Identification of TRIM25 and ZAP as Restriction Factors of Arenaviruses Through BioID-Derived Proximity Interactome Analysis of Arenavirus NP

Finley Ellis Otterburn Old

Thesis submitted for the degree of Doctor of Philosophy

The University of Nottingham

The School of Veterinary Medicine and Science

December 2024

Acknowledgements

First and foremost, I would like to express my eternal gratitude to my brilliant supervisor, Toshana Foster, for her incredible intelligence, unwavering support, and guidance throughout this journey. Her insights, encouragement, and infinite patience have been instrumental in shaping this thesis and my growth as a researcher. I am also deeply thankful to Ivan Campeotto for his incredible expertise in the complex and confusing field of structural biology. I owe the 5th chapter of my thesis to his generous providing of reagents and facilities as well as wonderfully insightful supervision. Janet Daly deserves much thanks and praise for her valuable expertise, guidance, and consistent support. Her encouragement and constructive feedback have been greatly appreciated and have significantly enriched this work. A very special thank you goes to Robert Stott-Marshall, whose readiness to help and seemingly infinite knowledge were a constant source of inspiration. Your availability and expertise have been invaluable, and I am incredibly grateful for your generosity in sharing your time, wisdom, and extremely useful experimental optimisations.

To my amazing fiancée, Rebecca, thank you for your endless love, patience, and encouragement. You have always believed in me, even when I doubted myself, and for that, I am forever grateful. This thesis is as much yours as it is mine. To my wonderful family, thank you for your unconditional love and support throughout this journey, without you this would not have been possible.

To my incredible friends, Joseph Hinds, Abbie Hinds, Joshua Morrison, and Liam Burton, I am forever grateful for your unwavering support, understanding, and cheerleading from the sidelines. Your faith in me and your constant encouragement have been a source of comfort and motivation.

Finally, I would like to thank everyone who contributed, directly or indirectly, to this work. Your support has meant the world to me.

Abstract

The *Arenaviridae* family of RNA viruses includes several important human pathogens, such as Lassa mammarenavirus (LASV), the causative agent of Lassa fever. A characteristic of arenavirus infection is the broad symptom range, from asymptomatic to fatal viral haemorrhagic fever. Understanding the mechanisms underlying this observed range in pathogenicity and symptoms is crucial to combatting the global health burden posed by human pathogenic arenaviruses.

This thesis aimed to identify host restriction factors of Old-World arenavirus nucleoproteins (NP) using BioID2 proximity labelling and to investigate their antiviral potential. Host interactomes for Lassa mammarenavirus, Lujo mammarenavirus and the non-human pathogenic Mopeia mammarenavirus NPs were generated. These highlighted critical immunological networks and conserved host proteins, including TRIM25 and Zinc antiviral protein (ZAP). Both of these host proteins were experimentally validated as potent restriction factors of arenavirus replication. Infection assays using a panel of TRIM25 mutants identified the SPRY domain of TRIM25 as essential for interaction with NP and effective viral replication inhibition. Structural studies utilising SEC-SAXS revealed distinct oligomerisation patterns between LASV, LUJV, and MOPV NP. Although challenges in protein purification hindered detailed characterisation of TRIM25 interactions, this research optimised methodologies for future investigations. These findings deepen the understanding of arenavirus-host interactions and identify key antiviral targets to inform therapeutic development.

Contents

List of figures

List of tables

List of abbreviations

µg	microgram
µl	microlitre
3pRNA	RNA with 5’-triphosphate modifications
aa	amino acid
ADP	adenosine diphosphate
ALT	alanine transaminase
AMSF	Advanced Mass Spectrometry Facility
AP	affinity purification
AST	aspartate transaminase
BIBD	Boid inclusion body disease
CARD	caspase recruitment domain
CC	coiled coil
CDK	cyclin dependent kinase
cDNA	complementary DNA
CETSA	cellular thermal shift assay
CHAPV	Chapare virus
CPR	centre for proteome research
cRNA	complementary RNA
DC	Dendritic cells
DCP1	decapping protein 1
DENV	Dengue virus
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
dsRNA	double stranded RNA
DTT	dithiothreitol
EBOV	Ebola virus
EDTA	ethylenediaminetetraacetic acid
mRNA	messenger RNA
MS	mass spectrometry
ng	nanogram
NK	natural killer
nm	nanometre
NMR	nuclear magnetic resonance
NP	nucleoprotein
NRP	neurophilin
NW	New World
ORF	open reading frame
OW	Old World
PAMP	Pathogen associated molecular patterns
PARN	polyadenylate-specific ribonucleas
PASC	pairwise sequence comparison
PBS	Phosphate buffered saline
PCA	principle component analysis
PCR	polymerase chain reaction
PDB	protein database
PDB-MS	proximity-based labelling coupled to Mass Spectrometry
PEI	Polyethylenimine
PFA	paraformaldehyde
PFU	plaque forming units
PI4P	phosphatidylinositol-4-phosphate
PKR	Protein kinase R
PPE	personal protective equipment
PPI	protein-protein interaction
ProS	host protein S

eIF4F	eukaryotic initiation factor 4Fs
EMSA	electrophoretic mobility shift assay
ER	endoplasmic reticulum
ESCRT	endosomal sorting complexes required for transport
EVAg	European Virus Archive goes Global platform
ExoN	exoribonuclease DEDDh motif
FBS	Fetal bovine serum
GA	Golgi apparatus
GaS6	growth arrest-specific gene 6
GGV	Golden Gate reptarenavirus
GPC	glycoprotein complex
GSEA	genome set enrichment analysis
GTOV	Guanarito virus
HEK	human embryonic kidney
HISV	Haartman institute snake virus
HR	Heptad repeat
HRP	horseradish peroxidase
HSV-1	Herpes simplex virus type 1
IAV	Influenza A virus
ICNV	International committee on nomenclature of viruses
ICTV	International committee on taxonomy of viruses
IF	Immunofluorescence
IFN	interferon
IGR	intergenic region
IKKe	IkB kinase e
IMAC	immobilised metal-affinity chromatography
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
PRR	Pattern recognition receptor
PRYSP RY	SPRY/SPIa and Ryanodine receptor
PSM	peptide spectrum match
PtdSer	phospholipid phosphatidylserine
RdRO	RNA-directed RNA polymerase
Rg	radius of gyration
RIG-I	Retinoic acid inducible gene I
RING	Really interesting new gene
RNA	ribonucleic acid
Rnase	ribonuclease
RNP	ribonucleoprotein complex
rSAP	recombinant shrimp alkaline phosphatase
RSV	respiratory syncytial virus
RTC	replication-transcription complexes
RT-qPCR	Real time quantitative Polymerase chain reaction
SAXS	small angle x-ray scattering
SBAV	Sabia virus
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SF	serum free
sgRNA	single guide RNA
siRNA	short interfering RNA
SPAV1	pescarenavirus 1
SPAV2	pescarenavirus 2
SSP	stable signal peptide
ssRNA	single stranded RNA
TAM	Tyro3/Axl/Mer -receptor tyrosine kinase members
TCPA	Thermal proximity co-aggregation

IRF	interferon regulatory factor
ISG	Interferon stimulated gene
ISGF3	Interferon-Stimulated Gene Factor 3
ISRE	interferon-sensitive response element
JAK/STAT	Janus kinase/signal transducer and activator of transcription
JUNV	Junin virus
kDa	kilodalton
KHNYN	KH and NYN domain-containing protein
LAMP1	lysosomal associated membrane protein 1
LASV	Lassa mammarenavirus
LC	liquid chromatography
LCMV	Lymphocytic choriomeningitis virus
LFQ	label free quantification
LUJV	Lujo mammarenavirus
Mab	monoclonal antibody
MACV	Machupo virus
MAVS	Mitochondrial antiviral-signaling **protein**
MCS	multiple cloning site
MDA5	melanoma differentiation-associated protein 5
MHC	major histocompatibility complex
ml	millilitre
mM	millimolar
MOI	multiplicity of infection
MOPV	Mopeia mammarenavirus
TCVR	Tacaribe virus
TF	Transcription factor
TLR	Toll like receptor
TMT	tandem mass tag
TRAF	Tumour necrosis factor associated factor
TRIF	TIR-domain-containing adapter-inducing interferon-β
TRIM25	Tripartite motif containing 25
TYK2	tyrosine-protein kinase
UBLs	ubiquitin-like proteins
UV	ultraviolet
VLP	virus like particle
vRNA	viral RNA
VSV	vesicular stomatitis virus-vectored vaccine
WlFAV1	Wēnlǐng frogfish arenavirus 1
WlFAV2	Wēnlǐng frogfish arenavirus
WT	wild type
XAPV	Xapuri virus
XL-MS	Cross-linking MS
YFV	yellow fever virus
Z	Matrix protein
ZAP	Zinc antiviral protein
ZIKV	Zika virus
ZRE	Zap responsive element
α-DG	α-dystroglycan

Chapter 1

Introduction to the Arenaviruses

1 Introduction to the Arenaviruses

1.1 General introduction

1.1.1 Discovery of Arenaviruses

The rodent-borne Lymphocytic Choriomeningitis mammarenavirus (LCMV) was the first discovered arenavirus and was isolated from monkey-to-monkey transfer in 1933 during a study of an epidemic of St. Louis encephalitis (Armstrong and Lillie, 1934). Despite not being the direct cause of the St. Louis encephalitis outbreak, LCMV was eventually found to be a cause of non-bacterial (aseptic) lymphocytic meningitis which can cause inflammation covering the brain and the spinal cord. Its viral nature was discovered by Rivers and Scott, who demonstrated that the causative agent of the disease could be filtered through an average pore size of 150µm and that no bacteria, fungal, or protozoal cells could be detected by histological studies of infected tissue or cultivated in bacterial growth media (Findlay et al., 1936, Rivers and McNair Scott, 1935). The natural reservoir host of LCMV was identified as the common house mouse (*Mus musculus)* when Traub (1935) recovered an infectious agent from white mice which then produced an illness in mice very similar to that described by Armstrong and Lillie. Two further similar agents were isolated from the cerebrospinal fluid of two men being treated for virus meningitis at the hospital of the Rockefeller Institute. One of the patients had worked with mice from the Rockefeller Institute's laboratory mouse colony which Traub showed were infected with LCMV (Traub, 1936), however the other individual was shown to be unlikely to have had contact with these same mice. Once the close similarity of these isolates was established, the name Lymphocytic Choriomeningitis mammarenavirus was adopted (Armstrong and Paul, 1935) and the original isolated strain was named the Armstrong strain. The two other LCMV strains were named Traub, after Eric Traub who

isolated the strain and WE strain which was named after the person from which it was originally isolated (Traub, 1935).

Tacaribe mammarenavirus (TCRV) from Trinidad and Tobago was discovered in 1956 when it was isolated from Jamaican fruit-eating bats (Downs et al., 1963) during a rabies surveillance survey being conducted in Trinidad from 1956 to 1958. However, the virus is not associated with severe human disease and the anecdotal reports suggest just a single human infection with mild febrile illness. Junin mammarenavirus (JUNV) from Argentina was discovered in 1959 and identified as the causative agent of Junin/Argentinian haemorrhagic fever with the reservoir host being the drylands laucha (*Calomys musculinus)* (Parodi et al., 1958, Parodi et al., 1959).

The "Tacaribe antigenic group" was established in 1963 when a serological relationship between JUNV and TCRV was shown using the complement fixation test (which determines the presence of antigen-specific antibodies by incubating patient serum with antigen and complement) and differences between the viruses was shown using a neutralisation assay in which they were easy to distinguish from each other (Mettler et al., 1963). Machupo mammarenavirus (MACV) was discovered in the remote areas of the Beni province in Bolivia in 1963 where it was isolated from a patient with Machupo/Bolivian haemorrhagic fever (Johnson et al., 1965). The reservoir host of MACV was identified as big Lauchas (*Calomys callosus)* (Johnson et al., 1966) and the virus was also shown to be antigenically similar to JUNV by complement fixation tests (Wiebenga et al., 1964).

It was then proposed that the LCMV and the Tacaribe complex viruses be grouped into a new taxonomic group of viruses with LCMV as the prototype (Pedersen, 1971); this new group was initially known as "Arenoviruses" and then later changed to "Arenaviruses" (Rowe et al., 1970a) and the proposal was based upon morphological and morphogenic similarities between LCMV and Tacaribe complex viruses (Murphy et al., 1969) as well as cross-

serological reactivity between them in indirect immunofluorescence assays (Rowe et al., 1970b).

In 1969 a novel arenavirus Lassa mammarenavirus (LASV) was first found in the small Nigerian town Lassa in Borno State, after which the virus is named (Frame et al., 1970). LASV was shown to be antigenically related to LCMV and some Tacaribe complex viruses and its morphology was similar to LCMV (Buckley and Casals, 1970, Speir et al., 1970). The criteria for grouping into arenaviruses included morphological and serological similarity but several viruses were also found to not only have similar limited geographical distributions but were also carried by specific rodent hosts (Except TCRV) and could infect people to induce clinically similar disease with similar symptoms of fever and potential haemorrhaging.

Because of this, in 1971, the taxon Arenavirus was approved at the genus level by the international Committee on Taxonomy of Viruses (ICTV) precursor; the International committee on nomenclature of viruses (ICNV) and the Arenaviridae family was created to include this genus with LCMV and Tacaribe complex viruses recognised (Fenner, 1976). Recently the ICTV has changed the way in which viruses are named, changing to a binomial naming convention for virus species. The ICTV now employs a genus-species format meaning a species name would be formed through combination of the genus name with the species-specific epithet. For example, Lassa mammarenavirus, whereas previously this would have been Lassa virus.

This taxonomic *arenavirus* genus got its name from the unique grainy electron-dense, sand-like granular appearance due to the presence of ribosomes that are acquired from the host cell. Since discovery and isolation of the Tacaribe mammarenaarenavirus from *Artibeus* fruit bats in Trinidad in 1956, there has been an average discovery rate of one new arenavirus every 1-3 years (Radoshitzky et al., 2015a). More recent additions to the list of discovered arenaviruses that cause human disease include Guanarito mammarenavirus (GTOV) found in Venezuela (1989),

Sabia mammarenavirus (SBAV) from Brazil (1993), Chapare mammarenavirus (CHAPV) also found in Bolivia (2004), and Lujo mammarenavirus in South Africa (2008) (Table 1.2).

With the discovery of several new arenavirus-like viruses in other natural reservoir hosts including reptiles and in fish, the(ICTV changed the previously recognised *Arenavirus* genus to *Mammarenavirus*. Thus, within the family of Arenaviridae, there are now four classified genera: *Antennavirus* (fish), *Hartmaniviruses* and *Reptarenaviruses* (reptiles), and *Mammarenaviruses* (rodents) (Stenglein et al., 2012b, Shi et al., 2018a, Maes et al., 2019a)

1.1.2 Classification of the *Bunyaviricetes class*

The ICTV decided to reorganise the virus taxonomy to come in line with how other biological taxonomies were organised. This resulted in the creation of the first viral realm, the Riboviria which comprised any and all viruses with RNA as the genomic material. The reorganisation also resulted in the creation of the order *Bunyavirales* into which the *Arenaviridae* family is classified. The aim of creating the *Bunyavirales* order was to group related viruses which possess a single-stranded, segmented RNA genome in either negative- or ambi-sense orientation (Maes et al., 2019b).

Recently, due to the rapidly increasing size of the order, *Bunyavirales* has been promoted to class *Bunyaviricetes*. This is because of frequent discovery of new related bunyavirals such as polyploviricotine viruses. Currently the *Bunyaviricetes* class is comprised of two major orders; *Elliovirales* which includes seven families (*Cruliviridae, Fimoviridae, Hantaviridae, Peribunyaviridae, Phasmaviridae, Tospoviridae, and Tulasviridae),* and *Hareavirales* which includes eight families (*Arenaviridae, Discoviridae, Konkoviridae, Leishbuviridae, Mypoviridae, Nairoviridae,*

Phenuiviridae, and Wupedeviridae). These fifteen families include hundreds of viruses (Kuhn et al., 2024).

1.1.1 Classification of the *Arenaviridae* family

The *Arenaviridae* family was established in 1976. This was to accommodate similar viruses which are bisegmented, ambisense single-stranded RNA genomes of about 10.5kb forming enveloped particles with a "sandy" or "granular" appearance from the latin *arenosus* meaning sandy (Fenner, 1976). Until recently it was a monogeneric family, meaning it included just a single genus, this being *Arenavirus*, with a constantly growing number species. After the discovery of several arenaviruses in snakes (Stenglein et al., 2012a, Bodewes et al., 2013, Hetzel et al., 2013), the *Arenavirus* family taxonomy was changed significantly with the *Arenavirus* genus being renamed to *Mammarenavirus,* and a second genus of *Reptarenavirus* being established in 2014 for these newly discovered snake viruses. In 2018, there were further changes and the *Arenaviridae* family was further expanded to include the addition of a novel genus, *Hartmanivirus*, for Haartman Institute snake virus (HISV) isolated from a captive snake in Finland. Further additions to the family genera include *Antennavirus* which infect fish, and *innmovirus* for arenaviruses which infect unknown hosts (Radoshitzky et al., 2023). Viruses that have been assigned to each of the five genera form a monophyletic clade based on phylogenetic analysis of the L protein/RdRP and nucleoprotein (NP) sequences and viruses from all five genera share one or more of the following traits: 1) enveloped spherical or pleomorphic virions; 2) segmented single-stranded, ambisense RNA genome that is lacking polyadenylated tracts at the 3'-end; 3) genomic sequence complementarity at both the 5'- and 3'-ends; 4) nucleotide sequences forming one or more hairpin configurations within non-coding intergenic regions (IGRS) of genomic and antigenomic segments; 5) capped but not polyadenylated virus mRNAs; 6) induction of persistent and often asymptomatic infection within reservoir hosts sometimes causing chronic viraemia (Radoshitzky et al., 2015a). For a novel arenavirus to earn

designation as a new species it must meet certain criteria set by the ICTV; the virus must be associated with a new natural and distinct main host or group of sympatric hosts and must be present in a new and distinct geographical region, the viral NP amino acid sequence must share less than 88% identity with other viruses (Salvato et al., 2005). The phylogeny of the arenavirus species is depicted in figure 1.1.

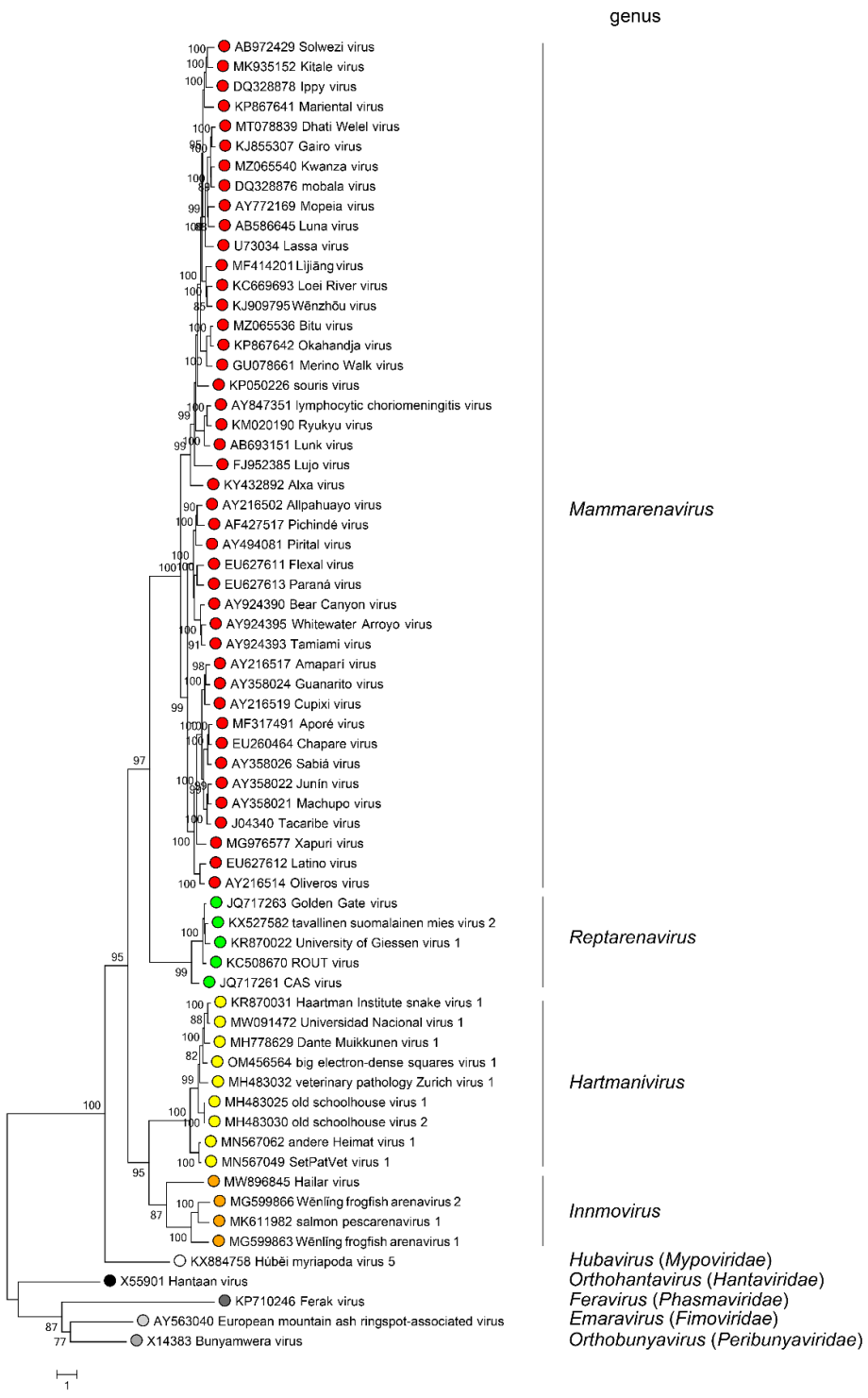


Figure 1.1 Phylogenetic tree of the Bunyavirales Class. Maximum-likelihood phylogenetic tree inferred from MAFFT alignment (Katoh and Standley, 2013) of the complete L protein amino-acid sequences of 61 arenavirids assigned to the genera Antennavirus (cyan dots), Hartmanivirus (dark green dots), Innmovirus (yellow dot), Mammarenavirus (red dots), and Reptarenavirus (light green dots), along with representative viruses of other bunyaviral families (dots of other colours). The tree was generated by the IQ-TREE software v.1.6.12 (Trifinopoulos et al., 2016).IQ-TREE selects the best-fitting evolutionary model and then generates a maximum-likelihood tree. Branch supports were calculated using the ultrafast bootstrap method (1,000 bootstraps). The tree was visualized using FigTree (http://tree.bio.ed.ac.uk) and is mid-point rooted. This tree was assembled by (Radoshitzky et al., 2023) and obtained from the ICTV website on.

The *Mammarenavirus* genus includes all arenaviruses from mammalian hosts. Mammarenaviruses possess a bi-segmented negative-sense RNA genome coding for just four proteins; L polymerase (L or RdRP), nucleocapsid protein (NP), glycoprotein complex (GPC) and the Z matrix protein (Z) in an ambisense orientation in which the negative-sense transcribes for L and NP whereas GPC and Z are transcribed in the positive-sense replicate of the vRNA (complement RNA; cRNA).

Species belonging to the *Mammarenavirus* genus are classified as either Old World (Lassa-lymphocytic choriomeningitis serocomplex) or New World (Tacaribe serocomplex). Viruses in the New World group are further classified into four clades; A, B, C, and D (also known as A/Recombinant as they are formed through recombination events occurring between arenaviruses). This genus organisation has existed since the 1970s based upon geographical distribution, serological analysis, and host association (Matthews, 1979) and this has only been supported by subsequent phylogenetic analysis of viral genes. The taxonomic organisation of the *Arenaviridae* family that is accepted and recognised by the ICTV is shown in table 1.1.

As mentioned, Mammarenaviruses within clade D appear to have come about through recombination events. There is much evidence for this such as the S segment of clade D viruses having chimeric origin with an NP sequence closely related to clade A arenaviruses and a GPC sequence most closely related to clade B arenaviruses (Charrel et al., 2003, Archer and Rico-Hesse, 2002, Charrel et al., 2002).

Xapuri mammarenavirus (XAPV) represents an interesting recent phylogenetic dilemma. XAPV is a mammarenavirus detected in Musser's bristly mouse found in the Amazon and was assigned to the New World group, but further phylogenetic analysis could not assign this virus to a specific clade. It is phylogenetically related to New World clade B and C viruses and as such has features that are interesting for the Tacaribe virus serocomplex group and requires it to be placed as a divergent but sister

group of Clade C S segments and Clade B L segments which could be indicative of reassortment between these two clades. No previous reassortment mammarenaviruses have been isolated from nature before this, despite much indication of viral diversification during arenavirus evolution due to high mutation rates from low fidelity viral RdRP, recombination, and reassortment events. As such, XAPV may represent the first natural reassortment of the *Arenaviridae* family and a new unrecognised clade within the Tacaribe serocomplex (Fernandes et al., 2018).

The *Reptarenavirus* genus was established in 2014 due to the discovery of several novel arenaviruses isolated from captive snakes presenting with boid inclusion body disease (BIBD) which is characterised by formation of cytoplasmic inclusion bodies (IBs) comprising reptarenavirus NP within the infected cells and although not directly fatal, snakes often die from secondary bacterial, fungal, or protozoal infections, or neoplastic processes (Hetzel et al., 2013, Stenglein et al., 2012a, Bodewes et al., 2013, Chang and Jacobson, 2010, Schilliger et al., 2011). These novel reptarenaviruses possessed the characteristics defining them as arenaviruses including genome organisation, ambisense orientation of arenavirus ORFs, and characteristic gene regulatory elements but were distinct from mammarenaviruses when it came to sequence with low NP and L amino acid homology across the two genus'; 23-26% and 17-19% respectively (Stenglein et al., 2012b). Importantly, the reptarenaviruses showed high similarity to each other, with NP and L sequences sharing 55% and 50% respectively. This supported the necessity of a new genus and independent phylogeny within the *Arenaviridae* family. Interestingly, the reptarenaviruses displayed even greater divergence from mammarenaviruses within the Z sequence sharing only 16% homology and the GPCs shared no structural features usually associated with mammarenavirus GPC and actually shared higher homology with filovirus

glycoproteins such as ebola virus (Stenglein et al., 2012a, Garry and Garry, 2019).

The *Hartmanivirus* genus was created in 2018 (Maes et al., 2019b) after pairwise sequence comparison (PASC) analysis was performed on the BIBD-isolates and the results suggested that one of the novel reptarenaviruses, Haartman Institute snake reptarenavirus (HISV) was distinct enough from the other identified reptarenaviruses to require a new genus (Hepojoki et al., 2018). Further analysis via next generation sequencing and *de novo* sequence assembly identified three additional novel species as well as providing more information about HISV1, and extensive sequencing showed that whilst hartmaniviruses possess S and L segments, the L segment only codes for the L protein and does not code for the Z matrix protein (Hepojoki et al., 2018). A key difference between members of the reptarenavirus family and the hartmanivirus family is in the glycoproteins. Glycoproteins of hartmaniviruses actually share many of the structural motifs theorised to be present within those of mammarenaviruses and share higher sequence homology of 23% when compared to 16% for the Golden Gate reptarenavirus (GGV) and also for LCMV (Garry and Garry, 2019). Although one key point is that, as of yet, no hartmanivirus has been associated with any pathological effect in snakes perhaps due to the lack of Z matrix protein or due to differences in NP which may affect function when compared to reptarenaviruses (Baggio et al., 2021).

The *Antennavirus* genus was established recently (Abudurexiti et al., 2019). So far all viruses have been found in fish and their discovery came from a large-scale metatranscriptomic study aimed at understanding the evolutionary history of vertebrate RNA viruses by analysing reptiles, amphibians, lungfish, ray-finned fish, cartilaginous fish, and jawless fish (Shi et al., 2018a). The study identified 214 vertebrate associated viruses including two novel arenaviruses found in frogfish (*Antennarius striatus)* and subsequently named Wēnlǐng frogfish arenavirus 1 (WlFAV1) and

Wēnlǐng frogfish arenavirus 2 (WlFAV2). There have only been two other Antennaviruses discovered since; salmon pescarenavirus 1 (SPAV1), and salmon pescarenavirus 2 (SPAV2) found in endangered wild salmon in the Northeast Pacific during metatranscriptomic sequencing and surveillance of dead and moribund (approaching death) cultured Chinook salmon (Mordecai et al., 2019). Interestingly, antennaviruses have genomes that consist of three genomic segments rather than two and appear not to encode a protein homologous with Mamm- or reptarenavirus Z protein. Antennaviruses have one ambisense and two negative-sense RNA segments in which the termini of the RNAs contain inverted complementary strands most likely coding for transcription and replication initiation signals (Shi et al., 2018b). Information from sequence data suggests that antennaviruses express four structural proteins: NP, GPC, L, and a fourth protein of currently unknown function. The Small (S) RNA of antennaviruses codes for NP, the medium (M) RNA codes for the GPC and the unknown protein and the L RNA codes for the L protein containing the RNA-directed RNA polymerase (RdRP). It was determined that antennaviruses belong to the *Arenaviridae* family due to sequence homology between the antennavirus and hartmanivirus NP and L sequences (Pontremoli et al., 2019).

The *innmovirus* genus is the most recent addition to the *Arenaviridae* family (Chen et al., 2021) with only one currently known innmovirus; Hailar virus (HLRV) detected from high throughput sequences of river sediment samples collected in China and with an unknown host(s) and no reported isolates. The innmovirus genome consists of three genomic segments and appears not to code for any homologue of the mammarenavirus and reptarenavirus Z protein. Based only on sequence data it seems likely that innmoviruses express three structural proteins; NP, GPC, and L, with the small (S) RNA coding for the NP, the medium (M) RNA coding for GPC and L RNA encoding for the L protein containing RdRP (Chen et al., 2021)

Genus	Species	Virus name	Abbrev.
Antennavirus	*Antennavirus hirsutum*	Wēnlǐng frogfish arenavirus 2	WlFAV2
	Antennavirus salmonis	salmon pescarenavirus 1	SPAV1
	Antennavirus salmonis	salmon pescarenavirus 2	SPAV2
	Antennavirus striale	Wēnlǐng frogfish arenavirus 1	WlFAV1
Hartmanivirus	*Hartmanivirus brazilense*	SetPatVet virus 1	SPVV1
	Hartmanivirus haartmani	Haartman Institute snake virus 1	HISV1
	Hartmanivirus haartmani	Haartman Institute snake virus 2	HISV2
	Hartmanivirus helvetiae	Dante Muikkunen virus 1	DaMV1
	Hartmanivirus patriae	andere Heimat virus 1	aHeV1
	Hartmanivirus quadrati	big electron-dense squares virus 1	BESV1

	Hartmanivirus scholae	old schoolhouse virus 1	OScV1
	Hartmanivirus scholae	old schoolhouse virus 2	OScV2
	Hartmanivirus turici	veterinary pathology Zurich virus 1	VPZV1
	Hartmanivirus turici	veterinary pathology Zurich virus 2	VPZV2
	Hartmanivirus unni	Universidad Nacional virus 1	UnNV1
Innmovirus	*Innmovirus hailarense*	Hailar virus	HLRV
Mammarenavirus	*Mammarenavirus abaense*	Abà-Miányáng virus	AMYV
	Mammarenavirus alashanense	Alxa virus	ALXV
	Mammarenavirus allpahuayoense	Allpahuayo virus	ALLV
	Mammarenavirus amapariense	Amaparí virus	AMAV
	Mammarenavirus aporeense	Aporé virus	APOV
	Mammarenavirus batangense	Batáng virus	BTTV

	Mammarenavirus bearense	Bear Canyon virus	BCNV
	Mammarenavirus beregaense	Berega virus	BEGV
	Mammarenavirus bituense	Bitu virus	BITV
	Mammarenavirus brazilense	Sabiá virus	SBAV
	Mammarenavirus caliense	Pichindé virus	PICHV
	Mammarenavirus cameroonense	souris virus	SOUV
	Mammarenavirus chapareense	Chapare virus	CHAPV
	Mammarenavirus choriomeningitidis	lymphocytic choriomeningitis virus	LCMV
	Mammarenavirus choriomeningitidis	Dandenong virus	DANV
	Mammarenavirus cupixiense	Cupixi virus	CUPXV
	Mammarenavirus dhati-welelense	Dhati Welel virus	DHWV
	Mammarenavirus flexalense	Flexal virus	FLEV

	Mammarenavirus gairoense	Gairo virus	GAIV
	Mammarenavirus ganziense	Ganzi virus	GNZV
	Mammarenavirus guanaritoense	Guanarito virus	GTOV
	Mammarenavirus ippyense	Ippy virus	IPPYV
	Mammarenavirus juninense	Junín virus	JUNV
	Mammarenavirus kitaleense	Kitale virus	KTLV
	Mammarenavirus kwanzaense	Kwanza virus	KWAV
	Mammarenavirus lassaense	Lassa mammarenavirus	LASV
	Mammarenavirus latinum	Latino virus	LATV
	Mammarenavirus lijiangense	Lìjiāng virus	LIJV
	Mammarenavirus loeiense	Loei River virus	LORV
	Mammarenavirus lujoense	Lujo mammarenavirus	LUJV
	Mammarenavirus lunaense	Luna virus	LUAV

	Mammarenavirus lunaense	Luli virus	LULV
	Mammarenavirus lunkense	Lunk virus	LNKV
	Mammarenavirus machupoense	Machupo virus	MACV
	Mammarenavirus mafigaense	Mafiga virus	MAFV
	Mammarenavirus marientalense	Mariental virus	MRLV
	Mammarenavirus mecsekense	Mecsek Mountains virus	MEMV
	Mammarenavirus merinoense	Merino Walk virus	MRWV
	Mammarenavirus mopeiaense	Mopeia mammarenavirus	MOPV
	Mammarenavirus mopeiaense	Morogoro virus	MORV
	Mammarenavirus ngerengerense	Ngerengere virus	NGEV
	Mammarenavirus okahandjaense	Okahandja virus	OKAV
	Mammarenavirus oliverosense	Oliveros virus	OLVV
	Mammarenavirus paranaense	Paraná virus	PRAV

	Mammarenavirus piritalense	Pirital virus	PIRV
	Mammarenavirus praomyidis	mobala virus	MOBV
	Mammarenavirus ryukyuense	Ryukyu virus	RYKV
	Mammarenavirus solweziense	Solwezi virus	SOLV
	Mammarenavirus songeaense	Songea virus	SOGV
	Mammarenavirus tacaribeense	Tacaribe virus	TCRV
	Mammarenavirus tamiamiense	Tamiami virus	TMMV
	Mammarenavirus tietense	Tietê virus	TIEV
	Mammarenavirus wenzhouense	Wēnzhōu virus	WENV
	Mammarenavirus whitewaterense	Whitewater Arroyo virus	WWAV
	Mammarenavirus whitewaterense	Big Brushy Tank virus	BBRTV
	Mammarenavirus whitewaterense	Catarina virus	CTNV
	Mammarenavirus whitewaterense	Skinner Tank virus	SKTV

	Mammarenavirus whitewaterense	Tonto Creek virus	TTCV
Reptarenavirus	*Mammarenavirus xapuriense*	Xapuri virus	XAPV
	Reptarenavirus aurei	Golden Gate virus	GOGV
	Reptarenavirus californiae	CAS virus	CASV
	Reptarenavirus commune	tavallinen suomalainen mies virus 2	TSMV2
	Reptarenavirus giessenae	University of Giessen virus 1	UGV1
	Reptarenavirus giessenae	University of Giessen virus 2	UGV2
	Reptarenavirus giessenae	University of Giessen virus 3	UGV3
	Reptarenavirus rotterdamense	ROUT virus	ROUTV
	Reptarenavirus rotterdamense	University of Helsinki virus 1	UHV1

Table 1.1 ICTV Taxonomic Organisation of the Arenaviridae Family (Radoshitzky et al., 2023, Radoshitzky et al., 2015b, Pontremoli et al., 2018, Pontremoli et al., 2019)

1.1.3 Hosts and Transmission

Reservoir hosts of a virus are defined as populations, specific species, or whole ecological communities that drive disease dynamics by their ability to become infected by and sustain an infectious agent without experiencing detrimental disease symptoms (Haydon et al., 2002, Salkeld et al., 2023). A key characteristic of each mammarenavirus is the association with a single or small number of specific rodent reservoir hosts with the exception of Tacaribe (TCRV) isolated from Jamaican fruit bats and great fruit eating bats in the Caribbean (Downs et al., 1963) and also lone star ticks in Florida (Sayler et al., 2014). Table 1.2 gives information on the geographic distribution of the current *mammarenavirus* genus, including rodent reservoir hosts and the year in which the virus was isolated. Despite much effort, the rodent host for TCRV has not yet been identified and although bats were proposed to be the reservoir host, no evidence has been found that chronic persistent infection is caused and TCRV can actually be fatal in bats (Cogswell-Hawkinson et al., 2012, Malmlov et al., 2017). The reservoir hosts for almost all mammarenaviruses are rodents of the superfamily Muroidea (Bowen et al., 1997, Hugot et al., 2001). It is thought that this close association of mammarenaviruses with a specific rodent host determines the limited geographic distribution of the arenaviruses. LCMV has global distribution as it is found in the house mouse whereas Old World mammarenaviruses found in Africa are mainly found in rodents belonging to the *Mastomys* and *Praomys* genera of the Muridae family and Murinae subfamily. New World mammarenaviruses found in North and South America are mostly found in rodents of the subfamily Cricetidae with the aforementioned exception of TCRV. There have been unclassified mammarenaviruses that have been discovered in a number of mammals including bats (Bentim Góes et al., 2022), hedgehogs (Reuter et al., 2023), dipodoid rodents (three toed Mongolian jerboas) (Wu et al., 2018), and

pikas (small mountain dwelling mammals resembling rabbits) (Cui et al., 2023, Luo et al., 2023).

Some mammarenaviruses share a host such as LASV and east African Mopeia mammarenavirus (MOPV). Although MOPV is non-human pathogenic, infections can still occur in humans which can be beneficial to humans in regions endemic for MOPV and subsequently LASV as MOPV infection can provide cross-protection for LASV due to their genetic similarity (Zapata and Salvato, 2013). The genetic similarity of LASV and MOPV also comes in useful during research as infectious MOPV is classed as a category 2 biological agent whereas LASV is category 4, this means infectious MOPV can be used in CAT-2 facilities for infection studies.

Mammarenaviruses typically establish chronic persistent and usually asymptomatic infections within their rodent reservoir hosts characterised by chronic viraemia (presence of virus in blood) and viruria (presence of virus in urine) (Bowen et al., 1997, Hugot et al., 2001). The chronic carrier state in rodents results from vertical transmission (exposure to infectious virus early in ontogeny (development)) (Webb et al., 1975, Mills et al., 1992) and these chronic infections persist due to suppressed host immunity resulting from inability to mount an effective immune response due to recognition of viral antigens as self (King et al., 2018, Mariën et al., 2017). Due to close physical contact, inhalation of infected material and events such as bites between rodent host populations, horizontal transmission of the virus occurs and is maintained (Tagliapietra et al., 2009).

It has been suggested that arenaviruses persist in chronically infected rodent populations even during periods of low population density (Mariën et al., 2020, Milazzo et al., 2011). For persistence to occur, it is important that the virus does not replicate uncontrollably and over-run the host, leading to immune detection and mounting of a proper immune response to clear the virus. LCMV cycles through a period of acute productive infection and persistent non-productive infection to achieve persistence, this was shown using fluorescent detection of single RNA molecules in

LCMV infected cells. The infected cells would cycle through periods of active viral replication and transcription and then to periods in which no viral RNA could be detected (King et al., 2018). This has also been shown *in vivo* (Francis and Southern, 1988), and whilst the mechanisms by which arenaviruses are able to regulate this cyclical system are currently unknown, models involving defective interfering particles or semi-competent systems of replication competent/transcription incompetent genomes have been suggested (King et al., 2018).

Although most mammarenaviruses do not usually infect mammals other than their primary reservoir host they are still able to cross species barriers. Humans can become infected through direct contact with infected rodents or their excreta (droppings and urine), ingestion of contaminated food or inhalation of aerosolised droplets from contaminated rodent excreta, secreta, or body parts (if rodent has been caught in machinery such as a harvester during farming) (Charrel and de Lamballerie, 2003). Individuals who handle or consume contaminated rodents are also at risk of infection (Keenlyside et al., 1983).

The most common cause of arenavirus transmission from rodents to humans is habitat disturbance of reservoir host rodents leading to establishment of new territories and potentially increasing likelihood of closer contact to humans, their homes, and their food sources/supplies. It has also been found that, in geographic regions that experience rainy seasons, sudden increases in rodent populations occur also resulting in habitat and territory expansion and subsequent increased chances of human contact as well as increased horizontal arenavirus transmission within the rodent population (Tagliapietra et al., 2009, Sarute and Ross, 2017).

Nosocomial routes of infection are the most common cause of person to person transmission and is usually because of direct contact with bodily fluids or excreta from infected patients (Fisher-Hoch et al., 1995, Johnson, 1965, Paweska et al., 2009). Though it is usually quite rare, nosocomial

transmission is more common with mammarenaviruses that can cause viral haemorrhagic fever such as LASV and thus poses higher risk to frontline healthcare workers at the highest risk of exposure especially in endemic areas where medical practices are poor, and preventative measures such as personal protective equipment (PPE) and isolation techniques are lacking (Kernéis et al., 2009, Fisher-Hoch et al., 1995, Ogbu et al., 2007).

Cases of New World South American mammarenavirus infection most commonly occurred in male agricultural workers during harvest season when rodent populations are highly active (Radoshitzky et al., 2018). Junin virus (JUNV) is the cause of Argentinian haemorrhagic fever, and the areas affected are constantly expanding. Machupo virus (MACV) causes Bolivian haemorrhagic fever in Bolivia although usually isolated outbreaks. Guanarito virus (GTOV) causes Venezuelan haemorrhagic fever in Venezuela. Chapare virus (CHAPV) and Sabia virus (SBAV) have been isolated from human cases resulting in fatality in Bolivia and Brazil respectively (Radoshitzky et al., 2018).

Old World group mammarenaviruses causing human disease include Lassa mammarenavirus (LASV) which causes Lassa fever in West Africa and Lujo mammarenavirus which caused a small but very severe disease outbreak in Southern Africa in 2008 resulting in the deaths of four of the five people infected including the original first and three healthcare workers (Paweska et al., 2009). The fourth infected healthcare worker received ribavirin treatment and survived.

Human infection with LCMV can occur in rural or urban areas that have high rodent populations and has even been acquired from hamsters kept as pets (Hirsch et al., 1974). Laboratory-acquired infections have occurred resulting in severe disease with a number of mammarenaviruses including Flexal virus(FLEV), JUNV, LASV, LCMV, MACV, and SBAV (Leifer et al., 1970, Vasconcelos et al., 1993, Coelho and García Díez, 2015, Gaidamovich et al., 2016, Rugiero et al., 1962).

LASV and LCMV specifically can be vertically transmitted between mother and foetus. LASV infection during pregnancy can result in high mortality for both the mother and foetus and this mortality is higher in women who contract infection later in pregnancy (30% mortality in 2nd and 3rd trimester vs 7% in 1st trimester) whereas mortality is higher for the foetus infected during an earlier trimester (92% in 2nd and 3rd trimester vs 75% in 1st) with a 100% fatality rate in neonates (Ogbu et al., 2007, Price et al., 1988). The reason behind these high mortality rates is high concentration of LASV within the Fetal and placental tissues due to a lack of major histocompatibility complex (MHC) presentation by the cells making up these tissues resulting in the inability of the mother's immune system to recognise and attack infected cells (Ogbu et al., 2007). LCMV infection can also be passed transplacentally from mother to foetus resulting in death of the foetus or significant congenital deformities (Wright et al., 1997, Bonthius, 2012, Barton and Hyndman, 2000, Barton and Mets, 2001). There has even been evidence of sexual transmission of LCMV and LASV due to these viruses persisting in semen for months after the virus has been cleared and no symptoms show (Raabe and Koehler, 2017).

Virus	isolated	Geographic Distribution	Host Species
Old World			
Alxa virus	2014	China	*Dipus sagitta*
Lijiang virus	2015	China	*Apodemus Chevrieri*
Gairo virus	2015	Tanzania	*Mastomys natalensis*
Ippy virus	1985	Central African Republic	*Arvicanthis spp.*
Lassa mammarenavirus	1969	West African countries	*Mastomys natalensis*
Loei River virus	2016	Thailand	*Bandicota indica, Bandicota savilei, Niviventer fulvescens*
Lujo mammarenavirus	2008	Zambia and Republic of South Africa	*Unknown*
Luna virus	2009	Zambia	*Mastomys natalensis*
Lunk virus	2012	Zambia	*Mus minutoides*

Lymphocytic choriomeningitis virus	1933	Globally spread	*Mus musculus*
Mariental virus	2015	Namibia	*Micaelamys namaquensis*
Merino Walk virus	1985	Republic of South Africa	*Myotomys unisulcatus*
Mobala virus	1983	Central African Republic	*Praomys spp.*
Mopeia mammarenavirus	1977	Mozambique and Zimbabwe	*Mastomys natalensis*
Okahandja virus	2015	Namibia	*Micaelamys namaquensis*
Ryukyu virus	2017	China	*Mus caroli*
Solwezi virus	2016	Zambia	*Grammomys sp.*
Souris virus	2014	Cameroon	*Praomys spp.*
Wenzhou virus	2015	China	*Rattus noervegicus, Rattus losea, Rattus tanezumi, Rattus rattus, Suncus murinus*
New World Clade A			

Allpahuayo virus	2001	Peru	*Oecomys bicolor, Oecomys paricola*
Flexal virus	1977	Brazil	*Oryzomys spp.*
Parana virus	1970	Paraguay	*Oryzomys buccinatus*
Pichinde virus	1971	Colombia	*Nephelomys albigulairs*
Pirital virus	1997	Venezuela	*Sigmodon alstoni*
New World Clade B			
Amapari virus	1965	Brazil	*Oryzomys goeldi, Neacomys guianae*
Apore virus	2008	Brazil	*Oligoryzomys mattogrossae*
Chapare virus	2008	Bolivia	*Unknown*
Cupixi virus	1965	Brazil	*Oryzomys capito*
Guanarito virus	1989	Venezuela	*Zygodontomys brevicauda*

Junin virus	1958	Argentina	*Callomys musculinus*
Machupo virus	1963	Bolivia	*Callomys callosus*
Sabia virus	1993	Brazil	*Unknown*
Tacaribe virus	1956	Trinidad, West Indies	*Unknown*
New World - No Defined Clade			
Xapuri virus	2015	Brazil	*Neacomys musseri*
New World Clade C			
Latino virus	1973	Bolivia and Brazil	*Callomys callosus*
Oliveros virus	1996	Argentina	*Necromys obscuri spp.*
New World Clade D			
Bear Canyon virus	1998	USA	*Peromyscus californicus*

Tamiami virus	1970	USA	*Sigmodon hispidus*
Whitewater Arroyo virus	1997	USA	*Neotoma spp.*
Big Brushy Tank virus	2002	USA	*Neotoma albigula*
Catarina virus	2007	USA	*Neotoma micropus*
Skinner Tank virus	2008	USA	*Neotoma mexicana*
Tonto Creek virus	2001	USA	*Neotoma albigula*

Table 1.2 Organisation, Isolation, Distribution and Hosts of the Mammarenavirus Genus

1.1.4 Geographical Distribution

Due to the close association with specific rodent hosts, arenaviruses are usually present in distinct geographical regions matching the rodent host population territories infections outside these regions are rare unless a new territory is established. Table 1.2 shows the rodent reservoir hosts of mammarenaviruses found in the Americas, the table also shows the habitat type and country in which they can be found and the diseases associated with infection. Old World arenaviruses are found in Africa and New World arenaviruses are found in South America with Clade D viruses being found in North America. LCMV has global distribution due to the global spread of the house mice (*Mus musculus)* which is their reservoir host. There have also been recent discoveries of novel mammarenaviruses now classed as Old World in Asia and Southeast Asia (Flannery et al., 2023). These viruses are Wenzhou Virus (WENV) and Plateau Pika Virus (PPV). Figure 1.2 displays the geographic distribution of mammarenaviruses and the clades to which they belong. Table 1.3 shows the organisation, isolation, distribution and hosts of the Mammarenavirus Genus which are depicted in figure 1.2.

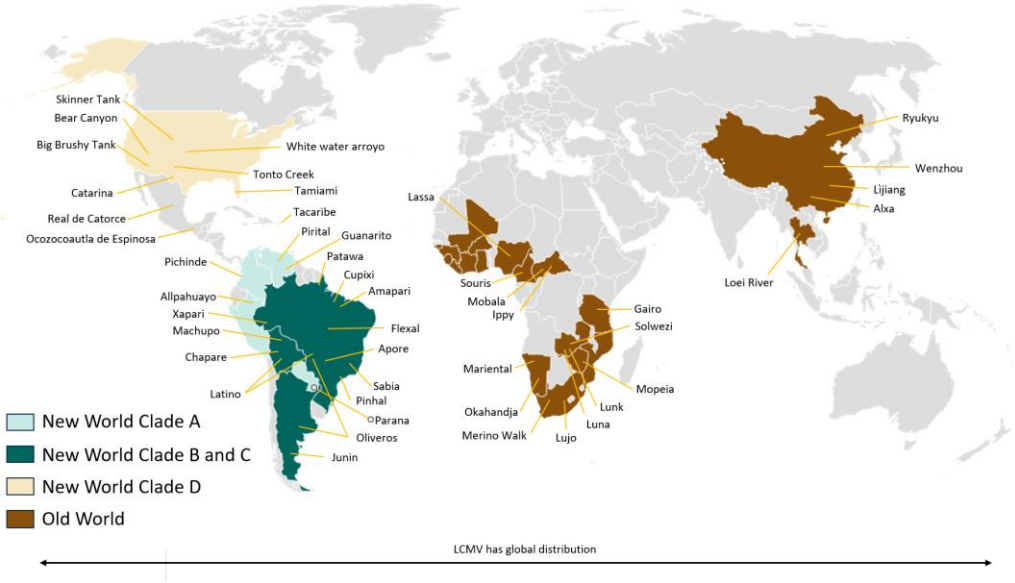


Figure 1.2 Map displaying geographic distribution of mammarenaviruses. Members of the Old-World group (brown) are mostly found in Africa with some more recently discovered species in China and Thailand. Members of the New World group (Blues and cream) are found in the Americas, with most members in South America and some species endemic to Northern America. This figure uses information from (Fehling et al., 2012, Radoshitzky et al., 2015b, Howley and Knipe, 2020, Sarute and Ross, 2017, Maes et al., 2019b)

1.1.5 Genetics and epidemiology of Lassa Fever

The most likely origin of LASV has been analysed using genetic epidemiology tools and it has been estimated that LASV originated over one thousand years ago in Nigeria (Andersen et al., 2015). It is then believed that LASV spread to neighbouring countries in West Africa recently over the past few centuries and has also undergone significant genetic divergence. LASV strains occupy a singly strongly supported large clade composed of four geographic lineages identified by phylogenetic analysis. Three distinct LASV genetic lineages are found within Nigeria; lineage I is found in the north-east and comprises the prototype LP strain of LASV originating from the village of Lassa. lineage II is found in the south region and Lineage III is most commonly found in the central Nigerian region (Ehichioya et al., 2019). Lineage IV can be found in Liberia, Sierra Leone, and Guinea (Bowen et al., 2000). The spread and diversification of LASV seems to still be occurring

with Nigeria reporting more cases than ever before in 2015 and an increase in reported cases with each subsequent year (Mofolorunsho, 2016, Akhiwu et al., 2018, Akpede et al., 2018, Ilori et al., 2019, Maxmen, 2018). Several new LASV lineages; V, VI, and VII have been discovered in the countries of Mali, Benin and Togo, Cote d'Ivoire, and also in Nigeria (Manning et al., 2015, Whitmer et al., 2018).

1.1.6 Pathogenesis and Disease

1.1.6.1 Lassa fever

Mammarenavirus infection in humans usually arises from direct contact with either an infected rodent or its excreta such as urine or faeces (Lapošová et al., 2013). Arenaviruses are a zoonotic disease that can also be transmitted between humans in healthcare associated infections otherwise known as nosocomial such as during pregnancy or via transplant with contaminated organs. A key trait of arenavirus infections in humans is the vast range of symptoms that can present from asymptomatic to fatal viral haemorrhagic fever. The most prevalent mammarenaviruses with the highest disease burden include LCMV and LASV from the Old-World group and JUNV, MACV, CHAPV, GTOV, and SABV all within Clade B of the New World group.

Lassa mammarenavirus infection causes a spectrum of clinical illness, ranging from asymptomatic infection to fatal haemorrhagic fever. Although reliable epidemiological data are lacking, it has been estimated that approximately 80% of infections with LASV are non-fatal or even asymptomatic, and only a small minority of persons develop severe disease (Garry, 2023). The relative contributions of these factors to disease severity are not well understood but may include differences in LASV strain, inoculation route and dose, genetic susceptibility of the host, or the presence of comorbid infections or conditions. The incubation period of LASV infection varies between 1 and 3 weeks, after which mild symptoms commonly begin. However, about 20% are believed to go on to develop severe haemorrhagic disease (VHF) within 4-7 days of initial mild illness.

Presenting symptoms may include fever, sore throat, vomiting, and coughing, common in most cases but can also differ in each case (Knobloch et al., 1980, McCormick and Fisher-Hoch, 2002, Richmond and Baglole, 2003). Approximately 40% of the patients with Lassa fever bleed from mucosal surfaces. Bleeding from mucosal surfaces shows poor prognosis. Case fatality rates for Lassa fever are very broad: while patients at the Lassa fever ward in Kenema, Sierra Leone, experience a case fatality rate of about 69%(Shaffer et al., 2014), case fatality rates in Nigerian settings lie between 20-30% under normal conditions, though outbreaks like the surge in 2015-2016 saw rates as high as 60% (Akpede et al., 2018, Akhuemokhan et al., 2017, Okokhere et al., 2018, Ehichioya et al., 2012). Death is seen around the 10th to 14th day from the beginning of symptoms; this process is usually shock due to reduced effective circulating volume leading to multi-organ failure (Buba et al., 2018, McCormick and Fisher-Hoch, 2002).

Laboratory findings have associated high LASV viral loads with increasing risks of fatality, along with elevated liver enzymes (indicating abnormal liver function); aspartate transaminase (AST) and alanine transaminase (ALT), and the kidney health marker, creatinine. Other common findings in the laboratory include thrombocytopenia and leukopenia, elevation of blood urea, and proteinuria (Shaffer et al., 2014). Certain factors predispose to this, such as age over 45 years in Nigeria and, in Sierra Leone, younger individuals; central nervous system manifestations and renal failure increase the fatality rate (Okokhere et al., 2018). Third-trimester pregnancy is associated with a particularly high mortality rate, but termination of pregnancy improves survival chances; however, infection with LASV is invariably fatal for the foetus (Okogbenin et al., 2019, Branco et al., 2011). The major long-term sequelae among survivors are neurological or visual deficits and permanent or temporary hearing loss, in approximately 30% of convalescents from Lassa fever. Though documented LASV recrudescence, similar to that described for Ebola virus, is not commonly

seen, it is an area for possible further investigation (Ficenec et al., 2019, Ficenec et al., 2020).

1.1.6.2 Clinical management of Lassa fever

The treatment for Lassa fever is largely supportive, including dialysis if renal injury has occurred. There are no approved drugs yet against Lassa fever; despite the uncertainties about its efficacy, ribavirin has remained a standard treatment. Early administration of ribavirin appears to improve recovery rates. The body mounts both innate and adaptive immunity after infection with LASV due to the ability of the innate immune system to recognize the presence of LASV double-stranded RNA, thereby initiating interferon responses. Of importance is that via different means, LASV manages to subvert the immune defense by inhibiting the activation of RIG-I and MDA5. The glycoprotein trimer of the virus protects it from humoral immunity, although adaptive immunity, with the chief help of T-cell responses, is important in virus clearance. In severe disease, delayed or impaired cellular immunity has often been recorded, and then it correlates with poor clinical outcomes.

1.1.6.3 Acute vs Chronic Persistent Arenavirus Infection

The wide range of disease severity and outcome can be attributed to differing IFN responses between individuals at the early stages of virus infection (Borrow et al., 2010). In mice it has been shown that a strong and effective IFN response against LCMV (Lymphocytic choriomeningitis virus) infection may inhibit early viral replication and induce CD8+ T cells which are virus specific and will clear the virus. However, If the IFN response at the early stages of viral infection is not potent enough this could result in a persistent infection and virus replication will not be suppressed to a level that would allow for CD8+ T cells to clear it during later infection (Wilson et al., 2013).

1.1.6.4 Lassa fever presents an unpredictable disease progression

As mentioned above, arenavirus infections can range from completely asymptomatic to inducing fatal viral haemorrhagic fever. The severity of a virus infection is often determined by how effective the host innate immune response is at successfully responding to and controlling the infection (Figure 1.3). Type I interferons (IFNs) are a crucial element of the host innate defence and are secreted from virally-infected cells and work to establish a non-specific antiviral state in these infected cells as well as nearby uninfected cells (Bowie and Unterholzner, 2008).

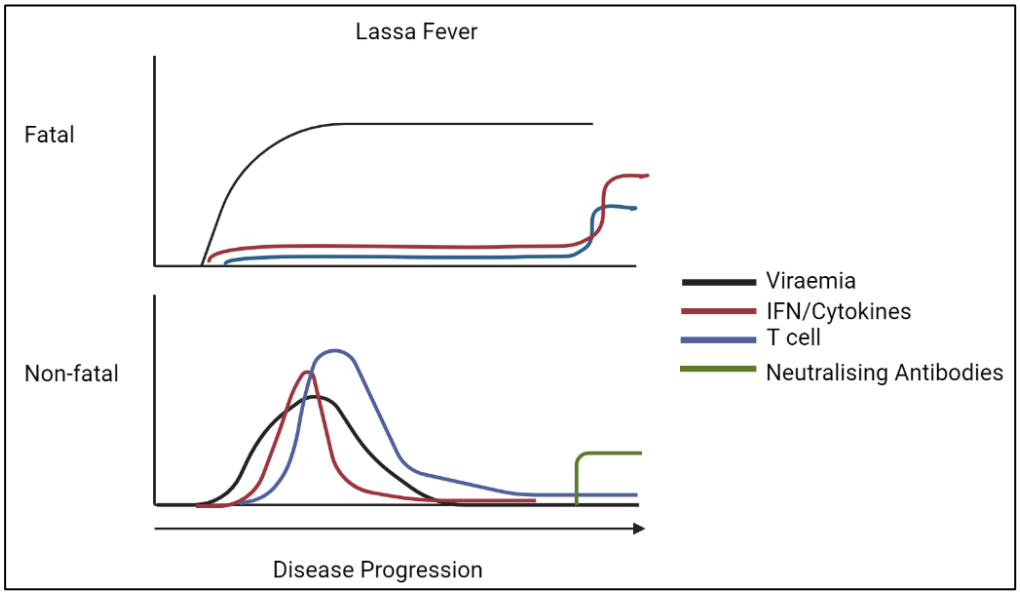


Figure 1.3: Comparison of immune responses in cases of fatal and non-fatal infection with Lassa mammarenavirus. Fatal Lassa fever infections show a lack of IFN and pro-inflammatory cytokines over the progression of infection. cases in which the infection was survived and/or completely cleared showed a strong T-cell response and much higher IFN levels early during infection although neutralising antibodies are not detected until later in infection. (Mantlo et al., 2019)

Interferons are also involved in stimulation and regulation of cells involved in both innate and adaptive immunity. These cells include natural killer (NK), NKT, T cells, dendritic cells, and macrophages. Interferons work in tandem with these immune cells as well as different chemokines and cytokines to control host innate immune response to virus infection (Medzhitov, 2007).

Innate immune responses are initiated by recognition of foreign ligands expressed by pathogens called PAMPS (pathogen-associated molecular patterns). This occurs through PRRs (pattern recognition receptors) which are expressed on host cells and have broad specificity allowing them to bind to many types of PAMPs (Medzhitov, 2007). PRR signalling results in IRFs (interferon regulatory factors) being activated as well as nuclear factor kB (NFkB) and other cofactors also being activated to trigger downstream regulators of innate immune responses and work to counter virus infection. Arenavirus proteins NP and Z inhibit and antagonise host innate immune function. For example, The NP protein of LASV has been shown to inhibit detection of PAMP viral dsRNA by the PRRs PKR, RIG-I, and MDA-5 which leads to immune suppression.

Induction of type I IFN expression is controlled by three different classes of PRR (Takeuchi and Akira, 2009, Honda et al., 2005): RLRs (Retinoic acid-inducible gene-1-like receptors), TLRs (Toll-like receptors) and NOD (nucleotide oligomerization domain)-like receptors. RLRs are made up of RIG-I (retinoic-acid inducible gene 1), MDA5 (melanoma differentiation-associated gene 5), and LPG2 (laboratory of genetics and physiology 2). These are cytosolic helicases that sense viral RNA (Yoneyama and Fujita, 2008a). RIG-1 and MDA5 become activated upon sensing and binding to viral RNA and then transduce signals downstream to IPS1 (IFN-beta promoter stimulator 1) found on the membranes of mitochondria. IPS-1 then acts to recruit the signal adapter TRAF3 (tumour necrosis factor-receptor-associated factor 3): the IPS-1, TRAF-3 complex is activated and in turn activates the IKKe/TBK-1 complex and the IKKa/beta complexes. The activated IkkE/TBK-1 complex works by phosphorylating IRF-3 (IFN regulatory factor 3) and/or IRF-7. IKKa/beta/gamma complex acts by phosphorylating NF-kB. Once activated, these transcription factors undergo nuclear translocation and start the expression of IFN-β and IFN-α as well as other proinflammatory cytokines.

RIG-I is a member of the RIG-I-like receptor family and is a PRR that detects positive and negative stranded RNA viruses in the cell cytoplasm (Yoneyama and Fujita, 2008a). RIG-I plays a key role in triggering host responses to many different viruses including Hepatitis C, Influenza A, and Measles virus. RIG-I recognises and is activated by RNA containing 5'-triphosphate modifications (3pRNA) (Hornung et al., 2006b). RIG-I's ATPase activity is kickstarted by ligand binding and this causes changes to structural conformation that subsequently allows RIG-I to interact via its amino-terminal tandem caspase-recruitment domain (CARD) with the adaptor protein MAVS (mitochondrial antiviral signalling protein). MAVS activates IRF3, IRF7, and NF-kB transcription factors which in turn produce type I IFNs and inflammatory cytokines resulting in innate antiviral immune response (Seth et al., 2005). There have been several studies that have shown that NP directly interacts with PRRs RIG-I and MDA-5 and downstream effectors IkB kinase e (IKKe) which causes inhibition of IRF3 activation (Meyer and Ly, 2016a, Martínez-Sobrido et al., 2007).

Arenaviruses have also been shown to activate the innate immune response through TLR activation once they bind viral PAMPs thus inducing an IFN1 response (Vidya et al., 2018). The NW virus JUNV induces activation of the TLR2/TLR6 complex which causes activation of AP-1 and NFkB transcription factors and subsequent activation of innate and adaptive immune responses (Cuevas et al., 2011) Interestingly, only TLR2 has been implicated in OW arenavirus infection and research suggests that OW LASV immunosuppression is cuased by suppression of TLR2-dependent cytokine induction (Hayes et al., 2012).

dsRNA accumulation during NW JUNV and MACV infections not only activates the RLRs and TLRs but also stimulates activation of the ubiquitously expressed PKR which is a host dsRNA sensor.

MOPV and LASV, despite being closely related induce differing T-cell responses during infection, (Pannetier et al., 2011). During a study, human dendritic cells infected with MOPV showed a strong a lasting CD8 and CD4

T cell response with high activation and proliferation. Whereas, human DCs infected with LASV underwent very weak memory responses (Pannetier et al., 2011, Huang et al., 2015a). The very different immune responses between these OW arenaviruses give clues to what causes the differing pathogenicity and areas to study further to build understanding of how arenaviruses interact with the human immune sytem.

1.1.6.5 Therapeutics and Treatments,

One of the main issues with arenaviruses is the lack of therapeutics or preventative measures such as vaccines. There are few options available and their effectiveness is debatable. Antivirals such as ribavirin have shown success in some cases but the efficacy is limited and the high toxicity and side effects can sometimes outweigh the benefits (McCormick et al., 1986). Favipiravir (T-705) is a small-molecule purine analogue (Furuta et al., 2017) with greater efficacy in treating LASV infection in several animal models than ribavirin (Safronetz et al., 2015, Oestereich et al., 2015, Rosenke et al., 2018). Two Lassa fever patients were treated with a combination of ribavirin and favipiravir and survived but developed transaminitis and viral RNA was able to be detected in their blood and semen samples for a long time after infection (Raabe et al., 2017).

Investigation into other therapeutics identified the novel drug LHF-535 which acts as an inhibitor of LASV viral entry by targeting the viral envelope glycoprotein. LHF-535 has been proven to be safe in a phase Ia human clinical trial (Larson et al., 2008)

Successful use of antivirals requires quick medical attention as well as rapid diagnosis of the correct virus at the early stages of infection. This is difficult in countries endemic for pathogenic arenaviruses that are also endemic for other diseases with similar symptoms and made nearly impossible if the patient is not aware of their infection before the onset of severe disease symptoms (Raabe and Koehler, 2017).

Another area worth exploring in the treatment of LASV and arenavirus infection, may be immunotherapeutics. Despite the humoral immune response appearing to be less involved in the host response to arenavirus infection, immunotherapy may still be provide substantial benefits. Sixteen (16) neutralising human monoclonal antibodies from survivors of LASV, some of which had long-term and multiple exposures to LASV have been successfully cloned and expressed (Robinson et al., 2016). Whilst investigation into immunotherapeutics such as neutralising antibodies is still ongoing, these recent successes present new and potentially effective options in the treatment of arenvirus infections.

Currently, supportive therapy is the most successful and important therapy strategy utilised and involves rest, hydration, pain relief medication, use of aspirin as a blood thinner, avoidance of intra-muscular injections, and sometimes sedation of the patient. In severe cases, presenting with viral haemhorragic fever, platelet transfusions and blood clotting factor replacements have shown to be useful (Howley and Knipe, 2020).

During the early stages of infection arenaviruses show low viraemia and so there is low risk of spread, as the infection progresses into later stages the level of viraemia increases and risk of spread also increases. Nosocomial spread is most likely with medical staff responsible for treating infected patients most at risk, especially if the patient has severe viral haemhorragic fever and during parenteral administration of treatments (treatment anywhere in the body other than the mouth and alimentary canal such as intramuscular, intravenous, and sub-cutaneous).

1.1.6.6 Vaccines

Candid #1 is currently the only successfully developed vaccine against arenaviruses and is most effective against the New World JUNV for which its licenced for use only in Argentina. The Candid #1 vaccine strain was created through serially passafing the JUNV XJ strain and once it was shown that it was safe and had high efficacy in preclinical studies using guinea pigs and rhesus monkeys, it was then trialled in agricultural workers in regions

endemic for JUNV such as Argentina (McKee et al., 2008). Interestingly, the mechanisms by which the Candid #1 strain underwent attenuation are not fully understood, making it difficult to recreate for other arenavirus vaccines (Enria and Barrera Oro, 2002).

LASV is on the World Health Organisation (WHO) revised list of diseases for which vaccine development is a priority. There are limiting stipulations for a successful vaccine candidate however, such as; low cost, highly effective, room-temperature stable, and "user-friendly". Recently the progress of twenty six LASV vaccine candidates listed by (Salami et al., 2019) has been reviewed (Isaac et al., 2022) as well as an additional eight also developed over the last eight years making a total of thirty-four. So far, thirty vaccines are still in pre-clinical stages and four have successfully progressed to clinical trials with the most promising candidates from 2019 being the vesicular stomatitis virus-vectored vaccine and the live-attenuated MV/LASV vaccine both currently in clinical trials (Isaac et al., 2022). Recently, three LASV vaccine candidates have demonstrated cross-protection against heterologous strains of LASV in non-human primates; the MeV-, VSV-, and modified VSV-based vaccines (Cross et al., 2020, Cross et al., 2022, Mateo et al., 2021). These vaccines work by targeting the GPC of LASV.

1.2 The mammarenavirus genome and virus structure

Arenaviruses are enveloped, bi-segmented, single-stranded RNA viruses that have an ambisense coding mechanism. Whilst it may appear they have a small and simple genome; the arenaviruses have complex replication and translational mechanisms and very multifunctional proteins that lead to a complex lifecycle during infection. The bi-segmented, ambisense genome codes for just four viral proteins: Nucleoprotein (NP), Zinc-finger containing matrix protein (Z), Glycoprotein precursor complex (GP1, GP2 and the

stable signal peptide (SSP), and an RNA-dependent RNA polymerase (L) (Figures 1.4 and 1.5) (McLay et al., 2014).

The viral ribonucleoprotein complex (RNP) that exists within the virus particle (Figure 1.5) consists of the most abundant of the four encoded proteins - the nucleoprotein (NP) which encapsidates the virus genomic segments NP and the RNA-directed RNA polymerase (L) protein, the least abundant, which mediates virus genome replication and transcription as well as cap-snatching for viral mRNA capping (Peng et al., 2020). The zinc-binding protein Z is a matrix protein responsible for the budding of nascent virus particles, and the glycoprotein precursor complexes (GPC) are cleaved by signal peptidase to generate a stable signal peptide and the heterotrimers consisting of GP1 and GP2. These three products form a tripartite heterotrimeric GP complex that inserts into the virion membrane. GP1 works to mediate cell-surface and internal receptor binding whereas GP2 is involved in virion-cell membrane fusion and thus cell entry (Figure 1.5) (York et al., 2004, Buchmeier et al., 1987).

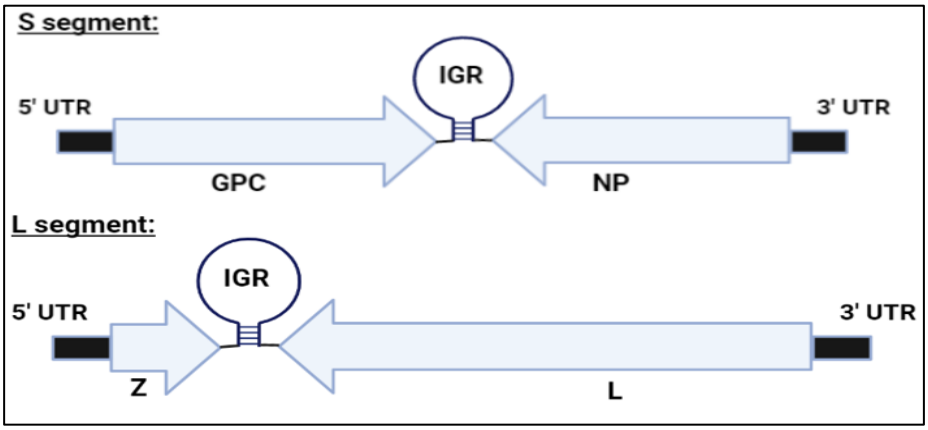


Figure 1.4 Arenavirus genomic structure. The genomic segments (vRNA; S and L) of mammarenaviruses encode four viral proteins: the nucleocapsid protein (NP) and glycoprotein complex (GPC) from the S segment, and the RNA-dependent RNA polymerase (L) and Z matrix protein (Z) from the L segment. The vRNA is bordered by untranslated regions (UTR), and the viral genes are separated by an intergenic region (IGR), which is thought to form a stem-loop structure.

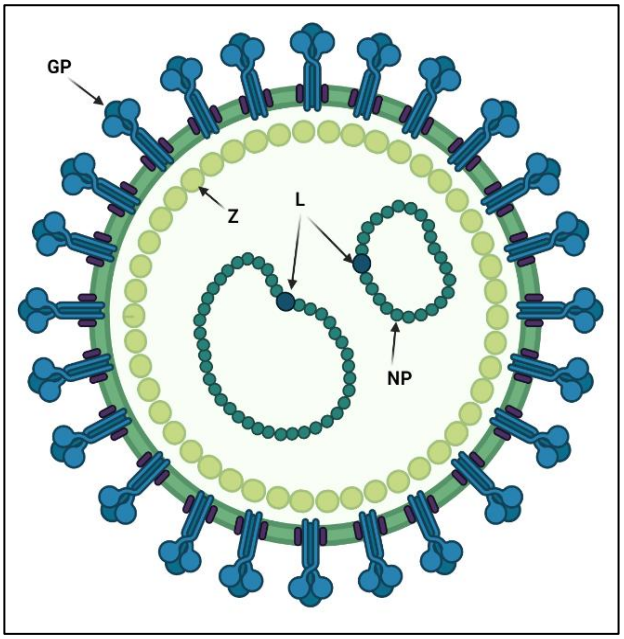


Figure 1.5 Arenavirus virion. The genomic segments are closely enclosed by the nucleocapsid protein (NP) and associated with the viral RNA-dependent RNA polymerase (L) to form the ribonucleoprotein (RNP) complex. The Z matrix protein (Z) oligomerizes to create a matrix layer on the interior side of the viral envelope. Glycoprotein spikes (GP) are embedded in the viral envelope, arranged in a trimeric structure. This figure was created with BioRender.

1.3 The Arenavirus Proteins

1.3.1 Glycoprotein Complex

The glycoprotein complex (GPC) is present on the viral surface and has a trimeric conformation. It is translated as a single polypeptide that is subsequently proteolytically cleaved into mature tripartite GPC made up of GP1, GP2, and the stable signal peptide (SSP) (Hastie and Saphire, 2018, Burri et al., 2012).

GPC mRNA translation is redirected to the endoplasmic reticulum (ER) when the ribosome reaches the SSP (York et al., 2004, Nunberg and York, 2012). A signal peptidase cleaves the GPC from the SSP although they remain associated. Interestingly, the SSP is different to most signal peptides in that it is much longer (58 amino acids compared to 18-30 amino acids) and also it has two hydrophobic regions instead of just one (von Heijne, 1985, York et al., 2004). The SSP becomes anchored to the

membrane by covalent binding of a myristoyl group to its second glycine residue (Nunberg and York, 2012, York et al., 2004). The two hydrophobic regions appear to be involved in membrane crossing which they do in an anti-parallel fashion, these hydrophobic regions are separated by a single positively charged lysine residue (K33) which is thought to interact with the GP2 ectodomain in the ER limen and thus is important for pH-mediated GP2 fusion (Froeschke et al., 2003). Arenavirus SSP also has roles in masking the ER retention motif of GPC which allows progression to the Golgi apparatus for downstream processing and subsequent trafficking to the plasma membrane (Agnihothram et al., 2006).

Immature GPC is translated as a trans-membrane protein that is embedded into the ER membrane. After translation GPC undergoes extensive N-glycosylation (Burri et al., 2012). Some of these N-glycosylation sites are required for correct proteolytic processing of GPC and this has been shown through mutation studies in which uncleaved GPC is trafficked to the plasma membrane instead of cleaved GPC (Eichler et al., 2006) although it is still unclear as to whether these mutations to N-glycosylation sites affect the overall structure thus preventing enzyme cleavage or if they are directly required for enzyme activity (Burri et al., 2012).

Glycosylation may be involved in masking the glycoprotein from the host immune response by way of forming a glycan shield preventing antibody access to the viral glycoprotein (Hastie and Saphire, 2018, Sommerstein et al., 2015).

The crystal structures of pre-fusion GPCs of LASV, LCMV, and UHV have been solved by X-ray crystallography (LCMV and LASV) or tomographic reconstruction (UHV and LASV) (Hastie et al., 2016b, Hetzel et al., 2013, Li et al., 2016). The resolution of these structures has revealed spikes on the viral membrane with trimeric conformation that are roughly 10nm wide and 9nm tall with leg like structures connecting the spikes to the membrane at 4nm intervals and could potentially correspond to the GP2 domain (Li et al., 2016).

The crystal structures give insight into the interactions between the GP1 and GP2 domains and the positions of residues necessary for receptor binding and regions that are exposed for antibody recognition (Hastie et al., 2016a, Hastie et al., 2017).

The GP1 domain (amino acids 59-265) is involved in mediating attachment to the host cell receptor and can be divided into three main structural features which are the N-terminal β-strand, the upper β-sheet face and the lower helix loop face (Hastie et al., 2016a). the interaction of GP1 and GP2 occurs when the N-terminal β-strand extends between the two to form a three-stranded, anti-parallel β-sheet.

α-dystroglycan (α-DG) is the host cell receptor for Old World arenaviruses but secondary receptors for specific viruses have been identified such as LASV requiring lysosomal associated membrane protein (LAMP1) (Jae et al., 2014) and LUJV requiring neutrophilin-2 (Raaben et al., 2017). It has been shown that LASV recognises a linear carbohydrate present on α-DG called matriglycan although the molecular mechanism by which matriglycan is recognised by LASV was unknown until recently (Katz et al., 2022). Matriglycan synthesis is mediated by LARGE1 (Kunz et al., 2005) and matriglycan is comprised of repeating units of xylose (Xyl) and glucuronic acid (GlcA) linked together to form $[\text{-3Xyl- } \alpha\text{1,3-GlcA- } \beta\text{1-}]_n$ (Inamori et al., 2012). These linear matriglycan chains can contain many repeats allowing them to project significant distances from the surface of the cell. LASV hijacks matriglycan to gain cellular entry via α-DG and recent structural data for both the LASV spike complex and LARGE1 have revealed how matriglycan is synthesised and how it is used as an arenavirus entry receptor (Katz and Diskin, 2024). Through the use of cryo-EM the complete, trimeric structure of the LASV spike complex has been determined (Katz et al., 2022). The receptor binding site where matriglycan binds has been shown to be quaternary and is formed through shared interactions between each of the three monomers making up the spike. This interaction occurs at the apex of the spike. In the centre of the binding site there are three C-

terminal regions of the GP1 domains from each adjacent monomer which undergo domain-swapping to create interlocked domains. These interlocked domains provide structural stability and create a cleft exposed to the surface which matriglycan binds to (Katz and Diskin, 2024).

