viral therapeutics aim to target viral proteins and pathways specifically, and it is therefore necessary to identify crucial elements of the viral life cycle in developing antivirals. One of the most successful anti-viral compounds is acyclovir, which is a modified nucleotide derivative, for the effective treatment of herpes simplex and varicella virus infections through specific targeting of the viral polymerase. However, unlike other nucleoside analogues such as ribavirin, it cannot be incorporated into human DNA, as it requires initial metabolism by a viral thymidine kinase prior to cellular processing. The final acyclovir tri-phosphate is then incorporated by viral DNA polymerase, inhibiting strand elongation (Gnann et al., 1983).

Most other antiviral agents confer a degree of an effect on host processes, with ribavirin's side-effects during LASV and HRSV treatment well documented, as discussed above (US National Library of Medicine, 2015). Novel therapeutics targeting viral processes such as sofosbuvir's targeting of HCV RNA polymerase enable greater specificity for the treatment of certain

RNA viruses, without the severe adverse effects observed during ribavirin therapy (Fung et al., 2014). The drawback for targeting viral processes is the possibility of resistance emergence. However, identifying a host process, not necessary for human life, which the virus must utilise could also allow for potent antiviral identification (Sayce et al., 2010). This strategy could also limit the potential for resistance development.

1.3. Studied viruses in detail

1.3.1. LUJV and NP

Lujo virus was first isolated from five patients from southern Africa in 2008, with the index case presenting to a clinic in the Zambian capital Lusaka with acute febrile illness. This patient was subsequently flown for treatment to Johannesburg in South Africa, with an attending paramedic (patient 2) on the flight becoming exposed before returning to Zambia. Three more people were exposed and subsequently became infected, one nurse (patient 3) attending to patient 1, a cleaner (patient 4) attending to patient 1's room and a nurse attending to patient 2 (patient 5). Patients 1-4 all died within 26 days of exposure. Patient 5 survived after high dose IV ribavirin – although the role of ribavirin in this recovery is unknown – as well as immunosuppressants and clotting factor administration in a specialist clinic in Johannesburg (Sewlall et al., 2014). Patient 5 was finally discharged after 42 days, with symptoms persisting for several months post-clearance. The severity of illnesses described in the doctors' reports indicated a rapid onset febrile infection, generating haemorrhagic complications. Initial diagnoses of alternative, endemic agents such as malaria or dengue virus may have contributed to the nosocomial transmission.

LUJV is a disease of acute research interest. The human population in areas of southern Africa, such as Zambia, is rapidly expanding, with increasing tourism and resources-trade stimulating growth and expansion into rural areas. Patient 1 was thought to have become exposed to LUJV at home, with

45

evidence of rodents on their property, although no rodent carrying LUJV has yet been identified (Ishii et al., 2012; Sewlall et al., 2014).

Since 2009, work has been performed to learn more about LUJV, with the establishment of a stable animal model for human infection in clone 13/N Guinea Pigs. LUJV was able to infect multiple organs, generating histological changes in the kidneys, liver and lungs. Persistent weight loss of between 3-5% per day, with all animals succumbing to the infection in between 11-16 days (Bird et al., 2012). The disseminated nature of infection mimicked the progression observed in humans, with similar clinical manifestations such as leukopenia, lymphopenia and coagulopathy (Bird et al., 2012). Individual tissue investigations showed a large degree of internal haemorrhaging and necrosis.

The individual differences of the functions of LUJV proteins to those of related arenaviruses is incompletely understood, with limited knowledge of several aspects of the LUJV life-cycle. Classified as an ACDP BSL4 pathogen there are significant restrictions on working with infectious LUJV, and identification of protein functionality is challenging in the context of a viral infection.

The roles of LUJV NP are likely to be similar to the NPs of other arenaviruses - predominantly involved in RNP formation, RNA stability and transcriptional anti-termination. However, other arenaviral NPs have roles, as discussed, outside these functions. LUJV NP is likely to possess similar functions, such as interferon induction ablation, as other arenaviruses are known possess. The nucleotide variation between LUJV and LASV NP is up to 38%, with LUJV an apparent outlier of the OW clade of viruses (Sewlall et al., 2014; Briese et al., 2009). This diversity could confer significant differences between LUJV and other members of the OW clade such as LASV and LCMV.

As NP is known to be crucial in the replication of other arenaviruses, this study will focus predominantly on the identification, and characterisation, of important interactions between LUJV NP and cellular proteins. The lack of knowledge regarding aspects of the arenaviral replication cycle such as

translation and RNA replication suggests that more functions for each arenavirus protein are likely to be established.

1.3.2. LCMV

LCMV is the most geographically widespread arenaviral infection, with an almost worldwide distribution through its association with the *Mus musculus* common house mouse. The life cycle of this host has become reliant on human dwellings, through millennial association with humans, with this association even instrumental in the domestication of the cat (Boursot et al., 1996; Cucchi et al., 2005). This reliance on humans has drawn them into co-habitation with humans, and thus exposes individuals to diseases carried by them.

LCMV infection generates a predominantly mild or asymptomatic illness in humans, and is thus generally undiagnosed. Squalid habitation most likely increases potential exposure to LCMV and, through the mouse, it was most likely transported worldwide through colonial era naval trade.

LCMV is a neglected pathogen with potential to cause human disease, chiefly among the immunocompromised and pre- and postnates. LCMV is capable of generating severe infections within immunocompromised patients and expectant mothers, and with the increase in the immunocompromised population, the risk that more patients will become terminally infected is growing. Developmental abnormalities in children infected with LCMV in a pre- or perinatal setting are also of distinct interest, with LCMV being a significantly undiagnosed pathogen of new-borns and young children (Barton et al., 2002; Jamieson et al., 2006; Bonthius, 2012). The recent and ongoing Zika virus outbreak (2015-16) gives evidence that a predominantly low-level, generally asymptomatic, infection can have devastating consequences on health in specific circumstances (Attar, 2016; World Health Organisation, 2016).

The NP of LCMV is perhaps the most well researched protein of all of the arenaviruses and it has been identified as a significant modulator of cellular interferon signalling, especially through IKK and IRF3 ablation, as shown in

figure 8 (Martinez-Sobrido et al., 2009; Pythoud et al., 2012). Characterising common strategies between LCMV NP and LUJV NP is key to identifying potential therapeutic strategies to combat arenaviral infections.

Identifying a common trait between two OW arenaviral species could lead to the identification of a pathway to target via a broad spectrum antiviral agent. There is a significant need to improve therapeutics beyond the contraindicated use of IV ribavirin, especially pre and during pregnancy (US National Library of Medicine, 2015). Vaccination is the ultimate aim in all infectious diseases, but given the huge diversity in arenaviral distribution – geographical, socio-economic factors and religious belief – the realities of a vaccination programme currently make it impractical. Thus, for emerging infectious diseases, effective therapeutics are vital.

1.4. <u>Nucleoprotein as a target for research and therapeutics</u>

The NP of arenaviruses is a crucial component of the viral replication cycle. It possesses the ability to modulate the immune response; assist in cap-snatching; act as an anti-terminator; stimulate transcription-replication switching; and maintain stable RNA species as part of an RNP.

These functions are evidence that NP's effective functionality is essential to the replication-cycle of arenaviral infections. Many of these abilities are shared between species – even between the two OW and NW serotypes. Given the lack of effective, targeted therapeutics and the need to identify novel, pan-tropic functions of NP, more research is needed in order to identify a pathway which could be targeted for therapy.

Several aspects of the arenaviral replication cycle are still poorly understood, including translation of viral mRNAs and where the replication centres, or RTCs, are derived from. Direct protein interacting partners for all arenavirus proteins are only known to a limited extent, with incompletely understood mechanisms of cell entry; how GP is translocated to the ER; which cytoskeletal network Z and NP exploit; translational enhancement and factor recruitment.

It is therefore necessary to identify novel and fundamental elements regarding the involvement of NP in viral replication cycle, and host cell manipulation. As discussed, interactions between NP and cellular proteins have been previously identified. In addition, the nucleocapsid proteins (N) of the related CCHFV and HAZV are known to interact with protein folding machinery (Surtees, 2014). The involvement of NP in this pathway could indicate that NP assists in the folding or glycosylation of arenaviral proteins. As the NP of LASV and CCHFV N are known to be structurally similar (Carter et al., 2012), there is the possibility that features of their functionality are also conserved.

One possible target for therapeutic development is the potential role of the 3' hairpins on arenaviral mRNA transcripts. These structures are, for a freely expressed RNA rather than one encapsidated within an RNP, illogical. dsRNA is a well known PAMP, recognised by a number of cellular surveillance proteins. IRES structures and other dsRNA elements present on may positive sense RNA viruses might similarly act as immune triggers for internal surveillance mechanisms, but the overall genome-ordered RNA structure (GORS) is thought to provide protection from such detection (Simmonds et al., 2004). The presence of a singular structured hairpin present on arenaviral mRNAs is unlikely to provide similar protection. The evolutionary selection pressures on expressing such a feature would indicate that, if unnecessary for replication, it should have been removed through selection. The knowledge of a dsRNA specific exonuclease, potentially digesting such a structure, would seemingly move to solve this problem. However, the dsRNA specificity of the exonuclease has been questioned (Hastie et al., 2012; Qi et al., 2010), and it is possible that the digestion of dsRNA occurs after the mRNA has served its purpose in the context of replication. As discussed, there is limited knowledge regarding the efficient translation of arenaviral mRNAs, excluding the known interactions between JUNV NP and eIF4A and eIF4G (Linero et al., 2013). How mRNA can be circularised without a poly(A) tail is an unanswered question regarding efficient translation of mRNA, and the 3' tail could have a role in facilitating this. Encoding a key feature of a highly multifunctional protein to simply degrade an encoded immune marker would be a convoluted

mechanism, and thus the 3' mRNA hairpin likely has another role in the arenavirus replication cycle.

Project Aims

The arenavirus NP is indispensable to virus multiplication. While its roles in RNP formation, RNA synthesis, segment packaging and innate immune regulation have been described, it is likely that additional roles exist. It is likely that these established and additional roles include important functional interactions with cellular factors.

The aim of this project was to elucidate the cellular interaction partners of the arenavirus NP. The two arenaviruses chosen were the HF-causing Lujo virus along with the congenital pathogen LCMV The aim of studying these two distinct viral species from across the OW clade was to identify common partners of NP that were potentially important across an entire clade. LUJV NP was used to initially catalogue a cellular interactome through the use of SILAC proteomics, in conjunction with immunoprecipitation (IP) and immunofluorescence microscopy (IF). In addition, in order to better visualise NP, a polyclonal antiserum against LUJV NP was generated. The use of LCMV then also allowed for the experimental examination of the observed interactions in the context of an infectious virus system. Finally, characterising the relationship between LCMV NP and identified interacting partners led to the final objective of the project which was to determine whether NP enhanced arenavirus mRNA translation through a suggested novel mechanism.

2. **Materials and Methods**

2.1. **Biological Materials**

2.1.1. **Vectors**

The LUJV NP cDNA corresponding to the LUJV NP ORF was synthesised by Dundee Cell Products and was provided cloned into an Enhanced green fluorescent protein (EGFP)-N1 expression vector, as a LUJV NP C terminal fusion protein. Expression of this protein is driven via a CMV promoter, generating strong and consistent expression upon entry into mammalian cell types. The parent EGFP-N1 also contains an SV40 poly(A) signal, increasing RNA stability, and enabling enhanced mRNA translation. The provided sequence was confirmed to match the GenBank JX017360.1 sequence via DNA sequencing.

The LUJV NP ORF from pEGFP-N1 vector was subcloned into a pET28a vector, as a his-SUMO- N terminal fusion protein. pET28a enables good expression levels of his-SUMO tagged proteins in *Escherichia coli (E. coli)*, enabling Ni affinity purification of fused proteins.

A plasmid designed to express an arenaviral-like mRNA was constructed by GeneArt in a pMK parental vector. The plasmid was comprised of 5' sequences of LCMV mRNA, from the 5' nucleotide of the LCMV antigenome up until the final nucleotide prior to the initiator AUG; the complete *Gaussia* luciferase sequence as obtained from GenBank AY015993.1; followed by the 3' mRNA sequences, as described by Meyer and Southern (1994). This plasmid was termed NPEG (NP Ends *Gaussia*).

See Appendix I for both his-SUMO- and –EGFP fused LUJV NP sequences, and NPEG plasmid sequences.

2.1.2. Bacterial Strains

Plasmid DNA constructs were amplified through transformation of *E. coli*, strain DH5α Competent Cells (Life technologies)*,* using the Inoue method for transformation. For protein expression, the *E. coli* BL21 Rosetta-2 strain was used.

2.1.3. Continuous cell lines

Human embryonic kidney 293T (HEK293T), Baby Hamster Kidney fibroblast clone 21 (BHK21), Vero E6 cells (Vero) and adenocarcinomic human alveolar basal epithelial cells (A549 cells) were utilised to express recombinant proteins or to permit viral replication. HEK293T cells constitutively express the SV40 large T antigen allowing for continuous growth. Both BHK21 and Vero cells show natural continuous growth with no external influence preventing senescence (Ammerman et al., 2008; Hernandez & Brown, 2010). A549 cells were originally explanted from an human adenocarcinoma (Giard et al., 1973).

2.1.4. LCMV strain.

LCMV strain Armstrong (LCMV Arm) was kindly provided by Prof. Roger Hewson of Public Health England (PHE) as an infectious cell culture supernatant from Vero cells. cRNA sequences for both S and L segments of LCMV Arm are available via GenBank with accession numbers AY847350.1 and AY847341.1 respectively.

2.2. Methods

2.2.1. Manipulation of cDNA

2.2.1.1. Bacterial Transformations

For routine plasmid amplification, DH5α cells (as in 2.1.2) were transformed according to the manufacturer's recommendations. Briefly, 50ng of plasmid DNA or 1μl of ligation reaction product was mixed with competent cells for 30 minutes on ice; followed by a 45 second 42°C heat shock and cooled on ice

for 2 minutes. 1000 μl of Luria-Bertani (LB) media (10g tryptone, 10g NaCl, 5g yeast extract, autoclaved in 1L ddH$_2$O) was added to the transformation mixture, and incubated, while shaking, at 37°C for 1 hour. Cells that had settled were then agitated by shaking, with 200μl being spread onto corresponding antibiotic LB agar plates and incubated at 37°C for 16 hours.

2.2.1.2. Agarose gel electrophoresis

The integrity, purity and size of plasmid DNA, PCR products, restriction digests and ligation reactions was determined using 1% agarose gels (0.5g analytical grade agarose (Sigma Aldrich) in 50 ml 1x TAE (40mM Tris-Acetate, 1mM ethylene-diamine-tetraacetic acid (EDTA); with SYBR Safe DNA stain (Life Technologies) at 1 1:10000 dilution). DNA samples were mixed to a concentration of 1x Orange G DNA loading dye (20g sucrose and 100mg Orange G dissolved in 50 ml ddH$_2$O. Samples were loaded alongside Hyperladder 1 (Bioline) for relative size comparison and subsequently run at 90V for 30-50 minutes in 1x TAE buffer. Blue light transillumination was used for DNA visualisation.

2.2.1.3. Restriction enzyme digestion

Plasmid DNA and PCR products were subjected to restriction enzyme digestion for cloning or diagnostic purposes. Reaction volumes of 50μl were used throughout, containing 1.2 μg of DNA; 1x CutSmart reaction buffer (New England Biolabs (NEB)) 2 units of appropriate enzymes (NEB), with nuclease free H$_2$O added to give a final volume of 50μl. Reactions were incubated using their appropriate reaction temperatures. Products were then purified by agarose gel electrophoresis, followed by gel extraction utilising a Zymoclean Gel DNA recovery kit (Zymo Research), following the manufacturer's recommendations.

2.2.1.4. Polymerase chain reaction (PCR)

Subcloning of cDNA was necessary for the generation of his-SUMO-LUJV NP within a pET28a vector for the purification of LUJV NP. The LUJV NP ORF from the pEGFP-N1 expression plasmid was amplified by PCR to incorporate terminal restriction endonuclease sites to facilitate subcloning into the pET28a plasmid. PCR reactions were performed in 20μl reaction volumes containing 50 ng Template DNA; 0.4μM of both forward and reverse primers; 0.3mM of dNTPs (Roche), 10x Termopol buffer – diluted to final concentration of 1x (NEB) and 1 unit of Vent polymerase (NEB). Reaction cycles were performed in a thermocycler (Eppendorf); 5min 95°C denaturation; 35 cycles of 95°C, 30 seconds – 50-65°C (dependent on primer T_m), 30 seconds – and a 90 seconds 72°C elongation. A final 5-minute extension at 72°C was performed, before being cooled to 4°C. Products were then purified via agarose gel electrophoresis and gel extraction.

2.2.1.5. Ligations

Ligation reactions were performed at three ratios of DNA (insert:vector) to ensure maximum efficiency; at 3:1, 6:1 and 9:1. 1x Ligase reaction buffer (Promega (30mM Tris-HCl (ph7.8); 10 mM $MgCl_2$; 10mM DTT; 1mM ATP)) and 1 unit of T4 DNA Ligase (Promega), with nuclease free H_2O added to give a final volume of 20μl. Reactions were incubated in a cooling water bath – from 21°C to 4°C – overnight.

2.2.1.6. Plasmid DNA amplification

Overnight cultures, grown at 30°C from either picked colonies, glycerol stocks or an appropriate starter culture for larger scale amplification, were pelleted by centrifugation at 4000 x*g* at 4°C for 10 minutes. Plasmid DNA was isolated and purified from bacterial cell pellets using appropriate Thermo Scientific Plasmid DNA extraction kits according to the manufacturer's instruction. Briefly, plasmid DNA was isolated via alkaline lysis, isolation on silica membranes under high salt concentrations and elution from the silica

membrane after ethanol based washing. After purification, a NanoDrop 1000 (Thermo Scientific) was used to quantify DNA by spectrophotometry.

2.2.2. Protein expression and purification in *E. coli*

2.2.2.1. Culture media and induction.

Initially, 50ml starter cultures were grown from picked colonies of pET28a-his-SUMO-LUJV NP transformed *E. coli* Rosetta-2 cells and used to inoculate 1L of LB containing 100 μg/ml kanamycin. This growth culture was then allowed to reach an OD_{600} of 0.6-0.8 by growth at 37°C. Cells were then induced with 100μM IPTG for 16hrs at 18°C.

Upon further experimental analysis, a shift to using auto-induction media was taken. Auto-induction media (20g Tryptone, 10g NaCl, 5g Yeast, autoclaved in 1L ddH$_2$O, supplemented with sterilised sugars (25 ml 8% w/v Lactose; 10 ml 60% v/v Glycerol; 5ml 10% w/v Glucose) and 100 μg/ml kanamycin) was inoculated with stabs from pET28a-his-SUMO-LUJV NP transformed *E. coli* Rosetta-2 glycerol stocks, grown at 18°C for 60 hours.

2.2.2.2. Bacterial lysis for protein purification

Bacterial cells from 1.2.2.1 were recovered by centrifugation at 4000 x*g* at 4°C for 15 minutes. The supernatant discarded and cells resuspended in lysis buffer (500mM NaCl; 20mM Tris-HCl pH 7.5; 0.5% v/v Triton X-100; 1mg/ml chicken egg white lysozyme (Sigma Aldrich); 1-unit DNase; 1-unit RNase with the addition of cOmplete Ultra protease inhibitor cocktail tablets, as per the manufacturer's instruction (Roche)). 15ml of lysis buffer was used per 1 L of growth culture. Cells were incubated in lysis buffer with agitation at 4°C for 30 minutes, followed by a sonication cycle of 10 cycles of 10 second 10μm amplitude bursts, followed by 10 seconds off – repeated three times. The samples were kept on ice throughout, with 2 minute incubations between cycles. Lysates were clarified at 40,000 x*g* at 4°C for 20 minutes. The

supernatant was retained for his-SUMO-LUJV NP purification via Ni affinity chromatography. Based on work by (Tanner et al., 2014).

2.2.2.3. Ni affinity chromatography

Super-NiNTA100 resin (Generon) was clarified with 4 column volumes (CV) of his-Binding Buffer (BB) (500mM NaCl; 20mM Tris-HCl pH 7.5; 20mM imidazole), with 2ml of resin slurry clarified per 1L of growth culture (1CV). Bacterial cell lysate was then allowed to flow through the resin at a flow rate of 0.5 ml/min. Washing buffers were prepared, with the composition identical to the BB, with increased imidazole concentrations. All buffers were prepared and used at $4°C$. Each washing step utilised 5 CV of corresponding wash buffer. Elution of affinity tagged protein was performed at 1M imidazole concentration (2 CV), with eluted protein concentration determined by adding 1 μl sample into a 1:5 dilution of Protein Assay Reagent (Bio-Rad) in ddH$_2$O and reading the absorbance at 595nm, verses an elution buffer control. Eluted his-SUMO-LUJV NP was then incubated overnight at $4°C$ under dialysing conditions with SUMO-Protease (expressed and purified in-house). The dialysis conditions were: (Gel Filtration Buffer (GFB)) 500mM NaCl; 20mM Tris-HCl pH 7.5; 1mM EDTA; 1mM DTT – prepared at $4°C$. Based on work by (Ariza et al., 2013)

In final purification stages, after size exclusion (2.2.2.4), samples were re-clarified over a 5ml His-Trap column, utilising the same buffer conditions as above. Elution was performed at 500mM imidazole.

2.2.2.4. Size exclusion chromatography

Size exclusion chromatography was used to further purify LUJV NP from both the his-Sumo tag and the SUMO protease. A 320ml HiLoad 26/600 Superdex S75 column (GE Healthcare) was used with and Akta Prime pump and collection system at $4°C$, with a 280nm absorbance sensor to determine protein concentration in eluted samples.

Prepared GFB was de-gassed and filtered using a 0.2μm sterile filter. The S75 column was equilibrated with degassed GFB. Sample protein was concentrated to a volume ≥5.1 ml using a 10 kDa MWCO Vivaspin column (Sartorius) and filtered using a 0.2μm filter. Sample was loaded into a clarified 5ml injection loop. The pump and collection system was instructed to collect 3ml fractions after the void volume ~90ml had passed through, with a flow rate of 0.3ml/min. Columns were routinely calibrated using a gel filtration buffer standard kit (Bio-rad). Fractions were subsequently collected, with samples analysed via SDS PAGE.

2.2.2.5. Immunochallenge

1.5 mg of purified NP was used to inoculate a sheep for the production of polyclonal antibodies. Inoculations were performed by Alta Bioscience, with a primary inoculation via intra-muscular injection followed by two booster doses. Antigens were administered in the presence of Freund's adjuvant. Pre-immune sera was collected, and subsequent collections of immune sera after primary, secondary and tertiary challenges by the antigen. It was clarified by centrifugation, aliquoted and frozen at -20°C.

2.2.3. Mammalian cell culture

2.2.3.1. Maintaining continuous cells in culture

All cells were maintained at 37°C in a humid, 5% CO_2 atmosphere. All cell types listed in 2.1.3, in the absence of viral infection, were maintained in "complete media": Dulbecco's modified Eagle's medium (DMEM, Sigma Aldrich), supplemented with 10% v/v foetal bovine serum (FBS, Sigma Aldrich); 100 IU penicillin/ml and 100μg streptomycin/ml (Life Technologies).

Cells were passaged using trypsin-EDTA (Sigma Aldrich) upon reaching 90-95% confluency. Cells were seeded into specific culture flasks and plates at appropriate densities after counting with a haemocytometer: 12 well plates (1x10^5 cells/well); 6 well plate (2x10^5 cells/well); 10cm dish (1x10^6 cells/dish)

and T175 flasks ($5\text{x}10^6$ cells/flask). For immunofluorescence experiments, autoclaved glass coverslips (VWR) were placed into wells prior to seeding.

2.2.3.2. Cryogenic storing and thawing of cells

Long-term storage of cells required freezing in liquid nitrogen. Cells were trypsinised and washed twice with cool, followed by ice-cold PBS, then centrifuged at 500 x*g* at 4°C for 10 minutes, followed by a further wash step and centrifugation. Cells were counted, and resuspended in ice-cold FBS, 10% DMSO to a final concentration of $1\text{x}10^6$ cells/ml. Cells were then aliquoted into cryovials, cooled to -80°C at a controlled rate of -1°C/minute (CoolCell, biocision). Cells were then transferred to liquid nitrogen storage. Frozen cells selected for re-animation were rapidly thawed at 37°C, centrifuged at 500x*g* at 4°C for 5 minutes, washed in complete media, resuspended, and transferred to 25cm^2 flask. Cells were subsequently expanded when reaching 80-90% confluency.

2.2.3.3. Cellular transfection using Lipofectamine 2000

HEK293T, Vero and A549 cells were transfected for varying experimental procedures, using Lipofectamine 2000 (Life Technologies). The utilisation of the lipofectamine 2000 method ensures minimal toxicity and high efficiency of DNA introduction. The cationic liposomes enable DNA to complex within liposomes, and enable plasma membrane fusion and introduction of genetic material into target cells (Dalby et al., 2004).

For the transfection of plasmid DNA, cells were allowed to reach 70% confluency (60% in the case of experiments involving IF), with transfection agent:DNA:Media ratio dependent on well size. For SILAC-based and Interaction validation transfections, 10cm dishes were used. For experiments involving IF, 12 well plates used. All volumes and ratios were as per the manufacturer's recommendation. Briefly – in the case of 10cm dish transfections – 20μg plasmid DNA was mixed 500μl DMEM and incubated for 5 minutes (Tube A); 64μl Lipofectamine 2000, mixed with 500μl DMEM and

incubated for 5 minutes (Tube B). Tube B was then added drop-wise into Tube A, and incubated for 20 minutes. Cells were supplemented with fresh DMEM, with 1ml of transfection mixture added per dish. After 4 hours, DMEM was aspirated and replaced with complete media.

2.2.4. Mammalian cell protein expression analysis

2.2.4.1. Preparation of whole cell lysate

Cells were harvested via physical scraping whilst in culture media. Cells were recovered via aspiration, and centrifuged at 1000xg for 10 minutes, washed 3x with ice-cold PBS, with centrifugation between each wash. Cells were then resuspended in lysis buffer (10 mM Tris-HCl pH 7.5; 150 mM NaCl; 0.5 mM EDTA; 0.5% NP40; 1x EDTA-free cOmplete protease inhibitor cocktail (Roche)) for IP. Cells were incubated on ice, with pipette mixing every 10 min, for 30 minutes. Lysates were then clarified at 20,000xg at 4°C for 10 minutes.

2.2.4.2. BCA assay determination of protein concentration

Total protein concentration was determined from whole cell lysates using the micro bicinchoninic acid protein assay system (BCA; Pierce), used as per the manufacturer's instruction. Bovine serum albumin (BSA) standards were prepared from the provided stock in order to prepare a standard curve.

2.2.4.3. Immunoprecipitations involving GFP-Trap and RFP-Trap

Immunoprecipitations (IP) were used to 'pull-down' EGFP or LUJV NP-EGFP for SILAC-based proteomics and subsequent validations. GFP-Trap IPs involve the use of alpaca-anti GFP antibodies, conjugated to agarose beads. The alpaca antibody, of the camelid family, consists of a single domain specific to GFP and EGFP, with the RFP-Trap control specific to only RFP. GFP/RFP-Trap beads were equilibrated in ice-cold dilution buffer (10 mM Tris-HCl pH 7.5; 150 mM NaCl; 0.5 mM EDTA; 1x EDTA-free cOmplete protease inhibitor cocktail (Roche)), via two washing and centrifugation steps (2000xg,

4°C). GFP-Trap and RFP-Trap beads were all obtained from Chromotek, and handled according to the manufacturer's recommendations.

In the case of SILAC IPs, only GFP-Trap beads were used. In the case of validation experiments, both GFP- and RFP-Trap beads were used, with lysates divided 50-50 and diluted as such.

Lysates from 2.2.4.1 were diluted to lower final NP40 concentration below 0.1% and a final volume of 1ml with dilution buffer and exposed onto 30µl per IP equilibrated GFP/RFP-Trap beads. Cell lysates were incubated with GFP-Trap beads for 90 minutes at 4°C, with rotation, to allow EGFP/LUJVNP EGFP to bind to GFP-Trap beads. Following the incubation period, the beads were sedimented at 2000xg and supernatant removed (known as flow through (FT)), with the beads washed twice in ice-cold wash buffer (10 mM Tris-HCl pH 7.5; 250 mM NaCl; 0.5 mM EDTA; 1x EDTA-free cOmplete protease inhibitor cocktail (Roche)). Immunoprecipitated proteins were then eluted from the beads via heating at 95°C in 50µl 2x NuPAGE LDS sample buffer (4x NuPAGE LDS sample buffer (Thermo Fisher Scientific) – diluted to final, 2x concentration in ddH$_2$O; 10mM DTT). Beads were then sedimented at 2700xg for 5 minutes, 4°C.

In the case of SILAC IPs, equal volumes of eluted proteins were combined, and sent to Dr Stuart Armstrong at the University of Liverpool for mass-spectrometry (MS) analysis.

2.2.4.4. SDS polyacrylamide gel electrophoresis (SDS PAGE)

5 ml SDS PAGE mini-gels were made as 10, 12 or 15% resolving gels, and 5% Stacking gels, as shown in table 2.

Reagent	5	10	12	15
	Volume (ml)			
ddH$_2$O	3.40	4.00	3.30	1.67
Polyacrylamide	0.83	3.30	4.00	5.63
1.5M Tris-HCl pH8.8	-	2.50	2.50	2.50
1M Tris-HCl pH6.8	0.63	-	-	-
10% (w/v) SDS	0.50	0.10	0.10	0.10
10% (w/v) APS	0.50	0.10	0.10	0.10
TEMED	0.05	0.01	0.01	0.01

Table 2 SDS poly-acrylamide gel recipes

Standardised protein samples were mixed 1:4 with 5x Laemmeli Buffer (60mM Tris-HCl pH 6.8; 10% (v/v) Glycerol; 2% (w/v) SDS; 5% (v/v) β-mercaptoethanol, 0.01% (w/v) bromophpenol blue), with samples subsequently denatured at 95°C for 5 minutes, prior to loading. Samples were loaded into wells at appropriate volumes for well size. Samples were loaded alongside ColourPlus Prestained protein ladder (10-230kDa) to assist in band size identification. Electrophoresis was performed with 1x SDS running buffer (25mM Tris; 192mM Glycine; 0.1% (w/v) SDS) at 200V for 45-60min.

2.2.4.5. Protein visualisation by Coomassie stain

Proteins resolved via SDS PAGE electrophoresis were visualised by incubating gels in 1x Coomassie stain (0.25% (w/v) Coomassie R250; 50% (v/v) CH$_3$OH; 10% (v/v) glacial acetic acid (CH$_3$COOH)) for one hour at RTP, with agitation. Gels were de-stained in destain buffer (40% (v/v) CH$_3$OH; 10% (v/v CH$_3$COOH) for two hours, with rehydration in ddH$_2$O prior to imaging.

2.2.4.6. Silver stain visualisation of proteins.

Silver staining of gels was performed using a SilverQuest staining kit (Invitrogen), as per the manufacturer's recommendations.

2.2.4.7. Western blot analysis

Proteins resolved via SDS PAGE electrophoresis gels were transferred to fluorescence compatible polyvinylindene fluoride membranes (PVDF;

Immobilon-P Transfer membrane, Milipore) using a semi-dry Tans-Blot (Bio-Rad) in Towbin buffer (25mM Tris, 192mM glycine, 20% methanol (v/v)) for 1 hour at 15V. PVDF membranes were subsequently blocked for 1 hour at room temperature in 1:1 Odyssey blocking Buffer-TBS (OBB; LiCor): TBS (50mM Tris pH 7.5, 150mM NaCl). Blocking buffer was then replaced with appropriate primary antibody solution in 1:1 OBB:TBS for either 1 hour at room temperature or overnight at 4°C. Membranes were washed three times in TBS and then incubated with appropriate fluorescently labelled secondary antibodies for 1hr at room temperature, protected from light. Membranes were then washed twice with PBS and finally once in water and dried, before visualising using LiCor Odyssey Sa infrared imaging system. A table of primary and secondary antibodies is provided in tables 3 and 4.

2.2.5. Stable isotope labelling of amino acids in cell culture (SILAC)

Frozen HEK293T cells were reanimated, as in 2.2.3.2, into isotopically distinct SILAC-DMEM (Dundee Cell Products), supplemented with dialysed, SILAC grade FBS (Dundee Cell Products), for a minimum of seven doublings to ensure ≥95% labelled amino acid inclusion. SILAC culture media recipes used were: 'light' R0K0 media for control samples; 'medium' R6K4 ^{13}C labelled arginine and 2D (^{2}H) labelled lysine and 'heavy' R10K8 of ^{13}C & ^{15}N labelled arginine with ^{13}C & ^{15}N labelled lysine – as depicted in figure 16. Cells were maintained as in 2.2.3.1, with the substitution of a PBS-EDTA based cell dissociation buffer (Life Technologies) trypsin for, to prevent introduction of unlabelled amino acids.

MIX AFTER PRECIPITATION
(MAP)
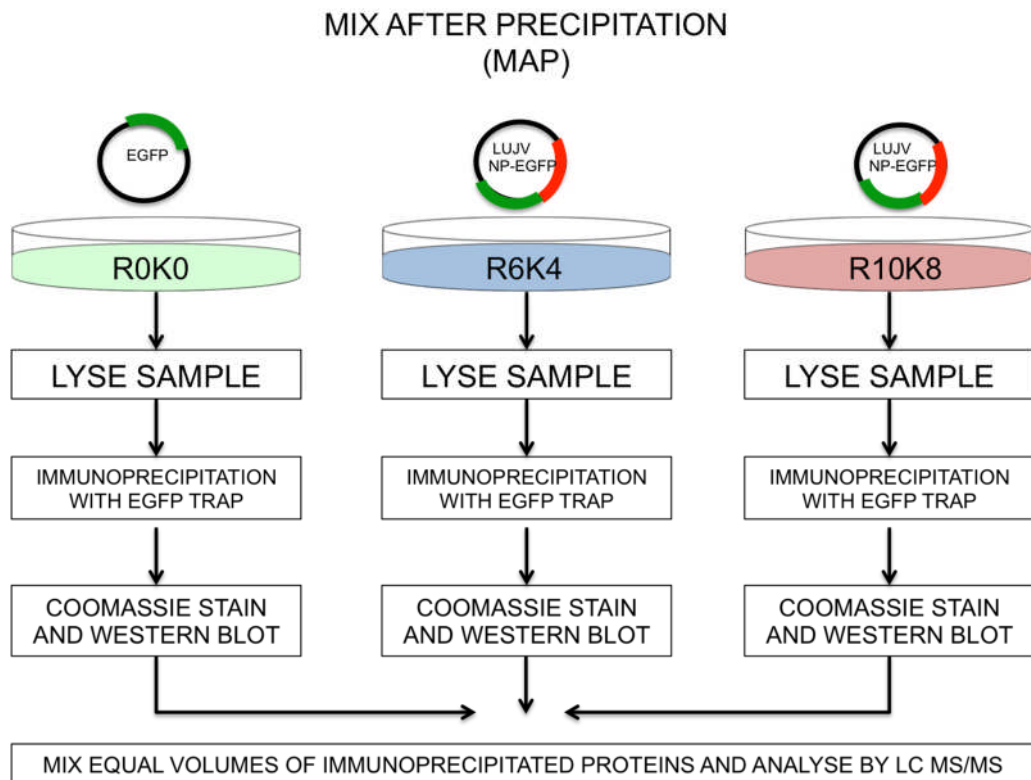
Diagram showing the mix after precipitation (MAP) SILAC IP procedure.

Figure 15. **Diagram illustrating SILAC MS procedure**

2.2.5.1. Immunofluorescence microscopy (IF)

IF experiments were performed to visualise various cellular and viral proteins and to determine their sub-cellular localisation. Cells were grown on sterilised 19mm glass coverslips within 12-well plates, and where appropriate, transfection or infection had already taken place prior to fixing. Cells were fixed in 4% paraformaldehyde (4% PFA) for 20 minutes at room temperature. Fixative was then removed, with fixed cells washed three times in PBS and either utilised immediately or stored in PBS at 4°C. Cells to be processed immediately were permeabilised in 0.5% (v/v) Triton-X 100 in PBS for 10 minutes at room temperature. Permeabilised cells were washed 3 times in PBS.

Primary and secondary antibody staining was performed separately with appropriate antibodies diluted to the recommended concentrations, in 2%

(v/v) FBS in PBS. Primary antibodies were then introduced onto coverslips, and incubated at room temperature for 1 hour. Primary antibodies were washed three times in PBS, before incubation with secondary antibodies at room temperature for 1 hour. A table of primary and secondary antibodies are provided in Tables 3 and 4. Cells were then washed three times in PBS. Coverslips were then mounted onto microscope slides using an appropriate anti-fade reagent. In experiments involving EGFP in addition to fluorescent antibodies, cells were mounted onto slides with ProLong Diamond anti-fade reagent with DAPI (Life Technologies). With experiments involving only fluorescent antibodies, cells were mounted with ProLong Gold antifade reagent with DAPI (Life Technologies). Mounted coverslips were cured at room temperature for 24 hours in the dark, and then stored at 4°C for long-term storage prior to visualisation. Mounted cells were then visualised using an inverted Laser Scanning Microscope (LSM) 700 with a 63x Plan-Apochromat objective (Carl Zeiss).

2.2.6. Virological techniques

2.2.6.1. LCMV infection

For LCMV-Arm propagation, BHK21 cells were seeded into T175 culture flasks and allowed to reach 50% confluency. Cells were then infected at a multiplicity of infection (MOI) of 0.001, and incubated in LCMV propagation media (DMEM; 50mM Glutamine; 10% heat inactivated FBS; 5% Tryptose phosphate; 5ml Pen/Strep) for 96 hours. After the incubation period, supernatant was aspirated and collected, centrifuged at 1000xg, 4°C for 10 minutes, aliquoted, and stored at -70°C.

2.2.6.2. Quantification of virus via plaque assay

Plaque assays were performed in duplicate in 6 well plates. 20 µl of LCMV Armstrong was diluted 10 fold into 200µl of SFM (10^{-1}) which was further serially diluted to create 10^{-2} - 10^{-6} dilutions. Virus dilutions were then used to infect BHK21 cells as in 2.2.6.1. BHK21 cells (1×10^5) BHK21 cells were

seeded onto 12 well plates and allowed to reach 90-95% confluency prior to infection. Dilutions were used to infect cells for 1 hour at 37°C. After adsorption, dilutions were removed and replaced with complete medium mixed 1:1 with 1.6% (w/v) carboxy-methyl cellulose (CMC; Sigma). The CMC overlay prevents virus from spreading throughout the well, only able to infect the direct neighbour cell.

Cells were incubated for 6 days, then fixed in 20% (v/v) formaldehyde for 20 minutes. Cells were washed thoroughly with ddH$_2$O and then stained with crystal violet solution (0.1% (w/v) crystal violet, 20% ethanol) for 5 minutes. The positively-charged chromophore within crystal violet binds the negatively charged membrane, meaning violet staining indicates living cells, with no staining an indication of cell death. Plaques of dead cells can then be counted, and the viral titre calculated via the following calculation:

$$Virus\ titre\ (PFU/ml) = \frac{Number\ of\ plaques}{Dilution\ factor\ x\ \ volume\ (ml)}$$

2.2.6.3. Purification of LCMV by iodixanol density gradient centrifugation.

A continuous 5-30%, or 20-45% iodixanol gradient was prepared at least 24 hours prior to LCMV purification. Briefly, sequential layering of 1.9ml 30%-5% iodixanol in TNE (100mM Tris-HCL pH 7.4; 100mM NaCl; 1mM EDTA) in 5% increments were loaded into a PolyClear 14 x 95mm Open-top centrifuge tube (Senton). In between layer additions, the gradient was frozen on dry-ice, with final, long-term storage at -70°C. An overnight thawing step was performed, prior to loading, to allow for layers to diffuse, creating a continuous gradient. Iodixanol was supplied as a 60% solution (OptiPrep, Sigma Aldrich).

LCMV was used to infect BHK21 cells, as in 2.2.3.5. However, after the incubation period, supernatant was collected and then centrifuged at 1000x*g*, 4°C for 10 minutes, to remove cellular contamination. Cell supernatant was then mixed at 1:4 dilution with 50% (w/v) PEG6000 in TNE. LCMV was then

precipitated overnight at 4°C with continuous mixing. Supernatant/PEG mix was then centrifuged at 3000x*g* for 40 minutes at 4°C, to recover precipitated virus. The pellet was then resuspended in 2ml TNE, and layered on-top of the continuous iodixanol gradient. The gradient was then ultracentifuged at 200,000x*g* for 90 minutes at 4°C in a SW40 swinging bucket rota (Beckman Coulter). After cycle completion, 500 μl fractions were recovered, from the top (least dense) to the bottom (densest). Fractions were then mixed, samples were taken for WB analysis and infectivity determination, and frozen at -70°C. Samples for WB analysis were inactivated at 65°C for 15 minutes.

After identifying fractions containing infectious virus, the fraction was inactivated at 65°C for 15 minutes, and sent for MS analysis. Figure 16 illustrates gthe rationale behind comparative MS analysis of LUJV-NP-EGFP vs LCMV NP interactome.
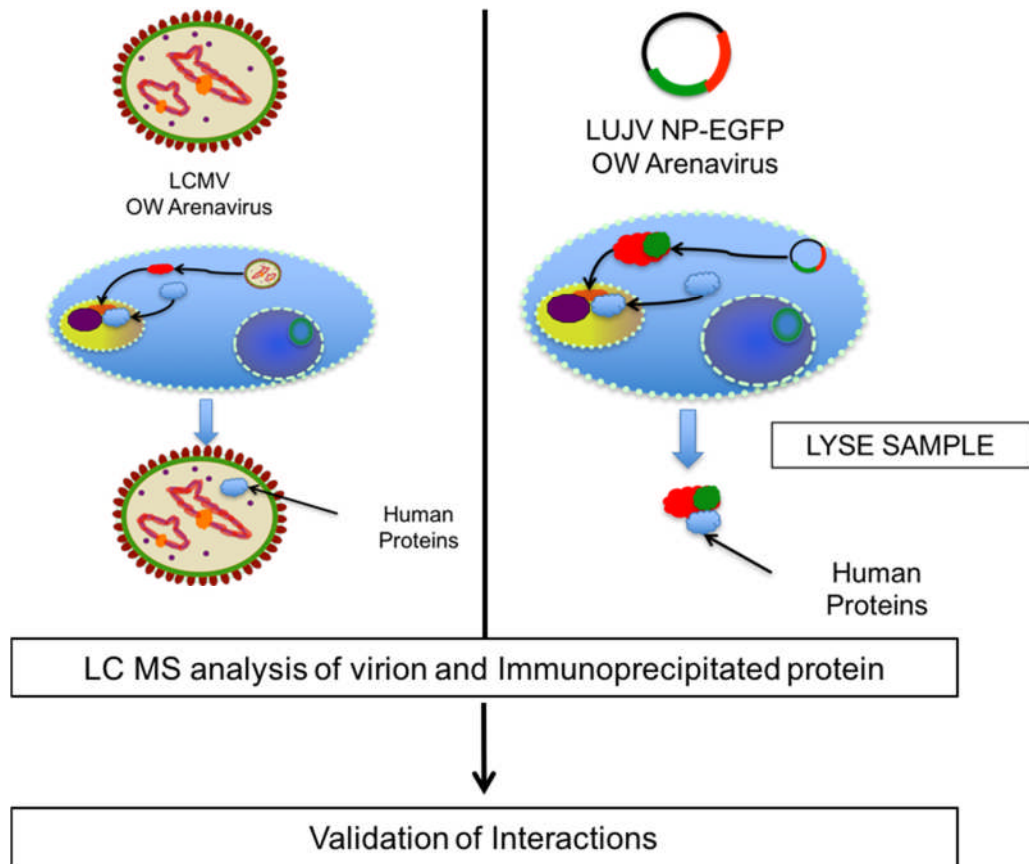
Diagram illustrating the immunoprecipitation, and inclusion within viral particles, of cellular proteins, and subsequent analysis via MS

Figure 16. **Diagram showing the comparative analysis of NP interactome and virion proteome**

2.2.7. Mass-spectrometry analysis of virus

MS analysis of the isolated fraction containing LCMV was performed in order to identify cellular proteins packaged within viral particles. MS analysis was performed by Dr Stuart Armstrong at the University of Liverpool.

2.2.7.1. NanoLC MS ESI MS/MS analysis

Protein samples generated by EGFP-trap immunoprecipitations were separated by 12% SDS-PAGE electrophoresis. The resulting separated proteins were cut from the gel in 10 slices and subjected to in-gel digestion with trypsin. Trypsin-digested peptides were separated using an Ultimate U3000 nanoflow liquid chromatography (LC) system (Dionex Corporation)

consisting of a solvent degasser, micro- and nanoflow pumps, a flow control module, a UV detector, and a thermostated autosampler. A sample volume of 10 µl (comprising 2 µg) was loaded at a constant flow rate of 20 µl/min onto a PepMap C18 trap column (0.3 mm by 5 mm; Dionex Corporation). After trap enrichment, peptides were eluted onto a PepMap C18 nanocolumn (75 µm by 15 cm; Dionex Corporation) with a linear gradient of 5 to 35% solvent B (90% acetonitrile with 0.1% formic acid) over 65 minutes at a constant flow rate of 300 nl/min. The high-pressure liquid chromatography system was coupled to an LTQ Orbitrap XL instrument (Thermo Fisher Scientific Inc.) via a nanoelectrospray ion source (Proxeon Biosystems). The spray voltage was set to 1.2 kV, and the temperature of the heated capillary was set to 200°C. Full-scan MS survey spectra (*m/z* 335 to 1,800) in profile mode were acquired in the Orbitrap instrument with a resolution of 60,000 after accumulation of 500,000 ions. The five most intense peptide ions from the preview scan in the Orbitrap instrument were fragmented by collision-induced dissociation (normalized collision energy, 35%; activation Q, 0.250; activation time, 30 ms) in the LTQ instrument after the accumulation of 10,000 ions. Maximal filling times were 1,000 ms for the full scans and 150 ms for the MS/MS scans. Precursor ion charge state screening was enabled, and all unassigned charge states as well as singly charged species were rejected. The dynamic exclusion list was restricted to a maximum of 500 entries with a maximum retention period of 90 seconds and a relative mass window of 10 parts per million (ppm). The lock mass option was enabled for survey scans to improve mass accuracy. The data were acquired using Xcalibur software.