There have been five key residues identified within GP1 (136, 153, 155, 190, 260) that mediate arenavirus binding to α-dystroglycan. The lower helix loop face of GP1 contains the Ser153 and Leu260 residues which are the reason for high affinity binding to α-dystroglycan and also crucial for establishment of persistent viral infection (Hastie et al., 2016a, Sevilla et al., 2000, Teng et al., 1996, Smelt et al., 2001). Binding to α-dystroglycan likely requires complete GPC to ensure correct conformation of GP1 which will allow receptor binding.

The GP2 domain (amino acids 266-498) is involved in directing fusion between virus and host cell membranes which facilitates the release of RNPs into the cell cytoplasm. The GP2 structure has been solved using x-ray crystallography for both the pre-fusion (Hastie et al., 2016a, Hastie et al., 2017) and post-fusion complexes (Igonet et al., 2011). GP2 has three structural regions; the N-terminal α-helix, a C-terminal α-helix and a T-loop region connecting the two α-helical structures (Igonet et al., 2011).

GP2 is a class I viral fusion protein despite possessing features of all three fusion protein classes (Eschli et al., 2006). The properties of GP2 which make it a class I fusion protein is because it needs to be proteolytically processed to release fusion peptide and it has characteristic class I α-helices forming the six-helix bundle post-fusion. Arenavirus GPC contains two heptad repeats which undergo significant conformational changes from pre- to post-fusion GP2. Pre-fusion, the first heptad repeat (HR1) appears as four segments (HR1a to HR1d) which all form helices with the exception of HR1b which presents as an extended loop. The T-loop appears as two anti-parallel β-strands in the pre-fusion form that has glycosylation sites (Hastie et al., 2016a)

Post-fusion the heptad repeats form the N- and C-terminal helices which are connected by the T-loop, also possessing α-helical structure. When the two heptad repeats join together along with the other two protomers, they form the anti-parallel six-helix bundle characteristic to class I viral fusion proteins (Hastie et al., 2016a)

1.3.2 Nucleocapsid Protein

The arenavirus nucleoprotein (Figure 1.6) (also sometimes referred to as nucleocapsid protein) (NP) is the most abundantly expressed of all the arenavirus proteins and accounts for over 70% of the protein present in purified virus and infected cells. The size of NP ranges between 558 and 570 amino acids with a molecular weight of 60-68kDa depending on the virus (Buchmeier, 2002). The viral gene coding for NP is located on the S gene segment and is encoded in a negative-sense for genome transcription to yield a coding mRNA (Meyer et al., 2002). Interestingly, during the persistent phase of arenavirus infection, as the expression levels of the other viral proteins decreases to near or totally undetectable amounts, the level of NP remains high (Ellenberg et al., 2002).

The function of the nucleocapsid protein (NP) of arenaviruses is to encapsidate the RNA segments within a ribonucleoprotein complex (RNP) which is vital for viral RNA synthesis and for facilitating viral transcription and replication.

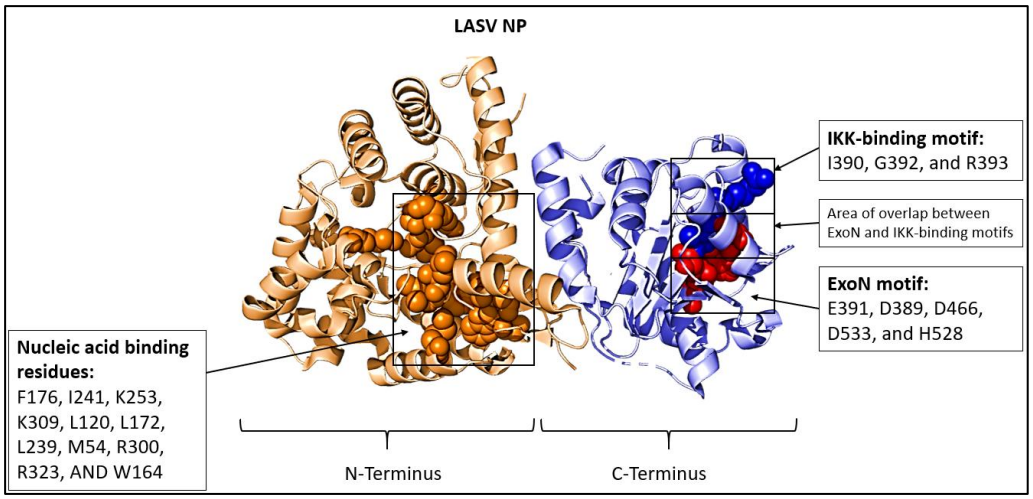


Figure 1.6 The overall protein structure of full-length LASV NP (PDB: 3MWP) shown as a cartoon with distinctions between the N-terminal (orange) and C-terminal (blue) domains. The N-terminal domain is shown with RNA bound in the deep, basic RNA binding groove. The C-terminal domain is the exoribonuclease (ExoN) domain which binds and degrades dsRNA. Adapted from (Stott et al., 2020)

The RNP is comprised of multiple NP copies which line up along the RNA segment in association with the L protein and NP acts as a scaffold within the RNP for a functional template for RNA synthesis by the L protein and to assist in Z matrix protein mediated recruitment of the segments into newly formed viral particles (Iwasaki et al., 2015).

NP is proposed to be an extremely multifunctional viral protein with roles in modulating host immune responses to infection and attenuating the host interferon type 1 (IFN-I) response as well as counter-acting host apoptotic responses and cellular factor recruitment to promote viral lifecycle (Martínez-Sobrido et al., 2007, Shao et al., 2018)

1.3.2.1 Structure of NP

The primary amino acid sequence of arenavirus NP is highly conserved across the whole *Arenaviridae* family even displaying high degrees of similarity between Old World group and New World group viruses (Peng et al., 2020, Qi et al., 2010b, Buchmeier, 2002). The crystal structures of NP from LASV, LCMV, JUNV, TCRV, and MOPV have been solved by a number of different groups to show that NP folds into two discrete domains, an N terminal domain from residues 1-388 approximately (LASV NP) and a C

terminal domain from residues 364-561 with an unstructured, flexible linking region connecting the two domains (Qi et al., 2010b). The structures (Figure 1.6) have been solved either as full length (Brunotte et al., 2011b, Qi et al., 2010b, Hastie et al., 2011c), or as individual N- and C-terminal domains, with the N-terminal domain being comprised mainly of α-helices forming head and body regions (Hastie et al., 2011b, West et al., 2014b, Yekwa et al., 2017, Zhang et al., 2013). Initially, the flexible N- and C-terminal linker was proposed as the groove involved in binding genomic RNA (Qi et al., 2010b), but since the structure of the N-terminal domain was solved in complex with single-stranded RNA, a deep cavity within the N-terminal domain is lined with basic residues and a patch of hydrophobic amino acids at the bottom of it is now thought to be the gRNA binding site (Hastie et al., 2011c). Initially was thought to be the site which binds the 7-methylguanylate (M^7G) cap structures, which arenaviruses uses for a unique mechanism of cap snatching cellular mRNAs to use as primers to initiate viral transcription. There have been conflicting studies into the cap snatching role of NP, residue mutation studies by (Qi et al., 2010a) in which a panel of NP mutants with alanine substitution of residues located inside and at the cavity entrance were tested for their ability to mediate cap-dependent viral RNA transcription showed either complete loss of RNA transcription activity or significantly decreased activity when compared to wild-type although none of these mutants affected NP function in IFN suppression. Whereas the study by (Brunotte et al., 2011b) showed an impaired ability to synthesis cRNA rather than affecting viral transcription suggesting that the N terminal cleft/groove does not actually bind M^7G caps and is instead responsible for the binding of single stranded viral RNA. Any attempts to co-crystallise LASV NP with cap-conjugated beads were unsuccessful and it has been shown that the basic amino acid residues that line the groove are key for mediation of the electrostatic interactions between NP and the phosphate backbone of single-stranded RNA molecules (Hastie et al., 2011c). Comparisons of full length LASV NP N-terminal domains that have bound RNA and those that are free from RNA

have shown structural rearrangements that likely represent a gating mechanism; in the RNA-bound domains, α-helix 5 is shortened to terminate before the RNA-binding groove and α-helix 6 and the connecting loop are shifted away from the groove allowing RNA access to the residues in the loop and the groove (Hastie et al., 2011c).

Insight into the roles of the NP C-terminal domain have been shown by studies of LASV NP crystal structures; two studies independently showed significant 3D structure similarity between the C-terminal domain of LASV NP and the characteristic fold of DEDD superfamily of exonucleases despite not sharing much primary sequence similarity (Hastie et al., 2011b). The C terminal domain is comprised of both α-helical structures and β-sheets in which the five-stranded β-sheet contains one anti-parallel strand and is surrounded by six α-helices all connected by flexible loops (Hastie et al., 2011b). A zinc atom is coordinated by residues present in the second β-sheet, the sixth α-helix, and the long loop containing basic residues that connects the fifth and sixth α-helical structures. The crystal structures also reveal a deep cleft at the top of the LASV NP C-terminal domain with catalytic residues occupying the same positions as in the active site of the human DEDD family 3' to 5' exonuclease TREX1.

Exonucleases are enzymes involved in catalysing cleavage of phosphodiester bonds between nucleotides at either the 3' or 5' end of the nucleotide chain and members of the DEDD superfamily are characterised by active sites containing Asp-Glu-Asp-Asp catalytic residues as well as having a neighbouring histidine or tyrosine residue and further sub-divides the superfamily into DEDDh or DEDDy (Zuo and Deutscher, 2001). Interestingly, the substrate for arenavirus NP exonuclease activity has been identified as double-stranded RNA (dsRNA), not DNA, which is digested in a 3' to 5' direction (Hastie et al., 2011b) and this activity is dependent upon a divalent Mn^{2+} cation coordinated by the C-terminal domain. It has also been shown that NP degraded both ssRNA and dsRNA with a clear preference for dsRNA.

LCMV, JUNV, and TCRV NP have also been shown to possess C-terminal 3' to 5' exoribonuclease domains, with TCRV NP having specificity for dsRNA ligands similar to LASV NP (Jiang et al., 2013). However, JUNV NP has been shown to lack this exoribonuclease activity through *in vitro* dsRNA degradation assays (Zhang et al., 2013). It has been suggested that the methods of protein crystallisation may have caused the protein to adopt an inactive conformation that lacked a vital Zn^{2+} cation conserved in all other crystallised NPs (West et al., 2014b, Zhang et al., 2013) but it's also entirely possible that JUNV does not actually possess the ability to degrade dsRNA substrates at all.

One of the other motifs present in arenavirus NP is a zinc finger motif located in the C-terminus (Tortorici et al., 2001, Qi et al., 2010b, Parisi et al., 1996) and is located between amino acids 500 and 530 in LASV and JUNV. The zinc finger motif position in LASV NP's crystal structure suggests it may have a role in coordination of the exoribonuclease domain (Tortorici et al., 2001).

Arenavirus NP has been shown to form homo-oligomers which are vital for encapsidation of the viral RNA genome and it is thought that the interactions between these individual NP monomers is mediated by the NP N-terminal domain (Ortiz-Riaño et al., 2012a, Levingston Macleod et al., 2011). C-terminal to N-terminal interactions have been shown to be important (Brunotte et al., 2011a). X-ray crystal structures for LASV NP have identified homo-trimers formed from monomers that associate in a ring-like structure (Brunotte et al., 2011a, Hastie et al., 2011c). It is thought that NP forms these trimers during viral transcription to prevent RNA binding and help to maintain levels of stored NP (Pattis and May, 2020) and when RNA replicates, a currently unknown signal is sent to initiate disassembly of the NP trimer thus shifting the C-terminal domain away from the N-terminal domain and opening up the RNA binding groove allowing NP to bind newly synthesised gRNA (Hastie et al., 2011c, Pattis and May, 2020). It has been shown that NP which is unable to trimerise, becomes defective and is

unable to effectively promote transcription and replication in minigenome reporter assays but does not affect interactions with other viral proteins or NP's role in antagonising the host IFN-I response (D'Antuono et al., 2014, Lennartz et al., 2013). Although the functional importance of these truncated NP cleavage products is as yet unknown. At late time points of infection (up to 20 days), staining of NP in nuclear inclusions can be observed within infected cells but only using monoclonal antibodies specifically able to recognise these short truncated forms (Young et al., 1987). There has been some work to elucidate the role of these truncated forms of NP using observations of JUNV in which JUNV NP was cleaved by caspases and expression of ectopic NP appeared to prevent effector caspase 3 activation which is a crucial event in apoptosis (Wolff et al., 2013) so it could be that these NP cleavage products are involved in the virus trying to avoid or prevent host cell apoptosis.

Although most NP expressed within infected cells has a molecular weight of 60-68kDa, itis possible to find and visualise 28kDa and 36kDa NP cleavage products (Harnish et al., 1981)

1.3.2.2 Functions of NP

The arenavirus NP is multifaceted within its roles in the cell during infection. It is essential in the formation of replication-transcription complexes (RTCs), which serve as discrete sites for viral RNA synthesis and are associated with specific cellular membranes, phosphatidylinositol-4-phosphate (PI4P), and various host proteins that include translation initiation factors and stress response proteins. A key role of NP is also in the recruitment of translation initiation factors, though this process varies between Old and New World arenaviruses, with New World viruses showing unique interactions that facilitate viral protein synthesis. Additionally, NP plays a key role in modulating the host immune response by degrading viral double-stranded RNA, thus evading host immune detection. Below we examine the mechanistic basis of these functions and highlights how NP's

immune suppressive abilities are pivotal for viral persistence and replication.

1.3.2.2.1 Modulation of host cell immune response

Perhaps the most important and least well understood role of arenavirus NP is in the modulation of the host immune response during infection. The binding of NP to viral RNA is necessary for viral nucleocapsid complex formation and for the characteristic suppression of host immune responses by acting as an interferon antagonist. It is because of its role as an orchestrater of immune suppression and its significance to arenavirus lifecycle that NP was chosen as the viral protein of focus for the research presented in this thesis. The mechanisms by which NP acts as an immune modulator and inhibitor are shown in figure 1.7.

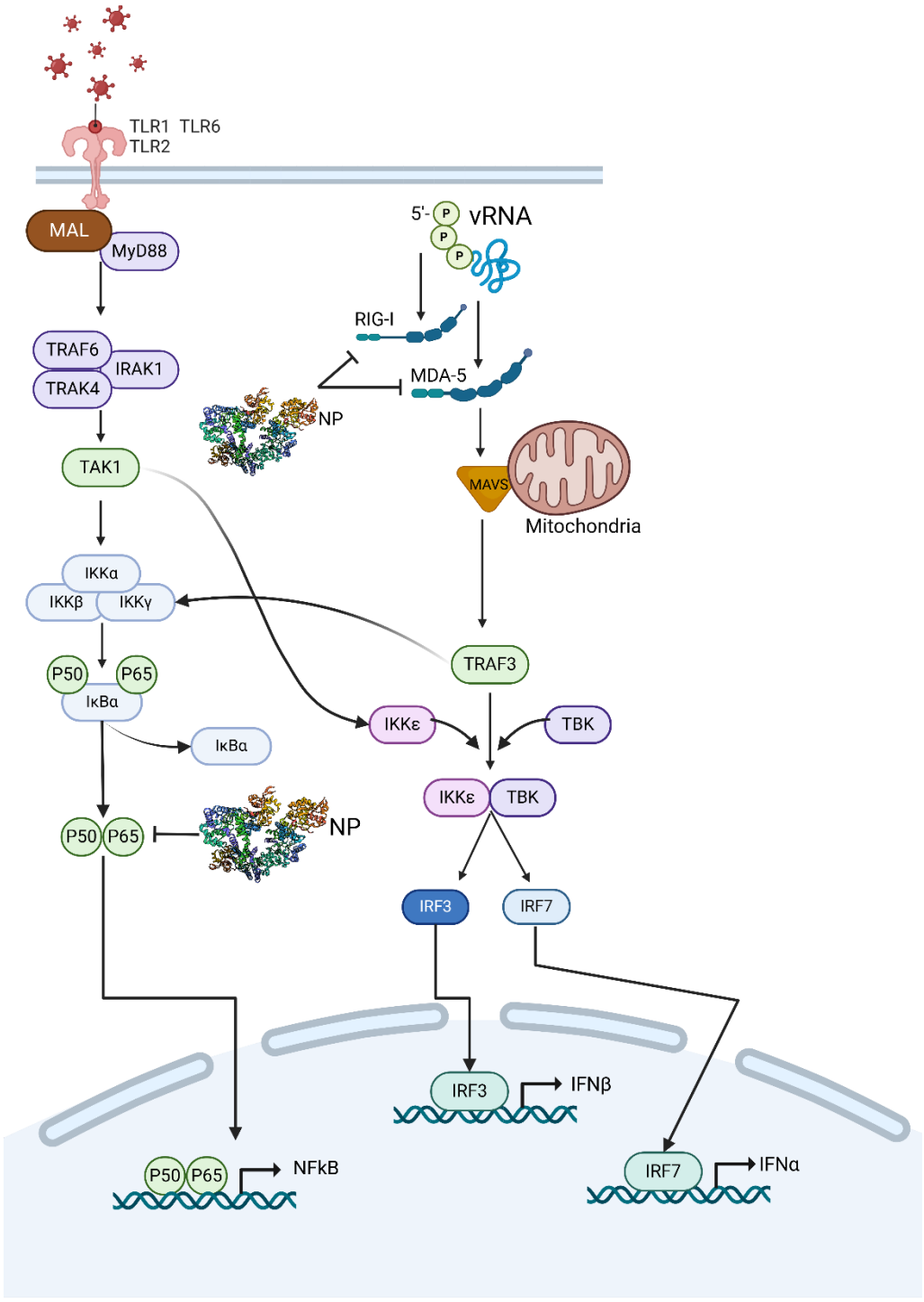


Figure 1.7 Inhibition of Innate Immune Response Steps by the Arenavirus Nucleocapsid Protein. This diagram illustrates the type I interferon (IFN) induction signalling pathways. Arenavirus non-polyadenylated RNA and double-stranded RNA are recognized by RIG-I and MDA5. The arenavirus NP disrupts this recognition and signalling by degrading RNA species and potentially binding directly to RIG-I. Additionally, arenavirus Z binds to RIG-I and MDA5 to block activation of the mitochondrial antiviral signalling (MAVS) protein. NP further inhibits later induction stages by binding to IKK-epsilon, thereby preventing the activation of IFN-responsive factor 3 (IRF3), and by blocking NF-κB activation, which normally occurs through Toll-like receptor (TLR) recognition of the arenavirus GPC. Adapted from (Meyer and Ly, 2016b)

NP acting as an IFN1 antagonist is a conserved characteristic across both OW LCMV and LASV as well as NW viruses, JUNV and MACV, it is also conserved in non-human pathogenic strains (Martínez-Sobrido et al., 2007). NP inhibits activation of IFN-β and IRF-3-dependent promoters and also inhibits the nuclear translocation of IRF-3 (Martínez-Sobrido et al., 2007). The immunosuppressive ability of NP can be linked to its C-terminal exonuclease activity (Hastie et al., 2011a). NP has been linked to double-stranded RNA degradation, and it has a high specificity for this in the 3'-5' direction. This mechanism of degrading viral dsRNAs, that would normally stimulate the host innate immune system via PRRs, is a common mechanism deployed by many arenaviruses to evade host innate immune responses (West et al., 2014a). The amino acid residues E391, D389, D466, D533, and H528 (LASV protein numbering style) make up the exoribonuclease DEDDh motif (ExoN) which is highly conserved. These residues are necessary for ExoN activity and thus dsRNA degradation (Mateer et al., 2020). It is interesting to consider the different mechanisms of dsRNA degradation utilised by OW and NW viruses. Use of the 9D5 monoclonal antibody (MAb) with a high affinity for dsRNA in conjunction with LASV, JUNV, and MACV minigenome replication systems by Mateer et al. showed that LASV NP ExoN activity was needed for limiting the accumulation of dsRNA and that mutations disrupting the LASV ExoN caused dsRNA to accumulate (Mateer et al., 2020). This inhibition of dsRNA was tested for in JUNV by co-infecting with LASV and it was shown that LASV NP expression alone did not limit dsRNA accumulation until it LASV L protein was also introduced, this suggests that there is cooperative degradation with activity of viral replication (Mateer et al., 2020). The action of NP to degrade viral dsRNA might also work to suppress IFN1 via preventing protein kinase R (PKR) signalling and the subsequent downstream eIF2α phosphorylation which will then inhibit host and viral protein translation. A study performed by Huan et al. using JUNV Romero strain showed that JUNV and MACV but now LASV infection resulted in high levels of activated PKR.

Research into the pathogenesis of other mammarenaviruses can often provide insight into how LASV NP might work. Due to the highly pathogenic nature of LASV and the strict laboratory and security measures around research involving live LASV, working on mammarenaviruses such as MOPV or LCMV is a necessary alternative and may shed light on similar mechanisms found in LASV. Mutagenesis studies performed on the NP of LCMV mapped the IFN1 inhibitory activity to the C-terminal region and identified the specific amino acid residues 382, 385, and 386 as crucial for antagonising IFN1 induction, these residues are located within a DIEG NP motif that is highly conserved among both OW and NW mammarenaviruses (Martínez-Sobrido et al., 2007). The study also showed that these LCMV mutants did not have any effect of viral RNA replication or infectious viral particle production which suggests that the IFN1 antagonistic ability of NP is separate from its roles within replication and the general lifecycle.

1.3.2.2.2 Formation of replication-transcription complexes

NP forms discrete sites for arenavirus RNA synthesis which are called replication-transcription complexes (RTCs) and are thought to be where vRNA and cRNA replication and transcription occurs (Baird et al., 2012). These RTCs are associated with a cytosolic membrane, a phosphatidylinositol-4-phosphate (PI4P), and a number of key host proteins that includes translation initiation factors, ribosomal proteins and the stress granule protein G3BP1 (Baird et al., 2012). The role of PI4P is to promote membrane dynamics and in the recruitment of cellular proteins for vesicular budding and in other viruses it has roles in membrane remodelling (Baird et al., 2012). These sites however have not been associated with any membrane-bound organelles and do not appear to export any viral mRNAs for translation, in fact, NP is thought to be the main effector for the formation of RTCs due to evidence from immunofluorescence staining and analysis of cells transfected with an NP-

expressing plasmid that reveals the characteristic discrete puncta seen in infected cells (Baird et al., 2012, Knopp et al., 2015). LCMV NP residue T206 is suggested to be involved in establishment of these sites, this corresponds to LASV NP T210 which is shown to be solvent exposed and located within α-helix 8 within the LASV NP N-terminal domain (Hastie et al., 2011c). Whilst it is solvent exposed, T206 is conserved across all mammalian arenaviruses highlighting the importance of the site to the replication-transcription functions of NP. These sites occur when T206 is transiently phosphorylated and mutation studies of T206 residues unable to be phosphorylated (T206A) show changes in NP distribution from discrete puncta to more diffuse cytoplasmic distribution (figure 1.8). Recombinant LCMV that has this T206 mutation is unable to be recovered from a reverse genetics system (Knopp et al., 2015) and mutating T206 to a phosphomimetic residue (T206E) which stimulates constitutive phosphorylation causes a phenotype similar to wild-type NP and mutant NP T206A in which the both discrete puncta and diffuse cytoplasmic distribution are shown (Knopp et al., 2015).

It is thought that the mechanism by which NP drives the formation of RTCs is by either aggregation of cellular proteins or interaction with them and although this is not yet fully understood, colocalization of NP at the RTCs with PI4P could be what drives formation of these sites and further investigation is required to understand the exact interactions of NP necessary for RTC formation and the role of phosphorylation in these interactions. Thus far there has been no evidence that the L protein is also found within the cytosolic puncta that form during arenavirus infection driven phosphorylation of T206. Why mRNAs are not also encapsidated in the puncta is also not yet understood and the exact mechanisms driving vRNA packaging mediate by NP are yet to be established (Sänger et al., 2023).

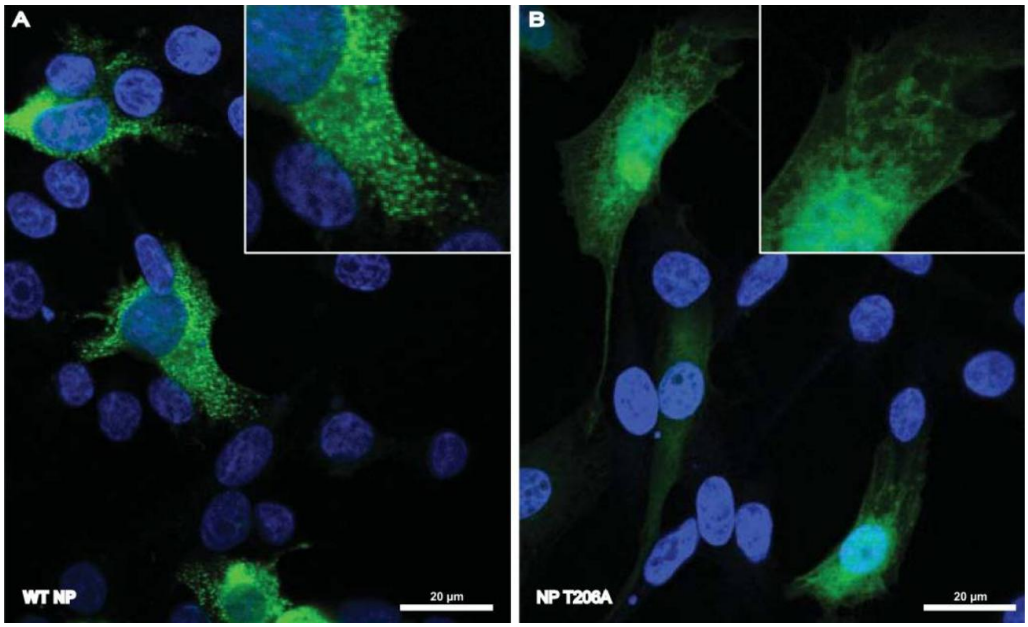


Figure 1.8 Distribution of WT NP and NP T206A in Cells : BHK-21 cells were fixed 48 hours after transfection with plasmids expressing either WT NP (A) or NP T206A (B). NP was visualized using an anti-LCMV NP antibody with Alexa Fluor 488 (green), and nuclei were stained with DAPI (blue). Figure has been obtained from (Knopp et al., 2015)

1.3.2.3 Role of virus-Host interactions that involve the LASV Replication Complex

One of the key host proteins found to play a role in the LASV replication complex is DDX3, a protein belonging to the DEAD (Asp-Glu-Ala-Asp) box RNA helicase family, which possesses ATPase and RNA helicase activity (Loureiro et al., 2018). During proteomic studies of arenavirus infected cells DDX3 has been identified as an interacting partner of not only LASV NP but also LCMV NP and the New World group Junin virus NP (King et al., 2017a). Deletion of the DDX3 gene using CRISPR/Cas9 systems results in a significant reduction in viral titre during infection for LASV, LCMV, and JUNV. Reconstituting DDX3 expression in these knockout cells causes recovery of viral replication of LASV, LCMV, and JUNV suggesting that DDX3 plays a significant role during viral replication as a proviral cellular factor (Loureiro et al., 2018).

Other cellular proteins with functions involved in RNA synthesis and assembly of the RNP complex include DDX5 (also known as RNA helicase p68) and DHX9 (also known as RNA helicase A), which are both members

of the DExD/Hbox protein family. These two proteins have been identified as potentially overlapping targets of LASV NP, JUNV NP, and both LCMV L and NP (Loureiro et al., 2018, King et al., 2017a, Khamina et al., 2017). The roles of these proteins during arenaviral infection has not yet been investigated but work has been done in other RNA viruses such as Influenza A, HCV, Chikungunya virus, and Porcine reproductive and respiratory syndrome virus (Goh et al., 2004, Bortz et al., 2011, Liu et al., 2016, Matkovic et al., 2019)

1.3.2.4 Variations in Translation Factor Recruitment by Arenavirus NP

It is thought that NP has other vital roles within the RTC which include interaction and recruitment of translation initiation factors although it has been shown that there is variation in how this works between different mammarenaviruses. LCMV studies have shown that NP colocalises with all of the three eukaryotic initiation factor 4Fs (eIF4F) but does not co-immunoprecipitate with any of them (Knopp et al., 2015) which suggests that the eIF4F complex gets recruited to the RTCs within LCMV infected cells due to NP interactions which then don't interact strongly with or replace any of the eIF4F components.

In New World mammarenaviruses JUNV, TCRV, PICV, it been proposed that NP replaces the eIF4F complex cap binding protein and studies have shown that NP colocalises and also co-immunoprecipitates with both EIF4A and eIF4G in infected cells (Linero et al., 2013). It has been shown that in cells with expression of eIF4A or eIF4G knocked down by siRNA the virus is unable to complete the lifecycle (Linero et al., 2013), this is not the case for eIF4E though in which knockdown doesn't have an effect on infection, nor does it show any colocalization or does it co-immunoprecipitate with NP in JUNV, TCRV, or PICV. A key difference between these New World mammarenavirus NP's and an Old World NP such as LASV is that the JUNV, TCRV, and PICV NPs were all retained by M^7G cap-conjugated beads which suggests a different interaction between Old and New world NP's and the

EiF4F complex though further research is needed to characterise this interaction difference (Hastie et al., 2011c, Linero et al., 2013)

1.3.3 L Polymerase

The arenavirus L protein (L) is a multi-domain protein crucial in governing viral genome transcription and replication for which the L RNA-dependent RNA polymerase (RdRp) is involved. Whilst the L's RNA-dependent RNA polymerase (RdRp) domain's main function appears to be in transcription and replication, it possesses additional vital residues for proper function of LASV L that exist both within and outside this domain (Hass et al., 2008, Morin et al., 2010, Lehmann et al., 2014).

The L protein is made up of a single polypeptide chain which comprises three domains: the N-terminal endonuclease-containing domain (PA-like region), the central RdRp-containing domain, and the C-terminal cap-binding domain (PB2-like region). The N- and C-terminal domains trans-complement each other after cleavage at specific positions (Brunotte et al., 2011b).

The N-terminal domain endonuclease region is highly structurally homologous across the *Arenaviridae* family as well as other bunyaviruses and also influenza A virus. This region possesses a two-lobed structure in which the second lobe contains the PD(E/D)K endonuclease active site residues involved in coordination of two divalent metal ions (Olschewski et al., 2020). Arenaviruses L differ from other viral polymerase endonucleases in that these other viruses usually use an additional histidine residue within the first lobe to coordinate the first metal ion in the active site whereas arenavirus endonucleases instead have a glutamate or aspartate residue. Interestingly arenaviruses exhibit reduced in vitro endonuclease activity due to the absence of this histidine residue potentially due to the inability of the N-terminal region to cleave RNA as shown in LASV L (Reguera et al., 2016, Vogel et al., 2019)

The C-terminal domain, associated with cap-snatching and replication, remains partially unresolved which suggests flexibility in its structure, but is suggested to contain the cap-binding domain (Peng et al., 2020, Lehmann et al., 2014). It is thought to contain the cap-binding domain due to mutational studies of LASV NP in which conserved region mutations resulted in inhibition of transcription but not replication which subsequently suggests a C-terminal role in cap-snatching (Morin et al., 2010, Olschewski et al., 2020).

In the RdRp region, the highly ordered finger-tip motif is stabilized without 5' vRNA activation, a unique feature compared to other viral polymerases despite containing typical motifs expected that are conserved among most viral polymerases (Peng et al., 2020). Arenavirus RdRP active site appears to not need to be switched on with 5’ vRNA activation like other viral polymerases in which the finger-tip motif structure could not be resolved without switching on and indicates a distinctive arenavirus regulatory mechanism (Peng et al., 2020).

Exploring the MACV L complex with 3' vRNA and Z reveals insights into L activity regulation. A positively charged groove facilitates 3' vRNA binding, similar to other viral polymerases. Z interaction with specific RdRp region residues may restrict replication and transcription by obstructing RNA entrance, providing insights into arenavirus L regulatory dynamics (Xu et al., 2021).

1.3.4 Z matrix protein

The Z matrix protein (Z) is a zinc-binding RING protein and is the smallest protein encoded by arenaviruses and is between 90 and 99 amino acids long with an 11kDa molecular weight. Similar to NP, it is an extremely multifunctional protein with its main role as the viral matrix protein forming a matrix layer on the inside of the viral envelope and linking the GPC spikes to the packaged RNP complexes. Z also has roles in recruitment and

packaging of all viral components into newly formed viruses. Z also recruits host cell machinery to promote viral budding from host plasma membranes. Z interacts with a multitude of host cell proteins and has roles in IFN antagonism.

The structure of Z has been solved using both nuclear magnetic resonance (NMR) and X-ray crystallography (Hastie et al., 2016b, Volpon et al., 2010). Z is comprised of three structural domains: the N-terminal domain, C-terminal domain, and a central domain. Both the N- and C-terminal domains are highly flexible. The second glycine residue in the N-terminal domain is myristoylated facilitating anchoring to the membrane and initiation of oligomerisation (Strecker et al., 2006, Loureiro et al., 2011). The N-terminal domain contains conserved basic residues with involvement in assisting membrane association via acidic phospholipid interactions (Fehling et al., 2012).

The N-terminal myristoylation shows conservation between Old and New world arenaviruses and mutation studies have shown it is vital for the Z-GP(SSP) interaction as they have affected the intracellular distribution and structure of Z as well as its ability to interact with GP (Capul et al., 2007, Strecker et al., 2006, Fehling et al., 2012).

The central domain of Z contains a RING domain which is 60 amino acids in length and has features typical for RING domains such as conserved $Cys_3HisCys_4$ motif responsible for coordinating two zinc cations at two different coordination sites: site I and II. This RING domain is crucial for self-assembly of Z and expression of just the RING domain is enough to drive formation of spherical cytoplasmic structures(Kentsis et al., 2002b, Kentsis et al., 2002a). The Z RING domain shows similarity to the globular type zinc-binding domains which interact with proteins whereas classical zinc-finger domains interact with nucleic acids (Fehling et al., 2012).

The C-terminal domain contains many late-domain motifs highly conserved between Old and New World arenaviruses. These late-domain

motifs mediate interactions between viral proteins and host proteins for endosomal sorting complexes required for transport (ESCRT) (Fehling et al., 2012).

An important role of Z in the arenavirus lifecycle is in stopping viral replication and transcription to promote formation of viral assembly complexes and kick-start viral budding from the host cell membrane and mini-genome assays have shown inhibitory effects of Z on replication and transcription (López et al., 2001, Cornu and de la Torre, 2001, Cornu and de la Torre, 2002). In TCRV it was shown that Z can directly bind and interact with L and there are to Z-binding sites present on arenavirus L at the N-terminus and one in the RdRp domain (Wilda et al., 2008a, Jácamo et al., 2003). It is thought that the residues required for Z to bind to L are the same as those necessary for Z inhibition of viral RNA synthesis and are found within the RING domain (Cornu and de la Torre, 2002, Loureiro et al., 2011). The direct interaction between L and Z inhibits catalytic activity of the L RdRp but does not have any effect on viral RNA binding important for packaging of the viral genome and L recruitment into new virions (Fehling et al., 2012, Kranzusch and Whelan, 2011a).

Z also inhibits RNA synthesis by L through interaction with NP via the RING domain and this has been shown through RING domain mutation studies which completely inhibit NP incorporation into Z-induced Virus like particles (VLPs) (Casabona et al., 2009, Fehling et al., 2012).

Therefore, a key role of Z is in facilitating interactions between NP and GPC which usually do not directly interact, thereby requiring Z to act as a bridge between them as shown by colocalization studies. Cells that only express GP and NP do not display colocalization and there in no incorporation of NP into the GP-directed VLPs. However, in cells that also express Z, there is colocalization of NP and GP and the VLPs now contain NP (Schlie et al., 2010).

Further Z interactions with NP are necessary for host IFN response antagonism which has been demonstrated in New World arenaviruses and is mediated by and interaction between Z and RIG-I (Fan et al., 2010). This interaction disrupts the normal interactions between MAVS and RIG-I necessary for cell signalling cascade which causes IFN production. In Old World arenaviruses such as LASV and LCMV there does not appear to be any interaction between Z and RIG-I, so it is likely there is np Z mediated IFN antagonism.

1.4 Introduction to the Arenavirus lifecycle

Mammarenaviruses all follow a very similar lifecycle (Figure 1.9) in which the first step is binding of GPC to the appropriate cellular receptor located on the cell plasma membrane. The virus will then enter the cells via receptor mediated endocytosis. The endosome undergoes acidification causing GPC mediated viral and host membrane fusion and release of RNPs into host cell cytoplasm. Transcription and replication of the vRNA then occurs (this creates cRNA which also undergoes transcription and replication), within the host cell cytoplasm directed by the arenavirus L. Viral protein mRNA is translated by host ribosomes in the cytoplasm, although GPC is instead translated into the endoplasmic reticulum (ER) where it matures through the Golgi apparatus (GA) towards the plasma membrane. All of the viral proteins assemble at the host plasma membrane and host complexes are recruited to facilitate viral budding from the membrane and the formation of a new arenavirus virion.

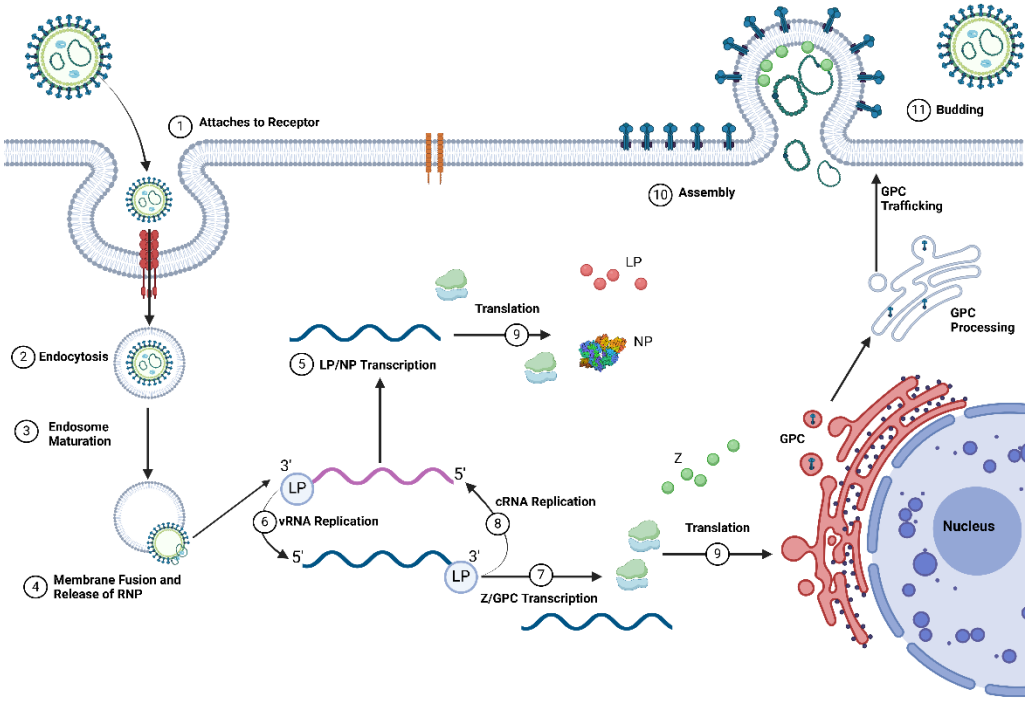


Figure 1.9 Mammarenavirus lifecycle. The mammarenavirus binds to its host cell receptor (1) and enters via an endocytic pathway (2). Acidification within the internalized endosome (3) triggers fusion of the host and viral membranes, releasing the ribonucleoproteins (RNPs) (4). The viral polymerase (L) then transcribes (5) the nucleocapsid protein (NP) and L polymerase (L) genes and replicates (6) the vRNA to produce cRNA (6). The cRNA is then bound by L, allowing transcription (7) of the glycoprotein (GPC) and Z matrix protein (Z) genes and replication of cRNA to generate more vRNA copies (8). The viral mRNAs are translated (9), and NP, L, and genome segments assemble to form ribonucleoprotein complexes. GPC is processed in the rough endoplasmic reticulum and Golgi apparatus. The Z matrix protein (Z) gathers RNPs and GPCs at the plasma membrane for assembly (10) and budding (11). This figure was created using BioRender.

1.4.1 Entry

It is thought that alveolar macrophages are one of the first cell types targeted due to the known routes of infection and that the virus is present in macrophages during the early stages of infection (González et al., 1980). These infected macrophages then move through the lymphatic system which results in migration of the virus and allows it to spread to other tissues (Buchmeier, 2007). This NP-Z interaction plays an important role in recruitment of genomic RNP complexes into new virions.

OW and NW viruses use different cellular receptors as well as different entry pathways into their target cells (Figure 1.10) For OW viruses such as LCMV and LASV as well as Clade C NW viruses, entry occurs when GP1

attaches to the appropriate cell receptor which is α-dystroglycan (α-DG), a widely found extracellular matrix protein found in the basement membrane (Spiropoulou et al., 2002, Shimojima et al., 2012). In contrast, LUJV of the OW complex has been found to utilise neurophilin (NRP)-2 receptor to mediate its entry into cells (Raaben et al., 2017). In LCMV and LASV entry there have been recent discoveries that implicate the TAM family, C-type lectins and Axl as candidate receptors (Shimojima et al., 2012, Fedeli et al., 2018). These recent findings suggest that there is a possibility of multiple entry methods and potential receptors for OW arenaviruses. Recent research into LCMV has uncovered that the viral envelope fuses with late endosomal membranes in a process requiring both low pH and interaction of the LCMV GP spike and secondary receptor CD164. Fulfilment of these two requirements causes uncoating of LCMV and for RNA genome-associated NP to separate from the Z protein matrix layer causing release of the viral genome into the cytosol. This process of arenaviral uncoating has been shown to be regulated by cellular endosomal potassium ion flux and inhibition of this process by blocking potassium ion channels resulted in arenavirus virions becoming trapped within Rab7 and CD164-positive endosomes (Shaw et al., 2024).

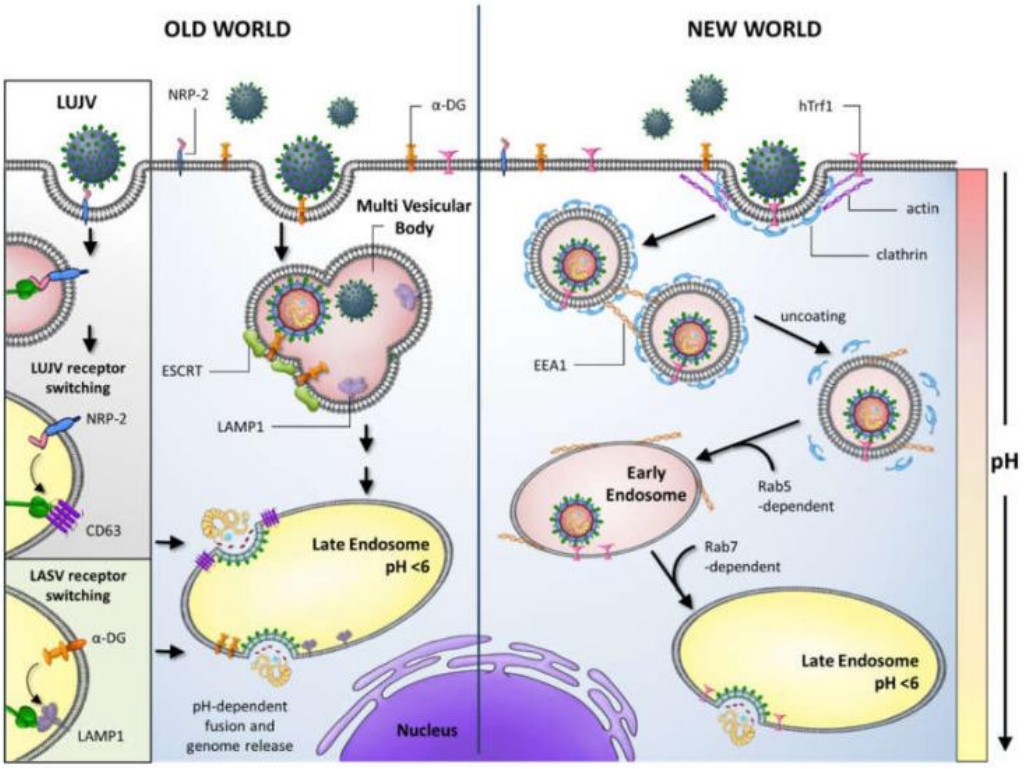


Figure 1.10 Entry pathways of Old World (OW) and New World (NW) arenaviruses. OW arenaviruses, except for Lujo mammarenavirus (LUJV), which binds NRP-2, attach to α-dystroglycan (α-DG). OW viruses enter cells through a clathrin-independent pathway involving multivesicular body formation and the endosomal sorting complex required for transport (ESCRT) pathway. Highly pathogenic NW arenaviruses use human transferrin receptor 1 (hTrf1), or species-specific counterparts, and enter via clathrin-mediated pathways. NW virus particles are trafficked to EEA1-positive endosomes and subsequently to late endosomal compartments through Rab5 and Rab7 dependence. In the late endosome, low pH triggers conformational changes in the arenavirus glycoprotein, leading to fusion and release of the viral genome into the cytoplasm. Lassa mammarenavirus (LASV) requires a receptor switch to lysosome-associated membrane protein 1 (LAMP1) in the late endosome, while LUJV similarly switches to CD63 to facilitate fusion. Figure obtained from (Stott et al., 2020).

The concept of LASV using alternative viral receptors to α-DG was originally suggested due to evidence that cell-types lacking functional α-DG, such as hepatocytes, were extremely susceptible to LASV infection. Mice lacking the LARGE gene crucial for synthesis of functional α-DG molecules (Kunz et al., 2005), are able to sustain replication of LASV at levels equivalent to that in wild-type mice (Imperiali et al., 2008, Barresi and Campbell, 2006). It has also been shown recently that LASV can utilise Phosphatidylserine receptors for cellular entry, cDNA library screens identified the Tyro3/Axl/Mer (TAM) receptor tyrosine kinase members Axl and Tyro3/Dtk as potential candidates as alternative LASV receptors (Shimojima et al., 2012). Members of the TAM family are tyrosine-kinases with roles in exposing

tandem immunoglobulin-related domains to the extracellular matrix. Once exposed, these domains interact with host protein S (ProS) and growth arrest-specific gene 6 (GaS6), which are both serum proteins that bind negatively charged phospholipid phosphatidylserine (PtdSer). PtdSer translocates from the inner to external leaflet of the plasma membrane in apoptotic cells and then acts as a signal for macrophages and dendritic cells (Lemke and Rothlin, 2008). TAM receptors are thought to utilise a mechanism called apoptotic mimicry during their role in virus entry. Apoptotic mimicry involves recognition of PtdSer that is exposed on the virion surface which has been incorporated from the cellular lipid bilayer during viral budding as a signal for uptake of virus (Amara and Mercer, 2015, Shimojima et al., 2006). Studies using a HIV-based lentiviral vector that had been pseudotyped with LASV GP showed the ability of Axl in facilitating entry to cells that either lacked fully functional α-DG (Shimojima et al., 2012).

In Clade B of the NW viruses, GP1 attaches to the transferrin receptor-1 to facilitate entry and the TIM protein family has also been shown to enhance NW infectivity (Helguera et al., 2012, Jemielity et al., 2013). The receptor used by Clade A NW viruses has not yet been discovered (Shi et al., 2018a)).

After GP1 has attached to the cellular receptor, the virion is then endocytosed. OW viruses enter through a currently undiscovered clathrin independent pathway whereas NW viruses are endocytosed through a clathrin-dependent pathway (Pasqual et al., 2011, Rojek et al., 2008). Because of this method of entry into the cell it suggests the involvement of Endosomal Sorting Complex Required for Transport (ESCRT) proteins for successful arenavirus entry (Pasqual et al., 2011). An unusual macropinocytosis pathway has been shown to be a potential route of entry for OW viruses too as demonstrated in LASV (Oppliger et al., 2016). Once the virus has undergone endocytosis and subsequent acidification of the endosome, conformational changes in the GP2 subunit (a class I fusion protein) act to facilitate membrane fusion and release of the viral genome

and replication complexes into the target cell (Eschli et al., 2006). In LASV cell entry, acidification of the endosome causes alterations in the GP1 receptor binding domain resulting in GP1 binding to Lysosomal-associated membrane protein-1 (LAMP1) which then assists the fusion activity of GP2 (Cohen-Dvashi et al., 2016). A similar strategy has been shown to be employed by LUJV but utilising CD63 rather than LAMP1 for fusion and endosomal escape (Raaben et al., 2017).

1.4.2 Replication and Translation

Upon release of the RNP into the target cell cytoplasm following viral envelope fusion, viral L and NP proteins begin replication of the viral genome. The association of NP, L, and the viral genome within the cell is considered the minimum requirement for replication initiation (Hass et al., 2004). The L protein contains four domains involved in viral replication (Brunotte et al., 2011b). Domain, one contains the endonuclease domain thought to facilitate 5' mRNA cap acquisition that is crucial to transcription as this cap enhances mRNA translation and stability. Domain three contains conserved sequences that have been identified as elements of RNA-dependent RNA-polymerase (Vieth et al., 2004). Currently the roles of domains two and four are unknown. Once NP and L protein have acquired 5' caps from host mRNA replication initiates and these caps prime initiation of early gene transcription. Arenavirus RNA is not able to be directly translated in its genomic form due to being of negative sense and so nascent RNA-dependant RNA transcription produced in the 5'-3' direction is necessary. The open reading frames (ORFs) of NP and L protein are found in this first transcribed region resulting in production as early products (Meyer et al., 2002). Termination of transcription is performed using intergenic hairpin structures located in the arenavirus L and S segment IGRs (Pinschewer et al., 2005). Newly translated NP binds viral RNA and facilitates more transcription, replication, and virion assembly.

Antigenomic RNA produced when the polymerase reads through the IGR hairpins and creates a complimentary RNA genome that allows early gene products to transition to late products (Meyer et al., 2002). This complimentary RNA genome is also used as a template for transcribing Z protein and GPC mRNAs. Because of this requirement of antigenomic RNA and later genomic RNA being produced, Z protein and GPC are considered the late gene products in the replication cycle.

Mutation studies of the central domain of the LASV L protein which acts as a putative RNA-dependent RNA polymerase have been used to map functionally relevant residues (Lelke et al., 2010). It was shown that seven residues (Asp-89, Glu-102, Asp-119, Lys-122, Asp-129, Glu-180, Arg-185) were selectively important for mRNA synthesis, with Asp-89, Glu-102, and Asp-129 having pronounced phenotypes necessary for transcription but not for replication (Lelke et al., 2010).

An arenavirus-specific unique pendant insertion located between the finger domain and the α-ribbon helices has been identified and linked to arenavirus polymerase activity regulation (Peng et al., 2020). By deleting this domain, the polymerase activity of LASV and MACV L proteins were significantly decreased as we transcription efficiency but the affinity for 3'-vRNA or 3'-cRNA binding was not affected.

Arenavirus polymerase has the unique characteristic among segmented negative-strand RNA virus polymerases in that it is only active in the presence of 3'vRNA. This is generally a property seen only in non-segmented RNA virus polymerases (Morin et al., 2013). It has been suggested that 5'-vRNA might act as a switch to transform the L protein from acting as a transcriptase to a replicase during the earlier stages of viral replication (Peng et al., 2020). This has also been seen to occur for Influenza virus polymerases in infected cells where small RNAs created by the Influenza A virus infection regulate this switching process (Perez et al., 2010).

Direct interaction between arenavirus L and Z protein has been shown experimentally, with subsequent allosteric regulation by the Z protein (Liu et al., 2023). Z was found to bind at the bottom of the interface between the core lobe of the PA-like domain and palm subdomain of the RdRp core. Insertion of the Z matrix protein into the L protein to form the L-Z complex causes L protein to undergo conformational changes and partially impedes the ability of nascent RNA products to exit which, in turn, inhibits RNA synthesis of L protein (Liu et al., 2023)

1.4.3 Virion Assembly and Budding:

The model for arenavirus assembly and budding has proposed that during the early stages of arenavirus replication, whilst Z is at low concentration, ongoing RNA synthesis is permitted by the Z protein. Once Z concentration has increased to high levels, binding to RNP occurs and causes inhibition of viral transcription and replication. The next stage of the arenavirus lifecycle involves matrix Z protein being transported to the cell membrane via interaction with ESCRT machinery. After transportation of Z and vRNP to the budding site, interaction between Z and the plasma membrane occurs through myristoylation which is proposed to shift Z conformation to allow for multimerization and binding to Tumour susceptibility gene 101 (Tsg101). This myristoylation of Z enables interaction between the SSP of GP and Z. During the process by which Z forms the membrane curvature in conjunction with the ESCRT complexes, GP that has been cleaved at either the ER or Golgi apparatus by S1P/SKI-1 is incorporated into the new virion (Urata and Yasuda, 2012)

The ESCRT pathway is conserved in all eukaryotes and consists of six hetero-oligomeric complexes (ESCRT-0, -I, -II, -III, ALIX/AIP1, and VPS4A/B). These complexes are sequentially recruited to membranes and have roles in a multitude of functions such as protein sorting, remodelling of membranes, and membrane fission. HRS within the ESCRT-0 complex recognises and binds ubiquitinated endosomal cargo creating a concentration at the endosomal membranes. This cargo is then bound by

ESCRT-I through the Ubiquitin E2 Variant (UEV) domain of the Tsg-101, and ESCRT-II is recruited to the membrane (Wilda et al., 2008a). ESCRT-I and -II complexes work together to induce formation of buds and the ESCRT-III complex and Vps4 activation then mediates membrane scission from the cytosolic side of the bud (Urata and Yasuda, 2012).

The accumulation of Z at the membrane results in many late-replication functions within the cell. As with most RNA viruses, sole expression of the viral matrix proteins are sufficient for production of virus-like particles (VLPs) (Urata and Yasuda, 2012). Z has been shown to initiate self-budding when other viral proteins and genomic RNA are absent, this results in virus like particle production and is evidence as to why Z protein is thought to be the budding factor for arenaviruses (Ziegler et al., 2016).

Importantly, L and Z proteins interact to inhibit polymerase action (Kranzusch and Whelan, 2011b) which results in the replication complex translocating to the cell membrane prior to virion budding. Structural research of Z and L proteins have identified highly conserved residues in the RdRP region of Tacaribe virus L polymerase (H1189 and D1329) that appear to play vital roles during Z protein interaction (Wilda et al., 2008b). Both of these residues are highly conserved in both OW and NW group arenaviruses and it is thought that Z protein may block binding of the NTP entrance to inhibit RNA synthesis and help initiate assembly of arenavirus virions (Peng et al., 2020).

A number of host factors impact and regulate viral assembly and budding. Ubiquitin and ubiquitin-like proteins (UBLs) have been identified as among these factors with functions involving arenavirus budding. Ubiquitin is a 76 amino acid multifunctional molecule that binds to substrate, usually at lysine (K), to mono- or poly-ubiquitinate the substrate. Mono-ubiquitination of a substrate causes intracellular trafficking whereas poly-ubiquitination results in degradation via the ubiquitin-proteasome which is an important pathway for the maintenance of cell homeostasis. ISG15 is a ubiquitin-like protein induced by IFN that has been implicated in regulating viral assembly

and budding for HIV-1 and EBOV (Okumura et al., 2008, Okumura et al., 2006). ISG-15 has been shown to not only tag HIV-1 and EBOV VP40 to regulate the budding process, but also tag ESCRT-III components (Seo and Leis, 2012). It has been shown that ubiquitination of HIV-1 Gag increases the strength of interaction with Tsg101 ten-fold (Garrus et al., 2001). Whilst the regulatory roles for ubiquitin and ISG15 have been demonstrated in HIV and EBOV, the nature of interaction they have with arenaviral budding is currently unknown.

Tetherin is another ISG with roles in the regulation and inhibition of viral budding. Tetherin has been reported as an inhibitory cellular factor during infection with HIV-1 (Martin-Serrano and Neil, 2011, Neil et al., 2008). Tetherin has also been shown to have antiviral activity in a number of other viruses such as retroviruses, filoviruses, and arenaviruses (Sakuma et al., 2009, Le Tortorec et al., 2011). Tetherin is constitutively expressed in a range of cells including terminally differentiated B cells, bone marrow stromal cells, and plasmacytoid dendritic cells. Treatment with IFN-I or -II has been shown to broadly induce Tetherin expression in various cell types (Neil et al., 2008). Tetherin is a type II transmembrane protein present on the cell surface and in perinuclear compartments with four domains: an N-terminal cytoplasmic domain, a transmembrane domain, an extracellular domain, and a putative C-terminal glycosyl phosphatidylinositol anchor domain. Tetherin is anchored in the cell membrane at both ends and X-ray crystallography and small angle X-ray scattering (SAXS) data suggest that it has an ectodomain forming a parallel coiled-coil homodimer. This structure and anchoring method is suggested to be required for the antiviral activity of Tetherin; by anchoring one end of the molecule on the cell membrane and the other on the viral envelope, progeny virus release is inhibited (Le Tortorec et al., 2011, Swiecki et al., 2011, Arias et al., 2011, Hinz et al., 2010, Yang et al., 2010). It has also been shown through electron microscopy that progeny virions released from cells can become directly tethered to each other via Tetherin (Hammonds et al., 2010).

An initial report on Tetherin activity during arenavirus assembly and budding showed that LASV Z-mediated VLP release was inhibited by Tetherin and that this restriction was not overcome by LASV GP but was able to be overcome by HIV-1 Vpu expression (Sakuma et al., 2009). It has been suggested that tethering of virions by Tetherin dimers is stronger than by Tetherin monomers due to the significantly stronger association with the membrane and that Tetherin inhibits release of a range of enveloped viruses via a similar mechanism. Research has demonstrated that Tetherin does inhibit infections LASV and that LASV does not possess any Tetherin antagonist (Radoshitzky et al., 2010). Thus far it is unclear whether Tetherin has an antiviral effect on arenavirus propagation *in vivo*.

1.5 Thesis Hypothesis and Aims:

The arenavirus nucleoprotein (NP) protein plays a critical role in the replication of the viral genome and evasion of the host immune response, and is likely to interact with multiple cellular proteins to promote virus life cycle steps. Understanding these interactions could lead to increased knowledge about the pathogenesis of arenaviruses and shed light on potential targets for therapeutic development. To unravel the arenavirus interactome and the differential mechanisms between human pathogenic and non-pathogenic OW arenavirus NP, this thesis aims to:

1) Develop and optimise proteomic screens for the identification of novel interacting partners of LASV, LUVJ, and MOPV NP. Optimising and expanding these methods will reveal novel co-factors that may define pathogenicity differences.
2) Validate identified interacting partners of NP as potential host restriction factors against arenaviral infection. The nature of interaction between identified host restriction factors and OW NP will be explored using techniques such as immunofluorescence, co-immunoprecipitation, viral infection assays in cells over-expressing host restriction factors, or with these factors knocked out by CRISPR/Cas9.
3) Linked to this will be the use of both experimental and theoretical biophysical and structural methods employed to elucidate the molecular details of the interactions of NP with identified host factors; exploiting these interactions is key in the application of this research to the design of therapeutics. The use of modern tools such as Alphafold3 in conjunction with experimental techniques such as X-ray crystallography and small angle X-ray scattering (SAXS) of host proteins and NP, both individually and in complex, will reveal detailed structural information will define the molecular details of novel host interactions.

Chapter 2

Materials and Methods

2 Materials and Methods

2.1 General Methods

2.1.1 VECTORS

Arenavirus nucleoproteins for LASV, LUJV, and MOPV (accession numbers AY628203.1, JX017360, DQ328874.1, respectively) were synthesised by GeneArt (Thermofisher) and then were cloned into pCAGGs vector with a C-terminal HA tag using *EcoRI* and *XhoI* enzyme restriction sites. ZAP Long and Short (accession numbers NM_020119.4 and NM_024625.4 respectively) were cloned into PCDNA3.1 with an N-terminal myc tag using EcoRI and NotI. TRIM25 (accession number NM_005082.5) was cloned into PCDNA3.1 with a C-terminal FLAG tag using EcoRI and XhoI restriction enzyme sites. Empty vectors used were pCAGGs, PCDNA3.1 (Both kindly provided by Professor Janet Daly, UoN), and BioID2-MCS-HA (Addgene).

2.1.2 Bacterial cell strains

The Escherichia coli (E. coli) Stbl3 strain from NEB was used for plasmid amplification: This cell line is manufactured by Invitrogen and derived from HB101 E. coli strain. Genotype:

F- mcrB mrr hsdS20(rB-, mB-) recA13 supE44 ara-14galK2 lacY1 proA2 rpsL20(StrR) xyl-5 lambda- mtl-1

2.1.3 Mammalian Cell Lines

Cell lines used in this work include the Human Embryonic Kidney 293T (HEK) cells (kindly provided by Professor Janet Daly), and A549 Lung carcinoma epithelial cells (ATCC), and Vero cells (kindly provided by Professor Kin Chow Chang, UoN).

All mammalian cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM), high glucose, GlutaMax™ Supplement (Gibco) with 10%

heat inactivated FBS and 200µg/ml Gentamicin (Sigma) and incubated at 37°C and 5% CO_2.

HEK293 cells were originally isolated by Alex Van der Eb in the 1970s, The HEK293T cells are a derivative of the original parent cell line which contains the SV40 large T antigen in its genome allowing production of recombinant proteins within plasmid vectors containing SV40 promoter (Graham et al., 1977)

The A549 (ATCC) cell line consists of hypotriploid alveolar basal epithelial cells and was originally developed in 1972 by removing and culturing pulmonary carcinoma tissue from the explanted tumour of a 58-year-old caucasian male (Giard et al., 1973).

A549 cells were also used in the creation of stably expressing CRISPR knockout cell lines for ZAP and TRIM25. These were additionally supplemented with 1 µg/ml puromycin every passage to select for only CRISPR knockout cells following transduction with the Esp.19 plasmid (Horizon Discovery) containing the sgRNA sequence of interest. The endogenous TRIM25 (target sequence: GTCGTGCCTGAATGAGACGT) and ZAP (target sequence: TATGAGCTGAGTTTCCAAGG) knockout synthetic sgRNA sequences were obtained from Horizon Discovery and VLPs were generated using pxPAX2 and pMD2.G plasmids (Addgene). From the generated knockout cells for ZAP and TRIM25, isolation and expansion of single-cell clones was performed to ensure only successfully knocked out cells were present in subsequent cultures. Selection of single-cell clones involved isolation by dilution plating; transduced cells were harvested by trypsinisation and counted to a concentration of 20 cells per 100µl. 200µl of diluted cells was added to the first row of a 96-well plate and from this first row, 100µl of medium containing cells was added to the second row which containing 100µl appropriate medium to create a two-fold dilution. This process was repeated down the rows of the plate resulting in a series of two-fold dilutions and some proportion of the wells containing just a single cell. These cells were given time to grow and expand prior to analysis

by SDS-PAGE and western blot to check that expression of either endogenous ZAP or TRIM25 was completely knocked-out. Cultures from these single-cell clones were used for all subsequent experiments requiring stably expressing knock-out cell lines.

The Vero cell line was established from African green monkey kidneys in 1962 by Yasamura and Kawakita and is useful for the growth of viruses, (Osada et al., 2014). Vero cells were used to grow our stocks of Mopeia mammarenavirus and also to perform plaque assays.

2.1.4 Virus Strains

The UVE/MOPV/UNK/MZ/Mozambique 20410 strain of Mopeia mammarenavirus was obtained from the European Virus Archive goes Global platform (EVAg) and Vero cells grown in DMEM supplemented with 2% FCS were used to generate mycoplasma free virus stocks. Plaque immunoassay technique was used to determine MOPV virus titre; details of which can be found in section 2.2.4.2.

2.1.5 Antibodies

Antibody	Species	Dilution for western blot (v/v)	Label	Company
Anti-LASV NP	Rabbit	1:10000	PA5-117437	Invitrogen
Anti-HSP90	Mouse	1:5000	MA1-10372	Invitrogen
Anti-TRIM25	Rabbit	1:5000	Ab167154	Abcam
Anti-TRIM25	Mouse	1:5000	Ab610570	Abcam
Anti-ZAP	Rabbit	1:5000	Ab105357	Abcam
Anti-ZAP	rabbit	1:5000	Ab154680	Abcam
Anti-FLAG	Mouse	1:2500	F1804-200UG	Sigma
Anti-FLAG	Rabbit	1:5000	14793S	CST
Anti-FLAG	Mouse	1:5000	8146S	CST
Anti-Myc	Mouse	1:5000	M5546-2ML	Sigma
StrepHRP		1:10000	Ab7403	Abcam
GAPDH	Mouse	1:5000	MA5-15738	Invitrogen
Anti-HA	Mouse	1:5000	Ab18181	Abcam
Anti-HA	Rabbit	1:5000	Ab137838	Abcam
HRP	Rabbit	1:5000	7074S	CST
HRP	Mouse	1:5000	7076S	CST
Anti-Mouse 594nm	Donkey	1:500	Ab150108	Abcam
Anti-mouse 488nm	Donkey	1:500	Ab150109	Abcam
Anti-rabbit 594nm	Goat	1:500	Ab150084	Abcam
Anti-rabbit 488nm	Goat	1:500	Ab150081	Abcam
(HRP)-conjugated horse anti-mouse IgC secondary	Horse	1:5000	7076S	CST
(HRP)-conjugated goat anti-rabbit IgC secondary	Goat	1:5000	7074S	CST

Table 2.1 Antibodies used

2.2 General Methods

2.2.1 Continuous Cell Culture

All mammalian cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM), high glucose, GlutaMax™ Supplement (Gibco) with 10% heat inactivated FBS and 200µg/ml Gentamicin (Sigma) and incubated at 37°C and 5% CO_2. Cells were routinely passaged once confluence was reached. Media was washed from cells using 1x PBS and cells were detached from the flask using Gibco Trypsin-EDTA (0.25%) phenol red and incubating at 37°C for 5 minutes. Cells were counted using a haemocytometer.

2.2.2 Transfection protocols

2.2.2.1 PEI

Polyethylenimine (PEI) was used for DNA transfection into host cells as it condenses DNA into positively charged particles that can then bind to anionic cell surfaces allowing for the PEI:DNA complex to be endocytosed by cells and DNA is subsequently released into the cytoplasm. It is a low cost and efficient transfection reagent that can be stored stably at 4°C. (Boussif et al., 1995, Sonawane et al., 2003). For PEI transfections a volume ration of 3:1 PEI:DNA was prepared in serum-free DMEM, e.g. Transfecting a single well of a 24 well plate: 3µg PEI-Max was diluted in 25µl SF media. Then 1µg of DNA was diluted in 25µl of media. Diluted PEI-Max to was added to diluted DNA and incubated at room temperature for 30 minutes. The PEI-DNA mixture was gently added to target cells and the plate returned to the incubator

2.2.2.2 Lipofectamine 3000

Lipofectamine transfection was performed according to the guidelines provided in the Invitrogen Lipofectamine 3000 transfection reagent kit. Transfection of a single well of a 24 well plate involved diluting 1.5µl Lipofectamine 3000 reagent in 25µl Opti-MEM medium. Then diluting 1µg

DNA in 25µl Opti-MEM medium with 2µl P3000 reagent (2µl/µg DNA). Diluted Lipofectamine 3000 was then added to the diluted DNA and the mixture incubated at room temperature for 5 minutes. The Lipofectamine 3000-DNA-P3000 mixture was added to wells containing target cells and the plate returned to the incubator.

2.2.2.3 Viafect

Viafect transfection was performed according to protocol provided by Promega. Before use the Viafect transfection reagent was allowed to reach room temperature. Viafect reagent was then mixed well by inverting and 45-50µl of serum-free medium added such that final volume after adding DNA was 50µl . 1µg of plasmid DNA was added and mixed with a 3:1 ratio Viafect: DNA (3µl Viafect for every 1µg DNA). The transfection complex was then incubated for 20 minutes at room temperature prior to careful addition to target cells and returning the plate to the incubator.

2.2.2.4 FuGENE 4K

FuGENE 4K transfection was performed according to the protocol provided by Promega. Before use the FuGENE 4K transfection reagent was allowed to reach room temperature. The reagent was mixed well by inverting and diluted in 25ul Opti-MEM using 4µl FuGENE 4K reagent per 1µg DNA. 1µg of DNA was diluted in 25µl Opti-MEM. Diluted FuGENE 4K reagent was added to diluted DNA and mixed well. The transfection complex was incubated for 20 minutes at room temperature and then carefully added to cells and returned to the incubator at 37°C.

2.2.2.5 siRNA Knockdown

Gene knockdown by siRNA was performed according to the DharmaFECT transfection reagent protocol from Horizon Discovery. ON-TARGETplus siRNA for ZAP and TRIM25 targeting the gene ORF from Horizon Discovery was used with a non-targeting ON-TARGETplus siRNA used as a control. In

tube 1, 0.5µl siRNA was diluted in 9.5µl serum-free medium. In tube 2, 0.5µl of DharmaFECT transfection reagent was diluted in 9.5µl of serum free medium. The contents of each tube was gently mixed by careful pipetting up and down and left to incubate for 5 minutes at room temperature. The contents of tube 1 were added to tube 2 for a total volume of 20µl and gently mixed by pipetting and then left to incubate for 20 minutes at room temperature. After incubation, 80µl of complete medium was added for a total reaction volume of 100µl which was then gently added to target cells and incubated for 48 hours.

2.2.3 Virological techniques for MOPV

2.2.3.1 Mopeia mammarenavirus infection and Interferon treatment assays

A549 cells were infected with MOPV at multiplicity of infection (MOI) 0.01 and incubated at 37°C. For interferon treatment, cells were treated with either 250, 500, or 1000 U/ml^{-1} IFN-I (universal type 1 IFN, PBL Interferon Source) for 16h before infection.

2.2.3.2 Plaque immunoassay

This technique was used to determine MOPV virus titre. Vero cells were infected with serial dilutions of MOPV and incubated in MEM with 2.5% low viscosity carboxymethylcellulose (Merck) for 4 days. The mouse monoclonal anti-Arenavirus (Old-World) rGPC, clone KL-AV-1B3 antibody was used to visualise plaques (BEI resources, 1:500 dilution), the secondary antibody was anti-mouse Alkaline Phosphatase antibody (Merck, 1:750 dilution) and addition of BCIP/NBT substrate to develop the plaques. Virus titres were calculated as plaque forming units per mL (PFU/mL)

2.2.4 Protein Analysis Techniques

2.2.4.1 Mammalian Cell Lysis

Mammalian cell lysis prior to protein expression analysis by western blot was performed using 2x Laemmli buffer. BIO-RAD premixed 4x Laemmli

protein sample buffer was used according to manufacturer guidelines. 100ul beta-mercapthoethanol was added 900ul BIO-RAD 4x Laemmli which was then diluted 1:2 using water.

2.2.4.2 SDS-PAGE and Western Blotting/Immunoblotting

Cell lysates (lysed in 2x reducing Laemmli buffer (Bio-Rad at 100°C for 10 min)) were resolved using 12% Precast gels (Bio-Rad), transferred onto nitrocellulose membranes (Bio-Rad), blocked in 5% milk in PBS with 0.1% Tween20 (PBS-T) prior to incubation with appropriate primary antibodies overnight at 4°C: mouse anti-FLAG (Sigma, F1804, 1:2000), rabbit anti-FLAG (sigma, F7425, 1:2000), mouse anti-HA (Abcam ab18181, 1:5000), mouse anti-HSP90 (Invitrogen, MA1-10372, 1:5,000), mouse anti-MYC (Sigma, M5546, 1:2000), rabbit anti-ZAP (Abcam ab154680, 1:5000), rabbit anti-TRIM25 (Abcam ab167154, 1:5000). Membranes were then washed with PBS-T and further probed with either horseradish peroxidase (HRP)-conjugated horse anti-mouse IgG (CST, 7076S, 1:5000) or goat anti-rabbit IgG (CST, 7074S, 1:5000) secondary antibodies in 5% milk solution in PBS-T for one hour at room temperature prior washing with PBS-T and subsequent detection using Supersignal™ West Pico PLUS chemiluminescent substrate and scanning with a iBright imaging system from Thermo.

2.3 Chapter 3 Specific Methods

2.3.1 BioID2 vector and cloning NP

Figure 2.1 shows the plasmid map of the BioID2 empty backbone vector (obtained from Addgene) into which NP genes were cloned to the N-terminus of in-vector BioID2 and HA tags. These constructs were then used in proximity-dependent biotin identification (BioID2). The multiple cloning site (MCS) is located upstream of the BioID2 tag and HA tag. Primers were designed to incorporate EcoRI and BamHI restriction enzyme sites into the LASV, LUJV, and MOPV NP sequences and these were then digested using these restriction enzymes. The BioID2 empty backbone was digested using the same enzymes as well as recombinant Shrimp Alkaline Phosphatase (rSAP) which is used to non-specifically catalyse the dephosphorylation of

5’ and 3’ ends of DNA phosphomonoesters. Once both the plasmid vector and the LASV, LUJV, and MOPV NP inserts were digested and the enzymes heat inactivated, they were run on a 1% agarose gel using gel electrophoresis. The DNA band at the appropriate size for the insert was excised from the gel and purified using the QIAquick Gel Extraction kit. The ligation protocol involved a 1:3 vector to insert ratio and use of T4 DNA ligase and 10x T4 ligase buffer and was performed overnight at 16°C.

Figure 2.1 Empty BioID2 vector. PCDNA3.1 backbone including a multiple cloning site (MCS) and HA tag (Obtained from Addgene)

Ligation products were transformed into DH5-Alpha E. coli competent cells and incubated overnight on ampicillin plates at 37C. Single colonies were picked and grown overnight in 2YT broth with ampicillin at a concentration of 50ug/ml using a shaking incubator at 37C. Overnight growths were then miniprepped prior to checking expression of constructs by transfecting into HEK293T cells using either Viafect or Lipofectamine 3000 transfection reagent, or into A549 cells using either Viafect or Lipofectamine 3000

transfection reagent. HEK293T cells transfected using PEI reagent had the best expression of LASV, LUJV, and MOPV NP within the BioID2-HA vector and as such this was chosen as the method for the BioID2 PDB-MS protocol.

2.3.2 Transfection of mammalian cells and BioID2 affinity purification

An overview of the experimental approach to identify human cellular protein partners of LASV, LUJV, and MOPV NP using BioID2 proximity labelling is shown in Figure 2.2. HEK293T and A549 cells were cultured in Dulbecco's Modified Eagle medium +/- GlutaMax and supplemented with 10% Fetal Bovine serum and gentamicin antibiotic at 50µg/ml. These cells were plated onto 6-well plates and transfected with 2ug of empty BiolD2-HA vector (MSC) or BioID constructs for expression of LASV, MOPV, and LUJV NP using either polyethylenimine (PEI, Polysciences), Lipofectamine 3000 (Invitrogen), or Viafect (Promega) transfection reagents to test each which transfection reagent would work best. After analysis, only PEI transfection in HEK293T cells was successful enough to be used for the full BioID protocol. LASV, LUJV, and MOPV NPs are fused to the N-terminus of a biotin protein ligase (BioID2) and used in proximity-dependent biotin identification (BioID) and HA-tag, creating an 89kDa protein (62kDa for NP and 27kDa for protein tags). For cells transfected with PEI, media was changed 16 hours post transfection. Cells were expanded into 6x10cm dishes per NP construct for selection with G418 antibiotic. 16 hours prior to performing the BioID pulldown, 100µM free-biotin (Invitrogen) was added to cells to induce biotinylation of potential interacting proteins. Biotin was prepared in warm media and filtered prior to addition to cells.

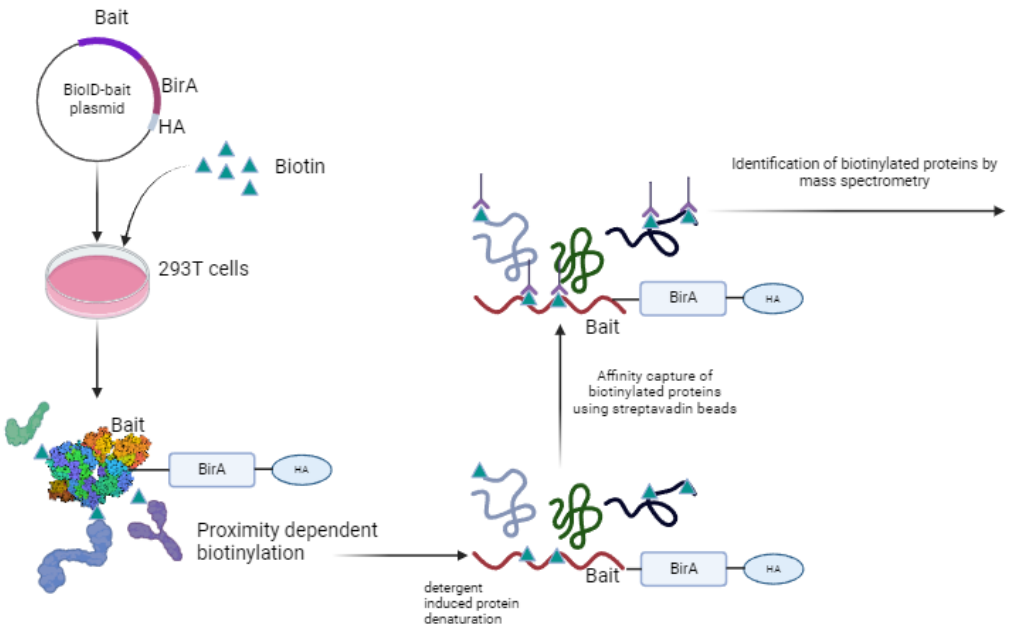


Figure 2.2 An overview of the BioID2 experimental approach. Subsequent example mass spectrometry spectral graph used to identify human cellular protein partners of the arenavirus nucleoprotein in the context of HEK293T cells transiently overexpressing BioID-HA MCS control, LASV NP, LUJV NP, or MOPV NP. Transfected cells were treated with free-biotin 16 hours prior to being lysed and added to streptavidin beads to immunoprecipitate any close proximity proteins to NP.