2.3. Role of Nucleoprotein in translation

2.3.1. Production of Arenavirus-like mRNA via *in vitro* transcription

mRNAs were synthesised from plasmid DNA by T7 promoter driven *in vitro* transcription through the use of an mMESSAGE mMACHINE® T7 Ultra Kit (Ambion) as per the manufacturer's recommendation. NPEG and 3'ΔNPEG plasmid DNA was linearised by *SfiI*. DNA was then transcribed by T7 RNA

polymerase for 1 hour at 37°C, prior to incubation with TURBO DNase for 15 minutes at 37°C to remove DNA templates from the synthesised RNA. RNA was then purified via the use of RNeasy Mini Kit (QIAGEN) and eluted into RNase free H_2O. RNA was subsequently quantified using a NanoDrop 1000 (Thermo Scientific) by spectrophotometry.

2.3.2. Transfection

RNA transfections were performed using lipofectamine 2000 (Invitrogen) reagent as per the manufacturer's recommendation. Briefly, 500ng RNA/ well (12-well plate) was incubated with appropriate volumes of lipofectamine 2000 and SF DMEM. Transfection mix was then added to A549 cells.

2.3.3. Luciferase assay

A 12-well plate was seeded with $1x10^5$ A549 cells, whereupon at 80% confluency they were infected/mock infected with LCMV-Arm at an MOI of 0.01, or MOI 1 for HRSV and subsequently incubated for 24 hours. A 0-hour time point was then collected prior to the differential transfection/mock transfection with appropriate mRNA, as shown in figure 17. Supernatants were then collected every two hours for the duration of the study, with media replaced with complete DMEM after each sample was collected.

The secreted luciferase was then detected using the Pierce *Gaussia* Luciferase Glow Assay Kit (Thermoscientific) in which a solution of coelenterazine in *Gaussia* Glow Assay Buffer is added to 20μl of each sample in a 96-well plate. Samples were then incubated for 10 minutes at room temperature to allow for signal stabilisation, with relative luciferase units (RLU) recorded accordingly, detected using a luminometre with RLU signal corresponding to the concentration of *Gaussia* luciferase in each sample.
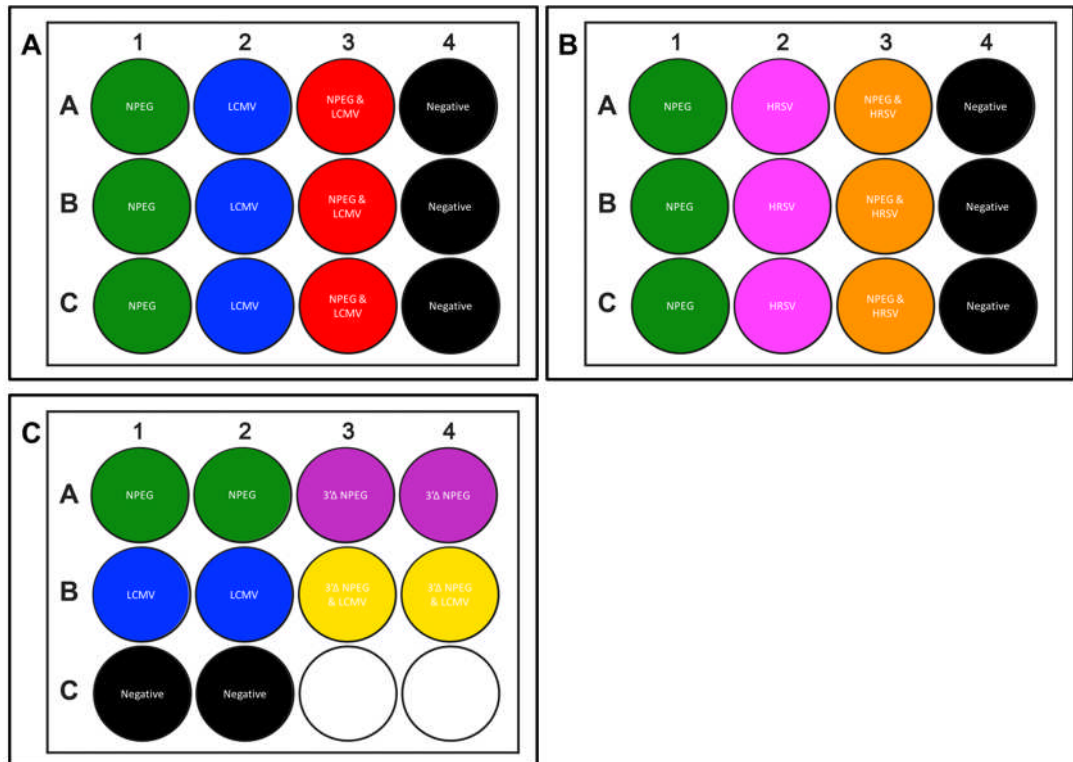
Diagram showing the layout of *Gaussia* luciferase assays performed in chapter 6. **A** shows NPEG and LCMV; **B** shows NPEG and HRSV; and **C** shows 3'Δ NPEG and LCMV.

Figure 17.   **Diagram depicting procedure of arenaviral-like mRNA translation assay**

2.3.4. List of primary antibodies

Target	Supplier	Catalogue Number	WB Dilution	IF Dilution
EGFP	Santa Cruz	sc-8334	1:2000	N/A
Actin	Sigma Aldrich	sc-966	1:2000	N/A
GAPDH	Abcam	ab8245	1:10000	N/A
LUJV NP	In house	N/A	1:5000	1:5000
LUJV NP - For use with LCMV NP	in House	N/A	1:2500	1:1000
HSP/HSC70	Abcam	ab2787	1:1000	1:100
eIF1A	Abcam	ab172623	1:1000	1:100
eIF2S1	Abcam	ab26197	1:1000	1:100
eIF3E	Abcam	ab36766	1:1000	1:200
eIF4E-Phospo (S209)	Abcam	ab131513	N/A	1:150
eIF4E	Abcam	ab1126	1:1000	1:100
eIF4A1	Abcam	ab31217	1:1000	1:100
eIF4G1	Abcam	ab2609	1:2000	1:500
eEF1A	Abcam	ab37969	1:1000	N/A
PABP	Abcam	ab21060	1:500	1:200
RPS11	Abcam	ab175213	1:2000	1:150
RPL10A	Abcam	ab174318	1:2000	1:200
RPL26L1	Abcam	ab181110	1:2000	1:150
RPS19	Abcam	ab57643	1:2000	N/A
PICV NP	in house	N/A	1:2000	N/A
LCMV NP	Abcam	ab31774	N/A	1:500

Table 3 Table of primary antibodies

73

2.3.5. Table of Secondary Antibodies

Target- primary antibody species	Conjugation	Vendor	Catalogue number	WB dilution	IF dilution
Mouse	Alexa Fluor 488	Life Technologies	A11001	N/A	1:500
Rabbit	Alexa Fluor 594	Life Technologies	A11061	N/A	1:500
Mouse	Alexa Fluor 594	Life Technologies	A21201	N/A	1:500
Sheep	Alexa Fluor 647	Life Technologies	A21448	N/A	1:500
Mouse	IRDye 680RD	LiCor	926-68072	1:10,000	N/A
Rabbit	IRDye 680RD	LiCor	926-32213	1:10,000	N/A
Sheep	IRDye 800CW	LiCor	926-32214	1:10,000	N/A
Rabbit	IRDye 800CW	LiCor	926-32213	1:10,000	N/A

Table 4 Table of secondary antibodies

3. Expression and purification of LUJV NP and Pichindé virus NP in BL21 Rosetta 2

3.1. Introduction

Lujo virus is a newly emerged haemorrhagic fever-causing infectious agent, capable of significant morbidity and high mortality rates in humans (Sewlall et al., 2014). In order to study the molecular and cellular biology of this pathogen, we sought to develop immunological tools that would allow identification of native LUJV NP in order to better understand its role in the virus lifecycle. The provision of LUJV NP antibodies would allow us to perform several important and revealing experimental procedures that would otherwise be problematic. These include determination of NP localization within infected or transfected cells, immunoprecipitation of NP from cell lysates, and identification of NP within cell lysates by way of western blotting.

An alternative approach to detecting a viral structural protein such as LUJV NP is to attach its ORF to that of EGFP or one of its derivatives, as described in chapter 4. In this way the LUJV moiety is detected by virtue of its linkage to the fluorescent EGFP molecule, for which many high affinity antibodies are already available should immunological detection be necessary. However, the use of EGFP, and other fusion proteins poses several issues.

Negative sense RNA viruses, such as those classified in the *Arenaviridae* family, encode relatively few proteins compared to positive sense RNA viruses such as HCV (Moradpour & Penin, 2013), or DNA viruses such as those in the *Herpesviridae* family (Jackson et al., 2011). Negative sense virus proteins, therefore, often perform multiple roles in the viral life cycle (Carter et al., 2012; Fehling et al., 2012; Hale et al., 2008). The arenavirus NP is known to be important in interferon modulation, transcriptional anti-termination, factor recruitment alongside encapsidating viral genomes (Pythoud et al., 2012; Tortorici et al., 2001; Baird et al., 2012; Pedersen & Konigshofer, 1976). The introduction of a tag – such as EGFP – is likely to either disrupt protein

structure, or interfere with assembly of multimeric complexes and therefore impinge the protein function *in vivo*. Therefore, using a tagged protein for research, while offering benefits in detection and purification, will likely limit the conclusions that can be made due to possible alteration or disruption of one or several functions of the native protein.

The tagging of proteins remains a useful method for tracking proteins and identifying interacting partners, and this technique is exploited to initially identify interacting partners of LUJV NP in chapter 4. However, the most authentic and functionally accurate method for determining a protein's location *in vivo* remains the use of un-tagged native protein. However, there is a lack of effective and commercially available antisera against NPs readily available for members of the *Arenaviridae* family. This chapter describes the expression and purification of LUJV NP and Pichindé virus (PICV) NP in *Escherichia coli*. PICV is a NW arenavirus, classified as a hazard group 2 organism which enables its use in BSL2 containment facilities. As it is asymptomatic in humans, it could act as a good surrogate species for the NW group pathogens in order to compare observations between NW and OW arenavirus. The generation polyclonal antisera in this manner should provide a useful tool for future research applications regarding the NW members of the *Arenaviridae* family.

Diagram illustrating the fusion protein comprising his-SUMO-LUJV NP. The 6xHis tag is illustrated by: **H**. The unboxed region between SUMO and NP ORFS represents the cleavage site of SUMO protease, after the C terminal of the SUMO ORF.

Figure 18.   Representation of SUMO- fusion proteins

After the purification of the NPs validation of the subsequently generated sheep polyclonal antibodies was also required. The strategy that was adopted was to express LUJV NP as a fusion protein, with the LUJV NP ORF linked to the ORF for his-SUMO within a pET28a expression vector. This attachment allows for SUMO-protease structure based recognition and cleavage, allowing

the efficient removal of the SUMO and 6xHis affinity tags from native LUJV NP (Figure 18) (Panavas et al., 2009).

In the pET system, expression of the protein of interest is driven by bacteriophage T7 RNA polymerase, which binds to the specific promoter region upstream of the target ORF for expression. pET systems can therefore be introduced into various strains of *E. coli* BL21 cells, which express T7 RNA polymerase. Inducing the expression of target protein is linked to the *lac* operator upstream of the target protein, with IPTG binding to the lac repressor, catalysing its removal from the operator, allowing for the synthesis of mRNA from lac operator linked genes. The auto-induction procedure utilised is based on providing growth cultures with low levels of glucose, enough to allow cells to reach saturation (Studier, 2005). Alongside this glucose is lactose, the metabolism of which is inhibited in the presence of glucose. The use of glucose by the growing culture then allows the bacterial culture to reach high concentrations, where the metabolism of lactose begins, inducing the expression of lac operon linked proteins. Finally, glycerol can be utilised by the saturated culture as a late stage energy source to prevent bacterial death from resulting in a loss of protein (Studier, 2005).

Purification of viral nucleocapsid proteins in *E. coli* from plasmid cDNAs is well documented, but it is has become apparent that each protein may behave differently during expression and purification. For example, despite sharing 60% amino acid sequence homology and striking structural similarity, the N proteins of HAZV and CCHFV require entirely different strategies for their effective purification (Surtees, 2014).

The establishment of a successful purification procedure required iterative testing of multiple different growth conditions, buffer preparations and the incorporation of various chromatography techniques in order to generate the required NP purity, whilst retaining a usable yield. The trialling of multiple different conditions allowed for incremental improvements in yield and purity of LUJV NP. The final, optimised, procedure is based on two separate

purification procedures – Ni affinity chromatography (IMAC) followed by size exclusion chromatography (SEC).

Ni affinity chromatography utilises polyhistidine tagged proteins – in this case his-SUMO – binding to resin embedded with immobilised metal ions, such as Ni^{2+}. Histidine, through its side chain, binds to these immobilised ions due to electron donor groups within the imidazole ring. The use of poly-histidine allows for the protein to effectively bind the matrix, through interaction with the Ni^{2+} ions. Sequential washing with buffers containing imidazole creates competition for binding sites with weaker binding partners, removing them from the resin matrix and facilitating the purification of target protein. This method does allow the potential for contaminant proteins to be eluted with the sample, as they may simply bind the matrix with equal efficiency as the poly-histidine tagged protein. Alternatively, the target protein may interact and bind with bacterial proteins, co-precipitating them when eluted from the matrix (Bornhorst & Falke, 2000).

In order to minimize this, his-SUMO cleavage acts as a further purification step. The cleavage of SUMO tag, comprising the 6xhis-SUMO portion of the fusion protein (figure 18), from the target protein generates two separate proteins – in the case of LUJV NP, one ~63kDa protein comprising LUJV NP, and the his-SUMO tag of ~15kDa are generated, in addition to the ~25kDa SUMO-protease. The SUMO tag is cleaved through a tertiary structure-driven recognition site, rather than a sequence-based cleavage site, which allows for increased specificity of the protease (Panavas et al., 2009). This allows the various components to be separated from each other by size exclusion chromatography on account of their relative size, increasing sample purity. As the specificity of the SUMO-protease is for the tertiary protein structure, any improperly folded proteins will not be cleaved, potentially allowing for un-cleaved proteins to be eluted alongside cleaved fractions after size exclusion (Panavas et al., 2009). By passing the SUMO-cleaved protein mixture through a naïve nickel resin column, unwanted his-tag containing contaminants can be removed from the sample (Panavas et al., 2009).

Size exclusion chromatography (SEC), unlike Ni affinity, relies on the target proteins not binding with the resin matrix. SEC utilises a bead matrix with varying microscopic pores present within beads. This physical structure of the matrix allows for proteins of different size to flow through the resin matrix at different rates (Wang et al., 2010). Depending on the chosen column type, this allows multimeric and large proteins to flow through without restriction, whilst smaller proteins must flow through more and more pores, slowing their progress. This procedure, alongside that of Ni affinity chromatography, has the potential to generate protein fractions of high purity.

The generation of antisera was performed in sheep, which provide a large yield of specific polyclonal antibody, enabling great flexibility in future applications. The procedure requires the use of purified protein in order to generate specific and reactive antibodies without undesired cross-reactivity with other host proteins. The use of an initial priming challenge, followed by boosters comprising secondary and tertiary immunisations allows for the generation of large volumes of antibody, dependent on the organism used in the immune-challenge procedure.

Finally, this chapter shows the application of the established purification procedure in order to purify the NW model arenavirus PICV NP. PICV is a well-established infectious model for several of the pathogenic NW arenaviral infections. Much like LCMV, it does not cause infection in immunocompetent adults, but serves as an effective model for arenaviral infection in animals and cells (Kumar et al., 2012; Zhang et al., 2001).

This chapter demonstrates a method for the purification of two arenaviral nucleoproteins – LUJV NP and PICV NP – and the subsequent specificity of LUJV NP antisera reactivity, and its cross reactivity with another arenaviral nucleoprotein. These antisera could prove to be an effective research tool regarding the study of arenaviral NPs and their intracellular functions.

3.2. <u>Results</u>

3.2.1. pET28a expression plasmid

A parental pET28a expression vector was digested and ligated with an LUJV NP insert, in order to generate a his-SUMO-LUJV NP fusion protein. *E. coli* DH5α were subsequently transformed with the pET28a his-SUMO-LUJV NP plasmid, and grown on selective media in order to isolate successfully transformed colonies. Plasmid DNA was extracted, purified and the construct confirmed as correct via sequencing.

3.2.2. Identification of effective bacterial expression vector

In order to identify the most appropriate strain of *E. coli* BL21 for the expression of LUJV NP, the expression properties of a panel of different *E. coli* BL21 strains were tested. Bacterial strains were transformed with his-SUMO-LUJV NP and single colonies individually selected for culturing. Subsequently, these cultures were grown in volumes of 300ml until reaching an OD_{600} of 0.6 and induced with IPTG. After growth, cells were pelleted, lysed and purified via IMAC to harvest soluble protein. The comparative initial purification profiles of both BL21 Rosetta 2 (R2) and BL21 Codon+ (C+) strains is shown in figure 19, with the position of his-SUMO-LUJV NP marked. The overall expression of LUJV NP was higher in R2, relative to C+, while also exhibiting a more effective purification profile. The R2 strain was thus chosen as the expression strain for LUJV NP expression and purification.
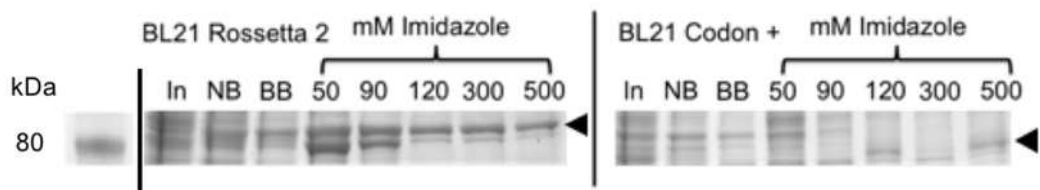
Coomassie stain of SDS PAGE gel showing the purification profile of his-SUMO-LUJV NP in two *E. coli* strains – BL21 Rosetta 2 and BL21 Codon+ via IMAC. The image shows a slice corresponding to a kDa range of approximately 90-50 kDa, with his-SUMO-LUJV NP evident in the centre of the sliced fractions at approximately 80 kDa. 'In' correlates to the input fractions; 'NB' to the non-bound fraction; 'BB' to the initial binding buffer wash.

Figure 19.   Comparative expression of his-SUMO-LUJV in *E. coli* BL21 Rosetta 2 and Codon+ strains

3.2.3. Growth Media

3.2.3.1. IPTG
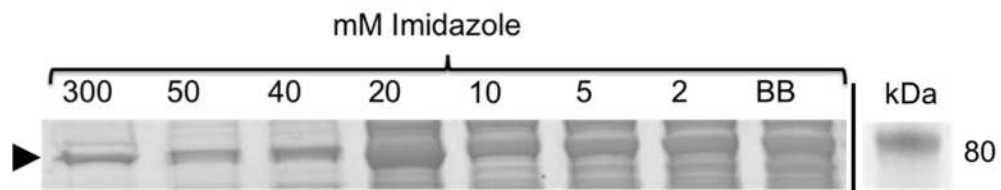
Coomassie stain of SDS PAGE gel showing the purification profile following IPTG induction of his-SUMO-LUJV NP via IMAC, with each imidazole concentration corresponding to the elution fraction. The image shows a slice corresponding to a kDa range of approximately 90-50 kDa, with his-SUMO-LUJV NP evident in the centre of the sliced fractions at approximately 80 kDa. "BB' indicates the initial binding buffer wash, with numerals indicating increasing concentration (mM) of imidazole in washes.

Figure 20.   Expression of SUMO-LUJV NP using IPTG induction

After establishing the most effective growth strain (R2), it was necessary to trial batch growth media and growth in a larger volume to enable optimal expression and stability of the required volume of protein. Cells were transformed, incubated, and subsequently grown in LB batch cultures of 6L, as routinely used in the laboratory. R2 cultures expressing LUJV NP were grown at 37°C until they reached OD_{600} of 0.6, at which point they were induced with 100μM IPTG and grown for 16 hours at 18°C. Cells were then pelleted and lysed as described. Initial expression attempts generated good yields of soluble LUJV NP, although following stringent washing a significant amount of protein was lost, as seen the protein eluted using relatively low

concentrations of imidazole, between 2 and 20mM (figure 20). These low concentration imidazole washes were performed to gradually remove protein contaminants, with the hope of minimizing the loss of his-SUMO-LUJV NP. However, experimental conditions that allowed the generation of a high yield of highly pure protein were not identified, and so alternative expression and purification strategies were investigated.

3.2.3.2. Auto-induction

Coomassie stain of SDS PAGE gel showing the purification profile of his-SUMO-LUJV NP via IMAC after induction via auto-induction based growth conditions. Each imidazole concentration (mM) corresponding to the elution fraction. The image shows a slice corresponding to a kDa range of approximately 90-50 kDa, with his-SUMO-LUJV NP evident in the centre of the sliced fractions at approximately 80 kDa. 'BB' indicates the initial binding buffer wash, with numerals indicating increasing concentration of imidazole. 2M* indicates a 2M NaCl wash.

Figure 21.   **Expression of SUMO-LUJV NP using auto-induction**

Expression of LUJV NP using the IPTG induction protocol resulted in a significant proportion of the expressed protein being unable to bind to the resin – suggestive of poor solubility or improper folding. In order to remedy this, an auto-induction protocol was adopted, involving 56 hours of bacterial growth at a set temperature after inoculation; in this case – after much optimisation – 18°C. In switching from IPTG-based induction to auto-induction based expression, it was hoped that proteins would be expressed at a slower rate, which may facilitate correct folding. 6L of auto induction media was inoculated and, after 56 hours, cells were pelleted. After changing growth conditions to auto-induction, and minor alterations to the imidazole washing stages, the previous observations regarding poor yield due to losses during washing were minimised. The addition of a 2M NaCl wash was adopted to remove any RNA bound to NP, which could have influenced purification by promoting the formation of high order RNA bound NP multimers. Taken together, these

protocol modifications resulted in improvements in the overall yield and purity of the resultant his-SUMO-LUJV NP (Figure 21).

3.2.4. Purification

After establishing the basic procedures for bacterial cell growth and identifying effective initial washing steps, it was necessary to improve the yield and enhance purity of the expressed his-SUMO-LUJV NP such that it would be suitable for further purification by size exclusion chromatography, and ultimately generate a soluble folded protein suitable for antibody production.

3.2.4.1. Purification conditions

The initial NP expression conditions were based on those used by Tanner *et al.* 2014 and Ariza *et al.* 2013 as described in 2.2.2.3. These experimental conditions were refined throughout the study, with changes implemented where required, as indicated in 2.2.2.3. Lysis was performed as described in 2.2.2.2 throughout. In order to effectively purify LUJV NP from its his-SUMO-fusion, it was necessary to establish the isoelectric point (pI) of both his-SUMO-LUJV NP and LUJV NP. Proteins held at pH conditions near their pI values are susceptible to precipitation, due to the lack of charge to aid with solubility. The predicted pI of LUJV NP was found to be 8.72 – using the ExPASY Compute pI/Mw software – with the predicted pI of the his-SUMO-LUJV NP fusion protein found to be 8.58. On the basis of these findings, the use of buffers at a pH of 7.5 was chosen.

3.2.4.2. Ni affinity purification of his-SUMO-LUJV NP

In the previous sections, a bacterial growth strain, growth media and induction procedure were selected for his-SUMO-LUJV NP expression. Enhancement of the IMAC purification procedure was then necessary to allow for greater final yield of purified LUJV NP. In order to achieve this, changes to the washing conditions were made, with differing concentrations of imidazole used in washing buffer preparations. In addition, larger column volumes of washing preparations were trialled in order to remove a higher proportion of contaminants. Proteins resolving via SDS PAGE electrophoresis to a similar

size as the eluted his-SUMO-LUJV NP band were eluted from 80mM imidazole fractions, as seen in figures 21 and 22. In order to try and minimise this loss, an increase in the number of washes of the lower imidazole concentration steps was trialled. Figure 22, compared to figure 21, shows increased contaminant protein levels eluted at lower imidazole concentrations, with minimal loss of the target protein at these stages also.
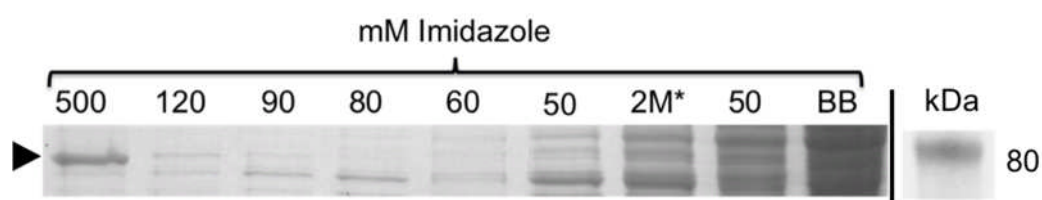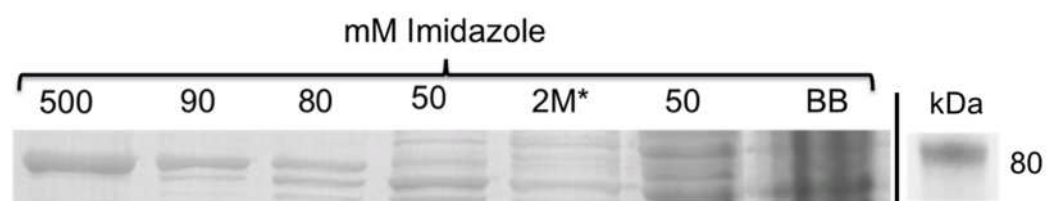

Coomassie stain of SDS PAGE gel showing the purification profile of his-SUMO-LUJV NP via IMAC, with each imidazole concentration corresponding to the elution fraction. The image shows a slice corresponding to a kDa range of approximately 90-50 KDa, with his-SUMO-LUJV NP evident in the centre of the sliced fractions at approximately 80 KDa. 'BB' indicates the initial binding buffer wash, with numerals indicating increasing concentration of imidazole. 2M* indicates a 2M NaCl wash.

Figure 22.   Purification of SUMO-LUJV NP via IMAC

Figure 23 shows the optimised IMAC purification profile of his-SUMO LUJV NP extracted from R2 cells – method as in 2.2.2.3 – including cleavage and dialysis, showing minimal loss of protein via dialysis into gel filtration buffer (GFB) and efficient cleavage of LUJV NP. The incremental improvements from initial attempts allowed for the purification of ~25mg of total protein prior to dialysis, after which this was subsequently concentrated prior to loading onto a SEC column. The shift in size from ~80kDa to ~63kDa after cleavage with SUMO protease is consistent with the eluted protein being the expected his-SUMO-LUJV NP, as the specificity of SUMO protease for the tertiary structure of the SUMO tag is highly specific and should not cleave any contaminating bacterial protein.
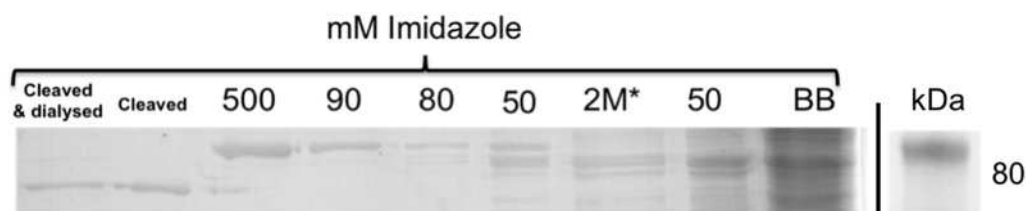
Coomassie stain of SDS PAGE gel showing the purification profile of his-SUMO-LUJV NP via IMAC, with each imidazole concentration corresponding to the elution of that fraction. The image shows a slice corresponding to a kDa range of approximately 90-50 kDa, with his-SUMO-LUJV NP evident in the centre of the sliced fractions at approximately 80 kDa. 'BB' indicates the initial binding buffer wash, with numerals indicating increasing concentration of imidazole. 2M* indicates a 2M NaCl wash. The 'cleaved' fraction shows the eluted protein fraction from '500' treated with SUMO protease.

Figure 23. **Purification of SUMO-LUJV NP via IMAC and cleavage by SUMO protease**

3.2.4.3. Purification of LUJV NP via size exclusion chromatography

After IMAC purification, SEC was chosen to separate cleaved LUJV NP from contaminant proteins from the IMAC procedure, including the SUMO protease and the his-SUMO tag. Dialysis of LUJV NP into GFB and subsequent concentration to a volume of 5ml from 20ml resulted in a protein loss, such that the total quantity of purified protein dropped from ~25mg to ~15mg prior to SEC loading, as determined via Bradford assay. The purification parameters are outlined in 2.2.2.4. Briefly, LUJV NP was loaded into a 5ml injection loop prior to loading onto an S75 column. The gel filtration system was allowed to run until completion at a flow rate of 0.3ml/min, with 3.3ml fractions collected and a 280nm trace utilised to identify fractions containing protein, as shown in figure 24. The 280nm trace identified a region of protein concentration from fractions 35-45, which correspond to the first two peaks visible. These fractions were collected and SDS PAGE electrophoresis performed in order to identify the protein present. Coomassie staining allowed for the identification of a band of ~ 63kDa – as expected for LUJV NP – in 10 of the 11 fractions. The final peak was identified as containing low MW bands, corresponding to the 25kDa SUMO protease and 15kD a his-SUMO tag.
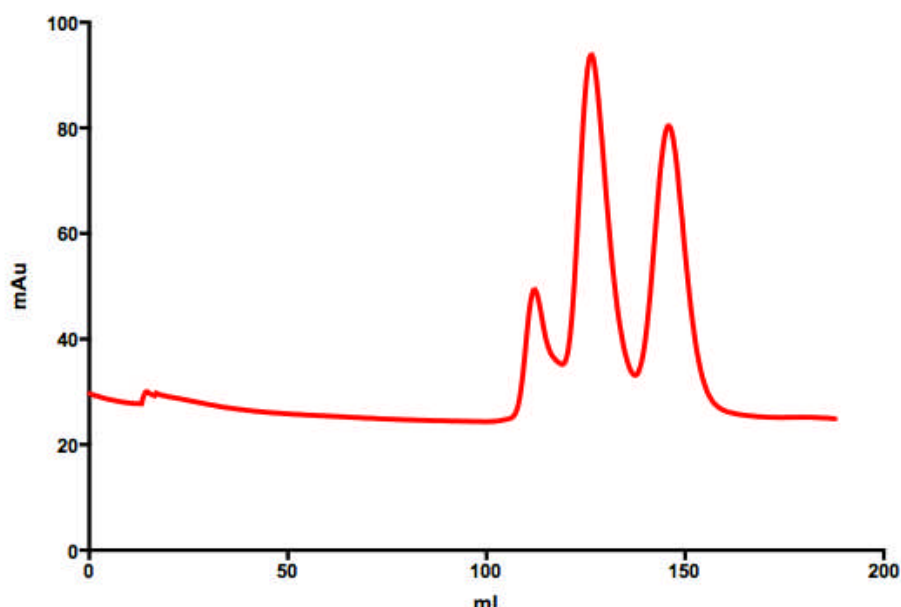
SEC 280nm mAu trace showing LUJV NP detected by absorption at 260mn. Elutions were collected in 3.3ml fractions.

Figure 24.   SEC purification profile of LUJV NP after detection of 280nm absorbance

Within fractions 36-39, a minor contaminant was detected with a molecular weight of ~45kDa. As it was not feasible to identify the protein via WB due to the lack of antisera, it was decided to only collect and pool fractions 40-44 that did not contain this contaminant for immunochallenge in sheep for polyclonal antibody generation. The total mass of purified protein was 1.5mg in a total volume of 1.5ml for sheep immunisation.
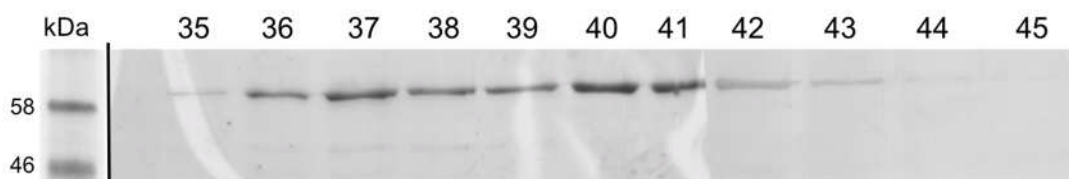
Coomassie stain of 12% SDS PAGE gel, showing fractions collected from SEC column containing LUJV NP, corresponding to the first two eluted peaks shown in figure 24.

Figure 25.   SEC purification profile of LUJV NP after SDS gel electrophoresis

3.2.5. Purification of PICV NP via IMAC and SEC via a his-SUMO PICV NP intermediary

After establishing the LUJV NP purification procedure described above, the purification of PICV NP was attempted utilising the same established method.

Protein expression was achieved using plasmid pET28a his-SUMO-PICV, which expressed the PICV NP appended to SUMO and His tags, in a similar design as described for LUJV NP expression (figure 23). Transformed *E. coli* BL21 R2 cells were grown in 6L of auto-induction media for 56 hours at 18°C, then pelleted and lysed, with purification following the optimised method established for LUJV NP purification. The expression of his-SUMO-PICV NP proved successful, with IMAC purification yielding approximately 50mg of un-cleaved total protein after elution. Subsequent cleavage and SEC purification yielded several fractions containing PICV NP, which were analysed by SDS-PAGE, as shown in figure 26.
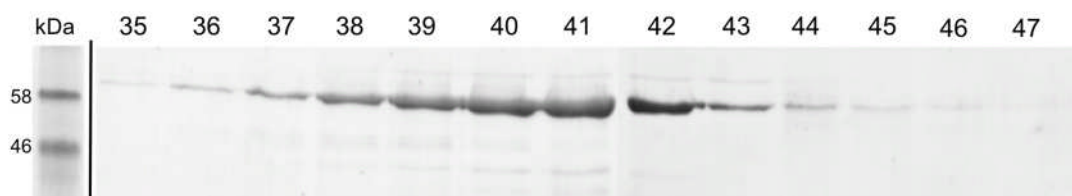
Coomassie stain of 12% SDS PAGE gel, showing fractions collected from SEC column containing PICV NP – as shown, and un-cleaved his-SUMO-PICV NP – "sPICV".

Figure 26.   SEC purification profile of PICV NP after SDS electrophoresis

Fractions 35 through 46 exhibited a band at ~62kDa corresponding to PICV NP, along with a band at ~78kDa in fractions 39-43, presumably corresponding to un-cleaved SUMO-PICV NP. In order to remove this contaminating band, collected fractions were exposed to a secondary IMAC purification step.

3.2.5.1. Secondary IMAC purification
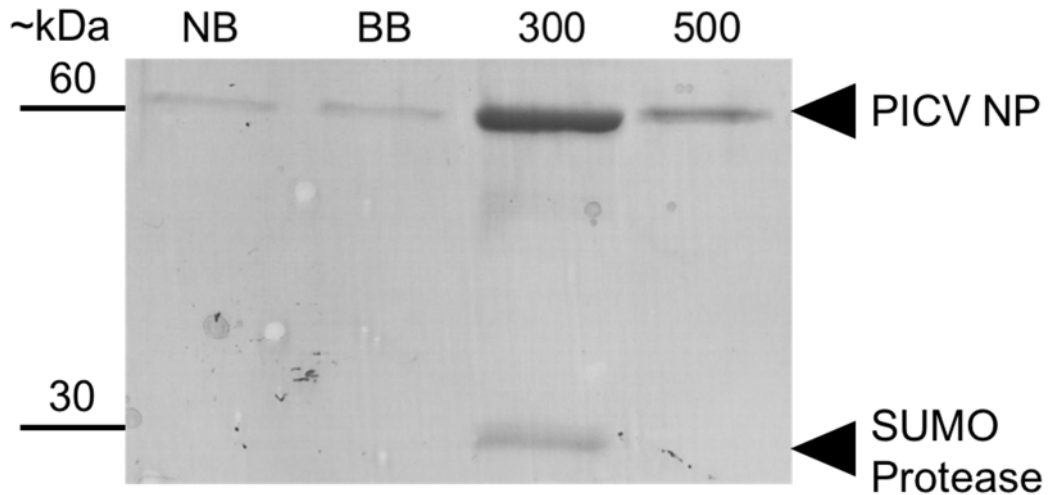
Coomassie stain of 12% SDS PAGE gel, showing fractions collected from SEC column containing PICV NP – as shown, and un-cleaved his-SUMO-PICV NP – "sPICV". ~kDa corresponds to the approximate kDa of resolved proteins; NB the non-bound fraction; BB the binding buffer wash; 300 a 300mM imidazole elution; and 500 corresponding to a second 500mM elution

Figure 27. **Secondary IMAC purification of PICV NP**

In order to complete the purification procedure, it was necessary to remove contaminating, un-cleaved his-SUMO-PICV NP from collected SEC fractions, achieved by a second IMAC step. Prior to loading the protein mixture onto a His-Trap IMAC column, SUMO protease was added to the fractions in order to cleave any remaining un-cleaved fusion proteins. After cleavage, fractions were loaded onto the IMAC resin. Figure 27 shows the fractions collected, with the majority of PICV NP eluted from the resin at 300mM imidazole, along with a contaminating band at approximately 25kDa, corresponding to the SUMO protease. The two eluted fractions were combined, and concentrated to 5ml, prior to a second SEC procedure.

3.2.5.2. Secondary Size exclusion chromatography

In the previous section, PICV NP was purified via a secondary IMAC purification procedure, resulting in a single fraction containing PICV NP alongside a contaminant corresponding with the SUMO protease at ~25kDa. In order to remove lower molecular weight contaminants, such as SUMO

protease, from the IMAC purification fraction it was necessary to perform SEC. A 5ml volume of PICV NP in GFB was loaded onto an S75 column, with fractions collected throughout. Figure 28 shows the collected fractions 41-47, with PICV NP visible at approximately 62kDa. There was no observable un-cleaved his-SUMO-PICV NP evident in collected fractions.
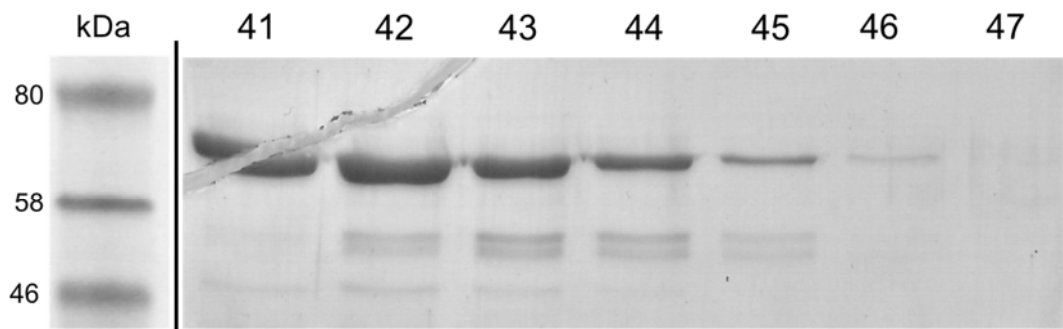
Coomassie stain of 12% SDS PAGE gel, showing fractions collected from SEC column containing PICV NP.

Figure 28.   **Secondary SEC pufrification of PICV NP**

However, the presence of some lower molecular weight proteins is indicative of protein degradation, due to their absence from the previous IMAC purification step. Had analysis of the IMAC procedure shown in figure 28 indicated a similar profile of additional protein species, then there would be cause to further purify the samples. However, as this was clear and only showed PICV NP and a single, approximately 25kDa species corresponding to SUMO Protease, the most likely cause for additional protein bands appearing after SEC is through protein degradation.

3.2.6. Immune challenge

After successfully purifying LUJV NP and PICV NP from R2 cells, 1.5mg of both proteins were sent to Alta Bisocience for the generation of polyclonal antibodies by immunisation of sheep. The immunisation programme was for an initial injection followed by boosting immunizations, which generated 3 bleeds, along with pre-immune sera for base-line comparison.

3.2.6.1. Validation of LUJV NP antibody specificity

In order to determine the specificity of the polyclonal LUJV NP antibody, it was necessary to confirm that the agent could detect, and was specific to the provided immunogenic sample. This was done in two ways: first, the antibody was used at a dilution of 1:2500 to detect the LUJV NP moiety of both his-SUMO-LUJV NP and the post cleavage LUJV-NP product (figure 29**B**), expressed and purified as described above (figure 23). The recognition of a single band corresponding to both the un-cleaved and cleaved, native LUJV NP showed the antibody was specific for the LUJV NP target. In addition, the ability of this antibody to detect these targets at a relatively high dilution suggested the titre of specific antibody was high.

Second, the LUJV NP antisera was used in immunofluorescence microscopy, using cells expressing arenavirus NP either by transient transfection with EGFP/LUJV NP-EGFP or infected/mock-infected with LCMV-Arm. It was expected that the antibody would cross-react with LCMV NP due to the high sequence similarity between LUJV NP and LCMV NP.
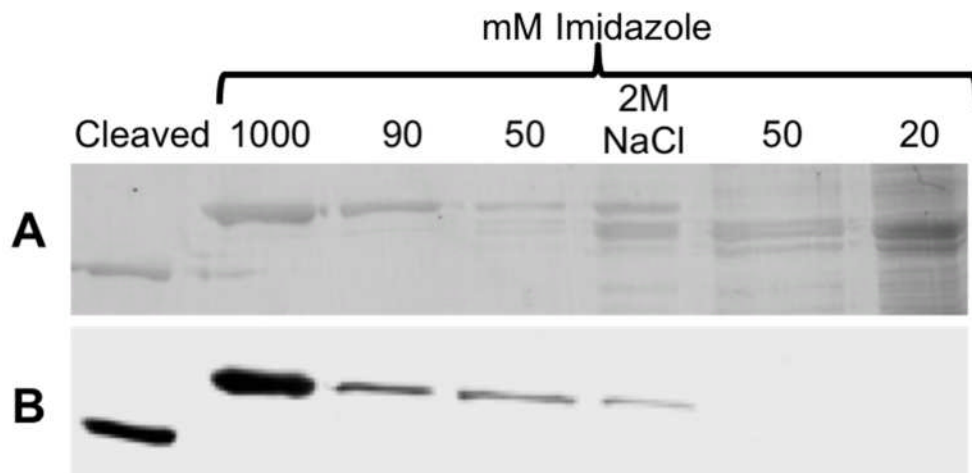
Comparison between a Coomassie stain of a 12% SDS PAGE gel in **A** and WB, shown in **B**. Both gels show the purification profile of LUJV NP via IMAC, with **B** demonstrating a WB staining by LUJV NP antisera at a dilution of 1:2500

Figure 29.   **Validation of anti-LUJV NP abtibody via western blot**

Cells were incubated for 36 hours after transfection/infection, fixed and stained using the anti-LUJV NP antisera at a dilution of 1:2000, and, where

required, anti-LCMV antibody. Cells were subsequently stained with appropriate secondary antisera, prior to mounting on microscope slides with DAPI-stain. Figure 30**A** shows LUJV NP staining exactly mimicking that of LUJV NP-EGFP, without any cross-reactivity with host cell proteins or EGFP in EGFP transfected cells. 30**B** also shows the staining of LCMV-Arm NP by LUJV NP antisera. In addition, the generated antibodies also show a good degree of reactivity against LCMV NP. This reactivity at low concentrations indicates that the immunological challenge was successful, and generated a specific antibody that detected two different arenaviral NPs under analysis.
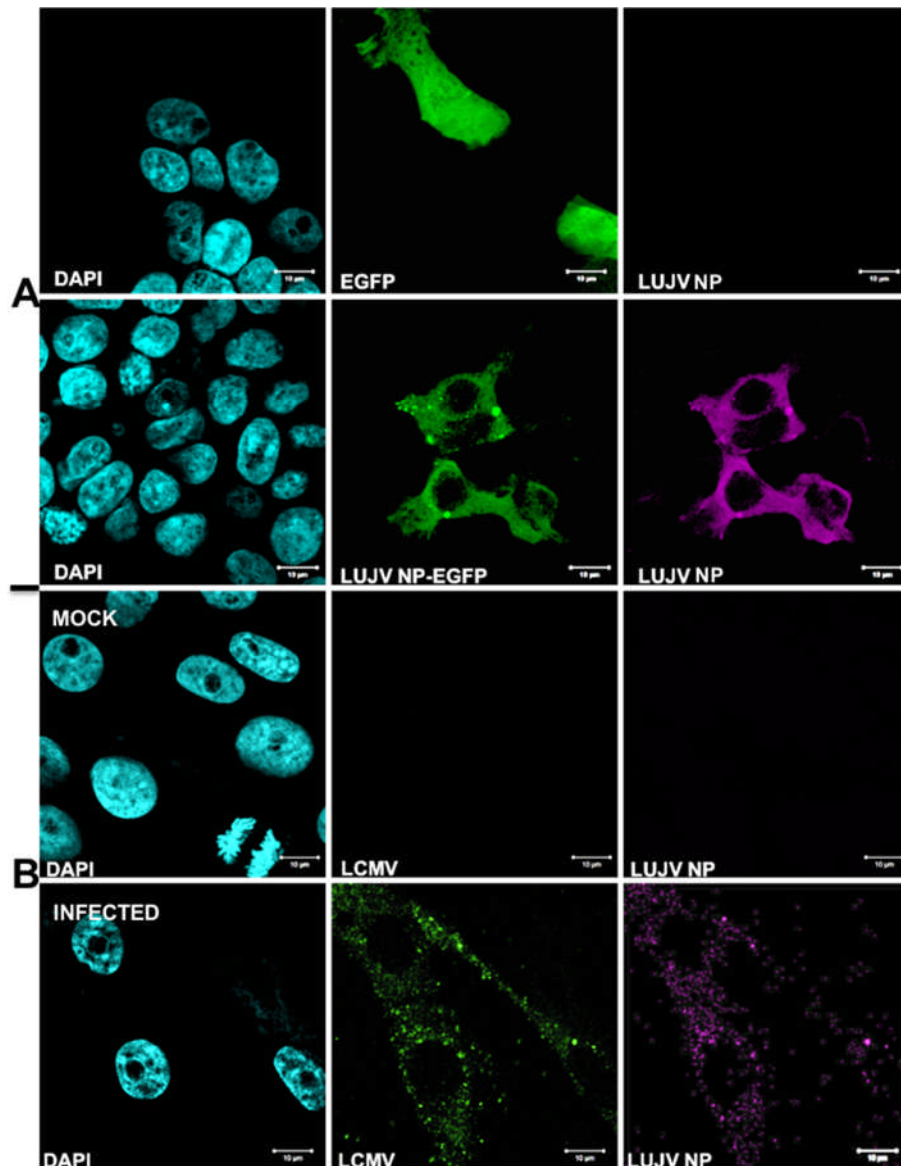
Immunofluorescence analysis of the specificity of LUJV NP antisera against LUJV NP-EGFP in **A;** and LCMV-Arm NP in **B**, against and EGFP and mock-infected control respectively. In both **A** and **B**, Vero cells were used. Cyan indicates DAPI staining of DNA; green indicates EGFP, LUJV-NP EGFP and LCMV NP – as stated; magenta indicates anti-LUJV NP staining. Scale bar indicates 10μm in all instances.

Figure 30. **Validation of anti-LUJV NP antibody reactivity against LUJV NP and LCMV NP via IF**

3.3. <u>Discussion</u>

The purification of LUJV NP and PICV NP and subsequent generation of specific antisera represents an important advance in the establishment of tools to further probe the molecular and cellular biology of arenaviruses.

These specific antisera were previously either not available, non-reactive or prohibitively expensive. As will be evident in later chapters in this thesis, the provision of these antibodies was a critical first objective that permitted the downstream analysis of native arenavirus NP in cells, and also in assembled virus particles. The NP expression and purification techniques described here could prove applicable for more arenaviral NPs, with further enhancements potentially allowing for the generation of large quantities of highly pure LUJV and PICV NP suitable for solving the high resolution crystal or solution structures.

The purification of LUJV NP initially proved challenging, with multiple incremental changes required to enable satisfactory purification and to meet the required yield. Initial usage of IPTG – as a standard induction technique – only generated a relatively poor yield of purified protein. The reasons for this are unknown, but initial problems with poor solubility, and inefficient IMAC purification as a result. Assessment of different wash buffer conditions resulted in an improvement in the overall purity of resulting eluted protein, but the overall yield of protein was unsatisfactory, with significant protein losses due to precipitation during purification. To this end, a change in approach to auto-induction allowed improvements on both yield and purity to be achieved.

The use of auto-induction medium generated culture which produced lysates with a higher proportion of soluble proteins, which were seemingly easier to separate from contaminants. While the evidence for this being solely as a result of the auto-induction media is circumstantial, the results were far more satisfactory than were observable than during IPTG based methods.

Imidazole washing steps were optimised during multiple purification attempts, with incremental improvements each time. The introduction of a 2M NaCl wash also aided in the likely removal of NP multimers bound to RNA, as observed via the 280nm absorbance sensor during SEC. This addition allowed for a greater yield of RNA-free protein. Finally, increasing the volumes of imidazole washes, combined with the optimised fractions, generated a good yield of both LUJV NP and subsequently with PICV NP. While the procedure

was successful, the switch from IPTG-based induction to auto induction contributed enormously to the advancements made. However, it is not possible to say that IPTG would not have subsequently been a more successful induction procedure, had the complete optimisation of the IMAC purification procedure been applied to cultures prepared from IPTG based induction.

SEC purification generally worked well throughout, with the only issue being a loss of overall protein from the loaded sample. However, the clean fractions observed (figures 25 & 28) illustrate that, overall, the procedure conferred the production of purified LUJV NP, and PICV NP. Trialling different flow rates to counter problems involving the Hagen-Poisuille principles of fluids flowing at different rates through closed systems to help generate high resolution clarification of proteins through the matrix. Setting the SEC flow rate higher would allow several different flow rates of the liquid within the system, lowering the accuracy with which the protein is clarified. Setting the flow rate to 0.3ml/min was chosen to balance this principle with practicalities surrounding protein stability, optimising the resolution of proteins through the bead matrix.

The generation of specific antisera against these two proteins will also allow for more in-depth analysis of LUJV NP, without the need for the expressed protein to be tagged. The reactivity of the LUJV NP antibody in immunofluorescence indicates that it is specific only for NP, with no-cross reactivity observable in mock-transfected/infected cells. LCMV cross-reactivity shown in figure 30 is greatly encouraging, and the ability to visualise LCMV NP via WB will be invaluable, given the lack of readily available appropriate antisera for this technique (Chapter 5; Figure 42). In addition, PICV antisera was validated, with the data not shown due to the lack of subsequent use of PICV and PICV NP throughout the study.

This chapter shows the generation of effective antisera against the NP of the novel OW pathogen LUJV, and its activity against LCMV-Arm NP in infected cells, conferring reactivity to both NPs in immunofluorescence and western

Chapter 3 Expression and purification of LUJV NP and PICV NP in
Escherichia coli BL21 Rosetta 2.

94

blot analyses. In addition, the purification procedure appears to be adaptable to other NP's – such as PICV NP and potentially LCMV NP, among others.

4. <u>Host interacting partners of LUJV NP-EGFP</u>

4.1. <u>Introduction</u>

LUJV is an Old World arenavirus that causes haemorrhagic shock in infected humans, in an outwardly similar manner to that of LASV or EBOV (Sewlall et al., 2014). In the single recorded LUJV outbreak in 2008, five patients were exposed – four nosocomially – with four subsequently dying due to the infection (Briese et al., 2009). The high lethality of LUJV makes it an important topic of research, especially in the context of the frequent emergence of other potentially pathogenic arenaviruses around the world. Indeed in 2015 alone there have been six newly described arenaviruses isolated from either rodents or snakes, all of which possess the potential to cause disease in humans (Hellebuyck et al., 2015; Witkowski et al., 2015; Lavergne et al., 2015; Aqrawi et al., 2015; Bisordi et al., 2015; Van Cuong et al., 2015; Gryseels et al., 2015; Mikesch et al., 2010). Arenaviruses are known to pose a significant public health threat, with LASV recently included on the World Health Organisation's Blueprint for R&D readiness (WHO, 2015). As with many recorded cases of virus emergence, the movement of humans into previously unpopulated regions exposes individuals to previously un-encountered arenaviruses. For example, this phenomenon appears to be partly responsible for the increased prevalence of GTOV disease within human populations of western Venezuela, a region that is experiencing rapid development due to the rich oil reserves (Salas et al., 1991; Milazzo et al., 2011; Talwani, 2002).

There is urgent need for improving our understanding of fundamental aspects of the arenavirus replication strategy, in order to identify potential novel therapeutic pathways. Utilising LUJV NP as a model for other newly emerging infectious arenaviral nucleoproteins could also help identify potential therapeutic strategies for other related arenaviruses of clinical importance.

Due to the limited coding capacity for RNA viruses, their proteins often perform multiple discrete functions, and the arenavirus nucleoproteins are excellent

examples of this. While the best characterised role of NP is that of a structural protein, encapsidating the RNA genome and interacting with other viral proteins during assembly (Pedersen & Konigshofer, 1976), the NP of several arenaviruses have been shown to be involved in non-structural processes. For example, viruses from both OW and NW clades have been shown to possess exoribonuclease activity (Huang et al., 2015; Qi et al., 2010; Hastie et al., 2012); inhibition of the IKK activation of IRF3 and NFκB translocation (Martínez-Sobrido et al., 2007; Rodrigo et al., 2012; Pythoud et al., 2012); direct inhibition of RIG-I (Pythoud et al., 2012); and transcriptional antitermination (Tortorici et al., 2001). Whilst some of these functions are only evident during infection, NP can accomplish several in the absence of any other viral proteins (Huang et al., 2015; Baird et al., 2012; Hastie et al., 2012). Another way that viruses are able to maximize their limited coding capacity is to subvert cellular functions for their own benefit, through interacting with cellular proteins. We hypothesise the arenavirus NP must interact with multiple cellular proteins in order to efficiently facilitate the infectious cycle.

The described body of research has established the role of NP in multiple stages of the infectious cycle, yet it is prudent to acknowledge that other undescribed functions may exist. Given the need for effective therapeutic strategies for arenavirus infections, establishing previously unknown essential roles for NP may offer additional targets for therapeutic intervention.

To aid in this process, the experiments described in this chapter aim to identify novel cellular interacting partners of LUJV NP. To facilitate this investigation, the strategy described here involved the expression of an EGFP (Enhanced Green Fluorescent Protein) NP fusion protein, which we precipitated from transfected cells using high affinity and specificity camelid single chain antibodies. This antibody constitutes a so-called 'GFP-Trap', and the high specificity of this immunoprecipitation (IP) protocol allows reduced non-specific precipitation of many cellular proteins that can result in the identification of false positive interactions. In order to quantify precipitated cellular proteins and thus determine their abundance, we utilised 'stable

isotope labelling of amino acids in cell culture' (SILAC) based mass spectrometry (MS). Initial MS-based identification of potential interacting partners then allowed for further investigation of related proteins using technical replicates and independent validation techniques, in order to establish a role of identified interactions in the virus life cycle. Our strategy involved the expression of NP in isolation of other viral proteins or RNAs, in order to examine cellular interactions that were mediated by NP alone.

4.1.1. Introduction to SILAC based MS utilising GFP-Trap co-immunoprecipitation

The unbiased identification of precipitated interacting cellular proteins is possible through the utilisation of SILAC based MS, in combination with the GFP-Trap based IP. The method utilises a target protein – LUJV NP-EGFP – and an EGFP control, with interacting cellular partners binding to the target and control proteins, and thus being precipitated once exposed to the EGFP monoclonal antibody conjugated to a bead matrix. SILAC represents a quantitative proteomic technique that offers a means to compare multiple cell populations; in this case one of which is expressing the LUJV-NP-EGFP fusion protein, and another population expressing EGFP alone. Comparison of the resulting MS data sets allows the identification of proteins precipitated by virtue of their interaction with EGFP or irrelevant interaction surfaces on the bead matrix; these proteins can then be eliminated as false positives.

The growth of cell populations in different stable isotopically labelled cell culture media allows for the differentiation of proteins. After a succession of passages, these isotopes become incorporated into proteins to levels greater than 95% (Ong et al., 2002). Due to the nature of MS sample preparation through in-gel trypsin digestion, precise SILAC isotope selection is important. The most common medium recipe includes heavy isotopes of both arginine and lysine, which ensures that at least one residue of labelled isotope will be included in each peptide, due to trypsin cleaving polypeptide chains after either arginine or lysine residues. These peptides can then be distinguished from one another, as peptide length should be identical for sister cell

populations after trypsin digestion, with the only anomaly being the presence of differing isotopes.

In the following use of SILAC based proteomics, three media 'weights' were used. The 'light' R0K0 media for control samples with no labelled amino acids included; the 'medium' R6K4 ^{13}C labelled arginine and D (^{2}H) labelled lysine; and finally the 'heavy' R10K8 of ^{13}C and ^{15}N labelled arginine with ^{13}C and ^{15}N labelled lysine. The use of this combination of isotopes enables the protein of interest to be run in duplicate.

When these digested peptides are analysed via MS, pairs of differentially labelled peptides can be identified due to a shift in the mass/charge ratio (m/z). The intensity of these peaks provides quantification of the samples – control versus target. All identified proteins are given a ratio, a ratio of 1 corresponds to equal abundance of an identified peptide in control and target precipitated proteins, with a higher ratio corresponding to increasing confidence that the interaction is with the target alone. This is, however, not without issue. Certain interacting partners may not be able to bind effectively with their partner once EGFP is bound to beads, or conversely, the identified protein may be non-specifically binding to motifs present on both the EGFP and LUJV – each individual not enough to bind after washing, but the fusion peptide may allow for the non-specific binding to occur.

As discussed above in 2.2.5 and shown schematically in figure 16, IPs were performed utilising the Mix After Precipitation (MAP) method. In this case, separate lysis and precipitation steps are carried out on cell populations, with the precipitated proteins mixed in equal volumes after elution from beads. This method allows greater identification of weaker or dynamic partners, as there is a greater availability of binding sites, which may still produce a lower abundance ratio. However, as there are more steps required for the MAP method, there is an increase in chance of experimental error, minimising this is essential to have confidence in the MS library of identified proteins.

4.2. <u>Results</u>

4.2.1. Expression and purification of LUJV NP EGFP

The LUJV NP ORF cDNA was synthesised and subsequently inserted into a pEGFP-N1 expression vector, such that it would direct the expression of LUJV NP fused with EGFP at the C-terminus with the addition of two alanine residues as linkers. In order to establish whether this plasmid directed effective expression of LUJV NP-EGFP, HEK293T cells were transfected using lipofectamine 2000 and harvested 24 hours post transfection. HEK293T cells were chosen due to their human origin and their tissue type, as the kidney is a common organ associated with arenaviral infection (Sewlall et al., 2014; Bergeron et al., 2012; Yun & Walker, 2012). In addition, a further important benefit of using human cells is that it allows the accurate identification of detected peptides, due to the extensive characterisation and annotation of the human proteome.
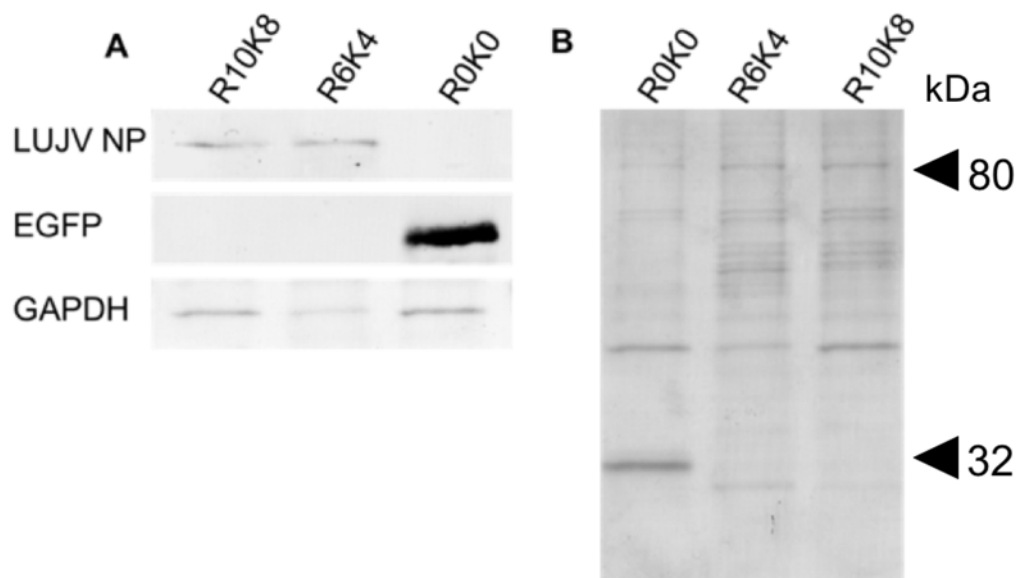
A Western blot analysis of GFP-Trap IPs from differentially labelled HEK293Ts. To visualise LUJV NP-EGFP (termed LUJV NP) an anti EGFP antisera was used to detect the EGFP fusion protein. **B** represents a coomassie stain of **A**, showing GFP-Trap IPs from differentially labelled HEK293T cells. R10K8 and R6K4 were transfected with LUJV NP-EGFP, and R0K0 transfected with EGFP only. Approximate kDa shown in **B**

Figure 31.   **Visulaisation of SILAC IP via western blot and coomassie stain**

Figure 31**A** shows a western blot of EGFP and LUJV NP-EGFP expression in HEK293T cells. Differentially labelled cells were transfected with EGFP in R0K0 control media, and LUJV NP-EGFP in both R6K4 and R10K8 media. While figure 31**B** shows the expression of both LUJV NP-EGFP and EGFP, along with other precipitated proteins via coomassie stain. The presence of additional proteins besides exogenously expressed EGFP or LUJV NP-EGFP within precipitated samples suggests cellular proteins were co-precipitated, and we anticipated these would be identified though MS.

4.2.2. SILAC LC-MS/MS identification of LUJV NP-EGFP interacting partners

In order to generate a library of cellular interacting proteins (from 4.2.1) which immunoprecipitated (IPd) alongside EGFP control or LUJV NP-EGFP, samples were identified and quantified by SILAC based MS, performed by Dr Stuart Armstrong at the University of Liverpool.

As part of the analysis through the Mascot search engine during data collection, each protein identity is assigned a score. This automatically generated confidence score is based on the protein score, with a higher score indicating that it is less likely that the identification took place by chance. This 'protein score' is algorithmically generated, based on protein coverage through peptide identification and the combined scores of observable mass spectra, which can be matched to known amino acid sequences within a given protein. This can also be used to derive approximate protein quantitation but is only a rudimentary representation of true abundance. Thus the confidence score correlates to the likelihood that the identification of individual proteins is not due to background contamination or experimental noise.

MS identification of IPd proteins resulted in the generation of unique SILAC ratios, or abundance ratios, for each identified protein; in essence, this ratio indicates the relative abundance of identified proteins within the sample (LUJV NP-EGFP) compared to the EGFP only control EGFP. In cases where proteins are not present within the control fraction, no ratio will be generated, but they could still be regarded as genuine interacting partners. Total unique

proteins identified numbered 276, with 86 generating an abundance ratio between sample and control. A protein frequency graph was generated for proteins which exhibited abundance ratios, shown in figure 32**A**. Among these, 32**B** shows a selection of the identified proteins, based on the peptide analysis described in 2.2.5 and its corresponding ratio for proteins discussed below. A complete list of immunoprecipitated proteins can be found in Appendix II.

A
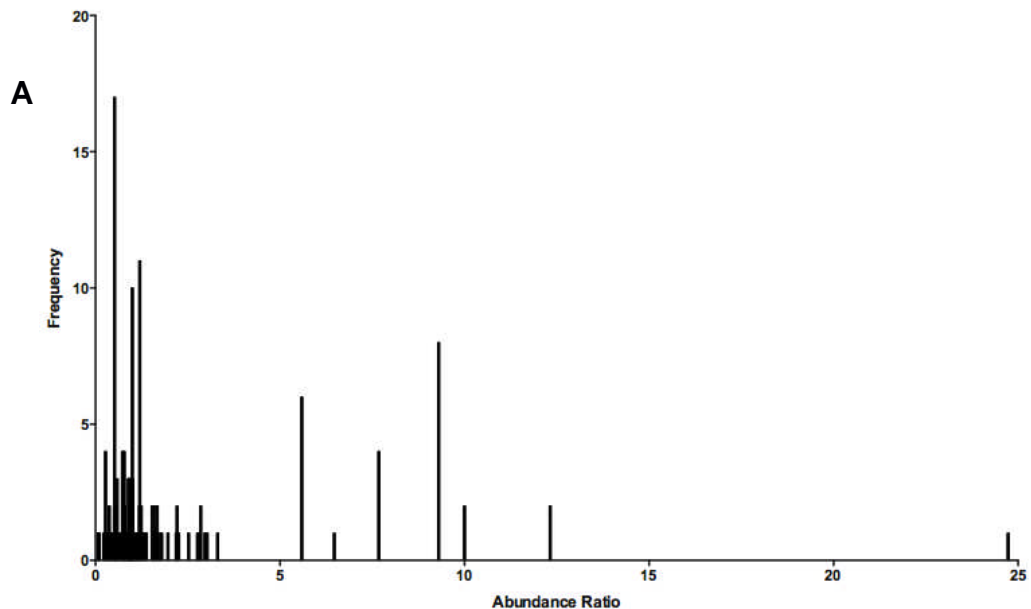
B

Protein	Abundance Ratio
HUMAN Zinc transporter SLC39A7	24.732
MYH9_HUMAN Myosin-9	9.9984
ACTB_HUMAN Actin	9.2999
RS3A_HUMAN 40S ribosomal protein S3a	2.9661
RL26_HUMAN 60S ribosomal protein L26	2.2099
RL11_HUMAN 60S ribosomal protein L11	2.1879
RS10_HUMAN 40S ribosomal protein S10	1.6716
RS19_HUMAN 40S ribosomal protein S19	1.5959
RL10_HUMAN 60S ribosomal protein L10	1.5385
HSP71_HUMAN Heat shock 70 kDa protein 1A/1B	1.3326
HSP7C_HUMAN Heat shock cognate 71 kDa protein	1.238
IF4A1_HUMAN EIF4A	0.94091
HS90B_HUMAN Heat shock protein HSP 90-beta	0.9378
DNJA1_HUMAN DnaJ homolog subfamily A 1	N/A
DNJA2_HUMAN DnaJ homolog subfamily A2	N/A

A) Graph showing a representation of abundance ratio of proteins against detection frequency following SILAC MS analysis. **B)** Table showing a selection of proteins identified via SILAC MS. Selection based on parameters outlined in section 4.3.1

Figure 32. **LC MS and MASCOT engine generated abundance ratios of LUJV NP-EGFP after SILAC IP**

4.3. <u>Validation of identified interacting partners</u>

4.3.1. Choosing interacting partners for further study

As discussed above, there are shortcomings with SILAC-based quantitative MS. While confidence in the validity of interactions can be good with incorporation of appropriate controls within experimental procedures, the likelihood for experimental contamination is great due to the high sensitivity of currently available MS equipment. MS identification alone does not confirm an interaction, and indeed the absence of a protein from the MS dataset does not exclude it as an interacting partner. In order to accurately establish an identified protein as a true binding partner, resulting data sets must be tailored using detailed knowledge of the limitations of the SILAC and GFP-Trap methodologies, as well as a well-informed understanding of relevant cellular biology and the virus lifecycle. In this context, the use of a single viral protein for the interaction target may create problems, as some viral components may only exhibit certain functions when in multi-component complexes. Futhermore, dynamic interactions, or important partners within a complex may not be identified due to the conditions of the IP. It would therefore be unwise to dismiss all identified proteins corresponding to peptides quantified below an arbitrary abundance ratio, and to assume that high abundance ratio partners are relevant or important interacting partners.

In order to identify possible partners, it is useful to look at the particular roles viral proteins may have within the viral replication cycle, in conjunction with the established roles of similar proteins in different viruses. Our approach was to treat the proteomics analysis as a means to apply an initial triage of potential interacting proteins, a snapshot of the global interactome. Firstly, identified cellular proteins with known roles in related virus nucleocapsid proteins – such as other members of the *Arenaviridae* family or other segmented negative sense viruses such as CCHFV. Secondly, proteins which exhibited a high abundance ratio from the identified SILAC library are likely,

but not certain, to be good candidates for further validation. The validation process is essential, as it helps give confidence to the library, whilst also opening new avenues for further validation of cellular proteins with related functions. By applying the criteria described above, we generated a priority listing of potential interacting partners that we sought to validate using multiple alternative and complementary means.

Actin is a well-studied protein responsible for multiple internal cell transport functions, muscle contractions and cell motility (Grummt, 2006; Pantaloni et al., 2001; Huber et al., 2013). The association of viral proteins with such an important and abundant cytoskeletal protein is not without precedent, with Epstein-Barr virus, hepatitis B virus and adenoviruses among those known to subvert the functions of actin for their own means. The utilisation of actin by arenavirus NP is not unexpected; although assembly of arenaviruses and internal transport is predominantly attributed to the small matrix protein Z (Urata & Yasuda, 2012; Urata et al., 2006), and the roles of NP are thought to predominantly surround the stability and facilitation of RNA replication. Viral proteins utilising the cytoskeletal network in RNP trafficking is a well-established function for nucleocapisd proteins from other RNA viruses, including the closely related N of CCHFV (Andersson et al., 2004). In the case of actin, identification of the utilisation of cytoskeletal proteins by NP could open an avenue for further validation of other members of this group, such as myosin-9 which generated the second highest abundance ratio.

The protein with the highest abundance ratio change between EGFP and LUJV NP-EGFP was Zn Transporter SLC39A7, a Golgi-ER Zn transporter protein. This protein was not investigated due to limited knowledge of the potential role of this protein interacting with NP, in addition to the lack of readily available antibody to the protein to characterise its relationship with NP. However, it would remain a good candidate for further investigation regarding the protein interactome of NP.

Eukaryotic initiation factor 4A (eIF4A) is an RNA helicase, responsible for the unwinding of mRNA, processing it from its secondary structure and therefore

facilitating codon scanning for the initiator codon (Rogers et al., 2002; Oberer et al., 2005). NW members of the *Arenaviridae* family have been shown to interact with components of the eIF4F complex (Linero et al., 2013) and thus the presence of eIF4A in the LUJV NP interactome would potentially indicate a cross-clade property of this protein.

The abundance ratio of eIF4A was 0.94, and given that this lies below the desired cut-off, it would be possible to dismiss it as a contaminant. However, as discussed above, dismissing potential partners based solely on the MS abundance ratios is not prudent. As eIF4F complex proteins are known to be involved in the replication strategies of the NW arenavirus JUNV and influenza A virus, eIF4A was chosen as a potential partner for further investigation. (Furic et al., 2010; Yanguez et al., 2011; Linero et al., 2013). eIF4A was chosen to also act as a representative for translation complex proteins – such as the ribosomal subunit proteins identified via SILAC MS, including Ribosomal Protein (RP) L10, L26, S11 and S19. While its identification as an interacting partner would not confirm the presence of other translation associated proteins, it would act as marker to test for other translation associated proteins, such as those described above. Similarly, if eIF4A was not identified, it would not negate the possibility that other translation associated cellular proteins do interact with LUJV NP.

Finally, HSP70 is a known chaperone protein, responsible for binding newly synthesised polypeptides as they emerge from the elongating ribosome (Mayer & Bukau, 2005). Host-cell chaperone proteins are of importance to viral protein maturation, given their total requirement for cell-based glycosylation enzymes, and potentially in order to maintain proteins in monomeric states. Without proper folding, glycosylation is impossible, and creates a rate limiting step in viral replication (Sayce et al., 2010; Dalziel et al., 2014). While NP, which is fundamentally an RNA binding protein and possesses no glycosylation sites, associating with a chaperone protein may seem circumstantial, recruiting multiple chaperone proteins to sites of replication could greatly increase functional viral protein expression. With a

single protein responsible for recruiting so many factors, efficiency could be increased, and thus infection spread also. Recent work regarding DENV replication has shown that chaperone proteins are essential co-factors throughout each stage of infection, with the DnaJ chaperone family essential in maximising viral output. If HSP70 were confirmed as a valid interacting partner of LUJV NP-EGFP, the DnaJ family proteins could then be pursued further, given their known association with HSP70 (Cyr et al., 1994). Given that they were identified solely within LUJV NP-EGFP transfected cells, they would be a good target for future analysis in the event of an HSP70-NP interaction.

4.3.2. GFP- and RFP-Trap co-immunoprecipitation of EGFP and LUJV NP-EGFP visualised via western blot
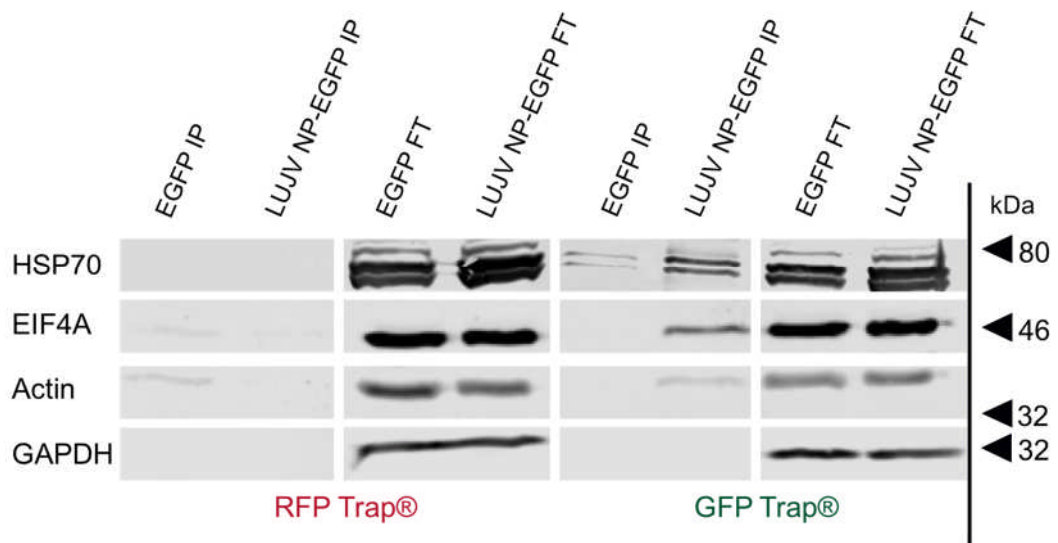
Western blot analysis of cellular proteins isolated via GFP- and RFP-Trap immunoprecipitations. In all instances, HEK 293T cells were transfected with EGFP or LUJV NP-EGFP expressing plasmids – as stated. 'EGFP FT' and 'LUJV NP-EGFP FT' indicates the flow through fraction of cell lysates exposed to GFP/RFP trap beads. 'EGFP IP' and 'LUJV NP-EGFP IP' indicates the immunoprecipitated proteins exposed to GFP/RFP-Trap beads.

Figure 33. **Western blot analysis of interaction between cellular proteins and LUJV NP-EGFP after GFP-Trap IP**

As described above, eIF4A, actin and HSP70 were selected for initial validation of the MS proteome analysis. EGFP and EGFP-fused proteins co-precipitate with their interacting cellular partners, enabling the detection of

those cellular proteins by western blot. The absence of proteins of interest in the EGFP IP fraction would indicate a lack of contamination or non-specific interaction with EGFP, with presence of the cellular interacting protein in the LUJV NP-EGFP fraction indicating a valid interaction between the cellular protein and LUJV NP. Any presence of interacting proteins in the RFP-Trap IP fractions would also indicate a false positive, with the protein interacting with the camelid antibodies rather than NP. HEK 293T cells were differentially transfected using lipofectamine 2000 with EGFP and LUJV NP EGFP and incubated for 48 hours. Cells were then lysed and separated into soluble and insoluble fractions, with soluble lysate loaded onto clarified GFP- and RFP-Trap agarose beads. A "non-bound" or flow-through (FT) fraction was taken, with beads then washed prior to elution of immunoprecipitated sample.

Figure 33 indicates that both actin and eIF4A show a valid interaction with LUJV NP-EGFP alone. For eIF4A, a strong signal was produced in the LUJV NP IP fraction, and was absent from the EGFP IP fraction. This indicates that the IP of eIF4A was due to the presence of LUJV NP. Actin, similarly, produced a signal in only the EGFP-LUJV NP IP fraction. For HSP70, a robust signal was detected in the LUJV NP-EGFP IP, although some precipitation of HSP70 bound to EGFP alone was evident. This suggests that some of this precipitated HSP70 may be due to a weak EGFP interaction. The HSP70 antiserum used in the analysis is known to also react with HSC70, thus producing two distinct bands, and also suggests that HSC70 is a potentially valid interacting partner, also being identified in the MS analysis (Appendix II). The lack of any observable protein precipitated during RFP-Trap immunoprecipitation for all three IPd proteins indicates that the proteins are not non-specifically binding to the conjugated antibodies. These findings suggest that eIF4A, Actin and HSP70/C70 are valid interacting partners of LUJV NP-EGFP.

4.3.3. Immunofluorescence microscopy

4.3.3.1. Localisation

In order to determine whether LUJV NP-EGFP is able to re-distribute these identified interacting partners within the cellular environment, it was necessary to determine the normal distribution of LUJV NP-EGFP. HEK293T cells were therefore differentially transfected with EGFP or LUJV NP-EGFP and subsequently fixed in 4% formaldehyde after 36 hours, permeabilised and mounted in DAPI mounting media for visualisation via confocal microscopy. Figure 34 shows the normal distribution of EGFP and LUJV NP-EGFP. LUJV NP-EGFP exhibits a predominantly cytoplasmic distribution, with distinct cytoplasmic puncta visible throughout. These puncta mimic those generated by NP and observed during JUNV infection (Baird et al., 2012) which were described as representing the sites of viral replication, and named as replication and transcription complexes (RTCs). The identification of a cross-clade utilisation of distinct cytoplasmic structures for replication could help identify potential avenues for therapy which target these pathways utilised by the virus.
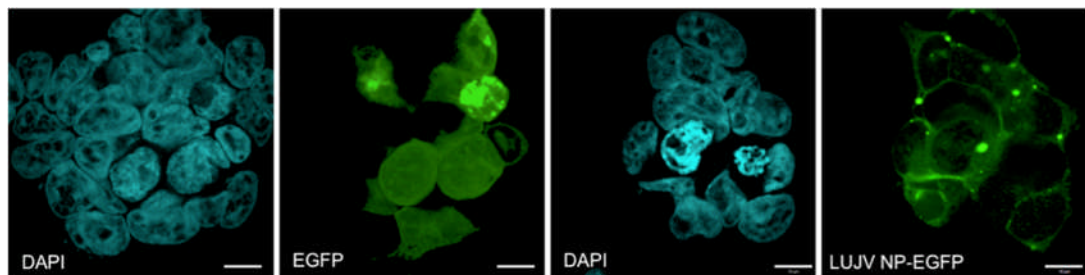
Diagram illustrating the localisation of LUJV NP-EGFP in HEK293T cells. Cyan indicates DAPI staining; Green indicates EGFP or LUJV NP-EGFP – as stated. The scale bar indicates 10μm.

Figure 34.   Localisation of LUJV NP-EGFP in HEK293T cells

4.3.4. Localisation of identified interacting partners

In the previous sections, LUJV NP-EGFP was found to be present within the cytoplasm, but also within distinct cytoplasmic puncta. In addition, Actin, HSP70 and eIF4A were found to immunoprecipitate with LUJV NP-EGFP. In order to determine whether these proteins were redistributed within cells and

co-localising with NP, HEK293Ts grown on cover slips were differentially transfected with EGFP and LUJV NP-EGFP, and incubated for 48 hours. Cells were then fixed, permeabilised and stained with relevant antisera prior to mounting in a DAPI containing mounting media for nuclear staining. Figure 35 shows the distribution of eIF4A, HSP70 and Actin in HEK293T cells, with and without the expression of LUJV NP-EGFP, against an EGFP control. The results indicate no significant redistribution of HSP70 in 35**A** within LUJV NP-EGFP expressing cells. Given the relatively strong indication of an interaction via western blot, this is surprising. While complete redistribution of HSP70 would not be expected due to the requirement of HSP70 as a cellular protein chaperone, but minor co-alignment might be expected. However, as discussed above, NP is expressed in the absence other viral proteins, within a living cell environment it may not encounter HSP70 due to the internal compartmentalisation of the cell.
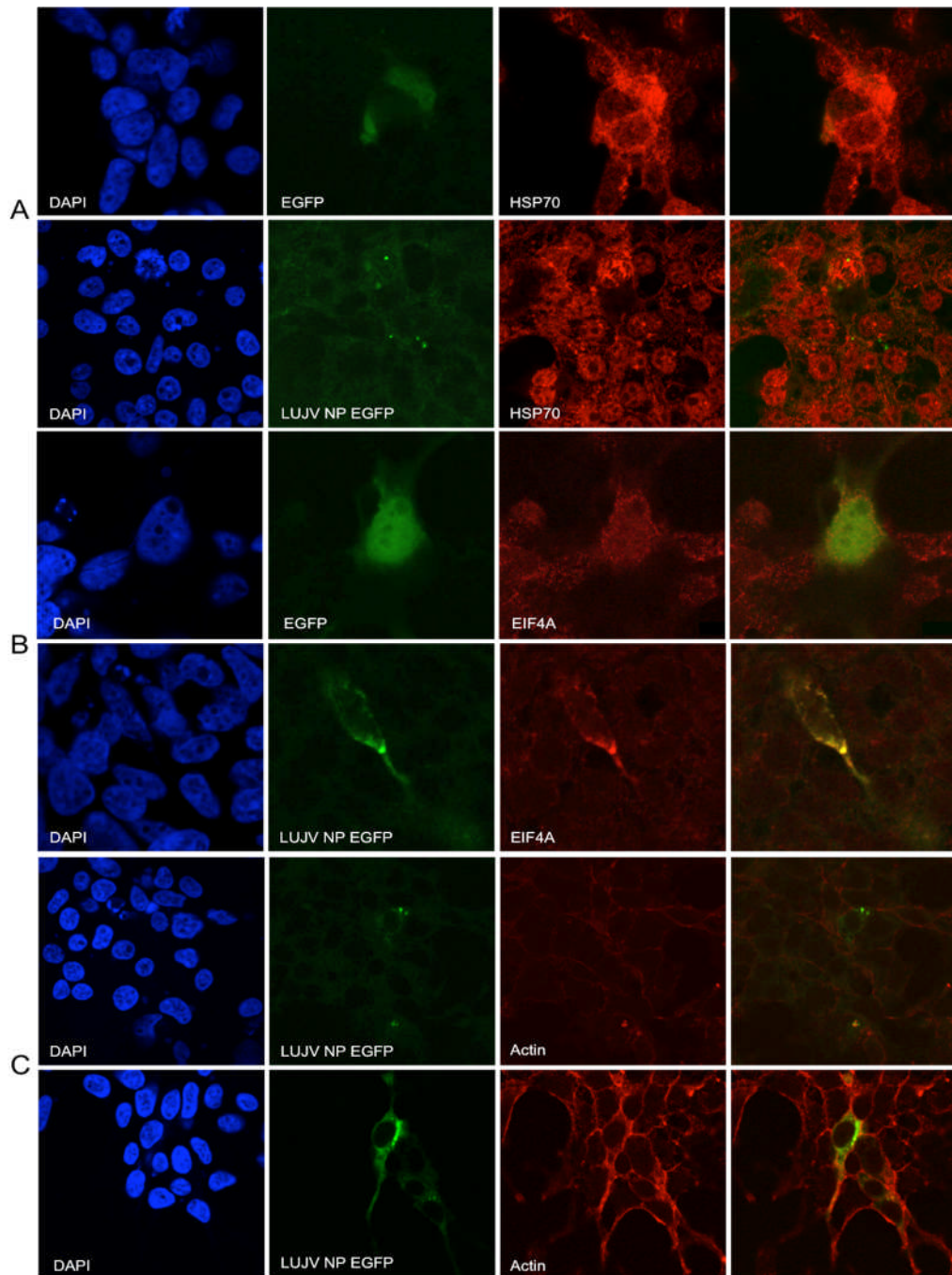
Immunofluorescence analysis of the alignment of LUJV NP-EGFP with specific cellular proteins, as compared with EGFP control, in HEK293T cells. **A:** HSP70 distribution, compared between EGFP and LUJV NP-EGFP; **B**: Distribution of eIF4A; **C:** Distribution of Actin; In **A-C**, Blue indicates DAPI stain of DNA, Green indicates EGFP/LUJV NP-EGFP (as stated), Red indicates cellular proteins – identified as shown.

Figure 35.   **Comparative localisations of LUJV NP-EGFP and HSP70, eIF4A and actin**

However, the observation of HSP70-NP interaction in Figure 34 could still be valid, as after lysis, NP would be able to interact with HSP70 irrespective of normal compartmentalisation. Actin shows some co-localization in 35**C**, into discreet cytoplasmic puncta, suggesting that NP might interact with actin. eIF4A also shows strong co-alignment with LUJV NP-EGFP, as shown in 35**B**, exhibiting a robust signal within observed cytoplasmic puncta. The mechanism responsible for this co-localisation is not possible to determine, but the presence of a protein responsible for mRNA processing in SILAC IP MS, western blot analysis and exhibiting co-localisation with LUJV NP-EGFP points toward a functional interaction with LUJV NP.

4.4. Investigating the interaction between LUJV NP-EGFP and multiple translation associated proteins

The results of the previous section showed LUJV NP co-localising with eIF4A, an important component of the translation initiation complex. Next whether LUJV NP interacted with other components of the translation initiation complex, as well as any ribosomal subunit proteins was investigated. Other cellular translation initiation factors and ribosomal subunit proteins were identified in the initial SILAC MS dataset, as shown in figure 32**B** and Appendix II, and these were thus selected for further interaction validation, along with other functionally-related proteins not identified via MS.

Ribosomal subunit proteins act as constituent parts within one of two of the major subunits of an active ribosome – the 40S and 60S subunits. The functions of these individual subunit proteins outside of general translation as part of a ribosomal complex is not wholly understood, with implications in tumour development (Takagi et al., 2005). However, collectively they enable the translation of mRNA into a polypeptide sequence (Ramakrishnan, 2002). Ribosomal subunit proteins selected for validation as LUJV NP interacting partners were those identified via SILAC MS; namely RPS19 and RPL10. In addition, eIF2G was identified via MS but solely in the LUJV NP IPs, thus not generating abundance ratios versus the EGFP control. The selection of these ribosomal subunit proteins would indicate that components of both the 40S

and 60S ribosomal subunits were IPd alongside LUJV NP, indicating that NP interacts with more than simply initiation factors but potentially ribosomal complexes. In addition, identifying the presence of an additional translation initiation factor - eIF2G – would confirm that the interaction of eIF4A and NP observed earlier was not the only translation associated protein to be IPd by LUJV NP.

4.4.1. LUJV NP association with elements of translation initiation complex proteins.

4.4.1.1. Immunoblot analysis
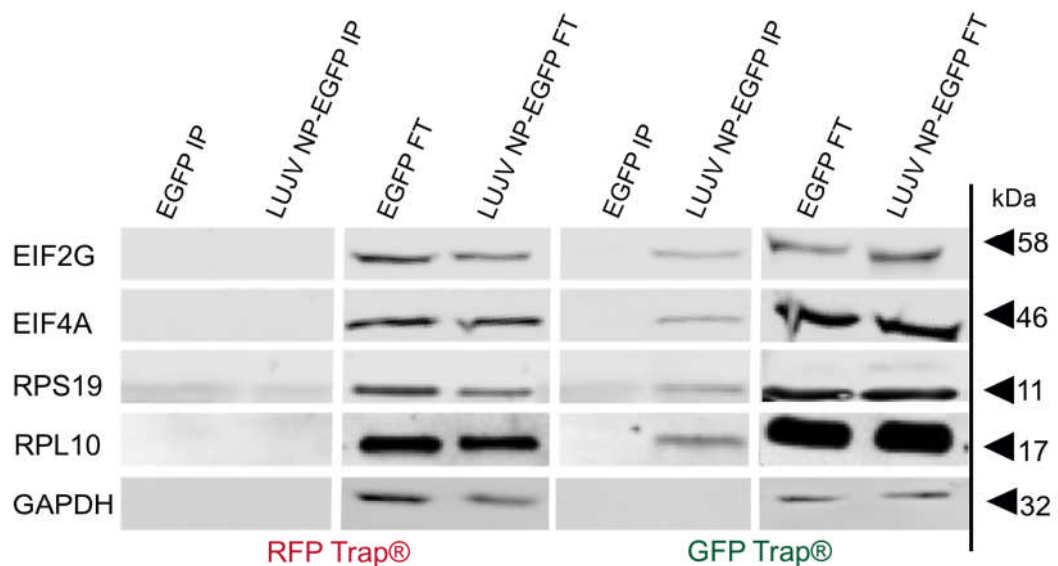
Western blot analysis of cellular proteins isolated via GFP- and RFP-Trap immunoprecipitations. In all instances, HEK 293T cells were transfected with EGFP or LUJV NP-EGFP expressing plasmids – as stated. 'EGFP FT' and 'LUJV NP-EGFP FT' indicates the flow through fraction of cell lysates exposed to GFP/RFP trap beads. 'EGFP IP' and 'LUJV NP-EGFP IP' indicates the immunoprecipitated proteins exposed to GFP/RFP-Trap beads.

Figure 36.   WB analysis of LUJV NP-EGFP co-IP with cellular translation associated proteins

HEK293T cells were differentially transfect ed with EGFP and LUJV NP-EGFP and incubated for 48 hours prior to lysis. Lysates were then separated into soluble and insoluble fractions, with the soluble fraction added to clarified GFP- and RFP-Trap beads and incubated at 4°C for one hour, in addition. A flow through fraction was collected, with beads then washed prior to elution of

IPd sample. eIF4A was included as a positive control due to its previous identification as a protein immunoprecipitated with LUJV NP-EGFP. In addition, eIF2G, which had been previously identified via SILAC based MS, was selected for study in order to determine whether NP interacts with different elements of the initiation complex, indicating an involvement of NP in complex recruitment. Figure 36 shows the presence of eIF2G, eIF4A, RPS19 and RPL10 all co-immunoprecipitated with LUJV NP-EGFP via a GFP-Trap and absent from EGFP and RFP-Trap controls. The presence of these factors indicates LUJV NP interacts with a number of translation-related factors, including elements of both the 40S and 60S ribosomal subunits. These findings point towards an association of NP with ribosomes and certain translation initiation complex proteins.