For the affinity purification, cells were washed 3x in ice-cold PBS and gently lysed on ice in 6ml of 50mM Tris pH 7.4, 500mM NaCl, 0.4% SDS, 0.5% IGEPAL, 1mM DTT and complete protease inhibitor (Roche). Samples were then sonicated using 3 x 20s cycles prior clarification by centrifugation at 13,000rpm for 10 minutes at 4°C. 200µl Streptactin Sepharose High Performance beads (GE Healthcare) were prepared for each construct. Beads spun at 1000g for 2 minutes and 2x washes in ice-cold PBS prior to resuspension in lysis buffer. An aliquot of the supernatant was used for Western blot analysis and the remaining supernatant was added to the beads and incubated on a rolling platform overnight at 4°C. Following incubation, the beads were washed 2x in ice-cold 50mM Tris pH 7.4, 500mM NaCl and 2 x in ice-cold 50mM Tris pH 7.4 to remove non-specific and unbound proteins. Beads were resuspended in 2x Laemmli buffer (Bio-Rad) or for downstream proteomic analysis were resuspended in 50mM ammonium bicarbonate and 1mM biotin. The bead suspension was immediately frozen at -80°C prior to shipment to the Advanced Mass

Spectrometry Facility (AMSF), University of Birmingham for proteomic analysis on an Orbitrap Eclipse Tribrid Mass Spectrometer. Samples of the affinity purification were analysed by Western blot.

2.3.3 Proteomic analysis

2.3.3.1 University of Birmingham AMSF MS data:

Two sets of samples were sent to the Advanced Mass Spectrometry Facility (AMSF) at Birmingham University. Group 1 represents the BioID2 control sample in which empty vector MCS-BioID2-HA was transfected into cells. Group 2 was the sample set in which LASV, LUJV, and MOPV NP-HA were transfected into cells for pull-down. Only the MOPV NP-HA sample from Group 2 was successfully pulled down and sent for mass spectrometry analysis. Both the MCS control (group 1) and MOPV NP (group 2) sample groups for this experiment only had one replicate each.

The following methodology has been obtained from the AMSF who performed the protocols for in-gel digestion and the subsequent LC-MS/MS experiment to generate mass spectrometry data:

2.3.3.2 In-gel digestion

Samples were run using gel electrophoresis and gel bands were excised and then they were destained and digested by a robot system. Briefly, Coomassie blue were destained with acetonitrile followed by 100 mM ammonium bicarbonate. This cycle was repeated if necessary, until gel pieces were fully destained. Destained gel pieces were dried (vacuum centrifugation; 5 min) and rehydrated in 100mM ammonium bicarbonate. Then the gel pieces were reduced with 10mM DTT, at 60°C for 15 min, the liquid was removed and replaced with 50mM iodoacetamide and 100mM ammonium bicarbonate for alkylation. Gel pieces were incubated at room temperature in the dark for 45 min, washed with 100mM ammonium bicarbonate and then dried for 5 min. Trypsin gold (Promega, WI, USA) solution was added in 100:1 ratio, shaken at room temperature for 20 min, before diluted with 100mM ammonium bicarbonate. Hydrolysis was

allowed to occur overnight (~16 h) at 37 °C. The gel pieces were first extracted with the solution of 2% acetonitrile and 0.1% formic acid in water and shaken for 30 minutes at room temperature. Extract the remaining peptides in the gel by using 70% acetonitrile with 0.1% formic acid in water shaken for 30 minutes at room temperature. The supernatant was pooled and dried in an evaporator. The samples were re-suspended in 0.1% formic acid in water for the LC-MS/MS analysis.

2.3.3.3 LC-MS/MS Experiment

UltiMate® 3000 HPLC series (Dionex, Sunnyvale, CA USA) is used for peptide concentration and separation. Samples are trapped on uPrecolumn Cartridge, Acclaim PepMap 100 C18, 5 um, 100A 300um i.d. x 5mm (Dionex, Sunnyvale, CA USA) and separated in Nano Series™ Standard Columns 75 µm i.d. x 15 cm, packed with C18 PepMap100, 3 µm, 100Å (Dionex, Sunnyvale, CA USA). The gradient used is from 3.2% to 44% solvent B (0.1% formic acid in acetonitrile) for 30 min. Peptides were eluted directly (~ 350 nL min^{-1}) via a Triversa Nanomate nanospray source (Advion Biosciences, NY) into a QExactive HF (QEHF) mass spectrometer (Thermo Fisher Scientific). The data-dependent scanning acquisition is controlled by Xcalibur 4.0 software. The mass spectrometer alternated between a full FT-MS scan (m/z 360 – 1600) and subsequent high energy collision dissociation (HCD) MS/MS scans of the 20 most abundant ions. Survey scans are acquired in the QEHF with a resolution of 120 000 at m/z 200 and automatic gain control (AGC) $3x10^6$. Precursor ions were fragmented in HCD MS/MS with resolution set up at 15,000 and a normalized collision energy of 28. ACG target for HCD MS/MS was $1x10^5$. The width of the precursor isolation window was 1.2 m/z and only multiply charged precursor ions were selected for MS/MS. Spectra were acquired for 56 mins with dynamic exclusion time of 20s.

The MS and MS/MS scans are searched against Uniprot database using Proteome Discoverer 2.2 (ThermoFisher Scientific). Variable modifications

are deamidation (N and Q), oxidation (M) and phosphorylation (S, T and Y). The precursor mass tolerance is 10 ppm, and the MS/MS mass tolerance was 0.02Da. Two missed cleavage is allowed and are accepted as a real hit protein with at least two high confidence peptides.

The mass spectrometry performed by the Birmingham was non-quantitative and can provide information on whether a protein was identified in the sample or not and whilst this does not provide direct information on details such as protein abundance changes, the number of peptide-spectrum matches (PSMs) is provided, from which we can infer protein abundance. PSM is a numerical value indicating the likelihood that a peptides fragmentation is recorded in an experimental mass spectrum and is used to identify the peptide sequences for a protein, the number of PSMs can be used as a measure of a protein's abundance. PSMs can also be used to measure the confidence of protein identification and provides important information regarding statistical significance. This data was used to perform analysis of protein function relationships and interactions identified using STRING (Szklarczyk et al., 2021), to generate a report of enriched molecular function pathways and a protein network with interaction connections established based only on confirmed previous experimental evidence and then used Cytoscape for further analysis and interaction network image generation (Shannon et al., 2003). (https://string-db.org/cgi/input?sessionId=byyjMfXT9CVQ&input_page_show_search=on)

2.3.3.4 Mass-Spectrometry analysis performed by Centre for Proteome Research at the University of Liverpool

Issues with the instrumentation at University of Birmingham, coupled with time pressures led us to explore label-free quantification offered by the Centre of Proteome Research, University of Liverpool. Label free quantification allows for a more detailed analysis of protein abundance and abundance changes between samples.

Four sample groups were included in this experiment, each with three replicates. Initially four replicates were set up but only three (replicates 1, 2, and 4) were successful. The sample groups were MCS control (Group 1), LASV NP (Group 2), LUJV NP (Group 3), and MOPV NP (Group 4).

The following protocols and methodologies for digestion, LC-MS analysis , and LC-MS method for sample runs were performed by the University of Liverpool Centre for Proteome Research:

Digestion

Proteins bound to streptavidin beads were washed with 200µL of 50mM Tris-HCl (pH 7.5), followed by two washes with 200µL 2M urea in 50mM Tris (pH 7.5) buffer.

The final volume of 2M urea in 50 mM Tris (pH 7.5) buffer was removed and the beads were incubated with 80µL of 2M urea in 50mM Tris-HCL containing 1mM DTT and 0.4µg trypsin at 25°C for 1 h while shaking at 1,000rpm. After 1h, the supernatant was removed and transferred to fresh lo-bind tubes.

The streptavidin beads were washed twice with 60µL of 2M urea in 50mM Tris (pH 7.5) buffer. The washes were then combined with the on-bead digest supernatant from the previous step.

Reduction of the disulfide bonds in the eluent was achieved by adding DTT to a final concentration of 4mM (11.1µL) and incubated at 25°C for 30 min with shaking at 1,000rpm. Alkylation of the eluent was achieved by adding iodoacetamide to a final concentration of 10 mM (8.43µL) and incubated at

25°C for 45 min in the dark while shaking at 1,000rpm. An additional 0.5µg of trypsin was added to the sample and digestion performed overnight at 25°C with shaking at 700 rpm. After overnight digestion, samples were acidified with formic acid (FA) such that the sample contained ~1% (vol/vol) FA and was at pH 3.

2.3.3.5 LC-MS analysis

Base peak intensity (BPI) target for the instrument that analysis was performed on (Q Exactive HF Quadrupole-Orbitrap™ mass spectrometer (Thermo Scientific)) is 1-2 e^9.

To determine injection volumes of each sample (target of BPI of 1-2e^9) survey runs of 15 min gradient (30 min program) were ran and final sample loading volumes were based on the BPI's observed from the ranging runs. A final sample loading volume of 7uL was used for each sample (neat) and ran on a 2hr program.

2.3.3.6 LC-MS Method for Sample Runs

Samples were analysed using an Ultimate 3000 RSLC™ nano-system (Thermo Scientific, Hemel Hempstead) coupled to a Q Exactive HF Quadrupole-Orbitrap™ mass spectrometer (Thermo Scientific).

The sample was loaded onto the trapping column (Thermo Scientific, PepMap100, C18, 300µm X 5mm),using partial loop injection, for seven minutes at a flow rate of 12µL/min with 0.1% (vol/vol) FA.

The sample was resolved on the analytical column (Easy-Spray C18 75µm x 500mm 2µm column) using a gradient of 96.2% A (0.1% formic acid) 3.8% B (79.95% acetonitrile, 19.95% water, 0.1% formic acid) to 50% A 50% B over 30 minutes at a flow rate of 0.3µL/min (1-hour program).

The data-dependent program used for data acquisition consisted of a 60,000-resolution full-scan MS scan in the orbitrap (AGC set to 3e^6 ions with a maximum fill time of 100ms). The 5 most abundant peaks per full scan were selected for HCD MS/MS (30,000 resolution, AGC set to 1e^5 ions with

a maximum fill time of 300ms) with an ion selection window of 2 *m/z* and a normalised collision energy of 30%. Ion selection excluded singularly charged ions, ions with equal to or a greater than +6 charge state and ions from a streptavidin based exclusion list.

To avoid repeated selection of peptides for fragmentation the program used a 60-second dynamic exclusion window.

2.3.3.7 Quantitative Analysis of Liverpool MS data: Fragpipe and Fragpipe-analyst

The raw data files were converted to .mzML file types with PeakPicking using MSConvert and analyzed using FragPipe to obtain protein identifications and their respective label-free quantification values using in-house standard parameters. Statistical analysis was performed based on the combined_protein.tsv file. First, contaminant proteins were filtered out. In addition, proteins that were not identified/quantified consistently in same condition have been removed as well. The MaxLFQ intensity values were converted to log2 scale, samples were grouped by conditions and missing values were imputed using the 'Missing not At Random' (MNAR) method, which uses random draws from a left-shifted Gaussian distribution of 1.8 StDev (standard deviation) apart with a width of 0.3. Protein-wise linear models combined with empirical Bayes statistics were used for the differential expression analyses. The limma package from R Bioconductor was used to generate a list of differentially expressed proteins for each pair-wise comparison. A cutoff of the adjusted p-value of 0.05 (Benjamini-Hochberg method) along with a |log2 fold change| of 1 has been applied to determine significantly regulated proteins in each pairwise comparison. Quick summary of parameters used: Tested pairwise comparisons = CONTROL_vs_LASV, CONTROL_vs_LUJV, CONTROL_vs_MOPV, LASV_vs_LUJV, LASV_vs_MOPV, LUJV_vs_MOPV

• Adjusted p-value cutoff £ 0.05

• Log fold change cutoff ^{3} 1

Results of the Fragpipe analysis output contains proteins groups of which 1164 proteins were reproducibly quantified, and 66 proteins were found to differ significantly between samples.

2.4 Chapter 4 Specific Methods

2.4.1 Molecular Biology

2.4.1.1 Polymerase Chain Reaction (PCR)

PCR was used for the amplification of the ORFs and to introduce restriction enzyme sites to allow for ligation with a complementary vector. The oligonucleotide primers used in the reactions are described in table below

Primer Name	Primer Sequence (5' – 3')
LASV NP PCAGGs Forward (EcoRI)	GCGCGAATTCATGAGTGCCTCAAAGGAAATAAAATCCTTTTTG
LASV NP-HA PCAGGs Reverse (XhoI)	GCGCCTCGAGTTAAGCGTAATCTGGAACATCGTATGGGTACAGA ACGACTCTAGGTG
LUJV NP PCAGGs Forward (EcoRI)	GCGCGAATTCATGTCCCAATCAAAAGAAGTGAAATCC
LUJV NP-HA PCAGGs Reverse (XhoI)	GCGCCTCGAGTCAAGCGTAATCTGGAACATCGTATGGGTACAGA CTCAGCTTGGCTG
MOPV NP PCAGGs Forward (EcoRI)	GCGCGAATTCATGTCCAATTCAAAGGAAGTG
MOPV NP-HA PCAGGs Reverse (XhoI)	GCGCCTCGAGCTAAGCGTAATCTGGAACATCGTATGGGTACAGG ACGAGCTTGGGTG
ZAP Long-myc PCDNA3.1 Forward (EcoRI)	GCGCGAATTCATGGAACAAAAACTCATCTCAGAAGAGGATCTGG CGGACCCGGAGGTGTGC
ZAP Long PCDNA3.1 Reverse (NotI)	GCGCGCGGCCGCCTAACTAATCACGCAGGCTTTGTC

ZAP Short-myc PCDNA3.1 Forward (EcoRI)	GCGCGAATTCATGGAACAAAAACTCATCTCAGAAGAGGATCTGG CGGACCCGGAGGTGTGC
ZAP Short PCDNA3.1 Reverse (NotI)	CAGATGAAGAGAGGGCCAGAGTAAGCGGCCGCGCGC
TRIM25 PCDNA3.1 Forward (EcoRI)	GCGCGAATTCATGGCAGAGCTGTGCCCCCT
TRIM25 Flag PCDNA3.1 Reverse (XhoI)	GCGCCTCGAGCTACTTGTCGTCATCGTCTTTGTAGTCCTTGGGGG AGCAGATGGAGA

Table 2.2 PCR primers used

PCR was performed using a volume of 50µL containing; 50-100ng template DNA, 10mM dNTPs, 1µM forward primer and 1µM reverse primer, 5x Q5 Reaction Buffer, H_2O, GC Enhancer, and 1U Q5 High Fidelity DNA Polymerase (Q5 High Fidelity Kit; NEB).

PCR protocol cycling was performed in a thermocycler, using the following conditions:

98°C for 30 seconds, then 35 cycles of: 98°C for 10 seconds, annealing temperature for 30 seconds, and 72°C for 30 seconds/kb, followed by a final extension step of 72°C for 2 minutes. Optimal annealing temperatures for primers was determined using temperature gradients ranging from 60°C to 70°C

2.4.1.2 Restriction Digest

Restriction digests were performed for either diagnostic digests or to create sticky ends for downstream ligation of PCR products into new vectors.

Up to 1µg of DNA was mixed with H_2O, 20 units/mL of each restriction enzyme (equivalent to 1µL), and 10X rCutSmart® Buffer (NEB). Reaction was incubated at 37°C for a minimum of 1 hour and enzymes were then either heat-inactivated or removed using DNA Gel Extraction Kit.

1µL of recombinant shrimp alkaline phosphatase (rSAP) (NEB) was included in restriction digest of vectors to dephosphorylate sticky ends and prevent relegation of linearised plasmid DNA

2.4.1.3 Agarose Gel Electrophoresis

To examine diagnostic restriction digests or be able to purify DNA samples by gel extraction, DNA samples were separated using agarose gel electrophoresis.

DNA samples were mixed with 6x gel loading dye (Gel Loading Purple (6x) B7024A NEB) and loaded into a 1% Agarose gel (1x TAE buffer (40mM Tris-acetate, 1mM EDTA; TAE) and stained with Nancy-red at 1:50,000 dilution. The protein ladder used was Quick-Load 1kb PLUS DNA (N0550S NEB). The gel was run at 100V for 1 hour in 1x TAE buffer and DNA was visualised using either UV blue light trans-illumination or on the Imager (find out exact model of new imager)

2.4.1.4 Ligation

After restriction enzyme digest of both vector and insert and subsequent gel extraction and DNA cleanup, vector and plasmid were mixed in a 1:3 molar ratio . T4 DNA ligase buffer and T4 DNA ligase were added and the reaction incubated at room temperature for a minimum of 10 minutes and maximum of 1 hour prior to heat inactivation at 65°C

2.4.1.5 Transformation

E. Coli STBL3 cells were used for propagation of plasmid DNA. 45µL of Bacterial cells were thawed on ice for 10 minutes prior to mixing with 5µL ligation product and incubated on ice for 30 minutes. The mixture was heat shocked at 42°C for 30-45 seconds and incubated on ice for a further 2 minutes. 150µL antibiotic free 2YT media was added to the cells which were incubated for 1 hour at 37°C with shaking. 100µL of this was pipetted onto and then spread around a 2YT agar plate containing 100µg/mL ampicillin antibiotic and was incubated overnight at 37°C.

2.4.1.6 DNA Amplification

A single bacterial colony was used to inoculate 3-5mL of 2YT media containing 100µg/mL ampicillin and incubated overnight with shaking at 37°C. 1.5mL was used to pellet bacterial cells by centrifugation and plasmid DNA extracted using QIAprep Spin Miniprep Kit (QIAGEN).

50µL of the overnight growth was used to inoculate a 50mL starter culture of 2YT broth containing 100µg/mL ampicillin. Plasmid DNA extraction was performed according to manufacturer's instructions using Midi-prep protocol and kit from Macherey-Nagel. DNA concentration and purity was determined by spectrophotometry using a NanoDrop and was confirmed as correct through Sanger sequencing, performed by Eurofins.

2.4.1.7 RNA extraction and cDNA synthesis

RNA was extracted from cells directly in assay plates using the QIAGEN RNeasy Plus Mini Kit and cDNA was reverse transcribed using the Applied Biosystems™ High-Capacity cDNA Reverse Transcription Kit according to manufacturer's instructions.

Temperature	25°C	37°C	85°C	4°C
Time	10 minutes	120 minutes	5 minutes	Hold

Table 2.3 PCR protocol for cDNA synthesis

2.4.1.8 CRISPR Knockouts – stable cell lines

2.4.1.8.1 VLP production

PEI transfection into 293T cells was used to produce virus like particles (VLPs) of CRISPR plasmids which would then be used to transduce A549 cells. A ratio of plasmids was transfected in to enable proper VLP production: 600ng of VSVG plasmid was co-transfected with 600ng of packaging plasmid and 990ng guide RNA plasmid. After 72 hours these were harvested by filtering media through 45-micron filters and stored at -80°C until use for transduction.

2.4.1.8.2 Transduction

A549s were plated at $6x10^5$ cells/well in a 6 well plate and 500ul of VLPs were added per well. The plate was then spun in a centrifuge at 1500g for 45

minutes then returned to the incubator at 37 °C. 48-72 hours post transduction cell selection with puromycin began (1ul per 10ml media)

2.4.2 TRIM25 mutants: Overlap PCR to Introduce Specific Mutations

Overlap PCR was utilised to generate the TRIM25 mutants used in the research presented in this thesis. These mutants include 7KA TRIM25 mutant, a mutant missing the RING domain (ΔRING), and a mutant missing the SPRY domain (ΔSPRY). Overlap PCR involved initial rounds of PCR to generate overlapping gene segments which are then used as template DNA for another PCR reaction to create a full-length product. Internal primers are used to generate overlapping, complementary 3'-ends on the intermediate segments and introduce the site directed mutagenesis required to create the TRIM25 mutants. Overlapping strands of these intermediate products then hybridise at the 3'-region in a subsequent PCR and extend to generate the full-length mutated product amplified by flanking primers that included restriction sites for inserting the product into the PCDNA3.1 expression vector. The details of primers used are listed in Table 2.4. The PCR protocol followed used the same optimisation parameters outlined in the initial TRIM25 cloning protocol in section 2.4.2.2. Forward and reverse primers for TRIM25 shown in Table 2.2 were also used in this protocol to generate intermediate strands.

Primer Name	Primer Sequence (5’ – 3’)
TRIM25 ΔRING_Fwd EcoRI	GCGCGAATTCGCCACCATGCGCGCCGTCTACCAGGCGCGACCGC
TRIM25 ΔSPRY_Rev XhoI	CGCGCGCTCGAGAATCTTGTCGTCATCGTCTTTGTAGTCTTTAATGTAATACTCCAGGAG
TRIM25_CR_7KA_Fwd 1	CGCAGAGGAAGCGGCATCCGCGGCACCTCCCCCTGTCCCTGCCTTA
TRIM25_CR_7KA_Rev 1	TGCCGCTTCCTCTGCGGAGACCGCCGCCACAGGGCGTGTGGATTTGTG
TRIM25_CR_7KA_Fwd 2	ACACGCCCTGTGGCGGCGGTCTCCGCAGAGGAAGCGGCATCC
TRIM25_CR_7KA_Rev 2	AGGGACAGGGGGAGGTGCCGCGGATGCCGCTTCCTCTGCGGAG

Table 2.4 Primers for overlap PCR to introduce specific mutations into TRIM25

2.4.3 Protein Analysis Techniques

2.4.3.1 Real Time qualitative PCR (RT-qPCR)

Primer Name	Primer Sequence (5’ – 3’)
GAPDH Forward	ACATCGCTCAGACACCATG
GAPDH Reverse	TGTAGTTGAGGTCAATGAAGGG
B-Actin Forward	CACCAACTGGGACGACAT
β-Actin Reverse	ACAGCCTGGATAGCAACG
MOPV L Forward	ATCTCCTCATGCAGCCACAC
MOPV L Reverse	GGACTGTTGGAGAGTTGCGA
MOPV NP Forward	CCCTGGCATGTCAAGACCAT
MOPV NP Reverse	CCCTGTGGAAGTTGCGATCT

Table 2.5 RT-qPCR primers used

For each qPCR reaction, 10ng of cDNA was used with the Applied Biosystems™ PowerUp™ SYBR™ Green Master Mix under the following conditions: 50°C 2 mins, 95 °C 2 mins then 40 cycles of 95 °C 15 secs, 55 °C 15 secs, 72 °C 1 min. Primers used were as follows: GAPDH forward (5'-ACATCGCTCAGACACCATG-3'); GAPDH reverse (5'-TGTAGTTGAGGTCAATGAAGGG-3'); B-Actin Forward (5'-CACCAACTGGGACGACAT-3'); B-Actin Reverse (5'-ACAGCCTGGATAGCAACG-3'); MOPV L forward (5'-ATCTCCTCATGCAGCCACAC-3'); MOPV L reverse (5'-GGACTGTTGGAGAGTTGCGA-3'); MOPV NP forward (5'-CCCTGGCATGTCAAGACCAT-3'); MOPV NP reverse (5'-CCCTGTGGAAGTTGCGATCT-3'); TRIM25 Forward (5'-GAGAAACAGGCAGCAGGATG-3'); TRIM25 Reverse (5'-GCTGTTGACCCTCTTCTCCT-3'); ZAP Long Forward (5'-TTTTATGCGACAAGCCGTGC-3'); ZAP Long Reverse (5'-ACAGCTGTCGAACTGTGGAG-3'); ZAP Short forward (5'-CGCCAGGAGAAAGACTGTGT-3'); ZAP short reverse(5'-TCCCCTCGGACAGGATTGAT-3').Primer specificity was confirmed by melting curve analysis. Relative fold expression of target genes was normalised to the GAPDH and B-Actin housekeeping gene by the $^{\Delta\Delta}$Ct method.

2.4.3.2 Immunoprecipitation (IP) and Co-Immunoprecipitation (Co-IP)

293T cells seeded in a 6-well dish at 4x10^5 cells per well were transfected using PEI using two wells per condition 2µg total DNA was transfected per well; in wells with two constructs co-transfected in, 1µg DNA of each was used. 48 hours post transfection these cells were harvested and lysed in a buffer of 50mM Tris, 100nM NaCl, 10mM MgCl2, 1% IGEPAL detergent, 2mM DTT, and 1 protease inhibitor tablet per 50m buffer (Roche). Buffer was made with H2O and pH 7.5. Cells were sonicated at 4°C 3x30s prior to clearing. Cells were cleared by centrifugation for 5 minutes at 10,000rpm at 4°C . 50µl of lysate was removed and stored in 50µl 2x Laemmli (input). The

remaining lysate was transferred to a new 1.5ml Eppendorf and either mouse anti-HA antibody was added at dilution of 1:150 or mouse anti-FLAG antibody at a dilution of 1:400 added and incubated whilst rolling for 1.5-2hrs at 4°C. Pierce protein G magnetic bead preparation involved equilibration in lysis buffer using 50ul neat beads per sample. A magnetic rack was used to separate the beads from any supernatant to allow for easy removal by pipetting. Bead wash steps involved removing the supernatant after the beads had magnetically stuck to the side of the tube placed in the magnetic rack and addition of 1ml PBS. The magnetic beads in PBS were then briefly vortexed to resuspend the beads and were incubated at 4°C with rolling for 5 minutes. After this 5-minute incubation the tube was placed back in the magnetic rack and beads allowed to stick to the side of the tube due to the magnet and the PBS was removed and replaced with 1ml of lysis buffer. The same steps were followed and 1 additional was in lysis buffer performed before resuspension of the magnetic beads in 2x Laemmli buffer for analysis by SDS-PAGE and Western blot.

2.4.4 Microscopy techniques: Fixing and staining cells

A549 cells transfected using Viafect with MOPV NP-BioID2-HA or NP-PCR3.1-HA were stained for co-localisation with TRIM25-FLAG, ZAPL-MYC, ZAPS-MYC. Cells were seeded onto glass coverslips in a 24-well plate followed by fixation using 4% paraformaldehyde and neutralisation in 10mM Glycine/PBS. Cells were permeabilised in 1% Bovine Serum Albumin, 0.1% Triton-X-100/ PBS before incubation with primary antibodies against the NP-HA-tag, TRIM25-FLAG tag, ZAP-MYC tag, overnight at 4°C. Cells were then washed and treated with Alexafluor 488 and 594 secondary antibodies (ThermoFisher) before washing and mounting the coverslips onto slides with DAPI mounting solution for visualisation of nuclear DNA. Images were taken using a Leica upright microscope

2.5 Chapter 5 Specific Methods

2.5.1 Protein expression and purification

2.5.1.1 Protein expression

Fragments for TRIM25 CC-PRYSPRY domain (residues aa189-630), and the N-terminal core domains of LASV (residues 1-340), LUJV (residues 1-344), and MOPV (residues 1-340) NP were synthesised by TWIST Biosciences. The LASV NP and TRIM25 fragments were subsequently cloned into a pET15b vector (Figure 2.3) and the LUJV and MOPV NP fragments cloned into a pET28b vector (obtained from Addgene) (Figure 2.4). LASV NP and TRIM25 in pET15b were transformed into Rosetta BL21(DE3) competent cells with ampicillin resistance and LUJV and MOPV NP in pET28b (obtained from Addgene) were transformed into NiCo21(DE3) (NEB) competent cells with kanamycin resistance.

BL21 competent E. coli is a widely used T7 expression E. coli strain. The Rosetta host strains are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in E. coli. Rosetta BL21 cells have the following genotype:

F-ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)

NiCo21(DE3) *E. coli* competent cells from NEB are derived from BL21(DE3) cells and share identical growth characteristics. They are considered a superior alternative to BL21(DE3) for protein expression and result in improved purity of target proteins isolated by IMAC. NiCo21(DE3) cells have the following genotype:

can::CBD fhuA2 [lon] ompT gal (λ DE3) [dcm] arnA::CBD slyD::CBD glmS6Ala ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5

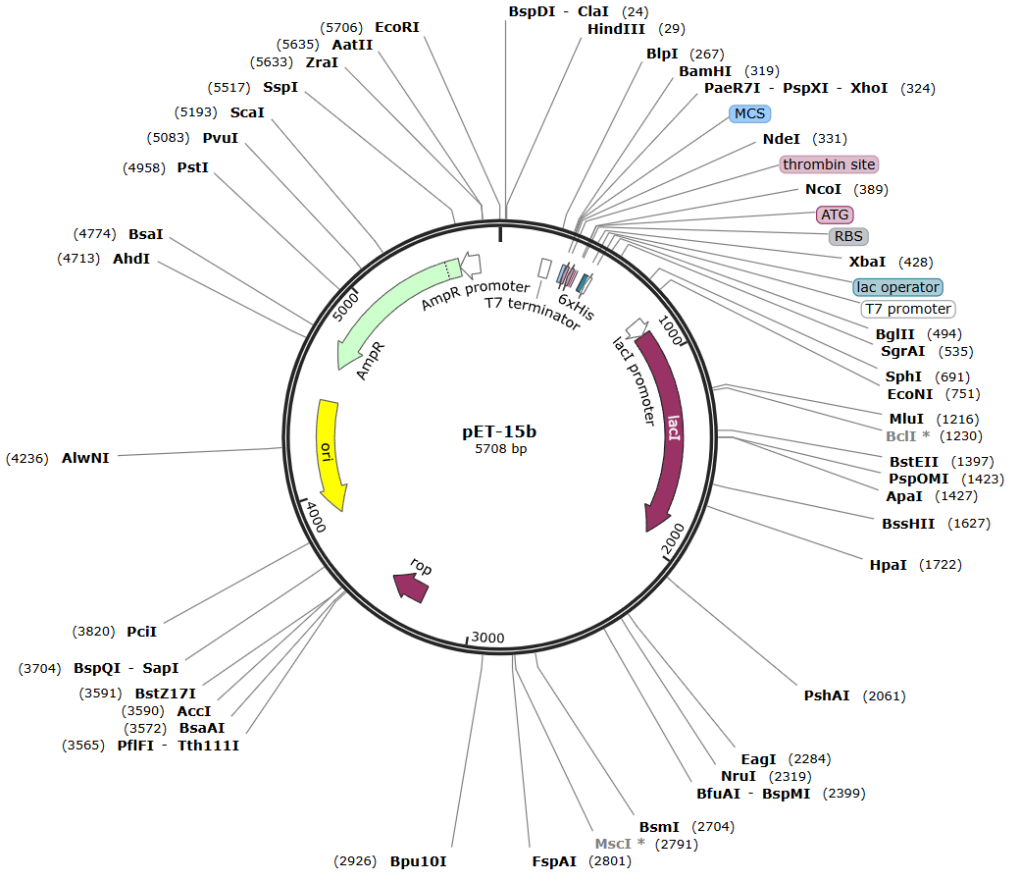


Figure 2.3 pET15B Vector map (obtained from Addgene)

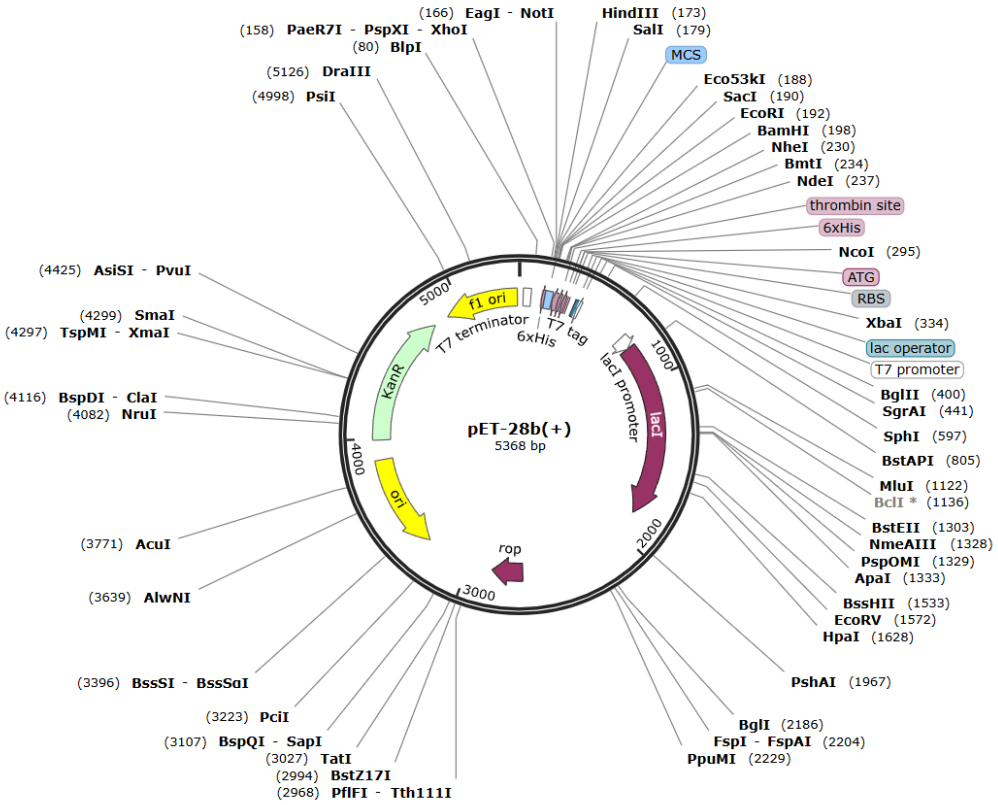


Figure 2.4 pET28b Vector map (obtained from Addgene)

Following transformation into an appropriate vector. A single colony from each construct was resuspended in 10ml liquid 2YT culture with appropriate antibiotic. This was cultured overnight and used to inoculate a larger culture for protein expression and induction with IPTG. For TRIM25 CC-PRYSPRY a 10L culture was prepared in a fermenter. For the LASV, LUJV, and MOPV NP N-terminal proteins, 3L cultures were prepared. These cultures were incubated at 37°C with shaking until the OD_{600} reached an ideal value of 0.6. Once this value was reached by each culture they were induced with IPTG (Isopropyl β-D-1-thiogalactopyranoside) which triggers the release of the lac repressor, inducing gene transcription and protein expression. Samples were taken for each uninduced construct at OD_{600} and resuspended in Laemmli buffer. IPTG at a concentration of 0.5mM was used for induction. Cultures were incubated overnight at 20°C with shaking. Samples were taken from the induced cultures after overnight incubation and OD_{600} checked and normalised to the uninduced sample OD_{600} by dilution in PBS and in Laemmli buffer. Cultures were pelleted by

centrifugation at 4,000rpm for 15 minutes at room temperature and the pellet collected and supernatant discarded. The pellet was then resuspended in appropriate lysis buffer (detailed in the table below) and incubated at 4°C whilst rolling for at least 30 minutes to allow for total lysis. Once lysed, sonication was applied to the cells at 15 second intervals with a 45 second pause a total of 6 times. A sample of this "crude extract" was resuspended into Laemmli buffer. Crude extract was then centrifuged at 20,000rpm for 45 minutes at 4°C. From this, the supernatant was taken for purification by immobilized metal affinity chromatography (IMAC). A sample of both the supernatant was taken into Laemmli buffer as well as a sample of the pellet which was first resuspended in lysis buffer and then taken into Laemmli buffer.

Protein	Lysis buffer
TRIM25 CC-PRYSPRY (189-630aa)	BugBuster Protein Extraction Reagent from Novagen
LASV NP N-terminal core domain (1-340aa)	50mM NaH_2PO_4 pH 8, 500mM NaCl, 10mM Imidazole, 0.5% IGEPAL, 1 protease inhibitor tablet (Roche)
LUJV NP N-terminal core domain (1-344aa)	50mM Tris-HCL pH 7.2, 150mM NaCl, , 20mM Imidazole, 1 protease inhibitor tablet (Roche)
MOPV NP N-terminal core domain (1-340aa)	50mM Tris-HCL pH 7.2, 150mM NaCl, , 20mM Imidazole, 1 protease inhibitor tablet (Roche)

Table 2.6 Lysis buffers used for each protein fragment during protein purification

2.5.1.2 Immobilised Metal Affinity Chromatography (IMAC)

Hispur Cobalt Resin from Thermofisher was used for all IMAC. Before applying protein supernatant to the resin, it was first washed in the column with MilliQ water to remove the 20% ethanol solution it is kept in as a slurry.

Once washed, the resin was equilibrated with the appropriate lysis buffer for the protein outlined in table x above. Once equilibrated, supernatant for each protein was added to the resin and allowed to drip through by gravity. This was collected as "flow-through". A wash buffer specific to each protein was then applied to the column (outlined in table 2.7 below). Each protein was washed 3 times, and each wash step was collected for analysis. Then a final elution buffer was applied to the cobalt resin and the elution collected for dialysis to remove any imidazole. Dialysis for each protein was performed in a 1:50 volume of elution buffer with no imidazole.

Protein Construct	Wash buffer	Elution buffer	Dialysis buffer
TRIM25 CC-PRYSPRY (aa189-630)	50mM HEPES pH 7.5, 300mM NaCl, 20mM Imidazole, 5% glycerol, 500µM TCEP	50mM HEPES pH7.5, 300mM NaCl, 300mM Imidazole, 5% glycerol, 500µM TCEP	50mM HEPES pH7.5, 300mM NaCl, 5% glycerol, 500µM TCEP
LASV NP N-terminal core domain (aa1-340)	50mM NaH_2PO_4 pH 8, 500mM NaCl, 30mM Imidazole	50mM NaH_2PO_4 pH 8, 500mM NaCl, 500mM Imidazole	50mM NaH_2PO_4 pH 8, 500mM NaCl
LUJV NP N-terminal core domain (aa1-344)	50mM Tris-HCL pH 7.2, 150mM NaCl, , 20mM Imidazole	50mM Tris-HCL pH 7.2, 150mM NaCl, , 500mM Imidazole	50mM Tris-HCL pH 7.2, 150mM NaCl
MOPV NP N-terminal core domain (aa1-340)	50mM Tris-HCL pH 7.2, 150mM NaCl, , 20mM Imidazole	50mM Tris-HCL pH 7.2, 150mM NaCl, , 500mM Imidazole	50mM Tris-HCL pH 7.2, 150mM NaCl

Table 2.7 Wash, elution, and dialysis buffers used for each protein fragment during purification

2.5.1.3 Coomassie staining

Coomassie staining was performed using Protein Ark's Quick Coomassie Stain according to manufacturer guidelines. SDS-PAGE protocol was followed, and the gel was washed with MilliQ water before being placed in

Coomassie stain and incubated at room temperature with rocking to allow stain to develop. Once stain had developed, the Coomassie stain was removed and the gel washed with MilliQ water before imaging.

2.5.1.4 Size-Exclusion Chromatography (SEC)

Size exclusion chromatography was performed on a Superdex 200 Increase 10/300 GL column (GE Healthcare) which had been equilibrated with filtered equilibration buffer specific to each protein which was the same as the dialysis buffer used and has been outlined in table 2.7 above. Each eluted and dialysed protein from IMAC was concentrated to 500µl (with a maximum concentration of 5mg/ml) using a 30kDa molecular weight cut-off concentrator (Amicon). This concentrated protein was injected onto the column using an AKTA with a flow rate of 0.3mL/min. Fractions were collected and the absorbance of the samples at 280nm (mAU) was measured. Samples were taken from each fraction corresponding to absorbance peaks observed and analysed by SDS-PAGE and Coomassie to assess which fractions contained target protein. Once the presence of the protein was confirmed, fractions were collated and concentrated as close to 8mg/ml as possible for future use in SEC-SAXS.

2.5.2 SEC-SAXS

SEC-SAXS was performed at the Oxford Diamond Light Source Synchrotron using Beamline B21 for high-throughput small-angle-X-ray scattering. This is a beamline with a bending magnet source in the 3 GeV storage ring at the Diamond Light Source Ltd Synchrotron. It utilises a double multi-layer monochromator and a toroidal focussing optic to deliver 2×10^{12} photons per second to a 34 × 40 µm (FWHM) focal spot at the in-vacuum Eiger 4M (Dectris) detector. A high-performance liquid chromatography system and a liquid-handling robot enable loading of solution samples into a temperature-controlled in-vacuum sample cell with a high level of automation. The beamline has a default scattering vector ranging from 0.0026 to 0.34 $Å^{-1}$ and low instrument background.

Analysis of sample data was initially performed using ScÅTTER and refined data from was them used to trim data in ATSAS CHROMIX and PRIMUS to generate macromolecule envelope data for visualisation in UCSF Chimera 1.18.

Small-angle X-ray scattering (SAXS) is a powerful tool in the analysis of macromolecules in a somewhat native environment. The basic principles of SAXS are relatively simple; dissolved macromolecules are illuminated with X-rays of defined energy at a specific X-ray wavelength, λ, usually about 1Å. The intensities of the scattered X-rays, *I*, are recorded by a detector (Figure 2.5). SAXS samples are isotropic, which means that the molecules are able to tumble in solution independently of each other. It's crucial to include measurements from an identical solution which does not contain the macromolecule of interest. This is usually referred to as a buffer sample and allows for subtractions of buffers to be made to isolate macromolecule information. The profile will contain scattering contributions from the solvent or buffer, the sample holder and also the machine set-up and these will all need to be taken into account for background subtraction. Buffer selection is important for generating good SAXS data as the overall electron density of proteins (which is about 0.43 electrons/$Å^3$) is just slightly higher than that of water (0.33 electrons/$Å^3$). This minor discrepancy in electron densities results in a relatively weak scattering signal which is difficult to generate good results from. This small discrepancy decreased even further with addition of components to the buffer resulting in even weaker scattering. Buffer selection is also important to prevent sample degradation from radiation damage. The addition of components such as TRIS-HCl or glycerol can protect proteins from damage from X-rays which will improve the overall quality of the scattering profile.

Sample purity is another key aspect in the generation of good and reliable data from SAXS. The X-rays will scatter from anything within the sample not just the target macromolecules so any impurities can greatly affect the data profile.

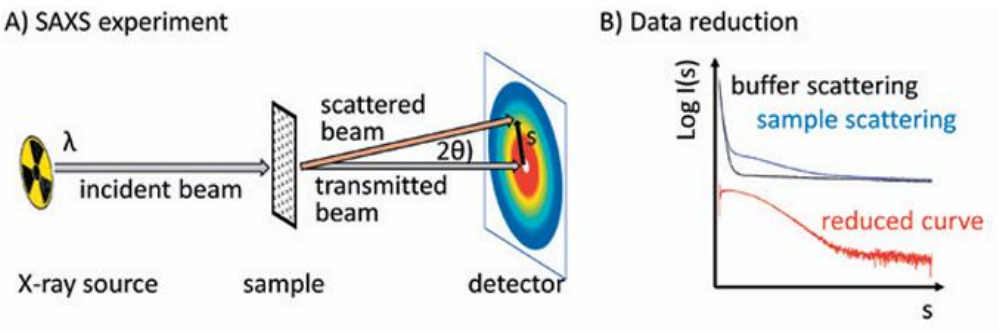


Figure 2.5 SAXS data collection schematic. A) Schematic drawing of the standard experimental set-up. B) Example presentation of data reduction from subtracting buffer from sample scattering. Figure obtained from (Gräwert and Svergun, 2020)

A major recent advancement in biological SAXS beamlines has been in the implementation of SEC-SAXS which involves the use of size-exclusion chromatography columns attached to the HPCL set-up. This allows for different components within the sample to be separated depending on size prior to analysis of each sample by SAXS. Figure 2.6 depicts a schematic of how SEC is implemented and an example of the X-ray scattering profile that may be generated. This improvement can greatly assist in studying macromolecules that may aggregate and form oligomers within the sample.

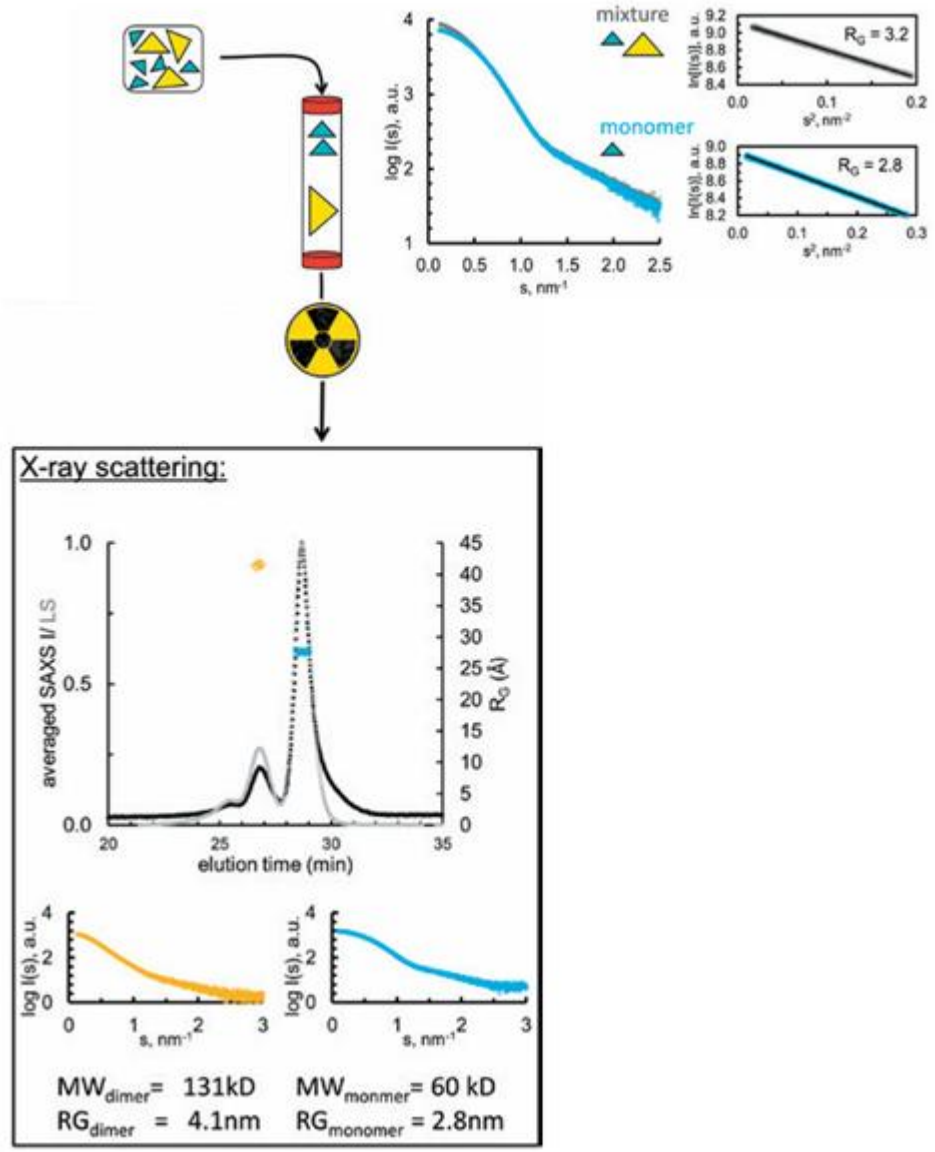


Figure 2.6 A SEC-SAXS/LS experiment. Schematic of data collection method where SEC is directly coupled to SAXS data collection. In parallel this has the advantage of additional characterization of freshly separated components. In the X-rat scattering profile, data of the separation of monomeric (blue) and dimeric BSA (yellow) is indicated. Figure obtained and edited from (Gräwert and Svergun, 2020)

From scattering experiments, overall molecular size and weight as well as the degrees of flexibility and compactness can be determined. As well as this analytical information, the scattering data can provide structural information.

Size of a molecule is determined by the radius of gyration (R_g) in which as R_g^2 and is the average squared distance of the scatterers from the centre of

the object. Graphing scattering intensity I(q) against the scattering vector, q, can give early information about the overall shape of the macromolecule, an example graph and correlating shapes can be seen in Figure 2.7. Radius of gyration is then calculated with a specific equation depending on the shape provided by this information.

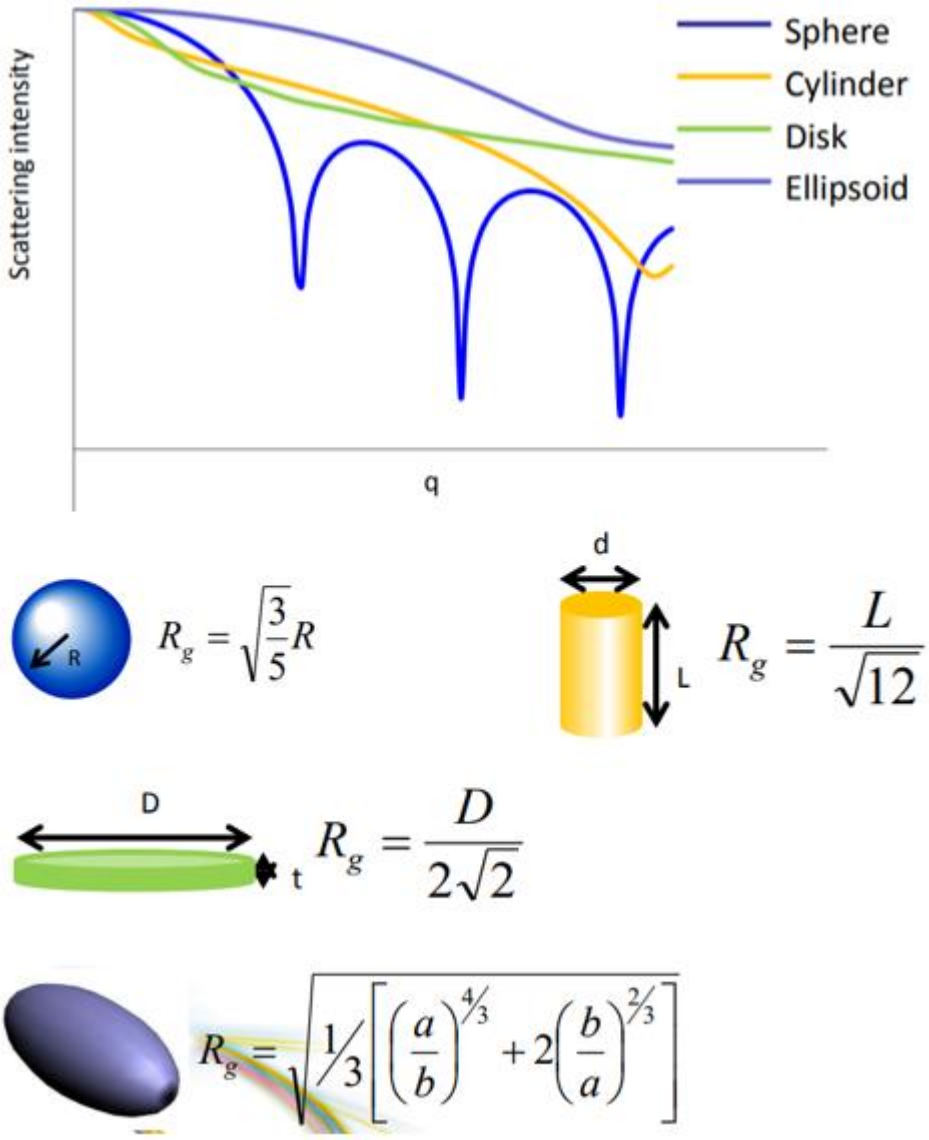


Figure 2.7 Graph of scattering intensity vs scattering vector, q. Graph pattern will depend on the SAXS data obtained experimentally and will inform about the shape of the macromolecule being targeted. The graph shape can then be used to inform the correct equation to use to calculate radius of gyration. These equations and the correlating shapes are shown below the graph. This figure was edited and obtained from the Scattering 101 lecture slides PDF from https://www.diamond.ac.uk/Instruments/Soft-Condensed-Matter/small-angle/Beginners-Guide.html accessed 17/12/2024.

Other information derived from SAXS data necessary to generate a final 3D envelope of the target macromolecule

include the Guinier approximation in which the low-q region of the scattering curve is characteristic for the overall dimension of the particle when intensity of scattering I(q) is plotted against the scattering vector, q. This can then be used to create a Guinier plot in which *I* is plotted against q^2 to give a straight line slope. Deviation from the straight line in the Guinier plot indicates intermolecular interactions or aggregation of the sample. To obtain an ideal scattering curve across the entire q-range, the obtained scattering profile has to be extrapolated to infinite dilutions at low resolutions and then merged with the scattering profile for larger angles. From this, accurate large angle data can be obtained through several means. Either by measuring higher concentrations, using increased exposure times, or decreasing the distance between the sample and the final detector.

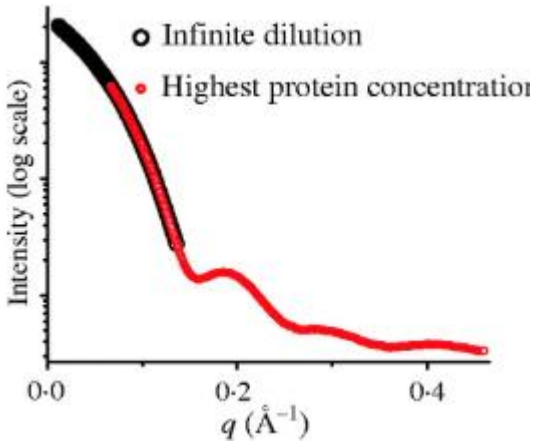


Figure 2.8 Example graph of Intensity vs scattering vector at low resolution depicting the ideal sample scattering. This figure was obtained from (Putnam et al., 2007)

If the SAXS data generated follows the trends expected and fits the graphs without much interference or detrimental noise, then subsequent analysis of the sample data can be performed.

Figure 2.9 shows an example graph of pair-distance (*P(r)*) function plotted against distance (*r*). The graph in the figure shows that the state of protein can be determined from the graph shape. And that globular macromolecules will have a P(r) function with only a single peak. Elongated macromolecules such as multidomain proteins will possess a longer tail at

large r values and can have multiple peaks. The D_{max} is the maximum length in the particle and is the position in which the P(r) function returns to zero at large r values.

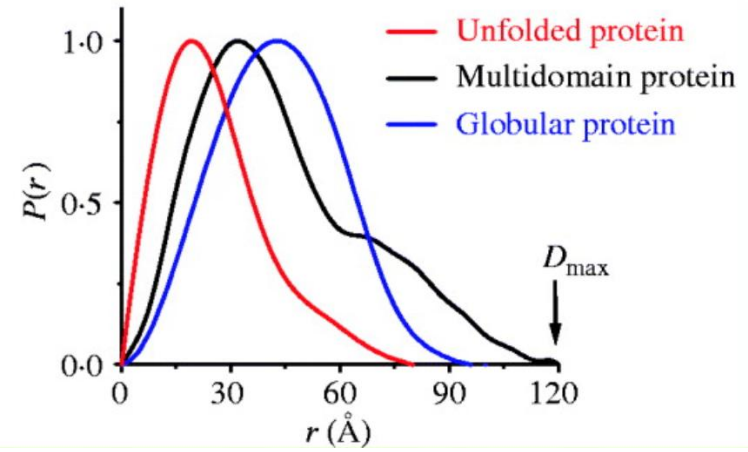


Figure 2.9 Example graph of pair-distance, P(r), plotted against distance, r. the graph shows expected graph shapes for different types of macromolecules. This figure was obtained from (Putnam et al., 2007).

Kratky analysis is another key step and involved plotting $I*q^2$ vs q. This analysis is sensitive to the particle morphology and also to compactness of a protein (Figure 2.10).

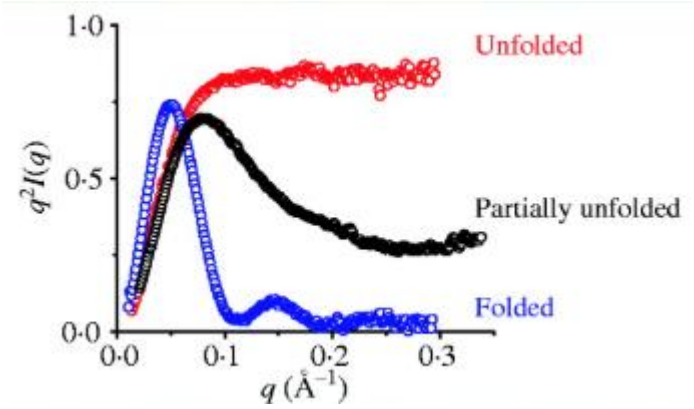


Figure 2.10 Example Kratky plot showing the different graph trends that inform on the final structure and morphology of the macromolecule being investigated. Globular macromolecules follow Porod's law and have bell shaped curves. Extended molecules such as unfolded proteins do not possess this peak and will have a plateau. This figure was obtained from (Putnam et al., 2007)

All of this data is then used to overall determination of the macromolecule structure which can be visualised through either envelope or bead models. Envelope models use spherical harmonics to produce molecular envelops which will fit the scattering data generated from SAXS. Bead models fit the scattering data using beads as scattering centres and these scattering

centres will represent distinct atoms of the residues identified known as “dummy residues”. Program packages such as ATSAS from the EMBL Hamburg Svergun group can be used create these final structure models. These can be incredibly powerful techniques in structural biology especially if combined with known and experimentally identified crystal structures and with the recent developments of predictive tools such as Alphafold3. It’s important to note that use of these techniques will always generate a structure, but this doesn’t guarantee they are correct or make sense.

Analysis of SEC-SAXS data was performed using SCATTER (obtained from http://www.bioisis.net/tutorial/9) and ATSAS CHROMIX and PRIMUS. Statistical testing of data included the CorMap test indicating location of non-fit for protein envelopes from randomness of residuals, the reduced $\chi 2$ test to standardise residuals (this is the gold standard statistical test for SAXS), and the Anderson-Darling test for checking whether the distribution of residuals is normal. UCSF Chimera 1.18 was used to visualise generated protein envelopes and fit these envelopes to protein structures.

Chapter 3

Identification of the Host Interactome of Old Word Arenavirus Nucleoproteins (NPs)

3 Identification of The Host Interactome of Old-World Arenavirus Nucleoproteins (NPs)

3.1 Introduction

Viruses exploit host cellular processes and factors, a characteristic that facilitates various steps in their viral lifecycle and helps them evade or suppress the host immune response. Investigating the interactions of arenavirus nucleoprotein (NP) with host factors is crucial for unravelling the molecular mechanisms underlying viral replication and transcription. This exploration provides insights into the co-evolutionary dynamics between the virus and its host, shedding light on the strategies employed by the virus to hijack cellular machinery.

Arenaviruses are categorised as neglected tropical diseases and are often highly pathogenic, such as LASV, requiring CL-4 facilities to handle and study the live virus. Without access to these facilities, alternative methods must be developed to study the key interactions of these viruses in lower containment level laboratories. These methods include using molecular tools such as recombinant viruses, study of individual viral proteins and using genetically similar but less pathogenic viruses as homologous models. Thus, we can see how research into unravelling the host proteomes of specific arenavirus proteins may have been limited prior to development of these molecular tools. Thus far, the only arenavirus NP interactomes identified using proteomics involving mass spectrometry have been LCMV and Junín NP in the context of infection of A549 and HEK293T cells (King et al., 2017a). Further non-viral, protein-specific proteomic research has been conducted in the context of persistent LCMV infection and the proteome response of HeLa cells (Benej et al., 2019). PICV was used as a biosafety level 2 model for LASV for investigation of the differential host cell nuclear proteome induced during arenavirus infection (Bowick et al., 2009). The proteomes identified have highlighted key pathways targeted by different

strains of arenaviruses and have provided key insights into their pathogenicity and most importantly begun building a foundation of data from which druggable targets can be identified for host-directed antivirals (Fang et al., 2022).

Arenavirus nucleoprotein (NP) is an extremely multifunctional viral protein that plays key roles in host immune suppression and modulation during arenavirus infection often leading to widespread and uncontrolled viral replication in various organs (Ma and Suthar, 2015). Disease can also vary from asymptomatic to fatal implying that the interplay between the host immune response and viral replication is a major predictive factor for disease outcome (Mazzola and Kelly-Cirino, 2019). Understanding how NP interacts with host components is pivotal for comprehending the host's defence mechanisms and the tactics employed by the virus to counteract or evade the immune response. One such example of NP modulating the host immune response is in its ability to target and counteract the host interferon response. Arenavirus NP inhibits detection and signalling of arenavirus RNA by RIG-I and MDA5 by degrading the RNA species and by directly binding to RIG-I. NP also binds to IKK-ε and NF-κB to prevent activation of interferon-responsive factor 3 (IRF3) which affects later stages of IFN induction.

In the context of arenavirus NP, studying these interactions may pinpoint specific host factors essential for viral replication. Targeting these factors presents potential therapeutic strategies, contributing to the development of antiviral drugs aimed at inhibiting viral replication and treating arenavirus infections. This is crucial in combatting the burden that arenaviruses present as a human pathogen as currently there are no licenced treatments, therapeutics or vaccines for Old World mammarenaviruses, those being investigated are non-specific and related to antigen or antibody levels rather than targeting NP-mediated mechanisms.

A detailed map of the cellular interactome of NP, considering both pathogenic and non-pathogenic strains, is essential to fully grasp their

multifunctional roles and identify key mechanisms as potential targets for therapeutic intervention. We have utilised biotin proximity labelling and mass spectrometry as our proteomic method of choice in the pursuit of characterising the molecular interactome of Old-World group arenavirus NP, specifically LASV, LUVJ, and MOPV NP.

3.2 Proximity based labelling and mass spectrometry - BioID

Over the past two decades, the integration of the two fields of proteomics and virology have transformed the ability to identify and understand the complex mechanisms underpinning the pathogenesis of many different viruses. Technological advances in mass spectrometry (MS) based proteomics and also in experimental work-flows for antibody-based immune-affinity purification (AP) of viral-protein complexes have allowed for the leaps and bounds to be taken in characterisation of virus-host protein interactions occurring during viral infection (Lum and Cristea, 2016). AP-MS based approaches allow for the identification of both direct and indirect interactions as well as enabling investigation of interactions during the progression of an infection. This information therefore provides important temporal information although, the information of most interest is that provided by overexpression of individual viral gene products (NP in this case) to elucidate the functions of single proteins. Another similar method often used is proximity-based labelling coupled to MS otherwise called proximity-dependent biotinylation MS (PDB-MS). Table 3.1 outlines the pros and cons of each method of MS based proteomics.

PDB-MS involves expressing a bait protein genetically fused to a biotin ligase (BioID), a horseradish peroxidase (HRP) or just a peroxidase (APEX). These fused enzymes can catalyse externally added biotins or phenolic biotins into reactive biotin intermediates which diffuse out to biotinylate proteins in close proximity to the bait protein. Once biotin labelling has occurred, cells are then lysed and a pulldown using streptavidin or

neutravidin is performed. The next step is identifying and quantifying with MS (Gingras et al., 2019, Roux et al., 2012, Rhee et al., 2013).

Crucial for PDB-MS is promiscuous biotinylation which is a covalent modification process dependent biotin intermediates randomly diffusing away from the biotin ligase enzyme. This process is limited by distance however, and proteins in close proximity to the bait/enzyme fusion are preferentially biotinylated and the strength of labelling diminishes with increasing distance. PDB-MS defines the area surrounding the bait within an effective labelling radius specific to the enzyme known as a neighbourhood. The proteins comprising this neighbourhood may constitute the actual physical contacts of the bait itself, or other proteins present in the bait/enzyme fusion vicinity by chance (Low et al., 2021). The effective labelling radius for BirA*, a mutant biotin ligase utilised in BioID has been estimated to be about 10nm (Kim et al., 2014).

BioID was developed using E. coli derived 35kDa BirA* which has a R118G mutation resulting in destabilisation of the catalytic domain. BirA* catalyses conversion of biotins to form the highly reactive biotinoyl-AMP intermediates which then dissociate prematurely and diffuse out and away and react with the neighbouring lysine residues in a promiscuous manner (Roux et al., 2012). An improved method has been developed called BioID2, this was created from *Aquifex aeolicus* derived biotin ligase which has an R40G mutation resulting in a smaller ligase which is 27kDa and more catalytically active. This increased catalytic activity results in more efficient biotinylation and decreased mis-localisation of the bait (Kim et al., 2016).

PDB-MS allows for the detection of protein-protein-interactions (PPIs) among both soluble and membrane proteins, aside from enriching for interactions that are transient, weak, of low abundance or with higher turnover (Gingras et al., 2019, Trinkle-Mulcahy, 2019, Hung et al., 2017).

Due to PDB-MS biotinylating proteins within cells, it enables labelling of fragile complexes or interactions as well as avoiding post-lysis artefacts.

Importantly, the affinity of biotin to streptavidin is probably one of the strongest (and yet reversible) biological interactions found in nature. Consequently, highly stringent conditions for sample denaturing, solubilization, capture, wash and extraction of biotinylated proteins can be used to maximise the recovery of hydrophobic proteins while minimising nonspecific background contaminants. Proximity labelling using BioID2 can be used in living cells as it does not have to require disruption of the cells and therefore preserves evidence of weak or short-lived interactions often not detectable utilising traditional techniques, however in the use of BioID, temporality of interactions was not being investigated and therefore the protocol chosen did disrupt the cells (Bosch et al., 2021)

Use of quantitative MS and dedicated bioinformatics algorithms such as SAINT allows for reduction and differentiation of background contamination by identifying significant differences in protein abundance between the experiment and negative controls (Choi et al., 2011)

A typical MS assay workflow is comprised of the following steps: (1) Pre-cleared protein lysate is incubated with beads conjugated with bait or epitope tag-specific antibodies (2) Washing of beads to reduce non-specific binding (3) Elution of purified complexes (4) using MS to identify eluted proteins.

There are pros and cons to all MS-based methodologies for the identification of and quantification of protein-protein interactions.

Affinity-purification MS is currently the most widely used high-throughput method for protein-protein-interaction investigations and involved a bait protein being selectively purified with either specific antibodies or affinity reagents which will pull any potential interaction partners, known as prey, down with the bait from tissue or lysate. These purified proteins are then identified and quantified by mass spectrometry and often the experiment is repeated with different bait proteins. It has the advantages that the co-immunoprecipitation can be done without tagged baits expressed at

physiological levels and still allow for identification of endogenous PPIs. Also, if no suitable antibodies are available, epitope tagging provides a useful alternative for protein purification, although this may interfere with bait functions and solubility and thus skew the results of the MS, ectopic expression of tagged baits also has the potential to result in misfolding and/or mislocalisation of bait proteins causing background contamination and unintended interactions.

Cross-linking MS (XL-MS) involves a selection protein or complex in a native state being chemically cross-linked with reagents able to covalently bind and tether amino acid residues in close spatial proximity. These cross-linked proteins then undergo proteolysis and peptide mixture is then separated and analysed by liquid chromatography-MS or standard MS. This process then requires subsequent interrogation of databases of the MS data to understand the sequence of crosslinked peptides and their crosslinked sites. It is useful because crosslinking reagents are able to covalently link two or more non-covalently interaction proteins, regardless of the strength of the interaction or its duration enabling very transient or weak PPIs to be analysed. However, this method is limited by crosslinking reagents often having very low efficiency (1-5%) causing marginal crosslinks in which only the top 20-30% of proteins can be detected. Crosslinking times can also be up to about 30 minutes in some cases and this long reaction time can cause large, crosslinked protein aggregates. When XL-MS is used in conjunction with structural biology techniques such as X-ray crystallography, Cryo-Electron-Microscopy (Cryo-EM), Nuclear Magnetic Resonance (NMR) or native MS, the spatial information provided by XL-MS can help elucidate molecular modelling allowing for construction of connectivity maps and determination of subunit topologies and further understanding of the protein complex's dynamic behaviour.

Co-fractionation coupled to MS (co-Frac-MS) analyses the spatiotemporal co-behaviour of biomolecules to identify functional or physical interactions between proteins (Feiglin et al., 2014). The idea that proteins

sharing similar co-fractionation profiles may also co-localise, due to the tendency of polypeptide constituents of shared assemblies to co-migrate in the same analytical column when in native conditions, underpins the reasoning behind this method. In co-Frac-MS, protein complexes from cell lysates are fractionated extensively under non-denaturing conditions with chromatographic or electrophoretic techniques. Each fraction then undergoes proteolysis and is analysed by LC-MS or native MS, after which the proteome composition is identified and quantified and the fractionation profiles of individual protein complexes can be built. Due to the necessity of preserving the intact protein complexes, cells or tissue are rapidly lysed under refrigerated and native conditions with minimal dilution and non-denaturing conditions are required. This method of investigating PPIs has high throughput and provides global identification and quantification of native protein complexes without the need for genetic modification or transient overexpression of proteins allowing for endogenous and physiologically relevant interactome data. A downside of this method is that false positives are a significant issue resulting from random co-elution.

Thermal proximity co-aggregation (TPCA) utilises the phenomenon that interacting proteins co-aggregate and co-precipitate after heat induced denaturation (Tan et al., 2018). Due to this phenomenon, these proteins will have high similarity in their thermal solubility when compared to non-interacting proteins and the assembly state of known protein complexes can then be inferred from this similarity. Changes in thermal solubility of proteins can also be used to identify those proteins modulated across cellular states or physiological conditions. Isobaric tandem mass tag (TMT) reagents are used to simultaneously quantify protein solubility across ten different temperatures from a cellular thermal shift assay (CETSA). This technique allows for system-wide profiling of protein complex dynamics and does not require use of antibodies or epitope tagging. Preparation time for this technique is minimal and enables high coverage study of protein complexes in situ and in vivo, as such, this method can be rapidly utilised

to elucidate the assembly states of protein complexes across cellular state, type, tissue, and physiological conditions. This provides crucial information about the functions of these protein complexes and can be used to compare normal and diseased cells. The major downside of this method is that, in its current state, TPCA is limited to only studying the dynamics of known or predicted protein complexes across cellular states and physiological conditions. Existing and extensive interaction data including graph and network clustering information is required to infer novel protein complexes.

Based upon this information and that outlined in Table 3.1 it was decided that PDB-MS utilising BioID2 would be the best method for identifying and characterising the interactome of Old-World arenavirus NP. The advantages of BioID over other methods to screen for protein interactions are that BioID is effective when applied to insoluble and membrane-associated proteins which are two classes of protein potentially refractory to screening with conventional approaches such as AP-MS. BioID also enables identification of weak, transient, or hydrophobic interactions in a relatively natural cellular environment and with modifications is effective in a wide variety of cell types and species.

Methods	Pros	Cons
AP-MS	Co-immunoprecipitation can be performed without tagged baits expressed at physiological levels to identify endogenous PPIs	Co-IP with untagged baits is limited by the availability of antibodies, and the low expression levels of baits
	Epitope tagging provides an alternative for purifying proteins lacking suitable antibodies	Epitope tags may interfere with the functions and solubility of the baits
	Transient transfection of tagged baits enhance their expression, thus improving the efficiency and throughput of the pulldowns	Ectopic expression of tagged baits may promote misfolding and mis-localisation of the baits, promoting background contamination and spurious interactions
PDB-MS	Allows detection of PPIs among both soluble and membrane proteins, as well as enriching for PPIs that are transient, weak, low abundance or have high turnover (Gingras et al., 2019,	May react with biotin-phenol and H_2O_2 to produce reactive radicals resulting in cellular toxicity (APEX) (Che and Khavari, 2017)

Methods	Pros	Cons
	Trinkle-Mulcahy, 2019, Hung et al., 2017)	
	Avoids post-lysis artefacts (Gingras et al., 2019)	The accessibility and labelling efficiency of the biotinylating enzyme are locality-dependent, as its orientation and topology within the protein complex may impede its performance
	The affinity of biotin to streptavidin is robust yet reversible. Hence, highly stringent conditions for sample denaturing, solubilization, capture, wash and extraction of biotinylated proteins can be employed to maximize the recovery of hydrophobic proteins while minimizing nonspecific background contaminants	The high affinity of the streptavidin–biotin interaction may hinder the recovery of highly biotinylated proteins. PDB-MS suffers from false positives in the forms of high-abundance background proteins or artefacts from

Methods	Pros	Cons
		endogenous biotinylation
		The labelling time for different enzyme varies from 1 min to 24 h (Roux et al., 2012, Rhee et al., 2013, Kim et al., 2016)
XL-MS	Crosslinking reagents can covalently connect two or more non-covalently interacting proteins, regardless of the duration and strength of the interaction. As such, even transient and weak PPIs can be preserved (De Jong et al., 2017, Walker-Gray et al., 2017)	The low efficiency (~ 1–5%) of crosslinking reagents, which often results in marginal crosslinks, where only the top 20–30% of proteins are detected
	When used in combination with X-ray crystallography, CryoEM, NMR and native MS, the spatial constraint data from XL-MS can guide	The crosslinking reaction time may be relatively long (~ 30 min). Excessively long reaction time

Methods	Pros	Cons
	molecular modelling, construct a connectivity map for determining subunit topology, and map the dynamic behaviour of the protein complex (Liu and Heck, 2015, Yu and Huang, 2018, Tang et al., 2021)	may result in large, crosslinked protein aggregates
	To expand the number and coverage of crosslinks, alternative modes of crosslinking can be employed, such as carboxyl-targeting reagents (Steigenberger et al., 2019a, Steigenberger et al., 2019b, Tan et al., 2016)	A crosslinker covalently links two linear peptides, giving rise to a hybrid dipeptide that can dramatically expand the search space during spectra matching, giving rise to the 'n-square problem' (Liu et al., 2020, Courouble et al., 2021)
Co-Frac-MS	CoFrac-MS has high throughput, and it provides global identification and quantification of native	False positives constitute a significant problem in the form of chance co-elution

Methods	Pros	Cons
	protein complexes in one setting	
	It can be operated without genetic manipulation and overexpression, thereby inferring endogenous, physiologically relevant interactome (Titeca et al., 2019)	
	CoFrac-MS combined with quantitative proteomics can delineate the relative distribution of a protein in multiple co-elution features. Thus, the stoichiometries and dynamics of a target protein within different co-isolated complexes can be simultaneously elucidated (Mallam et al., 2019)	
TPCA	TPCA permits system-wide profiling of protein complex dynamics, and it requires neither antibodies nor	The current version of TPCA is limited to studying the dynamics of known or predicted

Methods	Pros	Cons
	epitope tagging (Tan et al., 2018)	protein complexes across cellular state and physiological conditions. Need to incorporate existing interaction data with graph/network clustering algorithms to identify novel protein complexes (Tan et al., 2018)
	Little preparation time is required. It allows most of the study of protein complexes in situ and in vivo	
	TPCA profiling can be rapidly deployed to unravel the assembly state of protein complexes across cellular state, cell type, tissue and physiological conditions to provide insight into their functions in normal and diseased cells	

Table 3.1 List comparing the pros and cons of MS based methodologies for identifying and quantifying PPIs adapted from (Low et al., 2021)

3.3 Chapter Objectives

This chapter aimed to:

1. Identify host protein interaction networks for Old World (OW) arenavirus nucleoproteins (NPs), specifically those of the pathogenic Lassa virus (LASV) and Lujo virus (LUJV), and the non-pathogenic Mopeia virus (MOPV).
2. Determine which cellular functions and pathways are targeted by these viral NPs.
3. Elucidate mechanisms underlying OW arenavirus pathogenicity by comparing host protein interactions between pathogenic and non-pathogenic arenavirus NPs.
4. Address the current knowledge gap in OW arenavirus NP interactomes beyond the well-studied lymphocytic choriomeningitis virus (LCMV).

3.4 Results

3.4.1 BioID2 of MOPV NP

Despite the minimisation of nonspecific background contaminants, it is still crucial to design and implement appropriate controls which will enable confident identification true interactors from nonspecific binding. The experimental control used in the PDB-MS protocol BioID2, utilises expression of an empty vector, MCS-BioID2-HA .

Samples were analysed by western blot to evaluate protein expression levels of the transfected and immunoprecipitated arenavirus NPs. This was a necessary step to ascertain if the samples successfully expressed NP and if it was worthwhile to send them for mass-spectrometry analysis. Samples were analysed by western blots stained with Strep-HRP antibody, or anti-HA antibody to stain for biotinylated proteins. Figure 3.1 shows the results of these western blots and that of the three NP samples, only MOPV NP-BioID2-HA was successfully immunoprecipitated.