4.4.1.2. Immunofluorescence analysis

The results of the previous section indicated that LUJV NP interacts with multiple translation complex proteins. Next, the ability of LUJV NP to influence the cellular distribution of these identified interacting partners into characteristic cytoplasmic puncta was tested using immunofluorescence microscopy. HEK293T cells grown on poly-L lysine treated cover slips were differentially transfected with either EGFP or LUJV NP-EGFP and incubated for 48 hours. Cells were then fixed with formaldehyde, permeabilised and subsequently stained using appropriate antisera. In addition to the eIF4A, eIF2G, RPL10 and RPS19, the localisation of additional translation associated factors was observed against LUJV NP-EGFP, in order to establish whether NP co-localises with different components of the translation initiation complex. Thus, a factor from four of the major translation initiation factor groups was selected; eIF1A, eIF2G, eIF3E, and eIF4A, along with the 60S ribosomal subunit proteins RPL10 and L26, and the 40S subunit protein RPS11.
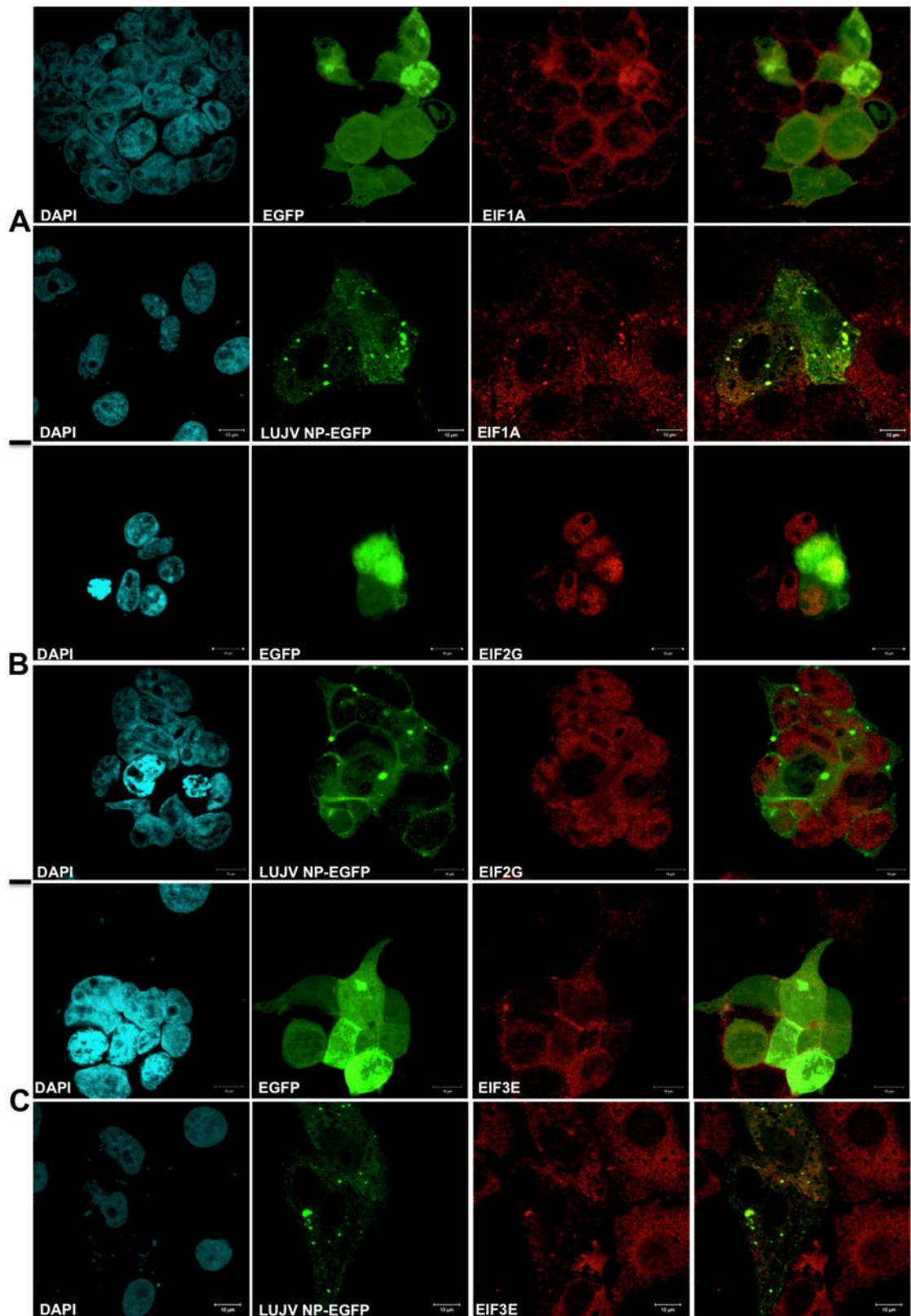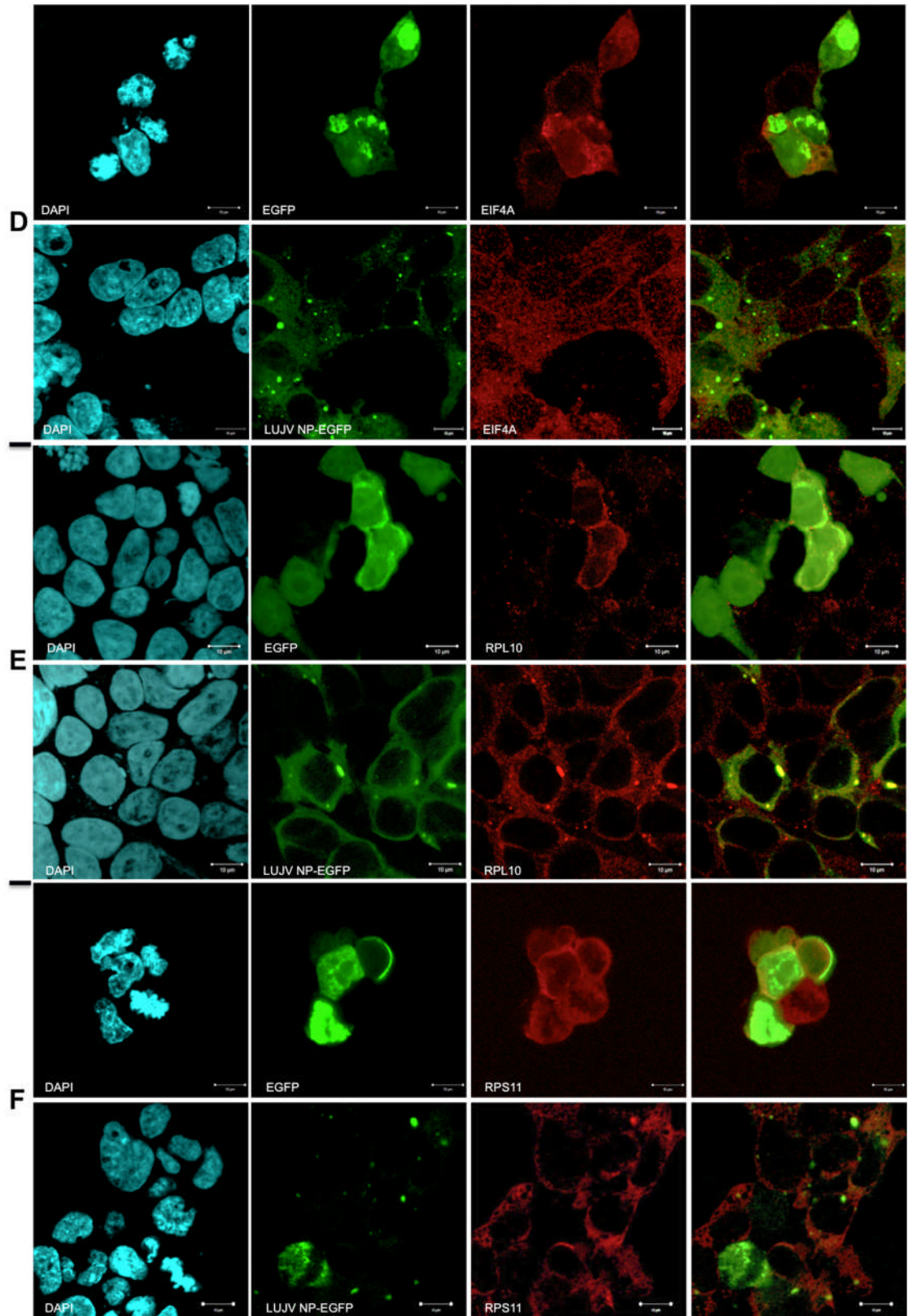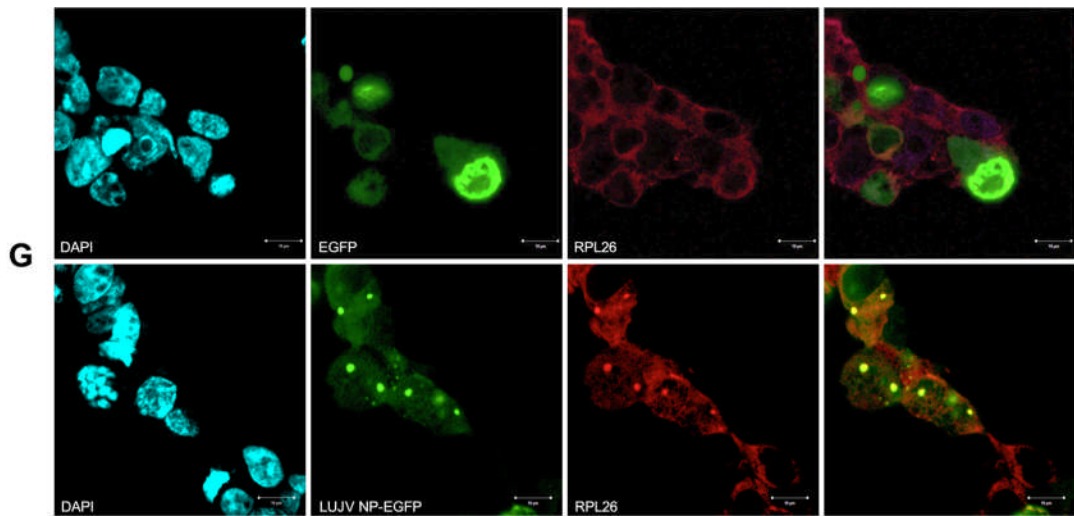
Immunofluorescence analysis the of alignment LUJV NP-EGFP with specific cellular proteins, as compared with EGFP control in HEK 293T cells. **A-G** show the comparison of a cellular protein localisation compared between EGFP and LUJV NP-EGFP expressing cells **A:** eIF1A **B**: eIF2G; **C:** eIF3E **D**: eIF4A; **E** RPL10A; **F:** RPS11; **G:** RPL26. In **A-G**, Cyan (channel 1) indicates DAPI stain of DNA, Green (channel 2) indicates EGFP/LUJV NP-EGFP (as stated), Red (channel 3) indicates host-cellular protein staining (as stated). The channel 4 in **A-G** shows the merge of channels 2 and 3. Scale bar indicates 10µm.

Figure 37. **Comparative localisation of LUJNV NP-EGFP and cellualar translation associated proteins**

As seen in figure 37**A** the initiator tRNA^{MET} transfer protein (Chaudhuri et al., 1997), eIF1A, was not identified via MS but did exhibit a degree of redistribution within LUJV NP-EGFP expressing cells, into the observed puncta, as opposed to a diffuse distribution observed in EGFP expressing cells. Conversely, although eIF2G was identified as an LUJV NP interacting partner via both MS and via specific WB immunostaining, no evident difference between its distribution was observed between EGFP and LUJV NP-EGFP expressing cells 37**B**. In 37**C** eIF3E, a component of the eIF3 complex responsible for positioning incoming mRNA by facilitating pre-initiation complex (PIC) formation (Jackson et al., 2010), shows a degree of re-distribution into cytoplasmic RTC-like structures. This is notably different to the normally diffuse arrangement in EGFP expressing and un-transfected cells. As in figure 35**B,** eIF4A exhibits a partial co-alignment with LUJV NP as seen in 37**D**. The ribosomal subunit proteins RPL10, RPS11 and RPL26 in 37**E, F** and **G** respectively all showed a degree of co-alignment, supporting

115

the identification of ribosomal subunit proteins via GFP-Trap IPs in figure 36. The use of RPS11 in this instance, as opposed to RPS19, was due to the unsuitability of the RPS19 antisera for IF microscopy. In addition, RPL26 antibody is not recommended for use in WB analysis, whereas visualisation via IF produces a strong indication that RPL26 co-localises with LUJV NP-EGFP in HEK293T cells. The presence of these three ribosomal subunit proteins within RTC like-structures, combined with identification by IP, indicates that whole ribosomes could be present within RTC like structures rather than individual constituent proteins. These structures mimic those observed by Baird *et al.* 2012 in process of JUNV infection.

4.4.2. The use of intercellular organelles in the formation of LUJV generated RTC-like structures

In the previous section, LUJV NP-EGFP was shown to co-localise with multiple components of the translation initiation complex. In order to better understand the localization of NP within cytoplasmic puncta reminiscent of arenavirus RTCs, the sub-cellular origin of these cytoplasmic inclusions was investigated. A panel of antibodies that recognised a variety of organelle-specific marker proteins was used to examine whether they localised with LUJV-NP. Vero cells were differentially transfected with either EGFP or LUJV NP-EGFP and incubated for 48 hours, where they were then fixed, permeabilised and stained with the appropriate antisera and secondary antibodies.

In addition to the antisera against organelle marker proteins, the LUJV NP antibody produced in the previous chapter was used. Figure 38**A** demonstrates the relative distribution of nuclear pore complex protein 98 (NUP98) a dynamic component of the nuclear pore complex. While overall distribution of NUP98 remains intact, there is a degree of co-alignment with LUJV NP-EGFP, potentially indicating that NP interacts with elements of the nuclear pore complex. Whilst NUP98 itself was not identified via MS, the complex protein importin-β1 was identified in a single sample fraction, which could indicate a transient or dynamic interaction. Figure 38**B** shows the

relative distribution of β-tubulin to identify whether LUJV NP-EGFP was utilising elements of the microtubule network to re-distribute factors. There was no change in the relative distribution of the network, and no apparent co-alignment, which indicates there is no direct utilisation of these structures by NP. However, NP may utilise these networks in the presence of other viral proteins during an infection, through its previously identified interaction with actin (figures 34&35). 38**C** shows the relative distribution of calnexin, a marker for the endoplasmic reticulum. The lack of any distinct co-alignment with LUJV NP-EGFP indicates that NP does not utilise ER derived membranes for the establishment of RTC like structures.
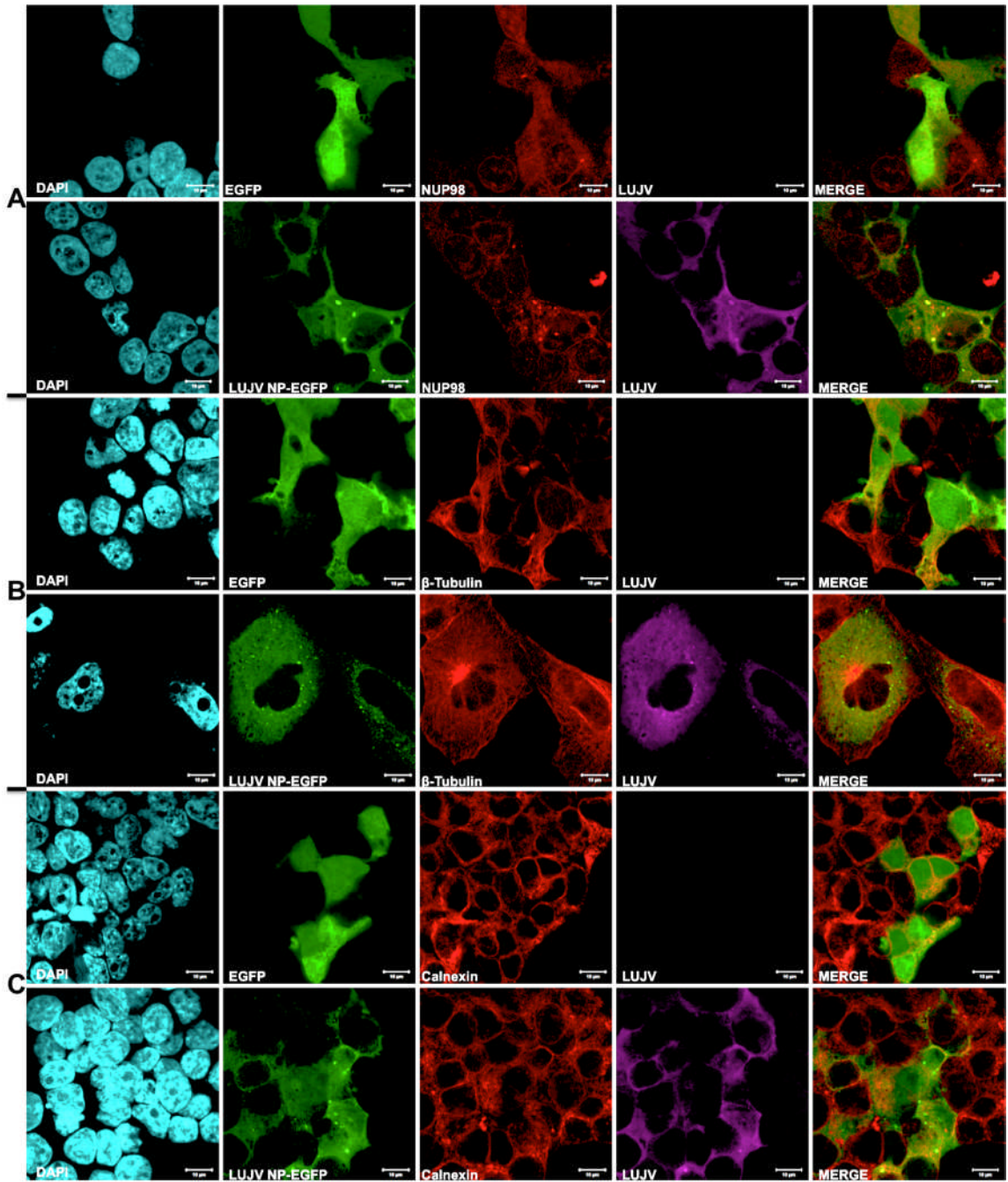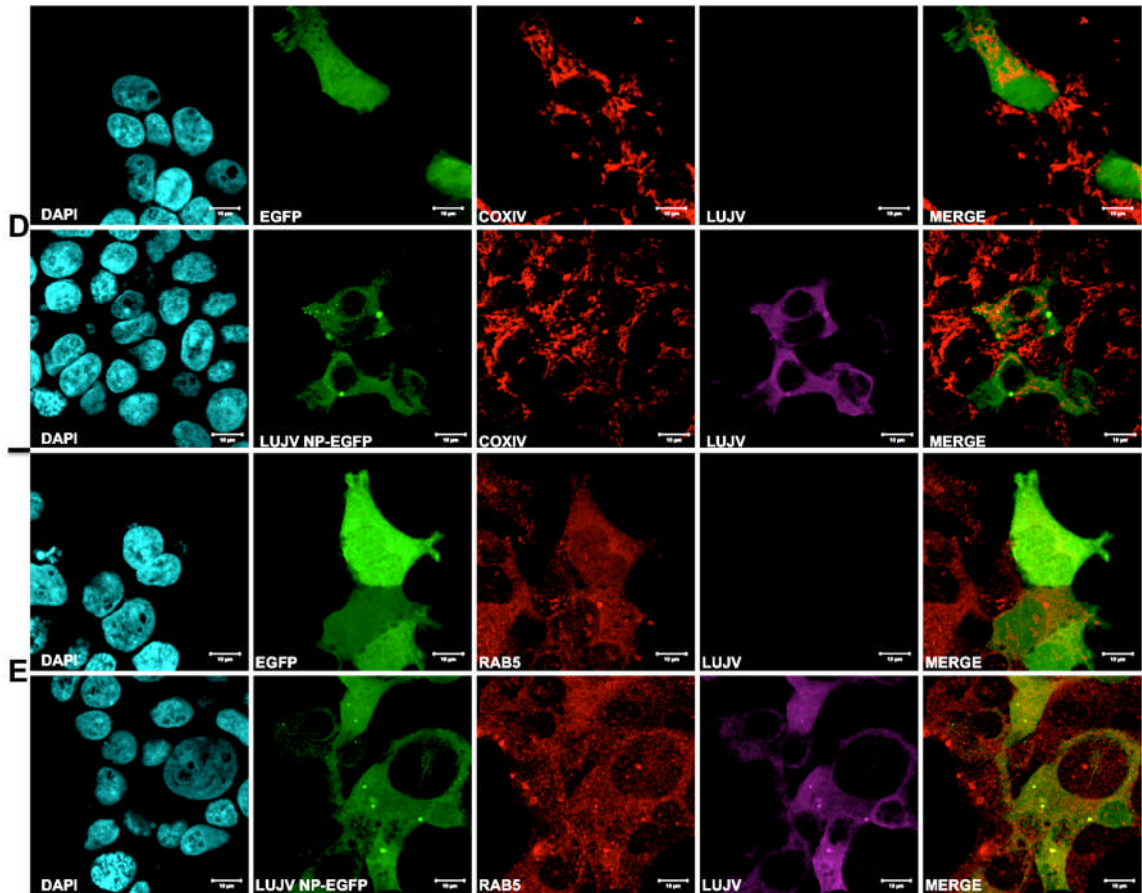
Immunofluorescence analysis of the alignment of LUJV NP-EGFP' with cellular organelle markers, as compared with EGFP control in Vero cells. **A:** NUP98 distribution, compared between EGFP and LUJV NP-EGFP; **B**: Distribution of β-Tubulin; **C:** Distribution of Calnexin; **D**: Distribution of COX IV; **E** distribution of Rab5. In **A-E**, Cyan indicates DAPI stain of DNA, Green indicates EGFP/LUJV NP-EGFP (as stated), Red indicates cellular organelle marker staining (as stated), and magenta indicates LUJV NP staining. Scale bar indicates 10µm.

Figure 38.   Utilisation of specific sub-cellular organelles by LUJV NP-EGFP

In 38**D**, an antibody specific for COX IV a mitochondrial marker indicates the relative distribution of COX IV appears unaltered in the presence of LUJV NP-EGFP vs untransfected or EGFP expressing cells.

Finally, in 38**E**, there is an apparent and distinct co-alignment of LUJV NP-EGFP with the early endosomal marker Rab5, suggesting the corresponding compartments are focal points for NP accumulation. With relevance to the arenaviral life-cycle, the presence of NP within early endosomes as viral factories – hijacked to form an RTC – is encouraging. The localisation of NP within early endosomes is consistent with the known utilisation of early

endosomes as sites of membrane fusion, with arenaviruses internalised and fusion occurring within endosomes (York et al., 2008). The apparent ability of NP to recruit a number of factors within early endosomes suggests that the role of NP in factor recruitment may be crucial to replication.

4.5. **Discussion**

This chapter describes the identities of interacting partners of the NP of the newly discovered pathogenic LUJV, when fused with EGFP. The aforementioned interacting partners – including eIF4A, RPL10 and RPL26 – can be functionally grouped as proteins involved in protein translation with diverse roles in ribosome assembly and translation initiation. These interactions suggest a potential role for NP in facilitating mRNA translation, presumably of virus-specific mRNAs. As observed in cells undergoing JUNV infection (Baird et al., 2012), LUJV NP-EGFP is able to form RTC-like, punctate structures, while co-localising alongside multiple components of the translation initiation complex within early endosomes. This is an observation consistent with the known use of endosomes as the point of arenavirus membrane fusion, resulting in exposure of the arenavirus RNPs to the cytosol (York et al., 2008; Nunberg & York, 2012).

The combination of host-cell factors described indicates a potential role of NP in facilitating the expression of viral proteins. Amongst these proteins identified via IP and IF, there are those that could directly bind NP. It has been shown that NP is able to re-localise a number of such translation-associated proteins, with this being indicative of a protein that has an important role in facilitating the translation of viral mRNAs, a role beyond its well established role as an RNA-binding protein.

The use of unbiased SILAC-based MS to identify cellular interacting partners of LUJV NP-EGFP via a GFP-Trap IP from differentially labelled and transfected HEK293T cells generated a library of potential partners. In total, 276 unique proteins were isolated at least once, with 86 proteins generating an abundance ratio between the control and duplicate LUJV NP sample. As

discussed above, this process is not without issue. The required technique may prevent certain interactions from being observable and introduce experimental contaminants. Indeed, although the relative abundance eIF4A was below that of the control, subsequent validations identify this interaction as genuine. This highlights the importance in conducting validations on the MS-generated library. Validating a number of results does increase confidence in the MS data, but viewing the library as a snapshot in time, rather than a true reflection of a protein's complete interactome during the extended period of the viral replication cycle, is prudent.

Not all proteins identified are feasible to validate in one study, due to time and funding constraints. Identifying patterns between previous knowledge and the identified proteins can help elucidate the nature of the interaction between identified cellular proteins and NP. Two of the identified proteins shown in 29**B** which were not validated further; DnaJA1 and A2 have recently been implicated in DENV infection, with depletion of DnaJA2 significantly depleting DENV production (Taguwa et al., 2015). The DnaJ protein family constitute the largest and most diverse chaperone protein group. The comprehensive work of Taguwa *et al.* identifies DnaJ members as essential in all stages of the DENV replication cycle, with nine DnaJ proteins found to have a significant effect on viral depletion when their expression is blocked (Taguwa et al., 2015). DnaJA2, identified via SILAC MS in LUJV NP-EGFP fractions alone, was found to be important in RNA synthesis during DENV infection. The apparent binding of this protein to LUJV NP should be explored further given the use of the cofactor HSP70 during DENV infection, a validated interacting partner of LUJV NP-EGFP in this study, and a known partner of CCHFV N (Surtees, 2014).

The initial screening of selected proteins – eIF4A, actin and HSP70 – from the MS library suggested that all three could be valid interacting partners via identification after GFP-Trap IP. However, the three proteins did not exhibit similar sub-cellular localisations when viewed via IF in the presence of LUJV NP-EGFP.

HSP70 is a protein with well-established roles in viral infection including the closely-related CCHFV N and HAZV N (Surtees, 2014), rabies virus (Lahaye et al., 2012) and HCV NS5A (Khachatoorian et al., 2014), and was thus a good candidate for further analysis. The interaction between actin and NP, and degree of co-localisation, was expected. Actin is a known partner of related RNA virus nucleocapsid proteins – notably CCHFV (Surtees, 2014; Andersson et al., 2004). This could indicate that LUJV NP utilises the cytoskeletal network to facilitate the recruitment of factors to the RTC-like structures, in a manner similar to related RNA viruses. The lack of major redistribution does not negate the importance of the interaction of NP with actin in efficient viral replication through enabling internal transport.

It was however, the identification of eIF4A as an interacting partner that was the most intriguing prospect for further analysis. Previously, the NP of JUNV has been shown to exhibit a co-aligning distribution pattern with eIF4A in infected Vero cells. This observation led to the hypothesis that sites of JUNV NP protein density constituted replication complexes (Baird et al., 2012). The cytosolic sites of NP density that LUJV NP-EGFP exhibited suggest that NP is able to form such structures alone, and indicates that NP could be directly involved in the translation of viral mRNAs. Figures 35**B** and 37**D** demonstrate the presence of eIF4A within these LUJV NP generated structures which, in addition to identification via IP, and this observation posed a question as to why eIF4A and NP would interact. The recruitment of the translation initiation factor eIF4A into sites of NP concentration suggests NP has a role in facilitating the translation of host mRNAs. This is supported by the observation that eIF4A showed a similar distribution pattern with JUNV NP in JUNV infected cells.

Due to this intriguing observation, proteins similar to eIF4A were thus chosen for further analysis, including other initiation factors, and ribosomal subunits. Initiation factors from the other major groups - eIF1, eIF2 and eIF3 - were thus chosen to identify whether NP recruited major elements of translation

complexes, rather than an individual protein, such as eIF4A. In addition to this, ribosomal subunits identified via MS were chosen for the same reason.

The findings in figures 35 and 37 support the theory that NP can redistribute the constituent elements of translational complexes into cytosolic structures, shown to be early endosomes (38**E**), in the absence of other viral factors. The presence of any one of these proteins does not itself increase confidence that LUJV NP might generate replication complexes. However, the identification of multiple cellular proteins within such structures indicates a role in infection not previously identified for the NP of OW arenaviruses, as it indicates an involvement of NP in facilitating translation through factor recruitment. Indeed, the ability of LUJV NP-EGFP to recruit such factors alone is evidence to implicate NP in having a tangible role in translation. To this end, it seems reasonable to suggest that LUJV NP-EGFP-induced structures are de facto RTCs, mimicking those observed during JUNV infection.

The specific factors identified as potential partners are of interest also, with implications for understanding the potential role of NP in translation. The initial action during translation of mRNA is the formation of eIF2-GTP-tRNA^{MET}, which is subsequently guided into forming the 43S PIC by eIF1A. eIF1A then catalyses the association of the initiator tRNA^{MET} complex with the 40S subunit, along with the 40S subunit inhibitor protein eIF3 (Chaudhuri et al., 1997; Hinnebusch, 2006). Together, these proteins prime the 40S subunit, forming the 43S PIC. eIF4A, bound to eIF4G, is subsequently responsible for the unwinding of any 5' secondary structures in order to facilitate mRNA import into the 43S PIC. Without this, recognition of the initiator AUG is not possible (Jackson et al., 2010). The utilisation of these factors by NP suggests that NP modulates translation through relocalisation of multiple initiation factors in order to facilitate the translation of viral mRNAs. The apparent interaction of NP with these factors – direct or indirect – suggests that its ability to recruit these proteins is a necessary step for the virus to replicate. In the absence of these 43S PIC proteins, translation would not occur (Jackson et al., 2010).

These identified factors, along with other eIFs, thus facilitate the translation of mRNAs into polypeptide strands. The absence of eIF2G from sites of NP concentration was surprising given other observations, and could indicate that NP could substitute the function of eIF2G as a component of the pre-initiation complex (PIC). The role of the eIF2 group is considered to be a rate-limiting step in translation, so it is conceivable that a viral protein would have evolved to increase the efficiency of PIC formation (Kimball, 1999). Indeed, the absence of eIF2 in certain stages of poliovirus infection illustrates that the presence of eIF2 is not essential for translation to occur (Welnowska et al., 2011). Given the conflicting data observed, with eIF2G IPd via GFP-Trap and its subsequent absence from RTC-like structures, further investigations are warranted.

If such structures are utilised by viruses to act as RTCs, then the utilisation of specific subcellular organelles by NP should correspond with known components of the viral life cycle. Identifying subcellular organelles utilised by NP to form RTCs produced two interesting findings. The presence of NUP98, a marker for nuclear pore complexes (NPC), at the site of NP concentration could help facilitate the entry of certain factors required for replication. The use of components of the nuclear pore complex could help facilitate entry of ribosomal proteins, tRNA and other translation mediating proteins into RTCs (Christie et al., 2015). This was unexpected, given the lack of NPC proteins identified via the SILAC-based MS, but could reveal a novel role of NP through further investigation.

The localisation of NP within early endosomes was also encouraging. Arenavirus paerticles are known to fuse with cellular membranes within endosomes indicates that NP can independently recruit multiple factors to a specific organelle, which has already been implicated in the virus life cycle (Tani et al., 2014; York et al., 2008). The utilisation of endosomes by LUJV NP to form what could prove to be RTCs in a similar manner as those observed by Baird *et al.* (2012) is evidence that NP is able to modulate the internal cellular environment to facilitate replication. Specifically, the ability of

NP to recruit a number of proteins related to translation cannot be viewed as circumstantial, and indicates direct involvement in facilitating translation. Utilising endosomes as a replication centre would isolate viral replication strategies from several internal surveillance mechanisms. Further investigation, as discussed below in chapters 5 and 6, is required to establish whether NP facilitation of translation is direct, or indirect, or indeed both.

In conclusion, LUJV NP-EGFP is able to recruit multiple members of the translation initiation and subsequent elongation complexes into early endosomes, forming RTCs independently of other viral proteins. Given the lack of therapeutics for these viruses, if such a role could be identified in multiple *Arenaviridae* family members, novel therapeutic strategies could be identified.

5. LCMV infection, purification and analysis of viral proteome

5.1. Introduction

LCMV is an Old World arenavirus, with a host reservoir organism of the common house mouse, *Mus musculus* (Childs et al., 1992; Knust et al., 2014). The presence of LCMV within this host and the frequent co-habitation of humans and *M. musculus* means the likelihood of LCMV transmission to humans is relatively high. Even though LCMV does not induce clinical signs in otherwise healthy adults, LCMV is an important global health concern, especially among the immunocompromised where it can cause cerebral complications and is often fatal (Wright & Fishman, 2014). In rare cases of LCMV transmission through solid organ or tissue transplantation, mortality rates can be 100%, depending on the donated tissue and stage of infection within the donor (Macneil et al., 2012; Jamieson et al., 2006; Fischer et al., 2006). LCMV also poses a significant risk to expectant mothers due to the mild immunosuppression associated with pregnancy, their unborn child, and during post-natal infections. (Barton et al., 2002; Jamieson et al., 2006; Noonan et al., 1979; Attar, 2016). Infection in these developmental stages has been indicated in substantial developmental abnormalities in infants (Jamieson et al., 2006; Barton et al., 2002). As with all other arenaviruses, therapeutic options are limited, with therapy restricted to supportive therapy to combat symptoms as they develop (CDC, 2015). Unlike instances of LASV infection, ribavirin is not routinely used clinically to treat LCMV infection; partly due to the rarity of cases and its contraindication during pregnancy and in children. A campaign for greater awareness of LCMV in paediatric healthcare centres is growing, with calls for it to be included in the 'TORCHS' group of routinely tested infections in infants and expectant mothers (Bonthius, 2012).

LCMV is of significant research interest in itself, and also holds additional research value in that it represents an effective model for studying immune

responses to viral infection in general, with historical implications in the discovery of the major histocompatibility complex restriction and T-cell memory (Doherty & Zinkernagel, 1975; Zinkernagel & Doherty, 1974; Lau et al., 1994). It is also a model for studying the molecular and cellular biology of arenaviruses, often being described as the prototypic arenavirus (de la Torre, 2009). The OW LCMV is particularly valuable in its role as a surrogate of the HF-causing viruses LASV and LUJV, due to its ability to be propagated under less restrictive biological containment conditions. However, the unique cerebral tropism of LCMV compared to HF arenaviral infections mean that findings regarding the pathogenesis of LCMV and disease progression may not be wholly representative of other arenaviral diseases. It is therefore crucial that studies utilising LCMV in this manner focus on common features between LCMV and the HF-causing virus in question, in order to ensure LCMV can act as an effective model of these diseases.

Wild-type LCMV is classified as an ACDP hazard level 3 pathogen, and as such, its propagation requires containment level 3 (CL3) facilities, this presents a significant barrier to research progress. However, an attenuated strain of LCMV known as Armstrong is routinely used in research laboratories due to its classification as an ACDP CL2 pathogen, allowing fewer containment restrictions. The Armstrong strain is originally a splenic clone from an infected mouse; it causes an acute infection which can normally be cleared within a few days in healthy mice (Chiller & Oldstone, 1984). A subtype of this strain, termed Armstrong Clone 13, was again a splenic isolate, but was isolated from a persistently infected mouse (Chiller & Oldstone, 1984; Wherry et al., 2003). Two single nucleotide changes were identified as being responsible for chronic vs acute infection outcome in mice; with a mutation in GP, F260L, and the polymerase L K1079Q responsible for the persistent form (Matloubian et al., 1990).

In terms of arenavirus molecular biology, one long-standing observation that remains unexplained is the presence of ribosome-like structures within the virions of multiple arenavirus species (Pedersen & Konigshofer,

1976)(Murphy & Whitfield, 1975). The reason for their inclusion and/or their role in the arenavirus life cycle is poorly understood. There is evidence for NW arenaviral NPs interacting with components of the translation initiation complex as well as ribosomal proteins (Baird et al., 2012), consistent with the presence of ribosome-like structures within the corresponding infectious virions. In addition, our findings presented in the previous chapter suggest that the recently discovered OW LUJV NP-EGFP is able to interact with a variety of cellular translation-associated proteins, and re-distribute them within cells. Taken together, these data suggest that robust and important interactions may exist between arenaviral components and the cellular translational machinery at multiple stages of the arenavirus life cycle. We reasoned that the identification of translation-associated proteins within infectious LCMV-Arm particles may help to both better characterise the large ribosome-like complexes observable by EM and also provide insight as to why they might be incorporated within virions. In addition, the availability of LCMV-Arm presented us with an opportunity to examine whether the association between NP and the translational machinery occurred during a productive virus infection of living cells.

It was necessary to assess whether LCMV-Arm NP was functionally analogous to LUJV NP-EGFP in terms of its ability to associate with translation machinery components. The association of LCMV-Arm NP with the identified interacting partners of LUJV NP (chapter 4), during both the intracellular and extracellular phases of the viral life cycle. The abundant presence of specific cellular proteins within particles, or their association with NP would be suggestive of a potential role in LCMV infection. In addition to this, the knowledge that differing members of the OW clade of arenaviral infections exhibit similar and essential intracellular interactions could provide avenues for pan-species therapeutic targeting.

Finally, experiments described in this chapter aimed to identify whether LCMV-Arm NP played a crucial role in the selective translation of viral mRNAs over those of the host. If LCMV NP appears to mimic the function of LUJV NP,

it may be possible to infer that certain elements of their fundamental interactions remain conserved.

5.2. Results

5.2.1. Infection of cells

5.2.1.1. Virus quantification via plaque assay

Prior to performing comparative studies between LCMV-Arm NP and LUJV NP-EGFP, it was necessary to establish a means to quantify infectious LCMV-Arm, which would allow quantitative measurement of both virus stocks as well as provide a quantitative assessment of virus growth and viability. This was achieved by performing a plaque assay in which Vero cells were infected with LCMV-Arm, overlaid with CMC-DMEM and incubated for 6 days, after which plaques were visualised via crystal violet staining. Figure 39 shows plaques identified, and subsequently used to quantify virus.
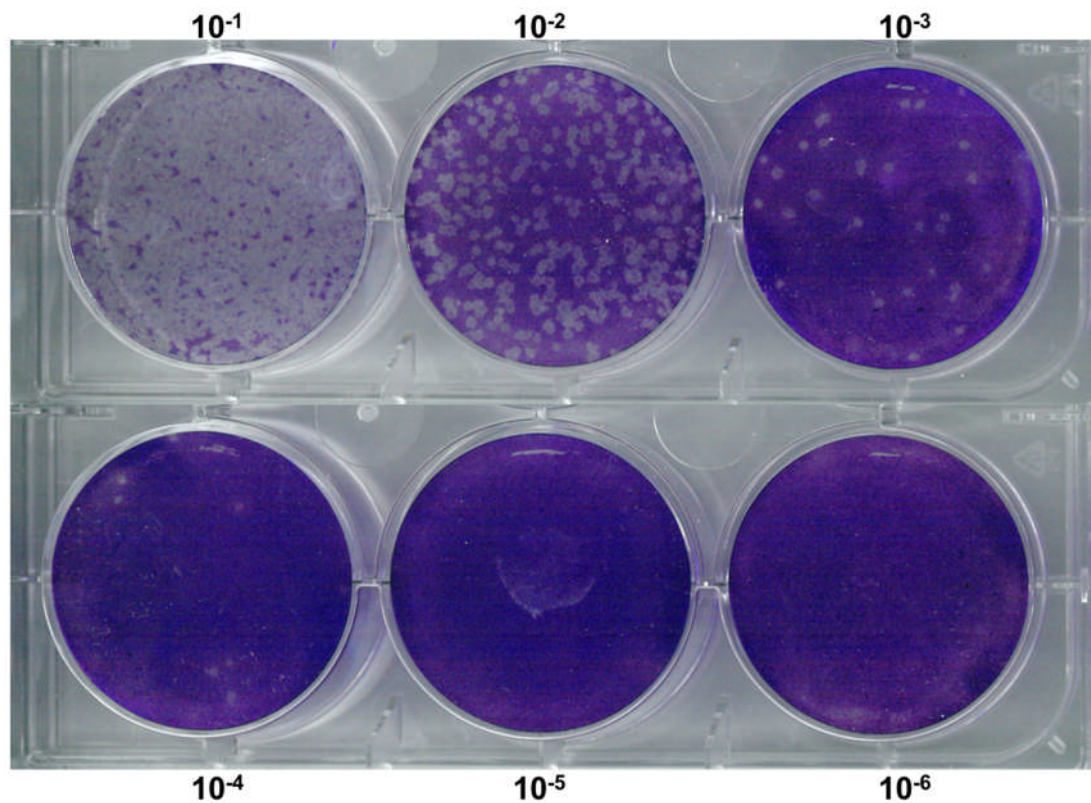
10^{-3} = 26 Plaques

Figure 39. **Quantification of LCMV titre via plaque assay**

Based on the number of accurately countable plaques in the 10^{-3} dilution, the virus titre could be quantified, with the titre found to be 1.3×10^5 PFU/ml, shown in the calculation below. This titre can then be used to infect cells at an appropriate MOI in subsequent analysis.

$$\text{Virus Titre} = \frac{26}{1 \times 10^{-3} \times 0.2}$$

$$1.3 \times 10^5 = \frac{26}{1 \times 10^{-3} \times 0.2}$$

5.2.2. Purification of virus

After determining viral stock concentration, it was necessary to amplify and purify LCMV-Arm in order to analyse the LCMV-Arm virion proteome. BHK21 cells were infected at an MOI of 0.001 and 6 days post infection the growth media was collected, with the LCMV-Arm purification protocol followed as in 2.2.6. Briefly, released virus was PEG precipitated overnight from clarified media, and then applied to the top of a 5-30% iodixanol gradient prior to ultracentrifugation at 200,000x*g*. Following centrifugation, a total of 25 fractions were collected and analysed via silver staining. As shown in figure 40, fraction 24 exhibited the highest concentration of protein - corresponding to LCMV particles. In addition, the distinct band within fraction 24 corresponds in approximate size to the 11 kDa Z protein, observable towards the bottom of the gel, corresponding to the arrow, in figure 40. This fraction was thus selected for analysis by LC-MS/MS, to determine viral proteome from the collected fraction. Prior to analysis, infectivity was confirmed via immunofluorescence microscopy, described below.
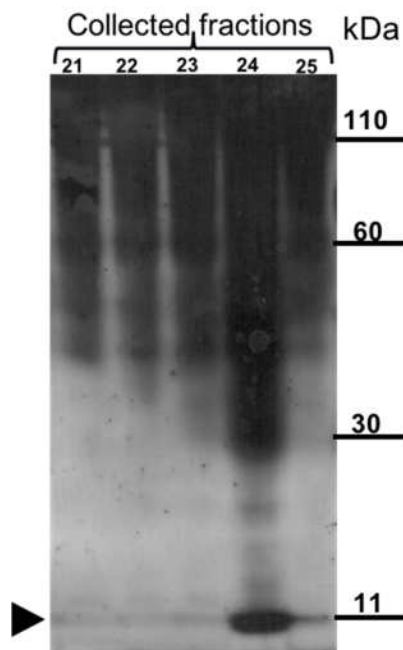
Silver stain of collected 5-30% iodixanol gradient fractions separated by SDS-PAGE after LCMV purification. Numbers above lanes indicate the numbered fraction collected after ultracentrifugation, with 25 being the densest fraction collected.

Figure 40.   **Iodixanol gradient purification of LCMV-Arm**

5.2.3. Confirmation of infectivity of iodixanol purified LCMV-Arm

As described in the previous section, LCMV-Arm was collected from BHK21 cell supernatant and purified by iodixanol gradient centrifugation. In order to determine whether the purification procedure yielded infectious LCMV-Arm, we determined whether these collected fractions contained any infectious particles by indirect immunofluorescence confocal microscopy using a monoclonal anti-LCMV antibody specific for the NP of LCMV. Aliquots of $2\mu l$ were taken from fractions 21-24, diluted in DMEM, and used to infect BHK21 cells seeded on cover slips. Control samples were mock-infected with fraction 1 from the gradient, in order to show that no viral particles were spread throughout the gradient. At 24 hrs post infection, cells were fixed, permeabilised and stained for the presence of LCMV NP. As shown in figure 41 the detection of abundant NP in cells showed all fractions analysed by IF contained infectious particles, with the fraction 24 showing the most abundant staining. The lack of NP staining in the mock infected sample indicated that

viral particles were not spread throughout the gradient. In addition to confirming infectivity of the purified gradient samples, LCMV-Arm was also found to generate cytosolic structures similar to those observed under LUJV NP expression and JUNV infection, with NP staining showing discreet cytosolic puncta throughout the cytosol.
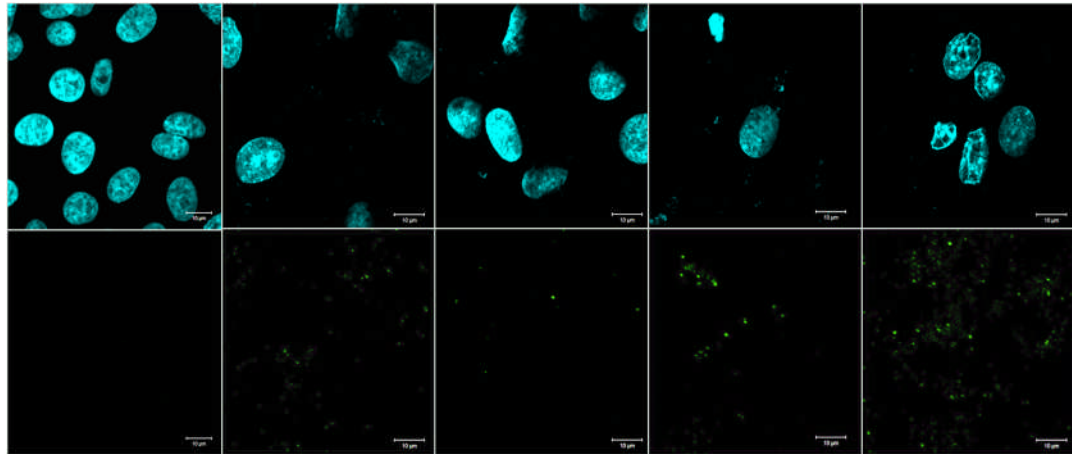
Immunofluorescence images of BHK21 cells infected with LCMV isolated from a 5-30% iodixanol gradient. In all images series, cyan indicates DAPI stain, green indicates LCMV-Arm NP, stained with LCMV NP antibody; scale bar indicates 10µm. The identifying number for each image series corresponds with the collected gradient fraction. Mock infected utilised fraction 1 in order to confirm infectious particles were not present throughout the gradient.

Figure 41.   Immunofluorescence analysis of purified LCMV-Arm infectivity

5.2.4. LC-MS/MS analysis of viral proteome

In the previous sections, infectious LCMV-Arm was purified from BHK21 cells, via an iodixanol gradient. In order to identify cellular proteins within infectious LCMV-Arm particles, MS analysis of fraction 24, which contained the highest level of infectious virus, was performed by Dr Stuart Armstrong at the University of Liverpool. This analysis was intended to provide an initial library of virion-associated protein identities, which would guide further examination of the LCMV-Arm proteome through immunological validation. The fraction containing LCMV-Arm was inactivated, denatured and processed for MS analysis. After Mascot analysis (as discussed in 4.1.1), a total of 44 unique proteins were identified within LCMV particles, and a selection of these and

their confidence values are shown in table 5, with the complete list available in Appendix III.

Accession Number	Confidence	Identifed Protein
P09992	1271.66	LCMV Arm NP
Q711N9	556.35	Actin
P09991	200.2	LCMV Arm GP
P86221	185.54	beta-Tubulin
P86204	152.6	HSP70 2
E2GMU8	72.23	eEF1a
P86237	49.05	HSP70 1
P18541	36.44	LCMV Arm Z

Table showing the MS identification of selected proteins from an iodixanol gradient fraction known to contain infectious LCMV particles. Analysis was performed by Dr Stuart Armstrong at the University of Liverpool. A complete list is available in Appendix III

Table 5 Identification of proteins within LCMV particles

The MS analysis detected three of the four viral proteins, namely NP, GP and Z, which is consistent with their known structural roles in assembled virions and the known infectivity of the virus-containing fraction 24. The confidence levels of these three structural proteins is high, with that for NP being the highest of all detected proteins. The L protein was not identified and this may be due to the known low abundance of this protein, possibly with only one copy of L associated with each of the two RNA segments. Of the cellular proteins detected, it was noticeable that the majority could be grouped into one of three functional classes – cytoskeletal proteins, cellular chaperones, and proteins associated with translation machinery. Of those identified, some were also described in the previous chapter as interacting with LUJV NP-EGFP, such as HSP70 and Actin.

Nevertheless, it was noteworthy that some of these previously identified LUJV NP-EGFP interacting partners – such as eIF4A, RPL10 and RPS19 – were not detected by the LCMV MS analysis. However, the absence of proteins – as discussed – does not preclude their presence within particles, as evidenced by the lack of L from the MS dataset. The lack of certain proteins from the data is most likely due to abundance, with there being only two copies

of L per particle. In addition, the annotation of the *Mesocricetus auratus* genome is also limited, therefore certain peptide sequences may not be characterised, leading to the poor annotation of identified peptides obtained from LCMV particles generated in BHK21 cells.

5.2.5. Validation of viral proteome

The results presented in the previous section revealed a listing of viral and cellular proteins that were identified by MS analysis within purified LCMV particles. In order to independently validate the presence and identities of these proteins, LCMV-Arm was again purified by iodixanol gradient centrifugation, and was then probed for the presence of cellular proteins using specific antibodies. In order to achieve better LCMV purification, the LCMV purification strategy was modified by the use of a 15-40% iodixanol gradient in place of the 5-30% used previously. The rationale behind this change was to position the LCMV-Arm containing fraction nearer to the middle of the gradient, rather than at its base – as had been achieved previously. It was reasoned this approach may better separate LCMV from cellular components, that sediment rapidly towards the bottom of the gradient.

5.2.5.1. Immunoblotting
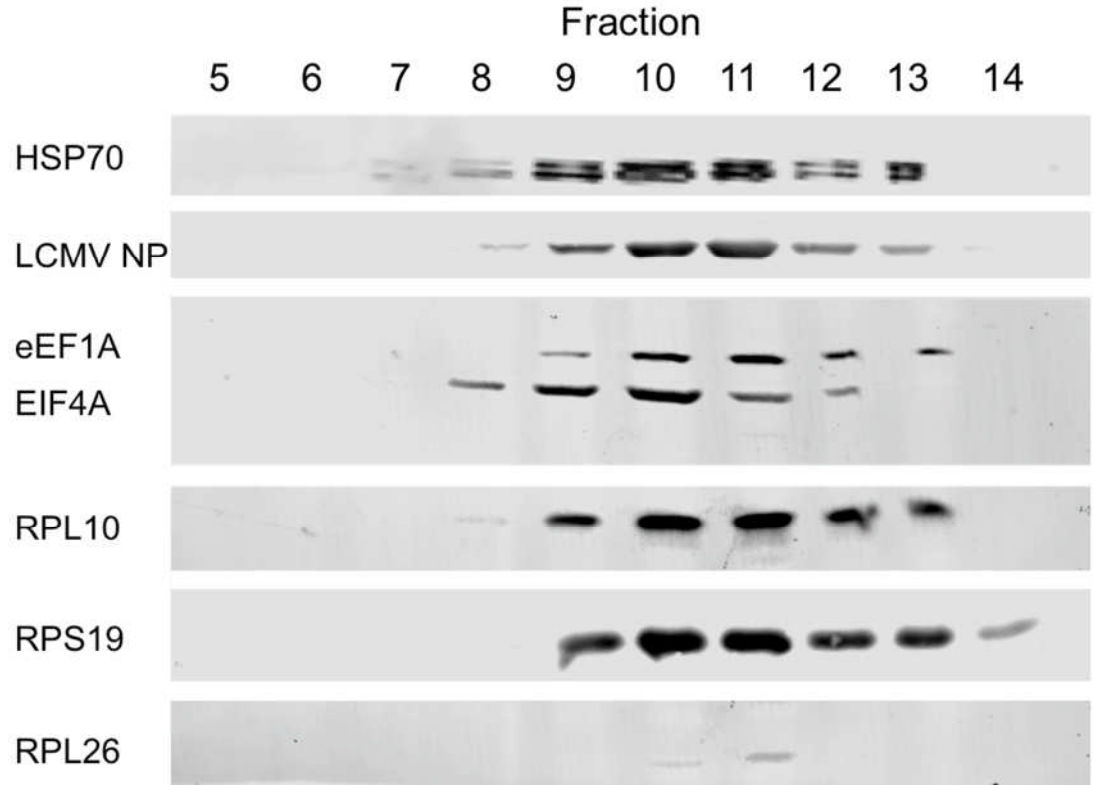
Western blot analysis of LCMV-Arm particles, separated on a 15-40% iodixanol gradient. Lane numbers correspond with collected gradient fractions. In all cases, relevant antiseruk was used. In the case of LCMV-Arm NP, LUJV NP antiserum was used.

Figure 42.   WB analysis of purified LCMV-Arm proteome

As described above, the presence of infectious virus could be tracked by identification of NP by western blotting within aspirated fractions taken from the gradient following centrifugation, with LCMV NP visualised through the use of anti-LUJV NP antibodies, as described in 3.2.6.1. After loading onto the 15-40% iodixanol gradient, the LCMV NP was predominantly concentrated into fractions 9, 10 and 11 (figure 42), and associated LCMV infectivity of fraction 11 was confirmed by IF using the anti-LUJV NP antibody, described in section 3.2.4.

Next, further western blot analysis of fractions was performed using antibodies specific for a variety of cellular proteins identified by the MS analysis. The cellular proteins chosen for immunological validation satisfied at least one of three criteria: firstly, that they were identified by MS analysis with high

confidence; secondly, they were known to be associated with, or comprise components of the cellular translation machinery; or thirdly, they had previously been identified as interacting partners with related viruses, and thus had a plausible functional role in the context of the LCMV life cycle. The cellular chaperone HSP70 satisfied two of these criteria, and thus HSP70-specific antisera was used to probe fractions. The identification of a 70kDa band in these peak fractions confirmed the presence of HSP70 in infectious LCMV particles, with the characteristic double band expected, given the antibody specificity to not only HSP70, but also HSC70 – as described in section 4.3.2. Cellular proteins eIF4A and eEF1A were also prime candidates for immunological validation due to their association with the cellular translation machinery as well as being previously reported to associate with LUJV NP, and other arenaviruses. The presence of these proteins within fractions 8 through 13 was assessed by western blotting using specific antisera, which again showed the consistent presence of both eIF4A and eEF1A in all virus-containing gradient fractions. Finally, the presence of ribosomal proteins RPL10, L26 and S19 was also tested, and western blotting using specific antisera identified bands of the corresponding molecular weight. It is noteworthy that RPL26 is present within peak virus fraction 11 in lower abundance that either RPL10 or RPS19, which may raise doubt over the specificity of its incorporation into virions. However, the distribution of RPL26 solely within peak fractions 10 and 11, rather than in other fractions that do not possess any LCMV-Arm NP, provides evidence that it is not a non-specific contaminant.

Taken together, these validations confirm the presence of multiple cellular proteins within LCMV virions. These data demonstrate that LCMV-Arm proteins interact with similar proteins that were previously identified via LUJV NP-EGFP, indicating that the effect is not limited to only LUJV.

5.2.5.2. Immunofluorescence imaging

In the previous section, cellular factors associated with translational initiation and elongation complexes, as well as ribosome subunit proteins, were

identified within LCMV particles. Based on these observations suggested that LCMV requires components of the cellular translation machinery for its efficient multiplication inside cells. This hypothesis would predict that during infection with LCMV, these cellular components would be recruited to sites where virus multiplication and gene expression were on-going and thus virus encoded proteins were abundant. One such location is the RTCs that are proposed to constitute the sites for arenaviral RNA synthesis. In order to test this prediction, the cellular location of multiple components of the cellular translation machinery was examined in LCMV-infected cells compared to mock-infected cells to determine whether redistribution to RTCs had occurred. Based on the findings of the MS analysis of LCMV virions, the subsequent immunological validation, as well as the interactome analysis of LUJV NP, a panel of antibodies specific for cellular translation components was assembled comprising; eIF1A, eIF4E-PHOS (S209), RPL10a, RPL26 and RPS11. In addition, cells were counterstained with the generated LUJV NP antibody (from section 3.2.4), this was due to the poor and expensive commercial LCMV antibody. eIF4E-PHOS has been implicated in enhanced 5' cap binding and tumour development (Furic et al., 2010). The potential use of eIF4E-PHOS in viral infection would allow for increased translational efficiency and therefore increased copies of viral proteins within an infected cell.

5.2.5.3. Re-localisation of identified partners to RTC-like structures during LCMV-Arm infection

In order to establish whether localisation patterns of specific cellular proteins were altered in response to LCMV-Arm infection, it was necessary to determine the relative distribution of these cellular proteins against the distribution of LCMV-Arm NP, visualised by both LCMV antibody and anti-LCMV NP antibody (3.2.6.1). A549 cells were infected with fraction 11 from the above (5.2.5.1) purification of LCMV-Arm with mock-infected cells using the non-infectious fraction 1 as above in figure 41. Figure 43 shows the relative distribution of a number of cellular factors known to interact with LUJV

NP-EGFP – with some identified within LCMV particles. The colocalisation observed between LCMV NP and eIF4E-PHOS in 43**B** indicates an interaction with the translation enhancing isoform of the cap binding protein. In addition, several other translation-associated proteins and ribosomal subunits such as eIF1A in 43**A**; RPL10 in 43**C**; RPL26 in 43**D**, and RPS11 in 43**E** were identified to co-localise alongside LCMV NP in cytoplasmic puncta. As shown for LUJV NP, ribosomal proteins RPS11, RPL10 and RPL26 all appear to show a degree of co-localisation with LCMV NP. This observation increases confidence in the proposal that LCMV NP might interact with these host-cell factors and subvert their usual function during virus multiplication and relocate them to sites of viral replication, as well as direct the packaging of these factors into the virion. As shown in figure 38, puncta of RTC-like structures are prominent in LCMV-Arm infected cells (figure 41), consistent with the identified use of early endosomes by LUJV NP-EGFP. In addition, these LCMV-Arm induced punctate structures resemble those observed with LUJV NP-EGFP expression and also detected in the course of JUNV infection of Vero cells (Baird et al., 2012).
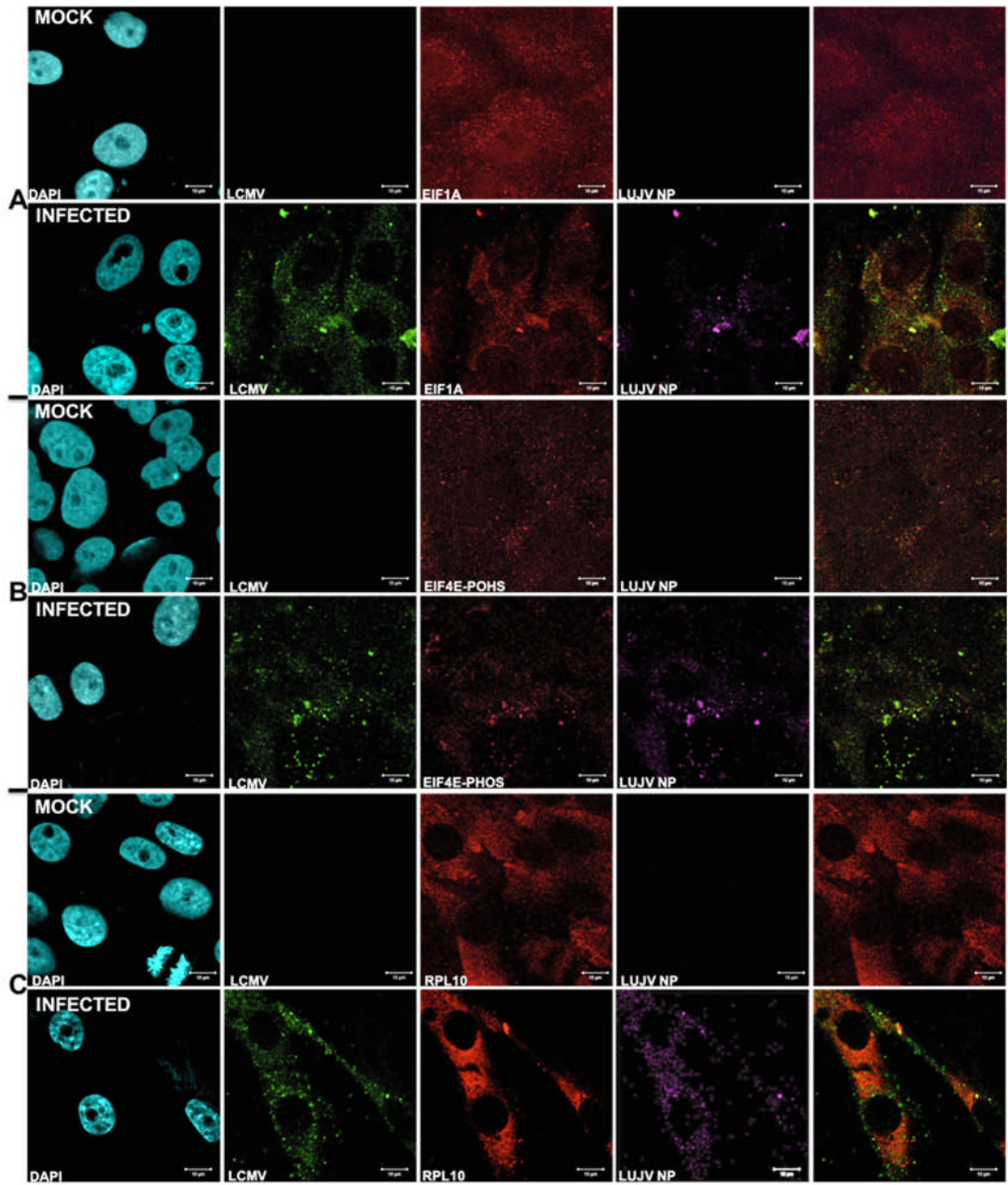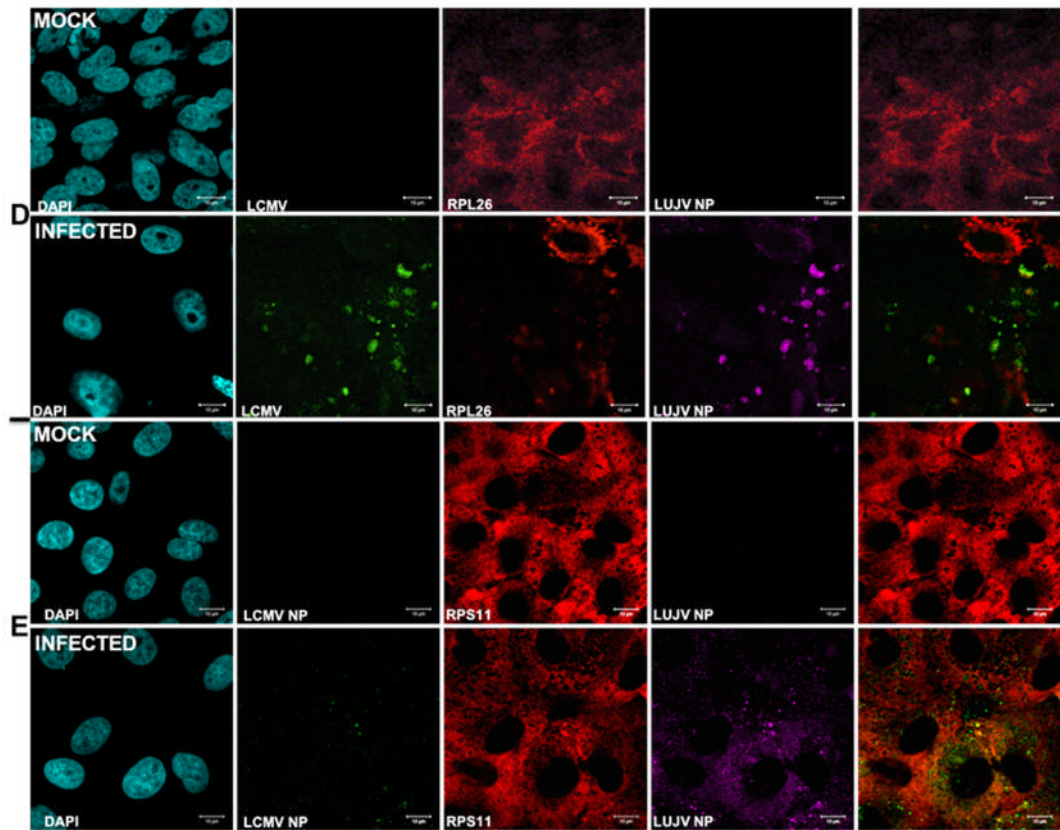
Immunofluorescence analysis of the localisation of cellular proteins against mock-infected control Vero cells and LCMV NP Infected Vero cells. **A** shows the localisation of EIF1A against LCMV NP staining and LUJV NP co-stain – reactive to LCMV NP; **B** shows the localisation of EIF4E PHOSPHO; **C** shows the localisation of RPL10; **D** shows the localisation of RPL26; **E** shows the localisation of RPS11; In **A-E**, cyan indicates DAPI staining of DNA; green indicates LCMV NP staining; red indicates the indicated cellular protein; magenta indicates LUJV NP antisera staining of LCMV NP and the final image in each series indicates the merge of red and green channels – in **E** the merge is between red and magenta (false coloured green), due to the poor LCMV NP staining.

Figure 43. **Comparative localisation of LCMV-Arm alongside identified interacting partners of LUJV NP-EGFP**

This section shows the relocalisation of a number cellular translation factors into punctate structures, closely resembling those theorised to constitute replication complexes – as observed during JUNV infection and LUJV NP-EGFP expression.

5.3. <u>Discussion</u>

The data discussed above provide evidence that multiple cellular translation-associated proteins are packaged within infections viral particles, including

eIF4A, eEF1A and constituent protein components of both the 40S and 60S ribosomal subunits. In addition, a subset of these proteins, as well as other translation machinery previously identified as being associated with the LUJV NP-EGFP, were identified as localising with LCMV-Arm NP within infected cells. In addition, the IP of LUJV NP alongside eIF2G, eIF4A and ribosomal subunit proteins shown in chapter 4 suggested an interaction between NP and translation complex proteins. Taken together, these observations provide compelling evidence for an interaction between arenavirus NPs and components of the cellular translation machinery in both intra- and extracellular phases of the virus life cycle.

Evidence of translation associated proteins being sequestered into proposed sites of RNA replication and viral mRNA translation is not unexpected, given the need for viral RNA species to be protected from PRRs within the cell in order to prevent detection from surveillance mechanisms. The knowledge that LUJV NP alone can re-localise factors into de-facto RTCs suggests that the NP of both LUJV and LCMV-Arm, besides being involved in RNA binding, to form an RNP to increase RNA stability, might play a crucial role in viral replication through a direct involvement in translation.

The number of translation-associated proteins identified within the LCMV-Arm virions is suggestive of an important role during the initial stages of virus multiplication, shortly following virus entry. Once RNPs are exposed to the cytosol after membrane fusion, the rapid transcription and subsequent translation of viral mRNAs could be achieved through the delivery of host-cell proteins within the virion. This would then allow for these components to perform a critical function immediately upon virus entry, without the need for their recruitment in the initial stages of infection. It does not seem reasonable to suggest that such a multitude of functionally-related cellular proteins would be included within viral particles by chance, especially given previous EM-based evidence that ribosome like-structures are present within arenavirus particles (Murphy & Whitfield, 1975).

The presence of eEF1A within the virion (figure 42), in addition to the previously discussed initiation factor eIF4A, is intriguing. The eEF1 complex is responsible for the delivery of aminoacylated-tRNAs (aa-tRNAs) into the A site of a ribosome during the elongation phase of polypeptide synthesis (Mateyak & Kinzy, 2010). If eEF1A is present within LCMV particles is associated with ribosomes, these are likely to be assembled in a translationally ready state, and may be competent to initiate translation immediately following entry of the virion contents into the infected host cell through membrane fusion within the maturing endosome. It is possible that there is a requirement of coupled transcription and translation, whereby an active ribosome allows for the continual strand elongation of nascent RNA from the viral polymerase; this coupling has been demonstrated for the transcription of Bunyamweravirus (BUNV) RNAs (Barr, 2007).

As discussed above, the roles of the eIF4F complex proteins are well established (Oberer et al., 2005; Linero et al., 2013; Furic et al., 2010; Lefebvre et al., 2006). The presence of the eIF4A RNA helicase within the viral particle (figure 42) is also intriguing. If as above for eEF1A, the eIF4A within virions is associated with ribosomes, these are also likely to be in a functional state that can perform translation immediately upon entry into a host cell. Indeed, it is the presence of not just one of these proteins (eIF4A, eEF1A, RPL10, RPL26 and RPS19), but the combination, which appears to suggest that translation complexes are packaged within LCMV particles.

The presence of translation initiation and elongation factors, along with ribosomal subunit proteins within virions leads to the question as to why they are packaged. The general mechanism of arenaviral mRNA synthesis is known, with primary mRNA transcription requiring cap-snatching of 5' mRNA caps from host mRNAs in order to prime nascent mRNA synthesis. Critically, this initial transcription occurs on the packaged vRNA templates, however, it is difficult to see how a ribosome would be beneficial in this context. One possibility is that transcription and translation is coupled, in a manner similar to that of BUNV, where the translocation of the 40S ribosome complex

alongside the active polymerase is required to assist in the elongation of mRNA transcripts (Barr, 2007). If 40S translocation is blocked, transcription is prematurely terminated, leading to a loss of viral protein synthesis. This would explain the presence of ribosomes, seemingly packaged with their associated translation factors such as eIF4A and eEF1A.

This hypothesis would require that the transcription of mRNA from the arenaviral genomes is not possible without the presence of an elongating ribosome. The presence of eEF1A within the virion particle, alongside the initiation factor eIF4A, is evidence of a potential functional complex. However, the observations of Leung and Rawls (1977) indicated that ribosomes packaged within virions were not required for productive arenavirus infection. These findings were recorded over a 48-hour period, recording plaques formed after inactivation of temperature sensitive ribosomes (Leung & Rawls, 1977). These findings do not preclude the possibility that infection is still possible without packaged ribosomes, but that their presence results in initial enhancement of translation.

A second possible reason for the packaging of ribosomes within virions is that it represents an efficient mechanism to deliver ribosome complexes into cells simply to enhance initial translation. The distinct co-alignment of NP with eIF4E-PHOS within RTC-like structures in infected BHK21 cells is most interesting, given that phosphorylated eIF4E is known to increase the efficiency of cap-binding – and thus mRNA expression, and is known to have a role in tumour development (Furic et al., 2010; Yanagiya et al., 2012; Scheper & Proud, 2002). The presence of eIF4E-PHOS within RTCs during LCMV infection could be in part due to the known sequestration of eIF4E by Z, and its subsequent suppression of cellular translation (Campbell Dwyer et al., 2000; Kentsis et al., 2001). Z and NP could work in conjunction in order to establish complete RTCs, however, as LUJV NP-EGFP also exhibited co-localisation with eIF4E, it is likely that NP in some way utilises eIF4E for translational enhancement, prior to Z stimulating a shut-off of protein expression in a manner to stimulate viral replication (Volpon et al., 2010). The

143

documented enhancement of mRNA expression in the presence of eIF4E-PHOS, known to increase protein expression in rapidly multiplying tumour cells, could similarly play a crucial role in enhancing viral replication. If mRNAs can be expressed more rapidly, the innate cellular defence mechanisms may be overwhelmed before they can become fully active.

The lack of 'free' NP immediately after infection within cells should prevent the effective recruitment of translation complexes due to the known association of NP within RNPs - an association which is still required within cells (Iwasaki et al., 2015). To this end, for NP to be able to effectively recruit factors in order to facilitate an exponential increase in translation, to coincide with increases in mRNA availability, there must be an abundance of free NP. Therefore, there should be a virally encoded mechanism in order to allow for NP to recruit cellular factors to the sites of replication.

If virions are able to deliver ribosomes which can immediately initiate translation upon entry, the availability of NP within the cell would quickly rise, given the transcription of both NP and L mRNAs from their respective S and L genome segments. The role of NP in NF-κB modulation would be of greater effect if there was an abundance of free protein available within the cell quickly (Martínez-Sobrido et al., 2007; Rodrigo et al., 2012).

In conjunction with the delivery of translation associated proteins, the presence of the chaperone HSP70 within LCMV-Arm particles could indicate infectious virus particles contain multiple host-cell proteins to aid in different aspects of protein expression, not simply limited to translation. The delivery of chaperone proteins alongside other factors could facilitate faster protein folding – again mitigating potential rate-limiting steps (Dalziel et al., 2014). HSP70 could also allow for NP to remain in a monomeric state, preventing it from multimerising and therefore facilitate NP in performing its roles beyond RNP formation.

The delivery of translation complex proteins and chaperone proteins into cells along with the genome would allow for fast, efficient translation and expression upon cell entry. Without the delivery of such factors, it is

reasonable to assume infection would still be possible, but, the presence of such factors could increase the rate at which the infection is established within the cell. Figure 44 shows a working hypothesis for how these factors might be arranged within an infectious viral particle. The 5' cap pre-loaded in the L protein could extend, with the translation initiation complex closely associated, or potentially coupled, with the primed polymerase.
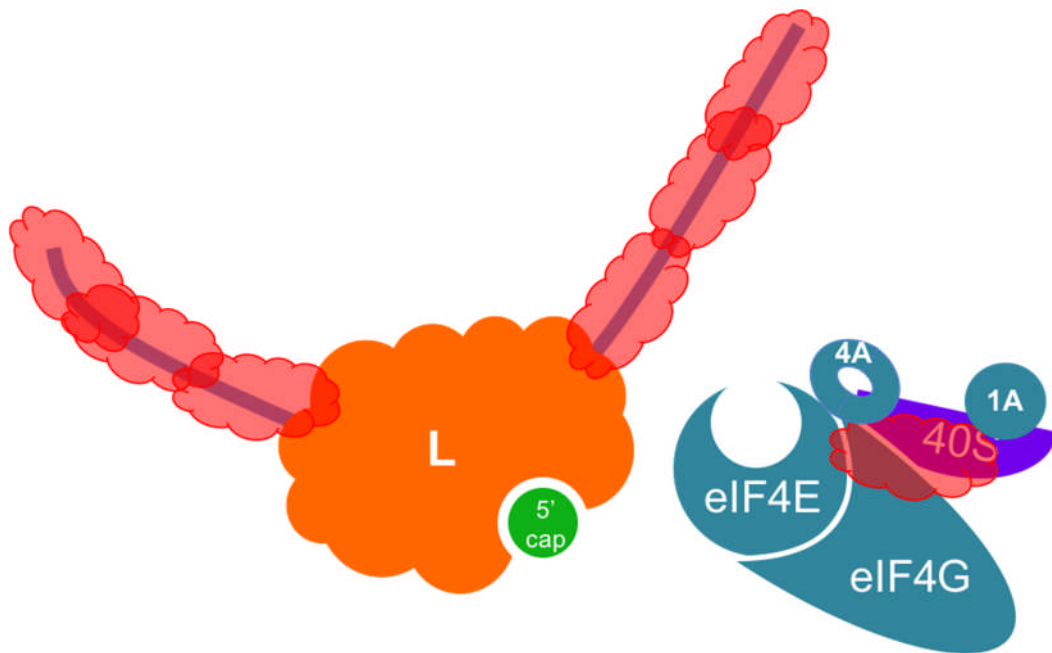
Diagram illustrating the proposed role of cellular proteins within LCMV particles. The presence of translation-associated proteins within LCMV particles indicates a role in infection. This proposed model indicates a 'pre-loading' of the viral particle in order to quickly facilitate infection. The purple, coated with NP (red) line indicates the viral genome associated in an RNP with NP.

Figure 44.   Diagram depicting cellular proteins included within LCMV-Arm particles

The arrangement shown in Figure 44 illustrates the complete 40S ribosomal sub-unit, rather than simply the identified proteins. As the individual roles of such proteins are not fully understood, the significance of their individual presence within particles is not possible to determine. Given that morphological analysis of LCMV and LASV has shown that there are ribosomal-like structures within particles, the presence of these sub-unit proteins points towards a full ribosome-like structure within infectious particles. This would coincide with the hypothesis that such factors are pre-

loaded within infectious particles in order to facilitate rapid translation upon cell entry. The apparent lower abundance of RPL26 within LCMV particles compared to the other subunits identified, namely RPS19 and RPL10, could be explained by a number of means. Most likely, specificity of the antiserum for RPL26 could be lower than that of RPS19 or RPL10 in the context of western blots.
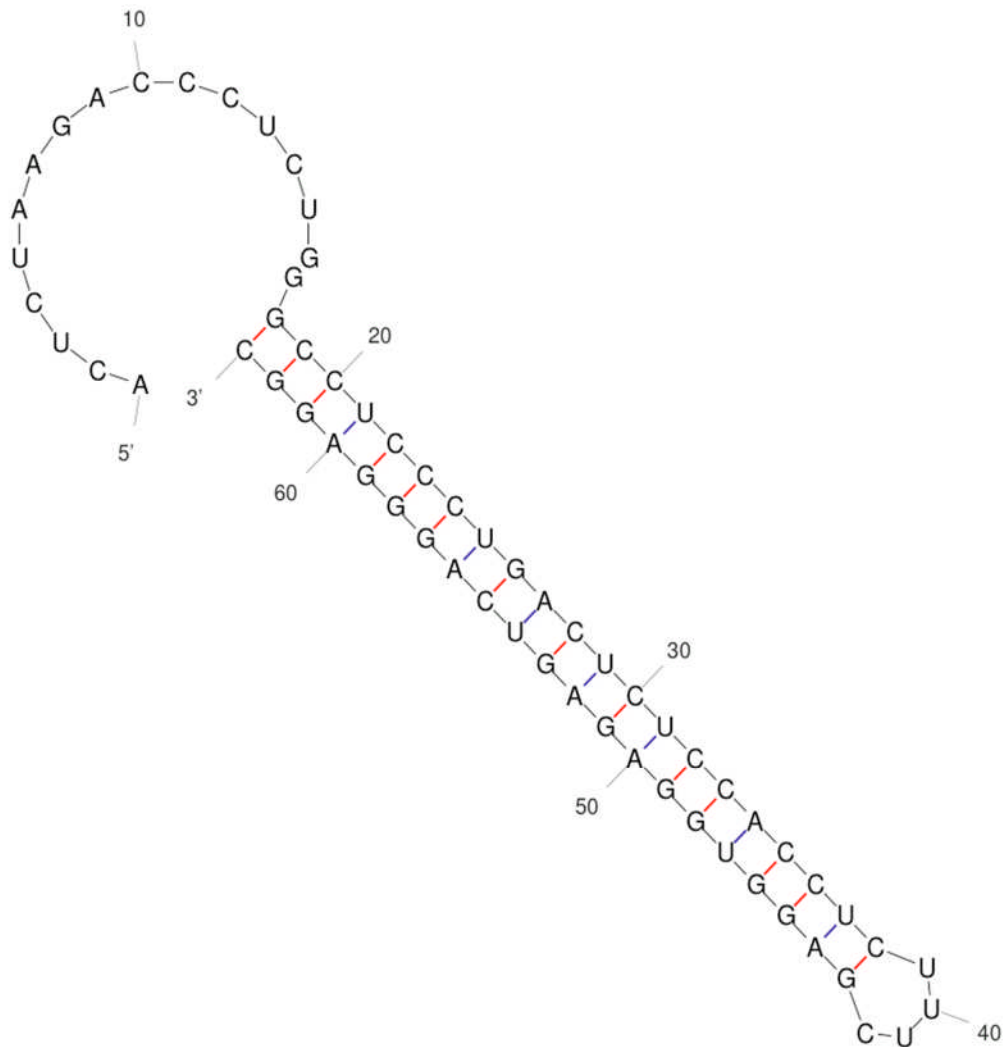
Diagram showing RNA hairpin structure, predicted to form from the 3' end of LCMV NP mRNAs (Meyer & Southern, 1994) using the mfold web server prediction software (Zuker, 2003).

Figure 45.   Predicted 3' hairpin of LCMV NP mRNA

Another question remains regarding how arenaviral mRNAs are translated efficiently. Arenaviral mRNAs are known not to contain a 3' poly(A) tail but instead contain a substantial, approximately 46 nucleotide, hairpin loop at the

3' end (figure 45) (Meyer and Southern 1993, Meyer and Southern 1994). Given its 21 base pair length, this could act as a PAMP for Mda5, RIG-I or PKR, thus triggering NF-κB signalling and a type I interferon response(Alexopoulou et al., 2001; Manche et al., 1992; Peisley et al., 2012). The identified roles of NP as an inhibitor of NF-κB signalling (Martinez-Sobrido *et al.* 2006, Pythoud *et al.* 2012, Rodrigo *et al.* 2012), it appears incongruous to expect a virus to trigger an innate immune response, merely to subsequently subvert it. The 3' hairpin structure, shown in figure 45, given its conservation, is likely to play a crucial role in a stage of the virus life cycle. NP has an already established role in transcription (Tortorici et al., 2001; Pinschewer et al., 2003), and the data shown in this, and the previous, chapter suggests that NP also plays an important role in mRNA translation through the recruitment and sequestration of translation initiation and elongation factors.

Certain components of mRNA translation complexes are essential, while others confer efficient translation. As discussed previously, efficient cap-binding and 43S PIC formation is an essential step in the initiation of translation (Jackson et al., 2010). The interaction of NP with multiple elements of this complex is already indicative that the NP of multiple arenaviruses is important in facilitating the formation of this complex around arenaviral mRNAs. A separate step that confers efficient translation is that of the circularisation of mRNA prior to initiation. This circularisation allows for efficient, stable complexes to form, and therefore can facilitate an increase in protein expression from any given mRNA molecule (Wells et al., 2012). Circularisation is aided by the mRNA poly(A) tail present on host transcripts, and the subsequent binding of this feature by PABP and its ensuing association with the eIF4F complex scaffold protein eIF4G (Prévôt et al., 2003). Given that there is no poly(A) tail on the viral mRNAs, the circularisation of mRNAs via PABP1 interaction with the poly(A) tail would not be possible.

Arenaviruses likely have developed an alternative strategy to enable efficient translation, and to preserve mRNA stability (Wells et al., 2012; Kahvejian et

al., 2012; Weill et al., 2012). Given the presence of multiple cellular factors within viral particles, and the interaction of LUJV NP-EGFP with multiple translation associated proteins shown above, the evidence suggests NP could facilitate efficient translation of mRNA.
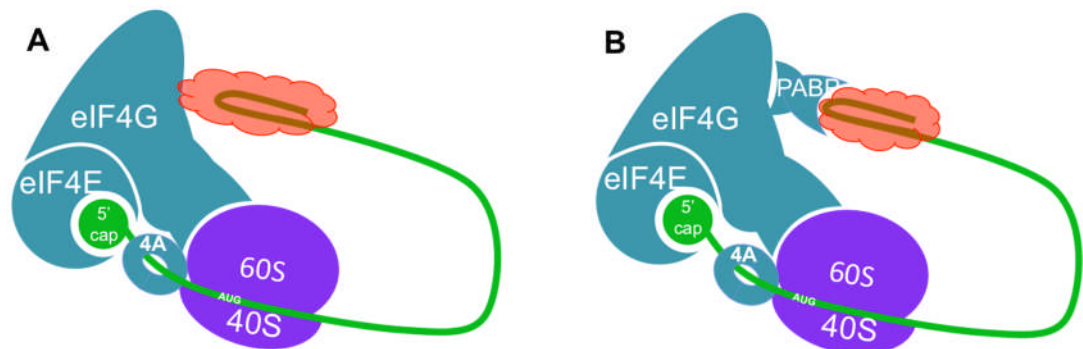
Diagram indicating two proposals for NP modulation of translation. **A:** Illustrates NP substituting PABP to bind eIF4G, facilitating the circularisation of mRNA. **B:** Illustrates NP binding to the 3' hairpin structure of arenaviral mRNAs, with mRNA circularisation facilitated by binding with PABP. Green colour shows the viral mRNA strand; red NP; teal cellular initiation factors; and purple showing ribosomal subunits.

Figure 46.   **Diagram depicting possible orientations of arenavirus mRNA circularisation via NP**

Detailed in figure 46 are two proposed models for how NP might facilitate circularisation and therefore efficient translation of viral mRNAs, through the identified interacting partners described above in addition to eIF4G. Figure 46**A** illustrates that NP might substitute the function of PABP, binding to the dsRNA hairpin and then binding with eIF4G. The affinity of NP for dsRNA is well established, with LASV NP known to have a strong affinity for dsRNA (Hastie et al., 2012; Jiang et al., 2013). While the exoribonuclease activity of multiple arenaviral NPs is well known (Huang et al., 2015; Jiang et al., 2013; Hastie et al., 2012), it seems evolutionarily illogical for a virus to encode, and constitutively express, a distinct PAMP with a major role of an encoded protein simply to remove it. To this end, there may be a second role for such a hairpin structure, and indeed the affinity of NP for it. The second hypothesis shown in 46**B**, is fundamentally similar, however rather than NP simply substituting the function of PABP, NP could bind to the hairpin structure and then subsequently PABP itself. PABP could then complete the circularisation through its interaction with eIF4G (Kahvejian et al., 2012).

As for the unknown function of particle-associated ribosomes, the exact mechanism of mRNA translation is unknown, with findings in this chapter and chapter 4 implicating NP in the efficient translation of mRNA though generation of replication complexes and the recruitment of translation factors. In the following chapter, the implications of th direct involvement of NP in mRNA translation is explored.

In conclusion, this chapter shows the purification of LCMV-Arm via iodixanol gradient purification, with such a procedure yielding infectious particles. Subsequent analysis identified multiple cellular proteins that are packaged within these viral particles – notably eIF4A, eEF1A and 40S and 60S ribosomal subunit proteins. This chapter also demonstrates that, just as for LUJV NP-EGFP, BHK21 cells infected with LCMV-Arm exhibit redistributed cellular proteins into RTCs, indicating that the observations seen for LUJV NP in chapter 4 are not limited to one OW arenaviral infection, but are more likely to be a common trait for this clade of infections. Finally, given the conclusions from this chapter, and chapter 4, two potential models for the role of NP in the translation of viral mRNAs are proposed. The identified interactions of LUJV NP-EGFP and LCMV-Arm NP point toward NP facilitating translation and directly enhancing the expression of viral proteins in order to allow for rapid infection establishment.

6 Functional investigation of the role of LCMV NP in translation of viral proteins

6.1 Introduction

The conclusions of the previous chapter led us to propose two potential models that describe the involvement of arenavirus NP in the process of arenaviral mRNA translation. These models have been constructed on the basis of four supporting pieces of evidence from ourselves and others: first, we have shown that arenavirus NP co-precipitates and co-localises in cells with several functionally-diverse components of the cellular translation machinery; second, the lack of 3' poly(A) tails on arenavirus mRNAs suggests that the canonical interaction between poly(A) binding protein (PABP) and the 3' poly(A) tail is likely replaced by an alternative protein-RNA interaction. Thirdly, the arenavirus NP has been previously shown to bind long RNA hairpin structures such as those found at the arenaviral mRNA 3' end; and finally, previous reports from others have identified interactions between other arenavirus NPs and components of the eIF4F cap binding complex, at the exclusion of PABP (Linero et al., 2013; Baird et al., 2012). These models, shown schematically in Figure 46**A&B,** suggest how arenavirus mRNA translation might proceed despite the lack of a poly(A) tail, through NP-mediated mRNA circularisation. According to these models, the arenavirus NP would facilitate the efficiency of mRNA translation, and consequently increase the yield of translated protein from viral mRNAs. In the context of a virus infection, both these models might reasonably be expected to result in preferential translation of viral mRNAs over those encoded by the host, thus mediating shutdown of host cell-specific protein synthesis.

As alluded to above, previous studies have suggested that arenavirus mRNA translation requires interplay between the virus-encoded NP, the 3' sequences of mRNAs, and components of the host cell translation machinery. JUNV NP has been shown to interact and exhibit colocalisation with the

proposed replication centres – RTCs – of eIF4G and eIF4A but not the cap binding protein eIF4E. It was proposed that NP acted as the cap-binding protein, substituting eIF4E within the eIF4F complex (Linero et al., 2013). Previous observations of arenavirus 5' cap binding also supported NP's role as a cap binding protein (Qi et al., 2010). Hantavirus NP has been implicated in the substitution of the entire eIF4F complex, and thus it is not without precedent that NP could substitute at least one of the eIF4 complex proteins (Mir & Panganiban, 2008). However, observations in chapter 5 suggested that the more efficient cap binding eIF4E-PHOS was present within RTCs during LCMV-Arm infection at greater levels than the unphosphorylated eIF4E. This indicates that NP could recruit the more efficient form into RTCs in order to increase translational efficiency – a divergent mechanism of eIF4F complex interaction to that proposed previously.

In addition, the work of others has identified mRNA 3' UTR structures in several other viruses, such as flaviviruses, bunyaviruses and rotaviruses. DENV translation initiation appears to revolve around the interaction between the 3' and 5' UTRs, whereby the conformation can bypass cap-dependent translation when eIF4E is depleted (Edgil et al., 2006). The strategy of rotavirus mRNA circularisation is similar to our arenavirus proposal (46**A**) in that an interaction between a virus-encoded protein and 3' RNA replaces the canonical interaction between PABP and the 3' poly (A) tail. For rotaviruses, this protein is NSP3, and it is able to interact directly with eIF4G (Vende et al., 2000) thus mediating mRNA circularization. Bunyamwera virus (BUNV) similarly generates mRNAs without poly(A) tails; instead they terminate with a conserved sequence that has the potential to form a strong stem loop structure, although no cellular or virus-encoded protein binding partner for this element has yet been identified. Interestingly, during BUNV infection, PABP is actively removed from the cytoplasm and imported into the nucleus likely providing or contributing towards an effective mechanism of host-cell translation shut-off by the protein NSs (Blakqori et al., 2009). In this scenario, the now abundant cytosolic translation factors are available for sequestration

by viral factors to enhance viral mRNA translation of poly(A) tail-less mRNAs. However, how BUNV mRNAs are circularised is still unknown.

A predicted consequence of these viral 3' mRNA secondary structures, unless disguised, is that they likely stimulate innate immune activation via cellular PAMP surveillance mechanisms, and would thus be an undesirable feature for a virus to generate during infection. However, the paradoxical conservation of structured 3' mRNA sequence elements in all known arenaviruses and many other RNA viruses as described above suggests these structures play important roles within the respective replication cycles. The benefit of virus fitness bestowed by these structures likely outweighs their inherent drawbacks, and contributing to this, it is likely that the ability of viral proteins to bind these structures in some way masks their dsRNA signatures, preventing their detection as a PAMP. In this chapter we examine the possibility that the viral component is NP.

In addition to the 3' secondary structure, and NP, a third viral component that could conceivably be involved in translational enhancement is the 5' mRNA un-translated region. While not containing any obvious secondary structural motifs, arenavirus mRNA 5' UTRs do possess conserved sequences, and their role in translation has not been established.

In order to test the roles of these three components – NP, 5' UTR and 3' UTR – in the enhancement of arenavirus mRNA translation, a system was developed that allowed the generation of an arenavirus mRNA that could be site-specifically modified, and its ability to be translated measured by detection and quantification of a reporter protein expressed from this mRNA. Previous findings by others have defined the precise 5' and 3' terminal sequences of arenaviral mRNAs, as well as confirmed the presence of a 5' cap structure that is used to prime translation (Fechter & Brownlee, 2005) and propose the formation of the 3' hairpin loop (Meyer and Southern, 1993). To this end, using this system it would be possible to see whether deletion of mRNA sequences or the presence of NP influenced reporter gene expression. In this experimental design, the reporter protein gene expressed by the

synthetic arenavirus mRNA was *Gaussia* luciferase, and the mRNA was transcribed and co-transcriptionally capped *in vitro*.

Another central tenet of the models we propose is that the interaction between the mRNA 5' and 3' ends does not require PABP, but does require eIF4G in order to provide a scaffold for assembly of other eIF4F components, and thus promote mRNA circularisation. This complete complex could then allow for the complex to associate efficiently with eIF3E and other elements of the 43S PIC, and confer efficient translation initiation. Having confirmed the requirement of the 3' secondary structure for efficient mRNA translation, it was necessary to assess whether these cellular components were localised to the RTCs during an arenavirus infection, which are proposed to be the sites of viral gene expression and protein synthesis. This was achieved by using IF, which supported previous work described in chapter 5, which showed localization of eIF4G within RTCs. These sites of NP are thought to be derived from early endosomes – as shown by the findings in chapter 4 – and to constitute replication centres and the sites of arenaviral mRNA translation (Baird et al., 2012).

The results presented in this chapter suggest that both the 3' RNA secondary structural element, as well as a virus protein, likely NP, enhances viral mRNA translation. In addition, the localisations of cellular proteins involved in translation are consistent with our proposed models of arenavirus mRNA circularisation.

6.2 Results

6.2.1 Functional analysis of arenaviral-like mRNA translation

In order to examine the role of RNA sequences in promoting arenavirus mRNA translation, it was necessary to develop a system in which the characteristics of a model arenavirus mRNA could be site-specifically modified, and its ability to be translated measured by detecting expression of a reporter gene. To establish such a system, a plasmid was designed that would express an RNA

possessing the 5' and 3' un-translated regions of the LCMV-Arm NP mRNA, surrounding the gene encoding *Gaussia* luciferase, shown in figure 47.

Diagram illustrating the synthesised arenaviral-like mRNA, descriptions proceed from left to right. The green circle illustrates the 5' ARCA cap (Promega); the first 61 nucleotides of LCMV NP mRNAs preceding the start codon shown in green; Gaussia Luciferase sequence (genbank AY015993.1) from the start codon until the final codon before stop codon – shown in gold, with the final 63 nucleotides, including the stop codon, of LCMV NP mRNA shown in green.