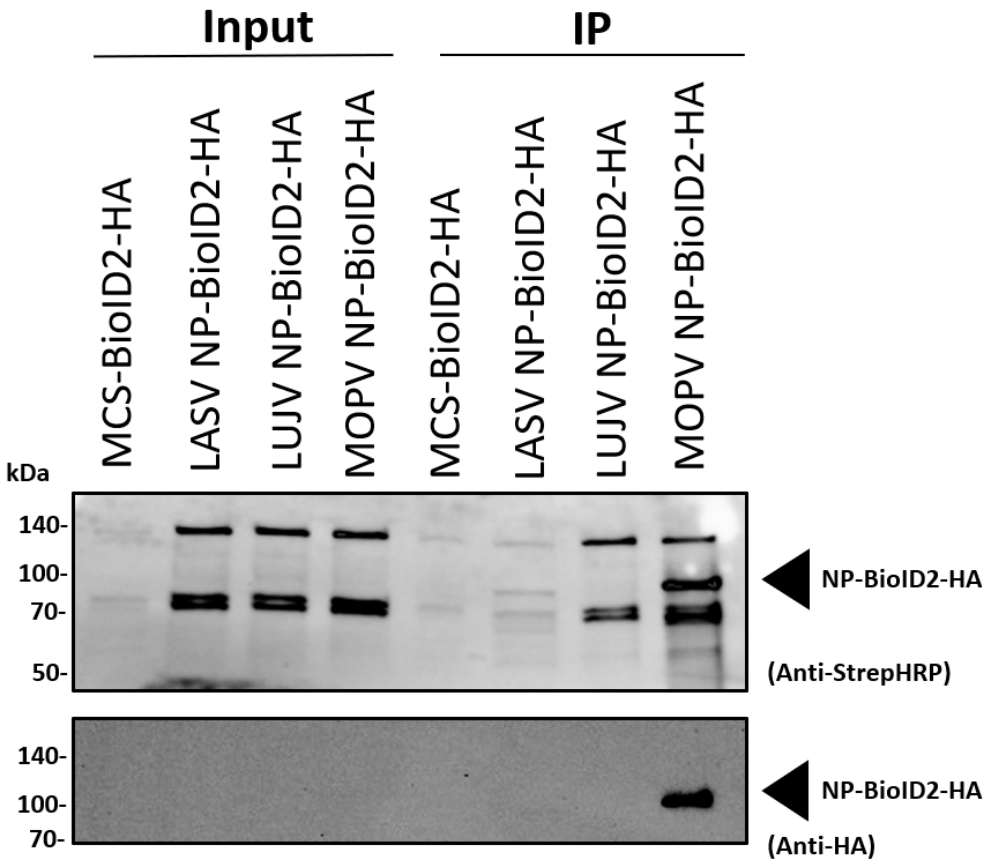


Figure 3.1 Western blot images of BioID2 samples comparing input samples to immunoprecipitated (IP) samples to evaluate expression of NP stained using mouse anti-HA specific for the HA tag on the BioID2 vector. Western blots of biotinylated proteins shown using StrepHRP staining and displaying good LASV, LUJV, and MOPV NP expression using mouse-anti-HA primary antibody (n=1)

This sample, and the control sample, were sent to ASMF at the University of Birmingham for proteomic analysis by mass spectrometry. The non-quantitative mass spectrometry performed by Birmingham generated data able to confirm or deny the presence of proteins in the sample compared to the control. Protein presence was decided based upon two factors: the presence of the protein within the MOPV NP sample and its absence in the control sample. The list of proteins identified as unique to the MOPV NP sample has been displayed as a STRING protein network in Figure 3.2. From this data, a list of host proteins interacting with MOPV NP was generated.

3.4.2 Identification of human proteins associating with MOPV NP

The mass spectrometry returned 187 host proteins interacting with MOPV NP as shown in Figure 3.2 which is a proteome network generated by inputting the filtered mass spectrometry gene list into STRING. Settings for the STRING network were set to the highest stringency with a minimum required interaction score of 0.70 (high confidence) and only proteins that have been mentioned together in all mined literature and then shown to interact biochemically through experimentation such as co-purification, co-crystallisation, or yeast2hybrid have their interactions displayed as connecting lines. Associations are meant to be specific and meaningful, and proteins are shown to be jointly contributing to shared function(s); this does not necessarily mean that these proteins are physically binding to each other and directly interacting.

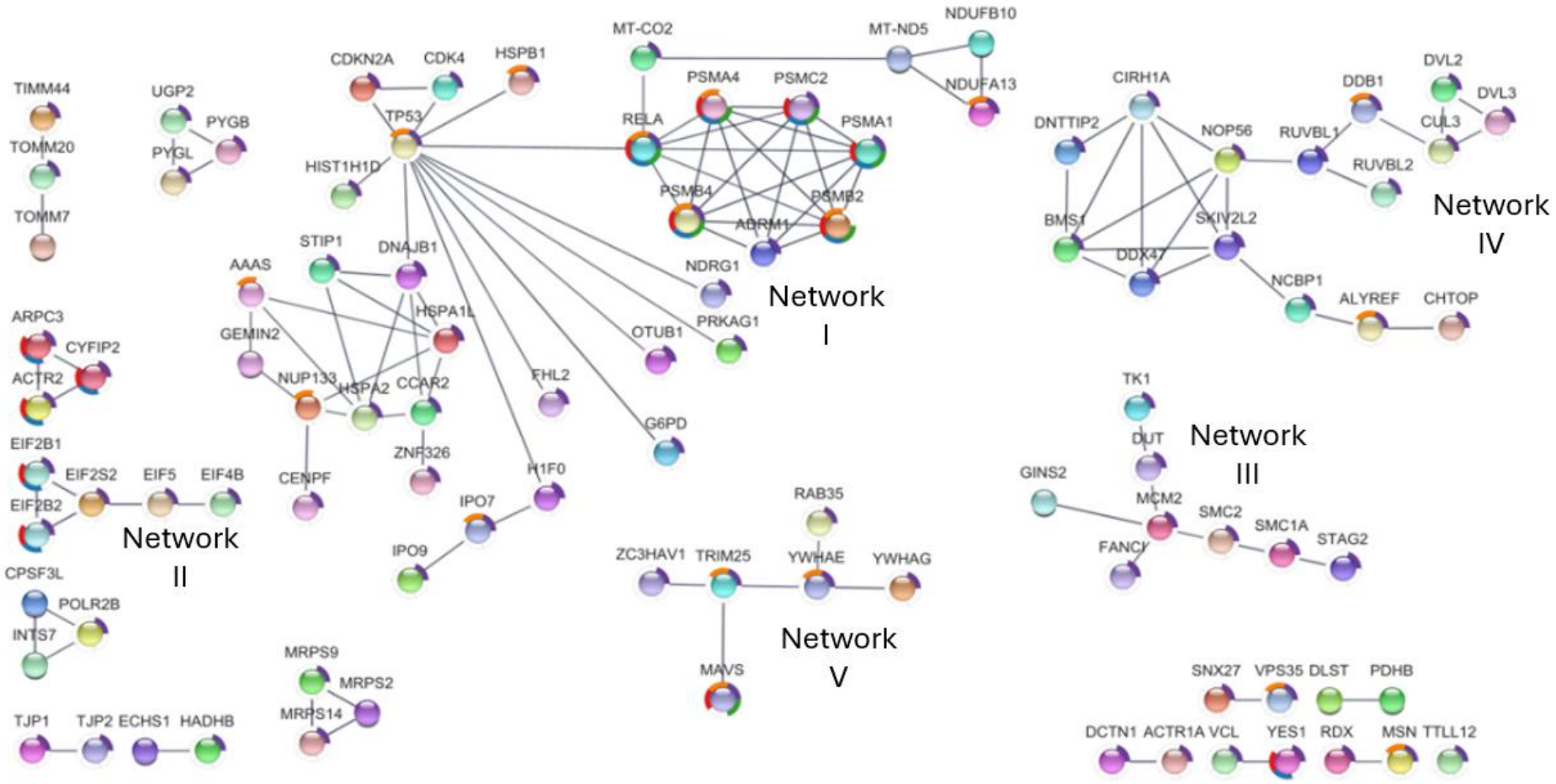

TIMM44
TOMM20
TOMM7
UGP2
PYGB
PYGL
CDKN2A
CDK4
HSPB1
TP53
HIST1H1D
MT-CO2
MT-ND5
NDUFB10
NDUFA13
PSMA4
PSMC2
RELA
PSMA1
PSMB4
PSMB2
ADRM1
NDRG1
Network I
STIP1
DNAJB1
AAAS
HSPA1L
GEMIN2
NUP133
HSPA2
CCAR2
OTUB1
PRKAG1
FHL2
G6PD
ZNF326
CENPF
H1F0
IPO7
IPO9
ARPC3
CYFIP2
ACTR2
EIF2B1
EIF2S2
EIF5
EIF4B
EIF2B2
Network II
CPSF3L
POLR2B
INTS7
TJP1
TJP2
ECHS1
HADHB
MRPS9
MRPS2
MRPS14
RAB35
ZC3HAV1
TRIM25
YWHAE
YWHAG
MAVS
Network V
CIRH1A
DNTTIP2
NOP56
RUVBL1
RUVBL2
DDB1
DVL2
DVL3
CUL3
BMS1
DDX47
SKIV2L2
NCBP1
ALYREF
CHTOP
Network IV
TK1
DUT
GINS2
MCM2
SMC2
SMC1A
STAG2
FANCI
Network III
SNX27
VPS35
DLST
PDHB
DCTN1
ACTR1A
VCL
YES1
RDX
MSN
TTLL12

Figure 3.2 STRING proteome protein network of host interacting partners of MOPV NP. Every node in the network represents a protein that interacts with MOPV NP based upon the BioID2 protocol and subsequent mass spectrometry and downstream bioinformatics data processing. Settings for the STRING network were set to the highest stringency feasible and only proteins that have been mentioned together in all mined literature and then proven to interact biochemically through experimentation such as co-purification, co-crystallisation, or Yeast2hybrid have their interactions displayed as connecting lines. Associations are meant to be specific and meaningful, and proteins are shown to be jointly contributing to shared function(s); this does not necessarily mean that these proteins are physically binding to each other and directly interacting. Network nodes represent proteins and splice isoforms, or post-translational modifications are collapsed so that each node represents all the proteins produced by a single, protein-coding gene locus. Edges represent protein-protein interactions and are colour coded according to type of interaction. Light blue denotes a known interaction from a curated database and purple is a known interaction which has been experimentally determined. Predicted interactions are either green for gene neighbourhood, red for gene fusions, or dark blue for gene co-occurrence. Other edges include light green for text-mined data, black for co-expression, and grey-blue for protein homology. In terms of nodes, those that are empty are for proteins of unknown 3D structure whereas those that are filled have some or all of their 3D structure known or predicted. Only proteins that are involved in interactions with other identified proteins and make-up a network have been visualised. Those proteins not currently known to be interacting with any other identified proteins have been omitted. These networks of interest have been labelled from I-VI.

Of the 187 identified proteins, 153 proteins are implicated in some form of general binding with the majority involved in RNA binding specifically. Further analysis of molecular functions in which these proteins are involved in revealed significant overlap between functions, for example, many of the proteins involved in RNA binding also had roles in enzyme binding. This is the same for the majority of proteins identified suggesting that NP is not only very multifunctional itself but utilises multifunctional host proteins to assist it in immune suppression and modulation as shown by the networks identified in Figure 3.2 which all contain proteins with some role in these pathways.

The earlier data from the AMSF for MOPV NP shows a number of different mini networks when displayed as a STRING proteome (Figure 3.2). Network I containing PSMB2, PSM, PSMA4, ADRM1, PSMA1, and PSMC2 is highly interconnected, and all of these proteins have key roles in immune response antigen presentation by MHC class I. They are also implicated in the CDK5 (Cyclin-dependent kinase) pathway and some DNA repair mechanisms. The likely interaction of NP in this pathway involved modulation of the host immune response to viral infection. CDKs are protein kinases with roles in cell division and transcriptional regulation

although recently they have been shown to have vital roles in various viral infections (Yan et al., 2022). The CDK5 pathway is involved in many systems but key here is likely to be its role in the IFNγ-induced programmed death ligand (PD-L1) upregulation, allowing for certain cells to evade detection by the immune system (Shupp et al., 2017). NP is potentially interacting with these proteins to inhibit them from marking infected cells as aberrant and infected so as to avoid detection by the host innate immune response and enable arenavirus replication to progress unhindered.

Another identified network with TP53 at its centre is network II which includes RELA, G6PD, CDK4, CDKN2A, HSPB1, DNAJB1, and STIP1, IPO9, IPO7, H1F0, NDRG1, HIST1H1D, FHL2, OTUB1, and PRKAG1. This subset has TP53 as its most multi-functional protein and connected protein here with roles in the Akt signalling pathway, apoptotic signalling pathway, MAPK signalling, and mTOR signalling; all of which involve at least one of the other proteins in this sub-network. The Akt signalling pathway also involves-PI3K and is an important mechanism through which some viruses have been shown to influence various cell functions to slow down apoptosis to prolong viral replication in both acute and persistent infections. There is growing evidence that Akt might be crucial for the survival of some viruses (Ji and Liu, 2008). Mitogen activated protein kinase (MAPK) is a major cell signalling pathway involved in converting extracellular stimuli into many different cellular responses and there is much evidence that it is activated by a broad range of viruses (Kumar et al., 2018). It has been shown that activation of the MAPK pathway is required for JUNV replication (Rodríguez et al., 2014) so it is likely that this pathway is also involved in promoting old world arenavirus replication. The mammalian target of rapamycin (mTOR) pathway is a key regulator of gene expression and translation as well as being involved in many metabolic processes. Viral infection involved often involves significant cellular stress which can activate, reduce, or inhibit the mTOR pathway to subvert this stress response and prolong viral replication within host cells with hospitable environments (Le Sage et al., 2016).

Network III including EIF5, EIF4B, EIF2S2, EIF2B1, and EIF2B2 is made up of different eukaryotic translation initiation factor subunits all involved in the EIF2 pathway. EIF2 binds GTP and Met-tRNAi to form what is called the ternary complex which then associates with the 40S ribosomal subunit and other EIF proteins to form the 43S pre-initiation complex. This signalling mediated by EIF2 also works to regulate pro-inflammatory cytokine expression and bacterial invasion (Shrestha et al., 2012).

Network IV comprises SMC2, SMC1A, STAG2, MCM2, GINS2, DUT, FANCI, TK1, and GINS2. These are proteins involved in either DNA metabolic processes such as repair, or mitotic cell processes. This again points to NP regulating host cell environments to prevent any interruptions to viral replication caused by cell stress or death or by being recognised as infected by the host immune response.

NOP56, RUVBL1, RUVBL2. DDX47, BMS1, DNTTIP2, CIRH1A, SKIV2L2, NCBP1, ALYREF, DVL1, DVL2, CUL3, DDB1, and CHTOP make up the larger network V with many proteins involved in different functions though the main function they all share is RNA processing in the nucleus and cytosol. In the context of interacting with arenavirus NP, hijacking of the host cell RNA processing pathways is crucial for viral RNA replication.

TRIM25, YWHAE, YWHAG, MAVS, and ZC3HAV1, RAB35 are proteins in network VI that all have roles in the innate immune defence against viruses. Some of these proteins are the products of Interferon Stimulated Genes (ISGs) such as TRIM25 and ZC3HAV1 otherwise known as Zinc Antiviral Protein (ZAP) and have specifically antiviral functions and are key to interferon mediated host immune response and viral mRNA targeting and degradation (Zhu and Gao, 2008). MAVS is the mitochondrial antiviral signalling protein involved in the toll-like receptor signalling pathways, NF-kappaβ signalling, and the intrinsic innate immunity signalling pathway. RAB35 is involved in autophagy and YWHAE and YWHAG have roles in the apoptosis and MAPK signalling pathways. Similar to many of the smaller two or three protein pathways, almost all the proteins in this pathway are

also involved in some form of host cell regulation and the role of NP appears to be to maintain a host cell environment that promotes viral replication whilst avoiding detection of infected cells by the host immune system.

Network VI contains two proteins of interest: ZAP (ZC3HAV1) and TRIM25. These proteins have significant antiviral roles in other important RNA viruses (Tang et al., 2017, Galão et al., 2022). ZAP is able to broadly target and bind directly to the viral mRNA of a broad range of many RNA viruses and TRIM25 is a very important cofactor of ZAP (Li et al., 2017), as well as being directly antiviral by itself due to its role in the RIG-I pathway to trigger IFN-I expression upon viral infection (Choudhury et al., 2022). The nuclear form of TRIM25 has been implicated in having direct antiviral activity against Influenza A vRNP (Meyerson et al., 2017). Both of these proteins also work in conjunction to target the Ebola virus ribonucleoprotein to mediate interferon-induced restriction (Galão et al., 2022). ZAP has also been shown to inhibit HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation (Zhu et al., 2011). Both ZAP and TRIM25 have been identified in other arenavirus NP interactomes but no further research has been done on them in the context of arenaviruses (King et al., 2017a). As such, this subset within my identified network is what I will focus on for my research and I will use live MOPV infection as a LASV infection homologue to validate their roles within arenavirus infection. To investigate the differences and/or similarities between both pathogenic arenavirus NP interactions and non-pathogenic arenavirus NP interactions, we performed BioID2 and LFQ-MS using the University of Liverpool Centre for Proteome Research to provide quantitative data on NP-host interactions.

3.4.3 BioID2 and mass spectrometry of LASV, LUJV, and MOPV NP using label-free quantification (LFQ)

Validation of the MOPV NP interactome was necessary, as was identifying host interactomes for LASV NP and LUJV NP. The BioID2 protocol was used again although optimisations were made to improve transfection efficiency of NP constructs and generate increased NP expression as well as decreasing the total time taken for the protocol as no expansion step was required. Optimisations included starting the process from a much larger foundation; 8x10cm dishes of HEK293T cells were set up per construct (Control, LASV, LUJV, MOPV NP) and 20µg of DNA was transfected into each dish. No G418 antibiotic selection was used. The lysis step was altered to be longer, and cells were lysed for an hour on a rolling platform at 4°C and the sonication step was removed.

Expression and successful immunoprecipitation of arenavirus NP is shown by western blot in figure 3.3. All three arenavirus NP samples (LASV, LUJV, and MOPV NP) showed good levels of expression and immunoprecipitation as shown by the HA blot in Figure 3.3 and biotinylation was successful as shown by the StrepHRP blot. Based on these results it was decided that the immunoprecipitated samples should be sent for proteomic analysis by mass spectrometry.

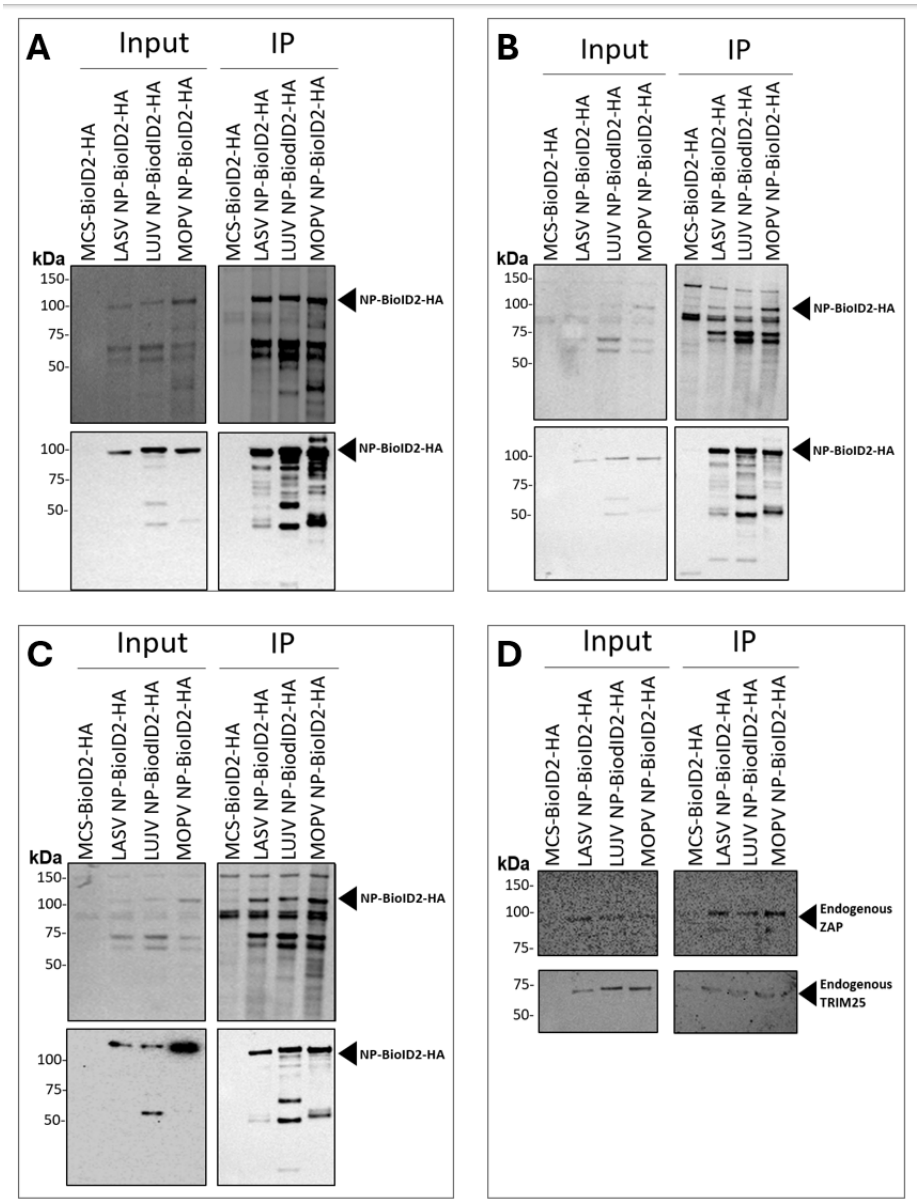


Figure 3.3 Western blots of BioID2 samples analysing NP expression by immunoprecipitation prior to sending off for mass-spectrometry. The top blot was stained using anti-StrepHRP conjugated antibody requiring no secondary antibody and the bottom blot was stained using anti-HA raised in mice and anti-mouse secondary antibody. The red arrows point to the band for the arenavirus NP within the BioID2-HA vector. The total molecular weight of the BioID2 vector (27kDa), HA tag (1.1kDa) and arenavirus NP (63-68kDa) is approximately 91.1-96.1kDa. The input samples were taken from the cleared supernatant after cell lysates were centrifuged but prior to addition of streptactin Sepharose beads and the IP samples are taken from the bead samples post pulldown. (A) Represents replicate 1 (B) Represents replicate 2 (C) Represents replicate 4. Samples were prepared in quadruplicate but poor NP expression in replicate 3 indicated that it was not worthwhile sending off for analysis by MS. (D) Western blot of BioID2 samples from Replicate 1 stained for endogenous ZAP or endogenous TRIM25 (n=3).

The subsequent raw mass spectrometry data from the Liverpool Centre for Proteomics was processed using MaxQuant which is a software package commonly used for quantitative proteomics, as the data was label-free, the MaxLFQ label free quantification module integrated into MaxQuant was used. This processed data was then analysed using FragPipe-Analyst which is an interactive web application allowing differential expression analysis compatible with LFQ data and based upon the original LGQ-Analyst code. This analysis resulted in 66 out of 1252 significant features with 5.27% of features differentially expressed across all conditions (Control, LASV NP, LUJV NP, and MOPV NP). This data and the variation both between replicates and between conditions can be seen in the exploratory analysis and quality control plots which are shown in figure 3.4A, B, and C. The PCA plot in figure 3.4A depicts feature abundance levels from across all sample runs used to determine the principal axes of abundance variation. The difference seen between conditions is expected due to the hypothesis that each NP strain will have different interactomes and each will be different from the control. What's interesting is the information the PCA plot between replicates, the LASV NP samples have the most variation. whilst the control sample replicates and the MOPV samples replicates have the lowest amount of variation. Although for the Control, LUJV, and MOPV NP sample replicates there exists an outlier replicate indicating significant difference to the other replicates in close proximity (and therefore close similarity). This difference observed between replicates for a given sample has been factored into the down-stream statistical analysis performed by the FragPipe-Analyst affecting the stringency applied to proteins identified as interacting with NP or not.

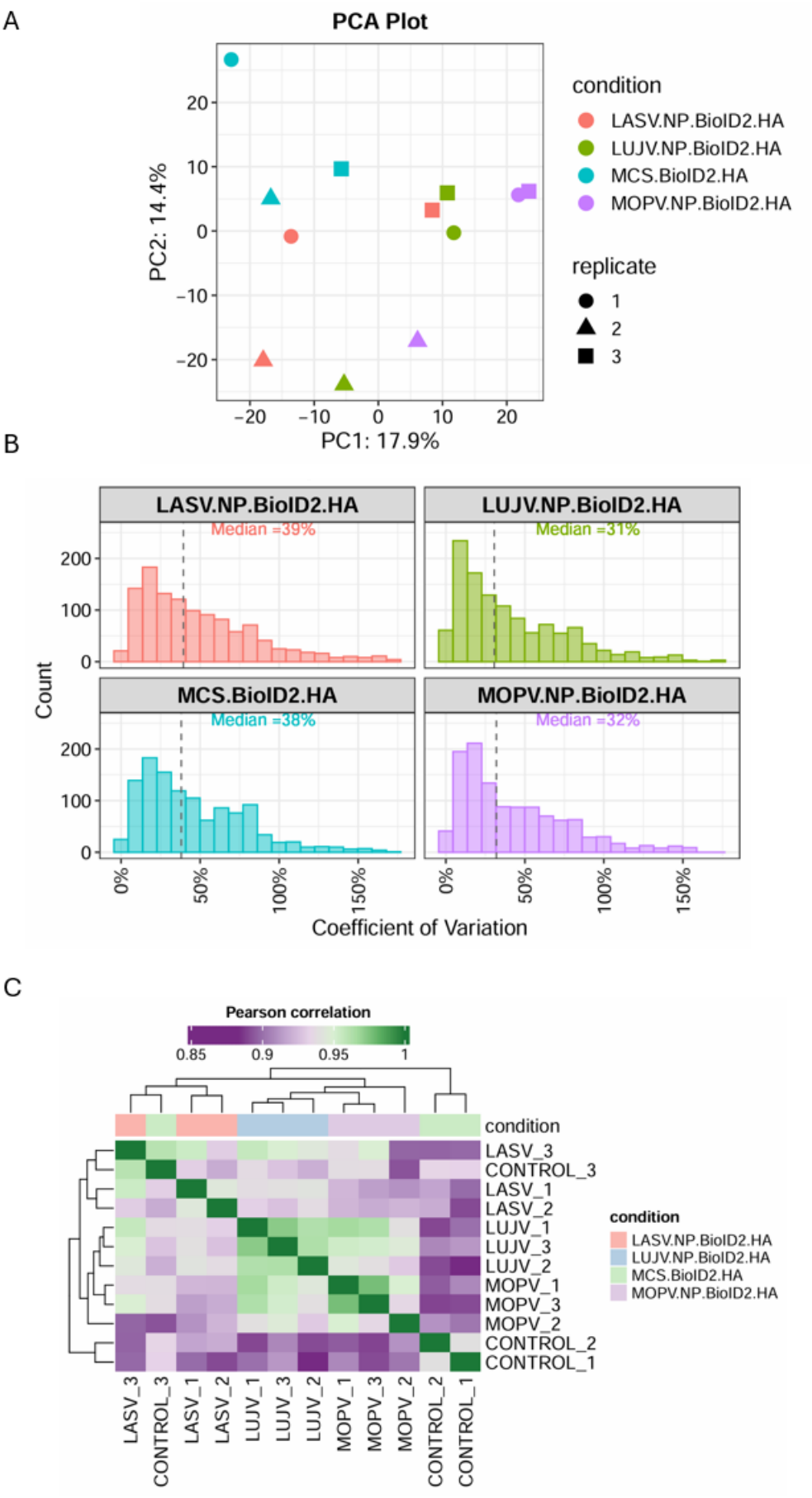

A
PCA Plot
PC2: 14.4%
PC1: 17.9%
condition
LASV.NP.BioID2.HA
LUJV.NP.BioID2.HA
MCS.BioID2.HA
MOPV.NP.BioID2.HA
replicate
1
2
3
B
LASV.NP.BioID2.HA
Median =39%
LUJV.NP.BioID2.HA
Median =31%
MCS.BioID2.HA
Median =38%
MOPV.NP.BioID2.HA
Median =32%
Count
Coefficient of Variation
C
Pearson correlation
condition
LASV_3
CONTROL_3
LASV_1
LASV_2
LUJV_1
LUJV_3
LUJV_2
MOPV_1
MOPV_3
MOPV_2
CONTROL_2
CONTROL_1
condition
LASV.NP.BioID2.HA
LUJV.NP.BioID2.HA
MCS.BioID2.HA
MOPV.NP.BioID2.HA

Figure 3.4 Exploratory analysis and Quality control plots. (A) Principal Component Analysis (PCA) plot, with principal component 1 on the x-axis and principal component 2 on the y-axis. feature abundance levels shown in the PCA plot are from across all sample runs and are used to determine the principal axes of abundance variation. This style of plotting the data allows for separation of the run samples according to abundance variations and is very useful in identifying outliers in the data in an easy to visualise style. Features that are close together do not show much variation and vice versa, with the greatest difference between samples represented by greater distance between features. By plotting replicates as well as features in a biplot we are able to show which features are contributing to the difference between runs and thus can more easily interpret the relationship between experimental groups and features with runs on the positive side of the axes having high abundance values for the feature whilst runs on the negative side have low abundance values. It's key to note that the closer the run is to the axis, the more the influence of this feature for that run position. (B) Sample Coefficient of variation (CVs) C) Sample matrix of Pearson correlation coefficients between all pairs of the 12 spectra generated for assessing analytical reproducibility presented in the form of a heatmap. Each column and row represents a spectrum. The Pearson correlation coefficient between two spectra is graphically indicated by a colour code which has a gradient ranging from green (highest correlation) to dark blue (lowest correlation).

The coefficient of variation plots shown in Figure 3.4B informs on the measure of dispersion and relative measure of variability of the samples. Comparisons of how close multiple measurements from different samples are to each other. Lower correlation of variance equate to lower deviation compared to the mean and thus mean increased reproducibility and increased data quality. The median across all of the samples in the data set is under 40% which can be considered a low value. This indicates that the data is of good quality. Figure 3.4C shows a Pearson correlation graph; in this graph the relationship between two variables is generally considered strong when their r value is larger than 0.7. As all minimum value for the graph displayed in Figure 3.4C is 0.85 this indicates that all the samples have high similarity. This is to be expected due to the large number of proteins identified that match across all samples, the data of interest here is that showing dis-similarity as this indicates differences in the data pertaining to differences in proteins identified to be interacting with each condition.

Figure 3.5 represents missing value handling within the data. It is rare that mass spectrometry returns observations for the same numbers of proteins in all replicated experiments. This means that more often than not there will

be missing values for some proteins when analysing all of the experiments together. There are a couple of options for handling these missing values: one can either; only analyse the proteins that we have observations for in all of our experiments; or one can impute values for the missing values from existing observations. Each method has its limitations and extensive study of various imputation and imputation free methods is ongoing (Wei et al., 2018, Taylor et al., 2022). In the experiments, a Perseus-type imputation was used to fill in the missing values. Figure 3.5A and 3.5B represent the missing data for each protein and also the difference that imputation made when applied. This method of data normalisation allows for statistical inference by accounting for run-to-run variations. Figure 3.5B is a graph of the protein data for all assays plotted as distributions; by considering our proteomics data as a distribution of values where there is one value for the concentration of each protein in our experiment that when put together forms the overall distribution in our sample. Figure 3.6 represents the number of features analysed across each sample and shows the difference between number of identified features. The western blots in Figure 3.1 may provide some explanation for the differences in identified feature numbers across samples it can be seen that there are differences in expression levels of proteins pulled down alongside NP during the BioID2 protocol. The relative differences in NP expression of each strain between different samples may cause variations in which interactions are identified. For example, Figure 3.1A shows stronger expression of MOPV NP in replicate 1 compared with 3.1B for replicate 2 and this correlates to the bars shown in Figure 3.6 where there are more features identified in MOPV replicate 1 than in replicate

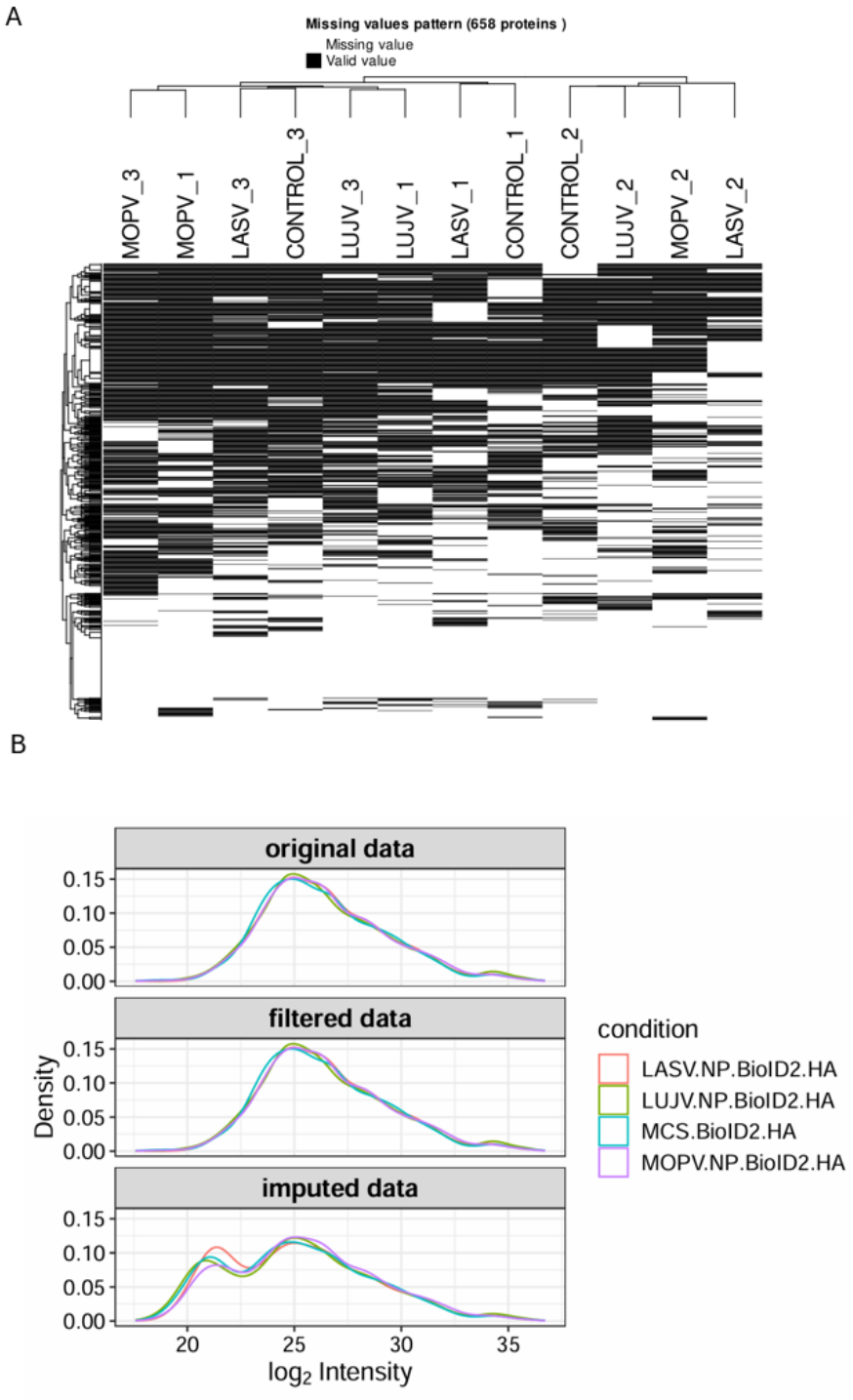


Figure 3.5 (A) Missing value Heatmap ; A heatmap for proteins with missing value in one or more dataset. Each row represent a protein with missing value in one or more replicate. Each replicate is clustered based on presence of missing values in the sample which are seen as zero values in the mass spectrometry data used to generate these figures. (B) Missing value distribution: Protein expression distribution before and after imputation. The plot showing the effect of imputation on protein expression distribution. (A) represents the values without imputed values B) represents imputed values taken into account.

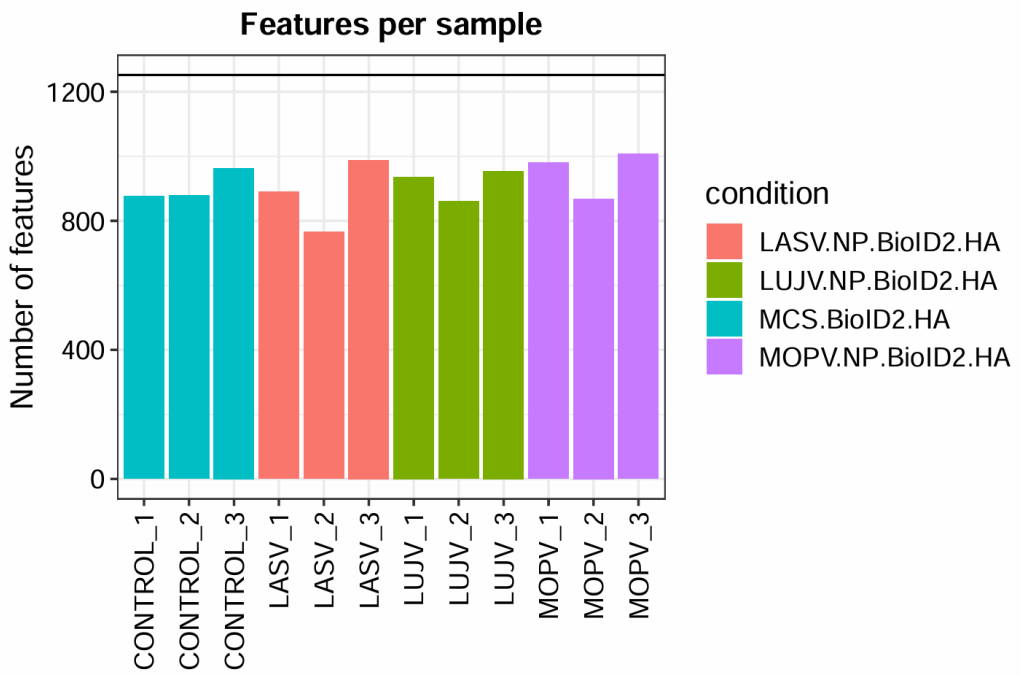


Figure 3.6 Proteomics Experiment Summary. (A) shows proteins quantified per sample (after pre-processing). Each bar represents a different biological replicate that was sampled and depicts the differences in peptide features identified between each sample.

3.4.4 Old World group arenavirus NP host interactome

Figure 3.7 shows a heatmap representing analysis of differential expression and results. Heatmaps allow for understanding of the relations between both proteins and between samples and can be used similar to PCA plots to evaluate sample similarity. However, heatmaps are different to PCA plots in that they are based directly on the distance between the untransformed protein quantifications within each sample and the heatmap dendrogram is built based on this information. The distances represent a measure of similarity where proximity represents similarity and dissimilarity increase with distance. The heatmap dendrogram includes branches of similar samples being grouped together and shows which groups of proteins in the samples have similar abundance patterns. From the heatmap displayed in Figure 3.7 there are both significant differences and similarities between the conditions. The clustering along the top of the heatmap indicates that the three NP samples are very different from the control samples shown by the low intensity cluster in the top right. This is likely due to the enrichment of proteins from interaction with NP and indicates the different

interactomes identified. Most of the bottom cluster of features have significantly lower abundances in LASV, LUJV, and the control samples when compared with MOPV NP.

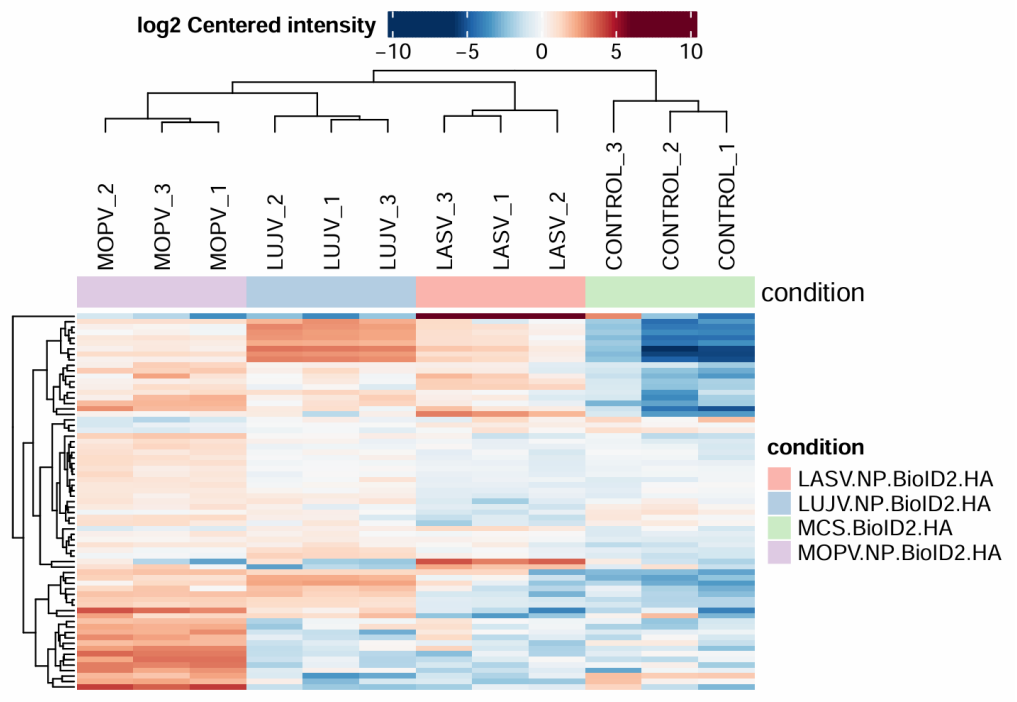
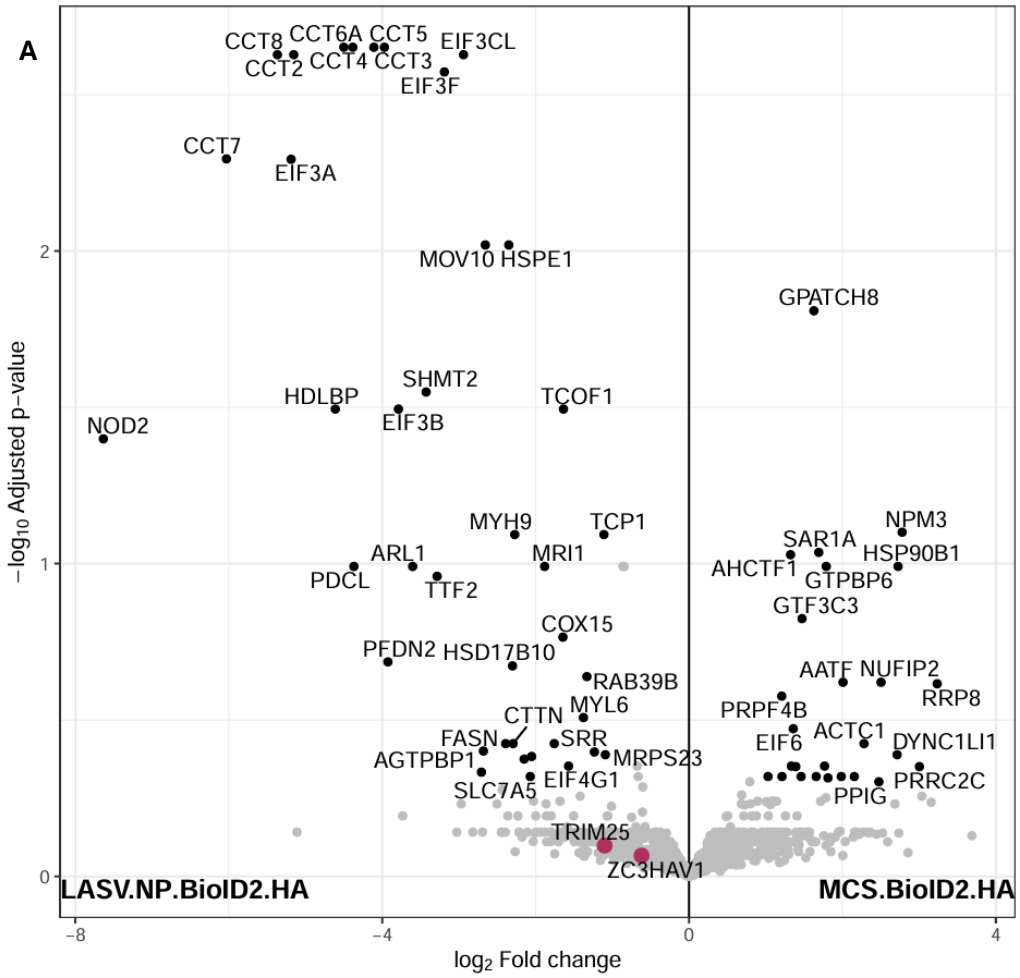
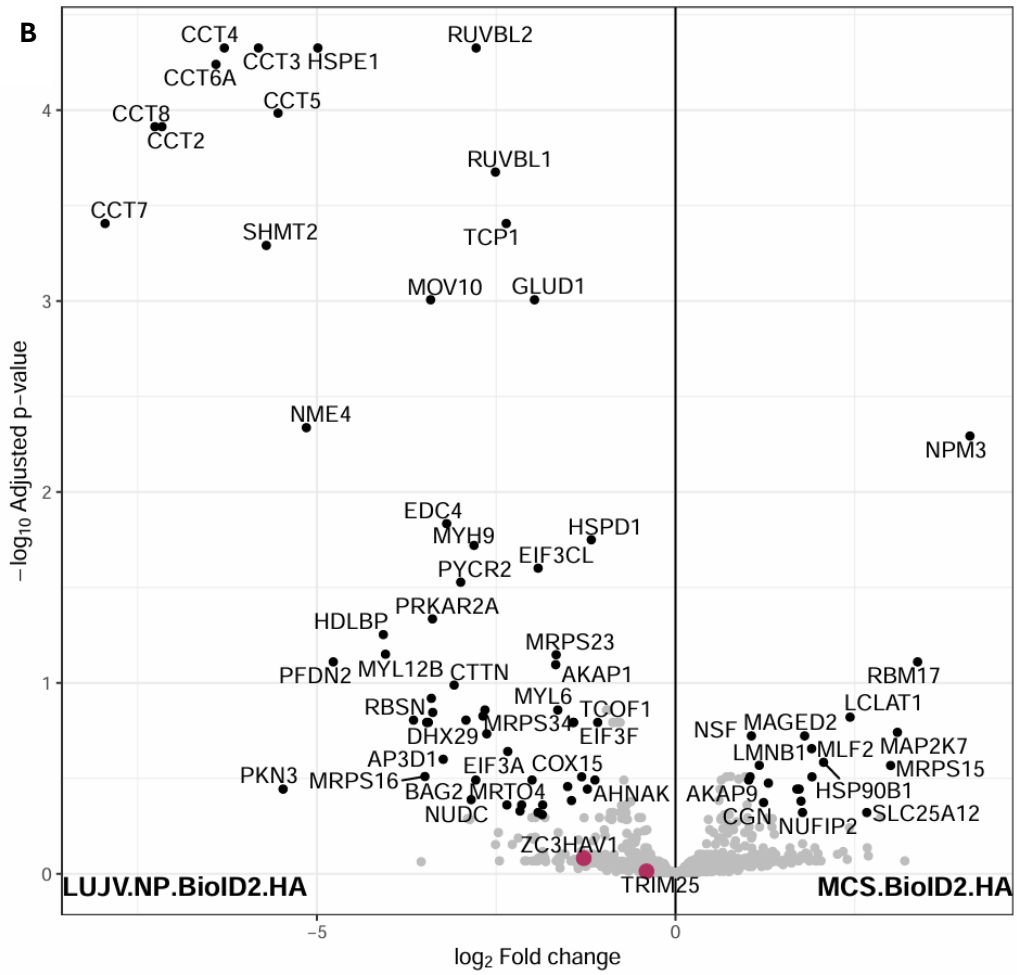


Figure 3.7 Differential expression analysis and results plots represented as a heatmap; a plot showing an overview of expression of all significant (differentially expressed) proteins (rows) in all samples comparing the MCS-BioID2-HA control, LASV-, LUJV-, and MOPV-NP-BioID2-HA with each other across all replicates (columns).

Figures 3.8 and 3.9 in this set are volcano plots comparing arenavirus strains against both control and each other. It is important to note that protein abundance levels are represented in these volcano plots, but that increased abundance may not correlate to increased importance of that identified host protein to the function of arenavirus NP. Proteins with low expression may still have potent effects and be highly impactful upon the host cell and/or virus. Volcano plots are used as visual tools because generally, the minimal comparative experiment spans two conditions with n biological replicates each. In the experiment, for example, three conditions each with three biological replicates were compared. The standard analysis workflow for this involves performing multiple hypothesis corrected two-sample (Student's T-) tests (Young et al., 1987, Krzywinski and Altman, 2014a, Krzywinski and Altman, 2014b). Thus, as the multiple

hypothesis correction is essential to any proteomics experiment, *p-values* can seem to be significant due to chance when thousands of comparisons from the same dataset are being made. By plotting the negative log10 of the corrected *p*-value against the difference in log-space for each protein a volcano plot is produced. The volcano plot y axis show how statistically significant the protein abundance differences are with more statistically significant protein abundances being towards the top of the plot (lower *p*-values). The x axis represents how big the difference in protein abundance is (fold change) in which a positive fold change means the protein is more abundant in the condition on the right of the graph (group A) compared to that on the left (group B). A negative fold change means the protein is more abundant in group A compared to group B and a fold change near zero means there is no significant difference in protein abundance. For example, in Figure 3.8A of Control vs LASV NP, the most statistically significant gene is CCT4 as it is the highest on the plot but the largest difference in protein abundance is NOD2 which has the largest differential fold change in the LASV NP sample when compared to the control sample but as is it a negative fold change it shows that NOD2 abundance is greatly increased in the LASV NP sample compared to its abundance levels in the control sample. In Figures 3.8 and 3.9 ZC3HAV1 (ZAP) and TRIM25 have been highlighted. ZAP and TRIM25 were selected as interacting host factors of arenavirus NP due to their conservation across multiple arenavirus proteomes (LASV, LUJV, etc.), known roles in antiviral responses (e.g. Ebola, Influenza), and experimentally validated interactions (Zheng et al, 2017; Li et al. 2017). Their unexplored role in arenavirus infection warrants further investigation which is the focus of later chapters.


A
CCT8
CCT6A
CCT5
EIF3CL
CCT2
CCT4
CCT3
EIF3F
CCT7
EIF3A
MOV10
HSPE1
GPATCH8
SHMT2
HDLBP
TCOF1
NOD2
EIF3B
MYH9
TCP1
NPM3
SAR1A
ARL1
MRI1
HSP90B1
AHCTF1
PDCL
TTF2
GTPBP6
GTF3C3
COX15
PFDN2
HSD17B10
AATF
NUFIP2
RAB39B
RRP8
MYL6
PRPF4B
CTTN
ACTC1
FASN
SRR
EIF6
DYNC1LI1
AGTPBP1
MRPS23
PRRC2C
EIF4G1
SLC7A5
PPIG
TRIM25
ZC3HAV1
LASV.NP.BioID2.HA
MCS.BioID2.HA
−log10 Adjusted p−value
log2 Fold change
2
1
0
−8
−4
0
4


B
CCT4
CCT3
HSPE1
CCT6A
CCT5
CCT8
CCT2
RUVBL2
RUVBL1
CCT7
SHMT2
TCP1
MOV10
GLUD1
NME4
NPM3
EDC4
MYH9
HSPD1
EIF3CL
PYCR2
PRKAR2A
HDLBP
MRPS23
PFDN2
MYL12B
CTTN
AKAP1
RBM17
MYL6
LCLAT1
RBSN
TCOF1
MRPS34
EIF3F
NSF
MAGED2
DHX29
LMNB1
MLF2
MAP2K7
AP3D1
EIF3A
COX15
MRPS15
PKN3
MRPS16
BAG2
MRTO4
AHNAK
AKAP9
HSP90B1
NUDC
CGN
NUFIP2
SLC25A12
ZC3HAV1
TRIM25
LUJV.NP.BioID2.HA
MCS.BioID2.HA
−log10 Adjusted p−value
log2 Fold change
4
3
2
1
0
−5
0

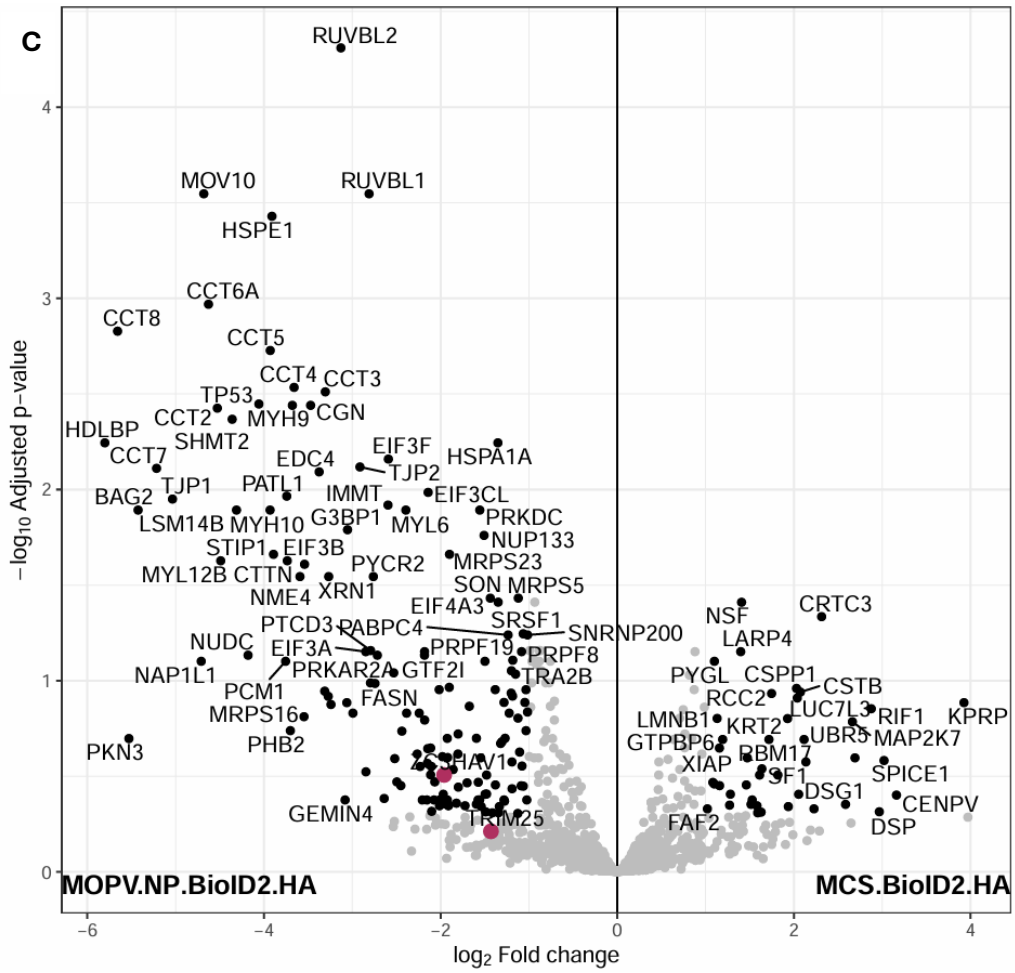
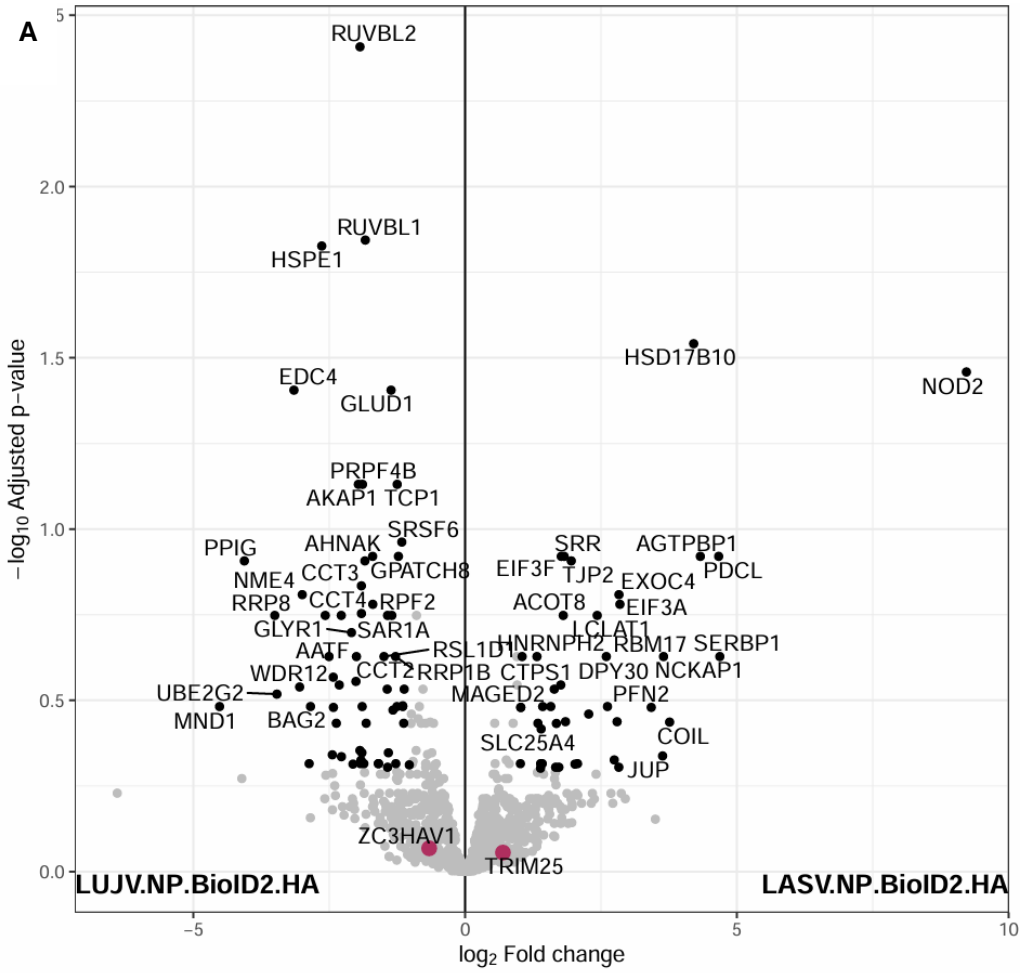
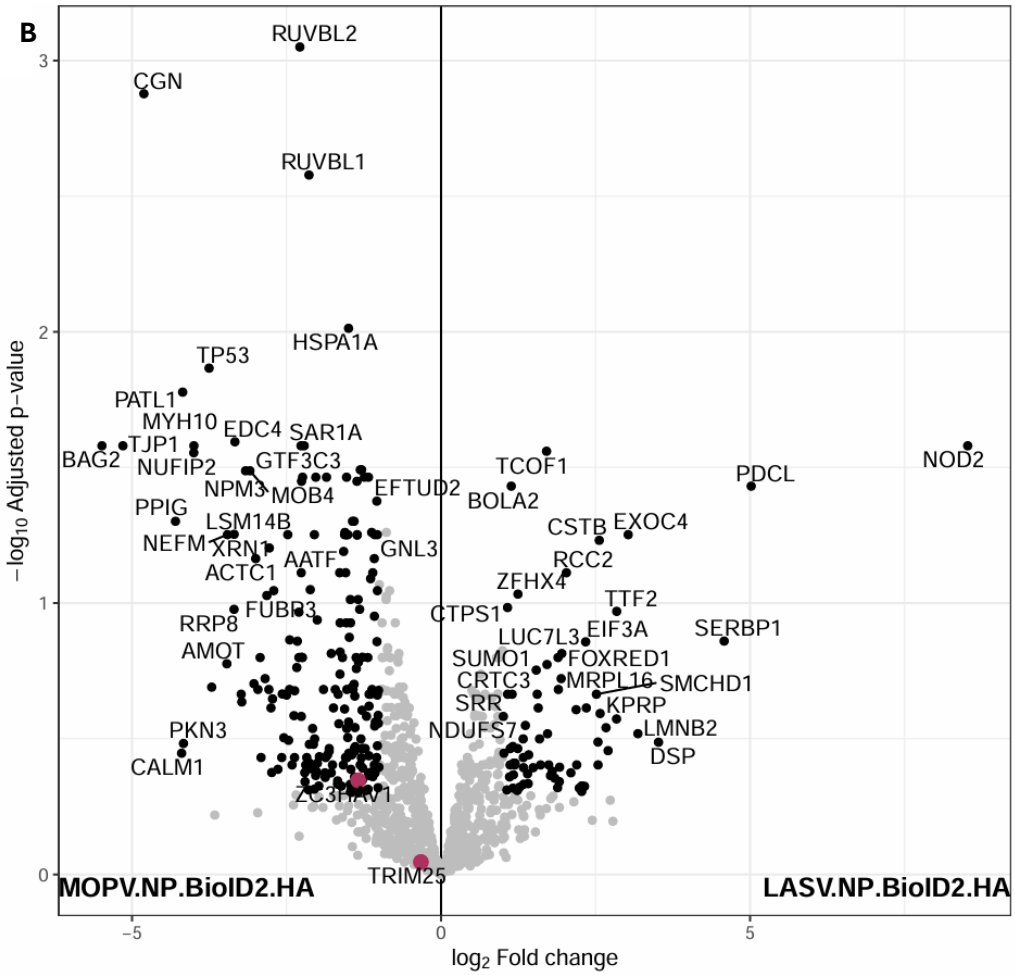


Figure 3.8 Volcano plots of differences from NPs to control with the x-axis as Log2 fold changes, revealing protein expression alterations and Y-axis is the negative logarithm of p-values, signifying statistical significance. The observations are signal intensity measurements from the mass spectrometer and these intensities relate to the amount of protein in each experiment and under each condition. Fold changes are ratios where a value larger than 1 for a protein implies that protein expression was greater after the treatment. It is important to note that ratios are not symmetrical around 1 so it can be difficult to observe both changes in the forwards and backwards directions however when transforming ratios on a log scale, the scale then becomes symmetrical around 0 and it is possible to observe the ratios in terms of positive, negative, or no-change. Plotting values onto a log scale allows for visualisation of small or large magnitude changes. (A) volcano plot comparing control sample vs LASV NP sample B) volcano plot comparing control sample vs LUJV NP sample A) volcano plot comparing control sample vs MOPV NP sample. TRIM25 and ZC3HAV1 (ZAP) are highlighted on the plots displaying the fold change compared between conditions.


A
RUVBL2
RUVBL1
HSPE1
EDC4
GLUD1
HSD17B10
NOD2
PRPF4B
AKAP1
TCP1
SRSF6
PPIG
AHNAK
SRR
AGTPBP1
NME4
CCT3
GPATCH8
EIF3F
TJP2
EXOC4
PDCL
RRP8
CCT4
RPF2
ACOT8
EIF3A
GLYR1
SAR1A
LCLAT1
AATF
RSL1D1
HNRNPH2
RBM17
SERBP1
WDR12
CCT2
RRP1B
CTPS1
DPY30
NCKAP1
UBE2G2
MAGED2
PFN2
MND1
BAG2
SLC25A4
COIL
JUP
ZC3HAV1
TRIM25
LUJV.NP.BioID2.HA
LASV.NP.BioID2.HA
−log10 Adjusted p−value
log2 Fold change
5
2.0
1.5
1.0
0.5
0.0
−5
0
5
10


B
RUVBL2
CGN
RUVBL1
HSPA1A
TP53
PATL1
MYH10
EDC4
SAR1A
TJP1
BAG2
NUFIP2
GTF3C3
NPM3
MOB4
EFTUD2
PPIG
LSM14B
NEFM
XRN1
AATF
GNL3
ACTC1
FUBP3
RRP8
AMOT
PKN3
CALM1
ZC3HAV1
TRIM25
TCOF1
BOLA2
PDCL
NOD2
CSTB
EXOC4
RCC2
ZFHX4
TTF2
CTPS1
LUC7L3
EIF3A
SERBP1
SUMO1
FOXRED1
CRTC3
MRPL16
SMCHD1
SRR
KPRP
NDUFS7
LMNB2
DSP
MOPV.NP.BioID2.HA
LASV.NP.BioID2.HA
−log10 Adjusted p−value
log2 Fold change
3
2
1
0
−5
0
5

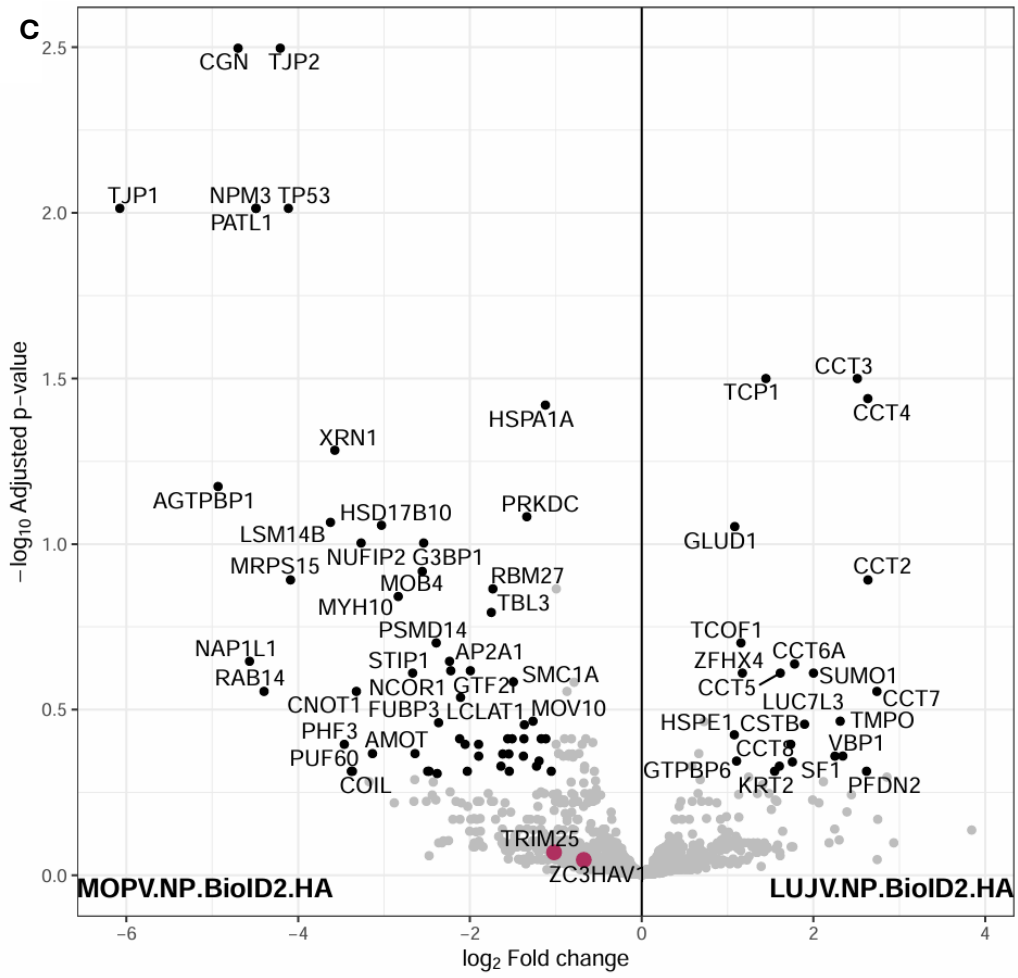


Figure 3.9 Volcano plots of differences between NPs with the x-axis as Log2 fold changes, revealing protein expression alterations and Y-axis is the negative logarithm of p-values, signifying statistical significance. A) volcano plot comparing LASV NP sample vs LUJV NP B) volcano plot comparing LASV NP sample vs MOPV NP C) volcano plot comparing LUJV NP sample vs MOPV NP sample. TRIM25 and ZC3HAV1 (ZAP) are highlighted on the plots displaying the fold change compared between conditions.

3.4.5 Conserved cellular processes targeted by arenavirus NP

The proteins showing conserved interactions with the pathogenic Old World arenavirus NP are of the most interest as these could potentially represent highly conserved aspects of arenavirus biology crucial to their pathogenesis. Similarly, the differences and non-conserved interactions between the non-pathogenic MOPV NP and the highly pathogenic LASV and LUJV could shed light on which mechanisms and pathways are required for arenaviruses to cause disease in humans. STRING was used to build networks from the MS data for each arenavirus NP strain (as shown in figure 3.11) and using these and the data from figure 3.10 shows how many proteins are unique to each strain compared to each other and the control. From this information a clearer picture of the similarities and differences of interest can be constructed.

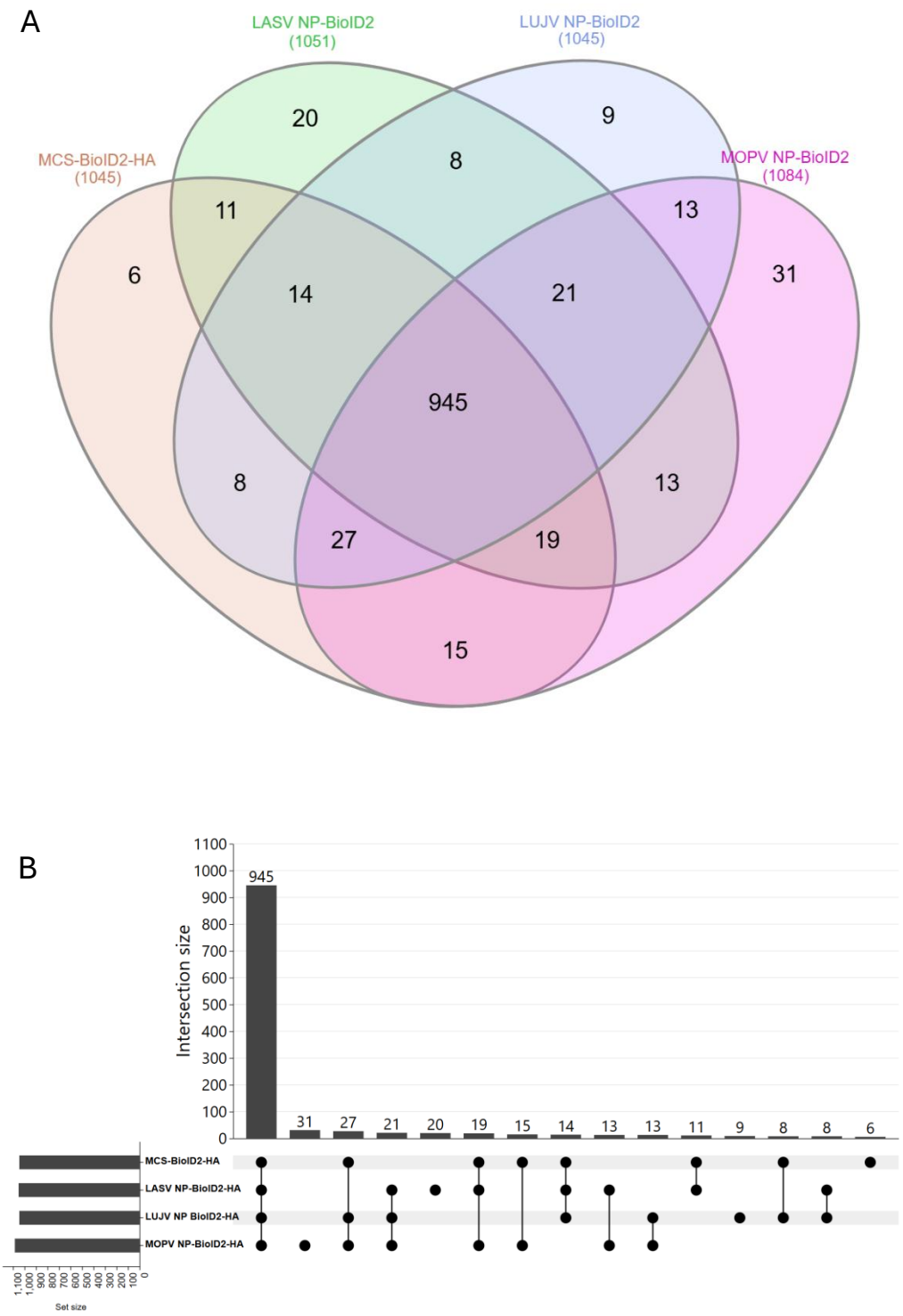
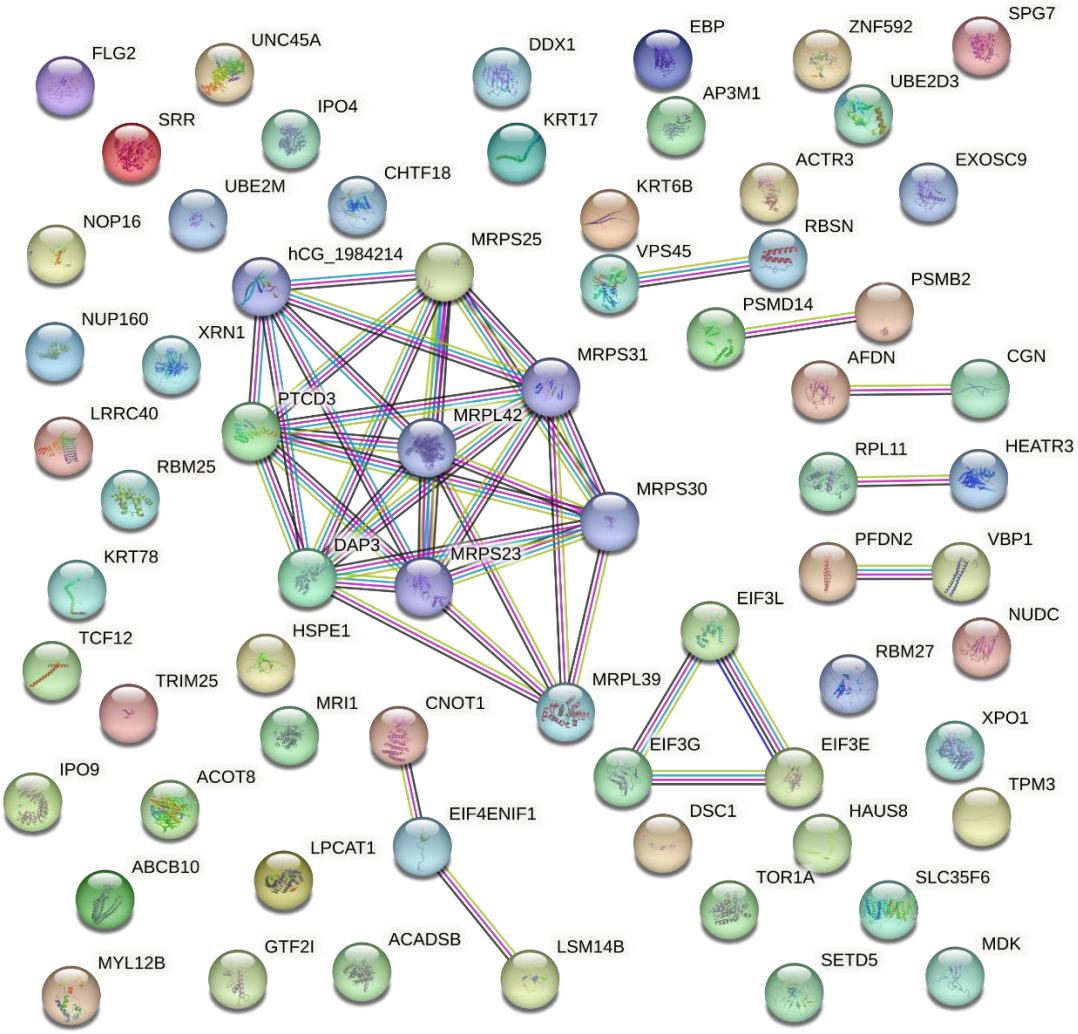
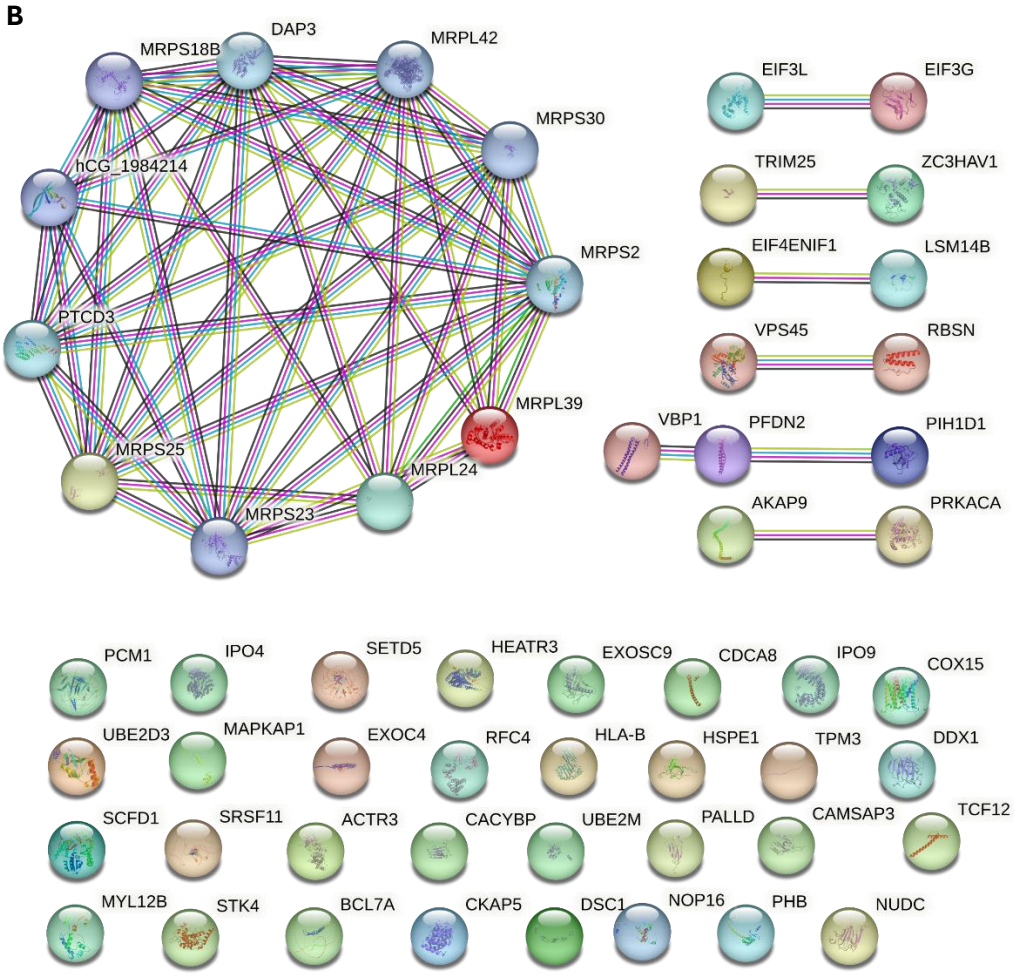


Figure 3.10 (A) Venn diagram showing overlap of protein sets identified from each condition. B) shows the number of proteins in each set that overlap B) shows the percentage of total proteins identified that overlap for each set (Heberle et al., 2015) (B) Upset plot visualising the intersections between our protein sets in the matrix; the columns correspond to the sets and the rows to the intersections. features and UpSet plot. UpSet plots have been shown to perform better than Venn diagrams when using larger numbers of sets and for when you want to show more contextual information about set intersections. The UpSet plot was created using https://www.chiplot.online/ (accessed 17/11/2024)

There is a highly conserved network across all three arenavirus NP strains: hCG_1984214, PTCD3, DAP3, MRPS31. MRPL42, MRPL39, MRPS30, MRPS23, and MRPS25 seen in figure 3.11. These proteins are all heavily involved in mitochondrial translation and have molecular functions in RNA binding and most of these proteins are structural constituents of mitochondrial ribosomes. Mitochondria are known to play vital roles in mediating innate immunity and so it is unsurprising that viruses evolve methods to counteract normal mitochondrial function. There is evidence that arenavirus NP enters the mitochondria of infected cells and affects their morphological integrity (Baggio et al., 2021). The proteins identified are involved in all aspects of mitochondrial translation, initiation, elongation, and termination. It is thought that targeting the mitochondria is done to complement interferon antagonism by NP and also to alter the metabolic state of infected cells.