Figure 47.   Diagram depicting constructed arenavirus-like reporter mRNA

This reporter gene was chosen as the corresponding reporter protein is secreted from cells, enabling a simple and rapid comparison of reporter expression at multiple time-points.
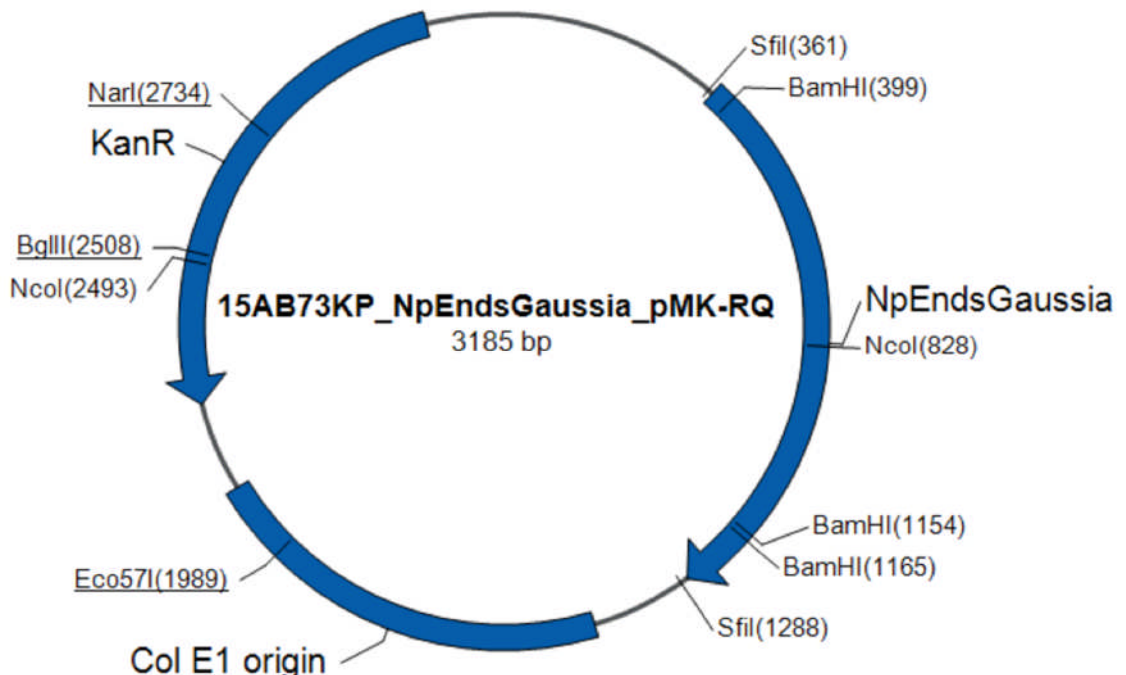
The synthesised NPEG gene was constructed by GeneArt, and inserted into a pMK-RQ vector. Full NPEG sequence can be found in Appendix I.

Figure 48.   pMK NPEG plasmid map

154

Figure 48 illustrates the constructed plasmid, with the gene – termed NPEnds*Gaussia* (NPEG) under control of the bacteriophage T7 RNA polymerase. The plasmid DNA was linearised with SfiI and transcribed *in vitro* to generate a run-off RNA transcript, which was co-transcriptionally capped using an ambion mMessage mMachine kit. RNA was subsequently purified via a Qiagen RNeasy purification kit, with mRNAs quantified via a NanoDrop1000 via spectrophometry. To ensure that the 3' hairpin structure of arenaviral mRNAs was unadulterated, a hepatitis delta virus (HDV) ribozyme sequence was incorporated after the final nucleotide of the hairpin loop. The presence of this HDV ribozyme sequence enabled the self-removal of the downstream RNA sequence, immediately after the hairpin, such that no additional non-viral 3' nucleotides were present (Chadalavada et al., 2007).

6.2.2 Effect of LCMV infection, and presence of NP, on *Gaussia* luciferase expression

To establish whether viral proteins, likely NP, were able to enhance the expression of arenavirus-specific mRNAs, A549 cells were subjected to lipofectamine-mediated mRNA transfection, and the levels of luciferase reporter activity secreted into the supernatant was assessed at various time points post transfection using the chemiluminescent coelenterazine compound, detected by a fluorescent plate reader. Figure 49 shows that for cells treated with mRNA transfection alone, the zero hour time point showed low secreted luciferase, at the same level as the un-transfected negative controls, and thus was considered to represent background signal. At 2 hours post transfection, cells transfected with NPEG mRNA showed increased luciferase secretion, and these levels increased up to a maximum at 6 hours post transfection. After this time point, secreted luciferase declined, such that after 12 hours, the luciferase signal was again approaching background levels. These findings show that the NPEG mRNA acts as a functional mRNA that can be translated in the absence of any viral proteins. As expected, for cells mock transfected with mRNA, the levels of secreted luciferase levels did not increase over the same time period.

We next assessed whether pre-infection with LCMV had any influence on the ability of arenaviral-like mRNA to be translated. The models presented in figure 47 would predict that pre-infection would stimulate luciferase expression due to the presence of abundant LCMV NP within the infected cells. In order to provide a sufficient cytosolic pool of NP, cells were pre-infected for 24 hrs with an MOI of 0.1 prior to mRNA transfection. The results shown in figure 49 indicate that at early time points of 2 and 4 hours post transfection, cells treated with both LCMV and NPEG mRNA generated increased levels of luciferase reporter compared to cells treated with NPEG alone. Student T-test analysis showed the level of increase at these time points was significant, although at later time points the differences were either negligible (Appendix IV, table 5) or insignificant. In addition, the signal remained at peak levels between 4 and 6 hours. When compared to the noticeable decline from the peak fraction immediately after the peak fraction for NPEG alone samples, it would appear that arenavirus like mRNA degradation within LCMV-infected cells was lower than uninfected cells. As expected, cells treated with LCMV alone secreted no luciferase. These results show that LCMV infection stimulates enhanced translation of the NPEG mRNA, and this finding is consistent with the proposed role of NP, although further work will need to be done to determine whether the supply of NP alone is responsible.
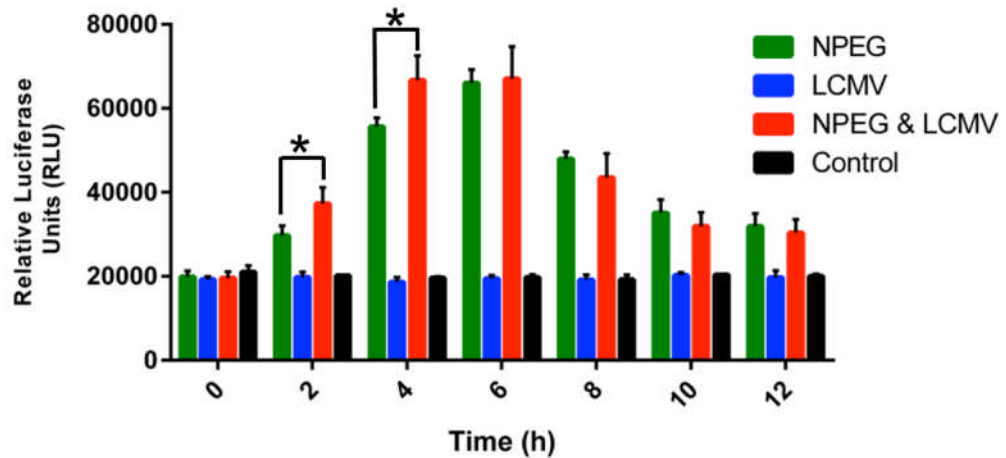
Graph showing Relative Luciferase Units (RLU) against time post-transfection of *Gaussia* expression, from A549 cells. NPEG (green) shows the RLU with only NPEG mRNA transfected into cells; LCMV-Arm (blue) shows RLU upon only LCMV-Arm infection; NPEG & LCMV-Arm (red) shows the RLU with NPEG transfection and LCMV-Arm infection; Control (black) shows the RLU of mock infected/transfected.

Figure 49.   **Enhancement of *Gaussia* luciferase reporter mRNA in presence of NP**

6.2.3 Effect on *Gaussia* Luciferase expression under hRSV infection.

In the above section, LCMV-Arm infection enhanced the expression of a *Gaussia* luciferase reporter that was transfected into cells as an arenaviral-like mRNA, and it was proposed this was due to the expression of a LCMV protein, likely NP. In order to determine whether the observed translational enhancement effect was not simply induced as a general cellular response to a viral infection, hRSV was used to infect A549 cells at MOI 1 for 24 hours prior to mRNA transfection, with the comparative expression of *Gaussia* luciferase (LUC) assessed as described in 6.2.2.1. Figure 50 shows the RLU detected following NPEG mRNA transfection (green bars), with the trend in LUC expression showing a steady increase before peaking at 6-8 hours, and was performed by Hayley Pearson. NPEG transfected cells infected with HRSV showed decreased NPEG expression compared to cells treated with NPEG alone. This indicates that the translational enhancement of the NPEG mRNA is specific to LCMV, and not to virus infection in general.
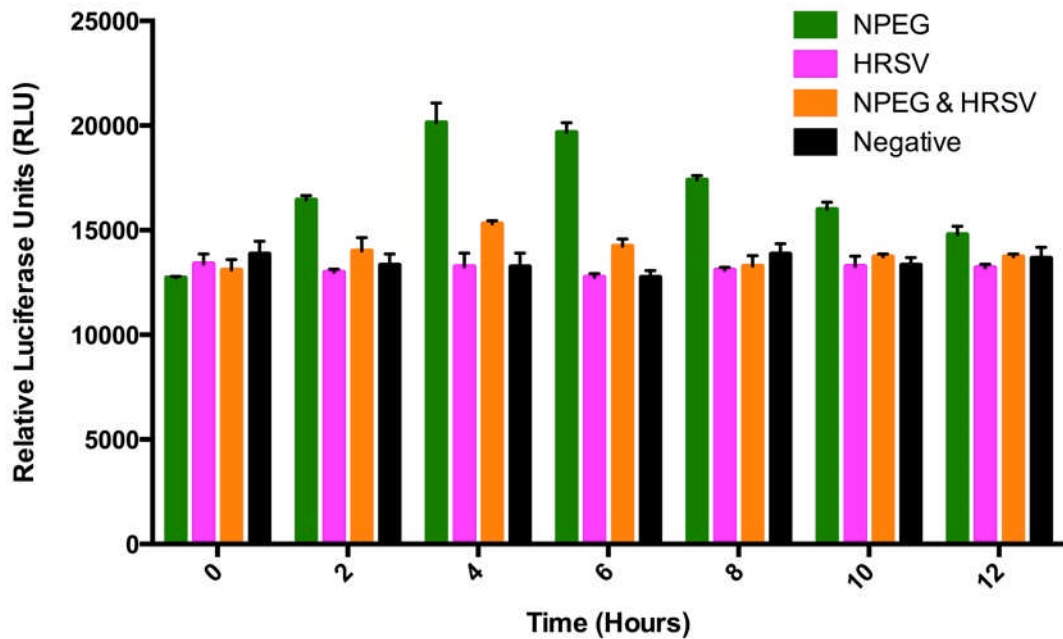
Graph showing Relative Luciferase Units (RLU) against time post-transfection of Gaussia expression, from A549 cells. NPEG (green) shows the RLU with only NPEG mRNA transfected into cells; HRSV (pink) shows RLU upon only HRSV infection; NPEG & HRSV (orange) shows the RLU with NPEG transfection and HRSV infection; Control (black) shows the RLU of mock infected/transfected.

Figure 50.   NPEG expression in presence of HRSV

6.2.4 LCMV mRNA 3' hairpin structure is responsible for efficient translation through association with NP

In the previous sections, LCMV-Arm was shown to enhance and HRSV shown to decrease arenaviral-like mRNA translation in A549 cells. In order to determine whether NP interacts with the 3' hairpin as proposed, a change was made to the NPEG plasmid, constructed by Hayley Pearson, generating a 3' deletion mutant (3'Δ) – the exact sequence alteration is shown in Appendix I, and demonstrated diagrammatically in figure 51.
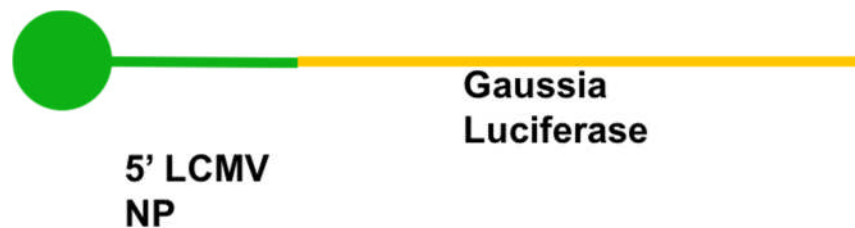
Diagram illustrating the synthesised arenaviral-like mRNA, descriptions proceed from left to right. The green circle illustrates the 5' ARCA cap (Promega). Gold indicates the *Gaussia* luciferase ORF followed by a stop codon.

Figure 51. **Diagram depicting truncated 3'Δ reporter mRNA**

Figure 52 shows the expression of *Gaussia* luciferase represented by RLU from differentially transfected/mock transfected and infected/mock infected A549 cells, as performed in 6.2.2.1. Cells transfected with NPEG and not LCMV infected presented a similar expression profile to the NPEG RLU detection in figure 49. After reaching a peak detection at 6 hours, the signal declines towards the background at 10 hours post transfection. As observed previously, LCMV infected and negative samples exhibited background levels of RLU throughout. The RLU detection of the 3'Δ and 3'Δ&LCMV samples behaved identically to the negative samples, with no increase from the background signal throughout the 10 hours of sample collection. This indicates that the observed enhancements NP exhibited above are not observed when the 3' hairpin structure is absent. The lack of RLU detection for 3'Δ samples indicates that the 3' hairpin structure is essential for the translation of arenaviral mRNAs, irrespective of the presence of NP or other viral factors.
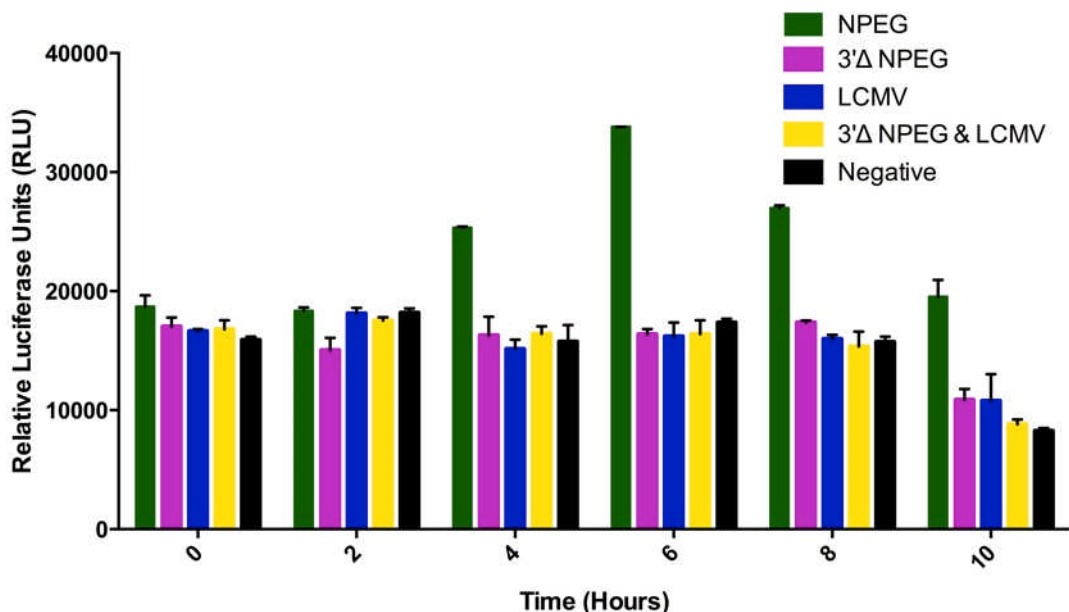
Graph showing Relative Luciferase Units (RLU) against time post-transfection of *Gaussia* expression, from A549 cells. NPEG (green) shows the RLU with only NPEG mRNA transfected into cells; 3' Δ NPEG (purple) shows RLU upon 3' Δ NPEG transfection; LCMV-Arm (blue) shows RLU upon only LCMV-Arm infection; 3' Δ NPEG & LCMV-Arm (yellow) shows the RLU with 3' Δ NPEG transfection and LCMV-Arm infection; Control (black) shows the RLU of mock infected/transfected.

Figure 52. **Lack of enhancement of 3'Δ NPEG reporter mRNA in presence of NP**

6.2.5 Presence of eIF4G and absence of PABP1 in RTCs and RTC-like structures

After establishing that the translation of arenaviral-like mRNA is enhanced during infection by LCMV-Arm through a mechanism proposed to be driven by 3' hairpin structures and the facilitation of circularisation by NP, it was necessary to determine whether either of the previously proposed models (fig 42) of circularisation could be correct. The recruitment of the circularisation components PABP or eIF4G within early endosome-derived sites of viral translation and replication – RTCs – would provide evidence that a viral factor aids in the circularisation of mRNA, and therefore contributes to the efficient expression of viral proteins.

6.2.5.1 Relative localisation of eIF4G and PAPB during LCMV infection of A549 cells

In order to provide further support for a functional role for NP in translation of arenavirus-specific mRNAs, it was necessary to establish whether eIF4G, or eIF4G and PABP1, were present within the theorised sites of viral translation, RTCs. The relative locations of eIF4G and PABP1 were thus determined during LCMV infection by IF microscopy, and compared against a mock-infected control. A549 cells were infected at MOI 0.01 or mock-infected, then fixed and permeabilised after 36 hours. Cells were then stained with relevant combinations of primary and secondary antibodies in order to visualise NP and the corresponding cellular markers. Figure 53 shows the diffuse localisation of both eIF4G and PABP1 under mock-infected conditions. Under LCMV-Arm infection, the distribution pattern of eIF4G (53**B**) shifts, appearing to co-localise with LCMV RTCs. As other elements of the eIF4F complex were previously identified as interacting partners of LUJV NP-EGFP and present within LCMV-Arm induced RTCs and particles, the role of eIF4G as the major scaffold protein of the eIF4F complex is indicative of organised and regulated amalgamation of translation factors. The presence of eIF4G in RTCs suggests that all the components of the eIF4F complex are present within RTCs, when

taken with the findings shown in chapters 4 and 5. Conversely, the relative localisation of PABP1 in LCMV-Arm infected A549 cells (53**A**) compared to mock-infected cells is unaltered, indicating its absence from RTCs, with no major evidence of PABP1 protein staining corresponding with that of NP. This suggests that PABP1 does not have a direct role in the circularisation of arenavirus mRNAs, and is thus not involved in viral mRNA translation, or a major role in LCMV-Arm translational enhancement.
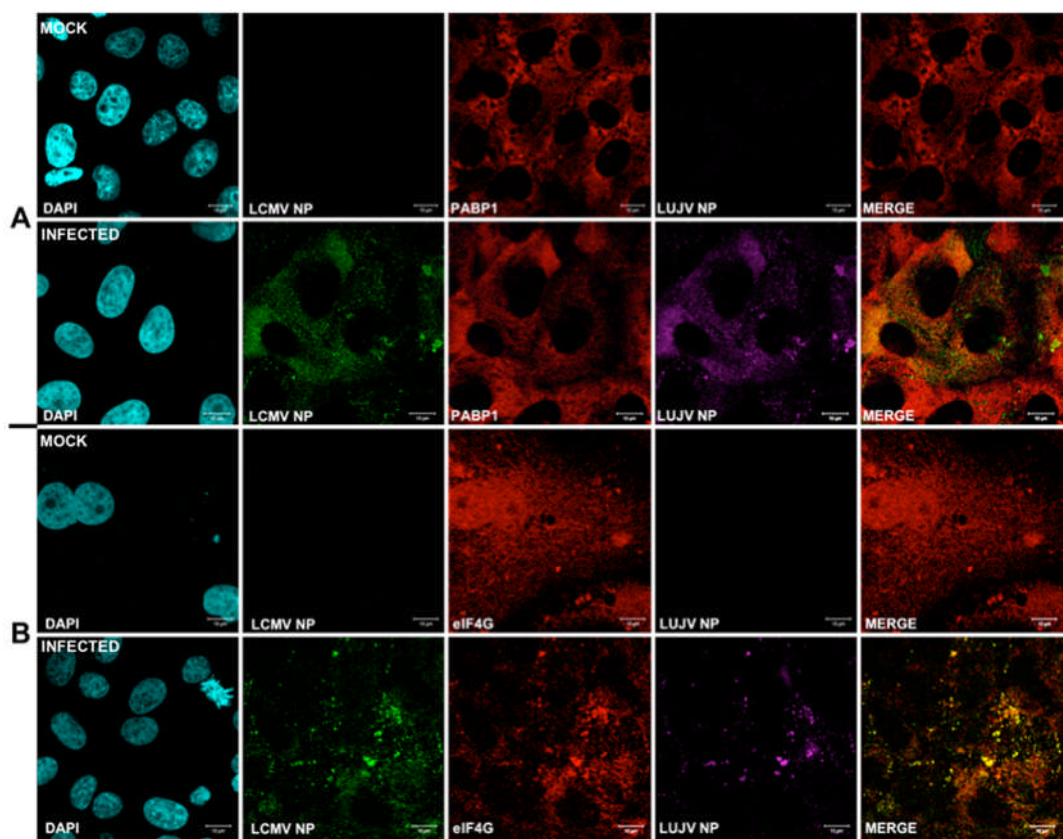
Immunofluorescence analysis of the relative distribution of PABP1 and eIF4G, against LCMV-Arm infected and mock infected A549 cells. **A** indicates the distribution of PABP1 against mock-infected and LCMV-Arm infected cells. **B** indicates the distribution of eIF4G against mock-infected and LCMV-Arm infected cells. In all image series, cyan indicates DAPI staining of DNA; green indicates LCMV-Arm antisera staining of LCMV-Arm NP; red indicates cellular protein of interest staining (as stated); magenta indicates LUJV NP antisera staining of LCMV-Arm NP. In all cases, the merged image is a composite of green and red channels.

Figure 53.   **Comparative localisation of LCMV-Arm NP against eIF4G and PABP**

6.2.5.2 Relative localisation of eIF4G with LUJV NP-EGFP

After establishing the presence of eIF4G within RTCs during LCMV infection, it was necessary to determine whether this effect was observed when NP was expressed in isolation from other viral components, or in other words whether LUJV NP-EGFP specifically could relocalise eIF4G on its own. Identifying a general effect in more than one OW virus would indicate a potential key role for NP in the arenaviral translation in general, rather than limited to solely LCMV-Arm. A549 cells were differentially transfected with LUJV NP-EFP or EGFP. These cells were then fixed, permeabilised and differentially stained with anti-eIF4G and PABP1 antibodies, with the relative localisation of eIF4G and PABP1 subsequently determined via IF microscopy. Figure 54 shows the localisation of eIF4G within RTC-like structures comprising LUJV NP-EGFP protein density, a shift from its normal localisation in un-transfected and mock transfected cells. The presence of eIF4G in these structures suggests that LUJV NP is responsible for the recruitment of several elements of the translation initiation complex into RTCs. The localisation of PAPB1 in LUJV NP-EGFP expressing cells showed a limited degree of co-alignment with LUJV NP-EGFP compared to the EGFP control, evidence of a potentially different strategy to that of LCMV-Arm, although the similarities between LUJV NP and LCMV-Arm NP observed previously indicate that the association of PAPB1 with LUJV NP-EGFP should be explored further.
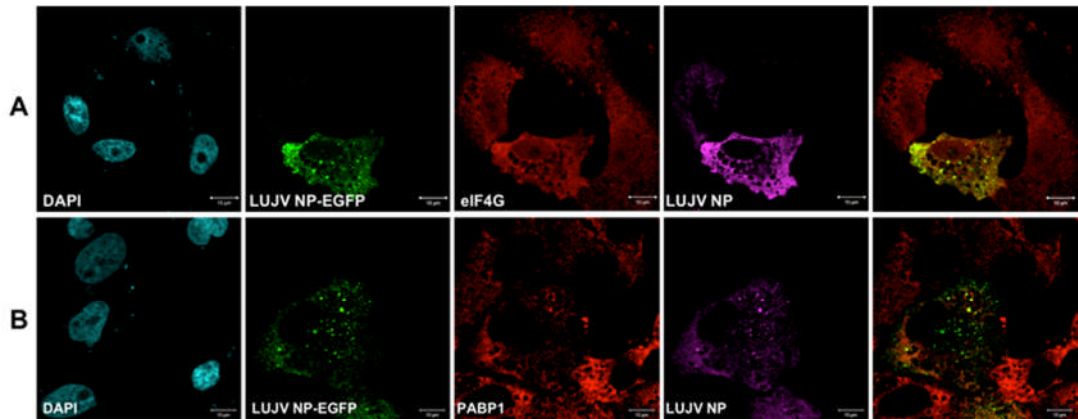
Immunofluorescence images showing the relative localisation of LUJV NP-EGFP with **A** eIF4G and **B** PABP1. Cyan shows DAPI staining of DNA; Green shows LUJV NP-EGFP; Red shows cellular protein staining (as indicated); Magenta shows LUJV NP antibody staining with the merge images showing both green and red channels. Scale bar indicates 10μm.

Figure 54.   Comparative localistation of LUJV NP-EGFP against eIG4G and PABP

6.3 Discussion

This chapter aimed to explore the role of NP in the translation of viral mRNAs, and to test a proposed model for how NP might achieve this. Previously identified interactions of LUJV NP-EGFP, and the subsequent localisation of such factors within RTCs in LCMV-Arm infected cells, suggested NP had a role in facilitating translation. This the experiments performed in this chapter demonstrated that, after infection by LCMV-Arm, the translation of arenavirus-like mRNAs is enhanced. The identification of these interacting partners within LCMV infectious particles also supported a model of 'pre-loading' of viral particles in order to facilitate rapid infection establishment within cells, as illustrated in figure 44.

Figure 46 illustrated our two proposals for efficient viral mRNA expression. The evidence discussed in chapters 4 and 5 indicated NP facilitates the formation of potentially functional RTCs. To this end, it is likely that this role is not limited to the recruitment of translation-associated proteins into RTCs – as described in chapters 4 and 5. The proposed models present the known interacting partners of LUJV NP and LCMV-Arm NP – eIF4F complex proteins and ribosomal subunit proteins – arranged in their orientations corresponding to how they are thought to operate during eukaryotic mRNA translation. As

discussed previously, PABP1 is required for circularisation of cellular mRNAs, and thus efficient translation, through association with the major scaffolding element of the eIF4F complex – eIF4G (Kahvejian et al., 2012; Prévôt et al., 2003; Jackson et al., 2010). In order for one of these models of viral mRNA circularisation to be justifiable, the presence of eIF4G or both eIF4G and PABP1 within RTCs would be necessary. Figure 53 indicates that eIF4G and not PABP1 is present within RTCs, during LCMV-Arm infection. The absence of PABP1 suggests that it plays no role in the circularisation of arenaviral mRNA, whereas the presence of eIF4G mimics that observed under infection by the NW arenavirus JUNV (Linero et al., 2013). These findings support our proposal that viral mRNA circularisation is facilitated by the binding of NP to the 3' hairpin structure, followed by eIF4G association – as shown in figure 55**A**, with the non-circularised from in the absence of NP shown in 55**B**.
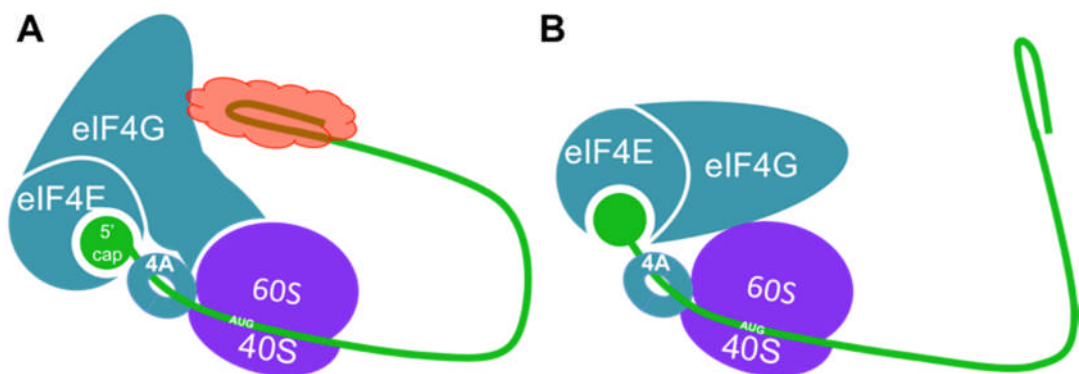
A Diagram illustrating the proposed model for viral mRNA circularisation in the presence of NP, with all elements of the eIF4F complex present, prior to the release of associated factors through GTP hydrolysis of eIF5B. **B** Diagram illustrating the proposal in the absence of NP. For diagrammatical purposes, not all elements of initiation complex are shown.

Figure 55. **Proposed role of NP in circularisation of arenaviral mRNAs**

If the proposed role of NP in translation enhancement is correct, then in the presence of NP there would be an increase in translational efficiency. In order to test this hypothesis, an arenaviral-like mRNA was synthesised, with 5' and 3' elements identical to that of LCMV with an internal reporter, *Gaussia* luciferase, enabling quantitative comparison between *Gaussia* expression in NP-containing cells and non-NP containing cells. Figure 49 shows the enhancement of mRNA translation in the presence of NP, through LCMV-Arm

infection. This observation supports the previous assertions that NP plays an effective role in translational enhancement, through its interactions with a wide array of cellular translation factors. While it cannot be included from this particular observation that NP is solely responsible for this enhancement, the presence of viral proteins enables increased translational enhancement of mRNAs lacking a poly(A) tail; possessing a 3' hairpin structure, identical in sequence to those found on WT LCMV mRNAs.

In order to assess whether this enhanced translation following virus infection was specific to infection by an arenavirus, the structurally unrelated HRSV was also used to co-infect the NPEG transfected cells. HRSV infection did not provide translational enhancement, suggesting arenaviral mRNA expression is not simply up-regulated through a non-specific virus infection-related effect. The lack of any enhancement suggests that the increase in reporter expression observed under LCMV-Arm infection is due to a specific LCMV driven event.

In addition, mutation of the NPEG plasmid, generating a 3' deletion mutant as shown in figure 51, indicated the importance of the 3' hairpin in facilitating the translational enhancement observed previously. The lack of RLU observed under 3'Δ transfection not only supports the theory that the 3' hairpin is a structure required for efficient translation, but that it is a necessity for the translation of arenaviral mRNA. The total lack of *Gaussia* luciferase signal upon the removal of the 3' mRNA hairpin structure strongly suggests that cellular components interact with the 3' hairpin structure of arenaviral transcripts, with viral components enhancing this relationship in some manner.

This chapter has demonstrated that the NP of both LUJV-Arm and LCMV colocalises with eIF4G and not PABP1 inside RTCs, indicating that arenaviral mRNAs are circularised via a novel mechanism. Due to the presence of dsRNA sequences at the 3' end of viral mRNAs, NP is the most likely viral protein to interact with dsRNA sequences on mRNAs (Meyer & Southern, 1994; Hastie et al., 2012). Translational enhancement of mRNAs possessing

3' hairpins during LCMV-Arm infection indicates that a viral factor does indeed facilitate efficient mRNA expression – most likely through circularisation by NP. This novel mechanism of arenaviral protein expression should be explored in related arenaviruses, with the potential to be exploited by therapeutics.

As ribosomes have been previously identified as not essential for replication (Leung & Rawls, 1977), their presence within arenaviral particles has been largely ignored. The identification of ribosomal subunits within viral particles in this study alongside a wide array of translation factors, indicated an involvement in the viral life cycle. As illustrated in figure 44, corresponding with the findings shown in figure 42, that multiple cellular proteins are incorporated into LCMV-Arm particles, our proposition is that translation-associated cellular components are 'pre-loaded' within LCMV-Arm particles. Virally transported ribosomes could then immediately transcribe newly translated mRNAs after membrane fusion and initial transcription has occurred. This immediate availability could then confer increased initial protein expression in order to increase the intracellular levels of viral proteins quickly. In order to test the activity of these ribosomes, it would be preferable to inactivate them irreversibly prior to infection in order to compare the protein expression of WT vs the inactivated state. Temporarily inactivating ribosomes through treatment with cyclohexamide or similar compounds would not achieve the desired result, as the effect would subside immediately upon removal of the inhibitor compound (Schneider-Poetsch *et al.* 2010). The use of an irreversible ribosomal inactivating protein (RIP) such as saporin could establish the required inactivation (Iglesias *et al.* 1993); however, the introduction of such an agent into an infectious viral particle would create a delivery vector for the RIP. Thus it was determined that it would not be possible establish whether virally-delivered ribosomes have any effect on establishing infection under the containment facilities available.

The previous findings discussed above enabled the generation of a number of functional models. Figure 44 proposed the presence of multiple cellular

translation factors, arranged within infectious LCMV particles to allow for immediate, potentially coupled, transcription and translation upon cell entry, within early endosomes – as shown in figure 38 – a proposition not without precedent given the coupled nature of bunyavirus transcription and translation (York et al., 2008; Barr, 2007). The presence of such an array of factors within infectious particles suggests they are unlikely be included by chance, and were theorised to therefore be important in the context of viral infection.

As discussed above, other viruses possessing 3' UTR structures within their mRNAs possess strategies for increasing their translational efficiency through the 3'UTR and a viral protein co-factor. The NSP3 or rotaviruses is a distinct example, whereby the interaction of the protein with the 3' UTR and eIF4G confers efficient translation of viral mRNAs. Arenaviruses, with their known and prominent 3' UTR hairpin, would require a similar strategy, and the combined findings of chapter 4, 5 and 6 indicate that NP of two OW arenaviruses aid in the generation of RTCs and the recruitment of translation associated proteins to them; LCMV-Arm infected cells exhibit an enhanced translation of arenaviral-like mRNAs and that 3' hairpin structure of mRNA is essential for the translation of mRNAs.

This experiments described in this chapter show that LCMV-Arm NP and LUJV NP co-localising with eIF4G within RTCs and RTC-like structures in A549 cells, with PABP1 absent from LCMV-Arm RTCs, indicating NP plays a significant role in mRNA translational efficiency, through mRNA circularisation. In assessing the proposed model of arenaviral mRNA translation, A549 cells expressing an arenaviral-like mRNA exhibited a translational enhancement in cells containing NP. The results presented here indicate a novel mechanism in arenaviral mRNA translation, through the interaction of a viral factor, theorised to be NP, circularising arenaviral-like mRNA through binding with the 3' hairpin displayed on arenaviral mRNAs.

7 General Discussion

The *Arenaviridae* family are a group of viruses of increasing global importance. Collectively they are by far the largest cause of haemorrhagic infections per year, with LASV infection the most prominent, causing in excess of the 20-50,000 reported deaths per annum. These neglected pathogens have the potential to cause major outbreaks on a scale similar to the 2013-16 West African Ebola virus outbreak, due to their similar transmission route, morbidity and symptom profile. Indeed, many presenting cases of severe febrile illness admitted to Ebola Treatment Units in Sierra Leone were in fact due to LASV, rather than EBOV (Personal Experience, 2015). The 2014/15 outbreak has highlighted the frailties in healthcare provision within developing states, and their reliance on overseas aid to manage severe public-health events. The incidence of emergence of new pathogens is still high, with several arenaviruses emerging in 2015 alone (Bisordi et al., 2015; Aqrawi et al., 2015; Lavergne et al., 2015; Witkowski et al., 2015; Gryseels et al., 2015); this trend highlights the danger that potentially devastating infections could emerge from isolated population centres.

The need to gain further understanding of highly pathogenic arenaviruses, along with other neglected tropical diseases, is apparent. There is potential for a novel pathogenic outbreak in a naïve population with poor healthcare surveillance leading to more large-scale outbreaks. Lessons learned from the West African Ebola virus outbreak have highlighted the necessity to have effective healthcare surveillance in rural communities (Crowe et al., 2015; Richards et al., 2015; Olugasa & Dogba, 2015), but without effective treatment strategies there will always be the risk for novel infections to cause more widespread outbreaks, outside of isolation.

Arenaviral replication strategies are incompletely understood. Identifying the roles of proteins such as NP in the replication cycle could unlock potential strategies for therapy. Establishing NP as a protein important in translation

suggests a novel role for a protein already known to be multi-functional. The roles of NP in encapsidation, transcription and immune modulation are well established, but its role in mRNA translation is less so (Tortorici et al., 2001; Pedersen & Konigshofer, 1976; Martinez-Sobrido et al., 2009).

LUJV NP and LCMV-Arm NP interact with, and redistribute, several components of translation complexes

The interaction of LUJV NP with translation-associated proteins presented here illustrates a new insight regarding OW arenaviral replication strategies. Facilitating translation through factor recruitment alone is of distinct interest. The ability of LUJV NP to seemingly recruit a number of translation factors and ribosomes to de-facto RTCs is evidence of NP redistributing cellular proteins and disrupting the cellular environment beyond previously described functions. Baird *et. al.* (2012) reported the establishment of cytosolic structures in cells during JUNV infection, and the localisation of numerous cellular factors to these structures suggested their role in replication. In this study, these so called RTCs have been identified as being derived from early endosomes (figure 35). The knowledge that LUJV NP generates cytosolic structures containing a collection of proteins involved in translation - as observed during JUNV infection - indicates that RTC generation is not limited to NW pathogens. The ability of NP to hi-jack early endosomes to gather cellular proteins is interesting, given the lack of other viral proteins to assist in subverting the cell's natural defences against such re-organisation. As early endosomes are the site of viral membrane fusion (York et al., 2008), the utilisation of these structures by NP is consistent with previous knowledge regarding the arenaviral life cycle.

LUJV NP-EGFP appears to be able to stimulate re-distribution of a multitude of cellular translation factors, even with an EGFP tag. Figures 35, 37 & 54, indicated the redistribution of multiple cellular proteins into distinct cytosolic puncta. The SILAC MS dataset used primarily to identify potential interacting partners initially identified three potential avenues – HSP70, actin (along with other cytoskeletal elements) and eIF4A. All three were identified via GFP-Trap IPs as valid interacting partners, but the redistribution and co-alignment of

eIF4A, allied with similar observations for the NW virus JUNV, indicated LUJV NP may act in a manner similar to its South American cousins (Baird et al., 2012). Further investigations established that LUJV NP-EGFP could redistribute a number of cellular proteins into distinct cytosolic structures. The alignment with several ribosomal proteins, including their immunoprecipitation via GFP-Trap, also supported these observations.

The subsequent application of these findings to LCMV-Arm NP and the re-distribution of similar translation factors indicates that this strategy is of importance to multiple members of the *Arenaviridae* family. Three separate viruses exhibit such features – shown in figures 35, 37, 43, 53 & 54, and Baird *et al.* 2012 – which suggests that recruitment by NP of translation factors straddles both NW and OW clades. This evidence led us to propose a model for NP in translational modulation that conflicts with that proposed by Linero *et al.* in 2013. The findings that JUNV NP interacted with eIF4G and eIF4A, but not eIF4E conflict with the strong correlation of LCMV-Arm NP with eIF4E-PHOS, as seen in figure 43. The findings presented here suggest a novel mechanism regarding the role of the arenavirus NP in viral mRNA translation.

Presence of NP enhances viral mRNA translation

Observations presented here also reveal that NP, in conjunction with its known recruitment of translation-associated complexes, enhances mRNA translation; most likely in order to enhance virus replication. The proposed mechanism shown in figure 46 was assessed by establishing whether the presence of NP generated an enhancement of translation. The increase in translational efficiency observed demonstrated that the involvement of NP is not incidental and that this highly multi-functional protein plays crucial roles in several parts of the viral life cycle. The mechanism underlying this enhancement is not fully understood, but the findings in figures 49 & 52 support the proposal in figure 55 that NP interacts with the 3' mRNA hairpin structure, facilitating circularisation.

The 3' end of LCMV NP mRNAs has been known for some time to generate hairpin structures (Meyer & Southern, 1994), with the corresponding

sequences at the 3' end of LUJV NP mRNA also predicted to generate a large 27bp hairpin structure, shown in figure 56, with this LUJV NP predicted hairpin exhibiting 96% canonical base pairing.
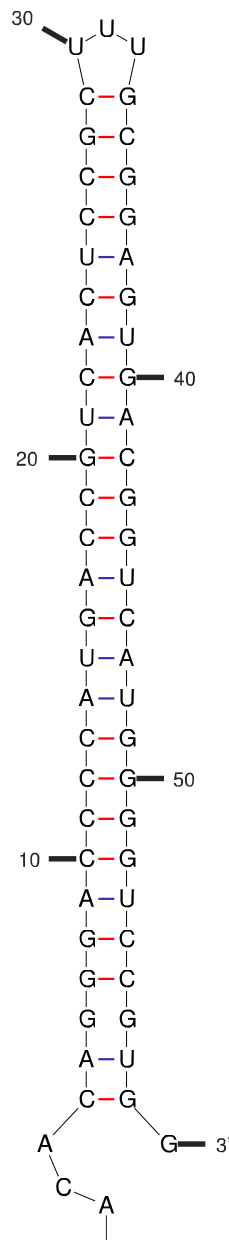
mFold predicted secondary RNA structure of 3' LUJV NP (Zuker, 2003). LUJV NP sequence obtained from GenBank – accession number NC_012776.1.

Figure 56.   Predicted 3' hairpin on LUJV NP mRNAs

This hairpin structure – as for LCMV – is of great interest. As discussed previously, it would seem incongruous for a virus to specifically encode and

display such a significant hairpin structure on mRNA sequences. Double stranded RNA is a known PAMP for immune surveillance proteins, with multiple viruses exhibiting strategies to avoid the detection of dsRNA replication intermediaries (Hastie et al., 2012; Oshiumi et al., 2013; Hale et al., 2008). The conspicuous presence of the 3' hairpin within mRNA sequences poses an intriguing question as to what its function is (Meyer & Southern, 1994).

The findings of this work indicate a role for NP in translation, and it is thus proposed that the role of NP could be to enhance mRNA expression, rather than merely limited to factor recruitment. This proposal was based on NP interacting with 3' mRNA hairpins to facilitate mRNA circularisation. In order to test this hypothesis, an arenaviral-like mRNA was generated, with LCMV NP sequences at both the 5' and 3' ends – shown in figure 47. Translational efficiency of this arenaviral-like mRNA was increased in the presence of LCMV-Arm NP against the -NP transfection control. The identification of a significant enhancement of reporter expression in cells expressing NP indicated that NP enhances translational efficiency. The model proposed in Chapter 6, figure 55, was supported by these findings. These findings indicate a novel role of NP in arenaviral mRNA translation.

Arenaviruses package ribosomal proteins alongside translation initiation and elongation factors within infectious viral particles.

The enhancement of mRNA translation shown here alone would be of interest, however, the identification of a multitude of translation factors – all of which appear to be involved in viral mRNA translation, facilitated by NP – within infectious viral particles is most intriguing.

As discussed above, the combination of a wide array of proteins identified within LCMV-Arm particles indicated a role for NP in translation. Pre-loading of viral particles with translation-associated machinery could potentially expedite initial infection establishment, as viral proteins could be rapidly expressed, without the rate-limiting step of recruiting factors in order to allow

translation to occur. If translation-associated machinery could be deployed alongside transcribing L, synthesised mRNAs could be translated rapidly.

This coupling could be similar to the closely related *Bunyaviridae,* whereby transcription of the genome to form mRNA requires the simultaneous translocation of ribosomes along nascent mRNA (Barr, 2007). Whilst there is no evidence to suggest that mRNA synthesis requires the presence of ribosomes at this time, the presence of multiple cellular proteins within viral particles, including ribosomal subunits, does indicate an involvement in replication strategies. The model proposes that the delivery of ribosomes and their associated initiation and elongation factors (figure 42) will contribute to efficient and rapid mRNA translation. In all likelihood, given the apparent interaction of NP with 3' hairpin structures and the likely mRNA circularisation as a result, the pre-loading and delivery of translational machinery is likely not directly coupled to transcription, but a situational efficiency step.

By delivering factors and removing the need for cellular proteins to be recruited immediately, rapid translation of mRNA could occur, facilitating the synthesis of increasing copies of viral proteins. This would, in turn allow more cellular factors to be recruited in order to keep up with the growing translational demands.

The enhancement of mRNA translation seen in figure 49 supports this theory, given that, in the presence of NP, overall signal induction is greater, and drops at a lower rate than the control sample. This might be due to NP being able to gather all the necessary components in one subcellular location, increasing efficiency. eIF4E-Phospho has for some time been known to exhibit increased 5' cap binding efficiency vs un-phosphorylated eIF4E, increasing translational efficiency as a result, and is contraindicated in tumour development (Furic et al., 2010). The distinct co-alignment of LCMV-Arm NP with eIF4E-phospho in RTCs acts as further evidence that NP facilitates translational enhancement.

Arenaviral nucleoproteins exhibit a novel role in establishing an environment conducive to replication and enhancing translation.

These findings overall have shown that arenaviral NPs are capable of orchestrating a significant re-organisation of cellular translation factors, in a relationship that ultimately increases mRNA translational efficiency. Figures 35, 37, 38, 43, 53 & 54 demonstrate the partial re-distribution of multiple components of translational machinery was demonstrated, most notably all subunits in the eIF4F complex. Translational machinery is also demonstrated to be packaged within viral particles, an observation which suggests that machinery could be 'delivered' to target cells, in order to facilitate rapid infection establishment, thus removing the need to recruit a huge array of host-cell factors prior to mRNA translation. Finally, an arenaviral factor proposed to be NP has also been demonstrated to directly enhance mRNA translation, potentially through the interaction with large dsRNA hairpin structures present on the 3' ends of viral mRNAs – as shown in figures 45 & 56. These observations suggest that NP has a greater role in replication than previously proposed, through the direct recruitment of translational machinery, and a proposed enhancement of viral mRNA translation through 3' interaction and mRNA circularisation.

Implications of findings

This study has shown that the NP of LUJV and of LCMV-Arm appears to expedite rapid viral replication through a novel mechanism of circularising the viral mRNA during the translational process, having first facilitated the recruitment and redistribution of several translational-associated complexes. Disruption of these processes may thus be a potential avenue of therapeutic attack during infection.

Bibliography

Abraham, J. et al. 2010. Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. *Nature structural & molecular biology*. **17**(4),pp.438–44.

Akira, S. et al. 2006. Pathogen recognition and innate immunity. *Cell*. **124**(4),pp.783–801.

Ambrosio, A. et al. 2011. Argentine hemorrhagic fever vaccines. *Human Vaccines*. **7**(6),pp.694–700.

Ammerman, N.C. et al. 2008. Growth and maintenance of Vero cell lines. *Current Protocols in Microbiology*. **Appendix 4**,p.Appendix 4E.

Andersson, I. et al. 2004. Role of actin filaments in targeting of Crimean Congo hemorrhagic fever virus nucleocapsid protein to perinuclear regions of mammalian cells. *Journal of medical virology*. **72**(1),pp.83–93.

Aqrawi, T. et al. 2015. Identification of snake arenaviruses in live boas and pythons in a zoo in Germany. *Tierarztliche Praxis. Ausgabe K, Kleintiere/Heimtiere*. **43**(4),pp.239–247.

Ariza, A. et al. 2013. Nucleocapsid protein structures from orthobunyaviruses reveal insight into ribonucleoprotein architecture and RNA polymerization. *Nucleic Acids Research*. **41**(11),pp.5912–5926.

Asogun, D.A. et al. 2012. Molecular diagnostics for lassa fever at Irrua specialist teaching hospital, Nigeria: lessons learnt from two years of laboratory operation. *PLoS Negl Trop Dis*. **6**(9),p.e1839.