A


FLG2
UNC45A
DDX1
EBP
ZNF592
SPG7
SRR
IPO4
KRT17
AP3M1
UBE2D3
UBE2M
CHTF18
ACTR3
EXOSC9
NOP16
KRT6B
RBSN
hCG_1984214
MRPS25
VPS45
PSMB2
NUP160
XRN1
PSMD14
MRPS31
AFDN
CGN
LRRC40
PTCD3
MRPL42
RPL11
HEATR3
RBM25
MRPS30
DAP3
MRPS23
PFDN2
VBP1
KRT78
EIF3L
NUDC
TCF12
HSPE1
RBM27
TRIM25
MRPL39
MRI1
CNOT1
XPO1
EIF3G
EIF3E
IPO9
ACOT8
TPM3
EIF4ENIF1
DSC1
HAUS8
LPCAT1
ABCB10
TOR1A
SLC35F6
GTF2I
ACADSB
LSM14B
MDK
MYL12B
SETD5


B
DAP3
MRPS18B
MRPL42
MRPS30
hCG_1984214
MRPS2
PTCD3
MRPL39
MRPS25
MRPL24
MRPS23
EIF3L
EIF3G
TRIM25
ZC3HAV1
EIF4ENIF1
LSM14B
VPS45
RBSN
VBP1
PFDN2
PIH1D1
AKAP9
PRKACA
PCM1
IPO4
SETD5
HEATR3
EXOSC9
CDCA8
IPO9
COX15
UBE2D3
MAPKAP1
EXOC4
RFC4
HLA-B
HSPE1
TPM3
DDX1
SCFD1
SRSF11
ACTR3
CACYBP
UBE2M
PALLD
CAMSAP3
TCF12
MYL12B
STK4
BCL7A
CKAP5
DSC1
NOP16
PHB
NUDC

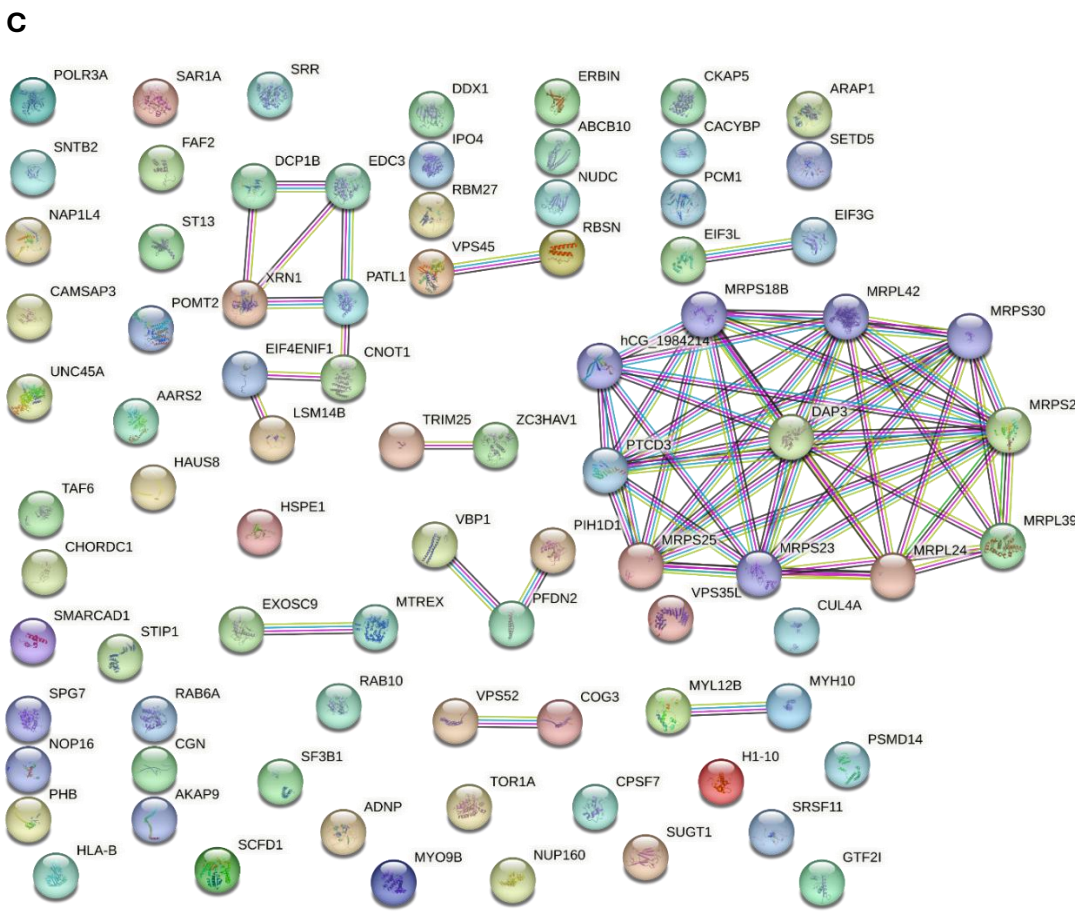
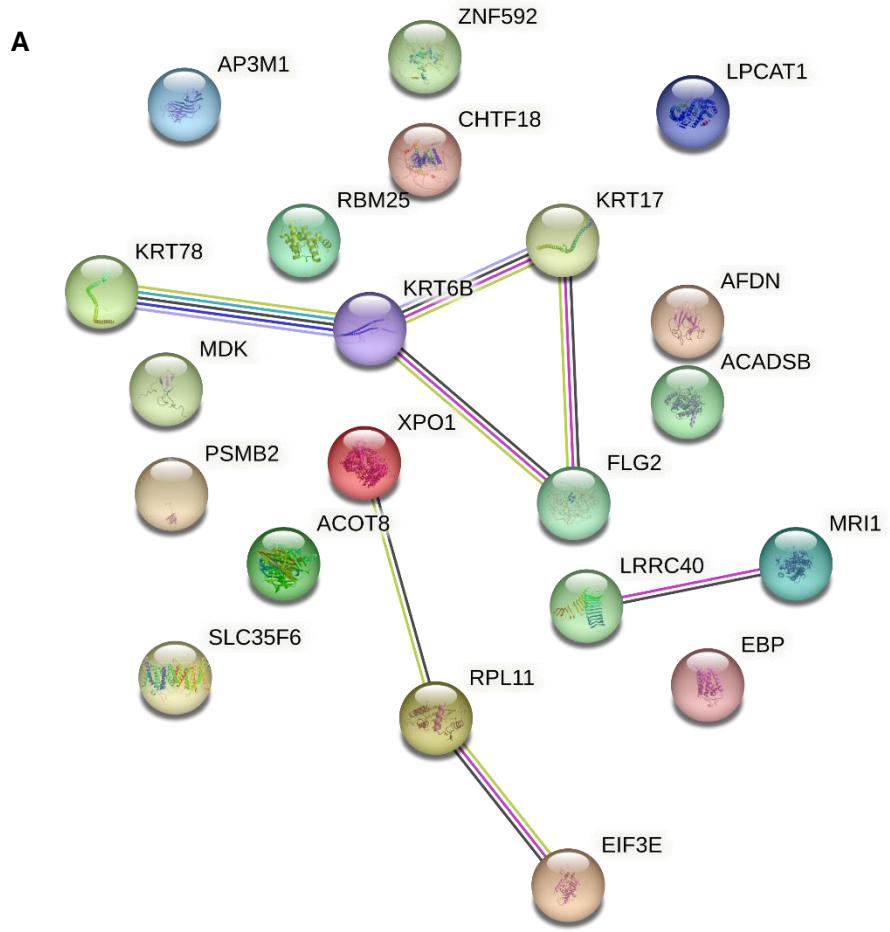


Figure 3.11 STRING proteome network of host interacting partners of; A) LASV NP B) LUJV NP and C) MOPV NP. Identified using data acquired by MS by Liverpool. All of these interacting proteins were present in the MOPV NP samples but not present in the Control samples Every node in the network represents a protein that interacts with MOPV NP based upon the BioID2 protocol and subsequent mass spectrometry and downstream bioinformatics data processing. Settings for the STRING network were set to the highest stringency feasible and only proteins that have been mentioned together in all mined literature and then proven to interact biochemically through experimentation such as co-purification, co-crystallisation, or Yeast2hybrid have their interactions displayed as connecting lines. Associations are meant to be specific and meaningful, and proteins are shown to be jointly contributing to shared function(s); this does not necessarily mean that these proteins are physically binding to each other and directly interacting. Network nodes represent proteins and splice isoforms, or post-translational modifications are collapsed so that each node represents all the proteins produced by a single, protein-coding gene locus. Edges represent protein-protein interactions and are colour coded according to type of interaction. Light blue denotes a known interaction from a curated database and purple is a known interaction which has been experimentally determined. Predicted interactions are either green for gene neighbourhood, red for gene fusions, or dark blue for gene co-occurrence. Other edges include light green for text-mined data, black for co-expression, and grey-blue for protein homology. In terms of nodes, those that are empty are for proteins of unknown 3D structure whereas those that are filled have some or all of their 3D structure known or predicted.

3.4.6 Differences between arenavirus NP strain host interactomes

Investigating the differences between which proteins and pathways are targeted by the different strains of arenavirus NP is just as important as understanding the similarities, as these differences may hold crucial information about what specifically causes increased pathogenicity. Figure 3.12 – 3.14 show STRING proteome networks built from proteins completely unique to each arenavirus NP. These also include information about the enriched biological pathways associated with each network.

LASV NP has 21 proteins uniquely interacting with it that are not present in the control samples, LUVJ NP samples, or MOPV NP samples (Figure 3.12). The rest of the proteins appear to have varied roles in promoting and maintaining cell health and homeostasis listed in table 3.2.


A
ZNF592
AP3M1
LPCAT1
CHTF18
RBM25
KRT17
KRT78
KRT6B
AFDN
MDK
ACADSB
XPO1
PSMB2
FLG2
ACOT8
MRI1
LRRC40
SLC35F6
EBP
RPL11
EIF3E

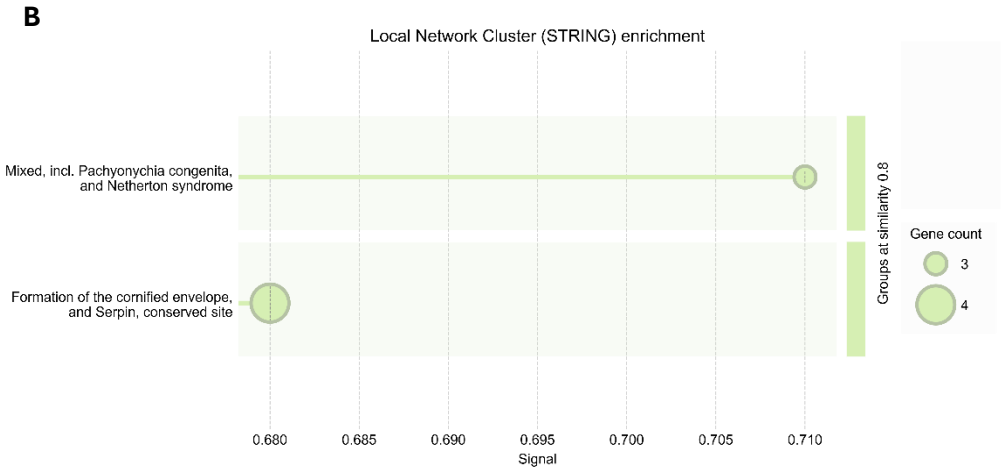

B
Local Network Cluster (STRING) enrichment
Mixed, incl. Pachyonychia congenita, and Netherton syndrome
Formation of the cornified envelope, and Serpin, conserved site
Groups at similarity 0.8
Gene count
3
4
0.680
0.685
0.690
0.695
0.700
0.705
0.710
Signal

Figure 3.12 Enriched network for unique interactors of LASV NP. A) STRING proteome network of totally unique host interacting partners of LASV NP identified using data acquired by MS by Liverpool. Every node in the network represents a protein that interacts with LASV NP based upon the BioID2 protocol and subsequent mass spectrometry and downstream bioinformatics data processing. Settings for the STRING network were set to the highest stringency feasible and only proteins that have been mentioned together in all mined literature and then proven to interact biochemically through experimentation such as co-purification, co-crystallisation, or Yeast2hybrid have their interactions displayed as connecting lines. Associations are meant to be specific and meaningful, and proteins are shown to be jointly contributing to shared function(s); this does not necessarily mean that these proteins are physically binding to each other and directly interacting. Network nodes represent proteins and splice isoforms, or post-translational modifications are collapsed so that each node represents all the proteins produced by a single, protein-coding gene locus. Edges represent protein-protein interactions and are colour coded according to type of interaction. Light blue denotes a known interaction from a curated database and purple is a known interaction which has been experimentally determined. Predicted interactions are either green for gene neighbourhood, red for gene fusions, or dark blue for gene co-occurrence. Other edges include light green for text-mined data, black for co-expression, and grey-blue for protein homology. In terms of nodes, those that are empty are for proteins of unknown 3D structure whereas those that are filled have some or all of their 3D structure known or predicted. B) Visualisation of functional enrichment for the protein network generated by STRING.

MRI1	Methylthioribose-1-phosphate isomerase; Catalyzes the interconversion of methylthioribose-1-phosphate (MTR-1-P) into methylthioribulose-1-phosphate (MTRu-1-P). Independently from catalytic activity, promotes cell invasion in response to constitutive RhoA activation by promoting FAK tyrosine phosphorylation and stress fibre turnover.
ACOT8	Acyl-coenzyme A thioesterase 8; Acyl-coenzyme A (acyl-CoA) thioesterases are a group of enzymes that catalyse the hydrolysis of acyl-CoAs to the free fatty acid and coenzyme A (CoASH), providing the potential to regulate intracellular levels of acyl-CoAs, free fatty acids and CoASH.
EIF3E	Eukaryotic translation initiation factor 3 subunit E; Component of the eukaryotic translation initiation factor 3 (eIF-3) complex, which is required for several steps in the initiation of protein synthesis.
RBM25	RNA-binding protein 25; RNA-binding protein that acts as a regulator of alternative pre-mRNA splicing. Involved in apoptotic cell death through the regulation of the apoptotic factor BCL2L1 isoform expression.
CHTF18	Chromosome transmission fidelity protein 18 homolog; Chromosome cohesion factor involved in sister chromatid cohesion and fidelity of chromosome transmission.

LPCAT1	Lysophosphatidylcholine acyltransferase 1; Possesses both acyltransferase and acetyltransferase activities.
SLC35F6	Solute carrier family 35 member F6; Involved in the maintenance of mitochondrial membrane potential in pancreatic ductal adenocarcinoma (PDAC) cells.
AP3M1	AP-3 complex subunit mu-1; Part of the AP-3 complex, an adaptor-related complex which is not clathrin-associated. The complex is associated with the Golgi region as well as more peripheral structures. It facilitates the budding of vesicles from the Golgi membrane and may be directly involved in trafficking to lysosomes.
AFDN	Afadin; Belongs to an adhesion system, probably together with the E- cadherin-catenin system, which plays a role in the organization of homotypic, interneuronal and heterotypic cell-cell adherens junctions.
ACADSB	Short/branched chain specific acyl-CoA dehydrogenase, mitochondrial; may play a role in controlling the metabolic flux of valproic acid in the development of toxicity of this agent
LRRC40	Leucine-rich repeat-containing protein 40
PSMB2	Proteasome subunit beta type-2; Component of the 20S core proteasome complex involved in the proteolytic degradation of most intracellular proteins. This complex plays numerous essential roles within the cell by associating with different regulatory particles.
FLG2	Filaggrin-2; Essential for normal cell-cell adhesion in the cornified cell layers. Important for proper integrity and mechanical strength of the stratum corneum of the epidermis
XPO1	Exportin-1; Mediates the nuclear export of cellular proteins (cargos) bearing a leucine-rich nuclear export signal (NES) and of RNAs.

MDK	Midkine; Secreted protein that functions as cytokine and growth factor and mediates its signal through cell-surface proteoglycan and non- proteoglycan receptors.
EBP	3-beta-hydroxysteroid-Delta(8),Delta(7)-isomerase; Catalyses the conversion of Delta(8)-sterols to their corresponding Delta(7)-isomers.
ZNF592	Zinc finger protein 592; May be involved in transcriptional regulation
RPL11	60S ribosomal protein L11; Component of the ribosome, a large ribonucleoprotein complex responsible for the synthesis of proteins in the cell.

Table 3.2 Unique host proteins interacting with LASV NP and their functions

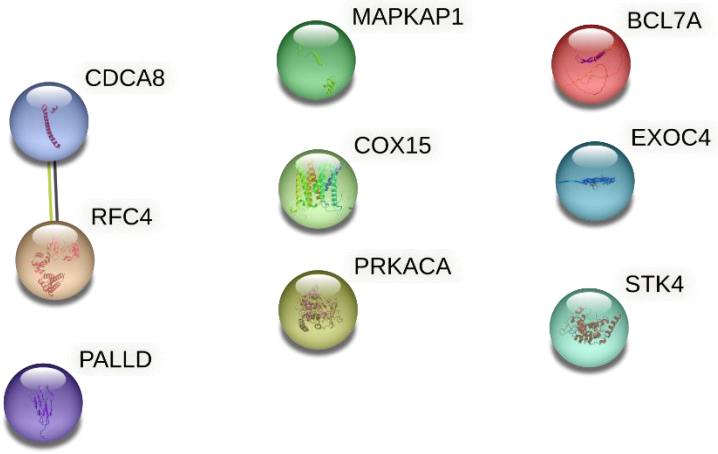


Figure 3.13 STRING proteome network of totally unique host interacting partners of LUJV NP identified using data acquired by MS by Liverpool.

COX15	Cytochrome c oxidase assembly protein COX15 homologue; May be involved in the biosynthesis of heme A.
EXOC4	Exocyst complex component 4; Component of the exocyst complex involved in the docking of exocytic vesicles with fusion sites on the plasma membrane.
MAPKAP1	Target of rapamycin complex 2 subunit MAPKAP1; Subunit of mTORC2, which regulates cell growth and survival in response to hormonal signals.
PRKACA	cAMP-dependent protein kinase catalytic subunit alpha; Phosphorylates a large number of substrates in the cytoplasm and the nucleus.
STK4	Serine/threonine-protein kinase 4 18kDa subunit; Stress-activated, pro-apoptotic kinase which, following caspase-cleavage, enters the nucleus and induces chromatin condensation followed by internucleosomal DNA fragmentation. Key component of the Hippo signaling pathway which plays a pivotal role in organ size control and tumor suppression by restricting proliferation and promoting apoptosis.
CDCA8	Borealin; Component of the chromosomal passenger complex (CPC), a complex that acts as a key regulator of mitosis.

RFC4	Replication factor C subunit 4; This subunit may be involved in the elongation of the multi-primed DNA template.
PALLD	Palladin; Cytoskeletal protein required for organization of normal actin cytoskeleton. Roles in establishing cell morphology, motility, cell adhesion and cell-extracellular matrix interactions in a variety of cell types. May function as a scaffolding molecule with the potential to influence both actin polymerization and the assembly of existing actin filaments into higher-order arrays.
BCL7A	B-cell CLL/lymphoma 7 protein family member A; BAF chromatin remodelling complex subunit BCL7A.

Table 3.3 Unique host proteins interacting with LUJV NP and their functions

LUJV NP only has 9 proteins uniquely interacting with it and no functionally enriched pathways or biological functions as yet discovered. Most of these proteins have roles in either DNA processing and mitosis, cell structure via regulation of actin cytoskeleton, or modulating cellular stress factors to keep the cell presenting as healthy and promote viral replication as seen in table 3.3.

MOPV NP has 31 uniquely interacting proteins identified with several functionally enriched processes and molecular functions (figure 3.14). Most of the proteins are involved in cellular component organisation or biogenesis and have roles in ADP and RNA binding as well as heterocyclic and organic cycling compound binding as seen in table 3.4. This can also be seen from the functionally enriched biological processes identified by STRING analysis; many of the identified proteins have roles in methyl-guanosine de-capping. Capping of mRNA is essential for efficient gene expression and cell viability and that MOPV NP interacts with so many proteins involved in this function indicates this may be a major target for the virus although this is a significant difference to what can be seen in the processes that LASV and LUJV NP appear to target.

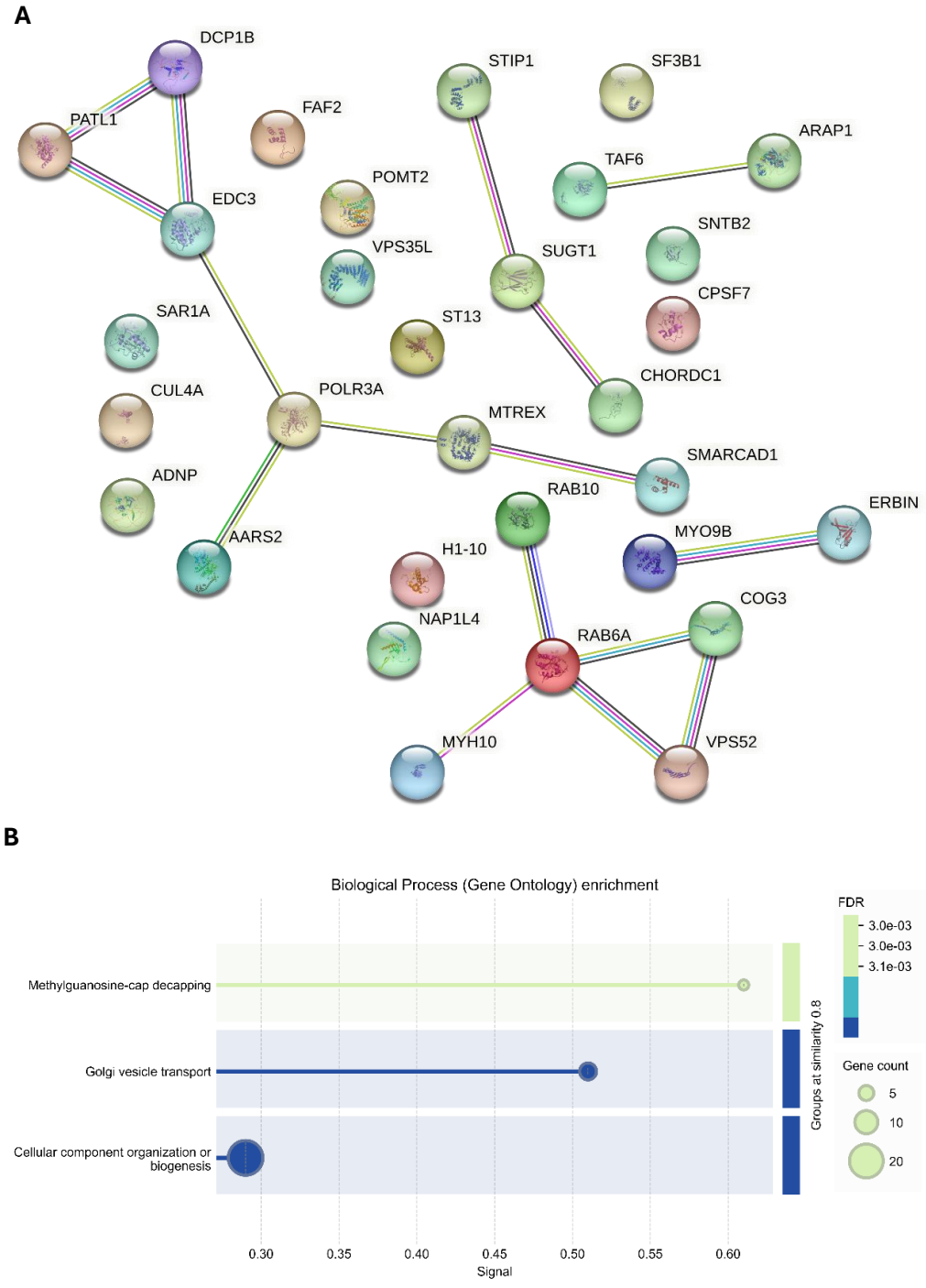


Figure 3.14 STRING proteome network of totally unique host interacting partners of MOPV NP. identified using data acquired by MS by Liverpool. B) Visualisation of functional enrichment for the protein network generated by STRING.

ST13	Hsc70-interacting protein; One HIP oligomer binds the ATPase domains of at least two HSC70 molecules dependent on activation of the HSC70 ATPase by HSP40.
MTREX	Exosome RNA helicase MTR4; Catalyzes the ATP-dependent unwinding of RNA duplexes with a single-stranded 3' RNA extension. Central subunit of many protein complexes, namely TRAMP-like, nuclear exosome targeting (NEXT) and poly(A) tail exosome targeting (PAXT).
AARS2	Alanine-tRNA ligase, mitochondrial; Catalyzes the attachment of alanine to tRNA(Ala) in a two- step reaction: alanine is first activated by ATP to form Ala-AMP and then transferred to the acceptor end of tRNA(Ala).
VPS35L	VPS35 endosomal protein sorting factor-like; Acts as component of the retriever complex.
COG3	Conserved oligomeric Golgi complex subunit 3; Involved in ER-Golgi transport.
POMT2	Protein O-mannosyl-transferase 2; Transfers mannosyl residues to the hydroxyl group of serine or threonine residues. Coexpression of both POMT1 and POMT2 is necessary for enzyme activity, expression of either POMT1 or POMT2 alone is insufficient.
FAF2	FAS-associated factor 2; Plays an important role in endoplasmic reticulum-associated degradation (ERAD) that mediates ubiquitin-dependent degradation of misfolded endoplasmic reticulum proteins.
RAB10	Ras-related protein Rab-10; The small GTPases Rab are key regulators of intracellular membrane trafficking, from the formation of transport vesicles to their fusion with membranes.
DCP1B	mRNA-decapping enzyme 1B; May play a role in the degradation of mRNAs, both in normal mRNA turnover and in nonsense-mediated mRNA decay.
PATL1	Protein PAT1 homolog 1; RNA-binding protein involved in deadenylation-dependent decapping of mRNAs, leading to the

	degradation of mRNAs. Acts as a scaffold protein that connects deadenylation and decapping machinery.
RAB6A	Ras-related protein Rab-6A; Protein transport. Regulator of membrane traffic from the Golgi apparatus towards the endoplasmic reticulum (ER). Ha
CHORDC1	Cysteine and histidine-rich domain-containing protein 1; Regulates centrosome duplication, probably by inhibiting the kinase activity of ROCK2. Proposed to act as co-chaperone for HSP90. May play a role in the regulation of NOD1 via a HSP90 chaperone complex.
H1-10	Histone H1x; Histones H1 are necessary for the condensation of nucleosome chains into higher-order structures.
SF3B1	Splicing factor 3B subunit 1; Involved in pre-mRNA splicing as a component of the splicing factor SF3B complex.
SNTB2	Beta-2-syntrophin; Adapter protein that binds to and probably organizes the subcellular localization of a variety of membrane proteins. May link various receptors to the actin cytoskeleton and the dystrophin glycoprotein complex. May play a role in the regulation of secretory granules via its interaction with PTPRN.
CPSF7	Cleavage and polyadenylation specificity factor subunit 7; Component of the cleavage factor Im (CFIm) complex that functions as an activator of the pre-mRNA 3'-end cleavage and polyadenylation processing required for the maturation of pre-mRNA into functional mRNAs.
STIP1	Stress-induced-phosphoprotein 1; Acts as a co-chaperone for HSP90AA1. Mediates the association of the molecular chaperones HSPA8/HSC70 and HSP90.
SMARCAD1	ATP-dependent helicase 1; DNA helicase that possesses intrinsic ATP-dependent nucleosome-remodeling activity and is both required for DNA repair and heterochromatin organization.
MYH10	Myosin-10; Cellular myosin that appears to play a role in cytokinesis, cell shape, and specialized functions such as secretion and capping.

ADNP	Activity-dependent neuroprotector homeobox protein; Potential transcription factor. May mediate some of the neuroprotective peptide VIP-associated effects involving normal growth and cancer proliferation.
POLR3A	DNA-directed RNA polymerase III subunit RPC1; DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates.
SAR1A	GTP-binding protein SAR1a; Involved in transport from the endoplasmic reticulum to the Golgi apparatus. Required to maintain SEC16A localization at discrete locations on the ER membrane perhaps by preventing its dissociation.
CUL4A	Cullin-4A; Core component of multiple cullin-RING-based E3 ubiquitin- protein ligase complexes which mediate the ubiquitination of target proteins. As a scaffold protein may contribute to catalysis through positioning of the substrate and the ubiquitin-conjugating enzyme.
SUGT1	Protein SGT1 homolog; May play a role in ubiquitination and subsequent proteasomal degradation of target proteins.
NAP1L4	Nucleosome assembly protein 1-like 4; Acts as histone chaperone in nucleosome assembly.
ARAP1	Arf-GAP with Rho-GAP domain, ANK repeat and PH domain-containing protein 1; Phosphatidylinositol 3,4,5-trisphosphate-dependent GTPase- activating protein that modulates actin cytoskeleton remodeling by regulating ARF and RHO family members.
TAF6	Transcription initiation factor TFIID subunit 6; TIIFD is multimeric protein complex that plays a central role in mediating promoter responses to various activators and repressors.
VPS52	Vacuolar protein sorting-associated protein 52 homolog; Acts as component of the GARP complex that is involved in retrograde transport from early and late endosomes to the trans-Golgi network (TGN).

ERBIN	Erbin; Acts as an adapter for the receptor ERBB2, in epithelia. By binding the unphosphorylated 'Tyr-1248' of receptor ERBB2, it may contribute to stabilize this unphosphorylated state. Inhibits NOD2-dependent NF-kappa-B signaling and proinflammatory cytokine secretion.
MYO9B	Unconventional myosin-IXb; Myosins are actin-based motor molecules with ATPase activity. Unconventional myosins serve in intracellular movements. Binds actin with high affinity both in the absence and presence of ATP and its mechanochemical activity is inhibited by calcium ions.
EDC3	Enhancer of mRNA-decapping protein 3; Binds single-stranded RNA. Involved in the process of mRNA degradation and in the positive regulation of mRNA decapping. May play a role in spermiogenesis and oogenesis.

Table 3.4 Unique host proteins interacting with MOPV NP and their functions

Important to include in this analysis is the ranked data provided by MS which provides information regarding relative protein abundance in our samples compared to each other. This will allow comparison of proteins present in different strains of NP with statistically significant changes in abundance which may indicate importance to the pathogenesis of that viral strain. Figure 3.8 and 3.9 (A-C) identifies many proteins of potential interest. When comparing LASV NP to MOPV NP, NOD2 immediately stands out as the protein with the biggest abundance increase in LASV as well as being the most statistically significant. This protein forms a multiprotein complex with NOD1 to create the NODosome which is crucial to interferon and NF-κB signalling following viral infection (Coutermarsh-Ott et al., 2016). It is possible that as MOPV is not human pathogenic that this pathway does not need to be stimulated to prevent infection hence the abundance difference however we see a similar abundance difference in Figure 3.12A between LASV and LUJV where NOD2 abundance is greatly increased in LASV. Despite only one known LUJV outbreak historically occurring, an 80%

fatality rate indicates that LUJV is much more pathogenic that LASV and perhaps is able to subvert the NOD2 immune stimulated pathways or modulate NOD2 to prevent this aspect of the host immune response therefore allowing much more uncontrolled viral replication and higher fatality rate.

Combining the information provided by non-quantitative data of unique interacting partners with different NP strains and the quantitative ranked data of protein abundance changes between strains allows for a much clearer picture of what may be behind arenavirus pathogenicity to be built.

3.5 Discussion

There has been relatively little research into arenavirus protein proteomes, and the majority of this research has focused on the arenavirus polymerase. So far, the only the OW group LCMV and NW group JUNV have any published research into the NP proteomes (King et al., 2017a, Iwasaki et al., 2018). The research by (King et al., 2017a) involved infection of cells with either LCMV strain Armstrong 53b or JUNV strain Candid# #1 and subsequent immunoprecipitation from cell lysates using an NP-specific monoclonal antibody. These immune complexes were then separated by SDS-PAGE, stained with Coomassie blue for NP visualisation, digested by trypsin and extracted for liquid chromatography-tandem mass spectrometry (LC-MS/MS). A total of 509 human proteins were identified as associating with JUNV NP and a total of 348 identified as associating with LCMV NP. Comparisons of the two groups of proteins showed that 275 were conserved interactors between both JUNV and LCMV NP (King et al., 2017b).

Analysis of the published research therefore identified a gap in knowledge. Whilst LCMV NP had available proteomic data, no other OW group mammarenavirus NP had been investigated, nor had any non-pathogenic

strains been analysed for pathogenicity comparisons. That there has been little research into arenavirus protein host interactomes also presents a challenge in selecting which proteins identified by the mass spectrometry to further investigate. Limited funding and time mean that careful selection of suitable proteins is crucial. The analysis identifying protein networks using STRING allowed for easy visualisation of subsets of proteins interacting with each other in mini networks related to molecular and biological function and proven by co-expression experiments. STRING analysis of the interaction between nodes also provided useful information about each protein and any relevant literature.

In this chapter, a total of 187 host proteins were initially identified as interacting partners of Mopeia mammarenavirus (MOPV) with further mass spectrometry analysis using label-free quantification methods to generate ranked lists validating the MOPV NP data and generating data for the host interactomes of LASV and LUJV NP. Through cross-referencing the proteomes for MOPV, LASV, and LUVJ NP we get the following proteins that are conserved:

DAP3, VBP1, MRPS25, PTCD3, LSM14B, SETD5, EIF4, ENIF1, VPS45, MRPL39, EIF3L, NUDC, NOP16, MYL12B, EIF3G, EXOSC9, TRIM25, IPO4, DDX1, RBSN, HSPE1, PFDN2, and ZC3HAV1 (ZAP). This still presents a large set of proteins to investigate. Through comparing this list of proteins with interactomes in published literature for LCMV and JUNV NP (which identified 275 conserved potential interactors of NP (King et al., 2017b)), we find that just three proteins are conserved; TRIM25, ZAP, and DDX1.

Only Zinc antiviral protein (ZAP) and tripartite motif 25 (TRIM25) have been selected as interacting host factors of arenavirus NP warranting further investigation. ZAP and TRIM25 were selected as potential candidates for further research for a number of reasons; both of these proteins been identified in the proteomes of LASV, LUJV, MOPV, LCMV, and JUNV NP, that both proteins are conserved across all arenavirus NPs for which proteomes have been identified indicates that they play a significant role during

infection. Both ZAP and TRIM25 have also been implicated in the antiviral response for other important human pathogenic viruses such as Ebola virus and Influenza virus and more (Galão et al., 2022, Zhu et al., 2011, Tang et al., 2017). Importantly, ZAP and TRIM25 have been experimentally determined as having known interactions as well as being successfully co-expressed (Zheng et al., 2017, Li et al., 2017). These factors create a strong foundation to begin investigating the role that these proteins may play in the context of arenavirus infection as this is an area in which no research has been done regarding these proteins. Another protein that would be a good candidate for investigation if the time and resources allowed would be DDX1 (ATP-dependent RNA helicase) as it was the only other protein identified in all of the arenavirus NP proteomes, and this high level of conservation indicates that it plays a significant role during both NW and OW group arenavirus infection. This protein is part of the DEAD box family characterised by the conserved motif Asp-Glu-Ala-Asp (DEAD), which have roles in many cellular processes such as translation initiation, nuclear and mitochondrial splicing, and ribosome assembly. DDX1 has been shown to be part of a cap-binding complex involved in activating mRNA translation (Chen et al., 2002). Further investigation of DDX1 could provide deeper understanding of the role of arenavirus NP in regulating mRNA translation within infected cells.

There have been a number of limitations that have occurred throughout the investigation of the arenavirus NP host interactome. One such limiting factor was only sending off one biological replicate of the control sample and MOPV NP sample to Birmingham for analysis. Although this limitation was overcome during the subsequent BioID2-MS experiment which had three biological replicates per sample and had samples for each of the arenavirus strain NPs of interest. However, it may be important to note that the time between sample preparation and actual running of the MS experiment by Liverpool took nearly three months to complete due to staffing and personnel issues at the University of Liverpool Centre for

Proteome Research. Without knowing exactly how long the samples took to be processed once delivered, whether the temperature within the packaging changed significantly, and how quickly they were stored at -80°C after delivery coupled with the potential impact of long-term storage, we cannot be sure that degradation did not occur. If there was degradation to samples in the interim between preparation and MS this could result in spurious interactions being identified, or interactions being missed entirely.

Much of the analysis of MS data has involved looking at host proteins that are not present in the control and only present in NP samples. The ranked data provided by the quantitative MS performed by Liverpool has not yet been fully analysed and still holds a wealth of proteomic information which will help to further understand the nature of how NP interacts with host proteins. Deeper analysis of this data requires use of software tools such as Genome Set Enrichment Analysis (GSEA) and DAVID NIH to functionally enrich genomes and highlight molecular pathways and functions that are statistically significantly enriched or upregulated. This information should allow for the differences between arenavirus NP strains to be studied in more detail and become better understood some software packages such as STRING and Cytoscape allow for druggable target networks to be created, and this could be a convenient next step in the search for potential therapeutics.

An improvement that could be made to the methodology is repeating the BioID2 protocol in the context of infection. Ideally using live LASV infection requiring collaboration with a facility with the means to work on live LASV. BioID2 in the context of infection would provide much more insight into the true proteome of LASV, the protocol using NP overexpression could be limited in that it induces a different response in the cell. This could mean differences in the levels of IFN-response and in turn affecting which ISGs have their expression induced and to what extent. An experiment in which NP is over-expressed in cells that have been pre-treated with IFN and had IFN levels maintained through-out the transfection to induce ISGs could be

a better way to investigate a host immune antiviral response within cells. Live MOPV infection is a useful tool that allows arenavirus infection to be studied in CL2 facilities but as it is non-human-pathogenic it also does not allow for a truly accurate host immune response and NP proteome to be captured. Recombinant LASV in which the virus has been modified for safe use in CL2 or even CL3 facilities could be a very useful tool for inducing an accurate host immune response to study NP proteomes in a context that could unveil previously unknown interacting host proteins crucial for the pathogenicity of LASV.

Figuring out the specific interactions of ZAP and TRIM25 with arenavirus NP and how they affect arenavirus replication during infection is a crucial step towards understanding the NP interactome as whole. To validate these interaction partners identified by mass spectrometry, investigation using co-immunoprecipitation and immunofluorescence studies to find protein localisation and subsequent co-localisation studies have been performed to try to understand where in the cell ZAP and TRIM25 are acting to create an antiviral environment. Modulating the level of ZAP and TRIM25 expression within live MOPV infected cells by either overexpression, siRNA knockdown, CRISPR-Cas9 knockout and RT-qPCR analysis of the effects on MOPV viral replication has also been performed and the results and discussion of these antiviral proteins is the subject of chapter four.

Chapter 4

TRIM25 and ZAP are Host Proteins with Key Roles in the Antiviral Defence Against Arenaviruses

4 TRIM25 and ZAP are host proteins with key roles in the antiviral defence against arenaviruses

4.1 Introduction

In the previous chapter, zinc antiviral protein (ZAP) and tripartite motif-containing protein 25 (TRIM25) were identified through BioID2 proximity-labelling methodology and mass spectrometry as host proteins that interact with arenavirus NP. These proteins have been implicated in the host immune response to a range of different viruses such as alphaviruses (Bick et al., 2003), retroviruses such as HIV-1 (Zhu et al., 2011), and filoviruses such as EBOV (Galão et al., 2022). Susceptibility of viruses to restriction by ZAP does not appear to be dependent on belonging to the same viral family as shown by herpes simplex virus type 1 (HSV-1), Zika virus (ZIKV), yellow fever virus (YFV) and dengue virus (DENV) which grow normally in the presence of ZAP (Bick et al., 2003, Chiu et al., 2018). Another example of this is two viruses from the Picornaviridae family behave which differently in the presence of ZAP; coxsackievirus B3 (Li et al., 2015) is inhibited by ZAP, but poliovirus remains uninhibited by ZAP (Bick et al., 2003).

To date there have been no investigations into the importance of ZAP and TRIM25 during infection with arenaviruses.

There are several ways in which cells sense and control infections by viruses. The main method of recognition is by detection of viral ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) (Chow et al., 2018). Host recognition of viral RNA or DNA triggers a series of intracellular signalling events which result in production of antiviral molecules (Schoggins, 2019). To successfully replicate, viruses must evolve mechanisms by which to evade or dampen the host immune response resulting in a molecular arms

race between host and virus in which various mechanisms of detection and subsequent restriction by the host and evasion by virus come about. One of these restriction mechanisms employed by the host to control viral infection is viral mRNA degradation involving a number of key antiviral proteins that bind to viral RNA, regulate its translation and then target it for degradation significantly interfering with multiple different stages of the viral replication cycle (Abernathy and Glaunsinger, 2015)

4.2 ZAP

4.2.1 Discovery

The zinc finger antiviral protein (ZAP), also known as poly(ADP-ribose) polymerase-13 (PARP13), Zinc finger CCCH-type, and also “antiviral 1” is encoded by the gene ZC3HAV1 (zinc finger CCCh-type containing, antiviral 1, chromosome 7). ZAP is a member of the PARP family which utilise nicotinamide adenine dinucleotide otherwise known as NAD^+ as a substrate for generating modifications in acceptor proteins but does not have poly(ADP-ribosylation) activity. ZAP is a type I interferon-inducible host factor originally discovered during a screen for interferon-stimulated genes (ISGs) that confer resistance in rats to infection of cells by Moloney murine leukaemia virus (MLV) (Gao et al., 2002, Vyas et al., 2013). During this screening, it was found that overexpression of ZAP resulted in a 30-fold increase in resistance to viral infection. ZAP is induced by both type I and type II interferons specifically inhibit certain viruses by binding directly to a region within the viral mRNA called the ZAP-responsive element (ZRE) (Zhu et al., 2011). For many viruses this ZRE are the cytosine-phosphate-guanine (CpG) dinucleotides found within viral RNAs as this is not present in host RNA and allows ZAP to distinguish between self and non-self RNA by this selective binding to CpG dinucleotides (Takata et al., 2017, Meagher et al., 2019, Luo et al., 2020). Once ZAP has bound to the CpG ZRE it represses translation and promotes degradation of the target viral mRNA (Gao et al.,

2002, Zhu and Gao, 2008). ZAP is involved in the host restriction of several negative-sense single stranded RNA viruses (Galão et al., 2022) and also positive-sense single-stranded RNA viruses (Yang et al., 2022) but the role ZAP has during double-stranded RNA viral infection has yet to be fully ascertained.

4.2.2 Isoforms

There are currently four known ZAP isoforms all of which contain an N-terminal RNA-binding domain and a central domain that binds poly(ADP-ribose)(Gao et al., 2002, Guo et al., 2004, Xue et al., 2022, Chen et al., 2012); long (ZAP-L), short (ZAP-S), medium (ZAP-M), and the extra-long isoform (ZAP-XL) (Li et al., 2019). ZAP-L and ZAP-S are the most abundant isoforms and ZAP-L being the endogenous isoform of ZAP (Li et al., 2019, Kerns et al., 2008). ZAP-L and ZAP-S are different as ZAP-L contains a catalytically inactive C-terminal poly(ADP-ribose) polymerase (PARP) domain and an associated C-terminal CaaX box motif crucial for ZAPs CpG-specific activity and binding to cofactors TRIM25 and KHNYN. The CaaX box mediates post-translational modification via a hydrophobic S-farnesyl. This process is called S-farnesylation which is a type of prenylation (addition of hydrophobic molecules to a protein) which attaches a farnesyl group to a cysteine residue to form a thioether bond and become membrane-associated due to the farsenyl's hydrophobic characteristic (Lackie, 2010). The attachment of the S-farnesyl group to ZAP-L causes re-localisation from the cytoplasm and increases its association with intracellular membranes (Kmiec et al., 2021a) which has been shown to be a crucial process in the ZAP-mediated inhibition of HIV-1, SARS-CoV-2 (Kmiec et al., 2021b), and Sindbis virus (SINV) (Schwerk et al., 2019, Charron et al., 2013). It is thought that the membrane-association of ZAP-L caused by the S-farnesylation process is why it shows a different cellular localisation such as plasma membranes or membranous compartments (endolysosomes and the endoplasmic reticulum) to ZAP-S which localises to the cytoplasm (De Andrade and Cirne-Santos, 2023).

ZAP is endogenously and consistently expressed in human cells and ZAP-L and ZAP-S isoforms occur through alternative splicing. ZAP-L is 902 amino acids in length and associates with the membrane. ZAP-S is 699 amino acids in length and localises to the cytosol. Both the ZAP-L and ZAP-S isoforms originate from the same exon (Kerns et al., 2008, Vyas et al., 2013)

Viral infection is the catalyst for ZAP activity due to the presence of binding sites for the signal transducer and activator of transcription (STAT) and Interferon Regulatory Factor 3 (IRF3) in its promoter (MacDonald et al., 2007). There are differences in isoform expression with ZAP-L being constitutively expressed in Huh7 cells and working quickly on present infections and ZAP-S is only expressed depending in IFN signalling(Hayakawa et al., 2011, Schwerk et al., 2019)

4.2.3 ZAP structure

ZAP possesses three distinct structural domains within humans as shown in Figure 4.1; an N-terminal RNA-binding domain (RBD) (amino acids 1-240) which contains four CCCH-type zinc-fingers; an integrated central domain (amino acids 241-700) containing TiPARP homology region (TPH) domain containing another zinc finger motif as well as two WWE modules (Domian in Deltex and TRIP13 homologues (Thyroid Hormone Receptor Interactor 12)); and a catalytically inactive C-Terminal poly(ADP-ribose) polymerase (PARP)-like domain (amino acids 701-902) that have regulatory functions (Kerns et al., 2008, Kmiec et al., 2021b, Gonçalves-Carneiro et al., 2021). ZAP-S does not possess the PARP-like domain characteristic of ZAP-L. This is due to ZAP-S lacking the histidine, tyrosine, and glutamate catalytic triad (HYE) (Goodier et al., 2015).

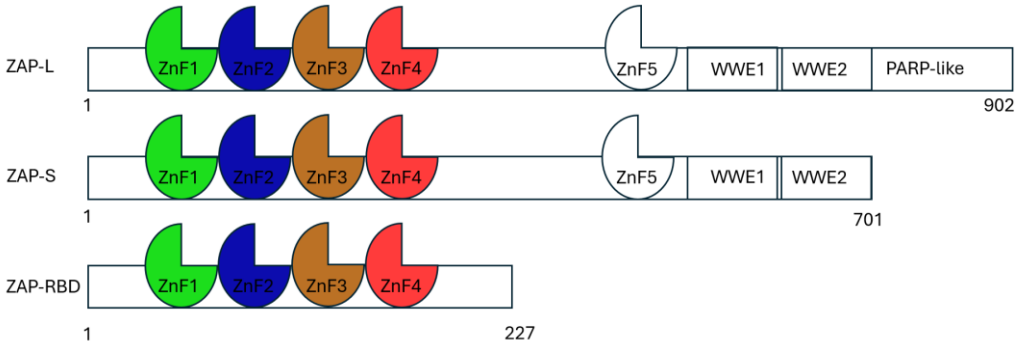


Figure 4.1 Schematic image of the protein domains of the two isoforms of human ZAP: ZAP-L and ZAP-S. ZnF1–4: four CCCH-type zinc finger motifs. TPH (or TiPARP Homology domain (conserved among ZAP paralogs and containing a fifth CCCH zinc finger motif). WWE motifs.

4.2.4 CpG dinucleotides and RNA recognition by ZAP

CpG dinucleotides work as a good selector between self and non-self RNA for ZAP because in they are generally present at higher frequency within viral genomes than they are in host vertebrate genomes in which CpG dinucleotides have relatively low frequency due to cytosine DNA methylation events as well as spontaneous deamination of the 5-methylcytosine to thymine (Holliday and Grigg, 1993). This results in the gradual replacement of CpG dinucleotides with TpG and CpA (Cooper and Gerber-Huber, 1985). In the context of RNA viruses, however, CpG dinucleotides do not undergo the same methylation pressure. Nonetheless, many vertebrate viruses, including those RNA viruses without DNA intermediates, often display suppression of these CpG dinucleotides within their genomes. There have been several experiments which have artificially increased the frequency of CpG dinucleotides within viral genomes which have subsequently resulted in increased inhibition of viral replication (Karlin et al., 1994, Greenbaum et al., 2008, Simmonds et al., 2013, Tulloch et al., 2014, Atkinson et al., 2014, Gaunt et al., 2016, Fros et al., 2017, Antzin-Anduetza et al., 2017, Ficarelli et al., 2020, Fros et al., 2021, Afrasiabi et al., 2022). These viruses include, HIV-1, SARS-CoV-2, echovirus 7, Influenza, and Zika. A high frequency of CpG dinucleotides

within viral genomes may have several negative and deleterious effects on viral replication (Caudill et al., 2020) with ZAP specifically binding to these CpG dinucleotides to restrict replication. It has been suggested that ZAP is somewhat responsible for this CpG dinucleotide frequency suppression in viruses (Kmiec et al., 2020, Gonçalves-Carneiro et al., 2021).

Whilst we know that ZAP targets and binds to CpG dinucleotides within viral mRNA, the specificity of the binding is still unclear, and some research shows that ZAP also exhibits antiviral activity by binding to another dinucleotide; UpA. It only occurs in certain viruses and whether there is direct binding between ZAP and UpA remains to be seen (Simmonds et al., 2013, Odon et al., 2019, Fros et al., 2021). It has also been shown that ZAP detection of non-self-genomes involves some RNA secondary structures which contain stem loops possessing conserved sequences such as “GGGUGG” and “GAGGG” (Huang et al., 2010). Experiments to mutate and alter these ZAP conserved regions results in a reduced ability to recognise RNA and a subsequent decrease in the antiviral ability of ZAP.

Another interesting point is that increased CpG dinucleotide frequency within viruses is not necessarily an indicator of guaranteed ZAP sensitivity with some research showing that ZAP can restrict viruses with high UpA dinucleotide frequency (Ficarelli et al., 2020)

4.2.5 ZAP-mediated vRNA translation inhibition

ZAP can inhibit the translation of vRNA through several mechanisms depending on the type of virus being targeted. ZAP inhibits translation of Sindbis virus mRNA through interaction with eukaryotic initiation factors (eIFs) which are involved in translation (Figure 4.2). These eIFs which ZAP interacts with include eIF4A which is a member of the DEAD box protein family necessary for decoding mRNA. ZAP can bind to eIFs? and inhibit eIF4F complex formation. The eIF4F complex is comprised of the eIF4A DEAD box RNA helicase, the eIF4E cap-binding protein, and the eIF4G scaffolding protein. ZAP Inhibiting this complex being formed causes a

block in translation (Zhu et al., 2012). Therefore, ZAP not only induces degradation of target vRNA of certain viruses through several different mechanisms but can also block translation in other viruses. The determining factors behind whether a virus undergoes degradation or translational blocking is still unknown and whether ZAP-mediated translational blocking is required for subsequent degradation of vRNA or vice versa is still to be understood.

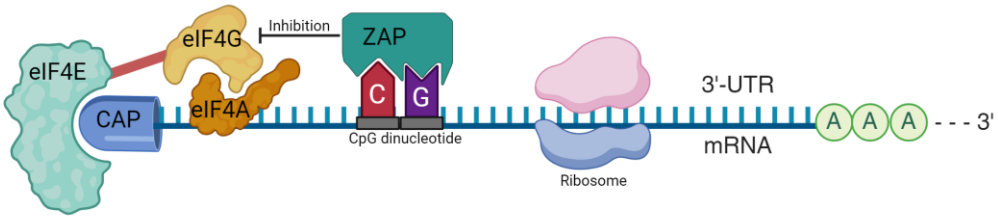


Figure 4.2 ZAP blocks viral mRNA translation by preventing the assembly of the eIF4F complex. ZAP, through its interaction with the initiation factor eIF4A, disrupts viral translation by inhibiting the formation of the eIF4F complex, which consists of the eIF4E cap-binding protein, eIF4A DEAD box RNA helicase, and eIF4G scaffolding protein. The CAP refers to a structure located at the 5' end of mRNA.

4.2.6 Cofactors

The role of ZAP within the host immune response is to recognise and bind to viral mRNAs and direct them towards degradation. It is one of the first steps in an antiviral pathway but requires assistance from other proteins called cofactors to degrade the viral mRNAs and successfully restrict and inhibit viral replication.

ZAP does not have any RNase activity by itself and relies completely on other mechanisms for any antiviral activity. ZAP utilises an interaction with a cellular polyadenylate-specific ribonuclease (PARN) to assist in RNAse activity by degrading the RNA poly(A) tail. ZAP also utilises an exosomal complex containing exoribonucleases that have 3’-5’ activity which cleave viral RNAs. These exoribonucleases included RNA-processing protein 46/exosome complex component (RRP46/EXOSC5) and ribosomal RNA-processing protein 42/exosome complex component (RRP42/EXOCS7) (Zhu et al., 2011, Guo et al., 2007).

ZAP binding to viral mRNA can also activate the de-capping complex of de-capping protein 1 (DCP1) and 2 (DCP2) via RNA helicase p72 (DDX17), resulting in removal of the 5' cap structure of the viral mRNA. ZAP also recruits the 5'-3' Exoribonuclease 1 (XRN1) to assist in viral mRNA degradation (Zhu et al., 2011).

ZAP also lacks any nuclease activity but recruits TRIM25 as probably one of its most important cofactors to assist in viral mRNA degradation (Li et al., 2017, Zheng et al., 2017). TRIM25 plays a crucial role within the host innate immune response to viral infections and is induced by type I IFN. TRIM25 is able to bind RNA but can also modify ZAP-L and ZAP-S via K48- and K63-linked polyubiquitin which causes enhanced antiviral activity of ZAP however this mechanism is not fully understood.

ZAP also recruits another cofactor called KH and NYN domain-containing protein (KHNYN) which is a cytoplasmic protein that contains an NYN ribonuclease domain and targets viral RNAs for degradation (Ficarelli et al., 2019). KHYNYN has been shown to interact with both the ZAP-L and ZAP-S isoforms. Research has shown that in cells which express high levels of KHNYN but do not express ZAP there is no significant inhibition of genomic RNA (gRNA) of HIV-1 with artificially increased levels of CpG dinucleotides within the Env protein (HIV-1EnvCpG86-561). (Ficarelli et al., 2019). It was also shown that HIV-1 had enhanced replication in the absence of KHNYN expression suggesting that KHNYN restricts HIV-1 RNA that contains CpG dinucleotides in a ZAP dependent manner. Within KHNYN, the KH-like domain and the NYN domain endonuclease are necessary for antiviral activity and KHNYN restricts and inhibits Env and Gag expression and the production of virions. KHNYN binds directly to ZAP to form a heterodimer and it has been shown that KHNYN-ZAP complex inhibition of HIV-1 also requires TRIM25 but TRIM25 is not required for KHNYN and ZAP to interact initially (Ficarelli et al., 2019).

P72 RNA helicase is another cofactor for ZAP which enhances and enables an antiviral state within the host cell. This protein is also known as p72

DEAD-box RNA helicase or DDX17 (Cordin et al., 2006) and is involved in RNA structure regulation. DDX17 contains a conserved motif of Asp-Glu-Ala-Asp (DEAD) and is involved in ATP-dependent RNA-helicase activity by catalysing the rearrangement of RNA structure as well as being involved in several metabolic processes such as transcription (Fuller-Pace, 2006), translation, and importantly for ZAP, RNA degradation (Wortham et al., 2009). ZAP binds to both the N-terminal and C-terminal domains of DDX17 in an RNA-independent manner (Zhu and Gao, 2008) and this interaction promotes and enhances the effectiveness of ZAP's ability to inhibit viruses through targeting mRNAs for exosomal-mediated degradation. Another antiviral mechanism of DDX17 is in the recruitment of the Dcp1:Ddcp2 de-capping enzyme to the 5' end of target vRNAs. This process inhibits the cap-dependent translation initiation and induces vRNA degradation (Chen et al., 2008). Whilst DDX17 does not directly interact with the exosome, it recruits the exoribonuclease complex XRN1, DDX17 then forms a complex with CpG activated ZAP (Chen et al., 2008).

Another E3 ubiquitin ligase protein; Riplet, has been suggested as a ZAP cofactor (Buckmaster and Goff, 2022). Overexpression of Riplet resulted in increased ZAP antiviral activity in cells infected with VSVG-pseudotyped HIV-luc reporter virus (Buckmaster and Goff, 2022). Although, there appears to be link between the E3 ubiquitin ligase function of Riplet and the antiviral abilities of ZAP. It was shown that ZAP can bind to the Riplet C-terminal P/SPRY domain and this interaction is necessary for HIV-1 reporter virus inhibition. Riplet has also been shown to interact with TRIM25 but whether an antiviral complex of Riplet/TRIM25/ZAP forms is still unknown.

4.2.7 Immune Pathways and ZAP

Cells developed surface pattern recognition receptors (PRRs) to recognise viruses to identify molecular structures on the surfaces of a variety of cell types including damaged cells, apoptotic host cells and pathogens. Figure 4.3 displays an overview of the interactions ZAP has in the host antiviral immune response. These crucial components of the host innate immune

system include members such as Toll-like receptors (TLRs) and cytosolic receptors such as acid-inducible gene I retinoic receptors (RIG-I, also called DDX58). Once activated, these receptors initiate the downstream signalling cascade. When TLRs are activated by viruses it stimulates adapter molecules such as myeloid differentiation primary response 88 (myD88), myD88 adapter-like protein (Mal), TIR-domain-containing adapter-inducing interferon-β (TRIF), and also TRIF-related adaptor molecule (TRAM) to subsequently activate transcription factors. These transcription factors include NF-κB, IRF3, and IRF7 resulting in proinflammatory cytokine and IFN expression. IFN-I and IFN-III signalling activates the Janus kinase/signal transducer and activator of transcription (JAK/STAT) pathway. This pathway is initiated with phosphorylation of JAK1 and non-receptor tyrosine-protein kinase (TYK2) on the cytoplasmic domains of the heterodimeric receptor subunits. This initial activation is followed by phosphorylation and dimerisation of STAT1 and STAT2.

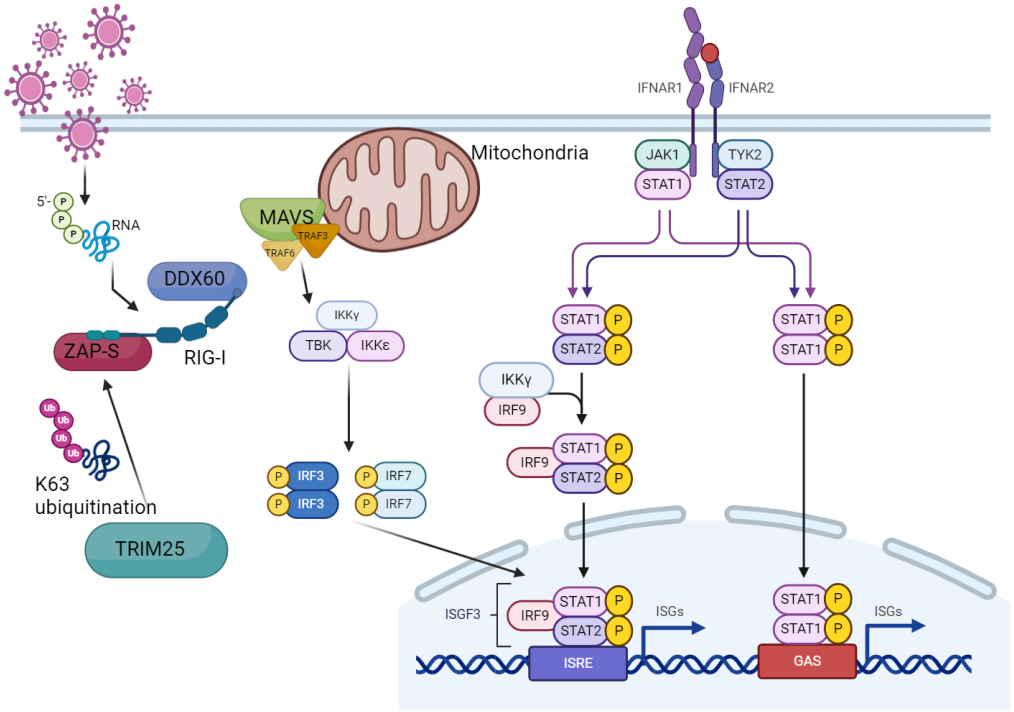


Figure 4.3 Interaction between ZAP, RIG-I, and IFN in the antiviral immune response : ZAP-S interacts with RIG-I to enhance the oligomerization and ATPase activity of RIG-I. This, in turn, increases the activation of IRF3 downstream when the RIG-I ligand, 3'pRNA, is present in human cells. DDX60 associates with RIG-I and is involved in RIG-I-dependent type I IFN production in response to viral RNA. TRIM25 activates the RIG-I pathway by facilitating K63-linked polyubiquitin chain formation through its E3 ubiquitin ligase activity. This ubiquitination promotes interaction with MAVS, leading to downstream signalling. DDX60 is a DEXD/H box helicase. Biorender used to create figure.

In the IFN-III signalling pathway, activation of STAT1 and STAT2 recruits IRG-9 to form a complex called the Interferon-Stimulated Gene Factor 3 (ISGF3) which translocates to the nucleus and binds to the interferon-sensitive response element (ISRE) located in the promoter regions of interferon-stimulated genes (ISGs). ISGs that are stimulated into transcription and translation include ZAP in which IRF3 binds to ISRE in the ZAP promoter during viral infection; and members of the TRIM family including TRIM25 (Crosse et al., 2017, Yang and Li, 2020, Wang et al., 2010). Research has shown that IFN-I has more of an effect on induction of ZAP-S expression than ZAP-L exactly why or how is still unclear (Li et al., 2019).

RIG-I is a member of the Dex(D/H) box helicase family and is comprised of two N-terminal CARD domains, then a central RNA helicase domain and finally, a C-terminal repressor domain with ATPase activity that is also involved in recognising 5'-triphosphorylated RNA. It has been shown that TRIM25 and K6-linked ubiquitin chains polyubiquitinate the second RIG-I CARD (CARD2) and cause oligomerisation and activation of RIG-I. Once activated, RIG-I is able to recruit and bind to MAVs via its CARD domain which subsequently activates TBK1- IKKε and IKKα-IKKβ complexes which in turn activate IRF-3/IRF7 and NF-κB respectively. Translocated IRF3 and IRF7 induce IFN-I synthesis which then bind to target receptors to induce intracellular signalling resulting in ISG transcription including ZAP (Rehwinkel and Gack, 2020, Gack et al., 2007). Another E3 ubiquitin ligase called Riplet promotes RIG-I signalling but in a manner independent of RIG-I oligomerisation (Wang et al., 2023) and instead causes K63-linked polyubiquitination of RIG-I RD (Oshiumi et al., 2013). Riplet has also been shown to induce TRIM25 mediated activation of RIG-I signalling and can act as a co-receptor assisting in oligomerisation of RIG-I. Interestingly, research has shown that TRIM25 is not required for full-length RIG-I signalling and instead, ectopic TRIM25 expression mildly stimulates signalling mediated by RIG-I CARD fragments (Cadena et al., 2019). ZAP-S

enhances IFN production in HEK293T cells when activated by 5'triphosphate RNA which directly binds to RIG-I (Hornung et al., 2006a).Expression of ZAP-S induces RIG-I-mediated IFN-I response in immune cells such as human primary $CD14^+$ and fibroblasts (Hayakawa et al., 2011). It has been shown that ZAP-S has multiple functions within stimulating RIG-I; ZAP-S can cleave vRNA with assistance from p27 RNA helicase or it can bind to RIG-I to stimulate activation of the innate antiviral immune system. The ZAP-S gene promoter region contains ISREs and IRF-binding elements but the many of the mechanisms behind ZAP-S regulation are still poorly understood and why ZAP-L does not regulate RIG-I in the same way is also not understood.

Another immune pathway ZAP plays a role in is the 2', 5'-oligoadenylate synthetase 1 (OAS1)-RNase1 antiviral pathway. OAS1 is an ISG found I the cytosol as an inactive monomer with low levels of endogenous expression that can be upregulated by IFN-I and IFN-III. OAS1 is activated by dsRNA and is oligomerised to form a tetramer that utilises ATP to synthesis 2',5'-oligoadenylate molecules which then bind to the inactive RNaseL monomers causing dimerisation and allosteric activation. RNaseL is an important inhibitor of viral replication, it is found in the cytoplasm and is constitutively expressed, once activated it functions to degrade a wide range of viral ssRNA and cellular RNAs (Schwartz et al., 2020). It has been found that ZAP and RNaseL are both involved in inhibiting E7 virus possessing a high CpG and UpA dinucleotide frequency within its genomic sequence. E7 viral inhibition depended on expression of both ZAP and RNaseL in vitro (Odon et al., 2019) and in the absence of ZAP there was an increase in RNaseL expression suggesting a collaborative relationship to inhibit viruses. It has been shown that RNaseL activity and activation can be targeted by viruses and that ZAP might act as an alternative antiviral pathway providing redundancy to the innate immune antiviral response (Bhattacharyya, 2014, Silverman and Weiss, 2014).

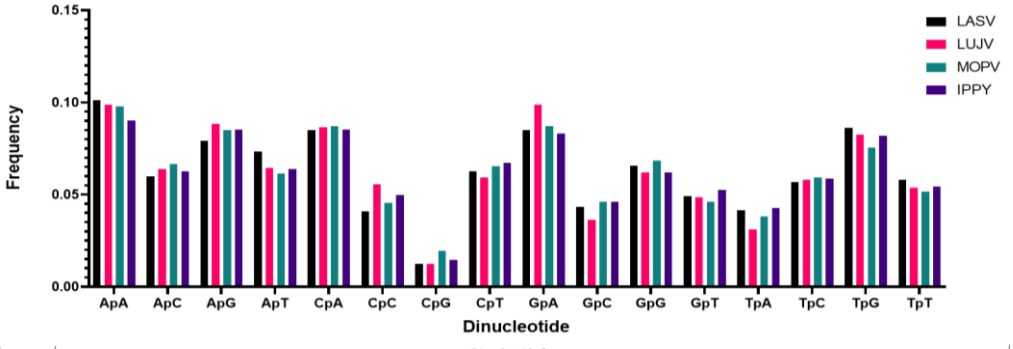
4.2.8 Sequence composition analysis of arenavirus strains reveals that they have greatly reduced CpG dinucleotide frequency

Many mammalian viral RNA genomes have evolved to have significantly reduced CpG dinucleotides (Takata et al., 2017) to suppress antiviral defence targeting. ZAP has been shown to directly bind viral CpGs and induce degradation of the RNA or repress translation. This has been shown in HIV-1 in which artificially increasing viral genome CpG concentration resulted in impaired replication via ZAP-mediated inhibition (Galão et al., 2022).

In figure 4.4, analysis has been performed on the CpG concentrations and frequencies of each of the OW arenavirus strains of interest and also the non-human-pathogenic Ippy virus. Another non-human-pathogenic arenavirus was included for analysis to ascertain with more confidence whether CpG dinucleotide frequency suppression is a characteristic associated with increased pathogenicity. In all four arenavirus strains, the frequency of CpG dinucleotides is greatly reduced in comparison to all other dinucleotides. Interestingly, both of the pathogenic strains (LASV and LUJV) do have an increase in CpG suppression compared to the non-pathogenic strains.

This decreased frequency of CpG regions will limit the number of CpG rich regions available for ZAP to bind to. This provides some insight into the need for multiple ZAP isoforms though as the main difference between the ZAPL and ZAPS isoform is the PARP-like domain that is spliced out from ZAPS. This domain's core function is in CpG-specific viral restriction (Kmiec et al., 2021a). As shown in figure 4.4, all analysed arenavirus strains have reduced CpG dinucleotide frequency which might have evolved to combat restriction factors such as ZAP (and thus ZAPL) targeting these regions to induce RNA degradation and inhibit replication and infection. ZAPS lacks CpG dependent antiviral activity due to missing the PARP-like domain

although it still retains and conserves the N-terminal RNA binding domain. The reduced CpG frequency of arenaviruses is not as much of a limiting factor for ZAPS activity as its antiviral activity is achieved through other mechanisms such as being a much more potent stimulator of the IFN response in cells. The impact this reduced dinucleotide frequency has upon TRIM25, whether direct or indirect (such as via ZAPL activity and recruitment of TRIM25) has yet to be characterised in the context of arenavirus infection and further analysis and experiments must be performed to explore the true magnitude of effect that CpG content plays and potentially how it can be exploited or avoided, as ZAPS may have found a way to do.



	Dinucleotide frequency															
NP	ApA	ApC	ApG	ApT	CpA	CpC	CpG	CpT	GpA	GpC	GpG	GpT	TpA	TpC	TpG	TpT
LASV	0.1013	0.0598	0.0791	0.0732	0.0849	0.041	0.0123	0.0627	0.0849	0.0434	0.0656	0.0492	0.0416	0.0568	0.0861	0.058
LUJV	0.0986	0.0639	0.0884	0.0645	0.0866	0.0556	0.0125	0.0591	0.0986	0.0364	0.0621	0.0484	0.0311	0.0579	0.0824	0.0538
MOPV	0.0977	0.0667	0.0848	0.0614	0.0871	0.0456	0.0193	0.0655	0.0871	0.0462	0.0684	0.0462	0.038	0.0591	0.0754	0.0515
IPPY	0.0901	0.0626	0.0854	0.0637	0.0854	0.0497	0.0146	0.0673	0.083	0.0462	0.062	0.0526	0.0427	0.0585	0.0819	0.0544

Figure 4.4 CpG dinucleotide frequency differences between arenavirus strains. (A) Clustered bar graph created using SSE (sequence editor, Database and Analysis Platform (Simmonds, 2012). The graph displays dinucleotide frequencies for each of the arenavirus strains LASV (Josiah strain), LUJV, MOPV (Mozambique strain), and Ippy virus (B) Data table for the bar graph above with colour conditional formatting based on frequency.

4.3 TRIM25

4.3.1 Introduction to TRIM25

Tripartite motif containing protein 25 (TRIM25) is an E3 ubiquitin ligase and is a member of the TRIM family of proteins. This family has over 80 members in humans and is a group of E3 ubiquitin ligase proteins all sharing a common domain structure. TRIM25 in humans is 630 amino acids in length and is widely expressed across human cell types and is also conserved among vertebrates including mammals, birds, and fish (Fagerberg et al., 2014). Similar to other members of the TRIM family, TRIM25 contains an N-terminal zinc-finger "really interesting new gene" (RING) domain which is required for E3 ubiquitin ligase activity; TRIM25 also contains two B-box zinc finger domains the function of which is still unclear; a coiled-coil domain (CCD) required for homo- and heterodimerisation; a linker domain leading on to a C-terminal that has an associated SPRY/SPIa and Ryanodine receptor (PRY/SPRY) domain that is responsible for any protein-protein interactions with TRIM25. Figure 4.5 depicts a schematic of the TRIM25 protein and domains.

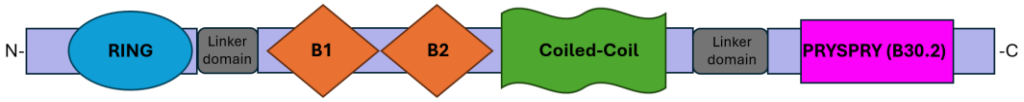


Figure 4.5 Schematic image of the protein domains of human TRIM25 : A RING domain followed by a linker domain that connects it to two B-box finger domains and then a Coiled-coil domain connected to the PRYSPRY domain by another linker.

The main function of TRIM E3 ubiquitin ligases is in catalysing the addition of polyubiquitin chains or single ubiquitin monomers (mono-ubiquitination) to lysine residues on target proteins. Ubiquitin is a small protein with myriad functions found in most tissues; it mainly helps to regulate the processes of other proteins in the cell. Ubiquitin is 76 amino acids in length and polyubiquitin chains are created by isopeptide bond formation between a C-terminal glycine and one of the 7 lysine residues present in the ubiquitin protein (Kerscher et al., 2006). Ubiquitination of a protein by addition of ubiquitin monomers or extension of an already existing polyubiquitin chain

requires three types of proteins; E1 activating enzyme, E2 conjugating and E3 ubiquitin ligase (Ye and Rape, 2009). Ubiquitin contains 7 lysine residues; the presence of these multiple residues allows for the formation of different types of polyubiquitin chains which all have differing functions. The most well understood and researched are straight, homogeneous K48-linked polyubiquitination chains involved in the targeting of proteins for proteasomal mediated degradation. K63-linked polyubiquitination chains are also well researched and involved in many intracellular signalling pathways including NF-κB activation and retinoic-acid-inducible-gene-I (RIG-I)/IKK/IFN-I induction (Komander and Rape, 2012).

TRIM25 dimerisation is mediated by the CCD and results in the formation of an antiparallel dimer with the RING domain of each monomer sitting at opposite ends of the dimer and the PRY/SPRY domains at the centre via the CCD-PRY/SPRY linker (Sanchez et al., 2014). Research has shown that for TRIM25 to successfully catalyse polyubiquitination chain formation, the RING domain of TRIM25 must dimerise, suggesting that higher-order assembly of TRIM25 dimers is required for proper function (Sanchez et al., 2016). Two different methods of higher-order assembly have been suggested; an end-to-end model with RING domains on either end of the dimer which then interact with RING domains from different dimers; the second method is a tetramer model with TRIM25 dimers stacking on top of each other with the RING domains on either end of one dimer then interacting with both RING domains from another dimer (Sanchez et al., 2016)

4.3.2 TRIM25 functions

TRIM25 is involved in a multitude of different roles within the cell and was first identified as a protein that responded to estrogen during a screen for regions of DNA bound by the estrogen receptor and subsequently shown to be upregulated in estrogen receptor-positive mammary cells (Inoue et al., 1993). TRIM25 has been shown to play important roles in the host defense against viral infection and is also involved in the activation of the RIG-I/IFN

pathway. TRIM25 also has other roles in innate immunity with one of these roles being in the enhancement of ZAP activity (Li et al., 2017). TRIM25 has also been shown to positively regulate Melanoma Differentiation-Associated gene 5 (MDA5) mediated signalling via Tumour Necrosis Factor Receptor-associated factor 6 (TRAF6) which is another E3 ubiquitin protein ligase, this positive regulation results in NF-κB activation (Lee et al., 2015). Although involved in the activation of the RIG-I pathway, TRIM25 has also been shown to be involved dampening of RIG-I signalling as well through formation of a complex between TRIM25/RIG-I and ubiquitin D (ubiquitin-like FAT10) which causes sequestering of RIG-I away from mitochondria and formation of insoluble aggregates inhibiting further signal transduction (Nguyen et al., 2016). FAT10 is usually unstable but TRIM25 prevents proteome-mediated degradation helping to stabilise FAT10.

TRIM25 plays a significant role in the RIG-I/IFN pathway. RIG-I is one of the host's first lines of defense against RNA viruses and acts as a pattern recognition receptor involved in recognising RNAs possessing a 5'-di or -triphosphate (5'pp or 5'/ppp) moiety. Once activated by detecting vRNA, RIG-I changes conformation to release a pair of signalling domains (Caspase recruitment domains or CARDs) which initiate downstream signalling pathways via the adaptor Mitochondrial Antiviral Signalling protein (MAVS also known as VISA, IPS-1, and Cardif). This signalling via MAVS results in activation of IRF-3 and 7, NF-κB, and IFN-I expression (Yoneyama and Fujita, 2008b, Yoneyama et al., 2015). The role TRIM25 plays is in ubiquitination of the RIG-I amino-terminal CARDs. The TRIM25 carboxy-terminal SPRY interacts with the N-terminal CARDs and delivers the Lys 63-linked ubiquitin moiety to the N-terminal CARDs causing significant increase in RIG-I downstream signalling (Gack et al., 2007).

It is important that RIG-I signalling is controlled to avoid excessive or inappropriate activation of the host innate immune responses which could then lead to unwanted inflammation and IFN responses. Thus, mechanisms exist to dampen RIG-I signalling when there is no 5'pp- or

5’/ppp-vRNA present. One of these mechanisms is targeting of TRIM25 for inhibition by the host cell. This is achieved via the Linear Ubiquitin Assembly Complex (LUBAC) which is comprised of Heme-oxidised IRP2 Ubiquitin Ligase 1 (HOIL-1) and HOIL-1 Interacting Protein (HOIP). LUBAC inhibits TRIM25 activation of RIG-I by competitively binding to RIG-I and also by targeting TRIM25 for proteosome mediated degradation (Inn et al., 2011).

4.3.3 Viral suppression and evasion of TRIM25-mediated host antiviral immunity

There are many examples of viruses hijacking the host antiviral responses to replicate efficiently, these viruses either evade or dampen the host immune response. There are several examples of this occurring in the TRIM25/RIG-I pathway with many RNA viruses that are recognised by RIG-I having developed mechanisms to inhibit RIG-I signalling at various stages of the pathway. Influenza A virus (IAV) NS1 protein can block ubiquitination of RIG-I by interacting with TRIM25 or even interacting with RIG-I directly (Gack et al., 2009, Mibayashi et al., 2007). NS1 achieves this by binding to the TRIM25 CCD and recent structural research suggests that this disrupts interactions between the TRIM25 PRY/SPRY domain and the CCD necessary for RIG-I ubiquitination activity (Koliopoulos et al., 2018). Several paramyxoviruses: Nipah, Sendai, measles, and parainfluenza viruses also show interaction with RIG-I and TRIM25. These paramyxovirus V proteins interact with the TRIM25 SPRY domain and also with the RIG-I CARDs to inhibit TRIM25 ubiquitination of RIG-I (Sánchez-Aparicio et al., 2018). Respiratory syncytial virus (RSV) NS1 (Ban et al., 2018) as well as the SARS-CoV N protein (Hu et al., 2017b) have also been shown to regulate TRIM25 activity. Dengue virus also interferes with TRIM25 by taking advantage of the RNA-binding functions through expression of a sub-genomic flavivirus RNA (sfRNA) which binds to and inhibits TRIM25 subsequently inhibiting RIG-I activation and IFN-I induction. Through this mechanism a new Dengue virus

clade PR-2B was able to increase its epidemiological fitness in comparison to the older PR-1 clade (Manokaran et al., 2015)

4.3.4 TRIM25 and RNA-binding

TRIM25 has been shown to have roles in RNA-binding originally discovered during a HeLa cell screen of mRNA binding proteins. During this screen, proteins were UV-cross-linked to RNA and mRNAs and were then isolated from cell lysates with oligo(dT) probes prior to analysis of bound proteins by mass spectrometry (Castello et al., 2012). TRIM25 binding to RNA was confirmed by immunoprecipitation (IP) of TRIM25 followed by radio-labelling of RNA, the signal from the radio-labelled RNA after TRIM25 IP was lower in cells which had TRIM25 knocked down by RNAi (Kwon et al., 2013). The same study also investigated the RNA binding of TRIM25 mutants. Mutants in which the N-terminal RING and B-box domains, or the C-terminal PRY/SPRY domains were deleted were still able to bind RNA but any mutants with the CCD deleted were no longer able to bind RNA suggesting that the CCD is crucial for TRIM25 RNA binding (Kwon et al., 2013). This research suggested that the PRY/SPRY domain was not necessary for RNA binding, but further research has shown that while TRIM25 is still able to bind RNA when the PRY/SPRY is deleted, this region actually targets to and binds hundreds of coding and non-coding RNA. This binding occurs specifically at amino acids 470-508 within the PRY/SPRY domain (known as the TRIM25 RNA-binding domain or RBD) (Choudhury et al., 2017). Further research into TRIM25 RNA binding revealed that a truncated TRIM25 mutant consisting of the RING, B-box, and CCD domains but not the PRY/SPRY domain could not be co-purified with RNA whilst full length wild-type TRIM25 successfully co-purified with RNA (Sanchez et al., 2018). TRIM25 possesses a motif containing 7 lysine residues at amino acids 381-392 within the CCD-PRY/SPRY linker region called 7K and this motif has also been implicated in RNA binding (Sanchez et al., 2018). Mutating all the lysine residues of the 7K within the linker region resulted in a substantial but not total loss in RNA-binding activity during

electrophoretic mobility shift assay (EMSA) experiments used for detecting protein-nucleic acid interactions.

4.4 ZAP and TRIM25

Viral replication has been shown to cause ZAP to be targeted to and localise with stress granules during viral infection. It has been suggested that this co-localisation of ZAP and stress granules is important for the antiviral activity of ZAP and that stress granules are an important cytoplasmic antiviral hub (Law et al., 2019, Lee et al., 2013), ZAP is also involved in the RIG-I immune pathway by interacting with and activating RIG-I and promoting oligomerisation of RIG-I (Hayakawa et al., 2011).

So far ZAP has not been found to have individual enzymatic ability and requires co-factors to achieve some many antiviral activity requiring enzymes (Kleine et al., 2008). It has been shown that TRIM25 interacts with ZAP and is potentially the most important cofactor of ZAP (Figure 4.6). This interaction occurs via the TRIM25 SPRY domain and the ubiquitin ligase and multimerization activity of TRIM25 are key steps in enhancing the antiviral activity of ZAP (Figure 4.4). Research has shown that only the K-68-linked polyubiquitination mediated by TRIM25 is responsible for enhancing the antiviral activity of ZAP (Li et al., 2017, Zheng et al., 2017). TRIM25 achieves this by assisting in the RNA binding ability of ZAP which subsequently leads to inhibition of viral translation. TRIM25 has also been shown to require cooperation from ZAP to successfully restrict SINV replication (Li et al., 2017). The E3 ligase activity of TRIM25 is required to facilitate and enhance ZAP-mediated inhibition of SINV RNA translation and, interestingly, TRIM25 ubiquitination of ZAP did not seem to have any significant effect on antiviral activity in this specific circumstance. TRIM25 also plays a key role as a ZAP cofactor during infection with HIV-1 virus modified to contain enriched and clustered CpG dinucleotides (Takata et al., 2017). This indicates that by artificially ‘restoring’ the expected CpG dinucleotide frequency of a virus that has evolved to suppress the number of CpG dinucleotides it

possesses, the virus becomes significantly more vulnerable to host antiviral factors; this information supports the hypothesis that viruses have evolved and are still evolving mechanisms to evade or suppress host innate antiviral immune responses mediated by restriction factors such as ZAP and TRIM25.

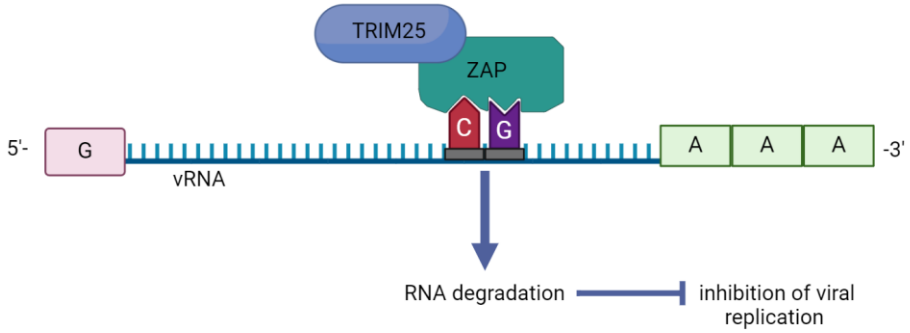


Figure 4.6 TRIM25 acts as a cofactor of ZAP. ZAP binds to the CpG sequence in viral RNA and, upon sequential binding of TRIM25 and catalytic activation, induces downstream signalling to inhibit viral replication. Figure created using Biorender

4.5 Chapter Objectives: Potential roles of ZAP and TRIM25 in the host antiviral response to arenavirus infection

This chapter aimed to:

1. Define the roles of ZAP and TRIM25 in the host antiviral response to arenavirus infection and assess their effectiveness in restricting viral replication.
2. Determine which ZAP isoform (Long or Short) is more critical for anti-arenaviral activity.
3. Identify the functional domains of TRIM25 required for interaction with arenavirus nucleoproteins using specific mutants (7KA ΔRING and ΔSPRY).
4. Characterize the cellular localization patterns and interaction dynamics between arenavirus NPs (LASV, LUJV, and MOPV) and the ZAP-TRIM25 axis.
5. Evaluate the impact of ZAP and TRIM25 overexpression on live MOPV replication.
6. Assess the consequences of ZAP or TRIM25 depletion (via siRNA knockdown or CRISPR knockout) on MOPV replication and interferon responses.
7. Establish ZAP and TRIM25 as key restriction factors in the innate immune response against arenaviruses.

4.6 Results

4.6.1 NP Exhibits Diffuse Cytoplasmic Distribution

To better understand the interactions of ZAP and TRIM25, both with each other and with NP, it is essential to know where NP localises within the cell. This allows for functional and well-informed comparisons to be made once we overexpress ZAP L or ZAP S or TRIM25 within the cell. The three strains of arenavirus NP exhibit a diffuse cytoplasmic localisation, as shown in figure 4.7, although, LUJV NP exhibits some differences and appears to show a higher concentration around the nucleus as well as perinuclear speckling not present in LASV NP and also not present in MOPV NP localisation either. MOPV NP also shows a higher concentration around the nucleus but lacks the speckling seen in LUJV, MOPV NP also has the characteristic cytoplasmic diffusion seen in the NP of the other strains.

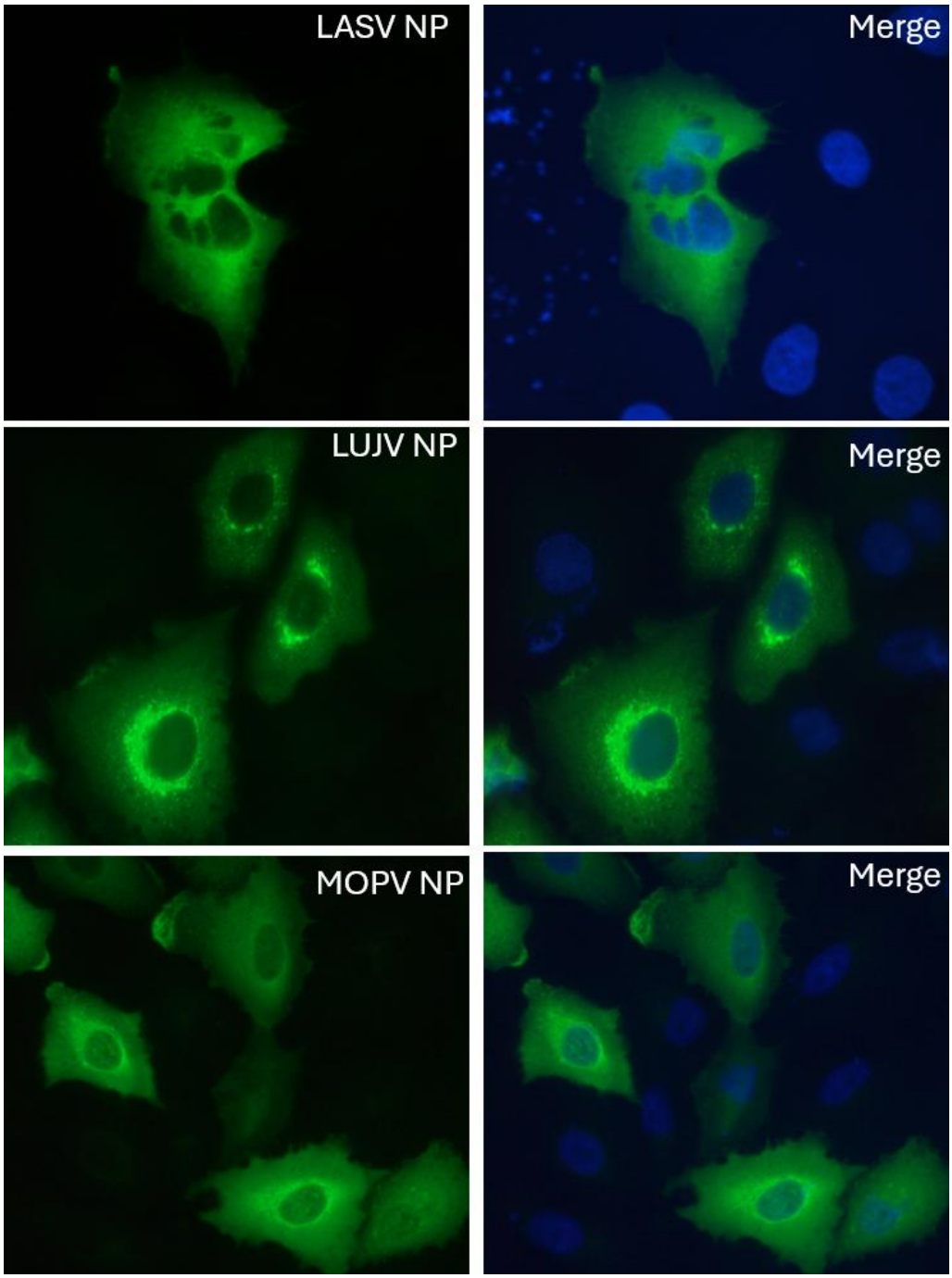


Figure 4.7 Immunofluorescence showing localisations of either LASV-NP-HA, LUJV-NP-HA, or MOPV-NP-HA. A549 cells seeded onto a coverslip at a density of 50,000 per well of a 24 well plate. Cells were transfected with either pCAGGs-LASV-NP-HA, pCAGGs-LUJV-NP-HA, or pCAGGs-MOPV-NP-HA at 500ng per construct per well of 24 well plate using Lipofectamine 3000 transfection reagent. Cells were then stained with mouse-anti-HA (Abcam Ab18181) at a dilution of 1:500 prior to staining with donkey raised-anti-mouse-488nm-fluorescent antibody (Abcam Ab150109) at a dilution of 1:500 to visualise the desired protein in a green colour. Slides were mounted using DAPI mounting solution to stain the cell nucleus for visualisation in blue colour (n=1).

4.6.2 Endogenous ZAP and TRIM25 Cellular Localisations

It was important to establish the baseline cellular localisations of endogenous ZAP and TRIM25. Understanding where these endogenous proteins localise without any external influences allows for easier comparisons to be made. Figure 4.8 shows the localisation of endogenous TRIM25 and endogenous ZAP and both seem to have diffuse cytoplasmic spread.

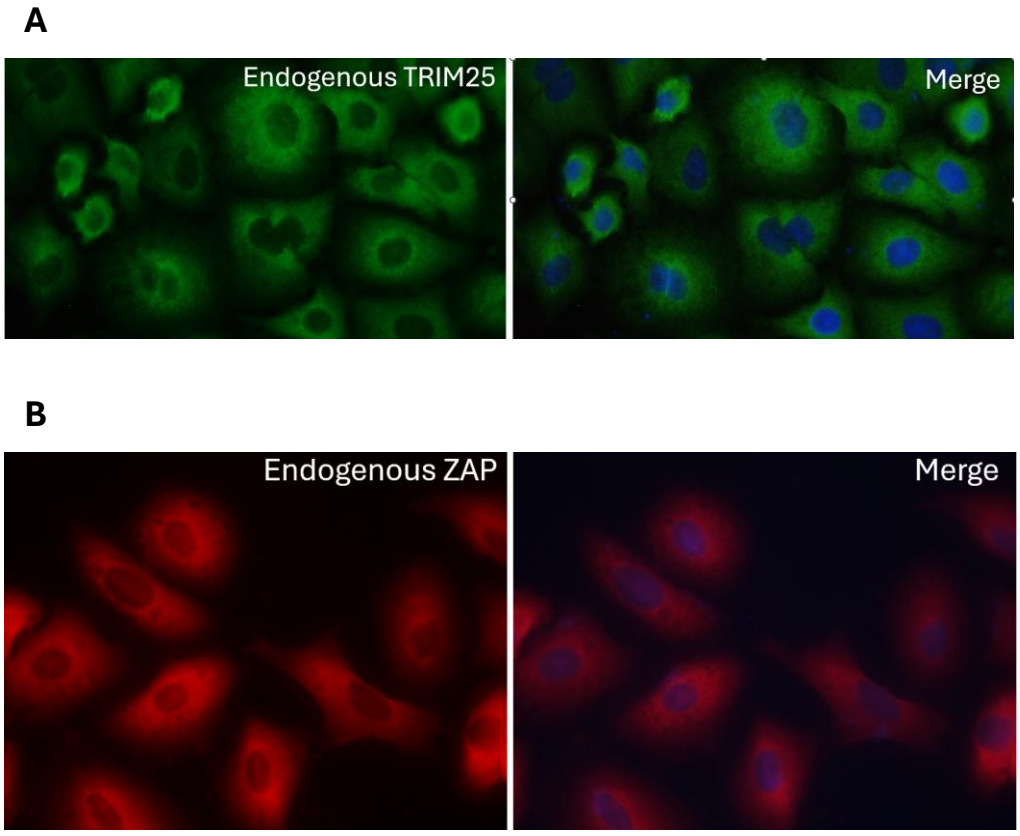


Figure 4.8 Immunofluorescence showing localisations of either (A) Endogenous ZAP, or (B) endogenous TRIM25 A549 cells seeded onto a coverslip at a density of 50,000 per well of a 24 well plate. Cells were then stained with either rabbit anti-ZAP antibody (Ab154680) or mouse anti-TRIM25 antibody (Ab610570) at a dilution of 1:500 prior to staining with donkey raised-anti-mouse-488nm-fluorescent antibody (Abcam Ab150109) at a dilution of 1:500 to visualise the desired protein in a green colour or staining with goat raised anti-rabbit 594nm-fluorescent antibody (Ab150084) to visualise the desired protein in a red colour. Slides were mounted using DAPI mounting solution to stain the cell nucleus for visualisation in blue colour. (n = 1)

4.6.3 TRIM25 Co-localises With Both ZAP Isoforms

Figure 4.9 depicts immunofluorescence of TRIM25 transiently over-expressed by transfection. This was performed prior to any co-localisation experiments to understand whether over-expression of TRIM25 affected its localisation. Comparing the immunofluorescence in Figure 4.9 to that of endogenous TRIM25 in Figure 4.7 we can see that, although over-expression has induced and increase in concentration and subsequent perinuclear speckling, the overall localisation remains fairly consistent with diffuse cytoplasmic distribution occurring in both the transfected TRIM25 and endogenous TRIM25.

Figure 4.10 shows the results of co-transfecting cells with either TRIM25-FLAG and ZAP-L-myc or TRIM25-FLAG and ZAP-S-myc. The results show that TRIM25 co-localises with both ZAP isoforms and that it appears to be targeted to ZAP rather than maintaining normal cellular localisation. This is shown by the merge from figure 4.10 which displays that despite the different sub-cellular localisations of the different isoforms of ZAP, TRIM25 still colocalises. It is known that ZAP interacts with TRIM25 via the SPRY domain (Li et al., 2017) and the structural differences between the ZAP isoforms (the loss of the PARP domain for ZAP-S) must not affect this interaction.

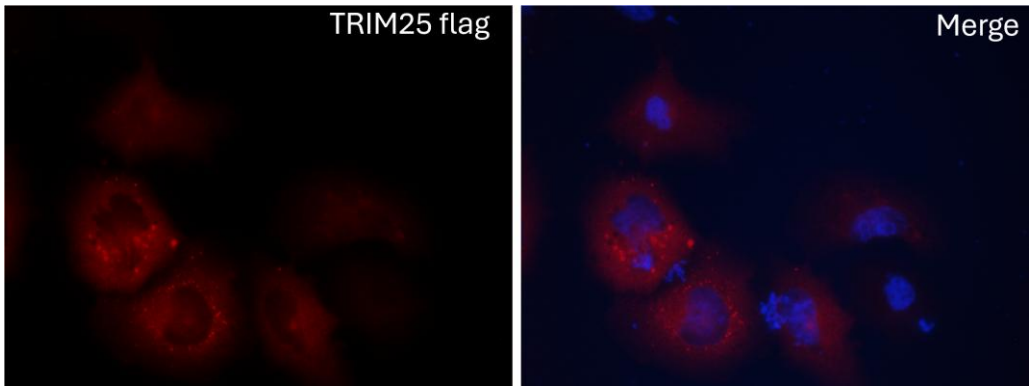


Figure 4.9 Immunofluorescence showing localisation of TRIM25-FLAG. HEK293T cells were seeded onto coverslips which were pre-treated with Poly-L-Lysine to increase the adherence of cells at a density of 50,000 per well of a 24 well plate. Cells were transfected with PCDNA3.1 TRIM25-FLAG at a DNA concentration of 500ng within a well of a 24 well plate using PEI transfection reagent. Cells were then stained with antibodies against mouse-anti-FLAG (Sigma F1804-200UG) at a dilution of 1:500 prior to staining with donkey raised-anti-mouse-594nm-fluorescent antibody (Abcam Ab150108) at a dilution of 1:500 to visualise the desired protein in a red colour. Slides were mounted using DAPI mounting solution to stain the cell nucleus for visualisation in blue colour. (n = 1)

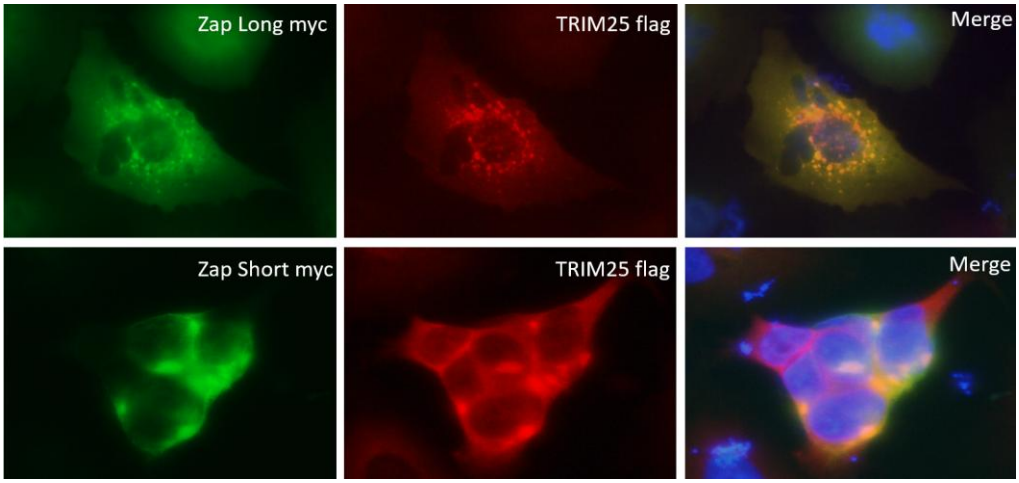


Figure 4.10 Immunofluorescence showing co-localisation of ZAP isoforms and TRIM25. HEK293T cells seeded onto a coverslip pre-treated with Poly-L-Lysine to increase the adherence of cells at a density of 50,000 per well of a 24 well plate. Cells were co-transfected with PCDNA3.1 ZAP-L Myc tag and TRIM25-FLAG or ZAP-S Myc and TRIM25-FLAG at 250ng per construct per well of 24 well plate using PEI transfection reagent. Cells were then stained with antibodies against mouse-anti-myc (Sigma M5546-2ML) and rabbit-anti-FLAG (Sigma F1804-200UG) at a dilution of 1:500 per antibody prior to co-staining with donkey raised-anti-mouse-488nm-fluorescent antibody (Abcam Ab150109) and donkey raised-anti-rabbit-594nm-fluorescent antibody (Abcam Ab150108) at a dilution of 1:500 to visualise the ZAP in green and TRIM25 in red. Slides were mounted using DAPI mounting solution to stain the cell nucleus for visualisation in blue colour. (n = 1)

4.6.4 TRIM25 and Arenavirus NP Co-localise

We observe a change in TRIM25 localisation in the presence of co-transfected arenavirus NP when compared to endogenous TRIM25 localisation seen in figure 4.8. Which way round this interaction occurs in arenavirus infection is not known. It has been suggested in studies of SARS-CoV that SARS N protein targets TRIM25 and competitively binds to the SPRY domain to modulate the RIG-I immune response and could also in turn prevent ZAP binding (Hu et al., 2017a).

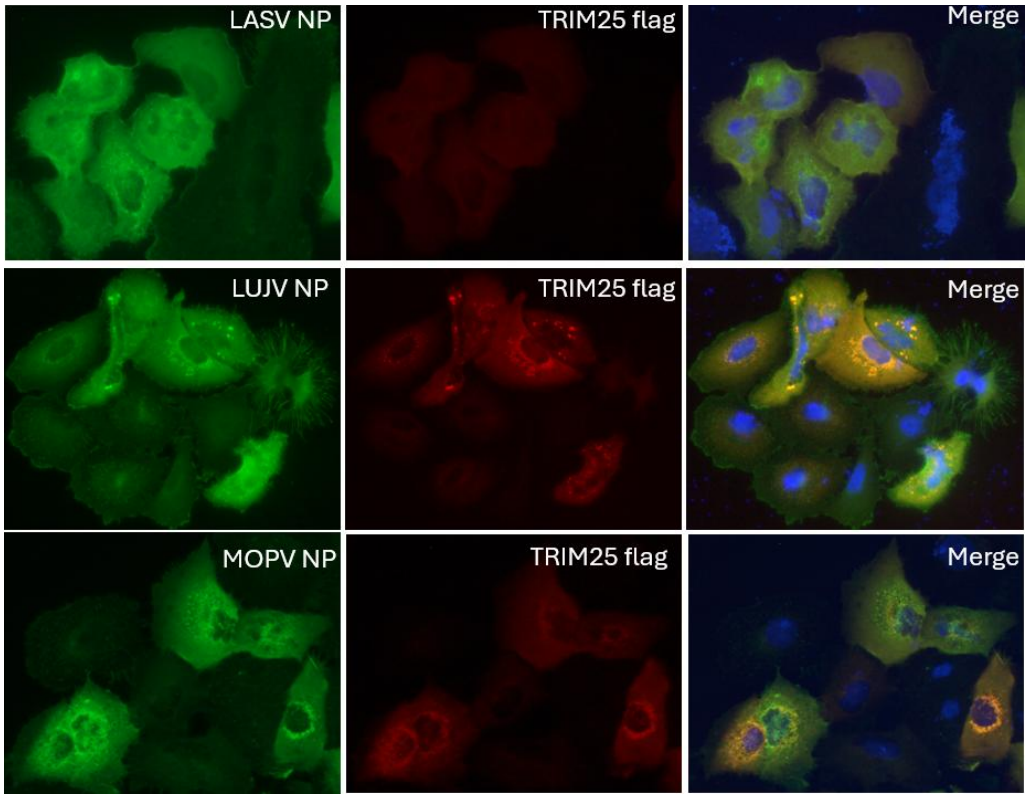


Figure 4.11 Immunofluorescence showing co-localisation of wild-type TRIM25-FLAG and either LASV-NP-HA, LUJV-NP-HA, or MOPV-NP-HA. A549 cells seeded onto a coverslip at a density of 50,000 per well of a 24 well plate. Cells were co-transfected with PCDNA3.1 wild-type-TRIM25-FLAG and either pCAGGs-LASV-NP-HA, pCAGGs-LUJV-NP-HA, or pCAGGs-MOPV-NP-HA at 250ng per construct per well of 24 well plate using Lipofectamine 3000 transfection reagent. Cells were then co-stained with mouse-anti-FLAG (Sigma F1804-200UG) and rabbit-anti-HA (Abcam Ab137838) at a dilution of 1:500 prior to co-staining with donkey raised-anti-mouse-594nm-fluorescent antibody (Abcam Ab150108) and goat-anti-rabbit-488nm (Abcam Ab150084) at a dilution of 1:500 to visualise the NP in green and TRIM25 in red. Slides were mounted using DAPI mounting solution to stain the cell nucleus for visualisation in blue colour. (n = 1)

4.6.5 Co-Immunoprecipitation Shows That TRIM25 Interacts with Both ZAP Isoforms

Whilst direct interaction between ZAP and TRIM25 has already been experimentally proven (Zheng et al., 2017), it was important to establish interactions in the cells lines we would be using and with the constructs designed. Figure 4.12 shows the results of a co-immunoprecipitation experiment confirming that both ZAP-L and ZAP-S isoforms directly interact with TRIM25. This supports the immunofluorescence evidence seen from co-localisations between the different ZAP isoforms and TRIM25 observed in Figure 4.10.

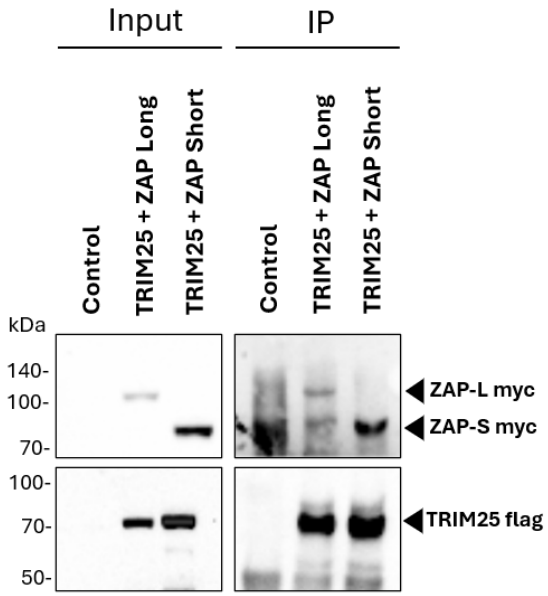


Figure 4.12 Western blot analysis for a co-immunoprecipitation experiment to investigate the interactions between both ZAP isoforms ZAP-L and ZAP-S and TRIM25. A549 cells were seeded at a density of 400,000 cells per well of a 6 well plate and subsequently co-transfected with a total concentration of 2ng DNA per well of either ZAP-L-myc or ZAP-S-myc and TRIM25-FLAG using Lipofectamine 3000 transfection reagent. Empty PCDNA3.1 was used as the control. 48 hours post transfection cells were washed once with ice cold PBS prior to lysis in a 50mM HEPES pH 7.2, 100mM NaCl, 1% IGEPAL buffer including a ROCHE protease inhibitor tablet. Lysis was left to occur whilst rolling at 4°C for 1 hour. During this time Protein G Pierce Magnetic beads were equilibrated in lysis buffer. Post-lysis, cells were centrifuged at 10,000rpm for 5 minutes at 4°C and 50µl supernatant added to 2x Laemmli buffer for the input sample and 950µl was added to 12.5µl neat beads per sample. Mouse-anti-FLAG antibody (SIGMA F1804-200UG) was added at a final dilution of 1:450 to each sample. Bead samples were left to bind whilst rolling overnight at 4°C. Samples were washed with lysis buffer 3 times and then 2 final washes in PBS before resuspension in 2x Laemmli buffer for Western blot analysis. Blots were stained for ZAP using a mouse-anti-myc (Sigma M5546-2ML) antibody, the other blot was stained for TRIM25 using a rabbit-anti-FLAG antibody (CST 14793S). All primary antibodies used at a dilution of 1:5000 and an anti-rabbit- or anti-mouse-secondary antibody was also used at a 1:5000 dilution. (n = 2)

4.6.6 Co-Immunoprecipitation Confirms That Arenavirus NP Interacts with Endogenous ZAP and TRIM25

The data from the previous chapter in which ZAP and TRIM25 were identified as interacting partners of LASV, LUJV, and MOPV NP was performed using BioID2 proximity labelling. This data indicates that ZAP and TRIM25 are in very close proximity to NP when pulled down and analysed by mass spectrometry. Whilst this indicates an interaction, it does not confirm a direct interaction between these proteins. To establish whether there is an interaction between ZAP and NP and TRIM25 and NP, a co-immunoprecipitation experiment was conducted, the results for which are shown in Figure 4.13. The results show that both endogenous ZAP and endogenous TRIM25 are pulled down with LASV, LUJV, and MOPV NP during the experiment which confirms an interaction between these proteins.

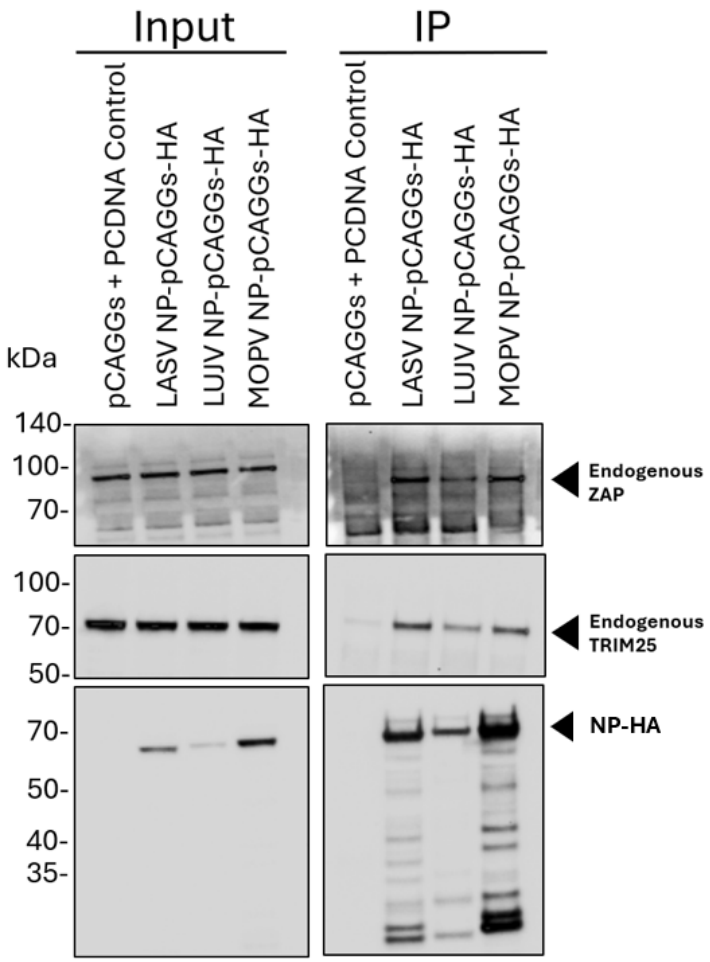
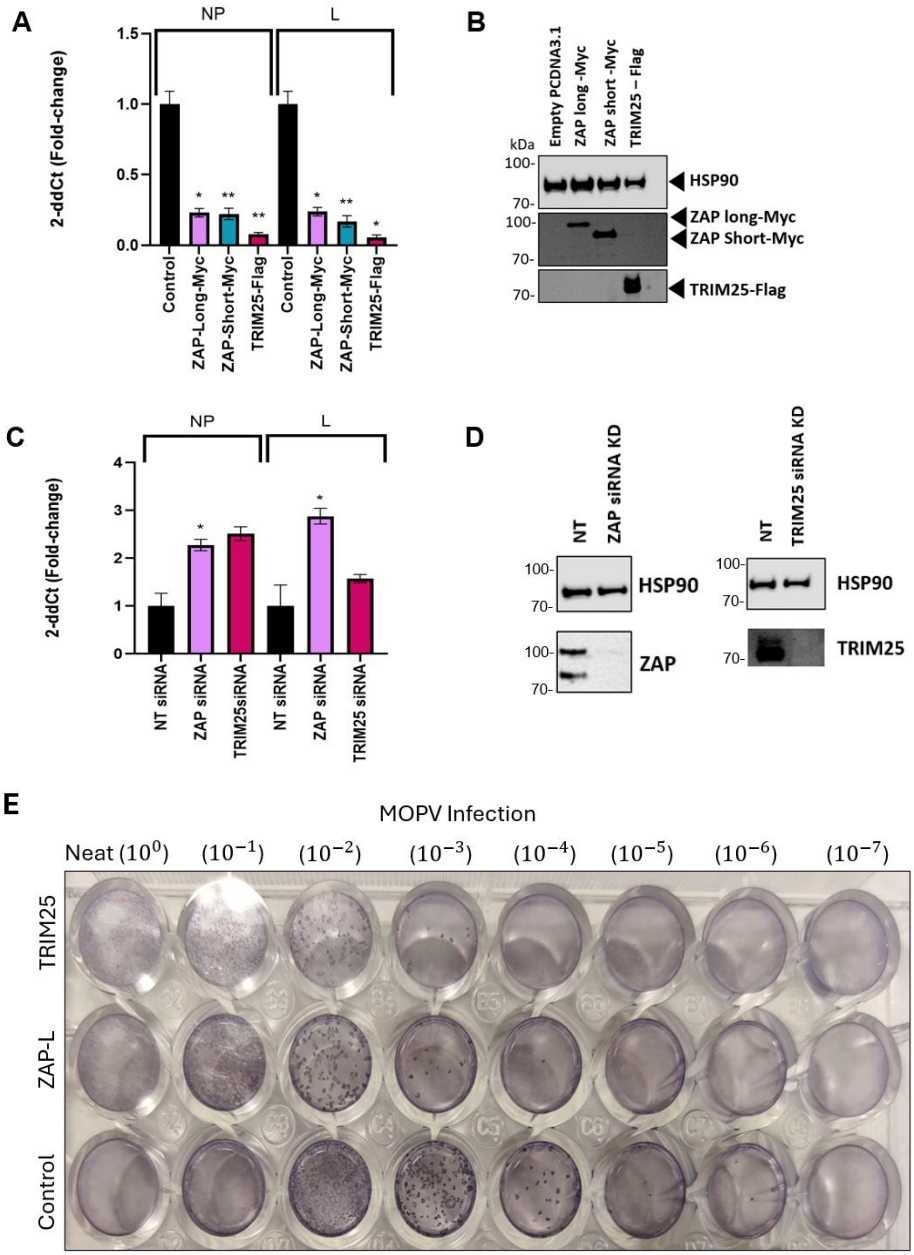


Figure 4.13 Western blot analysis for an immunoprecipitation experiment to investigate the interactions between different arenavirus strain NPs and endogenous ZAP or TRIM25. A549 cells were seeded at a density of 400,000 cells per well of a 6 well plate and subsequently transfected with a concentration of 2ng DNA per well of either LASV-, LUJV-, or MOPV-NP-pCAGGs-HA using Lipofectamine 3000 transfection reagent. 48 hours post transfection cells were washed once with ice cold PBS prior to lysis in a 50mM HEPES pH 7.2, 100mM NaCl, 1% IGEPAL buffer including a ROCHE protease inhibitor tablet. Lysis was left to occur whilst rolling at 4°C for 1 hour. During this time Protein G Pierce Magnetic beads were equilibrated in lysis buffer. Post-lysis, cells were centrifuged at 10,000rpm for 5 minutes at 4°C and 50µl supernatant added to 2x Laemmli buffer for the input sample and 950µl was added to 12.5µl neat beads per sample. Mouse-anti-HA antibody (Ab18181) was added at a final dilution of 1:150 to each sample. Bead samples were left to bind whilst rolling overnight at 4°C. Samples were washed with lysis buffer 3 times and then 2 final washes in PBS before resuspension in 2x Laemmli buffer for Western blot analysis. Blots were stained for endogenous ZAP using a rabbit-anti-ZAP (Ab105357) antibody, stained for endogenous TRIM25 using a rabbit-anti-TRIM25 antibody (), or stained for HA using a rabbit-anti-HA antibody (Ab137838). All primary antibodies used at a dilution of 1:5000 and an anti-rabbit-secondary antibody () was also used at a 1:5000 dilution. (n = 2)

4.6.7 Infection Assays Show Effectiveness of ZAP and TRIM25 as Inhibitors of Arenaviral Infection

To investigate how well TRIM25 and ZAP perform as arenavirus host restriction factors, various infection assays were performed using live MOPV. Experiments were initially carried out in normal A549 cells. Figure 4.14 shows the results from individually over-expressing either of the ZAP isoforms or TRIM25 in normal A549 cells. Over-expression of both ZAP-L and ZAP-S results in significant decreases in arenaviral replication as measured by the fold change differences for the arenaviral NP and L proteins from qPCR data. TRIM25 also exhibits potent anti-arenaviral characteristics, displaying even greater inhibition of MOPV replication than either of the ZAP isoforms.

To validate this, an inverse experiment was performed in which endogenous ZAP or TRIM25 was knocked down by siRNA in normal A549 cells infected with MOPV. Figure 4.14 shows that when the levels of ZAP or TRIM25 in a cell are decreased, arenavirus replication increases confirming that both ZAP and TRIM25 play important roles in the host antiviral defense against arenavirus infection.



F: Summary Table:

Condition	Dilution	Plaques	PFU/mL
Control	10^{-4}	28	**3.73×10^6**
TRIM25	10^{-3}	11	**1.47×10^5**
ZAP	10^{-3}	13	**1.73×10^5**

*Figure 4.14 Infection assays for overexpressed host proteins or siRNA knocked down host proteins. (A) A549 cells were transfected with 500ng of either ZAPL-MYC, ZAPS-MYC or TRIM25-FLAG or a control of empty PCDNA3.1 vector. These cells were subsequently infected with live MOPV at an MOI of 0.01 and harvested 72 hours later for RNA extraction and subsequent cDNA synthesis before analysis of their effects on viral replication by RT-qPCR looking at MOPV NP and L gene with GAPDH as our reference gene. (B) Success of the overexpressions and protein expression was confirmed by western blot staining for HSP90, MYC, or FLAG. (C) Knocked down expression of endogenous ZAP or TRIM25 in A549 cells using siRNA and Dharmafect protocol. The success of these knockdowns is shown in the western blot image. These A549 cells were infected with live MOPV at an MOI of 0.01 and harvested 72 hours post infection for RT-qPCR analysis. The stars in the graph represent a statistically significant difference in fold change from the control, where there are two stars this represents indicates a P value of 0.01 whereas a single start indicates a P value of less than 0.05. where no star exists, it indicates that the difference was not statistically significant. (n = 3). E) Plaque assay experiment comparing A549 cells transfected with either wild-type TRIM25, ZAP-Long, or empty PCDNA3.1 (Control) subsequently infected with MOPV. **Vero cells** were plated at **1.5×10^5 cells/well** in **48-well plates** and infected with **75 µL** of serially diluted **MOPV virus stock** (prepared in serum-free DMEM via 10-fold dilutions). After adsorption, cells were overlaid with **MEM containing 2.5% low-viscosity carboxymethylcellulose (CMC)** to restrict viral spread and incubated for **4 days at 37°C**. Plaques were visualized using **immunostaining**: cells were fixed and probed with **mouse monoclonal anti-Arenavirus rGPC antibody (clone KL-AV-1B3, 1:500 dilution; BEI Resources)**, followed by **anti-mouse alkaline phosphatase secondary antibody (1:750; Merck)** and developed with **BCIP/NBT substrate**.*

The creation of knockout cells lines was attempted for both ZAP and TRIM25. TRIM25 knockouts were successfully created but whilst initial knockouts for ZAP showed promise, these cell lines eventually recovered ZAP expression and can only be considered as having ZAP expression knocked down. This is displayed in the western blot image for ZAP expression in figure 4.15 in which bands corresponding to the ZAP S isoform are still present and visible indicating unsuccessful knockout. These cell lines were developed in normal A549 cells using a CRISPR-Cas9 system. These cell lines were used to repeat the infection assays from Figure 4.14 in a cell environment with knockout of endogenously expressed ZAP or TRIM25 rather than knockdown by siRNA. Figure 4.15 shows results from qPCR analysis of either ZAP or TRIM25 knockout cells infected with live MOPV. These results confirm the importance of both ZAP and TRIM25 as antiviral host restriction factors of arenavirus infection as when the cell is missing either of these proteins, MOPV replication is greatly enhanced.

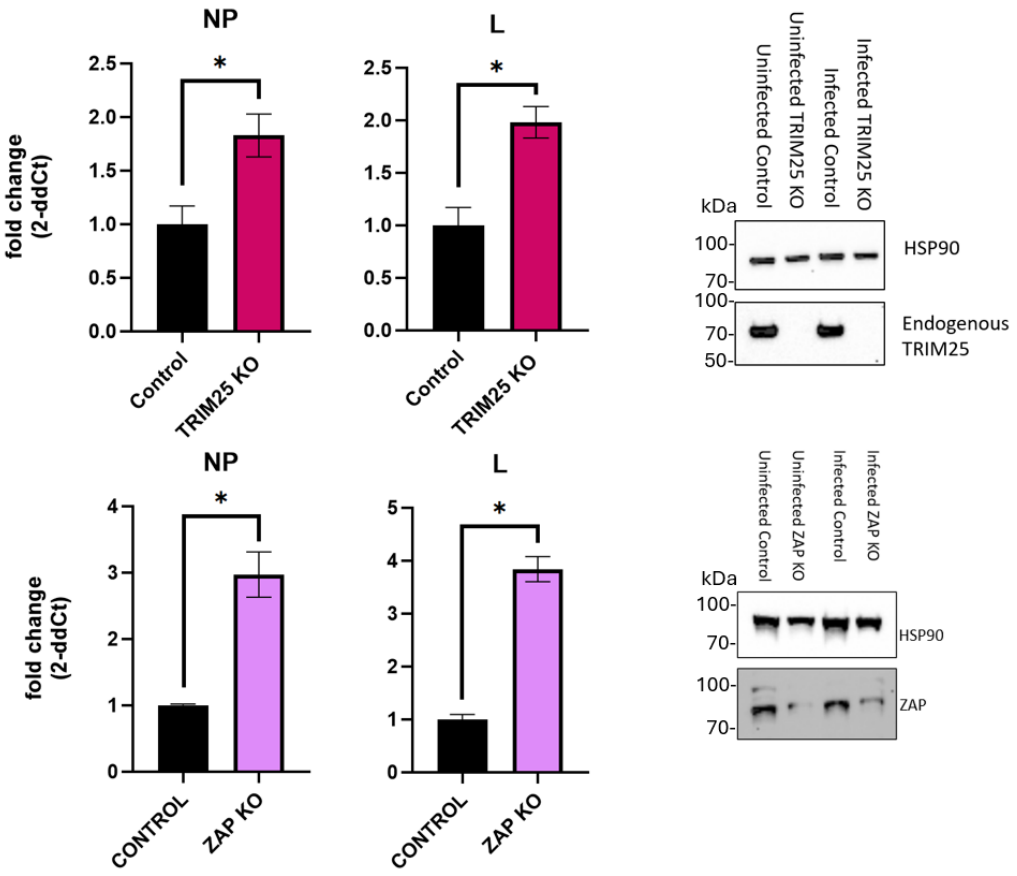


Figure 4.15 Infection assays for A549 knockout cell lines with either endogenous ZAP knocked down or TRIM25 knocked out were generated using a CRISPR-Cas9 system. These knockout/knockdown cells were infected with live MOPV at an MOI of 0.01 and harvested 72 hours later for RNA extraction and subsequent cDNA synthesis before analysis of how knocking out/down expression of these proteins' effects viral replication by RT-qPCR looking at MOPV NP and L gene with GAPDH and ACTIN as our reference genes. Overall cell health was determined by western blot and through staining for HSP90 using a mouse-anti-HSP90 antibody (Invitrogen MA1-10372). The success of the knockouts was determined using either an anti-rabbit-endogenous ZAP antibody (Abcam Ab154680) or an anti-mouse-endogenous-TRIM25 antibody (Abcam Ab610570). Either an anti-rabbit or anti-mouse secondary was used for staining prior to imaging. The presence of bands corresponding to ZAP S in both the uninfected and infected ZAP KO sample indicates that total knockout by CRISPR was unsuccessful and that this can only be considered a knock-down cell line. A band corresponding to the ZAP L isoform is almost visible in the infected ZAP KO sample indicating that the infection has stimulated ZAP expression which would not be possible in a true KO cell line therefore this must also be considered a ZAP L knockdown. The stars in the graph represent a statistically significant difference in fold change (p-value ≤ 0.05). (n = 3)

The availability of knockout cell lines allowed for improved investigations into the effectiveness of TRIM25 as an inhibitor of arenaviruses. By transfecting TRIM25 into cells in which endogenous TRIM25 expression was knocked out it means there can be no unintended interactions occurring between endogenous TRIM25 and transiently overexpressed TRIM25 and that we can be confident of the antiviral effect that TRIM25 has on MOPV arenavirus infection. Figure 4.16 shows the results of a MOPV infection

assay performed in TRIM25 knockout A549 cells in which TRIM25-flag has been transfected in at a concentration of 500ng using Lipofectamine 3000 transfection reagent and then infected with live MOPV at an MOI of 0.01 for 48 hours. The results are consistent with those from Figure 4.14A in which TRIM25 is a potent inhibitor of MOPV replication.

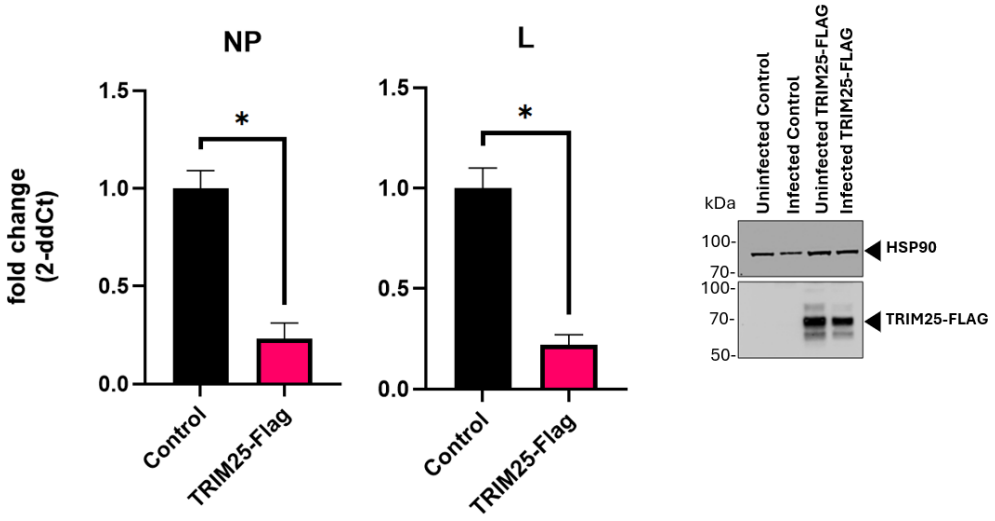


Figure 4.16 (A) Infection assays for A549 knockout cell lines with endogenous TRIM25 expression knocked out were generated using a CRISPR-Cas9 system. These knockout cells were transfected with either empty PCDNA3.1 as a control or PCDNA3.1 TRIM25-FLAG at a DNA concentration of 500ng per well of a 24 well plate using Lipofectamine 3000 transfection reagent. 48 hours post transfection; cells were infected with live MOPV at an MOI of 0.01 and harvested 72 hours later for RNA extraction and subsequent cDNA synthesis before analysis of their effects on viral replication by RT-qPCR looking at MOPV NP and L gene with GAPDH and Actin as our reference genes. (B) Success of the knockouts, transfections, cell health were determined by western blot staining for mouse-anti-HSP90 (Invitrogen MA1-10372) at a dilution of 1:5,000, rabbit-anti-endogenous-TRIM25 at a dilution of 1:5,000. The stars in the graph represent a statistically significant difference in fold change (p-value ≤ 0.05). (n = 3)

Both ZAP and TRIM25 are proteins coded for by interferon stimulated genes so investigation of the level to which arenavirus infection induces expression was performed. Figure 4.17 shows the results from densitometry analysis of ZAP-L, ZAP-S, and TRIM25 expression changes caused by both MOPV infection and treatment with differing levels of IFN-I. The results show that there are minor increases to the levels of ZAP-L and TRIM25 expression and also that ZAP-S expression appears to undergo a much greater increase in expression. This could suggest that the ZAP-S isoform is the important isoform for effective ZAP inhibition of arenaviral infection. It has been shown that different viruses are inhibited more

successfully by different ZAP isoforms depending on the virus characteristics and available ZAP targets. ZAP-L possesses a PARP-Like domain and targets CpG dinucleotide rich vRNA sequences to induce an antiviral response. Because of this, many viruses have evolved to have low frequency of CpG dinucleotides within their sequences to avoid detection by ZAP. Arenaviruses are among this subset of viruses with lower-than-expected CpG dinucleotide frequency as shown by analysis of sequences in Figure 4.4. It is possible and likely that the different ZAP isoforms came about to allow for targeting of non-standard viral epitopes to allow ZAP to continue to work as an effective host antiviral factor.

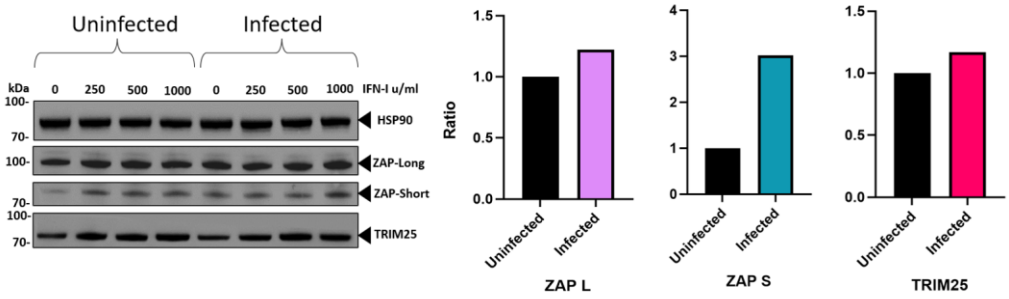


Figure 4.17 ZAP and TRIM25 induction due to infection. A549 cells were transfected with 500ng of either PCDNA3.1 ZAPL-myc, ZAPS-myc or TRIM25-FLAG or a control of empty PCDNA3.1 vector. 16 hours prior to infection, cells were pretreated with IFN-I at varying concentrations. These cells were subsequently infected with live MOPV at an MOI of 0.01 and re-spiked with the same IFN-I concentration and samples were harvested 72 hours later for analysis of protein expression and induction of ZAP-L, ZAP-S, or TRIM25 by western blot. Success of the overexpressions and protein expression was confirmed by western blot. Blots were stained for mouse-anti-HSP90 (Invitrogen MA1-10372), ZAP using a mouse-anti-myc (Sigma M5546-2ML) antibody, the other blot was stained for TRIM25 using a rabbit-anti-FLAG antibody (CST 14793S). All primary antibodies used at a dilution of 1:5000 and an anti-rabbit- or anti-mouse-secondary antibody was also used at a 1:5000 dilution. (n = 3)

Figure 4.18 also shows the results from an interferon assay performed in A549 cells with levels of endogenous ZAP or TRIM25 knocked down by siRNA. The results of this assay show that by decreasing cellular expression of these antiviral proteins, the IFN antiviral response to arenavirus infection was significantly attenuated. Figure 4.19 shows the results of a more robust version of the same experiment performed in A549 cells in which endogenous ZAP or TRIM25 had been knocked out by CRISPR-Cas9. The data from both Figures 4.18 and 4.19 allows us to characterise the antiviral

activity of ZAP and TRIM25 and clarifies that both proteins are strong inhibitors of arenavirus infection crucial to mediating restriction of MOPV and potentially of human pathogenic arenaviruses.

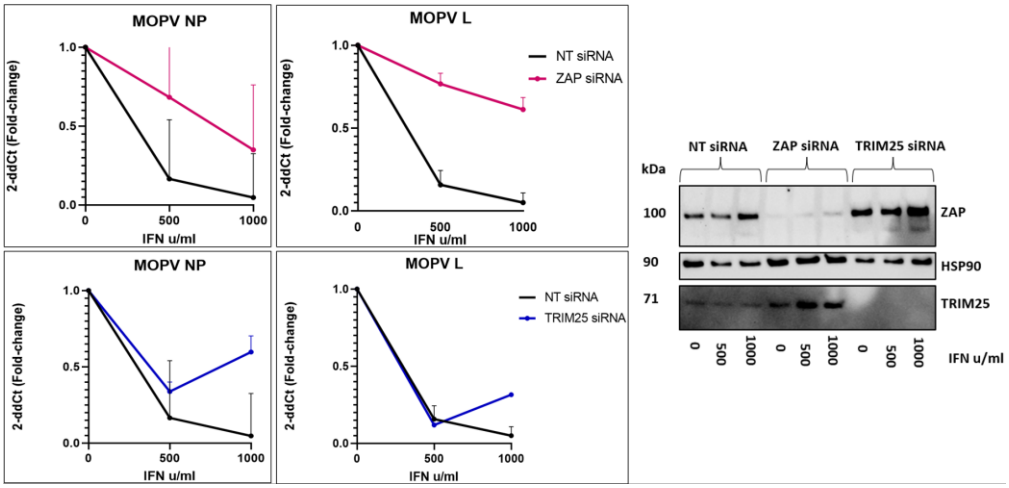


Figure 4.18 RT-qPCR data for IFN assays from siRNA knockdowns. Protein levels of HSP90, ZAP and TRIM25 were determined by western blot on A549 cells treated with non-targeting siRNA, and corresponding cells treated with either ZAP or TRIM25 siRNA at concentration of 100nm. Post Dharmafect transfection to knock down target gene, cells were pre-treated with IFN-I at increasing concentrations 16-hours prior to a 72-hour infection with live MOPV at MOI of 0.1. (n = 3)

Interestingly, whilst TRIM25 CRISPR-Cas9 knockouts have been extremely successful, maintaining knockout of ZAP has been difficult. The effects of interferon on ZAP expression can be seen in both Figures 4.18 and 4.19. Although only knocked down through siRNA in the experiment shown in Figure 4.18, we can see that increasing the concentration of IFN leads to detection of ZAP by western blot. This is not the same for TRIM25, inferring efficient knock out. Figure 4.19 shows that the ZAP knockout cells have already started show expression for ZAP-L, or endogenous, isoform although the ZAP-S isoform appears to stay knocked out. Another point of interest is when comparisons are made between the control samples with different levels of IFN treatment where it is evident that by increasing the concentration of IFN, we increase the induction of ZAP-S indicating it is a more IFN sensitive isoform than ZAP-L.

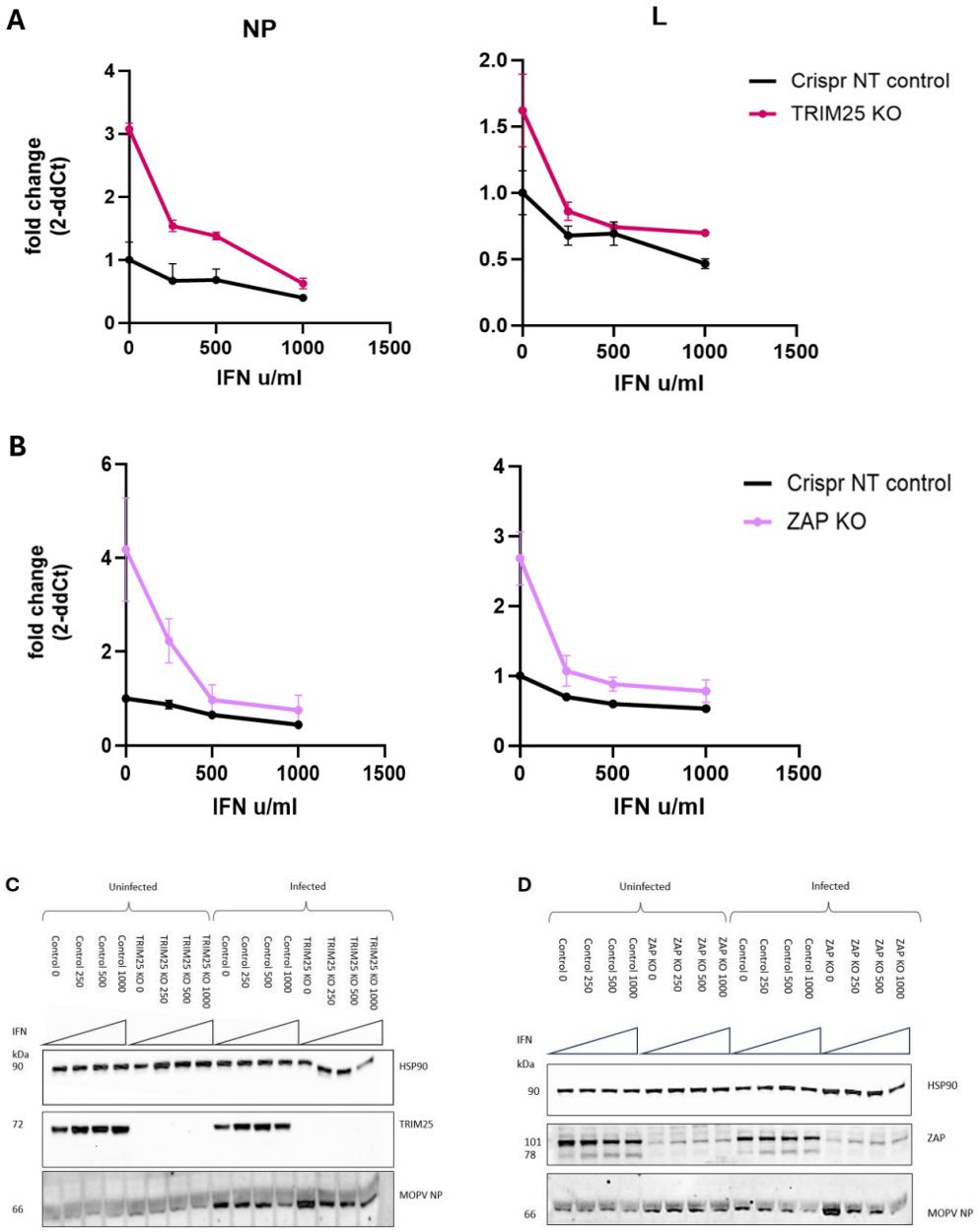


Figure 4.19 RT-qPCR data for IFN assays in knockout cell lines. Protein levels of HSP90, ZAP and TRIM25 were determined by western blot on A549 cells treated with non-targeting siRNA, and corresponding cells treated with either ZAP or TRIM25 siRNA at concentration of 100nm. Post Dharmafect transfection to knock down target gene, cells were pre-treated with IFN-I at increasing concentrations 16-hours prior to a 72-hour infection with live MOPV at MOI of 0.1. (n = 3)

4.6.8 TRIM25 Mutants

Four TRIM25 mutants were designed and cloned into PCDNA3.1 vector using overlap PCR: TRIM25-7KA, TRIM25-ΔRING, and TRIM25-ΔSPRY. Mutants missing these key TRIM25 domains were chosen to further investigate interactions with NP and deduce which TRIM25 domain is necessary for interaction with arenavirus NP. The TRIM25 7K motif is thought to be involved in RNA binding and substitution of all 7 lysines with alanine has been shown to decrease but not totally inhibit the ability of TRIM25 to bind RNA (Sanchez et al., 2018). It is likely that this retained ability of the 7KA mutant to bind RNA is due to the mutant still possessing an intact SPRY domain. The SPRY domain of TRIM25 has been shown to be necessary for many protein-protein interactions and it also has key functions in the ability of TRIM25 to bind RNA. Because of these roles the SPRY domain plays in the activity of TRIM25, a mutant protein missing this domain should have decreased effectiveness as an antiviral factor. Research has shown that TRIM25 interacts with a number of viral proteins via its SPRY domain including the SARS-CoV nucleocapsid protein (Choudhury et al., 2017). Because both arenavirus NP and the TRIM25 SPRY domain are crucial for respective RNA binding, it is possible that this domain is where TRIM25 and NP will interact, therefore experiments involving a -ΔSPRY TRIM25 mutant should shed light on the nature of these interactions and the mechanisms by which TRIM25 acts as an antiviral factor. The TRIM25-ΔRING mutant was selected due to the role that the catalytic RING domain plays in the synthesis of factors crucial to the ubiquitination activity of TRIM25. The RING domain dimerises to interact with Ub-conjugated E2 enzymes and has a crucial function in the k63-poly-Ub synthesis pathway (Sanchez et al., 2016). A TRIM25 mutant missing this RING domain would allow for investigation of the importance of TRIM25's ubiquitination activity to its role as an antiviral factor.

4.6.9 Cellular Localisation of TRIM25 Mutants

The immunofluorescence localisations from Figure 4.20 and the collated images in Figure 4.21 were performed in A549 TRIM25 knock out cells into which either wild type TRIM25-FLAG or the TRIM25 mutants were transfected. Figures 4.20 and 4.21 show that the TRIM25 mutants do localise differently to wild-type TRIM25. Whilst the localisation observed for wild type TRIM25 remains consistent with previous immunofluorescence as seen in Figure 4.9, the 7KA mutant appears to lose the cytoplasmic speckling and possesses a more diffuse distribution. The 7KA mutant disrupts the ability of TRIM25 to interact with RNA but does not completely abolish it (Sanchez et al., 2018). This diminished ability to bind RNA is likely altering the localisation of the 7KA mutant compared to wild type.

The ΔRING-TRIM25 mutant has localisation much more similar to wild type with the same speckling seen throughout the cytoplasm as well as a diffuse cytoplasmic distribution. This indicates that despite lacking the RING domain, this does not impact the localisation of TRIM25 within the cell.

The ΔSPRY-TRIM25 mutant does have different localisation to wild type as seen in Figures 4.20 and 4.21. Whilst there is still diffuse cytoplasmic distribution, the speckling appears to be much more concentrated around the cell nucleus. There is none of the speckling in the cytoplasm away from the nucleus that we observe in the wild type TRIM25 immunofluorescence. The SPRY domain of TRIM25 is involved in many protein-protein interactions as well as crucial to the RNA binding activity of TRIM25 and we can see that these functions also have an impact on where TRIM25 localises within the cell.

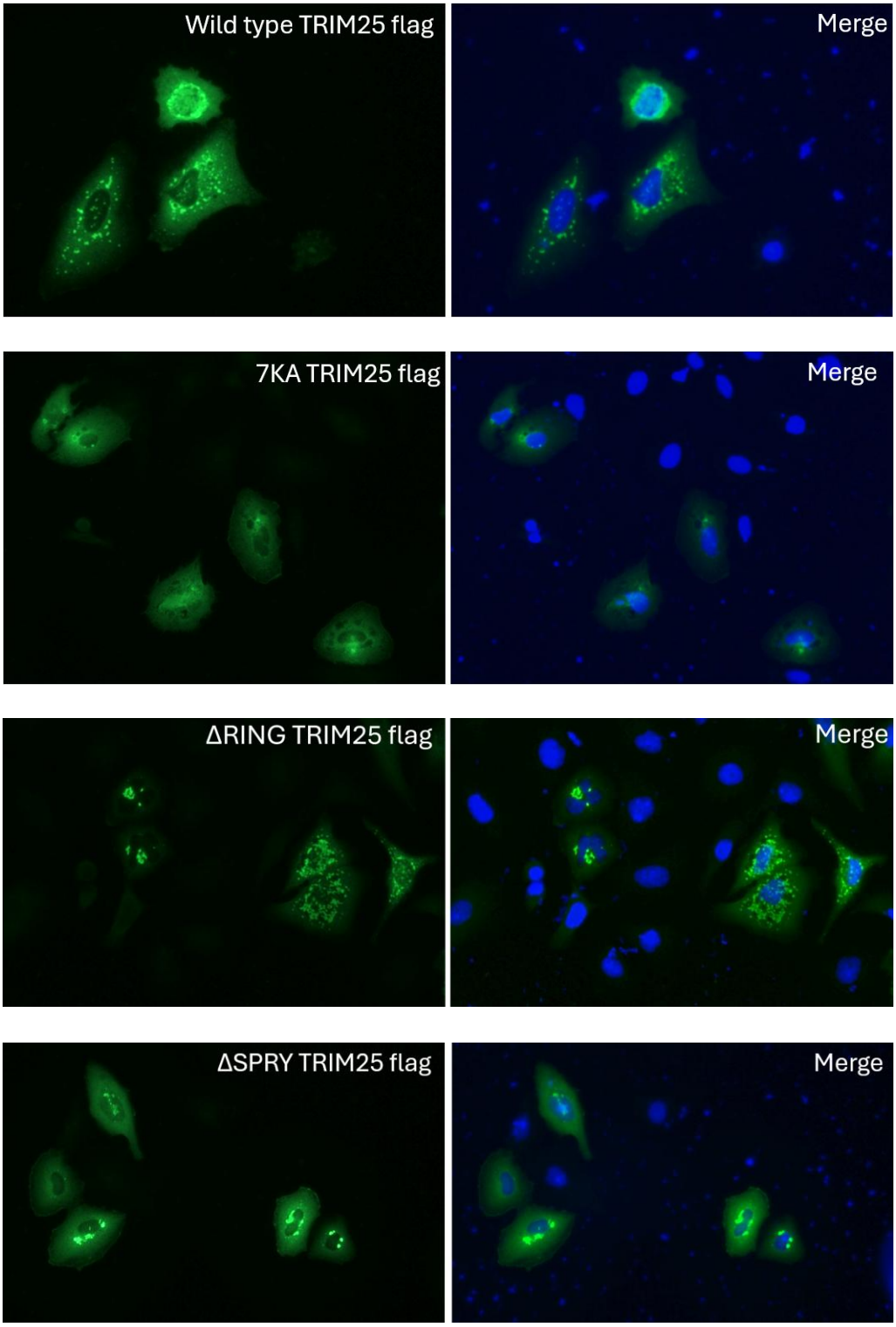


Figure 4.20 Immunofluorescence showing localisation of TRIM25 wild type and several mutants ; 7KA, ΔRING-domain, and ΔSPRY-domain. A549 cells seeded onto a coverslip at a density of 50,000 per well of a 24 well plate. Cells were transfected with PCDNA3.1 wild-type-TRIM25-FLAG or 7KA-TRIM25-FLAG or ΔRING-TRIM25-FLAG or ΔSPRY -TRIM25-FLAG at 500ng per construct per well of 24 well plate using Lipofectamine 3000 transfection reagent. Cells were then stained with mouse-anti-FLAG (Sigma F1804-200UG) at a dilution of 1:500 prior to staining with donkey raised-anti-mouse-488nm-fluorescent antibody (Abcam Ab150109) at a dilution of 1:500 to visualise the desired protein in a green colour. Slides were mounted using DAPI mounting solution to stain the cell nucleus for visualisation in blue colour. (n = 1)

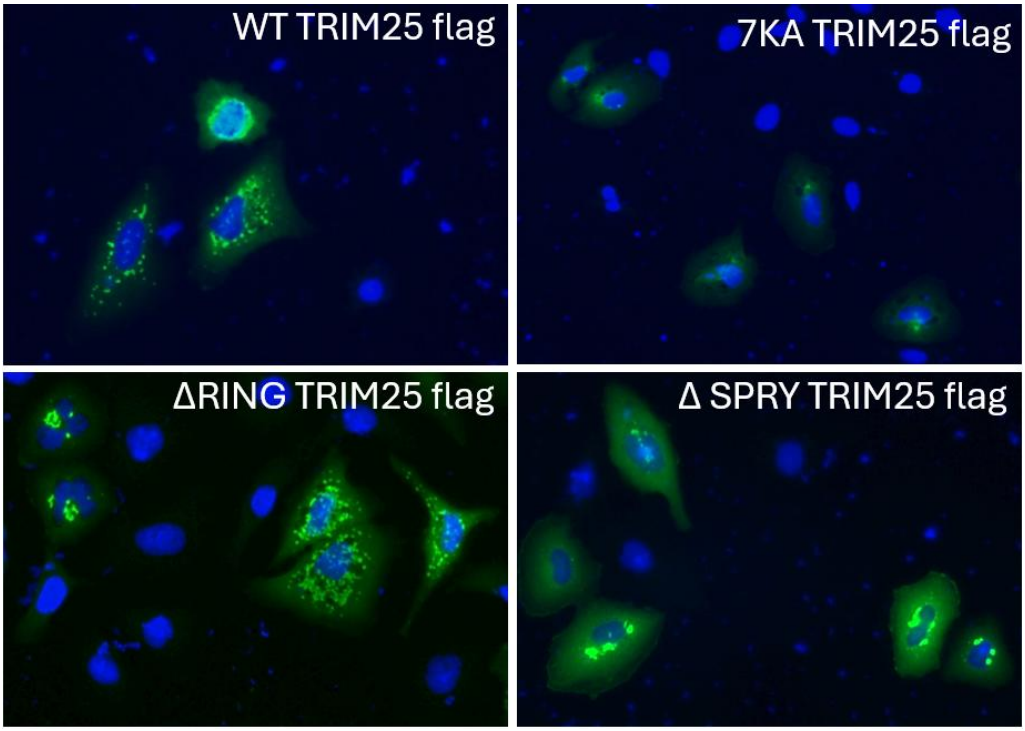


Figure 4.21 Collated immunofluorescence showing localisation of TRIM25 wild type and several mutants ; 7KA, ΔRING-domain, and ΔSPRY-domain. A549 cells seeded onto a coverslip at a density of 50,000 per well of a 24 well plate. Cells were transfected with PCDNA3.1 wild-type-TRIM25-FLAG or 7KA-TRIM25-FLAG or ΔRING-TRIM25-FLAG or ΔSPRY -TRIM25-FLAG at 500ng per construct per well of 24 well plate using Lipofectamine 3000 transfection reagent. Cells were then stained with mouse-anti-FLAG (Sigma F1804-200UG) at a dilution of 1:500 prior to staining with donkey raised-anti-mouse-488nm-fluorescent antibody (Abcam Ab150109) at a dilution of 1:500 to visualise the desired protein in a green colour. Slides were mounted using DAPI mounting solution to stain the cell nucleus for visualisation in blue colour. (n = 1)

An infection assay comparing the TRIM25 mutants was performed to identify which TRIM25 domain plays the most significant role in the role of TRIM25 as an anti-arenaviral host protein. The results of this assay are shown in figure 4.22; the ΔSPRY-TRIM25 mutant is immediately of interest as the level of viral replication for both NP and L are nearly at the same levels as seen in the TRIM25 knockout cells. This suggests that the ΔSPRY is crucial for TRIM25 to inhibit MOPV replication and likely plays a key role in the TRIM25 mediated anti-arenaviral response. The other TRIM25 mutants also show a reduced ability to inhibit MOPV replication as can be seen from the NP levels from the qPCR data in figure 4.22. Both the 7KA and ΔRING TRIM25 mutants show about a two-fold increase in NP levels when

compared to wild-type TRIM25. The 7KA mutant has been shown to have decreased ubiquitination activity when compared to wild type TRIM25 (Sanchez et al., 2018) and also to cause TRIM25 to have a diminished but not totally abolished ability to bind RNA. This decrease in both ubiquitination and RNA binding reported for the 7KA mutant explains the observed result in figure 4.22 in which the 7KA mutant still displays some anti-arenaviral activity but does not have the same potency as wild type TRIM25. The RING domain of TRIM25 is responsible for the E3 ubiquitin ligase activity. Thus, the ΔRING mutant would have an inhibited ability to target NP for ubiquitination and degradation although it would still have the capacity to bind the viral RNA and work as a ZAP cofactor to achieve antiviral activity. This perhaps suggests why it shows a diminished ability to inhibit MOPV replication compared to the wild type TRIM25 but still has some anti-arenaviral capability.

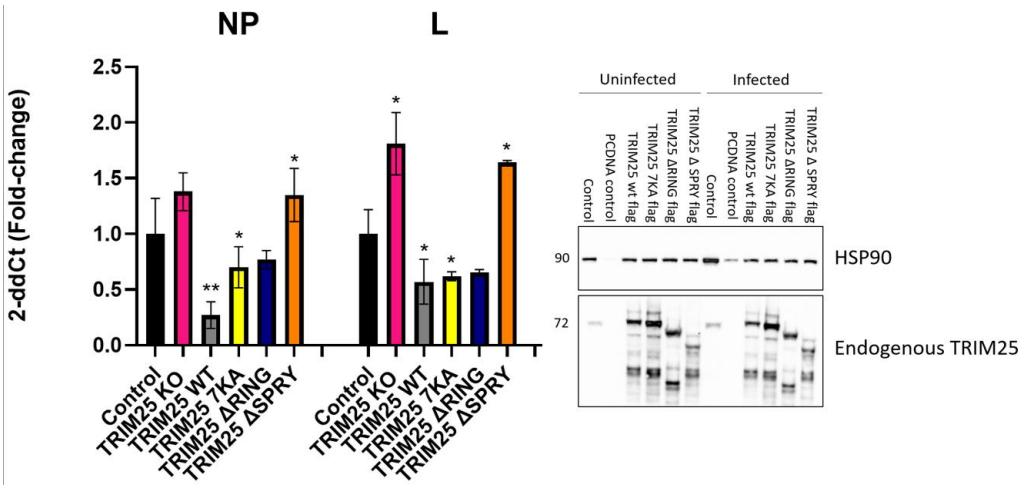


Figure 4.22 qPCR analysis and corresponding western blot of Wild-type TRIM25 and mutants 7KA, -ΔRING, and -ΔSPRY transfected into TRIM25 knockout A549 cells at a concentration of 500ng DNA per well of a 24 well plate using Lipofectamine 3000 transfection reagent. 24 hours post transfection; these cells were then infected with live MOPV at an MOI of 0.01. 48 hours post infection cells were harvested and samples taken for analysis by SDS-Page and western blot and for RNA extraction. Extracted RNA was then synthesised into cDNA which was then used for qPCR analysis of MOPV NP and L gene expression levels to determine the effect of TRIM25 wild type and mutants on MOPV infection. Western blotting was performed using either a mouse-anti-HSP90 antibody to stain for HSP90 expression as a measure of cell health, or a rabbit-anti-TRIM25 antibody. An anti-mouse and anti-rabbit secondary antibody were used respectively and blots were imaged using an iBright system (Thermo). Success of infection was determined by plaque assay of samples of infected cell media The stars in the graph represent a statistically significant difference in fold change from the control, where there are two stars this represents indicates a P value of 0.01 whereas a single start indicates a P value of less than 0.05. where no star exists, it indicates that the difference was not statistically significant. (n = 3).

Co-immunoprecipitation of MOPV NP with wild-type TRIM25 and the 7KA, ΔRING and ΔSPRY mutants was performed to investigate which domain is interacting between TRIM25 and NP (Figure 4.23). The interaction between the TRIM25 mutants which are missing the RING or SPRY domain is weaker than that between either the wild type or the 7KA mutant as shown by the comparative predominance of the wild type and 7KA bands for NP in figure 4.23. Given that less NP was immunoprecipitated with the ΔRING and ΔSPRY mutants suggests that the interaction between those mutants and NP is not as strong and that these domains likely have a role in the interaction between TRIM25 and NP and require further investigation. These results are in agreement with those from the infection assay in figure 4.22 in which the absence of the SPRY domain enhances MOPV replication.