Attar, N. 2016. ZIKA virus circulates in new regions. *Nat Rev Micro*. **14**(2),p.62.

Baird, N.L. et al. 2012. Arenavirus Infection Induces Discrete Cytosolic Structures for RNA Replication. *Journal of Virology*. **86**(20),pp.11301–11310.

Baird, N.L. et al. 2013. Arenavirus Infection Induces Discrete Cytosolic Structures for RNA Replication (vol 86, pg 11301, 2012). *Journal of virology*. **87**(5),p.2983.

Bannister, B. 2010. Viral haemorrhagic fevers imported into non-endemic countries: risk assessment and management. *British Medical Bulletin*. **95**(1),pp.193–225.

Barnett, S. a and Dickson, R.G. 1989. Wild mice in the cold: some findings on adaptation. *Biological reviews of the Cambridge Philosophical Society*. **64**(4),pp.317–40.

Barr, J.N. 2007. Bunyavirus mRNA synthesis is coupled to translation to prevent premature transcription termination. *RNA*. **13**(5),pp.731–736.

Barton, L.L. et al. 2002. Lymphocytic choriomeningitis virus: emerging fetal teratogen. *Am J Obstet Gynecol.* **187**(6),pp.1715–1716.

Bausch, D.G. et al. 2000. Diagnosis and Clinical Virology of Lassa Fever as Evaluated by Enzyme-Linked Immunosorbent Assay, Indirect Fluorescent-Antibody Test, and Virus Isolation. *Journal of clinical microbiology.* **38**(7),pp.2670–2677.

Bausch, D.G. et al. 2010. Review of the literature and proposed guidelines for the use of oral ribavirin as postexposure prophylaxis for Lassa fever. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* **51**(12),pp.1435–1441.

Bergeron, E. et al. 2012. Reverse Genetics Recovery of Lujo Virus and Role of Virus RNA Secondary Structures in Efficient Virus Growth. *Journal of Virology.* **86**(19),pp.10759–10765.

Bird, B.H. et al. 2012. Severe hemorrhagic fever in strain 13/N guinea pigs infected with Lujo virus. *PLoS Negl Trop Dis.* **6**(8),p.e1801.

Bisordi, I. et al. 2015. Pinhal Virus, a New Arenavirus Isolated from Calomys tener in Brazil. *Vector borne and zoonotic diseases (Larchmont, N.Y.).*

Blakqori, G. et al. 2009. Bunyamwera orthobunyavirus S-segment untranslated regions mediate poly(A) tail-independent translation. *Journal of virology.* **83**(8),pp.3637–46.

Bodewes, R. et al. 2014. Updated Phylogenetic Analysis of Arenaviruses Detected in Boid Snakes. *Journal of Virology .* **88** (2),pp.1399–1400.

Bonthius, D.J. 2012. Lymphocytic choriomeningitis virus: an underrecognized cause of neurologic disease in the fetus, child, and adult. *Semin Pediatr Neurol.* **19**(3),pp.89–95.

Bornhorst, J.A. and Falke, J.J. 2000. Purification of Proteins Using Polyhistidine Affinity Tags. *Methods in Enzymology.* **326**(16),pp.245–254.

Boursot, P. et al. 1996. Origin and radiation of the house mouse: Mitochondrial DNA phylogeny. *Journal of Evolutionary Biology.* **9**(4),pp.391–415.

Briese, T. et al. 2009. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS pathogens.* **5**(5),p.e1000455.

Briese, T. et al. 2009. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS Pathogens.* **5**(5),pp.1–8.

Burri, D.J. et al. 2012. Envelope glycoprotein of arenaviruses. *Viruses.* **4**(10),pp.2162–81.

Burri, D.J. et al. 2013. The role of proteolytic processing and the stable signal peptide in expression of the Old World arenavirus envelope glycoprotein ectodomain. *Virology.* **436**(1),pp.127–133.

Cajimat, M.N.B. et al. 2007. Catarina virus, an arenaviral species principally associated with *Neotoma micropus* (southern plains woodrat) in Texas. *The American journal of tropical medicine and hygiene*. **77**(4),pp.732–6.

Campbell Dwyer, E.J. et al. 2000. The lymphocytic choriomeningitis virus RING protein Z associates with eukaryotic initiation factor 4E and selectively represses translation in a RING-dependent manner. *Journal of virology*. **74**(7),pp.3293–300.

Cao, W. et al. 1998. Identification of α-Dystroglycan as a Receptor for Lymphocytic Choriomeningitis Virus and Lassa Fever Virus. *Science*. **282**(5396),pp.2079–2081.

Capul, A.A. et al. 2007. Arenavirus Z-glycoprotein association requires Z myristoylation but not functional RING or late domains. *Journal of virology*. **81**(17),pp.9451–9460.

Carter, S.D. et al. 2012. Structure, function, and evolution of the Crimean-Congo hemorrhagic fever virus nucleocapsid protein. *Journal of virology*. **86**(20),pp.10914–10923.

CDC 2015a. Arenavirus Fact Sheet. *CDC* [online]. Available from: http://www.cdc.gov/ncidod/dvrd/spb/mnpages/dispages/Fact_Sheets/Ar enavirus_Fact_Sheet.pdf [Accessed January 1, 2015].

CDC 2015b. Lymphocytic Choriomeningitis. *CDC* [online]. Available from: http://www.cdc.gov/vhf/lcm/.

Centers for Disease, C. et al. 2012. Possession, use, and transfer of select agents and toxins; biennial review. Final rule. *Federal register*. **77**(194),pp.61083–61115.

Centers for Disease, C. and Prevention 2006. Survey of lymphocytic choriomeningitis virus diagnosis and testing--Connecticut, 2005. *MMWR. Morbidity and mortality weekly report*. **55**(14),pp.398–399.

Chadalavada, D.M. et al. 2007. Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions. *RNA*. **13**,pp.2189–2201.

Charrel, R.N. et al. 2011. Arenaviruses and hantaviruses: from epidemiology and genomics to antivirals. *Antiviral research*. **90**(2),pp.102–114.

Charrel, R.N. et al. 2008. Phylogeny of the genus Arenavirus. *Current opinion in microbiology*. **11**(4),pp.362–8.

Chaudhuri, J. et al. 1997. Function of eukaryotic translation initiation factor 1A (eIF1A) (formerly called eIF-4C) in initiation of protein synthesis. *J Biol Chem*. **272**(12),pp.7883–7891.

Chevaliez, S. and Pawlotsky, J.-M. 2006. *HCV Genome and Life Cycle*.

Childs, J.E. et al. 1992. Lymphocytic choriomeningitis virus infection and house mouse (Mus musculus) distribution in urban Baltimore. *The American journal of tropical medicine and hygiene*. **47**(1),pp.27–34.

Chiller, J.M. and Oldstone, M.B.A. 1984. Selection of genetic variants of

lymphocytic choriomeningitis virus in spleens of persistently infected mice. Role in suppression of cytotoxic T lymphocyte response and viral persistence. *The Journal of experimental medicine*. **160**(2),pp.521–540.

Christie, M. et al. 2015. Structural biology and regulation of protein import into the nucleus. *Journal of molecular biology*.

Cogswell-Hawkinson, A. et al. 2012. Tacaribe virus causes fatal infection of an ostensible reservoir host, the Jamaican fruit bat. *Journal of virology*. **86**(10),pp.5791–5799.

Cornu, T.I. and de La Torre, J.C. 2001. RING finger Z protein of lymphocytic choriomeningitis virus (LCMV) inhibits transcription and RNA replication of an LCMV S-segment minigenome. *Journal of virology*. **75**(19),pp.9415–26.

Covarrubias, S. et al. 2009. Host Shutoff Is a Conserved Phenotype of Gammaherpesvirus Infection and Is Orchestrated Exclusively from the Cytoplasm. *Journal of Virology*. **83**(18),pp.9554–9566.

Crowe, S. et al. 2015. A Plan for Community Event-Based Surveillance to Reduce Ebola Transmission — Sierra Leone, 2014–2015. *Morbidity and Mortality Weekly Report (MMWR)*. **64**(3),pp.70–73.

Cucchi, T. et al. 2005. First occurrence of the house mouse (Mus musculus domesticus Schwarz & Schwarz, 1943) in the Western Mediterranean: A zooarchaeological revision of subfossil occurrences. *Biological Journal of the Linnean Society*. **84**(3),pp.429–445.

Van Cuong, N. et al. 2015. Rodents and risk in the Mekong Delta of Vietnam: seroprevalence of selected zoonotic viruses in rodents and humans. *Vector borne and zoonotic diseases (Larchmont, N.Y.)*. **15**(1),pp.65–72.

Cyr, D.M. et al. 1994. DnaJ-like proteins: Molecular chaperones and specific regulators of Hsp70. *Trends in Biochemical Sciences*. **19**(4),pp.176–181.

Dalby, B. et al. 2004. Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods (San Diego, Calif.)*. **33**(2),pp.95–103.

Dalziel, M. et al. 2014. Emerging principles for the therapeutic exploitation of glycosylation. *Science*. **343**,p.1235681,1 1235681,8.

Dobec, M. et al. 2006. High prevalence of antibodies to lymphocytic choriomeningitis virus in a murine typhus endemic region in Croatia. *Journal of medical virology*. **78**(12),pp.1643–1647.

Doherty, P.C. and Zinkernagel, R.M. 1975. H-2 compatibility is required for T-cell-mediated lysis of target cells infected with lymphocytic choriomeningitis virus. *The Journal of experimental medicine*. **141**(2),pp.502–7.

Dylla, D.E. et al. 2008. Basolateral entry and release of New and Old World arenaviruses from human airway epithelia. *Journal of virology*. **82**(12),pp.6034–6038.

Edgil, D. et al. 2006. Dengue virus utilizes a novel strategy for translation initiation when cap-dependent translation is inhibited. *Journal of virology.* **80**(6),pp.2976–86.

Eichler, R. et al. 2006. The role of single N-glycans in proteolytic processing and cell surface transport of the Lassa virus glycoprotein GP-C. *Virology journal.* **3**,p.41.

Emond, R.T. et al. 1982. A case of Lassa fever: clinical and virological findings. *Br Med J (Clin Res Ed).* **285**(6347),pp.1001–1002.

Fan, L. et al. 2010. Z proteins of New World arenaviruses bind RIG-I and interfere with type I interferon induction. *Journal of virology.* **84**(4),pp.1785–1791.

Faye, O. et al. 2015. Use of Viremia to Evaluate the Baseline Case Fatality Ratio of Ebola Virus Disease and Inform Treatment Studies: A Retrospective Cohort Study. *PLoS Medicine.* **12**(12).

Fechter, P. and Brownlee, G.G. 2005. Recognition of mRNA cap structures by viral and cellular proteins. *Journal of General Virology.* **86**(5),pp.1239–1249.

Fehling, S.K. et al. 2012. Multifunctional nature of the arenavirus RING finger protein Z. *Viruses.* **4**(11),pp.2973–3011.

Fichet-Calvet, E. et al. 2007. Fluctuation of abundance and Lassa virus prevalence in Mastomys natalensis in Guinea, West Africa. *Vector Borne Zoonotic Dis.* **7**(2),pp.119–128.

Fichet-Calvet, E. et al. 2014. Lassa serology in natural populations of rodents and horizontal transmission. *Vector Borne Zoonotic Dis.* **14**(9),pp.665–674.

Fichet-Calvet, E. and Rogers, D.J. 2009. Risk maps of Lassa fever in West Africa. *PLoS neglected tropical diseases.* **3**(3),p.e388.

Fischer, S. a et al. 2006. Transmission of lymphocytic choriomeningitis virus by organ transplantation. *The New England journal of medicine.* **354**(21),pp.2235–49.

Fung, A. et al. 2014. Efficiency of incorporation and chain termination determines the inhibition potency of 2???-modified nucleotide analogs against hepatitis C virus polymerase. *Antimicrobial Agents and Chemotherapy.* **58**(7),pp.3636–3645.

Furic, L. et al. 2010. eIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression. *Proceedings of the National Academy of Sciences of the United States of America.* **107**(32),pp.14134–9.

Garcia, J.B. et al. 2000. Genetic diversity of the Junin virus in Argentina: geographic and temporal patterns. *Virology.* **272**(1),pp.127–136.

Garcin, D. et al. 1993. The Tacaribe arenavirus small zinc finger protein is

required for both mRNA synthesis and genome replication. *Journal of virology.* **67**(2),pp.807–12.

Geisbert, T.W. et al. 2005. Development of a new vaccine for the prevention of Lassa fever. *PLoS medicine.* **2**(6),p.e183.

Georges, A.J. et al. 1985. Antibodies to Lassa and lassa-like viruses in man and mammals in the Central African Republic. *Transactions of the Royal Society of Tropical Medicine and Hygiene.* **79**(1),pp.78–79.

Gerlach, P. et al. 2015. Structural insights into bunyavirus replication and its regulation by the vRNA promoter. *Cell.* **161**(6),pp.1267–1279.

Giard, D.J. et al. 1973. In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *Journal of the National Cancer Institute.* **51**(5),pp.1417–1423.

Gnann, J.W. et al. 1983. Acyclovir: mechanism of action, pharmacokinetics, safety and clinical applications. *Pharmacotherapy.* **3**(5),pp.275–83.

Gomez, R.M. et al. 2011. Junin virus. A XXI century update. *Microbes and infection / Institut Pasteur.* **13**(4),pp.303–311.

Gosert, R. et al. 2003. Identification of the hepatitis C virus RNA replication complex in Huh-7 cells harboring subgenomic replicons. *Journal of virology.* **77**(9),pp.5487–5492.

Grosfeld, H. et al. 1995. RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. *Journal of Virology .* **69** (9),pp.5677–5686.

Grummt, I. 2006. Actin and myosin as transcription factors. *Current opinion in genetics & development.* **16**(2),pp.191–196.

Gryseels, S. et al. 2015. Gairo virus, a novel arenavirus of the widespread Mastomys natalensis: Genetically divergent, but ecologically similar to Lassa and Morogoro viruses. *Virology.* **476**,pp.249–256.

Hadi, C.M. et al. 2010. Ribavirin for Lassa fever postexposure prophylaxis. *Emerging infectious diseases.* **16**(12),pp.2009–2011.

Hale, B.G. et al. 2008. The multifunctional NS1 protein of influenza A viruses. *Journal of General Virology.* **89**(10),pp.2359–2376.

Hastie, K.M. et al. 2011. Crystal structure of the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding. *Proceedings of the National Academy of Sciences of the United States of America.* **108**(48),pp.19365–70.

Hastie, K.M. et al. 2016. Crystal structure of the oligomeric form of the Lassa virus matrix protein Z. *Journal of Virology .*

Hastie, K.M., Bale, S., et al. 2012. Hiding the evidence: two strategies for innate immune evasion by hemorrhagic fever viruses. *Current opinion in virology.* **2**(2),pp.151–156.

Hastie, K.M. et al. 2012. Structural basis for the dsRNA specificity of the Lassa virus NP exonuclease. *PloS one*. **7**(8),p.e44211.

Hellebuyck, T. et al. 2015. Detection of arenavirus in a peripheral odontogenic fibromyxoma in a red tail boa (Boa constrictor constrictor) with inclusion body disease. *Journal of veterinary diagnostic investigation: official publication of the American Association of Veterinary Laboratory Diagnosticians, Inc.* **27**(2),pp.245–248.

Hensley, L.E. et al. 2011. Pathogenesis of Lassa fever in cynomolgus macaques. *Virology journal*. **8**(1),p.205.

Hernandez, R. and Brown, D.T. 2010. Growth and maintenance of baby hamster kidney (BHK) cells. *Current protocols in microbiology*. **Chapter 4**(May),p.Appendix 4H.

Hinnebusch, A.G. 2006. eIF3: a versatile scaffold for translation initiation complexes. *Trends in Biochemical Sciences*. **31**(10),pp.553–562.

Huang, C. et al. 2015. Highly Pathogenic New World and Old World Human Arenaviruses Induce Distinct Interferon Responses in Human Cells. *Journal of virology*. **89**(14),pp.7079–7088.

Huang, Q. et al. 2015. In Vitro and In Vivo Characterizations of Pichinde Viral Nucleoprotein Exoribonuclease Functions. *Journal of Virology* . **89** (13),pp.6595–6607.

Huber, F. et al. 2013. Emergent complexity of the cytoskeleton: from single filaments to tissue. *Advances in physics*. **62**(1),pp.1–112.

Iapalucci, S. et al. 1991. The 3' end termini of the Tacaribe arenavirus subgenomic RNAs. *Virology*. **182**(1),pp.269–78.

Ishii, A. et al. 2012. Molecular surveillance and phylogenetic analysis of Old World arenaviruses in Zambia. *The Journal of general virology*. **93**(Pt 10),pp.2247–2251.

Ishii, A. et al. 2012. Molecular surveillance and phylogenetic analysis of Old World arenaviruses in Zambia. *Journal of General Virology*. **93**(PART 10),pp.2247–2251.

IUCN 2015a. Mastomys natalensis. *IUCN 2015*.

IUCN 2015b. Mus musculus. *IUCN 2015*.

IUMS 2015. Virus Taxonomy: 2014 Release. *International Committee on Taxonomy of Viuses* [online]. Available from: http://www.ictvonline.org/virusTaxonomy.asp.

Iwasaki, M. et al. 2015. Efficient Interaction between Arenavirus Nucleoprotein (NP) and RNA-Dependent RNA Polymerase (L) Is Mediated by the Virus Nucleocapsid (NP-RNA) Template. *Journal of virology*. **89**(10),pp.5734–5738.

Jácamo, R. et al. 2003. Tacaribe virus Z protein interacts with the L polymerase protein to inhibit viral RNA synthesis. *Journal of virology*.

77(19),pp.10383–93.

Jackson, B.R. et al. 2011. An Interaction between KSHV ORF57 and UIF Provides mRNA-Adaptor Redundancy in Herpesvirus Intronless mRNA Export. *PLoS Pathog.* **7**(7),p.e1002138.

Jackson, R.J. et al. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature reviews. Molecular cell biology.* **11**(2),pp.113–127.

Jamieson, D.J. et al. 2006. Lymphocytic choriomeningitis virus: an emerging obstetric pathogen? *Am J Obstet Gynecol.* **194**(6),pp.1532–1536.

Jiang, X. et al. 2013. Structures of arenaviral nucleoproteins with triphosphate dsRNA reveal a unique mechanism of immune suppression. *The Journal of biological chemistry.* **288**(23),pp.16949–59.

Kahvejian, A. et al. 2012. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes & development.* **19**(1),pp.104–13.

Kentsis, a et al. 2001. The RING domains of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E. *Journal of molecular biology.* **312**(4),pp.609–623.

Kentsis, A. et al. 2002. Control of biochemical reactions through supramolecular RING domain self-assembly. *Proceedings of the National Academy of Sciences of the United States of America.* **99**(24),pp.15404–15409.

Kernéis, S. et al. 2009. Prevalence and risk factors of Lassa seropositivity in inhabitants of the forest region of Guinea: a cross-sectional study. *PLoS neglected tropical diseases.* **3**(11),p.e548.

Khachatoorian, R. et al. 2014. The NS5A-binding heat shock proteins HSC70 and HSP70 play distinct roles in the hepatitis C viral life cycle. *Virology.* **454-455**,pp.118–27.

Kimball, S.R. 1999. Eukaryotic initiation factor eIF2. *Int J Biochem Cell Biol.* **31**(1),pp.25–29.

Knust, B. et al. 2014. Lymphocytic choriomeningitis virus in employees and mice at multipremises feeder-rodent operation, United States, 2012. *Emerging infectious diseases.* **20**(2),pp.240–7.

Kranzusch, P.J. et al. 2010. Assembly of a functional Machupo virus polymerase complex. *Proceedings of the National Academy of Sciences of the United States of America.* **107**(46),pp.20069–20074.

Kranzusch, P.J. and Whelan, S.P.J. 2011. Arenavirus Z protein controls viral RNA synthesis by locking a polymerase-promoter complex. *Proceedings of the National Academy of Sciences of the United States of America.* **108**(49),pp.19743–8.

Kumar, N. et al. 2012. Characterization of virulence-associated determinants in the envelope glycoprotein of Pichinde virus. *Virology*. **433**(1),pp.97–103.

de la Torre, J.C. 2009. Molecular and cell biology of the prototypic arenavirus LCMV: implications for understanding and combating hemorrhagic fever arenaviruses. *Annals of the New York Academy of Sciences*. **1171 Suppl**(Lcmv),pp.E57–64.

Lahaye, X. et al. 2012. Hsp70 protein positively regulates rabies virus infection. *Journal of virology*. **86**(9),pp.4743–51.

Lalis, A. et al. 2012. The impact of human conflict on the genetics of Mastomys natalensis and Lassa virus in West Africa. *PLoS ONE*. **7**(5),p.e37068.

Lau, L.L. et al. 1994. Cytotoxic T-cell memory without antigen. *Nature*. **369**(6482),pp.648–652.

Lavergne, A. et al. 2015. Patawa Virus, a New Arenavirus Hosted by Forest Rodents in French Guiana. *EcoHealth*. **12**(2),pp.339–346.

Lefebvre, A.K. et al. 2006. Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *Journal of Biological Chemistry*. **281**(32),pp.22917–22932.

Lehmann, M. et al. 2014. Role of the C terminus of Lassa virus L protein in viral mRNA synthesis. *Journal of virology*. **88**(15),pp.8713–7.

Lele, S.M. et al. 2003. Pathology of Whitewater Arroyo Viral Infection in the White-throated Woodrat (Neotoma albigula). *Journal of Comparative Pathology*. **128**(4),pp.289–292.

Leung, W.-C. and Rawls, W.E. 1977. Virion-associated ribosomes are not required for the replication of Pichinde virus. *Virology*. **81**(1),pp.174–176.

Levy, D.E. and Garcia-Sastre, A. 2001. The virus battles: IFN induction of the antiviral state and mechanisms of viral evasion. *Cytokine & growth factor reviews*. **12**(2-3),pp.143–156.

Li, K. et al. 2015. Isolation and characterization of a novel arenavirus harbored by Rodents and Shrews in Zhejiang province, China. *Virology*. **476**,pp.37–42.

Linero, F. et al. 2013. Participation of eIF4F complex in Junin virus infection: blockage of eIF4E does not impair virus replication. *Cellular Microbiology*. **15**(10),pp.1766–1782.

Loureiro, M.E. et al. 2011. Molecular Determinants of Arenavirus Z Protein Homo-Oligomerization and L Polymerase Binding. *Journal of Virology* . **85** (23),pp.12304–12314.

Macneil, A. et al. 2012. Solid organ transplant-associated lymphocytic choriomeningitis, United States, 2011. *Emerging infectious diseases*. **18**(8),pp.1256–1262.

Manara, a. R. et al. 2012. Donation after circulatory death. *British Journal of*

Anaesthesia. **108**(SUPPL. 1),pp.108–121.

Martin, V.K. et al. 2010. Investigation of clade B New World arenavirus tropism by using chimeric GP1 proteins. *Journal of virology.* **84**(2),pp.1176–82.

Martinez-Sobrido, L. et al. 2009. Identification of amino acid residues critical for the anti-interferon activity of the nucleoprotein of the prototypic arenavirus lymphocytic choriomeningitis virus. *Journal of virology.* **83**(21),pp.11330–11340.

Martinez-Sobrido, L. et al. 2006. Inhibition of the type I interferon response by the nucleoprotein of the prototypic arenavirus lymphocytic choriomeningitis virus. *Journal of virology.* **80**(18),pp.9192–9199.

Martínez-Sobrido, L. et al. 2007. Differential inhibition of type I interferon induction by arenavirus nucleoproteins. *Journal of virology.* **81**(22),pp.12696–703.

Mate, S.E. et al. 2015. Molecular Evidence of Sexual Transmission of Ebola Virus. *The New England journal of medicine.* **373**(25),pp.2448–54.

Mateyak, M.K. and Kinzy, T.G. 2010. eEF1A: thinking outside the ribosome. *The Journal of biological chemistry.* **285**(28),pp.21209–13.

Matloubian, M. et al. 1990. Genetic basis of viral persistence: single amino acid change in the viral glycoprotein affects ability of lymphocytic choriomeningitis virus to persist in adult mice. *The Journal of experimental medicine.* **172**(4),pp.1043–1048.

Mayer, M.P. and Bukau, B. 2005. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences.* **62**(6),pp.670–684.

McEvoy, G.K. 2005. *American Hospital Formulary Service - Drug Information 2003.* Bethesda: MD: American Society of Health-System Pharmacists, Inc. 2005.

MeSH 2015. Ribavirin. *Medical Subject Headings* [online]. Available from: http://www.ncbi.nlm.nih.gov/mesh/68012254 [Accessed December 5, 2015].

Meyer, B.J. and Southern, P.J. 1993. Concurrent sequence analysis of 5' and 3' RNA termini by intramolecular circularization reveals 5' nontemplated bases and 3' terminal heterogeneity for lymphocytic choriomeningitis virus mRNAs. *Journal of virology.* **67**(5),pp.2621–2627.

Meyer, B.J. and Southern, P.J. 1994. Sequence heterogeneity in the termini of lymphocytic choriomeningitis virus genomic and antigenomic RNAs. *Journal of virology.* **68**(11),pp.7659–7664.

Mikesch, J.H. et al. 2010. Discovery of novel transcriptional and epigenetic targets in APL by global ChIP analyses: Emerging opportunity and challenge. *Cancer Cell.* **17**(2),pp.112–114.

Milazzo, M.L. et al. 2010. Antibodies to tacaribe serocomplex viruses (family

Arenaviridae, genus Arenavirus) in cricetid rodents from New Mexico, Texas, and Mexico. *Vector borne and zoonotic diseases.* **10**(6),pp.629–37.

Milazzo, M.L. et al. 2011. Novel arenavirus infection in humans, United States. *Emerging Infectious Diseases.* **17**(8),pp.1417–1420.

Milazzo, M.L. et al. 2011. Transmission of guanarito and pirital viruses among wild rodents, Venezuela. *Emerging Infectious Diseases.* **17**(12),pp.2209–2215.

Mir, M. a and Panganiban, A.T. 2008. A protein that replaces the entire cellular eIF4F complex. *The EMBO journal.* **27**(23),pp.3129–3139.

Mocarski Jr, E.S. 2007. Comparative analysis of herpesvirus-common proteins *In*: *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.*

Mohamed, M.R. et al. 2001. Interaction between the J3R subunit of vaccinia virus poly(A) polymerase and the H4L subunit of the viral RNA polymerase. *Virology.* **280**(1),pp.143–52.

Monath, T.P. et al. 1974. Lassa virus isolation from Mastomys natalensis rodents during an epidemic in Sierra Leone. *Science.* **185**(4147),pp.263–265.

Moradpour, D. and Penin, F. 2013. Hepatitis C virus proteins: from structure to function. *Curr Top Microbiol Immunol.* **369**,pp.113–142.

Moraz, M.L. et al. 2013. Cell entry of Lassa virus induces tyrosine phosphorylation of dystroglycan. *Cell Microbiol.* **15**(5),pp.689–700.

Morin, B. et al. 2010. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. *PLoS pathogens.* **6**(9),p.e1001038.

Muckenfuss, R. et al. 1934. ETiology of the 1933 epidemic of encephalitis. *Journal of the American Medical Association.* **103**(10),pp.731–733.

Murphy, F.A. 2008. Emerging zoonoses: the challenge for public health and biodefense. *Prev Vet Med.* **86**(3-4),pp.216–223.

Murphy, F.A. and Whitfield, S.G. 1975. Morphology and morphogenesis of arenaviruses. *Bulletin of the World Health Organization.* **52**(4-5 6),pp.409–419.

Noonan, F.P. et al. 1979. Early pregnancy factor is immunosuppressive. *Nature.* **278**(5705),pp.649–651.

Nunberg, J.H. and York, J. 2012. The curious case of arenavirus entry, and its inhibition. *Viruses.* **4**(1),pp.83–101.

Oberer, M. et al. 2005. Structural basis for the enhancement of eIF4A helicase activity by eIF4G. *Genes and Development.* **19**,pp.2212–2223.

Ogbu, O. et al. 2007. Lassa fever in West African sub-region: An overview. *Journal of Vector Borne Diseases.* **44**(1),pp.1–11.

185

Okokhere, P.O. et al. 2009. Sensorineural hearing loss in Lassa fever: two case reports. *Journal of medical case reports.* **3**,p.36.

Olugasa, B.O. and Dogba, J.B. 2015. Mapping of Lassa fever cases in post-conflict Liberia, 2008-2012: a descriptive and categorical analysis of age, gender and seasonal pattern. *Ann Afr Med.* **14**(2),pp.120–122.

Ong, S.-E. et al. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP.* **1**,pp.376–386.

Oshiumi, H. et al. 2013. Multi-step regulation of interferon induction by hepatitis C virus. *Arch Immunol Ther Exp (Warsz).* **61**(2),pp.127–138.

Palacios, G. et al. 2008. A new arenavirus in a cluster of fatal transplant-associated diseases. *The New England journal of medicine.* **358**(10),pp.991–998.

Panavas, T. et al. 2009. SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. *Methods Mol Biol.* **497**,pp.303–317.

Pantaloni, D. et al. 2001. Mechanism of actin-based motility. *Science (New York, N.Y.).* **292**(5521),pp.1502–1506.

Paweska, J.T. et al. 2009. Nosocomial outbreak of novel arenavirus infection, southern Africa. *Emerging infectious diseases.* **15**(10),pp.1598–1602.

Pedersen, I.R. and Konigshofer, E.P. 1976. Characterization of ribonucleoproteins and ribosomes isolated from lymphocytic choriomeningitis virus. *Journal of virology.* **20**(1),pp.14–21.

Perez, M. et al. 2004. Myristoylation of the RING finger Z protein is essential for arenavirus budding. *Journal of virology.* **78**(20),pp.11443–11448.

Perez, M. and de La Torre, J.C. 2003. Characterization of the genomic promoter of the prototypic arenavirus lymphocytic choriomeningitis virus. *Journal of Virology.* **77**(2),pp.1184–1194.

Personal Experience, _ 2015. *Personal Experience.* Makeni, Sierra Leone.

Peterson, A.T. et al. 2014. Mapping transmission risk of Lassa fever in West Africa: the importance of quality control, sampling bias, and error weighting. *PLoS ONE.* **9**(8),p.e100711.

Pinschewer, D.D. et al. 2003. Role of the virus nucleoprotein in the regulation of lymphocytic choriomeningitis virus transcription and RNA replication. *Journal of virology.* **77**(6),pp.3882–7.

Prévôt, D. et al. 2003. Conducting the initiation of protein synthesis: The role of eIF4G. *Biology of the Cell.* **95**(3-4),pp.141–156.

Pythoud, C. et al. 2012. Arenavirus nucleoprotein targets interferon regulatory factor-activating kinase IKKepsilon. *Journal of virology.* **86**(15),pp.7728–7738.

Pythoud, C. et al. 2012. Arenavirus nucleoprotein targets interferon regulatory

factor-activating kinase IKKε. *Journal of virology.* **86**(15),pp.7728–38.

Qi, X. et al. 2010. Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature.* **468**(7325),pp.779–83.

Radoshitzky, S.R. et al. 2007. Transferrin receptor 1 is a cellular receptor for New World haemorrhagic fever arenaviruses. *Nature.* **446**(7131),pp.92–6.

Ramakrishnan, V. 2002. Ribosome structure and the mechanism of translation. *Cell.* **108**(4),pp.557–572.

Reynard, S. et al. 2014. Exonuclease domain of the Lassa virus nucleoprotein is critical to avoid RIG-I signaling and to inhibit the innate immune response. *Journal of virology.* **88**(23),pp.13923–13927.

Richards, P. et al. 2015. Social pathways for Ebola virus disease in rural Sierra Leone, and some implications for containment. *PLoS neglected tropical diseases.* **9**(4),p.e0003567.

Rodrigo, W.W.S.I. et al. 2012. Arenavirus nucleoproteins prevent activation of nuclear factor kappa B. *Journal of virology.* **86**(15),pp.8185–97.

Rogers, G.W. et al. 2002. eIF4A: the godfather of the DEAD box helicases. *Progress in nucleic acid research and molecular biology.* **72**,pp.307–331.

Rowe, W.P. et al. 1970. Arenoviruses: proposed name for a newly defined virus group. *Journal of virology.* **5**(5),pp.651–652.

Ruggiero, H.A. et al. 1986. [Treatment of Argentine hemorrhagic fever with convalescent's plasma. 4433 cases]. *Presse medicale.* **15**(45),pp.2239–2242.

Russier, M. et al. 2012. Immune responses and Lassa virus infection. *Viruses.* **4**(11),pp.2766–85.

Saito, T. and Gale Jr., M. 2008. Differential recognition of double-stranded RNA by RIG-I-like receptors in antiviral immunity. *The Journal of experimental medicine.* **205**(7),pp.1523–1527.

Salas, R. et al. 1991. *Venezuelan haemorrhagic fever.*

Salvato, M. et al. 1989. The primary structure of the lymphocytic choriomeningitis virus L gene encodes a putative RNA polymerase. *Virology.* **169**(2),pp.377–384.

Sayce, A.C. et al. 2010. Targeting a host process as an antiviral approach against dengue virus. *Trends in Microbiology.* **18**(7),pp.323–330.

Schafer, I.J. et al. 2014. Notes from the field: a cluster of lymphocytic choriomeningitis virus infections transmitted through organ transplantation - Iowa, 2013. *MMWR. Morbidity and mortality weekly report.* **63**(11),p.249.

Scheper, G.C. and Proud, C.G. 2002. Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation? *European Journal of Biochemistry.* **269**(22),pp.5350–5359.

Schlie, K. et al. 2010. Viral protein determinants of Lassa virus entry and release from polarized epithelial cells. *Journal of virology*. **84**(7),pp.3178–3188.

Sewlall, N.H. et al. 2014. Clinical features and patient management of lujo hemorrhagic Fever. *PLoS neglected tropical diseases*. **8**(11),p.e3233.

Shaffer JG; Grant DS; Schieffelin, J. 2014. After the 'blood diamond' conflict: lassa fever in Sierra Leone. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*. **59**(12),pp.iii–iv.

Shaffer, J.G. et al. 2014. Lassa fever in post-conflict sierra leone. *PLoS neglected tropical diseases*. **8**(3),p.e2748.

Shah, W.A. et al. 2006. Role of non-raft cholesterol in lymphocytic choriomeningitis virus infection via alpha-dystroglycan. *The Journal of general virology*. **87**(Pt 3),pp.673–678.

Shao, J. et al. 2015. Human hemorrhagic Fever causing arenaviruses: molecular mechanisms contributing to virus virulence and disease pathogenesis. *Pathogens*. **4**(2),pp.283–306.

Shtanko, O. et al. 2010. A role for the C-terminus of Mopeia virus nucleoprotein in its incorporation into Z-induced virus-like particles. *J Virol*.

Simmonds, P. et al. 2004. Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: Implications for virus evolution and host persistence. *RNA*. **10**(9),pp.1337–1351.

Sogoba, N. et al. 2012. Lassa Fever in West Africa: Evidence for an Expanded Region of Endemicity. *Zoonoses and Public Health*. **59**(SUPPL.2),pp.43–47.

Souter, M. and Van Norman, G. 2010. Ethical controversies at end of life after traumatic brain injury: defining death and organ donation. *Critical care medicine*. **38**(9 Suppl),pp.S502–S509.

Stenglein, M.D. et al. 2012. Identification, characterization, and in vitro culture of highly divergent arenaviruses from boa constrictors and annulated tree boas: candidate etiological agents for snake inclusion body disease. *mBio*. **3**(4),pp.e00180–12.

Stephenson, E.H. et al. 1984. Effect of environmental factors on aerosol-induced Lassa virus infection. *Journal of medical virology*. **14**(4),pp.295–303.

Strecker, T. et al. 2006. The role of myristoylation in the membrane association of the Lassa virus matrix protein Z. *Virology journal*. **3**,p.93.

Studier, F.W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification*. **41**(1),pp.207–234.

Surtees, R. 2014. Structural and cellular biology of the nairovirus

nucleocapsid protein.

Taguwa, S. et al. 2015. Defining Hsp70 Subnetworks in Dengue Virus Replication Reveals Key Vulnerability in Flavivirus Infection. *Cell.* **163**(5),pp.1108–1123.

Takagi, M. et al. 2005. Regulation of p53 translation and induction after DNA damage by ribosomal protein L26 and nucleolin. *Cell.* **123**(1),pp.49–63.

Talwani, M. 2002. *The Orinoco Heavy Oil Belt in Venezuela (Or Heavy Oil to the Rescue?).*

Tani, H. et al. 2014. Analysis of Lujo virus cell entry using pseudotype vesicular stomatitis virus. *Journal of virology.* **88**(13),pp.7317–30.

Tanner, S.J. et al. 2014. Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America.* **111**(4),pp.1580–5.

Teijaro, J.R. et al. 2013. Persistent LCMV infection is controlled by blockade of type I interferon signaling. *Science.* **340**(6129),pp.207–211.

Tortorici, M. a et al. 2001. Arenavirus nucleocapsid protein displays a transcriptional antitermination activity in vivo. *Virus research.* **73**(1),pp.41–55.

United Nations Populations Division 2015. *World Population Prospects.*

Urata, S. et al. 2006. Cellular factors required for Lassa virus budding. *Journal of virology.* **80**(8),pp.4191–5.

Urata, S. and Yasuda, J. 2012. Molecular mechanism of arenavirus assembly and budding. *Viruses.* **4**(10),pp.2049–79.

US National Library of Medicine 2015. Ribavirin. *NIH* [online]. Available from: https://www.nlm.nih.gov/medlineplus/druginfo/meds/a605018.html [Accessed December 5, 2015].

Vende, P. et al. 2000. Efficient translation of rotavirus mRNA requires simultaneous interaction of NSP3 with the eukaryotic translation initiation factor eIF4G and the mRNA 3' end. *Journal of virology.* **74**(15),pp.7064–71.

Volpon, L. et al. 2010. Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E. *Proceedings of the National Academy of Sciences of the United States of America.* **107**(12),pp.5441–6.

Wang, Y. et al. 2010. A Theoretical Study of the Separation Principle in Size Exclusion Chromatography. *Macromolecules.* **43**,pp.1651–1659.

Weaver, S.C. et al. 2000. Guanarito virus (Arenaviridae) isolates from endemic and outlying localities in Venezuela: sequence comparisons among and within strains isolated from Venezuelan hemorrhagic fever patients and rodents. *Virology.* **266**(1),pp.189–195.

Weill, L. et al. 2012. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nature Structural & Molecular Biology.* **19**(6),pp.577–585.

Wells, S.E. et al. 2012. Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell.* **2**(1),pp.135–140.

Welnowska, E. et al. 2011. Translation of Viral mRNA without Active eIF2: The Case of Picornaviruses. *PLoS One.* **6**(7),p.e22230.

Welsch, S. et al. 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell host & microbe.* **5**(4),pp.365–375.

Wherry, E.J. et al. 2003. Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. *Journal of virology.* **77**(8),pp.4911–27.

WHO 2015. WHO Blueprint for R&D Readiness. Available from: http://www.who.int/medicines/ebola-treatment/WHO-list-of-top-emerging-diseases/en/.

Wilda, M. et al. 2008. Mapping of the Tacaribe Arenavirus Z-Protein Binding Sites on the L Protein Identified both Amino Acids within the Putative Polymerase Domain and a Region at the N Terminus of L That Are Critically Involved in Binding. *Journal of Virology* . **82** (22),pp.11454–11460.

Wilson, E.B. et al. 2013. Blockade of chronic type I interferon signaling to control persistent LCMV infection. *Science.* **340**(6129),pp.202–207.

Witkowski, P.T. et al. 2015. Novel Arenavirus Isolates from Namaqua Rock Mice, Namibia, Southern Africa. *Emerging infectious diseases.* **21**(7),pp.1213–1216.

World Health Organisation 2015. Ebola data and statistics.

World Health Organisation 2016. *WHO to convene an International Health Regulations Emergency Committee on Zika virus and observed increase in neurological disorders and neonatal malformations.*

Wright, A.J. and Fishman, J. a 2014. Central Nervous System Syndromes in Solid Organ Transplant Recipients. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* **59**,pp.1–11.

Yanagiya, A. et al. 2012. Translational Homeostasis via the mRNA Cap-Binding Protein, eIF4E. *Molecular Cell.* **46**(6),pp.847–858.

Yanguez, E. et al. 2011. Functional impairment of eIF4A and eIF4G factors correlates with inhibition of influenza virus mRNA translation. *Virology.* **413**(1),pp.93–102.

York, J. et al. 2008. pH-induced activation of arenavirus membrane fusion is antagonized by small-molecule inhibitors. *Journal of virology.*

82(21),pp.10932–9.

York, J. et al. 2004. The signal peptide of the Junín arenavirus envelope glycoprotein is myristoylated and forms an essential subunit of the mature G1-G2 complex. *Journal of virology*. **78**(19),pp.10783–92.

York, J. and Nunberg, J.H. 2006. Role of the stable signal peptide of Junín arenavirus envelope glycoprotein in pH-dependent membrane fusion. *Journal of virology*. **80**(15),pp.7775–80.

Yun, N.E. et al. 2008. Pathogenesis of XJ and Romero strains of Junin virus in two strains of guinea pigs. *The American journal of tropical medicine and hygiene*. **79**(2),pp.275–282.

Yun, N.E. and Walker, D.H. 2012. Pathogenesis of Lassa fever. *Viruses*. **4**(10),pp.2031–48.

Zapata, J.C. and Salvato, M.S. 2013. Arenavirus variations due to host-specific adaptation. *Viruses*. **5**(1),pp.241–278.

Zhang, L. et al. 2001. Reassortant analysis of guinea pig virulence of pichinde virus variants. *Virology*. **290**(1),pp.30–38.

Zinkernagel, R.M. and Doherty, P.C. 1974. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature*. **248**,pp.701–702.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*. **31**(13),pp.3406–3415.

8 Appendix I

Sequences of LUJV NP-EGFP, his-SUMO-LUJV NP; his-SUMO-PICV NP; NPEG and NPEG 3'Δ.

LUJV NP-EGFP

```
ATGTCCCAGTCAAAGGAAGTCAAGTCATTTCTGTGGCTCCAAACACTGCGCAGAGAACTCT
CTCCTTTCTGTACCGATGTGAGGGCAAAGGTGGTGAAGGACGCCGTGTCCCTGATCAACGG
ACTGGACTTCTCAATGGTGTCTGATGTGCAACGCCTCATGCGGAAGGACAAGCGGAACGAT
GAGGACCTCATGAAACTGAGGGAGCTGAATCAGACTGTTGACGGACTGGTTGACCTGAAGT
CTTCCAACAAGAAGAATAGAGTCGGCGTGGGAAAACTGACATCAGACGAACTGATGATACT
GGCCACCGACCTGGAGAAGCTGAAGAAGAAGGTGACCAGGACCGAAGCACGCGGACCAGGT
GTTTACAGAGGCAACCTCTCTCAGGACCAGCTGGGAAGGCGCTCAGAGCTGCTGAATATGA
TCGGGATGGGCACACCACGGCCAACAAGAAATACAGTGGTGCGCGTCTGGGATGTGAAGGA
CAGTTCCCTCCTGAACAACCAGTTCGGGACAATGCCCTCCCTGACTCTGGCTTGTCTGACC
CGCCAAACCCGCGTGGACCTGAACGATTCCGTCCAGGCTTGCGTGGATCTGGGCCTCATCT
ACACCGCCAAGTTCCCTAACATGGATGATCTCGACAAGCTCAAGAATAAGCATCCAGTGCT
GGACTACGTGAGCAACTGCGATTCTGCCATTAACATCAGCGGGTATAACCTGTCTCTCGCC
TCCCTGGTGAAGGCTGGCGCCAGCCTGATGAAAGGAGGTGACATGCTGGAAACCATTGAAC
TGAATTCTAGAAACATCGACGACGTGATAAAGGCTACTCTCACCGCTCGCAACAAGGTGCA
GATGTTTGTTTCTGAGGTGCCTGGAGAGAGGAACCCTTATGAGAACCTGCTGTATAAGATT
TGCCTGAGCGGCGAGGGGTGGCCCTACATCTCATCACGGACCTCTATTAAAGGCCGGAGTT
GGGATAACACTGTGATCGATATGACCCCTAAAGATCCCACACCTCCCCAGAATGAAAGAGC
CAAGGCTCCCCATCAATTTCCTGTGGGCGTGAGCTTCAGTCAGTCTCAGCTCCTCGATGAC
ATTATGAAGAATCTCAATCCAAAAGGCAGGACATGGATGGACATCGAGGGACGGCCTGATG
ACCCTGTCGAAATCGCCATCTTCCAGCCTGAGGAAAGACTCTGCCTCCACTTCTATCGGGA
GCCCACCGACCAAAAGCAGTTCAAGAATGACAGCAAATATTGTCATGGTATGGATTTCACC
CAGCTCTGCAGCACACAGCCTGGGCTCACTACTGCCGTGCTGGAGAGACTGCCTCTGGGCA
TGGTTATTACATGTCAGGGGAAGGACGATATAGAAAAGCTGCTGCATTCTCAAGGAAGGAG
AGATGTCAAGTTCATCGACATACAGATGAGCAAAGAAGCTTCCAGAAAGTTTGAAGACCAG
GTGTGGGACTCTTATAAAACCTTTTGCAATCAGCACACTGGCATTGTCGTTACAAAATCAA
AGAAAGGCAAGAAGGAGATTACACCACACTGTGCACTCATGGATTGCATCATGTACGAATC
TGCTGTTAATGGCCAGCTCTATCAGGAACCCATCAGGAACCTGCTGCCAGCTGACATGATC
TTTCGCACTGCCGCTAAGCTGAGCCTGGCAGCCGTGAGCAAGGGCGAGGAGCTCTTCACCG
GCGTGGTGCCCATTCTCGTGGAGCTGGACGGGGACGTCAACGGCCACAAATTCTCTGTTTC
CGGAGAAGGAGAGGGAGACGCAACATATGGCAAACTGACCCTCAAATTCATCTGCACCACT
GGCAAGCTGCCAGTGCCCTGGCCCACACTGGTGACCACACTCACATATGGGGTCCAGTGTT
TTAGCCGGTATCCCGACCATATGAAACAGCACGACTTCTTTAAGTCCGCAATGCCAGAGGG
TTATGTCCAGGAACGCACAATATTCTTTAAGGATGATGGTAATTACAAAACACGGGCCGAG
GTCAAATTTGAGGGAGACACCCTCGTGAACAGAATCGAGCTCAAAGGAATAGACTTCAAGG
AAGACGGGAATATTCTGGGCCATAAACTGGAGTACAACTACAATTCACATAATGTTTACAT
CATGGCTGATAAGCAGAAGAATGGCATCAAGGTTAACTTCAAGATTAGACATAATATAGAG
GACGGCTCCGTTCAGCTCGCCGACCACTATCAGCAGAACACACCCATAGGAGATGGCCCTG
TGCTGCTGCCAGACAATCACTATCTGAGCACCCAGAGTGCCCTGTCCAAAGACCCCAATGA
GAAGAGAGATCACATGGTGCTGCTGGAGTTCGTGACAGCCGCTGGGATTACCCTCGGTATG
GATGAGCTGTACA
```

CTG: Final LUJV NP residue.