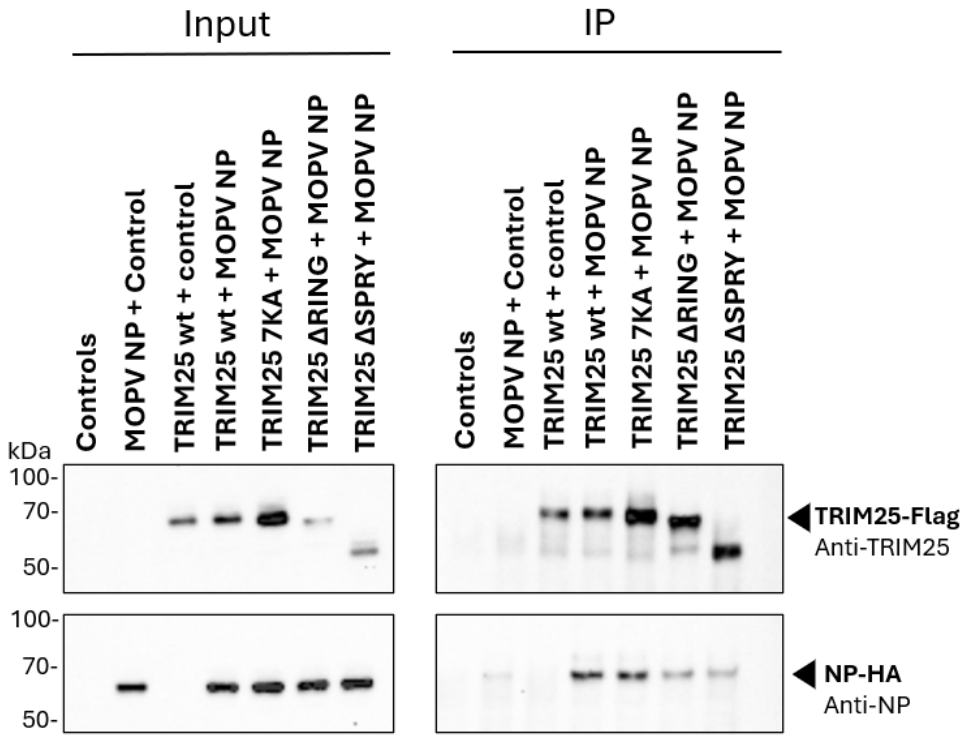


Figure 4.23 Co-immunoprecipitation experiment investigating interactions between wild type TRIM25 or mutants, and MOPV NP. The controls include co-transfected PCDNA3.1 and pCAGGs, the vectors into which MOPV NP and TRIM25 are cloned. MOPV NP was co-transfected with PCDNA3.1 and wild type TRIM25 with pCAGGs. A549 TRIM25 knockout cells were used for this experiment and co-transfection performed with Lipofectamine 3000 transfection reagent and protocol. Co-IP protocol was performed 48 hours post transfection using Pierce Protein G magnetic beads and pulling down with mouse-anti-Flag (Sigma F1804-200UG) at a dilution of 1:450) was performed. SDS-PAGE and western blot analysis was performed using a rabbit-anti-TRIM25 antibody (Ab167154) at a dilution of 1:5000, or anti-NP antibody at a dilution of 1:10,000. (n = 2)

4.7 Discussion

The findings presented in this chapter show that both ZAP and TRIM25 play important roles in the cellular defence against arenaviruses. Infection assays using MOPV have shown that TRIM25 appears to be a more potent inhibitor of viral replication. Whilst the ZAP-Short isoform shows increased induction due to infection with MOPV, infection assays do not show a significant difference in inhibition of viral replication between either isoform. Therefore, it is still unclear which isoform is the more potent restriction factor for viral replication and further work involving isoform specific CRISPR-Cas9 knockouts and over-expression repeats will need to be performed to elucidate this. This research has also validated that TRIM25, and ZAP interact within the cell through the implementation of immunofluorescence co-localisation experiments and co-immunoprecipitation. TRIM25 has been shown to co-localise to both the ZAP-Long and -Short isoforms and directly interacts with both. How the different isoforms affect the interaction of ZAP with TRIM25 and subsequent influence this may have on arenavirus infection has not been explored although provides an interesting idea for future research.

In future research, it may be necessary to perform proximity-labelling (both using transfected NP alone and also in the context of live viral infection) and proteomic analysis on all three arenavirus strain NPs to not only validate the MOPV NP results but confirm that TRIM25 and ZAP are interacting with LASV and LUJV NP as these are the human pathogenic strains for which a clearer understanding of pathogenicity is crucial. Analysis of TRIM25, ZAP, and NP by using functional mutants will also help characterise the key domains and their functions regarding arenaviruses.

A main limitation of the research presented was the availability of reliable antibodies against arenavirus NP. Despite numerous attempts at optimisation of both primary and secondary antibody dilution, incubation times, blocking time, the use of 5% milk vs BSA, reproducibility remained

difficult, and results were often inconclusive as to whether an infection had been successful. Often success of infection had to be determined through qPCR, and later by plaque immuno-assay. The anti-myc tag antibody used to assess success of ZAP isoform transfection and also used for ZAP isoform immunofluorescence also proved unreliable. This antibody worked as intended during early phases of the research, but subsequent batches of the same antibody often showed non-specific binding making western blot analysis difficult despite numerous attempts at optimisation. This was often circumvented through use of the reliable endogenous ZAP antibody although this did not solve the issue for immunofluorescence. ZAP was also difficult to knock out using CRISPR-Cas9. ZAP knockout cells would often show ZAP expression over time by western blot. For future studies, qPCR would need to be performed to detect significant reductions in mRNA expression compared to NT control for clonal cell line expansion.

In conclusion, this research has confirmed that both ZAP and TRIM25 are antiviral in the context of arenavirus infection, however, the exact mechanisms and specific biological and structural interactions remain largely unexplored. Investigation into the structural interactions between TRIM25 and NP are the subject of chapter 5. The experiments performed so far give many useful clues as to the nature of these interactions and help to inform future avenues of research. For example, this research has shown that there is a direct interaction between TRIM25 and NP through co-immunoprecipitation, and subsequently, the SPRY TRIM25 domain stands out as the main candidate for structural interaction analysis with NP. This is due to the greatly reduced ability of TRIM25 to act as an inhibitor of MOPV replication when the SPRY domain is missing as it is in this mutant. This mutant potently inhibits the ability of TRIM25 to bind RNA, a process crucial for the antiviral activity of TRIM25 (Álvarez et al., 2024). Research has shown that the PRYSPRY domain is predominantly responsible for TRIM25 RNA binding and that a region encompassing residues 470-508aa specifically binds RNA. Deletion of this region unfolds the PRYSPRY domain and inhibits

the RNA binding ability of TRIM25. Deletion of this region also results in a TRIM25 mutant unable to bind ZAP at all which would impact the overall antiviral activity of the ZAP-TRIM25 co-factor complex. Interestingly, the CC domain of TRIM25 (which also has roles in RNA binding) has been shown to be more promiscuous in its RNA binding sequence preference than the PRYSPRY domain (Álvarez et al., 2024). These findings indicate the importance of the SPRY domain in the anti-arenaviral response and indicate this may be the domain through which the TRIM25-NP interaction occurs, a hypothesis supported by the reduced level of interaction observed from the co-immunoprecipitation experiment between TRIM25 and MOPV NP.

Chapter 5

Elucidating the co-structure of TRIM25 and Arenavirus NP

5 Elucidating the co-structure of TRIM25 and Arenavirus NP

5.1 Chapter introduction

5.1.1 TRIM25

5.1.1.1 Introduction

Full length wild-type TRIM25 comprises 634 amino acids, for a total molecular weight of 70.6 kDa, and includes a RING domain, two B-box domains, a coiled-coil domain (CCD), and a PRYSPRY domain (B30.2 domain) (Figure 5.1). TRIM25 possesses two linker domains of currently unknown structure that separate both the RING and B-box domains and the coiled-coil and PRYSPRY domains. Thus far, the full length TRIM25 protein structure has not been successfully determined via experimental techniques including X-ray crystallography, however, the PRYSPRY domain, CCD domain and CC-PRYSPRY domain structures have been determined by X-ray crystallography (Koliopoulos et al., 2018).

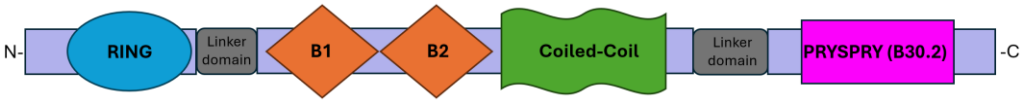


Figure 5.1 Schematic of TRIM25 protein structure and domains. TRIM25 is comprised of a RING domain connected to two B-box domains by a linker. The B-box domains are followed by a Coiled-coil domain which is connected to the PRYSPRY (otherwise known as B30.2) domain at the C-terminal end of the protein by another linker domain.

5.1.1.2 PRYSPRY

The structure of the PRYSPRY domain (Figure 5.2) is comprised of a hydrophobic core with thirteen β-strands forming two antiparallel β-sheets in a configuration known as a "bent β-sandwich" (Koliopoulos et al., 2018). The structure contains two α-helices located at the N-terminal end of the PRYSPRY packed against one face of the β-sheet core with α-helix 1 (α1)

running near parallel and α2 near perpendicular to the direction of the β-strands. The PRY section is formed by residues 440-512aa and the SPRY section by residues aa513-633 and together form a single structural module.

The TRIM25 PRYSPRY structure elucidated by (D'Cruz et al., 2013) revealed eight additional residues at the N-terminus of the PRY region compared to earlier B30.2 domain structures. These newly resolved residues form an extra α-helix (α1), which engages in hydrophobic interactions with the adjacent β-sheet while displaying polar residues on the solvent-exposed surface. In particular, Leu446, Phe449, and Leu450 from α1 participate in hydrophobic interactions with Phe592, Ile594, and Phe596 from β10, as well as Leu604 and Lys607 from β11. As a result, α1 protects a hydrophobic patch on the underlying β-sheet from solvent exposure. The total solvent-accessible surface area provided by the three-dimensional structure of the TRIM25 B30.2 domain is 8,850 $Å^2$ and this difference in volume is due to the extra alpha helix.

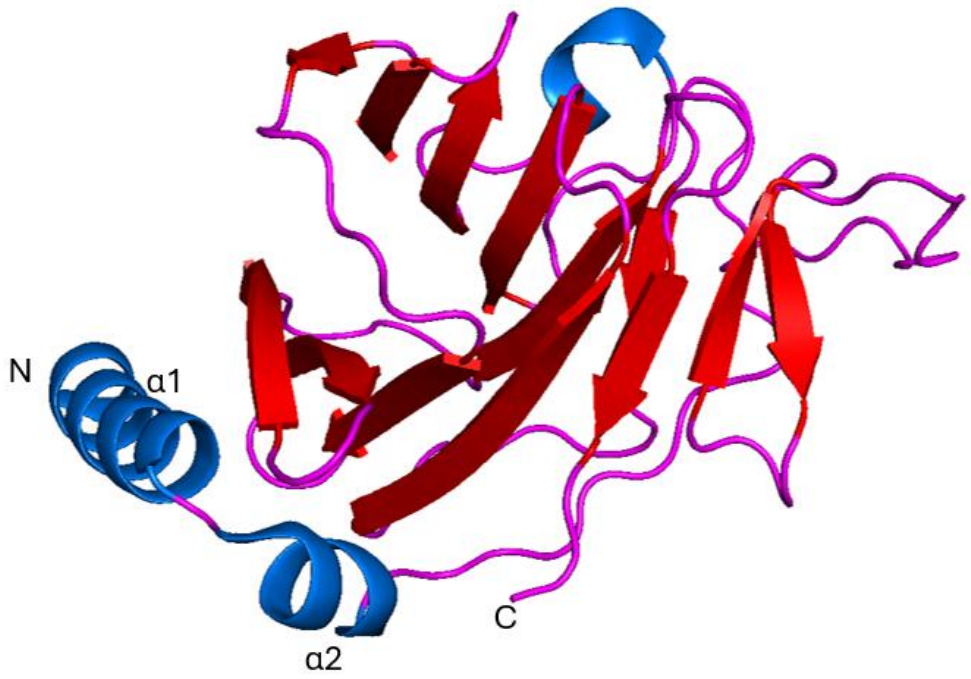


Figure 5.2 Cartoon schematic of TRIM25 PRYSPRY domain with colouring depending on structure; Helices are blue, ribbons are pink, and sheets are red. The N- and C-terminal ends of the PRYSPRY TRIM25 fragment have been labelled as have the α-helices α1 and α2. Schematic visualised using PyMOL.

5.1.1.3 TRIM25 Coiled-coil domain

The human TRIM25 coiled-coil (CC) domain is formed by residues 189-379aa and has had its structure determined by (Sanchez et al., 2014) (Figure 5.3). The 189-379aa residue includes the entire coiled-coil region and the N-terminal half of the second linker region that connects the CC and PRYSPRY domains. Analysis of the structure revealed that this construct is a stable dimer, and crystallisation determined the structure to 2.6Å and showed that the asymmetric unit was comprised of a single, elongated dimer which was about 17nm long. Each subunit of the symmetric dimer was found to fold back into a hairpin configuration possessing both short and long arms. The linker region and the coiled-coil residues are structurally distinct with the coiled-coil residues forming the long arm of each subunit (α1 helix coloured cyan in figure 5.3) and the linker region residues forming the short arm shown to fold back and sit against the α1 helix (α2, α3, and an irregular loop region all coloured orange in figure 5.3). The two subunits dimerise in an antiparallel manner and almost all of the hydrophobic sidechains within the structure are involved in packing interactions occurring across the entire length of each hairpin and burying a total surface area of 5,102Å^2. The coiled-coil/linker domain polar and charged side-chain also forms many hydrogen bonding and salt bridge interactions (Sanchez et al., 2014).

Originally, the TRIM family of proteins were predicted to contain two distinct coiled-coil domains that were separated by a helical but non-coiled? segment (Javanbakht et al., 2006, Sardiello et al., 2008). The TRIM25 CC structure revealed by (Sanchez et al., 2014) shows that the segments described actually form a single contiguous coil with the α1 helix forming the long arm of each subunit. The mechanism by which the interaction occurs between the α1 helix of each subunit underpins why the coiled coil

forms a dimer. The interaction is mediated by knobs-into-holes packing of both heptad repeats and hendecad repeats which produce a supercoil and these hendecads mediate interactions occurring at the centre of the coiled-coil and an amphipathic platform is created by super helical underwinding which causes the α1 helix of each subunit to arrange side by side. The terminal α3 helix from the short arm of each hairpin has been shown to pack tightly against one side of the platform which forms a bundle of four helices.

Research has shown that hydrophobic residues located at the centre of the coiled-coil domain are required for TRIM25 dimerisation (Sanchez et al., 2014). This central region seems to be necessary for dimer stability and has a high density of intermolecular interactions due to the tight packing of the α1 helices against each-other and also against the α3 helix.

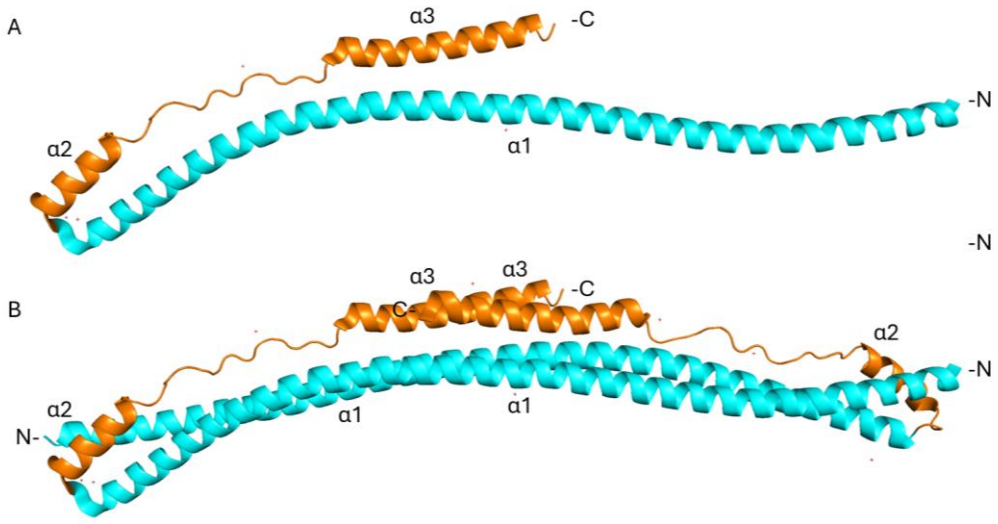


Figure 5.3 Cartoon representations of TRIM25 Coiled-coil domain with colouring depending on secondary structure elements, Long helices are cyan, whilst the shorter helices and the connecting loops region are orange. The N- and C-terminal ends of the fragment have been labelled as have the α-helices. A) shows the monomeric CC domain. B) Shows the dimeric CC domain. Schematic visualised using PyMOL.

5.1.1.4 TRIM25 CC-PRYSPRY

Research by (Koliopoulos et al., 2018) into the molecular mechanisms by which influenza A NS1 mediates TRIM25 recognition and inhibition resulted in identification of the structure of the human TRIM25 CC-PRYSPRY domain (residues 189-630aa) at a resolution of 3.6Å. (Koliopoulos et al., 2018) solved the structure using molecular replacement using the CCD structure already available (PDB:4LTB) and a de novo determined structure of the PRYSPRY domain of TRIM25 residues 434-630aa at a resolution of 2Å (Figure 5.4). It was found that the CC-PRYSPRY structure has one and a half CC dimers in an asymmetrical unit and a crystallographic two-fold that completes the second dimer. Each of the bow-shaped CC dimers has two PRYSPRY domains that are bound near the end of which maintains the twofold symmetry. The two bound PRYSPRY domains do not significantly alter the CC structure when compared to CC alone. There is a linker region connecting the CC and PRYSPRY domains which is 73 amino acid residues in length (residues 361-433aa), but it is not visible in the electron density of the solved structure so the exact connectivity between the CC and PRYSPRY is still unknown.

The impact of the PRYSPRY domain upon the CC spans across amino acid residues 269-287aa on the α1 helix and residues 320-326aa on the extended connection between α2 and α3, all on the same monomer. These CC domain regions contact the PRYSPRY domain surface loops that span across amino acid residues 460-464aa (loop 1), residues 472-476aa (Loop 2), residues 488-494 (Loop 3), and residues 504-506 (loop 4). Research performed by small angle X-ray scattering on the CC-PRYSPRY and NMR titrations between the separate CC and PRYSPRY domains revealed information about the domain dynamics. Interaction between the two domains was shown to be weakened through mutations of Y463S and Y476S (Koliopoulos et al., 2018). When these residues were mutated, RIG-I-2CARD ubiquitination capabilities were almost completely lost which shows the importance of this interface for correct PRYSPRY positioning

necessary for substrate ubiquitination. Interestingly, mutation of these residues does not affect the overall fold of the PRYSPRY domain. These results, along with other recent research (D'Cruz et al., 2018), suggest that when alone, the PRYSPRY domain is in a dynamic equilibrium between a bound and a constrained diffusing state tethered to the CC via the flexible linker region but the bound state is critical for ubiquitination of substrates.

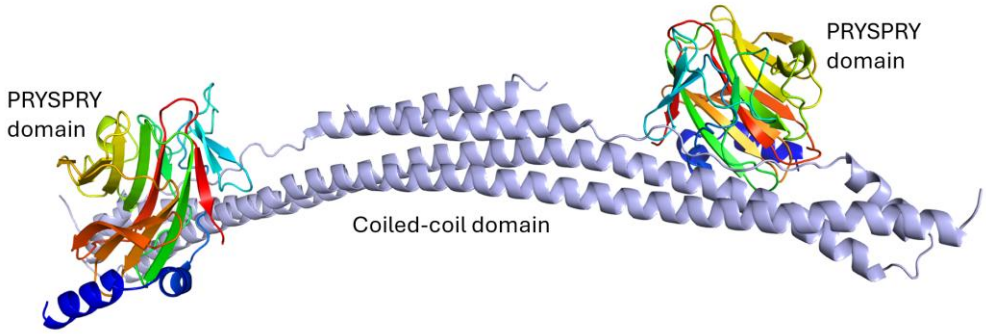


Figure 5.4Cartoon schematic of TRIM25 Coiled-coil-PRYSPRY overall final structure determined experimentally by (Koliopoulos et al., 2018). Colouring is dependent on structure domain; the helices making up the CC domain are grey and the PRYSPRY domain has been coloured using the PyMOL rainbow function which colours each segment of the cartoon molecule differently. This has been done to make visualisation of the distinct domains and where they interact clear. Schematic visualised using PyMOL

5.1.2 LASV NP Crystal Structure

The monomeric full-length crystal structure of LASV NP displays a protein possessing two domains separated by a 30 amino acid flexible linker domain (Figure 5.5). This linker domain has remained structurally uncharacterised because of its intrinsic flexibility. Both protein domains are in tight contact with each other. The LASV NP N-terminal domain is a globular α/β domain comprised of 14 α-helices and 6 β-strands which are able to undergo substantial structural conformation changes. Within all NP structures, the first 50 amino acids making up the first two α-helices (α1 and α2) are involved in the multimerization mechanism are folded over the NP core domain (Ortiz-Riaño et al., 2012b). Originally, NP was observed to be trapping NTP within the N-terminal domain suggesting a gating

mechanisms that allows access to a potential cavity (Qi et al., 2010b) the function of which was mis-identified as involved in cap-binding to provide host-derived primers for initiation of transcription by the viral polymerase. Later research cleared up this mis-conception and showed that the NP-cap-binding domain actually corresponds to the NP RNA binding site and the true cap-binding domain was located within the C-terminal part of the arenavirus L protein (Hastie et al., 2011c, Morin et al., 2010, Lehmann et al., 2014, Reguera et al., 2016, Vogel et al., 2019). In arenavirus NP, the RNA binding site is covered by two α-helices (α5, and α6) and a loop (coloured as orange in figure 5.5 C). The RNA binding site alters structural conformation in the presence of ssRNA and the α5 and α6 helices create a cleft to accommodate the RNA with the α6 helices being observed to have up to 7 different positions when RNA is bound (Hastie et al., 2011c). The creation of the cleft due to RNA binding causes the C-terminal domain to reposition although this has not yet been structurally characterised.

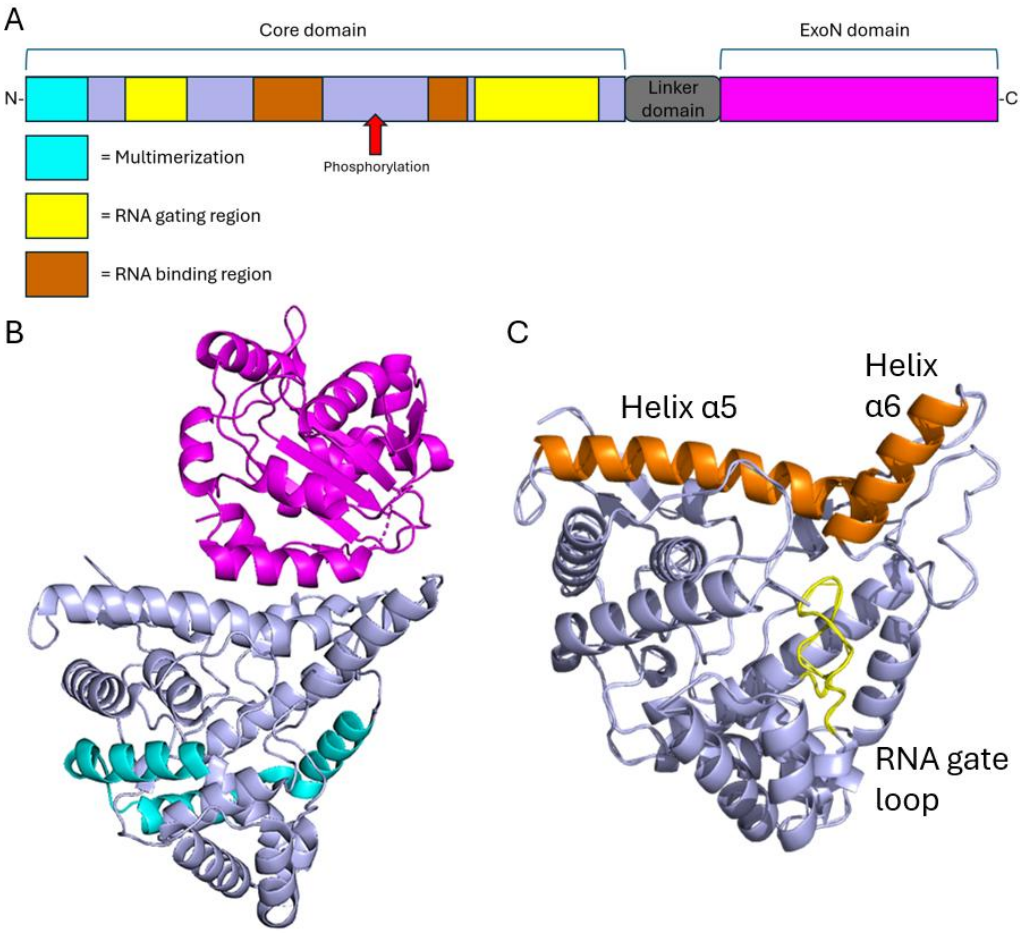


Figure 5.5A) Schematic of LASV NP structure. The colours correspond to the domains in the protein cartoons. At the N-terminal end of LASV NP there is a domain comprised of two α-helices involved in multimerization followed by another region of two α-helices and a loop region involved in RNA gating. The N-terminal core domain of LASV NP is connected to the C-terminal Exonuclease domain by a linker domain. B) Cartoon representation of the LASV NP protein highlighting the N-terminal multimerization in cyan and the C-terminal ExoN domain in pink. C) cartoon representation of LASV NP which highlighting in orange the α-helices and loop region involved in RNA gating mechanisms.

Arenavirus NP contains a 3'-5' exonuclease domain within its C-terminal region that has specificity to dsRNA and which functions to interfere with the IFN-I immune response (Martínez-Sobrido et al., 2009). It has been shown that this C-terminal exonuclease domain isn't required for viral genome replication or transcription in most arenavirus strains although in LASV, PICV, and LCMV NP mutated to lack exonuclease activity this resulted in greatly impaired replication (Huang et al., 2015b, Carnec et al., 2011).

The 3'-5' exonuclease domain (ExoN) is a member of the DEDDH family which is a family of nucleases known to process and cleave ssRNA or DNA through a two-metal-ion catalytic mechanism. Although LASV NP utilises

its 3'-5' exonuclease to cleave dsRNA. This two-metal-ion mechanism involves the two ions being coordinated by residues of the ExoN motif (Steitz and Steitz, 1993). The ExoN domain has a fold characteristic of the DEDDH family comprised of two β-sheets with two antiparallel strands and six mixed strands, and eight α-helices connected via a series of loops. The structural arrangement of these secondary structures form a central β-sheet surrounded on one side by three α-helices and seven on the opposite side with a zinc binding site, all of these structure usually have one metallic ion present in the catalytic site with the second metallic ion, which allows the catalytic reaction to initiate, comes from the RNA to be cleaved by the ExoN (Jiang et al., 2013, Zhang et al., 2013, Hastie et al., 2012). Within the LASV NP ExoN there are two flexible regions; the basic loops (residues 514-526aa) which sometimes structures itself as two anti-parallel strands located above the active site and the C-terminal arm (residues 549-570aa).

5.1.3 Arenavirus NP core domain and research focus

The decision to focus on the NP N-terminal core domain for interaction studies with TRIM25 was made because of the role this domain has in RNA binding, a characteristic shared with the TRIM25 CC-PRYSPRY domain. Similarly, the TRIM25 CC-PRYSPRY domain was chosen because not only is it involved in the RNA binding activity of TRIM25 but the CC-PRYSPRY domain has been shown to be the domains through which TRIM25 interacts with many other proteins, including viral, and recognises substrates (Koliopoulos et al., 2018). Examples of this include TRIM25 interactions with Influenza A NS1 in which interaction occurs at the TRIM25 CC domain and can disrupt formation of the CC-PRYSPRY complex formation subsequently inhibiting TRIM25 antiviral activity through RIG-I pathway induction (Koliopoulos et al., 2018). The SARS-CoV-2 N protein has also been shown to target TRIM25 to inhibit RIG-I activation and suppress innate immune antiviral responses. The interaction between the SARS-CoV-2 N

protein and TRIM25 has been shown to occur as the SPRY domain of TRIM25 (Gori Savellini et al., 2021).

Whilst published crystal structures for full length LASV NP and the TRIM25 CC-PRYSPRY construct exist, thus far no crystal structures for full length LUJV or MOPV NP have been determined experimentally. The ExoN domain of MOPV has been determined in complex with various metal ions, (Yekwa et al., 2017, Nguyen et al., 2022) but not the N-terminal core domain. Due to this lack of published experimental structural data, predictive modelling with techniques such as Alphafold3 must be utilised to gain insight into the potential structures of these NP domains. Figure 5.6 B and C depict the Alphafold3 predictions for the LUJV- and MOPV NP N-terminal core domains respectively. The colouring in the figure has been chosen to ease differentiating and comparing the tertiary and overall structures of the different NP N-terminal core domains. Figure 5.6 A shows the LASV NP N-terminal core domain fragment from the solved and published crystal structure (3MWP on PDB) (Qi et al., 2010b). Comparisons between the protein cartoon models shows that the main differences are in the location of the α-helix coloured in bright green which appears to be helix-α6 as depicted in orange in Figure 5.5 C. As expected due to the high genetic similarity between LASV and MOPV, there is little difference in the overall structure except for a difference in length of the helix in dark blue. LUJV displays a much different structure for the helix-α6. Whilst this difference is a possibility that it is due to limitations in the Alphafold3 prediction of very flexible regions such as the loops connecting the helix-α6, it has been shown in LASV NP that this helix-α6 has a different conformation when bound to RNA (Hastie et al., 2011c). This conformation closely matches that seen in the LUJV structure which is also likely to have different conformations to enable RNA binding.

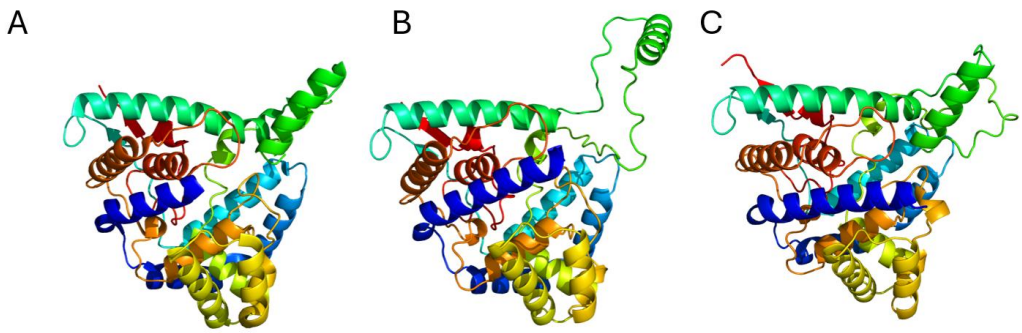


Figure 5.6 Cartoon representations of Arenavirus NP N-terminal core domain fragments visualised in PyMOL. A) LASV NP (PDB: 3MWP (Qi et al., 2010b)) B) LUJV NP C) MOPV NP

Chapter Objectives:

This chapter aimed to:

1. Structurally characterize the molecular interaction between arenavirus nucleoprotein (NP) and TRIM25.
2. Purify the specific protein domains hypothesized to mediate this interaction: the TRIM25 CC-PRYSPRY domain and the NP N-terminal core domain.
3. Define the molecular envelopes of individual protein domains and the NP-TRIM25 complex using SEC-SAXS.
4. Validate the structural models by fitting AlphaFold3 predictions and available crystal structures to the experimentally determined SAXS envelopes.
5. Integrate computational modelling with experimental data to elucidate the binding interface and structural basis of TRIM25-mediated restriction of arenavirus replication.
6. Establish a structural framework for understanding how TRIM25 recognizes and binds arenavirus NPs across different viral strains (LASV, LUJV, and MOPV).

5.2 Results

5.2.1 TRIM25 CC-PRYSPRY and LASV NP N-terminal fragment complex prediction using Alphafold3

AlphaFold analysis of the TRIM25 CC-PRYSPRY construct and the LASV NP N-terminal core domain (Abramson et al., 2024) predicted protein-protein interaction to occur at the TRIM25 PRYSPRY domain and at the C-terminal end of the NP core domain (Figure 5.7).

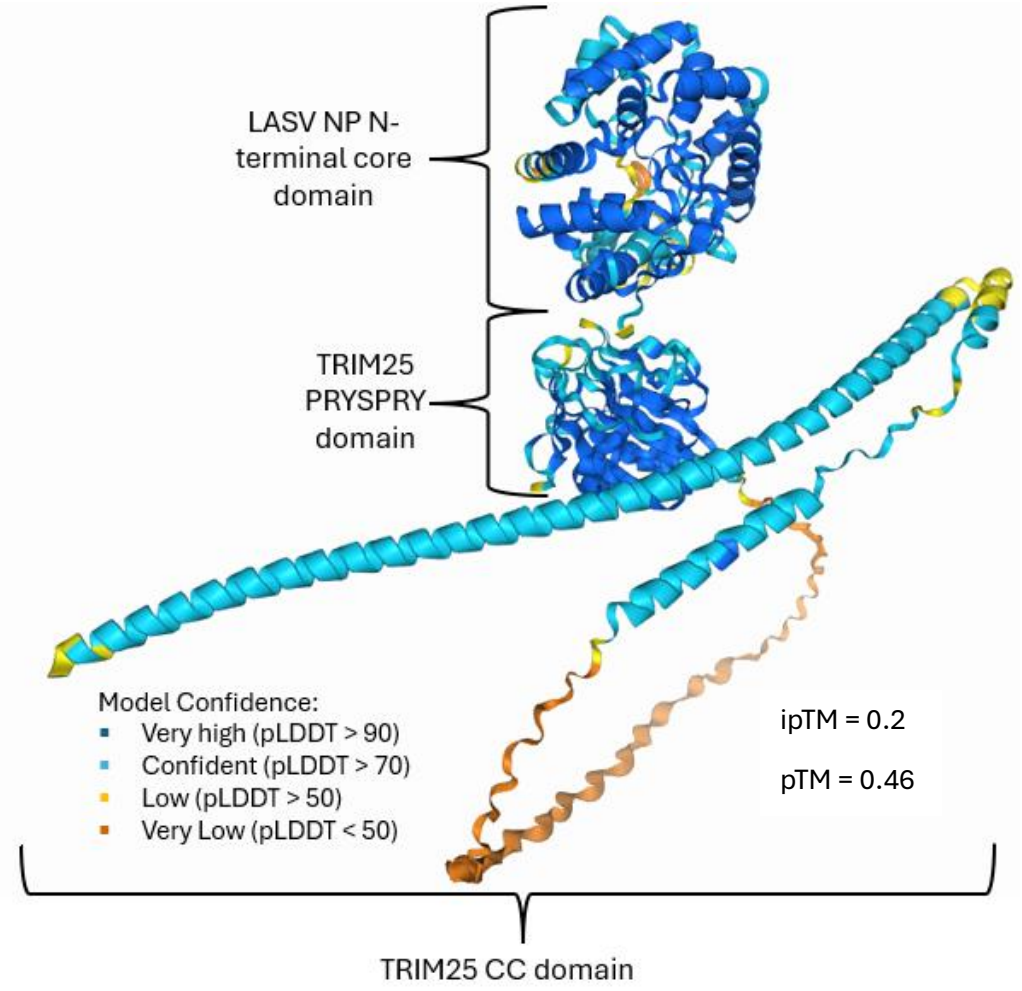


Figure 5.7 Alphafold3 prediction of complex formation between LASV NP N-terminal region residues 1-340aa and the TRIM25 CC-PRYSPRY domain residues 189-630aa. pLDDT is the predicted local difference test and is a per-reside measure of local confidence on a scale from 1-100. Higher pLDDT scores indicate higher confidence and this often correlates to a more accurate prediction (Abramson et al., 2024). the predicted template modelling (pTM) score and the interface predicted template modelling (ipTM) score are both derived from a measure called the template modelling (TM) score. This measures the accuracy of the entire structure, A pTM score above 0.5 means the overall predicted fold for the complex might be similar to the true structure. ipTM measures the accuracy of the predicted relative positions of the subunits within the complex. Values higher than 0.8 represent confident high-quality predictions, while values below 0.6 suggest likely a failed prediction. The ipTM score for this prediction = 0.2 and the pTM = 0.46.

Based on the predicted pose, the RNA binding region of NP faces the TRIM25 PRYSPRY domain (Figure 5.8) and that the C-terminal end of the NP is stretching towards the TRIM25 PRYSPRY. The CC-domain is not involved in any direct interaction with NP based on this prediction.

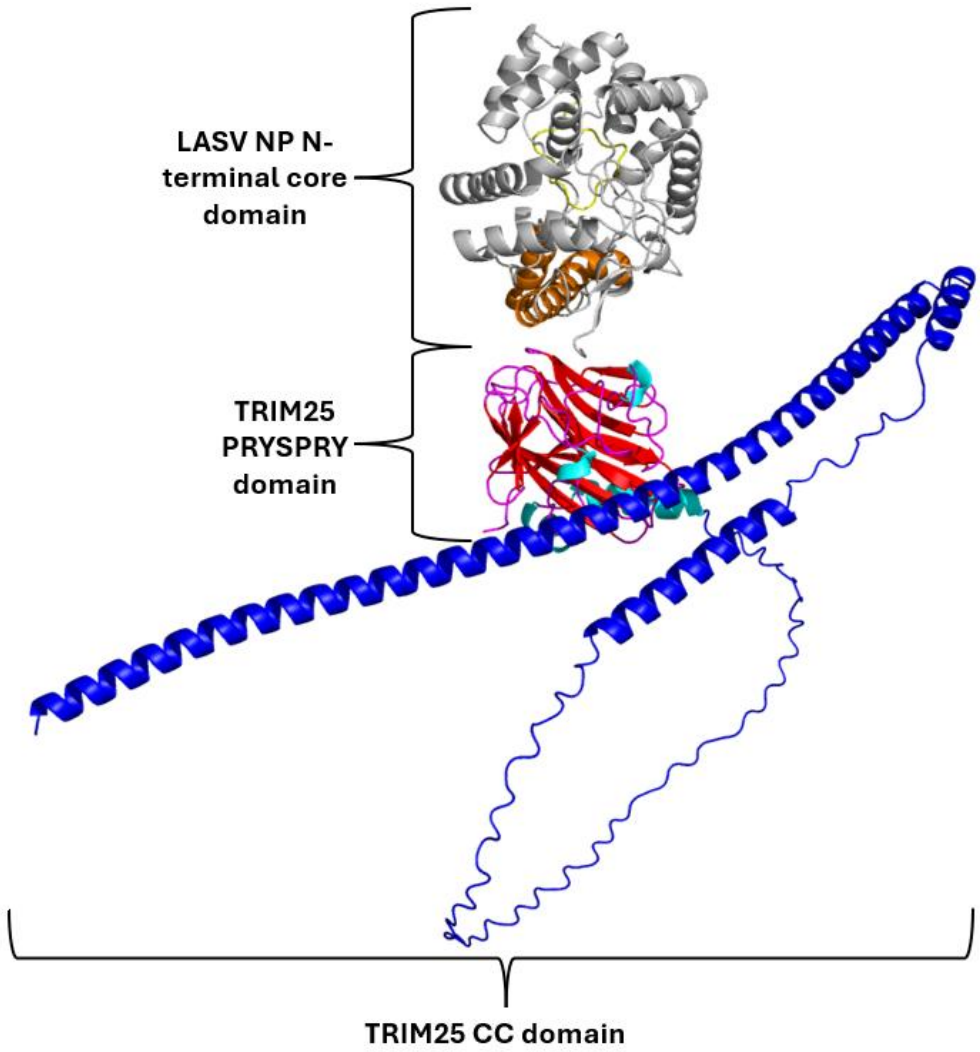
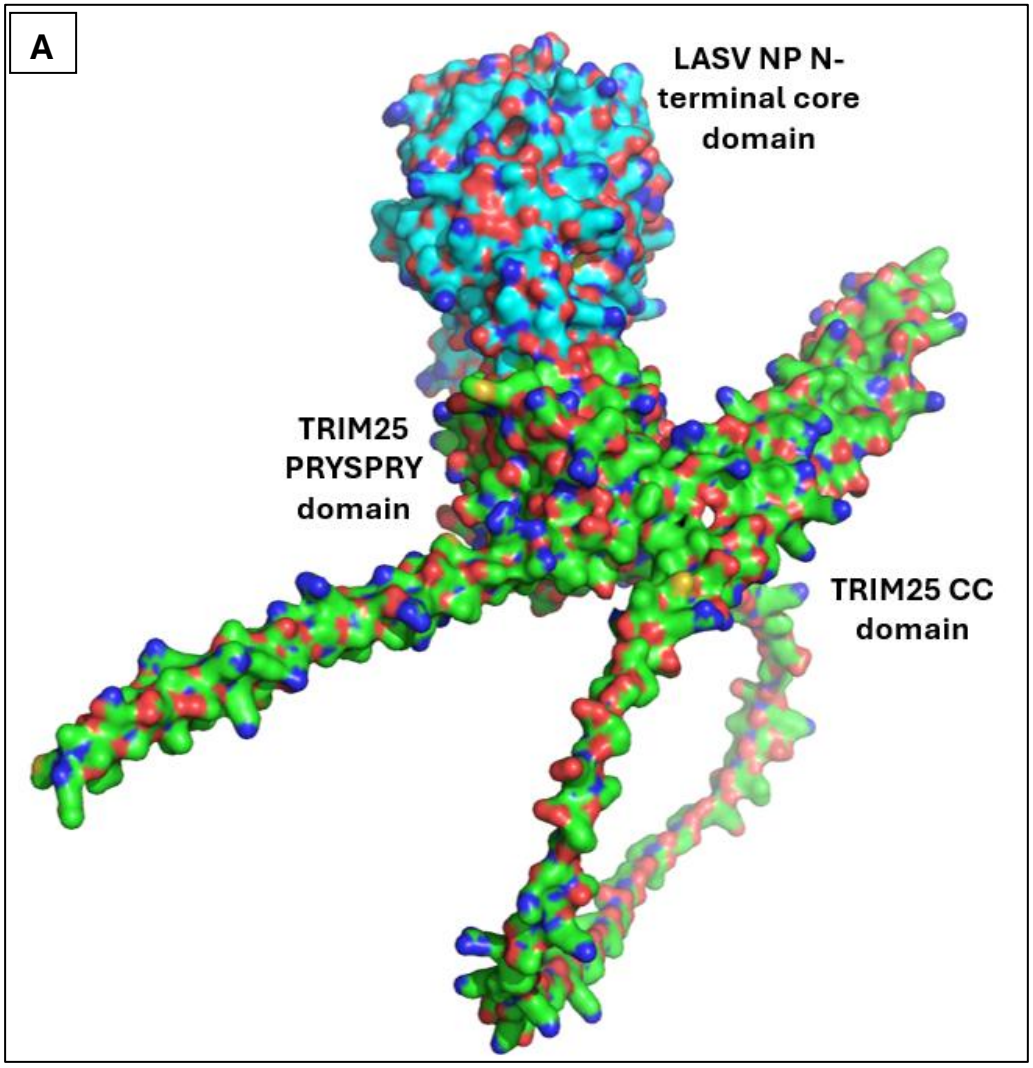
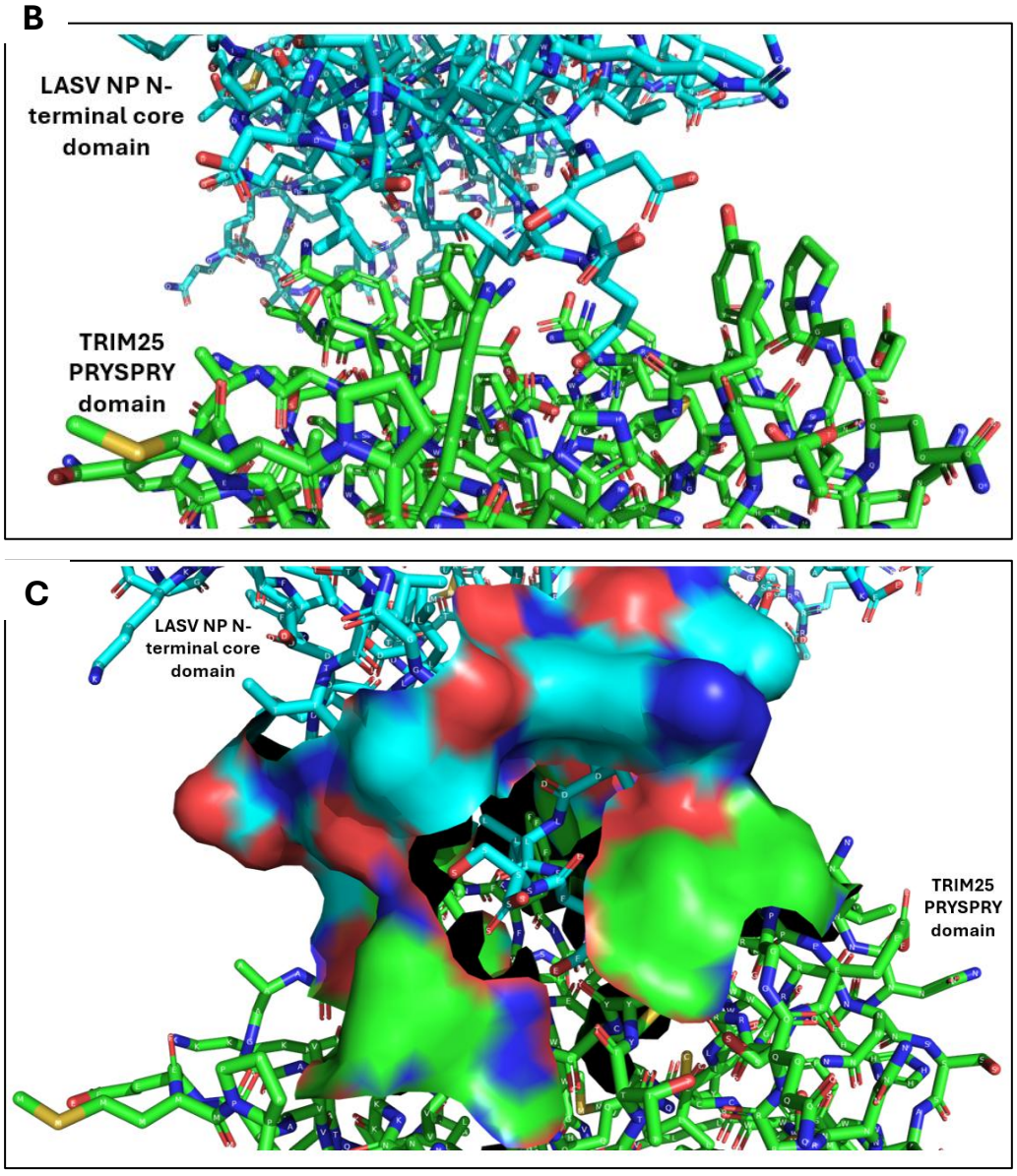


Figure 5.8 Alphafold3 prediction of complex formation with distinct colouring. Complex formation between LASV NP N-terminal region residues 1-340aa and the TRIM25 CC-PRYSPRY domain residues 189-630aa. The LASV NP N-terminal core domain follows the same colouring in figure 5.5 with the helices involved in RNA binding coloured orange and the RNA gating loop in yellow and the rest of the protein in grey. TRIM25 PRYSPRY has had the loops coloured magenta, the α-helices in cyan and the β-sheets in red. TRIM25 CC domain has been coloured in dark blue. This was done in PyMOL.

A closer view of the predicted protein-protein interaction can be seen in Figure 5.9. Although, it may appear as though only the C-terminal end of the NP fragment was involved in the predicted interaction with the TRIM25 PRYSPRY domain, Figure 5.9 A shows that when the surface structure is

visualised, the N-term domain of NP binds into TRIM25 PRYSPRY domain pocket. .The RNA binding region of the NP N-terminal core domain is shown to be involved in the interaction with the PRYSPRY domain. Utilising the ability to visualise protein bonds that PyMOL offers gives a more detailed picture of the nature of the protein-protein interaction occurring. Figure 5.9 D shows which amino acid residues are involved in this predicted interaction.




B
LASV NP N-terminal core domain
TRIM25 PRYSPRY domain
C
LASV NP N-terminal core domain
TRIM25 PRYSPRY domain

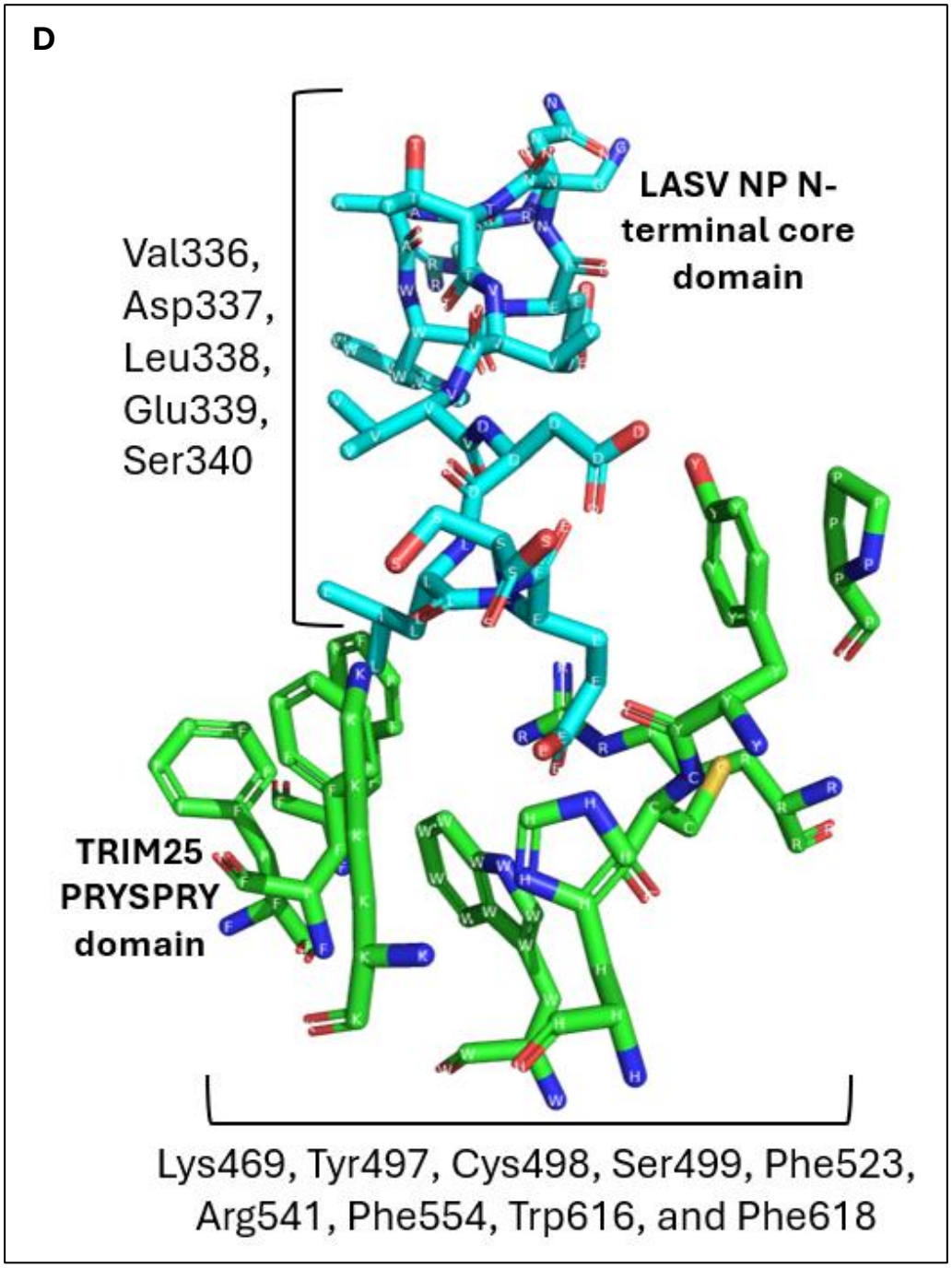


Figure 5.9 Alphafold3 prediction of complex formation showing interaction specifics. Complex between LASV NP N-terminal region residues 1-340aa and the TRIM25 CC-PRYSPRY domain residues 189-630aa. A) This depicts the PyMOL predicted surface structure of the NP-TRIM25 complex. B) amino acid side chains have been shown in PyMOL C) Surface structure prediction showing TRIM25 binding pocket and C-terminal end of NP core domain sitting in pocket. D) Just the amino acid residues represented as sticks showing the binding pocket of TRIM25 in which NP appears to insert

From the Alphafold3 prediction it appears as though the LASV NP N-terminal core domain binds to TRIM25 PRYSPRY in a pocket formed by amino acid residues Lys469, Tyr497, Cys498, Ser499, Phe523, Arg541, Phe554, Trp616, and Phe618. The region of LASV NP which appears to be bound in this pocket is a 5aa long chain at the C-terminal end of the domain spanning residues Val336, Asp337, Leu338, Glu339, and Ser340. It also appears that the LASV NP N-terminal core domain RNA binding region is involved in the predicted complex interaction point as seen in the surface model in Figure 5.9. Whilst LASV NP binds within a pocket in the TRIM25-PRYSPRY domain, the outer regions also interact to bind the proteins together. The outer region of the LASV NP binding to the outside of the pocket formed by TRIM25 is made up of the amino acid residues, Iso89, Arg91, Val92, Ser97, Asp98, Leu101, Glu108,and Lys111 and the surface structure for these is shown in Figure 5.9 C in cyan.

5.2.2 Full length TRIM25 and LASV NP complex prediction using Alphafold3

Based on the predicted pose for the interaction between the full-length proteins, the interaction occurring between LASV NP and TRIM25 is via the CC domain and the N-terminal core domain (Figure 5.10). Interestingly, the PRYSPRY domain is not involved in any direct interaction with NP in this prediction unlike that for the fragment interaction between the NP N-terminal core domain and the TRIM25 CC-PRYSPRY fragment. The RNA binding region of NP also appears not to be involved in the interaction based on this prediction.

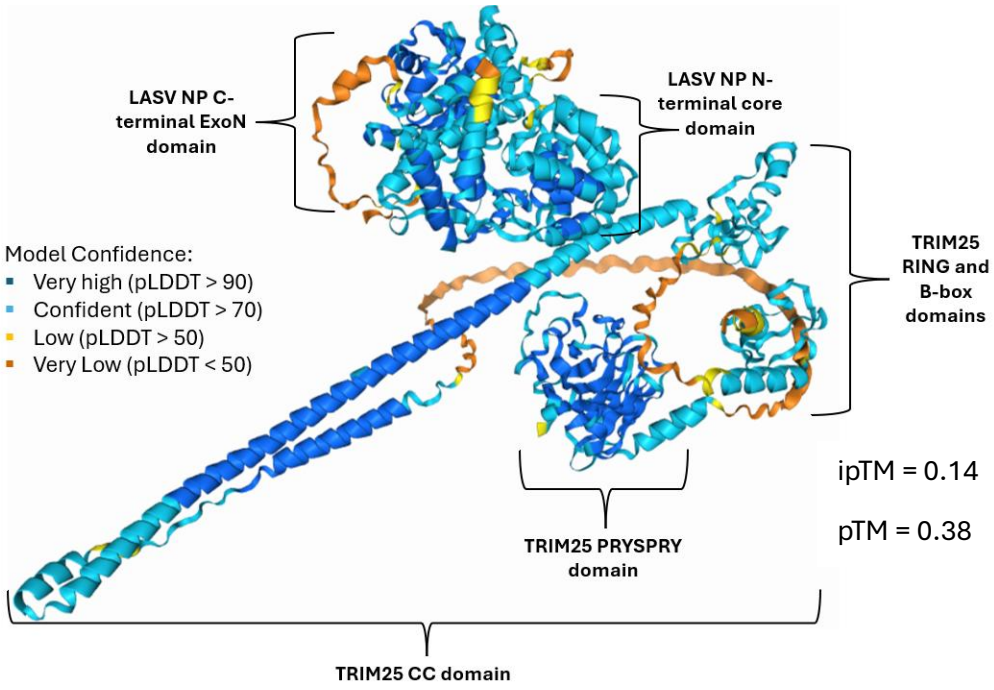


Figure 5.10 Alphafold3 prediction for complex formed by full length wild type LASV NP and full-length wild type TRIM25.

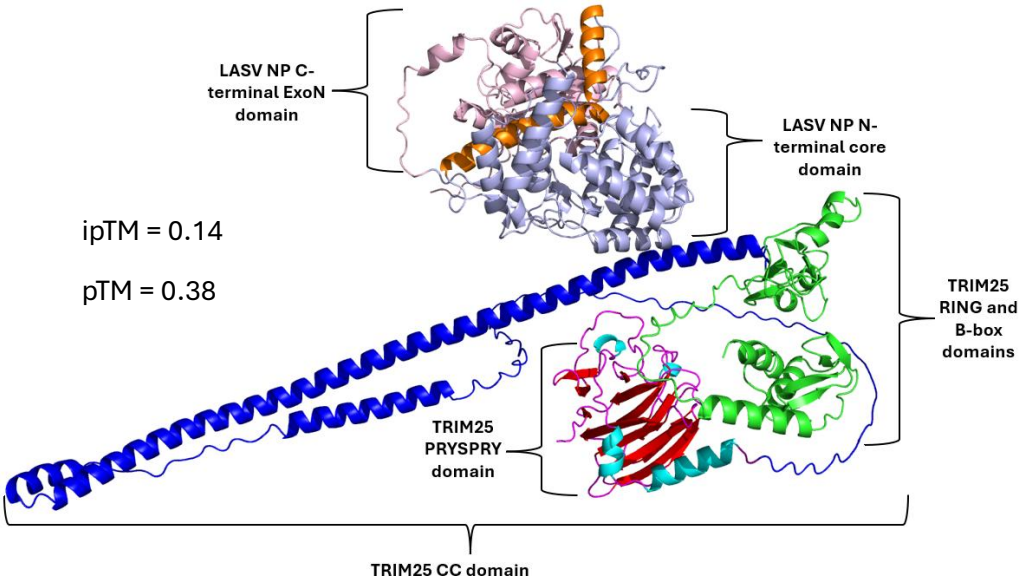


Figure 5.11 Alphafold3 prediction for complex formed by full length wild type LASV NP and full-length wild type TRIM25. The different domains of each protein have been coloured. The C-terminal ExoN domain of LASV NP is in light pink, The N-terminal core domain is in grey with the RNA binding domain in orange. The CC domain of TRIM25 is in dark blue and the RING and B-box domains in green. The PRYSPRY domain is coloured in red (β-sheets), cyan (α-helices), and pink (loops). the predicted template modelling (pTM) score and the interface predicted template modelling (ipTM) score are both derived from a measure called the template modelling (TM) score. This measures the accuracy of the entire structure. A pTM score above 0.5 means the overall predicted fold for the complex might be similar to the true structure. ipTM measures the accuracy of the predicted relative positions of the subunits within the complex. Values higher than 0.8 represent confident high-quality predictions, while values below 0.6 suggest likely a failed prediction. The ipTM score for this prediction = 0.14 and the pTM = 0.38.

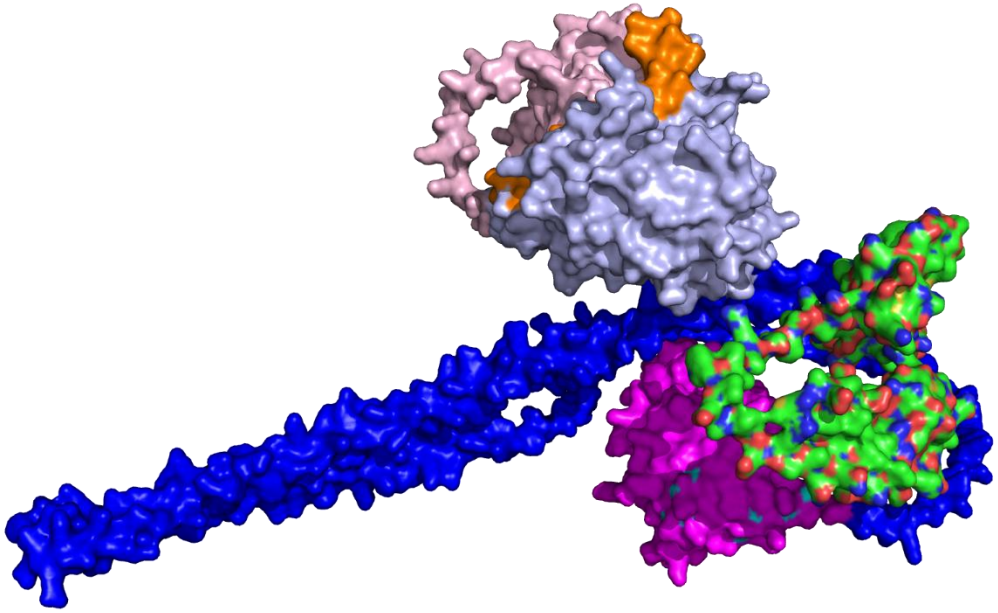


Figure 5.12 Surface structure prediction by PyMOL for complex of full length LASV NP and TRIM25 following the same colouring as in figure 5.12

The interaction in the Alphafold3 predicted pose occurs between LASV NP residues Lys33, Ala37, His40, Gly41, Asp43, Glu72, Ala73, Asp76, Leu77, Glu79, Lys81. And between TRIM25 CC domain Arg203, His204, Leu206, Thr207, Tyr210, Ser211, Ile213, Asn214, Ser217, Arg218. The interaction appears to be occurring between the CC domain of TRIM25 and the α2-, and α4-helix of the LASV NP N-terminal core domain Figure 5.13.

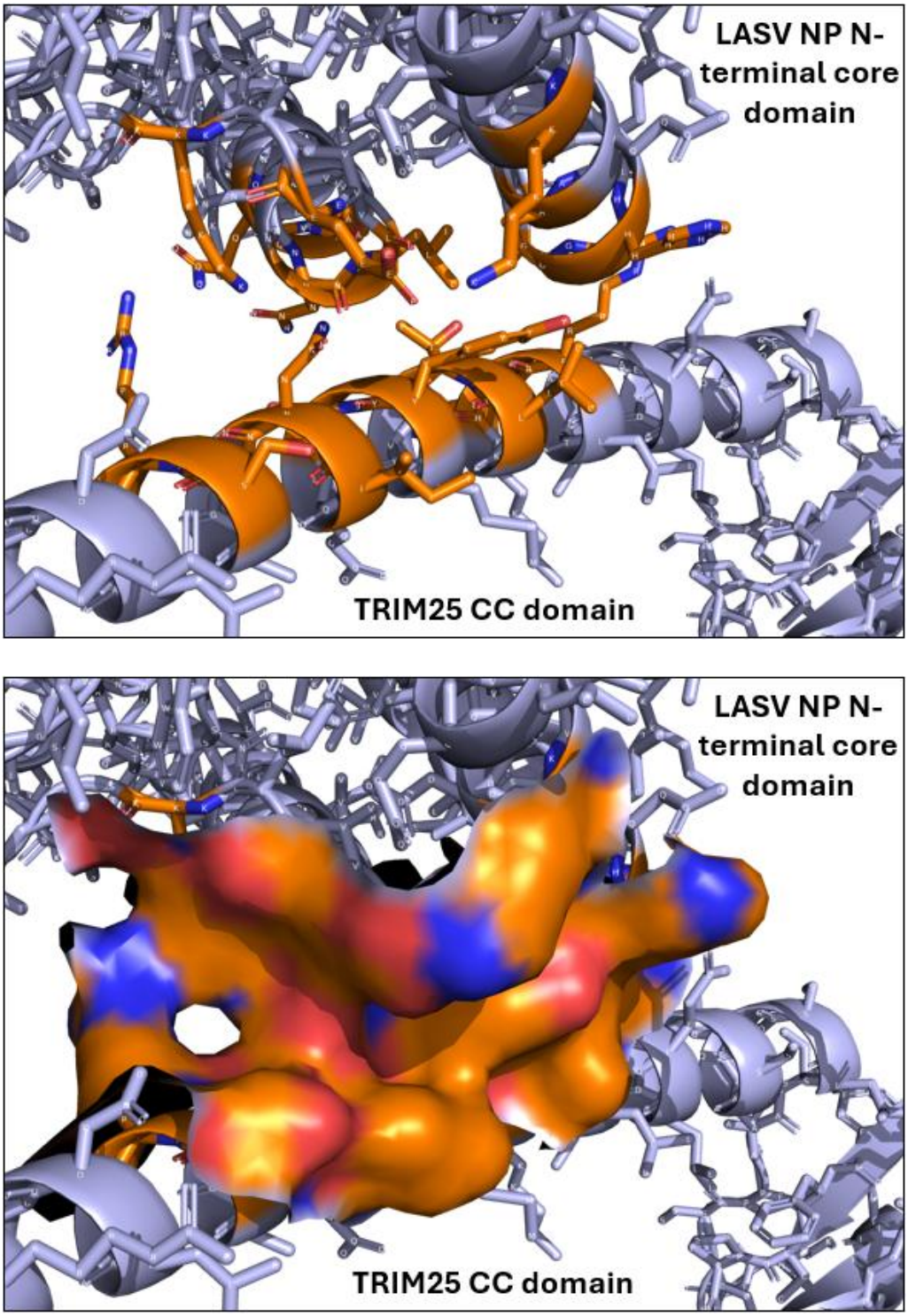


Figure 5.13 Close-up view of the predicted region of interaction between full length LASV NP and full length TRIM25. The amino acid residues involved in the interaction have been coloured in according to element colour and the non-interacting protein for both LASV NP and TRIM25 has been coloured grey. The surface prediction has also been visualised.

5.2.1 Protein Purification and SEC

5.2.1.1 LASV NP

The LASV NP fragment 1-340aa was expressed and purified according to Materials and Methods sections 2.5.1 Protein expressions and Purification and 2.5.2 IMAC, resulting in an SEC profile seen in Figure 5.14 A. A different buffer set was used compared to the LUJV and MOPV NP fragments due to storage issues once purified and the tendency of the LASV NP fragment to precipitate out of solution in the Tris-HCL buffer but not in the sodium phosphate buffer. The molecular weight of the fragment based upon the amino acid sequence is 37.68kDa. SDS-PAGE analysis of the samples across the purification stages, shows a predominant band at the expected size for the LASV NP fragment in the elution lane indicating successful purification by IMAC resin. Figure 5.14 A shows that the SEC chromatogram exhibited multiple peaks suggesting that the N-terminal LASV NP fragment is not monodisperse in solution and tends to form oligomers and/or aggregates. Protein eluted from 8 to 18 mL elution volume (Ve) in SEC 10/300 (Fig. 5.14), with multiple peaks indicating lack of monodispersion and aggregation and/or potential oligomerisation as 8mL corresponds to the void volume (Vd, ~ 8mL). A single peak at ~16mL Ve indicated isolation of LASV NP monomer based on the calibration of SEC 10/300 (16mL corresponds to Ovalbumin, as globular protein of 44kDa, from Cytiva calibration source). Although the SEC profile shows a non-monodisperse profile, samples were pooled in order to achieve the minimum protein concentration required for SEC-SAXS. As the protein sample would undergo another round of SEC during the SEC-SAXS run, the peaks produced will be analysed separately based on oligomeric state. Protein samples from C8 to E5 were pooled, as they eluted around the expected molecular weight for the protein as shown by SDS-PAGE analysis in figure 5.14 C.

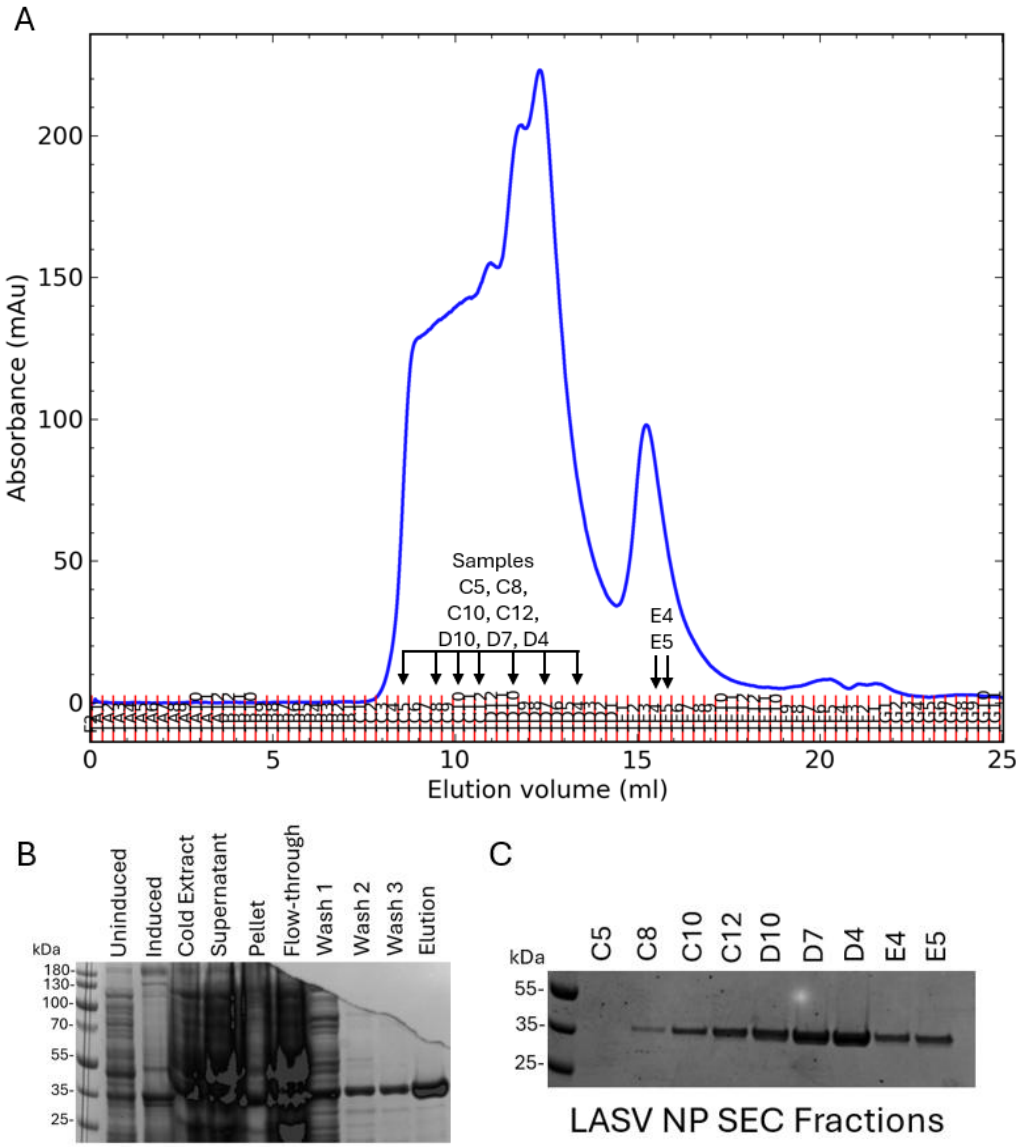


Figure 5.14 LASV NP fragment purification A) SEC graph of UV 280nm from AKTA created using PyCORN-Web. 500µl of purified LASV NP N-terminal fragment 1-340 was run through a Superdex 200 Increase 10/300 GL column (GE Healthcare) using a buffer of 50mM NaH_2PO_4 pH 8, 500mM NaCl B) Coomassie of LASV NP purification stages C) Coomassie of LASV NP fractions after SEC to identify which fractions contain pure LASV NP fragment to use for SEC-SAXS. (n = 1)

5.2.1.2 LUJV NP

The LUJV NP fragment 1-340aa was expressed and purified according to Materials and Methods sections 2.5.1 Protein expressions and Purification and 2.5.2 IMAC resulting in an SEC profile seen in Figure 5.15 A. The predicted molecular weight of this LUJV fragment made up of residues 1-344 based upon the amino acid sequence is 37.5kDa. Figure 5.15 B depicts the Coomassie of stages of the protein purification process and shows that a protein at the predicted molecular weight is well eluted from the resin used in the metal affinity chromatography. Figure 5.15 A show that the SEC chromatogram which exhibits only a single clear peak at Ve 12.5-14mL. This suggests that the N-terminal LUJV NP fragment might stay monomeric. SDS-PAGE analysis shown in Figure 5.15 C shows the Coomassie of the fractions from SEC corresponding to the peaks seen in A. Protein samples from C1 to C6 were pooled, as they eluted around the expected molecular weight for the protein.

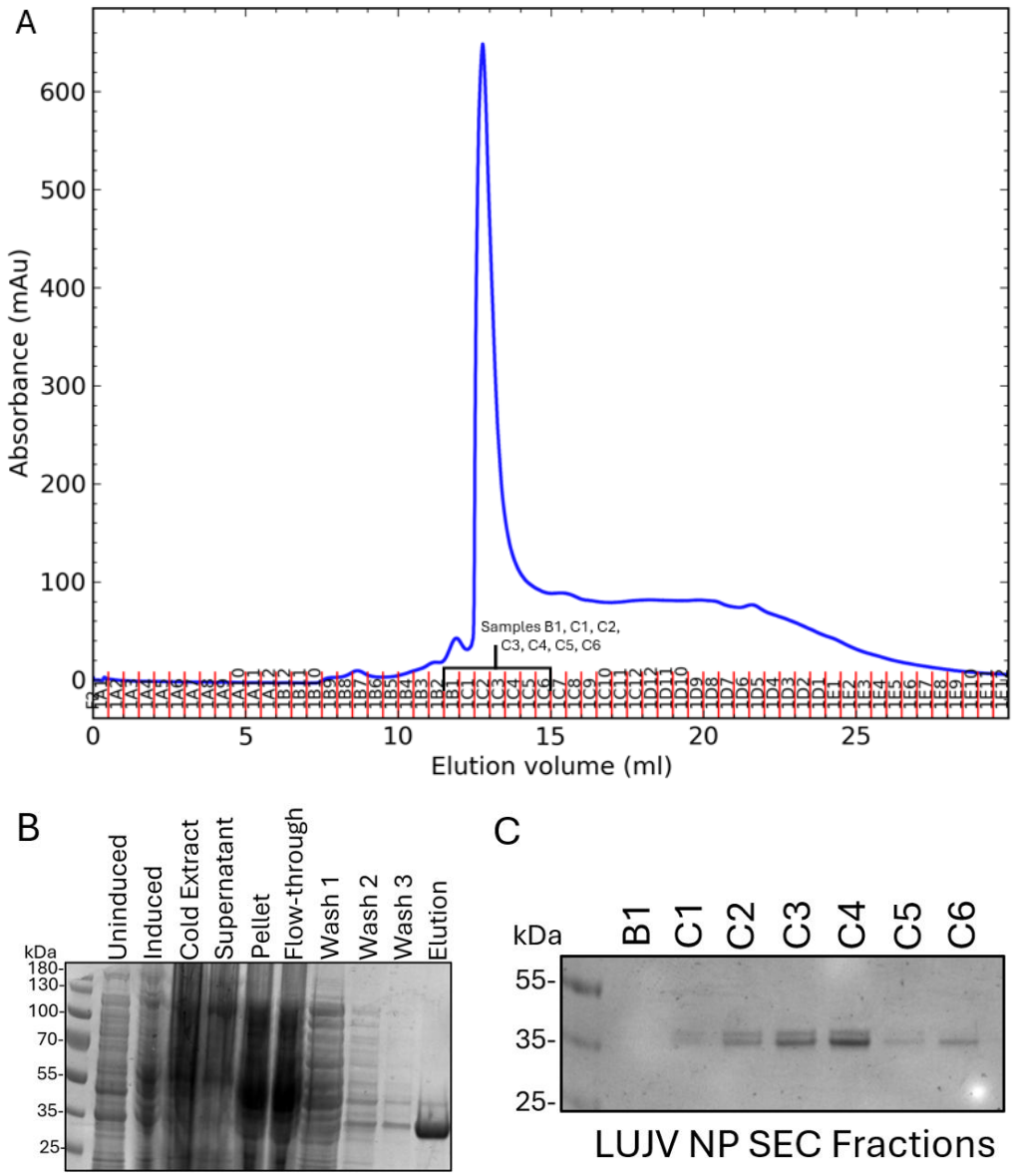


Figure 5.15 LUJV NP fragment purification. SEC graph of UV 280nm from AKTA created using PyCORN-Web. 500µl of purified LUJV NP N-terminal fragment 1-340 was run through a Superdex 200 Increase 10/300 GL column (GE Healthcare) using a buffer of 50mM Tris-HCL pH 7.2, 150mM NaCl B) Coomassie of LUJV NP purification stages C) Coomassie of LUJV NP fractions after SEC to identify which fractions contain pure LUJV NP fragment to use for SEC-SAXS (n = 1).

5.2.1.3 MOPV NP

The MOPV NP fragment 1-340aa was expressed and purified according to Materials and Methods sections 2.5.1 Protein expressions and Purification and 2.5.2 IMAC resulting in an SEC profile seen in Figure 5.16 A. The predicted molecular weight of the fragment based upon the amino acid sequence is 37.95kDa. Figure 5.16 B depicts the Coomassie of stages of the protein purification process and shows that a protein at the predicted molecular weight at around a 35kDa has been eluted from the resin. This Coomassie image also includes the dialysed protein, in which some of the MOPV NP fragment appears to have been lost to the overnight dialysis process. Figure 5.16 A show that the SEC output exhibited multiple peaks suggesting that the N-terminal MOPV NP fragment might be producing oligomers. Four elution peaks can be clearly distinguished at Ve 8-10mL (peak 1), 12 mL (peak 2), 15 mL (peak 3) and a broad 20-28 mL (peak 4). Interestingly, as expected, the Ve for peak 2 is the same as in LASV NP. Although the SEC profile shows a non-monodisperse profile, samples were pooled in order to achieve the minimum protein concentration required for SEC-SAXS. As the protein sample would undergo another round of SEC during the SEC-SAXS run, the peaks produced will be analysed separately based on oligomeric state. Protein samples from B5 to C8 were pooled, as they eluted around the expected molecular weight for the MOPV NP fragment as shown by SDS page analysis in figure 5.16 C.