GTG: Initial EGFP residue.

LUJV NP sequence corresponds to accession number NC_012776. EGFP sequence is obtained from parental EGFP N1 expression plasmid.

His-SUMO-LUJV NP

```
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATA
TGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGA
GACTCACATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAG
ACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACT
CCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGA
CATGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGATCCATGTCCCAG
TCAAAGGAAGTCAAGTCATTTCTGTGGCTCCAAACACTGCGCAGAGAACTCTCTCCTTTCT
GTACCGATGTGAGGGCAAAGGTGGTGAAGGACGCCGTGTCCCTGATCAACGGACTGGACTT
CTCAATGGTGTCTGATGTGCAACGCCTCATGCGGAAGGACAAGCGGAACGATGAGGACCTC
ATGAAACTGAGGGAGCTGAATCAGACTGTTGACGGACTGGTTGACCTGAAGTCTTCCAACA
AGAAGAATAGAGTCGGCGTGGGAAAACTGACATCAGACGAACTGATGATACTGGCCACCGA
CCTGGAGAAGCTGAAGAAGAAGGTGACCAGGACCGAAGCACGCGGACCAGGTGTTTACAGA
GGCAACCTCTCTCAGGACCAGCTGGGAAGGCGCTCAGAGCTGCTGAATATGATCGGGATGG
GCACACCACGGCCAACAAGAAATACAGTGGTGCGCGTCTGGGATGTGAAGGACAGTTCCCT
CCTGAACAACCAGTTCGGGACAATGCCCTCCCTGACTCTGGCTTGTCTGACCCGCCAAACC
CGCGTGGACCTGAACGATTCCGTCCAGGCTTGCGTGGATCTGGGCCTCATCTACACCGCCA
AGTTCCCTAACATGGATGATCTCGACAAGCTCAAGAATAAGCATCCAGTGCTGGACTACGT
GAGCAACTGCGATTCTGCCATTAACATCAGCGGGTATAACCTGTCTCTCGCCTCCCTGGTG
AAGGCTGGCGCCAGCCTGATGAAAGGAGGTGACATGCTGGAAACCATTGAACTGAATTCTA
GAAACATCGACGACGTGATAAAGGCTACTCTCACCGCTCGCAACAAGGTGCAGATGTTTGT
TTCTGAGGTGCCTGGAGAGAGGAACCCTTATGAAAACCTGCTGTATAAGATTTGCCTGAAC
GGCGAGGGGTGGCCCTACATCTCATCACGGACCTCTATTAAAGGCCGGAGTTGGGATAACC
CTGGGATCGAATTGACCCCTAAAGATCCCCACCCTCCCCAGAATGAAAAAGCCAAGGGTCC
CCATCAATTTCCTGTGGGCGTGAGCTTCAGTCAGTCTCAGCTCCTCGATGACATTATGAAG
AATCTCAATCCAAAAGGCAGGACATGGATGGACATCGAGGGACGGCCTGATGACCCTGTCG
AAATCGCCATCTTCCAGCCTGAGGAAAGACTCTGCCTCCACTTCTATCGGGAGCCCACCGA
CCAAAAGCAGTTCAAGAATGACAGCAAATATTGTCATGGTATGGATTTCACCCAGCTCTGC
AGCACACAGCCTGGGCTCACTACTGCCGTGCTGGAGAGACTGCCTCTGGGCATGGTTATTA
CATGTCAGGGGAAGGACGATATAGAAAAGCTGCTGCATTCTCAAGGAAGGAGAGATGTCAA
GTTCATCGACATACAGATGAGCAAAGAAGCTTCCAGAAAGTTTGAAGACCAGGTGTGGGAC
TCTTATAAACCTTTTGCAATCAGCACACTGGCATTGTCGTTACAAAATCAAAGAAAGGCA
AGAAGGAGATTACACCACACTGTGCACTCATGGATTGCATCATGTACGAATCTGCTGTTAA
TGGCCAGCTCTATCAGGAACCCATCAGGAACCTGCTGCCAGCTGACATGATCTTTCGCACT
GCCGCTAAGCTGAGCCTGGCTAGACTCGAGCACCACCACCACCACCACTGAGATCCGGCTG
CTAACAAAGCCCG
```

TCC: Final SUMO residue.

LUJV NP sequence corresponds to accession number NC_012776.

His-SUMO-PICV NP

```
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATA
TGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGA
GACTCACATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAG
ACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACT
CCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGA
```

CATGGAGGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTGGA<mark>TCC</mark>ATGTCCGAC
AATATCCCATCGTTCCGCTGGGTGCAATCCCTTAGGAGGGGCTTGTCCAACTGGACCCATC
CTGTGAAGGCTGATGTGCTGTCAGACACAAGAGCACTGTTGTCTGCTCTTGACTTTCACAA
AGTTGCTCAAGTCCAAAGAATGGTGCGCAAAGATAAGAGGACTGATTCTGATCTGACCAAG
CTGAGGGACATGAACAAAGAGGTTGATGCTCTGATGAATATGAGATCAGTCCAAAGAGATA
ATGTACTTAAAGTGGGAGGCTTAGCTAAAGAGGAGCTCATGGAGCTTGCATCTGATTTGGA
CAAGTTAAGGAAGAAAGTCACCAGAACCGAAGGCCTGTCTCAGCCTGGTGTCTATGAAGGC
AATCTTACAAACACTCAGCTGGAGCAAAGGGCAGAGATTCTCCGCTCGATGGGGTTTGCTA
ATGCCAGACCCGCAGGTAACAGAGATGGTGTTGTGAAGGTCTGGGACATCAAGGATAACAC
ACTATTGATTAATCAATTTGGATCAATGCCAGCCTTGACCATCGCCTGTATGACAGAGCAA
GGGGGTGAGCAGCTTAATGATGTTGTCCAAGCACTGAGTGCACTTGGTTTGCTTTACACTG
TCAAGTTCCCGAACATGACAGATCTGGAAAAGCTCACACAACAACACAGCGCCCTAAAAAT
CATCAGCCATGAGCCATCAGCCTTGAATATCTCTGGGTATAACCTTAGTCTGTCTGCAGCA
GTCAAAGCAGCTGCTTGTATGATTGATGGTGGCAACATGCTTGAGACCATCCAAGTGAAGC
CCTCTATGTTTAGCACCCTCATAAAAAGTCTACTGCAGATTAAGAACCGTGAGGGTATGTT
TGTGAGTACTACACCCGGGCAAAGGAACCCTTATGAAAACCTGCTGTATAAAATCTGTCTC
TCAGGGGATGGCTGGCCCTACATTGGCTCAAGATCCCAAGTTCAAGGGAGGGCTTGGGACA
ACACCACTGTGGATTTGGACTCAAAGCCAAGTGCTATCCAGCCACCAGTAAGAAACGGTGG
ATCACCAGATCTCAAACAAATCCCTAAGGAAAAGAAGACACTGTTGTGTCCTCAATTCAG
ATGCTTGATCCAAGAGCCACTACGTGGATTGATATTGAGGGGACACCAAATGACCCGGTGG
AAATGGCCATCTACCAACCTGATACAGGCAACTACATACACTGTTATAGGTTTCCCCATGA
TGAAAAATCCTTCAAAGAGCAGAGCAAGTATTCACATGGTCTCCTTTTAAAGGACTTGGCT
GATGCTCAACCGGGTCTGATCTCCTCAATCATCAGACATTTGCCCCAGAACATGGTTTTCA
CTGCTCAAGGTTCAGATGATATAATCAGACTGTTCGAAATGCATGGAAGAAGAGATCTAAA
AGTACTTGACGTGAAGCTCAGTGCAGAGCAGGCACGCACCTTTGAGGATGAGATCTGGGAG
AGATACAACCAACTCTGCACCAAGCACAAGGGCTTAGTCATAAAGAAGAAGAAGAAAGGAG
CTGTACAAACCACTGCAAACCCCCACTGTGCATTGCTTGACACCATCATGTTTGATGCAAC
AGTGACAGGCTGGGTCAGAGATCAGAAACCCATGAGATGTTTACCTATTGACACACTGTAC
AGGAACAACACAGACTTGATCAACCTCTGA

<mark>TCC</mark>: Final SUMO residue.

PICV NP sequence corresponds to accession number AF081555.1

NPEG – LCMV NP *Gaussia* mRNA plasmid

<mark>TTAATACGACTCACTATA</mark>GG<mark>CGCACAGTGGATCCTAGGCATTTGATTGCGCATTTGTCTTG
AGAAACCATTGAGCAACAAG</mark><mark>ATGGGAGTCAAAGTTCTGTTTGCCCTGATCTGCATCGCTGT
GGCCGAGGCCAAGCCCACCGAGAACAACGAAGACTTCAACATCGTGGCCGTGGCCAGCAAC
TTCGCGACCACGGATCTCGATGCTGACCGCGGGAAGTTGCCCGGCAAGAAGCTGCCGCTGG
AGGTGCTCAAAGAGATGGAAGCCAATGCCCGGAAAGCTGGCTGCACCAGGGGCTGTCTGAT
CTGCCTGTCCCACATCAAGTGCACGCCCAAGATGAAGAAGTTCATCCCAGGACGCTGCCAC
ACCTACGAAGGCGACAAAGAGTCCGCACAGGGCGGCATAGGCGAGGCGATCGTCGACATTC
CTGAGATTCCTGGGTTCAAGGACTTGGAGCCCATGGAGCAGTTCATCGCACAGGTCGATCT
GTGTGTGGACTGCACAACTGGCTGCCTCAAAGGGCTTGCCAACGTGCAGTGTTCTGACCTG
CTCAAGAAGTGGCTGCCGCAACGCTGTGCGACCTTTGCCAGCAAGATCCAGGGCCAGGTGG
ACAAGATCAAGGGGGCCGGTGGTGAC</mark>ACTCTAAGACCCTCTGGGCCTCCC<mark style="background-color:red">TGA</mark>CTCTCCAC
<mark style="background-color:lime">CTCTTTCGAGGTGGAGAGTCAGGGAGGC</mark>GGGTCGGCATGGCATCTCCACCTCCTCGCGGTC
CGACCTGGGCATCCGAAGGAGGACGCACGTCCACTCGGATGGCTAAGGGAGGGGATCCAGCT
CGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCA
ATAA<mark style="background-color:olive">CTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGC</mark>

Yellow – T7 Promoter; GG non viral; Blue – LCMV S seg NP 5' UTR (; Pink – *Gaussia* Lucif; Green – LCMV S seg NP genome 3' UTR; Grey – HDV ribozyme; no colour – junk; Sag Paneer – T7 term

LCMV 5' and 3' sequences based on sequence corresponding with accession number AY847350.1, with 3' sequences corresponding to those observed by Meyer *et. al* (Meyer & Southern, 1994). *Gaussia* luciferase sequence corresponds to accession number AY015993.1.

3'Δ NPEG

```
TTAATACGACTCACTATAGGCGCACAGTGGATCCTAGGCATTTGATTGCGCATTTGTCTTG
AGAAACCATTGAGCAACAAGATGGGAGTCAAAGTTCTGTTTGCCCTGATCTGCATCGCTGT
GGCCGAGGCCAAGCCCACCGAGAACAACGAAGACTTCAACATCGTGGCCGTGGCCAGCAAC
TTCGCGACCACGGATCTCGATGCTGACCGCGGGAAGTTGCCCGGCAAGAAGCTGCCGCTGG
AGGTGCTCAAAGAGATGGAAGCCAATGCCCGGAAAGCTGGCTGCACCAGGGGCTGTCTGAT
CTGCCTGTCCCACATCAAGTGCACGCCCAAGATGAAGAAGTTCATCCCAGGACGCTGCCAC
ACCTACGAAGGCGACAAAGAGTCCGCACAGGGCGGCATAGGCGAGGCGATCGTCGACATTC
CTGAGATTCCTGGGTTCAAGGACTTGGAGCCCATGGAGCAGTTCATCGCACAGGTCGATCT
GTGTGTGGACTGCACAACTGGCTGCCTCAAAGGGCTTGCCAACGTGCAGTGTTCTGACCTG
CTCAAGAAGTGGCTGCCGCAACGCTGTGCGACCTTTGCCAGCAAGATCCAGGGCCAGGTGG
ACAAGATCAAGGGGGCCGGTGGTGACTGACTCGAGGGGTCGGCATGGCATCTCCACCTCCT
CGCGGTCCGACCTGGGCATCCGAAGGAGGACGCACGTCCACTCGGATGGCTAAGGGAGGGA
TCCAGCTCGGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCG
CTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGC
```

Yellow – T7 Promoter; GG non viral; Blue – LCMV S seg NP 5' UTR (; Pink – *Gaussia* Lucif;; Grey – HDV ribozyme; no colour – junk; Sag Paneer – T7 term

3'Δ NPEG

GenBank accession numbers for sequences used in neiugbour joining phylogenetic analysis of arenaviruses shown in figure 4:

ALLV CLHP2472 (AY216502, AY012687); AMAV BeAn70563 (AF512834); BCNV AVA0070039 (AY924390, AY922491), A0060209 (AY216503); CATV AVA0400135 (DQ865244), AVA0400212 (DQ865245); CHPV 810419 (EU, 260464, EU260463); CPXV BeAn119303 (AY216519, AF512832); DANV 0710-2678 (EU136039, EU136038); FLEV BeAn293022 (EU627611, AF512831); GTOV INH-95551 (AY358024, AF485258), CVH-960101 (AY497548); IPPYV DakAnB188d (DQ328878, DQ328877); JUNV MC2 (AY216507, D10072), XJ13 (AY358022, AY358023), CbalV4454

(DQ272266); LASV LP (AF181853), 803213 (AF181854), Weller (AY628206), AV (AY179171, AF246121), Z148 (AY628204, AY628205), Josiah (U73034, J043204), NL (AY179172, AY179173); LATV MARU10924 (EU627612, AF485259); LCMV Armstrong (AY847351), ARM53b (M20869), WE (AF004519, M22138), Marseille12 (DQ286932, DQ286931), M1 (AB261991); MACV Carvallo (AY619642, AY619643), Chicava (AY624354, AY624355), Mallele (AY619644, AY619645), MARU222688 (AY922407), 9530537 (AY571959); MOBV ACAR3080MRC5P2 (DQ328876, AY342390); MOPV AN20410 (AY772169, AY772170), Mozambique (DQ328875, DQ328874); NAAV AVD1240007 (EU123329); OLVV 3229-1 (AY216514, U34248); PARV 12056 (EU627613, AF485261); PICV (K02734), MunchiqueCoAn4763 (EF529745, EF529744), AN3739 (AF427517); PIRV VAV-488 (AY216505, AF277659); SABV SPH114202 (AY358026, U41071); SKTV AVD1000090 (EU123328); TAMV W10777 (EU627614, AF512828); TCRV (J04340, M20304); WWAV AV9310135 (AY924395, AF228063)

9 Appendix II

SILAC GFP-Trap immunoprecipitation

SILAC MS IP interactome analysis of EGFP and LUJV NP-EGFP

Protein IDs	Protein Name	Peptides	PEP	Ratio M/L normalized	Ratio H/L normalized	Intensity L	Intensity M	Intensity H
238890539	Nucleocapsid protein [Lujo virus]	12	1.95E-43	NaN	NaN	0	1385400	14801000
A6NFI3	Zinc finger protein 316	1	0.0071437	2.5558	0.69218	50811	322060	230000
O00148	ATP-dependent RNA helicase DDX39A	2	2.99E-05	NaN	NaN	0	0	930870
O00159	Unconventional myosin-Ic	10	4.09E-26	NaN	NaN	0	0	2953100
O15116	U6 snRNA-associated Sm-like protein LSm1	1	0.037709	NaN	NaN	5368400	0	0
O15144	Actin-related protein 2/3 complex subunit 2	1	0.014107	NaN	NaN	0	0	0
O15260	Surfeit locus protein 4	1	0.030274	NaN	NaN	0	0	212770
O15523	ATP-dependent RNA helicase DDX3Y	3	4.49E-06	NaN	NaN	0	0	570670
O43175	D-3-phosphoglycerate dehydrogenase	4	6.70E-33	1.2142	1.0342	204290	171100	2382900

O43707	Alpha-actinin-4	2	0.00058141	NaN	NaN	0	0	163280
O43795	Unconventional myosin-Ib	3	6.76E-07	NaN	NaN	0	0	1659000
O43813	LanC-like protein 1	1	0.0013375	NaN	NaN	0	0	356150
O60361	Putative nucleoside diphosphate kinase	2	3.58E-05	0.62651	0.59043	52348	35229	424140
O60506	Heterogeneous nuclear ribonucleoprotein Q	1	0.00044816	NaN	NaN	0	0	163170
O60884	DnaJ homolog subfamily A member 2	1	0.011885	NaN	NaN	0	0	109910
O75323	Protein NipSnap homolog 2	1	0.01752	NaN	NaN	0	439250	0
O94832	Unconventional myosin-Id	18	2.36E-54	NaN	NaN	0	0	8499000
O94905	Erlin-2	2	2.78E-06	NaN	NaN	0	666610	0
O95831	Apoptosis-inducing factor 1, mitochondrial	4	3.26E-11	NaN	NaN	0	0	604610
P00330	Alcohol dehydrogenase 1	9	4.89E-25	NaN	NaN	9163300	0	0
P00338	L-lactate dehydrogenase A chain	6	6.10E-38	0.81566	0.77229	490320	536510	2301600
P00505	Aspartate aminotransferase, mitochondrial	1	0.00079264	NaN	NaN	0	0	258170
P00558	Phosphoglycerate kinase 1	5	1.89E-16	0.94095	0.76988	117720	147050	1031800

P04075	Fructose-bisphosphate aldolase A	8	4.89E-26	1.0899	0.99837	181280	245480	3525000
P04406	lyceraldehyde-3-phosphate dehydrogenase	7	4.44E-32	0.58568	0.51231	2486100	2149100	6776000
P04843	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	5	3.97E-13	2.3102	0.9397	181780	341740	1106800
P05023	Sodium/potassium-transporting ATPase subunit alpha-1	2	4.81E-09	1.246	0.73396	49358	75710	404490
P05109	Protein S100-A8	2	5.60E-05	NaN	NaN	479430	0	0
P05141	ADP/ATP translocase 2	4	1.42E-13	0.38364	0.27351	961580	702380	1816200
P05388	60S acidic ribosomal protein P0	4	6.52E-12	1.2319	1.6088	100010	201850	3003700
P05455	pus La protein	3	1.39E-18	0.90389	1.1375	79991	133230	780200
P06576	ATP synthase subunit beta, mitochondrial	5	5.36E-15	0.18097	0.071425	10289000	319070	1276400
P06733	Alpha-enolase	9	6.94E-55	0.78425	0.74893	1120500	1134600	7081400
P06748	ucleophosmin	2	1.88E-07	1.5289	1.796	128560	238470	2945600
P07195	L-lactate dehydrogenase B chain	9	6.26E-28	0.68318	0.90323	847710	531260	5948600
P07355	Annexin A2	4	2.33E-13	1.8488	0.37351	138030	766610	1095400
P07437	Tubulin beta chain	10	2.38E-51	1.0036	1.2012	1706000	2746400	14702000

P07900	Heat shock protein HSP 90-alpha	16	6.38E-75	1.095	1.0118	778150	958780	7292100
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	2	1.37E-06	NaN	NaN	0	0	329710
P08107	Heat shock 70 kDa protein 1A/1B	14	2.68E-94	2.1893	1.3326	2323600	6790300	22955000
P08133	Annexin A6	6	5.75E-14	NaN	NaN	0	1549800	482830
P08195	F2 cell-surface antigen heavy chain	1	0.00055029	NaN	NaN	0	0	185580
P08238	Heat shock protein HSP 90-beta	17	2.49E-104	1.0009	0.9378	2760000	3860400	18920000
P08670	Vimentin	18	2.14E-75	0.34123	0.51766	2720400	748990	11102000
P08754	Guanine nucleotide-binding protein G(k) subunit alpha	1	0.018091	NaN	NaN	0	0	182380
P08758	Annexin A5	1	0.0059068	NaN	NaN	0	0	347870
P08865	40S ribosomal protein SA	2	2.05E-08	NaN	NaN	0	0	802460
P09651	Heterogeneous nuclear ribonucleoprotein A1	3	1.80E-16	2.1637	1.1898	22142	109280	806980
P09661	U2 small nuclear ribonucleoprotein A	2	5.55E-07	0.36891	0.3364	81347	109240	481330
P09874	Poly [ADP-ribose] polymerase 1	1	1.69E-05	NaN	NaN	0	0	408210
P10599	Thioredoxin	1	0.00031617	NaN	NaN	368410	0	0

P10809	60 kDa heat shock protein, mitochondrial	6	3.49E-36	1.3638	0.71148	1298100	2377200	4906500
P11021	78 kDa glucose-regulated protein	8	4.04E-40	0.92944	0.43071	2036300	493400	1771500
P11142	Heat shock cognate 71 kDa protein	19	1.82E-140	1.9749	1.238	6556500	12155000	35096000
P11532	ystrophin	1	0.0025418	NaN	NaN	6057600	0	0
P11586	C-1-tetrahydrofolate synthase, cytoplasmic	1	0.0077035	NaN	NaN	0	0	96828
P11940	Polyadenylate-binding protein 1	4	3.82E-12	NaN	NaN	0	0	1326500
P12268	Inosine-5-monophosphate dehydrogenase 2	1	0.00089879	NaN	NaN	0	0	463590
P12277	Creatine kinase B-type	3	2.92E-08	0.67959	1.0187	137720	190930	1642500
P12956	X-ray repair cross-complementing protein 6	2	3.95E-05	NaN	NaN	0	0	210210
P13010	X-ray repair cross-complementing protein 5	1	0.033143	NaN	NaN	0	0	0
P13639	longation factor 2	14	3.34E-42	0.8134	0.93243	744160	819830	8101300
P13797	Plastin-3	1	0.015139	NaN	NaN	0	0	149880
P14174	acrophage migration inhibitory factor	2	5.86E-06	NaN	NaN	331570	0	211640
P14618	Pyruvate kinase isozymes M1/M2	6	4.55E-16	0.75999	0.7908	202230	166320	2099800

P14625	Endoplasmin	3	2.22E-11	NaN	NaN	0	0	536950
P14868	Aspartate--tRNA ligase, cytoplasmic	2	0.00010136	NaN	NaN	0	0	355250
P14923	Junction plakoglobin	2	0.00014268	NaN	NaN	398360	0	0
P15880	0S ribosomal protein S2	3	1.68E-10	NaN	NaN	0	0	728490
P17066	Heat shock 70 kDa protein 6	6	5.32E-48	NaN	NaN	0	0	1621000
P17844	Probable ATP-dependent RNA helicase DDX5	5	1.24E-12	1.1928	1.3813	115180	228750	1687100
P17987	T-complex protein 1 subunit alpha	19	5.07E-73	7.2754	3.029	84553	1821400	14344000
P18077	60S ribosomal protein L35a	1	0.015747	NaN	NaN	0	0	109660
P18124	0S ribosomal protein L7	2	2.12E-05	NaN	NaN	0	0	797210
P18621	60S ribosomal protein L17	1	0.0066318	NaN	NaN	0	0	86397
P19105	Myosin regulatory light chain 12A	1	0.0036507	NaN	NaN	0	0	1021800
P19338	Nucleolin	6	1.41E-22	0.51881	0.49505	627790	377260	3410800
P20929	Nebulin	1	0.0047187	NaN	NaN	3074600	0	0
P21796	Voltage-dependent anion-selective channel protein 1	2	5.58E-06	NaN	NaN	0	1153300	0
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	6	1.27E-17	2.6464	1.9656	38378	130340	2385600

P23246	Splicing factor, proline- and glutamine-rich	14	1.73E-54	NaN	NaN	0	0	30277000
P23396	0S ribosomal protein S3	8	1.65E-33	0.74082	1	627100	267880	3761300
P23528	Cofilin-1	3	2.52E-09	NaN	NaN	0	0	1921000
P25705	ATP synthase subunit alpha, mitochondrial	9	6.01E-87	0.7528	0.34238	3221900	1116600	2379600
P26373	60S ribosomal protein L13	3	4.27E-08	NaN	NaN	0	0	1535200
P26599	Polypyrimidine tract- binding protein 1	2	4.51E-05	NaN	NaN	0	0	237230
P26641	Elongation factor 1- gamma	2	2.34E-10	0.90816	1.1409	180960	268430	1308600
P27348	14-3-3 protein theta	3	2.69E-06	NaN	NaN	0	0	0
P27635	60S ribosomal protein L10	2	1.87E-05	1.4225	1.5385	18562	30080	383390
P27824	Calnexin	4	1.83E-22	1.141	0.60868	271400	453780	1592700
P28288	ATP-binding cassette sub- family D member 3	1	0.036281	NaN	NaN	0	0	0
P28289	Tropomodulin-1	1	0.032546	NaN	NaN	0	0	65091
P30041	Peroxiredoxin-6	1	0.011617	NaN	NaN	0	0	194410
P30049	ATP synthase subunit delta, mitochondrial	1	0.00023439	NaN	NaN	244690	0	0
P30050	60S ribosomal protein L12	2	6.83E-10	NaN	NaN	614490	0	0

P30153	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	1	0.0014952	NaN	NaN	0	0	99268
P31689	DnaJ homolog subfamily A member 1	1	0.00030598	NaN	NaN	0	0	194700
P31944	Caspase-14	1	0.00017719	NaN	NaN	0	0	0
P31947	14-3-3 protein sigma	3	2.70E-06	0.74038	0.89705	73372	71496	369530
P31948	Stress-induced-phosphoprotein 1	3	1.20E-07	0.33873	0.33194	59710	17269	354250
P32969	0S ribosomal protein L9	1	0.010233	NaN	NaN	0	0	448260
P34897	Serine hydroxymethyltransferase, mitochondrial	4	1.34E-10	2.0119	1.6623	55332	153090	1314000
P34932	Heat shock 70 kDa protein 4	1	0.00035539	NaN	NaN	0	0	313400
P35232	rohibitin	1	0.0028581	NaN	NaN	0	0	190550
P35268	60S ribosomal protein L22	3	2.89E-13	0.28814	0.25022	144210	75081	543420
P35520	ystathionine beta-synthase	1	0.0041126	NaN	NaN	0	0	57129
P35579	Myosin-9	70	0	3.2353	9.9984	167240	542150	78072000
P35580	Myosin-10	52	2.57E-246	NaN	NaN	0	0	30107000
P36578	0S ribosomal protein L4	2	3.67E-05	NaN	NaN	0	0	767920

P36873	Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	2	0.0001022	NaN	NaN	0	0	356940
P38646	Stress-70 protein, mitochondrial	4	6.36E-12	NaN	NaN	449510	0	974000
P38919	Eukaryotic initiation factor 4A-III	2	6.45E-07	NaN	NaN	0	0	152090
P39019	40S ribosomal protein S19	3	3.00E-07	1.3208	1.5959	51723	81103	1013700
P39023	0S ribosomal protein L3	2	4.68E-07	NaN	NaN	0	0	1348300
P39656	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	1	0.0049495	NaN	NaN	0	0	287330
P40227	T-complex protein 1 subunit zeta	14	3.22E-49	6.4889	2.8557	535080	4870400	19104000
P40926	Malate dehydrogenase, mitochondrial	2	0.00037261	NaN	NaN	0	0	510030
P45880	Voltage-dependent anion-selective channel protein 2	2	4.37E-09	4.9879	0.41194	53620	530850	141800
P46777	0S ribosomal protein L5	6	3.08E-13	NaN	NaN	0	0	1624100
P46778	60S ribosomal protein L21	1	0.0049021	NaN	NaN	0	0	875630
P46779	60S ribosomal protein L28	1	0.0073929	NaN	NaN	263170	0	0
P46782	0S ribosomal protein S5	1	1.29E-10	NaN	NaN	0	0	1726200

P46783	40S ribosomal protein S10	2	9.60E-05	1.2633	1.6716	56898	84267	983380
P47914	60S ribosomal protein L29	1	1.01E-10	0.39905	0.86743	88133	44291	446790
P48643	T-complex protein 1 subunit epsilon	13	4.59E-39	NaN	NaN	0	2422000	12966000
P49368	T-complex protein 1 subunit gamma	14	3.78E-45	6.9248	2.7664	413180	4783400	10439000
P49411	Elongation factor Tu, mitochondrial	1	0.0014751	NaN	NaN	0	0	185560
P49755	Transmembrane emp24 domain-containing protein 10	1	0.0031442	NaN	NaN	0	0	0
P50914	60S ribosomal protein L14	1	0.00011154	NaN	NaN	0	0	765350
P50990	T-complex protein 1 subunit theta	26	2.02E-99	7.044	2.7919	496550	5909800	26036000
P50991	T-complex protein 1 subunit delta	16	8.28E-74	6.2234	2.5249	1117300	9851800	18270000
P51572	B-cell receptor-associated protein 31	4	1.35E-07	1.2351	0.4969	496130	156880	691600
P52179	Myomesin-1	1	0.0042913	NaN	NaN	0	0	5696300
P52272	Heterogeneous nuclear ribonucleoprotein M	7	8.08E-25	0.84575	0.87184	129620	122690	1884000
P52907	F-actin-capping protein subunit alpha-1	3	1.98E-16	NaN	NaN	0	0	1419200

P53985	Monocarboxylate transporter 1	3	2.71E-08	NaN	NaN	0	0	1714600
P54577	Tyrosine--tRNA ligase, cytoplasmic	1	0.0098503	NaN	NaN	0	0	0
P55060	Exportin-2	1	0.013963	NaN	NaN	0	0	188930
P55072	Transitional endoplasmic reticulum ATPase	2	1.38E-05	0.91366	1.2746	28152	54409	377210
P55084	Trifunctional enzyme subunit beta, mitochondrial	1	0.037102	NaN	NaN	0	0	0
P55795	Heterogeneous nuclear ribonucleoprotein H2	1	8.16E-06	NaN	NaN	0	0	565120
P59998	Actin-related protein 2/3 complex subunit 4	1	0.0012306	NaN	NaN	0	0	0
P60174	Triosephosphate isomerase	2	1.18E-13	NaN	NaN	0	0	499880
P60660	Myosin light polypeptide 6	7	2.55E-19	NaN	NaN	0	0	5742600
P60842	Eukaryotic initiation factor 4A-I	5	1.46E-16	0.75834	0.94091	176150	214840	1758100
P60953	Cell division control protein 42 homolog	1	0.016719	NaN	NaN	0	0	112870
P61158	Actin-related protein 3	3	7.51E-07	NaN	NaN	0	0	582160
P61247	40S ribosomal protein S3a	3	3.81E-10	1.4965	2.9661	12220	32449	860630
P61254	60S ribosomal protein L26	4	1.18E-10	1.5074	2.2099	32424	58285	1914900
P61353	60S ribosomal protein L27	2	0.00019502	NaN	NaN	0	0	870840

P61513	60S ribosomal protein L37a	1	1.69E-05	NaN	NaN	152280	0	0
P61604	10 kDa heat shock protein, mitochondrial	1	6.90E-07	NaN	NaN	465970	0	0
P61978	Heterogeneous nuclear ribonucleoprotein K	2	1.22E-07	NaN	NaN	0	0	1408100
P61981	14-3-3 protein gamma	3	2.64E-06	NaN	NaN	0	0	47379
P62195	26S protease regulatory subunit 8	1	0.017487	NaN	NaN	0	0	187760
P62249	40S ribosomal protein S16	3	4.81E-08	0.8428	1.1157	44867	46831	758760
P62258	14-3-3 protein epsilon	3	3.13E-06	NaN	NaN	0	0	274930
P62277	40S ribosomal protein S13	2	1.28E-11	0.88541	1.0779	242540	285250	1511200
P62280	40S ribosomal protein S11	2	0.00025699	NaN	NaN	0	0	490190
P62424	60S ribosomal protein L7a	1	0.0020585	NaN	NaN	0	0	665860
P62701	40S ribosomal protein S4, X isoform	4	7.96E-11	NaN	NaN	0	0	2112500
P62753	0S ribosomal protein S6	3	4.43E-07	NaN	NaN	0	0	1355600
P62805	stone H4	1	0.0022684	NaN	NaN	147460	0	0
P62826	TP-binding nuclear protein Ran	2	2.85E-07	1.2363	0.57936	124360	209820	391840
P62829	60S ribosomal protein L23	2	0.00026706	1.3332	0.91427	54660	94805	773160
P62851	40S ribosomal protein S25	2	1.02E-05	NaN	NaN	0	0	902610
P62899	60S ribosomal protein L31	2	1.83E-06	NaN	NaN	0	0	1351100

P62906	60S ribosomal protein L10a	1	0.0019691	NaN	NaN	0	0	404960
P62913	60S ribosomal protein L11	3	2.97E-56	2.7353	2.1879	142330	438650	1819300
P62917	0S ribosomal protein L8	2	3.58E-05	NaN	NaN	0	0	2003300
P62937	Peptidyl-prolyl cis-trans isomerase A	2	0.00010342	NaN	NaN	0	0	956050
P62987	Ubiquitin-60S ribosomal protein L40	3	8.91E-28	0.58509	0.78138	947290	2180700	4241000
P63104	14-3-3 protein zeta/delta	4	1.61E-08	NaN	NaN	0	0	1599000
P67809	Nuclease-sensitive element-binding protein 1	1	1.63E-05	NaN	NaN	0	0	219170
P68133	Actin, alpha skeletal muscle	13	3.23E-107	2.8824	7.6779	543620	1498500	20792000
P68371	Tubulin beta-4B chain	8	5.32E-40	0.98747	0.99468	293370	325730	1199100
P78371	T-complex protein 1 subunit beta	20	5.42E-105	5.6687	2.2561	842360	6889700	21806000
P78426	Homeobox protein Nkx-6.1	1	0.0096283	NaN	NaN	482520	0	0
P81605	ermcidin	1	0.0012778	NaN	NaN	948600	0	0
P83731	60S ribosomal protein L24	2	7.32E-06	NaN	NaN	0	0	781700
P84085	ADP-ribosylation factor 5	1	0.0039433	NaN	NaN	0	0	203080
Q00839	Heterogeneous nuclear ribonucleoprotein U	2	2.57E-05	2.6264	1.1759	41749	224340	304260
Q01469	Fatty acid-binding protein, epidermal	1	0.005731	NaN	NaN	160840	0	0

Q02413	Desmoglein-1	2	5.85E-05	NaN	NaN	530590	0	0
Q02543	60S ribosomal protein L18a	1	0.00050872	NaN	NaN	0	0	348750
Q02790	Peptidyl-prolyl cis-trans isomerase FKBP4	1	0.03161	NaN	NaN	0	0	101890
Q02878	0S ribosomal protein L6	1	0.00027608	NaN	NaN	0	0	0
Q06830	Peroxiredoxin-1	3	2.19E-11	0.47735	0.54294	350050	266860	1798400
Q07020	60S ribosomal protein L18	1	0.011375	NaN	NaN	0	0	0
Q07021	Complement component 1 Q subcomponent-binding protein, mitochondrial	1	1.89E-08	NaN	NaN	263950	0	0
Q08188	Protein-glutamine gamma-glutamyltransferase E	2	0.00058225	NaN	NaN	183820	0	0
Q0D2K2	Kelch-like protein 30	1	0.024261	NaN	NaN	454900	0	0
Q12792	Twinfilin-1	1	0.00024424	NaN	NaN	0	0	238700
Q12905	Interleukin enhancer-binding factor 2	1	0.0023142	NaN	NaN	0	0	352010
Q13162	Peroxiredoxin-4	2	1.93E-07	NaN	NaN	0	0	173580
Q13283	Ras GTPase-activating protein-binding protein 1	1	0.0074658	NaN	NaN	0	0	274210
Q13310	Polyadenylate-binding protein 4	4	3.15E-12	NaN	NaN	0	0	0

Q14103	Heterogeneous nuclear ribonucleoprotein D0	1	0.0001989	NaN	NaN	0	0	492080
Q14140	SERTA domain-containing protein 2	1	0.027867	NaN	NaN	5129100	0	0
Q14697	Neutral alpha-glucosidase AB	1	0.0016445	1.4999	0.85036	40337	64412	144770
Q14974	Importin subunit beta-1	3	5.16E-08	NaN	NaN	0	0	701470
Q15084	Protein disulfide-isomerase A6	3	3.66E-10	2.5558	0.69218	50811	322060	868310
Q15233	Non-P	12	9.44E-53	0.72151	6.4697	140930	126650	24313000
Q15366	Poly(rC)-binding protein 2	1	0.0030262	NaN	NaN	0	0	308220
Q15572	TATA box-binding protein-associated factor RNA polymerase I subunit C	1	0.026831	0.33259	0.10911	4043200	2308000	2371500
Q15758	Neutral amino acid transporter B(0)	2	1.36E-05	NaN	NaN	0	0	581110
Q16563	Synaptophysin-like protein 1	1	0.0020673	NaN	NaN	0	0	326310
Q16576	Histone-binding protein RBBP7	1	0.0098765	1.352	0.94173	15801	27959	89772
Q16643	Drebrin	8	1.37E-21	NaN	NaN	0	0	4980300
Q2VIR3	Putative eukaryotic translation initiation factor 2 subunit 3-like protein	1	0.01249	NaN	NaN	0	0	120050

Q3SYG4	Protein PTHB1	1	0.013167	1.8233	1.3174	331400	834590	2500800
Q562R1	Beta-actin-like protein 2	7	5.93E-50	NaN	NaN	0	0	69915000
Q58FF8	Putative heat shock protein HSP 90-beta 2	6	3.10E-17	NaN	NaN	0	0	507560
Q5THK1	Protein PRR14L	1	0.016542	NaN	NaN	592250	0	0
Q5VTE0	Putative elongation factor 1-alpha-like 3	7	1.08E-43	0.95943	0.91038	4274600	4205800	21289000
Q684P5	Rap1 GTPase-activating protein 2	1	0.041051	NaN	NaN	1502300	0	0
Q6NXT2	istone H3.3C	1	0.014121	NaN	NaN	161710	0	0
Q86TJ2	Transcriptional adapter 2-beta	1	0.013113	NaN	NaN	2142600	0	0
Q86WA8	Lon protease homolog 2, peroxisomal	1	0.03869	NaN	NaN	0	0	996110
Q8IWC1	MAP7 domain-containing protein 3	1	0.038695	NaN	NaN	0	0	180080
Q8IZP2	Putative protein FAM10A4	1	0.012812	NaN	NaN	0	0	107450
Q8N960	Centrosomal protein of 120 kDa	1	0.020107	NaN	NaN	111080	0	0
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein	3	2.31E-24	0.96584	1.7466	70836	103170	984980
Q92485	Acid sphingomyelinase-like phosphodiesterase 3b	1	0.0091832	NaN	NaN	0	94595	0

Q92504	Zinc transporter SLC39A7	1	0.013713	5.1273	24.732	11893	64673	1402300
Q92734	rotein TFG	2	1.52E-09	NaN	NaN	0	0	1279200
Q92841	Probable ATP-dependent RNA helicase DDX17	4	5.30E-12	NaN	NaN	0	0	101600
Q96KD3	Protein FAM71F1	1	0.01615	NaN	NaN	938570	0	0
Q99623	Prohibitin-2	1	2.00E-05	0.67946	0.23133	291060	231880	427100
Q99832	T-complex protein 1 subunit eta	12	2.99E-48	8.0658	3.3106	561950	2619400	12316000
Q9BR76	Coronin-1B	1	0.035398	NaN	NaN	0	0	0
Q9BUJ2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	1	0.0010379	NaN	NaN	0	0	108120
Q9BVK6	Transmembrane emp24 domain-containing protein 9	1	7.49E-08	NaN	NaN	0	0	221480
Q9C0B2	Uncharacterized protein KIAA1751	1	0.0061334	0.92596	0.91981	63021	143410	305410
Q9H489	Putative testis-specific Y-encoded-like protein 3	1	0.035627	NaN	NaN	1301300	0	0
Q9HAV0	Guanine nucleotide-binding protein subunit beta-4	1	0.0039972	NaN	NaN	0	0	216690
Q9NYL9	Tropomodulin-3	1	0.0015437	NaN	NaN	0	0	243430
Q9NZI8	Insulin-like growth factor 2 mRNA-binding protein 1	2	1.50E-05	NaN	NaN	0	0	278800

Q9P035	Very-long-chain (3R)-3-hydroxyacyl-[acyl-carrier protein] dehydratase 3	1	0.00026613	NaN	NaN	0	0	186260
Q9P0K7	Ankycorbin	1	0.02731	NaN	NaN	0	0	0
Q9UEU0	Vesicle transport through interaction with t-SNAREs homolog 1B	1	0.01525	NaN	NaN	0	0	82763
Q9UG63	ATP-binding cassette sub-family F member 2	1	0.016292	NaN	NaN	0	0	34830
Q9UHB6	LIM domain and actin-binding protein 1	11	2.08E-26	NaN	NaN	0	0	5240000
Q9ULV4	Coronin-1C	14	1.61E-37	3.2517	12.324	29349	117870	20424000
Q9UM54	Unconventional myosin-VI	5	2.68E-10	NaN	NaN	0	0	761200
Q9UQ80	Proliferation-associated protein 2G4	4	1.79E-10	NaN	NaN	0	0	2040000
Q9Y265	RuvB-like 1	1	0.015002	NaN	NaN	0	0	293060
Q9Y3I0	tRNA-splicing ligase RtcB homolog	2	0.0004902	NaN	NaN	0	0	334850

10 Appendix III

LCMV Particle MS analysis from BHK21 cells.

Accession	Peptide count	Confidence score	Description	Raw abundance
P09992	35	1271.66	Nucleoprotein	95913946.2
Q711N9	15	556.35	Actin, cytoplasmic 1	18381609.14
A9XDF3	14	426.14	Serine protease	5099672.694
G8IFB9	14	326.67	Clathrin heavy chain	1462694.652
O08570	6	216.49	Fritz	19657334.6
P09991	7	200.2	Pre-glycoprotein polyprotein GP complex	2662623.988
P86221	5	185.54	Tubulin beta-4B chain	2522698.128
Q91Y69	6	174.58	Matrix metalloproteinase-2	26967677.16
P97279	7	156.13	Inter-alpha-trypsin inhibitor heavy chain H2	5101193.058
P86204	5	152.6	Heat shock-related 70 kDa protein 2	3110992.727
P86210	4	143.18	Alpha-enolase	1372682.103
Q0QET9	4	125.1	Glyceraldehyde-3-phosphate dehydrogenase	1324996.126
C0HJG9	5	115.57	Annexin A2	999802.9267
P86234	5	113.86	Tubulin alpha-3 chain	3144298.408
E2GMV3	3	104.32	Milk fat globule-EGF factor 8	8697570.245
Q925Q6	2	97.01	Tissue inhibitor of matrix metalloproteinase-2	17757786.88
Q80Z94	3	91.98	Gap junction protein	1579046.793
D0G7D9	4	87.42	Collagen type VI alpha 3 subunit	927789.7218
D7RXW0	3	80.43	Alpha-tubulin I	2966797.393
P51640	3	76.37	Glyceraldehyde-3-phosphate dehydrogenase	1673663.472
P97280	4	73.18	Inter-alpha-trypsin inhibitor heavy chain H3	2205684.314
E2GMU8	3	72.23	Eukaryotic translation elongation factor 1 alpha	2216206.695
P86246	2	69.25	Keratin, type I cytoskeletal 19	1564647.47

P86245	2	50.81	T-complex protein 1 subunit beta	1236180.263
P02544	2	50.29	Vimentin	122132.0219
P86237	2	49.05	Heat shock 70 kDa protein 1-like	255273.7396
B6V7E0	2	47.37	Decorin	6957576.72
P01945	2	46.58	Hemoglobin subunit alpha	12319848.96
D0G6X1	2	46.58	Collagen type VI alpha 1 subunit	872339.4204
P70110	2	41.88	Platelet glycoprotein 4	218104.4245
P86247	2	38.9	Keratin, type II cytoskeletal 8	5537907.721
P18541	2	36.44	RING finger protein Z	2520837.402
Q60546	1	31.81	Coagulation factor X	517630.2584
Q5KTJ7	1	27.72	Ras-related protein Rab-3B	108930.9741
P86208	1	26.2	T-complex protein 1 subunit alpha	169559.7865
Q60522	1	25.18	CD44 antigen	365024.1273
K7WL83	1	25.04	Beta-globin	118395069.6
A6YF56	1	24.77	Serum albumin	4280235.095
Q8VIB7	1	23.29	Attractin	12779869.83
Q7M090	1	23.03	Annexin II	971075.2667
P97277	1	22.41	Alpha-1-antitrypsin	1187194.317
Q5KTJ6	1	22.25	Ras-related protein Rab-13	352767.1339
P13540	1	22.01	Myosin-7	327422.7109
Q7M0B0	1	21.33	Alpha-1-inhibitor III	4374146.615
Q9Z115	1	13.54	Complement C3	23371426.67

11 Appendix IV

Time (h)	NPEG			LCMV			NPEG & LCMV			Negative			P value
	1	2	3	1	2	3	1	2	3	1	2	3	NPEG : NPEG& LCMV
0	18315	20525	20885	19840	19490	18465	18480	18980	21330	19290	22100	21730	0.806417
2	31465	27030	30590	19030	18930	21245	40925	37760	33300	20085	20045	20390	0.0422957
4	57535	56020	53370	17755	18310	19900	70595	60025	69575	19700	19190	19910	0.0362849
6	69290	65855	62870	18940	20385	19010	75900	62945	62355	20295	19075	19975	0.835561
8	49900	47210	46710	17895	19460	20220	39895	40405	50140	18345	18945	20500	0.268915
10	32715	33850	38745	20940	19660	20150	35670	29250	30830	20525	20065	20420	0.299313
12	35120	29095	31555	18885	21605	18515	27365	29870	33790	19695	19875	20610	0.570151

Table showing the total values for RLU detection.

218

The end.