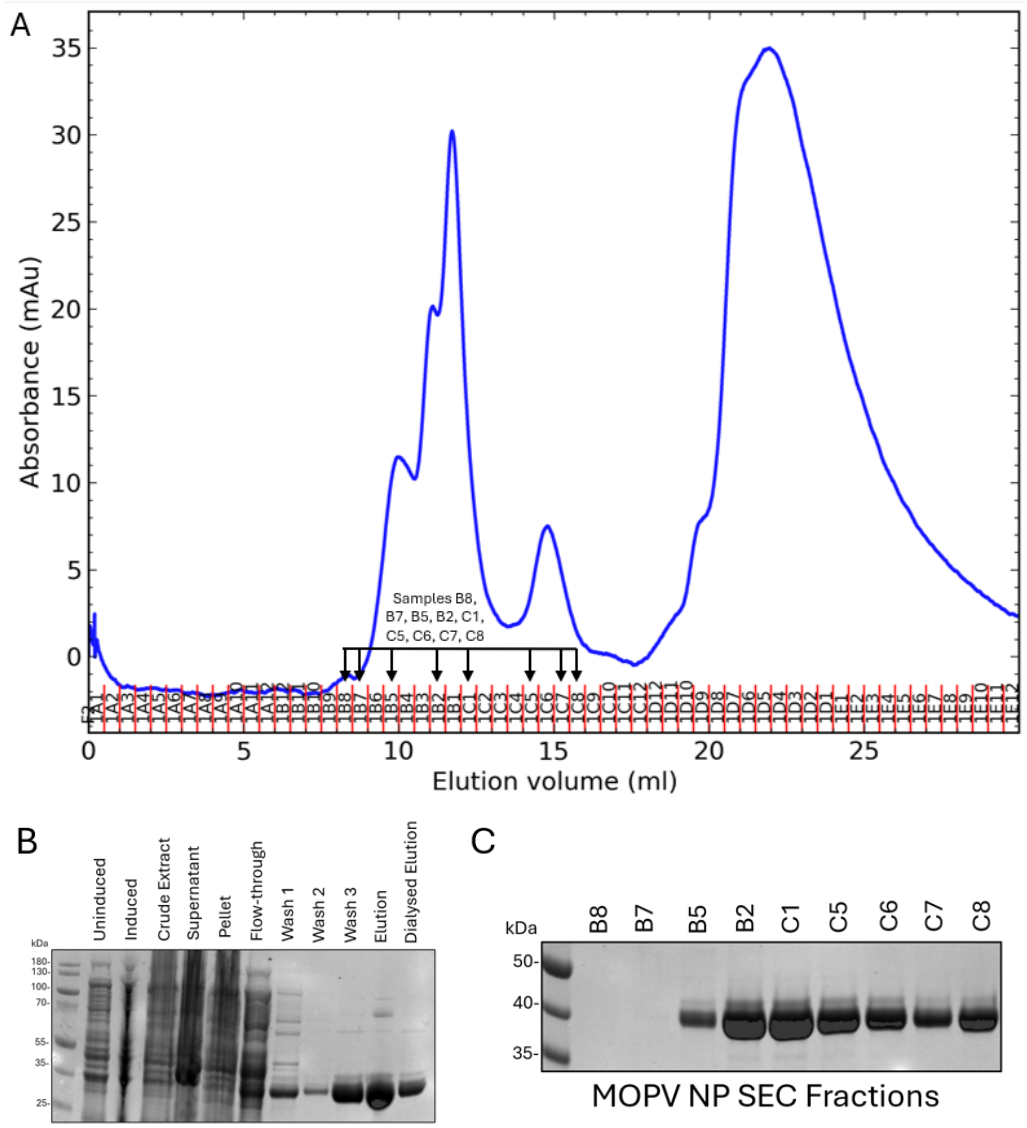


Figure 5.16 MOPV NP fragment purification. SEC graph of UV 280nm from AKTA created using PyCORN-Web. 500µl of purified MOPV NP N-terminal fragment 1-340 was run through a Superdex 200 Increase 10/300 GL column (GE Healthcare) using a buffer of 50mM Tris-HCL pH 7.2, 150mM NaCl B) Coomassie of MOPV NP purification stages C) Coomassie of LASV NP fractions after SEC to identify which fractions contain pure MOPV NP fragment to use for SEC-SAXS (n = 1).

5.2.1.4 TRIM25

The TRIM25 CC-PRYSPRY fragment 189-630aa was expressed and purified according to Materials and Methods sections 2.5.1 Protein expressions and Purification and 2.5.2 IMAC resulting in the SEC profiles seen in Figure 5.17 A due to multiple runs being required. The predicted molecular weight of the fragment based upon the amino acid sequence is 50.22kDa. Figure 5.17 B depicts the SDS-PAGE and Coomassie of stages of the protein purification process and shows that a protein at the predicted molecular weight appears just under the ladder band for 55kDa and has been eluted from the resin used in the metal affinity chromatography. A significant amount of the TRIM25 fragment is present in the wash stages of purification. Figure 5.17 A shows that the SEC output exhibited multiple peaks suggesting that the TRIM25 fragment might be producing oligomers or aggregating, as the column void is ~ 8mL. It has been shown experimentally that the TRIM25 CC-PRYSPRY domain dimerises so an additional peak should be expected, and it is likely the unexpected extra peaks are due to bacterial protein contaminants in the sample. Figure 5.17 C shows the Coomassie of the fractions from SEC corresponding to the peaks seen in A. These fractions have the expected molecular weight we predict for the domain at 50kDa but also contain a high content of impurities as seen by the extra bands present in the elution lane at above 100kDa, at 70kDa, 35kDa, and 25kDa. The fractions with the lowest amount of impurity and high concentration of target protein were taken and collated and concentrated. These fractions were taken from fourth peak in the SEC graph seen in Figure 5.17 A which correlate to protein samples from E1 and E2 which were pooled , as they eluted with a predominant band at around the expected molecular weight for the TRIM25 CC-PRYSPRY fragment as shown by SDS page analysis in figure 5.17 C.

All of the protein samples purified from SEC analysis were flash cooled in liquid nitrogen after pooling and concentrating by centrifugation. The aimed

for concentration for SEC-SAXS was 8mg/ml and a minimum volume of 40µl. Due to the low concentration of some of the samples, the purity stringency applied to samples such as the LASV NP fragment and TRIM25 CC-PRYSPRY fragment was relaxed to ensure an appropriate concentration was achieved. These samples were then transported on dry ice for analysis at the B21 beamline at the Oxford DLS synchrotron (see methods and materials section 2.5.5).

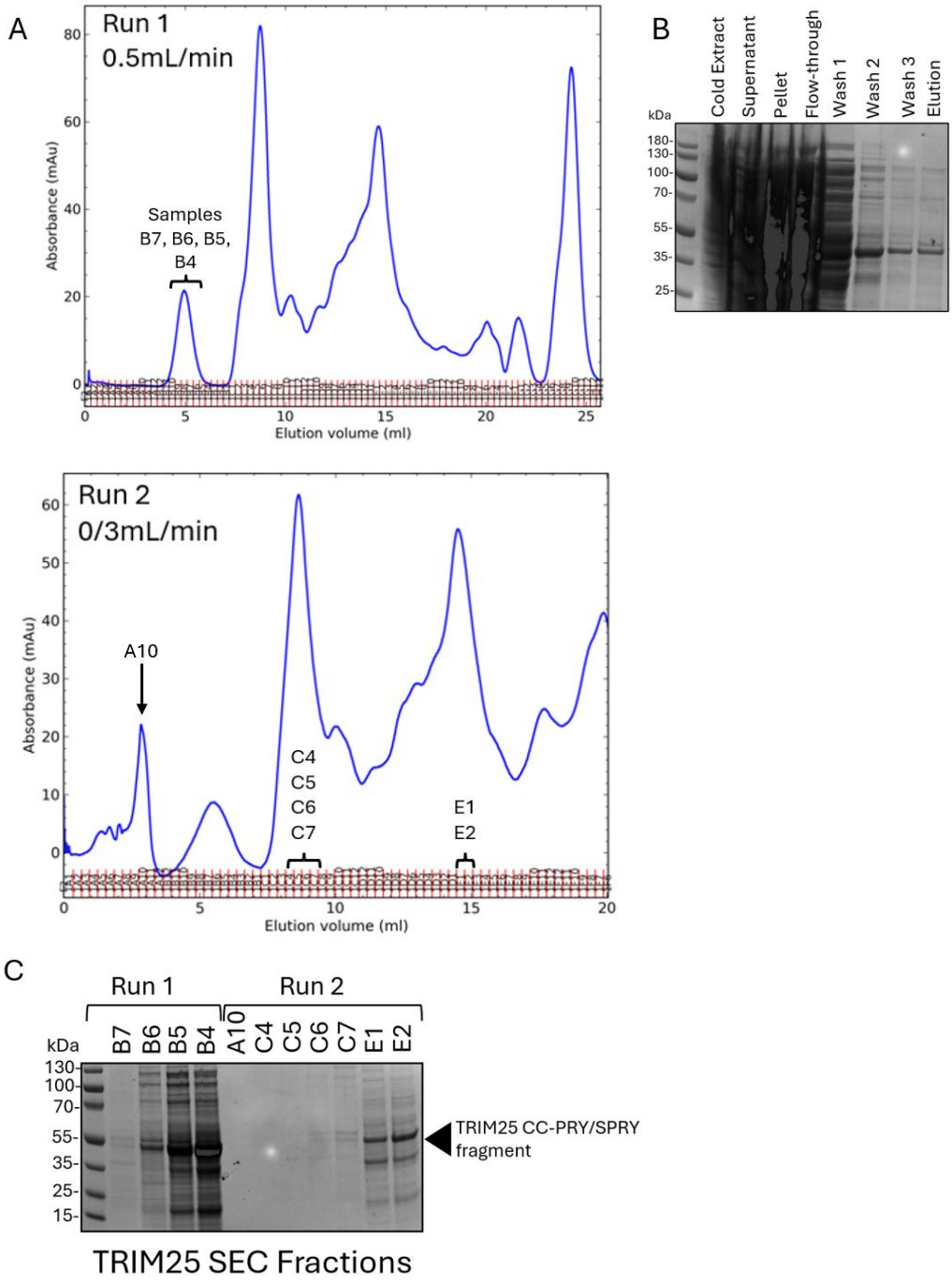


Figure 5.17 TRIM25 CC-PRYSPRY fragment purification. A) SEC graphs of UV 280nm from AKTA created using PyCORN-Web. 500µl of purified TRIM25 CC-PRYSPRY was run on a Superdex 200 Increase 10/300 GL column (GE Healthcare) using a buffer of 20mM HEPES pH 7.5, 100mM NaCl, and 500µM TCEP with 1% glycerol. Due to the high concentration of the sample (8mg/ml) it was diluted to 4mg/ml and run twice. Due to the number of peaks seen in run 1, the flow rate for run 2 was altered to improve resolution B) Coomassie of TRIM25 fragment purification stages C) Coomassie of TRIM25 fragment fractions after SEC to identify which fractions contain pure TRIM25 CC-PRYSPRY fragment to use for SEC-SAXS. Samples from both SEC runs were taken and run on the same gel (n = 1).

5.2.2 SEC-SAXS

5.2.2.1 LASV NP

SCATTER and subsequent ATSAS CHROMIX/PRIMUS assessment of the 600 data frames generated from SEC-SAXS analysis of the LASV NP N-terminal core domain (1-340aa) identifies an envelope for a much larger than expected globular protein. Figure 5.18 A shows the SEC graph of the 600 data points or frames from the sample. The data points highlighted in black represent the buffer and the sample data selection was done by method of fine slicing from sections of the large peak observed. Analysis performed on the left, middle, and right sides of the peak gave results for proteins of significantly different molecular weights; the left side of the peak resulted in a protein with a molecular weight of around 318 kDa in size, whereas that from both the middle and right side of the peak gave a globular protein of about 200 kDa. Although upon further comparisons of the envelopes generated from these peak sections, the smaller protein from the middle and right peaks did not pass the χ^2 or the CorMap statistical tests. Analysis of the proteins forming these sections of the peak actually produced a protein envelope with larger dimensions than expected despite the smaller molecular weight. Thus, analysis on the larger globular protein from the left side of the peak was used for LASV NP which passed all of the statistical testing. Analysis was also performed on the small peak seen from frames 560-600 although this did not yield any useful results and predicted a protein with a molecular weight of 0.35 kDa. From the left side of the large peak, which was selected for analysis, only frames with high similarity were chosen; frames 427-452. This similarity is indicated through SCATTER analysis which matches frames with high similarity. SCATTER performed merge analysis on this frame set using the selected buffer frames as background. The aggregation of samples seen on the far-left side of the graph in Figure 5.18 B is noise generated from damage caused by the X-rays so these samples will need to be trimmed prior to analysis. Further noise

builds in the graph as the graph curve progresses indicating either further X-ray damage to samples or a lack of similarity between sample frames. Figure 5.18 C shows the table for Durbin Watson and Shapiro Wilks tests (Grant et al., 2015, Gräwert and Svergun, 2020) for statistical analysis of the merged buffer and sample frames. Each frame undergoes statistical testing, and these are represented on the graph. The result from the statistical tests is used to inform Figure 5.18 B. The frame points on the graph do not have totally random distribution and does appear to possess a slight trend. This may indicate either some radiation damage to the samples or a lack of similarity. In Figure 5.18 D the I(0) trends towards decreasing as the sample frames measured progress indicating that the sample frames selected here have actually have lower radiation damage and as expected, the I(0) is decreasing as the particle is eluted as the concentration decreases. E shows only marginal displacement between the black and cyan curves indicating good buffer subtraction.

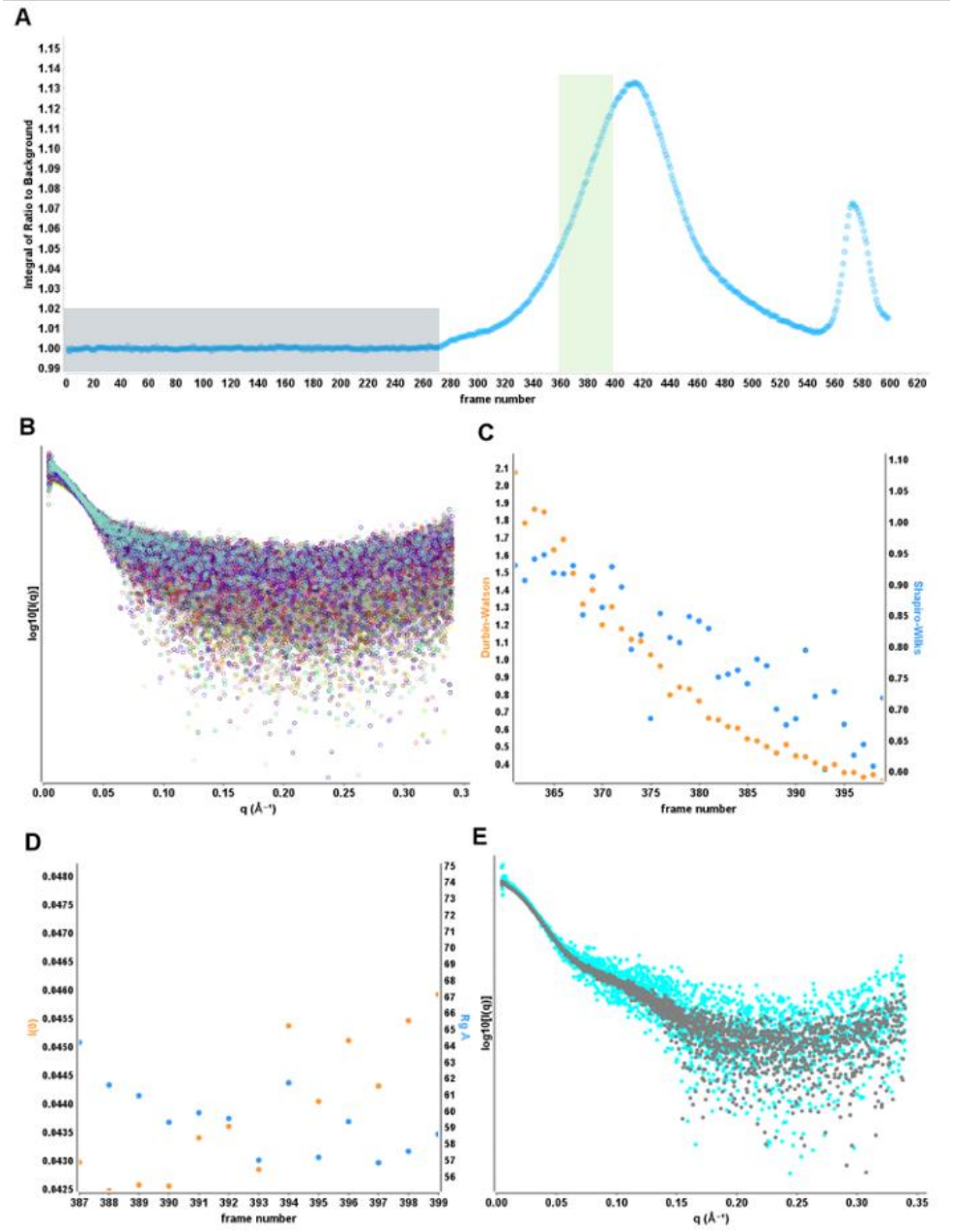


Figure 5.18 LASV NP fragment SCATTER SEC-SAXS analysis A. SEC-SAXS Signal Plot. Each point represents the integrated area of the ratio of the sample SAXS curve to the estimated background. The frames used as the buffer background are gray with the average represented as the gray horizontal line. The frames and peak region selected for subsequent analysis is highlighted in the green box in (A). Elevation of the baseline after peak elution suggests capillary fouling. B. Overlay of SAXS curves of subtracted frames. Each frame is colored based on the following table. C. Durbin-Watson and Shapiro-Wilks tests examining the distribution of the residuals between two frames. In this case, comparisons are made in reference to the first frame. Radiation damage or lack of similarity can be observed as a trend in either statistic across the frame set. Likewise, similarity is demonstrated by a random distribution of the statistics. D. Double Y plot with I (0), orange, and Rg, cyan, estimated from the Guinier region for each subtracted frame. For a single concentration measurement made over several frames, radiation damage will be observed as an increase in I (0) and Rg. For SEC-SAXS, I (0) should change with the concentration of the particle during elution. E. Log 10 intensity plot of subtracted and merged SAXS frames. Black represents averaged buffer frames subtracted from averaged sampled frames. Cyan represents the median of the buffer frames subtracted from the averaged sample frames. Poor buffer subtraction leads to a displacement between the two curves at high-q.

The SAXS intensity vs scattering vector graph seen in Figure 5.19 A shows a relatively sharp drop in I(q) indicating a large structure identified. This agrees with the large D_{max} value observed. Figure 5.19 B shows the Kratky plot which is sensitive to the overall protein morphology and compactness and informs on the final shape of the protein. The shape of this plot indicates that the protein is globular and potentially partially folded. This is confirmed by the intersection with the crosshair marks indicating the Guinier-Kratky point. In Figure 5.19 C, the majority of data points are at positive intensities as indicated by the line above the X-axis. This indicates that there has not been over-subtraction of background. Figure 5.19 D has a shape with multiple peaks and a large D_{max} value which suggests a large, multidomain protein.

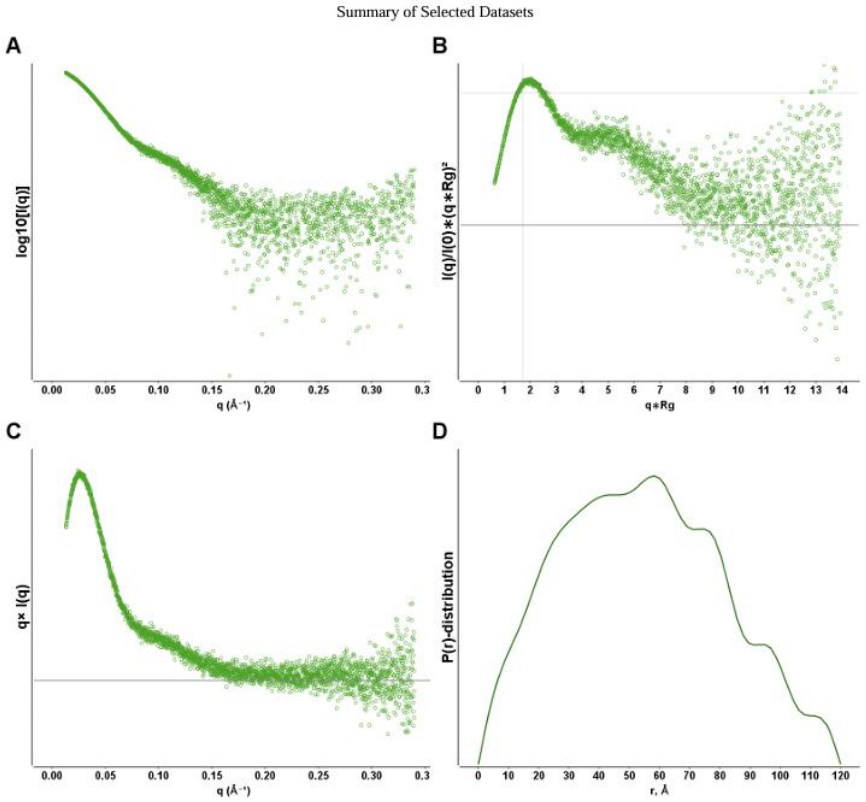


Figure 5.19 A) LASV NP fragment SCATTER SEC-SAXS analysis summary Log10 SAXS intensity versus scattering vector, q. Plotted range represents the positive only data within the specified q-range. B. Dimensionless Kratky plot. Crosshair marks the Guinier-Kratky point (1.732, 1.1), the main peak position for globular particles. C. Total scattered intensity plot. Plot readily demonstrates negative intensities at high-q. Over-subtraction of background leads to significant negative intensities. Likewise, under-subtraction can be observed as an elevated baseline at high-q. Horizontal line is drawn at y=0. D. Pair-distance, P(r), distribution function. Maximum dimension, dmax, is the largest non-negative value that supports a smooth distribution function.

Processed data from SCATTER was used to inform data selection in the ATSAS suite of programs used for modelling of the SAXS data into a final protein envelope (point to materials and methods). This final envelope for the LASV NP fragment is shown in Figure 5.20. Processed data from CHROMIX was analysed in ATSAS PRIMUS which was used for calculation of molecular weight and for dummy atom modelling used to generate the envelopes. Molecular distances were determined using UCSF Chimera 1.18. Interestingly, the molecular weight calculated from the data is 318.45kDa whereas the predicted molecular weight for the LASV NP fragment is 37.68kDa. This suggests that the LASV NP N-terminal fragment is forming oligomers, potentially as large as an octamer. ATSAS PRIMUS statistical analysis of the SEC data used to generate the envelope for the LASV NP N-terminal core domain fragment seen in Figure 5.20 had a reduced χ^2 p value of 0.0855 when modelling as a globular protein indicating that this is the correct shape of the LASV NP structure identified during SEC-SAXS.

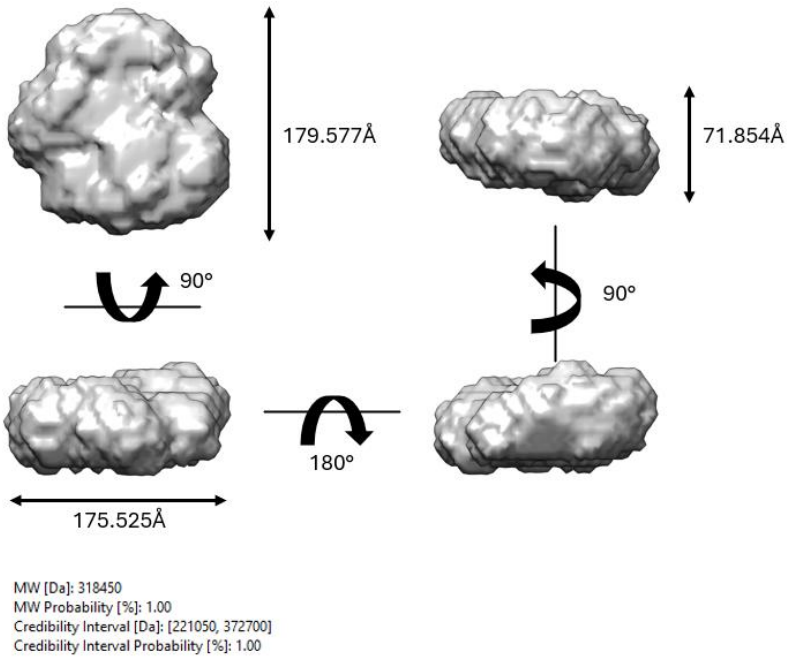
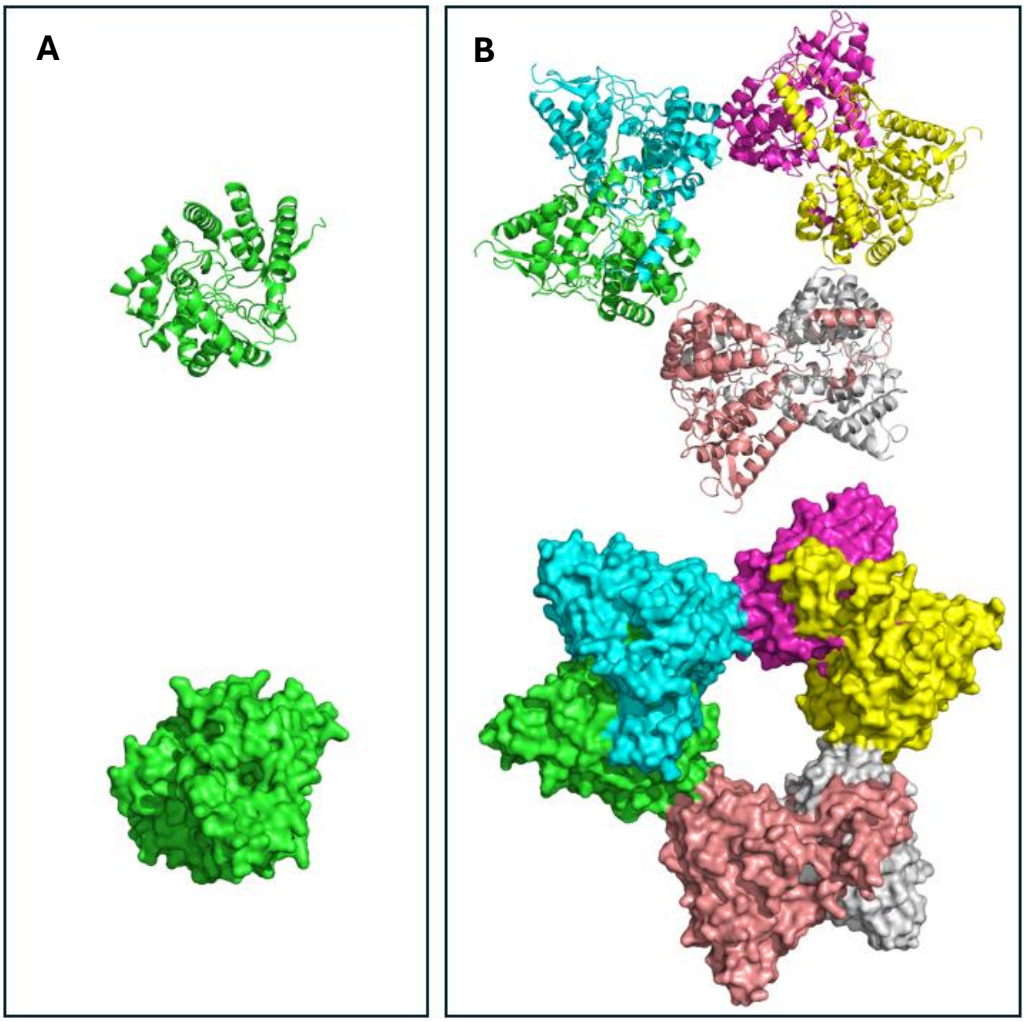


Figure 5.20 LASV NP N-terminal protein molecule envelope as predicted from SAXS data. Data were trimmed and refined in SCATTER prior to processing in ATSAS and PRIMUS from the CHROMIX suite. Dummy atom modelling processing performed to generate 10 fits, which were averaged into a DAMSTART global fit and imported into UCSF Chimera 1.18 to generate a 3 D envelope. n. B. Post processing in ATSAS and PRIMUS interface showing calculated molecular weight of the predicted protein molecule from SAXS data.

Because the calculated molecular weight of the LASV NP envelope is 318.45kDa and the envelope has a much larger than expected size and shape this suggests that LASV NP may be forming large oligomers. To determine whether this is true and if so, which oligomer is being formed, analysis of oligomers has been performed in Figure 5.21. This shows different LASV NP oligomers generated in PyMOL using Alphafold3 predictions for the complexes making up the LASV NP oligomers. The LASV NP N-terminal core domain oligomer comparison has been made between only the larger oligomers that might match the size and shape of the envelope generated from SAXS, although the monomeric protein fragment has been included for reference. Interestingly each of the oligomers appears to take on very different final conformations, each of which would drastically affect the shape of the protein envelope from SAXS data. Whilst

the size of the protein from the SAXS data most closely matches the molecular weight of the octamer, it is possible that due to the star-like shape of the hexamer (Figure 5.21 B) that this might also be a fit. However, model fitting between the envelope seen in Figure 5.20 and the oligomers in Figure 5.21 has been omitted due to the noise in the calculated envelope.



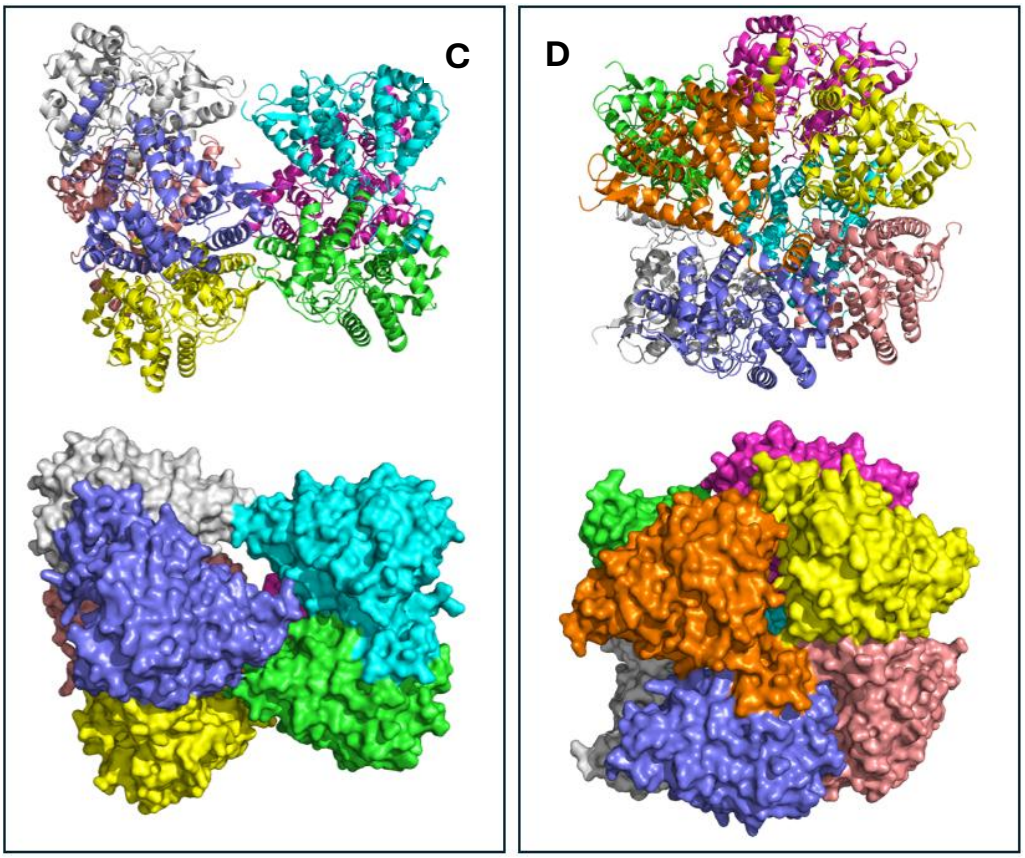


Figure 5.21 Monomer and Oligomers of LASV NP N-terminal protein residues 1-340 generated using an Alphafold3 prediction imported into PyMOL for visualisation and generation of surface structures. Each monomer making up an oligomer has been defined as a different distinct colour. A. Monomeric protein B. Hexameric protein. C. Heptameric protein. D. Octameric protein

5.2.2.2 LUJV NP

SCATTER analysis of the SEC-SAXS data generated by the LUJV NP N-terminal fragment generated the data seen in figure 5.22. The data points highlighted in black in Figure 5.22 A represent the buffer and the sample data was selected from the middle of the first peak, of which only frames with high similarity were chosen (frames 466-494). SCATTER performed merge analysis on this frame set using the selected buffer frames as background, generating Figure 5.22. Figure 5.22 B shows that the data is very noisy meaning that across multiple reads of similar frames there is a lot of difference. Figure 5.22 C shows the table for Durbin Watson and Shapiro Wilks tests for statistical analysis of the merged buffer and sample

frames. The frame points on the graph do not have totally random distribution in either statistical test and both also appear to possess a trend indicating either some radiation damage to the samples or a lack of similarity. Figure 5.22 D shows no data despite numerous attempts at re-analysing the data or even selecting different data sets from the same sample, this either means the calculations to generate the data points on this graph failed or evidence of radiation damage was not detected. Figure 5.22 E shows a reasonably high amount of displacement between the black and cyan curves indicating that buffer subtraction has not been ideal. Figure 5.22A shows that there is a lot of difference between each buffer sample frame and this could be the cause of the disparity seen in E, this could indicate that buffer optimisation is required for this specific sample.

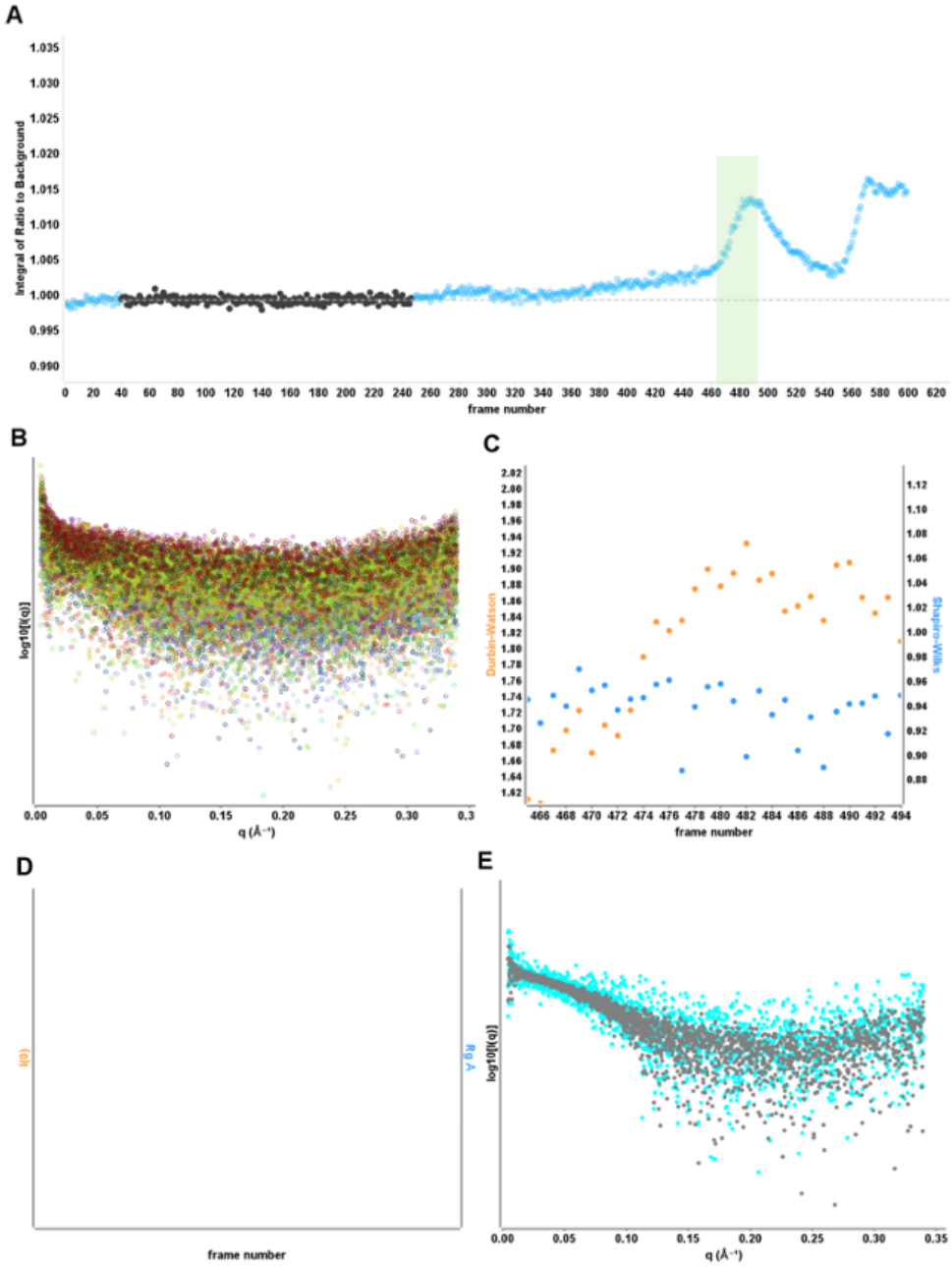


Figure 5.22 LUJV NP fragment SCATTER SEC-SAXS analysis . LUJV NP A) SEC-SAXS Signal Plot. Each point represents the integrated area of the ratio of the sample SAXS curve to the estimated background. The frames used as the buffer background are gray with the average represented as the gray horizontal. The frames and peak region selected for subsequent analysis is highlighted in the green box in (A). Elevation of the baseline after peak elution suggests capillary fouling. B) Overlay of SAXS curves of subtracted frames. Each frame is colored based on the following table. C) Durbin-Watson and Shapiro-Wilks tests examining the distribution of the residuals between two frames. In this case, comparisons are made in reference to the first frame. Radiation damage or lack of similarity can be observed as a trend in either statistic across the frame set. Likewise, similarity is demonstrated by a random distribution of the statistics. D) Double Y plot with I(0), orange, and Rg, cyan, estimated from the Guinier region for each subtracted frame. For a single concentration measurement made over several frames, radiation damage will be observed as an increase in I(0) and Rg. For SEC-SAXS, I(0) should change with the concentration of the particle during elution. E) Log 10 intensity plot of subtracted and merged SAXS frames. Black represents averaged buffer frames subtracted from averaged sampled frames. Cyan represents the median of the buffer frames subtracted from the averaged sample frames. Poor buffer subtraction leads to a displacement between the two curves at high-q.

The SAXS intensity vs scattering vector, also known as the Guinier approximation, seen in Figure 5.23A shows no anomalies or extra peaks in the graph and the I(q) value drops slowly as the q value progresses suggesting a small and monomeric protein. This data is in agreement with the lower D_{max} value observed. Figure 5.23 B shows the Kratky plot which is sensitive to the overall protein morphology and compactness and informs on the final shape of the protein; the shape of this plot indicates that the protein is globular and potentially fully folded as the Y value does go to zero. That the protein is globular is confirmed by the intersection with the crosshair marks indicating the Guinier-Kratky point. In Figure 5.23 C, the overall majority of data points are at positive intensities as indicated by the line above the X-axis, although there are still a significant number of data points at negative intensities. This indicates that whilst there has not been a significant over-subtraction of background, the threshold for this has nearly been reached indicating a potential issue with the buffer. Figure 5.23 D has a shape with multiple peaks although it has a much lower D_{max} value than seen for the LASV NP fragment. Whilst this may suggest a multidomain protein, this needs further investigation albeit it also suggests a globular protein.

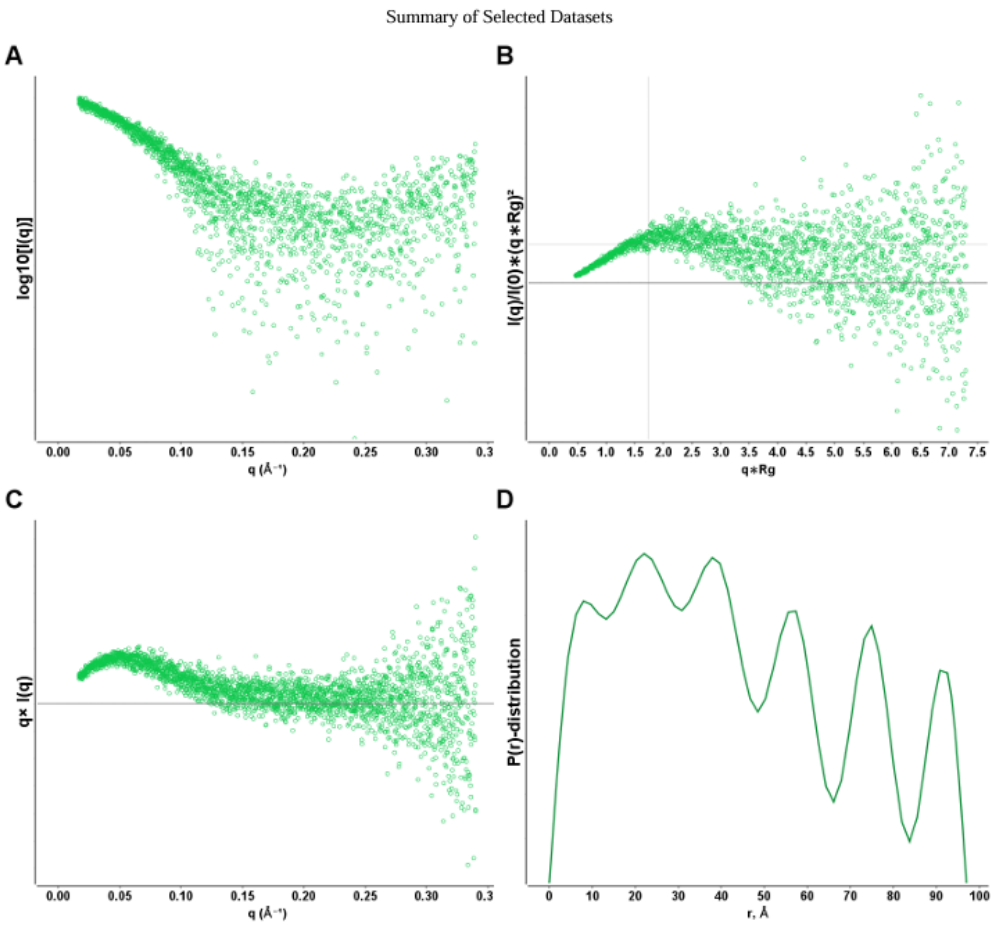


Figure 5.23 LUJV SEC-SAXS data summary plots ; Log10 SAXS intensity versus scattering vector, q. Plotted range represents the positive only data within the specified q-range. B. Dimensionless Kratky plot. Crosshair marks the Guinier-Kratky point (1.732, 1.1), the main peak position for globular particles. C. Total scattered intensity plot. Plot readily demonstrates negative intensities at high-q. Over-subtraction of background leads to significant negative intensities. Likewise, under-subtraction can be observed as an elevated baseline at high-q. Horizontal line is drawn at y=0. D. Pair-distance, P(r), distribution function. Maximum dimension, dmax, is the largest non-negative value that supports a smooth distribution function.

The final protein envelope produced from processing the SAXS data for the LUJV NP fragment suggests that this protein fragment is a monomer in solution. SCATTER analysis resulting in Figure 5.24 A depicts just the envelope surface at different rotations and with the calculated molecular distances. Figure 5.24 B shows the fit of the predicted LUJV NP fragment generated from Alphafold3 into the envelope. The predicted protein shows a near perfect fit to the envelope, and this is further confirmed by Figure 5.24 C. By modelling the Alphafold3 prediction surface and fitting this to the envelope the match becomes even closer and is a remarkable fit indicating

that this is the true structure and envelope for the LUJV NP fragment and that this fragment remains as a monomer.

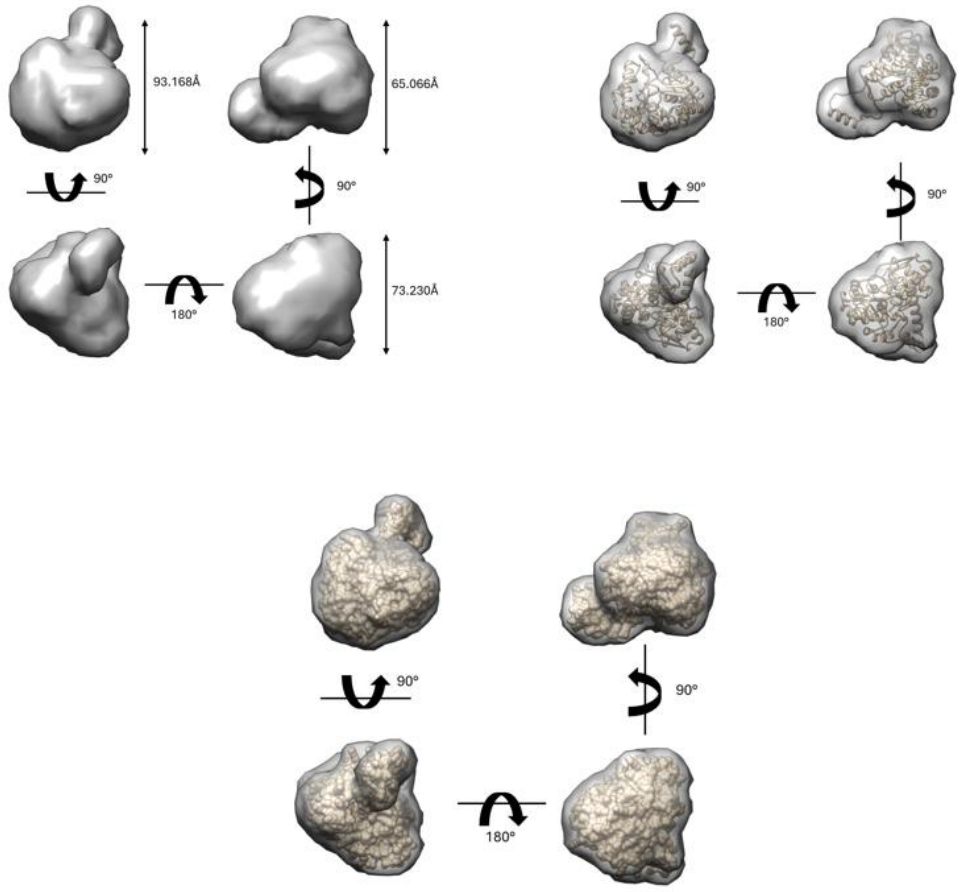


Figure 5.24 LUJV NP N-terminal protein molecule envelope predicted from SAXS data. Data trimmed and refined in SCATTER prior to processing in ATSAS and PRIMUS from the CHROMIX suite. Dummy atom modelling processing performed to generate 10 fits which were averaged into a DAMSTART global fit and imported into UCSF Chimera 1.18 for visualisation. Envelope generated using the command molmap #0 15 and molecule distances calculated between selected atoms using in-built program functions. A. Shows just the surface of the envelope and overall shape with height/width of envelope depicted in angstroms. B. A structure for LUJV NP N-terminal fragment residues 1-344 was created using Alphafold3 predictions and imported into UCSF Chimera 1.18 where a map fit was performed to match the predicted protein shape to the envelope generated from SAXS data. C. This depicts the envelope matched to the surface structure of the Alphafold3 LUJV NP prediction. The surface of the LUJV NP protein has been generated using UCSF Chimera 1.18.

5.2.2.3 MOPV NP

SEC-SAXS data analysis performed using SCATTER in conjunction with ATSAS CHROMIX and PRIMUS reveals a large, potentially multidomain protein envelope for the MOPV NP N-terminal core domain fragment. Figure 5.25 shows graphs generated from analysing the MOPV NP fragment SEC-SAXS data in SCATTER. Figure 5.25 A shows the SEC graph of the 600 data points or frames from the same sample. The data points highlighted in black represent the buffer and the sample data was selected from the large peak, of which only frames with high similarity were chosen which in this case were frames chosen were from 442-452. Other peaks were analysed such as around frame 460 which returned similar protein envelopes as the 442-552 frame peak, but the statistical testing didn't return as good values. SCATTER performed a merged analysis on this frame set after background subtraction. This generated the SCATTER report seen in Figure 5.25. Figure 5.25 B shows the overlay of SAXS curves of selected frames and also informs us about how "noisy" the data set is and how much trimming and data refinement may be required. The aggregation of samples on the far-left side of the graph could be attributed to noise that is most likely due to damage caused by the X-rays, so frames selection needs to be reduced prior to analysis. A lot of noise builds in the graph as we move from X to Y in the curve which indicates either further X-ray damage to samples or a lack of similarity between sample frames which will also need trimming and refinement. Figure 5.25 C shows the table for Durbin Watson and Shapiro Wilks tests for statistical analysis of the merged buffer and sample frames. Each frame undergoes statistical testing, and these are represented on the graph. The result from the statistical tests is used to inform Figure 5.25 B. The frame points on the graph do not have totally random distribution and does appear to possess a slight trend. This also indicates either some radiation damage to the samples or a lack of similarity. In Figure 5.25 D the I(0) trends towards decreasing as the sample frames measured progress

indicating that the sample frames selected here have actually have lower radiation damage and as expected, the I(0) is decreasing as the particle is eluted as the concentration decreases. Figure 5.25 E shows only marginal displacement between the black and cyan curves indicating good buffer subtraction.

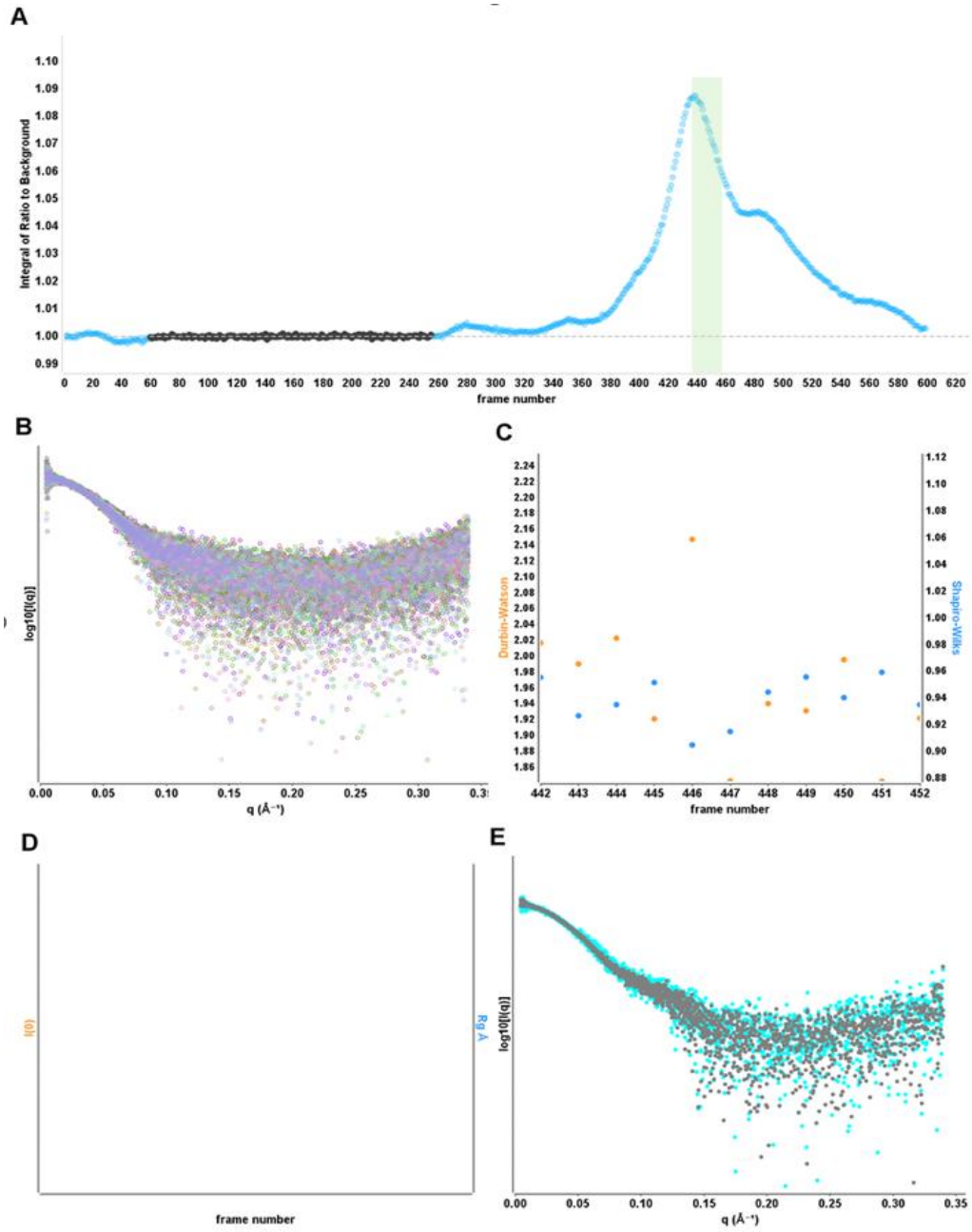


Figure 5.25 MOPV NP N-terminal fragment SEC-SAXS. A. SEC-SAXS Signal Plot. Each point represents the integrated area of the ratio of the sample SAXS curve to the estimated background. The frames used as the buffer background are gray with the average represented as the gray horizontal line. The frames and peak region selected for subsequent analysis is highlighted in the green box in (A). Elevation of the baseline after peak elution suggests capillary fouling. B. Overlay of SAXS curves of subtracted frames. Each frame is colored based on the following table. C. Durbin-Watson and Shapiro-Wilks tests examining the distribution of the residuals between two frames. In this case, comparisons are made in reference to the first frame. Radiation damage or lack of similarity can be observed as a trend in either statistic across the frame set. Likewise, similarity is demonstrated by a random distribution of the statistics. D. Double Y plot with I(0), orange, and Rg, cyan, estimated from the Guinier region for each subtracted frame. For a single concentration measurement made over several frames, radiation damage will be observed as an increase in I(0) and Rg. For SEC-SAXS, I(0) should change with the concentration of the particle during elution. E. Log 10 intensity plot of subtracted and merged SAXS frames. Black represents averaged buffer frames subtracted from averaged sampled frames. Cyan represents the median of the buffer frames subtracted from the averaged sample frames. Poor buffer subtraction leads to a displacement between the two curves at high-q.

The SAXS intensity vs scattering vector for the MOPV NP fragment sample shows a good scattering profile indicating that the sample was pure and that our buffer selection suitable (Figure 5.26 A). Figure 5.26 B shows the Kratky plot which is sensitive to the overall protein morphology and compactness and informs on the final shape of the protein. The shape of this plot indicates that the protein is globular although could be potentially partially folded. The amount of noise as the graph progresses makes it hard to confirm whether the value would have reached zero on the Y axis as the full trend cannot be seen. That the protein is a globular protein is confirmed by the intersection with the crosshair marks indicating the Guinier-Kratky point. In Figure 5.26 C, the majority of data points are at positive intensities as indicated by the line above the X-axis. This indicates that there has not been over-subtraction of background. Figure 5.26 D has a shape with multiple peaks and a large D_{max} value which suggests a multidomain protein or aggregation, likely to hamper the generation of a reliable envelope. Although the D_{max} is lower than that for the LASV NP fragment.

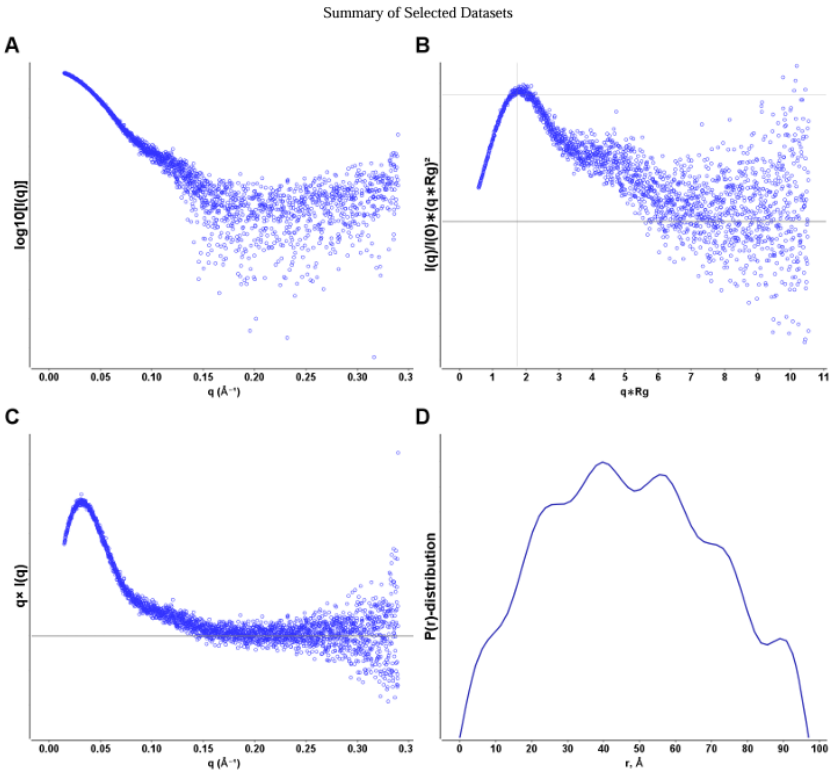


Figure | **A.** Log_{10} SAXS intensity versus scattering vector, q. Plotted range represents the positive only data within the specified q-range. **B.** Dimensionless Kratky plot. Cross-hair marks the Guinier-Kratky point (1.732, 1.1), the main peak position for globular particles. **C.** Total scattered intensity plot. Plot readily demonstrates negative intensities at high-q. Over-subtraction of background leads to significant negative intensities. Likewise, under-subtraction can be observed as an elevated baseline at high-q. Horizontal line is drawn at y=0. **D.** Pair-distance, P(r), distribution function. Maximum dimension, d_{max}, is the largest non-negative value that supports a smooth distribution function.

Figure 5.26 MOPV NP N-terminal fragment SEC-SAXS summary plots | A. Log10 SAXS intensity versus scattering vector, q. Plotted range represents the positive only data within the specified q-range. B. Dimensionless Kratky plot. Crosshair marks the Guinier-Kratky point (1.732, 1.1), the main peak position for globular particles. C. Total scattered intensity plot. Plot readily demonstrates negative intensities at high-q. Over-subtraction of background leads to significant negative intensities. Likewise, under-subtraction can be observed as an elevated baseline at high-q. Horizontal line is drawn at y=0. D. Pair-distance, P(r), distribution function. Maximum dimension, dmax, is the largest non-negative value that supports a smooth distribution function.

Processed data from SCATTER was used to inform data selection in the ATSAS suite of programs used for modelling of the MOPV NP fragment SAXS data into a final protein envelope which is shown in Figure 5.27. Processed data from CHROMIX was analysed in ATSAS PRIMUS which was used for calculation of molecular weight and for dummy atom modelling used to generate the envelopes. Molecular distances were determined using UCSF Chimera 1.18. The reduced χ^2 p value of the data processed in ATSAS PRIMUS was 0.104 indicating that this shape given for the MOPV NP N-terminal core domain is correct and suggesting that the fragment oligomerises.

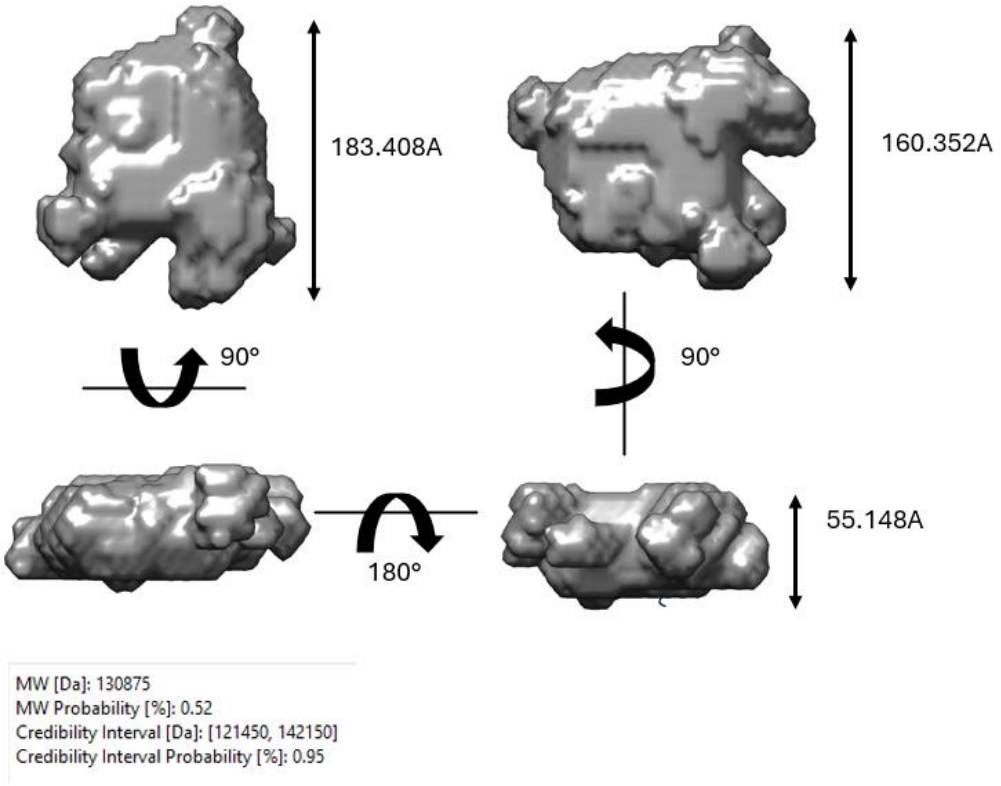


Figure 5.27 MOPV NP N-terminal protein molecule envelope predicted from SAXS data. Data trimmed and refined in SCATTER prior to processing in ATSAS and PRIMUS from the CHROMIX suite. Dummy atom modelling processing performed to generate 10 fits which were averaged into a DAMSTART global fit and imported into UCSF Chimera 1.18 for visualisation. Envelope generated using the command molmap #0 15 and molecule distances calculated between selected atoms using in-built program functions. Post processing in ATSAS and PRIMUS interface showing calculated molecular weight of the predicted protein molecule from SAXS data.

Because the calculated molecular weight of the MOPV NP envelope is 130.875kDand the envelope has a much larger than expected size and shape this suggests that MOPV NP may be forming an oligomer. To determine whether this is true and if so, which oligomer is being formed, analysis of oligomers is necessary. Figure 5.28 shows different MOPV NP oligomers generated in PyMOL using Alphafold3 predictions for the complexes making up the MOPV NP oligomers. The calculated molecular weight from the SAXS data divided by the predicted molecular weight of the MOPV NP fragment (37.95kDa) gives a value of 3.45 which suggests that MOPV may be forming either a trimer or a tetramer. However, due to the noise in the calculated envelope, model fitting has been omitted.

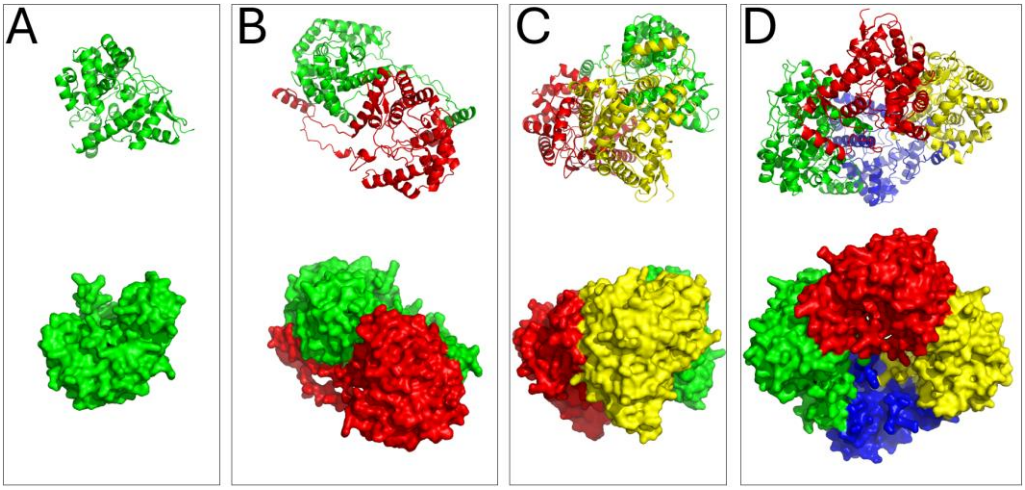


Figure 5.28 Monomer and Oligomers of MOPV NP N-terminal protein residues 1-340 generated using an Alphafold3 prediction imported into PyMOL for visualisation and generation of surface structures. Each monomer making up an oligomer has been defined as a different distinct colour. A. Monomeric protein B. Dimeric protein. C. Trimeric protein. D. Tetrameric protein

5.2.2.4 TRIM25

Due to the lack of sample purity for the TRIM25 CC-PRYSPRY fragment, this sample would not normally be suitable for analysis by SEC-SAXS. However, as this protein was critical for research of potential complexes, it was decided to attempt SEC-SAXS analysis anyway. It was hoped that further runs of SEC would allow for what TRIM25 CC-PRYSPRY was in the sample to be isolated enough for subsequent SAXS data to yield goo results. From the analysis, none of the peaks, or slices of individual peaks seen in Figure 5.29 A yielded any sensible results. The extra peaks that were not selected for analysis in the SEC graph are most likely data from the impurities in the sample seen from Figure 5.17. From the selected peak (highlighted in green in Figure 5.29 A), only frames with high similarity were chosen which in this case included only two frames (433 and 434). That only a small number of frames displayed high enough similarity to warrant choosing coupled with the multiple peaks observed in Figure 5.29 A correlates with the impurity of the sample seen in Figure 5.17. SCATTER performed merge analysis on this frame set using the selected buffer frames as background. This generated the SCATTER report seen in Figure 5.29 and despite only working from a

small set of two frames, the data appears to be good. Figure 5.29 B shows the overlay of SAXS curves of selected frames and tells us that there is a moderate amount of noise to the data and that some degradation may have occurred due to damage from the X-rays. Figure 5.29 C shows the table for Durbin Watson and Shapiro Wilks tests for statistical analysis of the merged buffer and sample frames. Due to the low number of frames this graph does not tell us too much as there are not enough data points for any trend to form or for distribution to be judged as random or not. Therefore, the amount of radiation damage or similarity between sample frames cannot be determined. The same can be said for the data shown in Figure 5.29 D where the lack of data makes inferring from the graph somewhat dubious. Although, the I(0) does trend towards significantly decreasing across the two sample frames measured indicating that the frames selected here have lower radiation damage and that the particle is eluted as the concentration decreases. Figure 5.29 E shows almost total fit with no displacement early on in the curve and as the graph curve progresses, only then do we start to see displacement between the black and cyan curves. This indicates very good buffer subtraction.

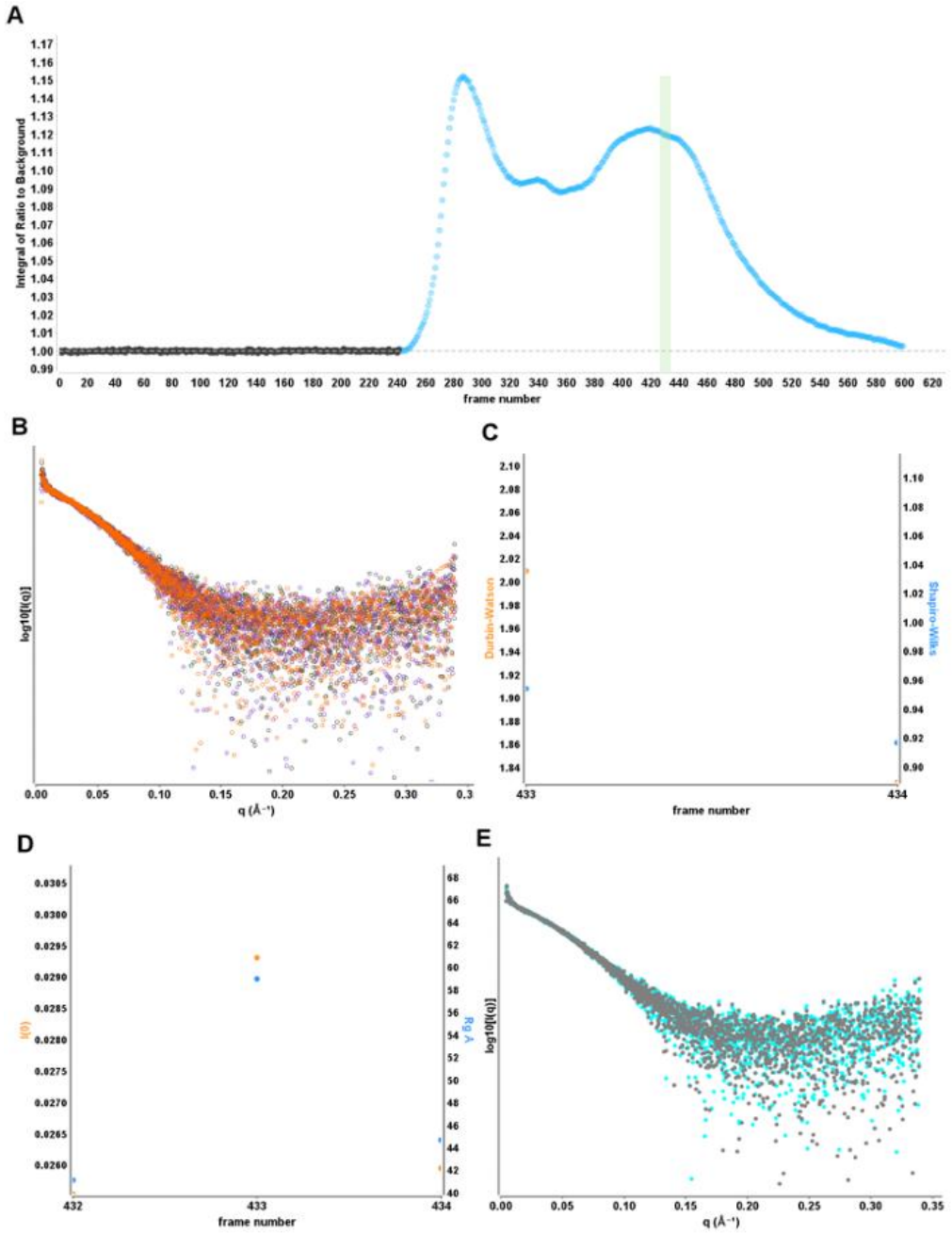


Figure 5.29| TRIM25 CC-PRYSPRY fragment SCATTER SEC-SAXS analysis. A. SEC-SAXS Signal Plot. Each point represents the integrated area of the ratio of the sample SAXS curve to the estimated background. The frames used as the buffer background are gray with the average represented as the gray horizontal line. The frames and peak region selected for subsequent analysis is highlighted in the green box in (A). Elevation of the baseline after peak elution suggests capillary fouling. B. Overlay of SAXS curves of subtracted frames. Each frame is colored based on the following table. C. Durbin-Watson and Shapiro-Wilks tests examining the distribution of the residuals between two frames. In this case, comparisons are made in reference to the first frame. Radiation damage or lack of similarity can be observed as a trend in either statistic across the frame set. Likewise, similarity is demonstrated by a random distribution of the statistics. D. Double Y plot with I(0), orange, and Rg, cyan, estimated from the Guinier region for each subtracted frame. For a single concentration measurement made over several frames, radiation damage will be observed as an increase in I(0) and Rg. For SEC-SAXS, I(0) should change with the concentration of the particle during elution. E. Log 10 intensity plot of subtracted and merged SAXS frames. Black represents averaged buffer frames subtracted from averaged sampled frames. Cyan represents the median of the buffer frames subtracted from the averaged sample frames. Poor buffer subtraction leads to a displacement between the two curves at high-q.

The SCATTER analysis resulting in the SAXS intensity vs scattering vector seen in Figure 5.30 for the TRIM25 CC-PRYSPRY fragment sample indicates a larger than monomeric and rod-like protein envelope. This graph shows a good scattering profile despite the impurity of the initial sample indicating that the sample still contained TRIM25 CC-PRYSPRY at high enough concentration and that a suitable buffer was selected. Figure 5.30 B shows the Kratky plot which is sensitive to the overall protein morphology and compactness and informs on the final shape of the protein. The shape of this plot indicates that the protein is potentially globular although the peak does not match the crosshair marks indicating the Guinier-Kratky point so could actually possess a different shape. In Figure 5.30 C, the majority of data points are at positive intensities as indicated by the line above the X-axis.as expected for adequate background subtraction. Figure 5.30 D has a shape with multiple peaks and a large D_{max} value which suggests a multidomain protein.

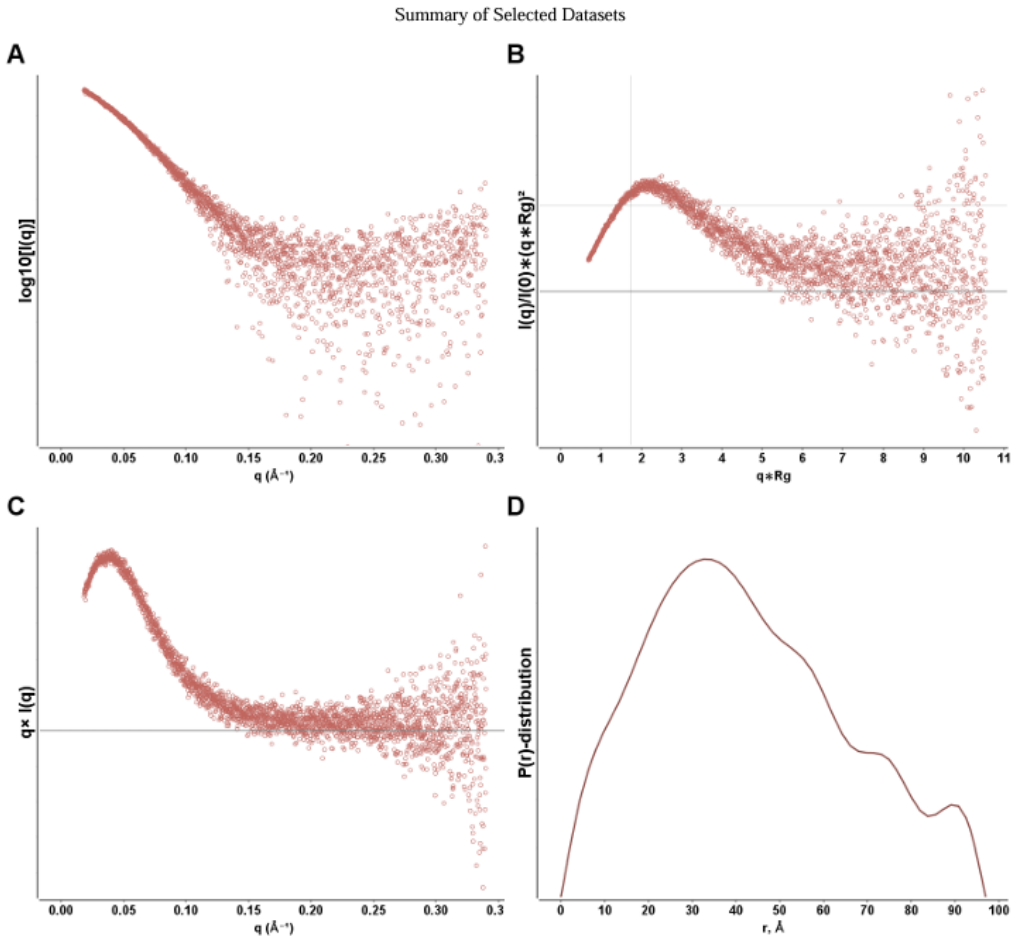


Figure 5.30 |TRIM25 CC-PRYSPRY fragment SEC-SAXS summary plots. A. Log10 SAXS intensity versus scattering vector, q. Plotted range represents the positive only data within the specified q-range. B. Dimensionless Kratky plot. Crosshair marks the Guinier-Kratky point (1.732, 1.1), the main peak position for globular particles. C. Total scattered intensity plot. Plot readily demonstrates negative intensities at high-q. Over-subtraction of background leads to significant negative intensities. Likewise, under-subtraction can be observed as an elevated baseline at high-q. Horizontal line is drawn at y=0. D. Pair-distance, P(r), distribution function. Maximum dimension, dmax, is the largest non-negative value that supports a smooth distribution function.

The final protein envelope for the TRIM25 CC-PRYSPRY fragment generated from the SAXS data is shown in Figure 5.31 A. This also displays the height and width of the envelope, and Figure 5.31 also shows the calculated molecular weight as 88.325kDa. The generated envelope is a rod or ellipsoid shape. Although the ATSAS PRIMUS analysis passed on the Anderson-Darling statistical test (indicating that the sample of data is drawn from a population of data with a normal distribution) with a p-value of 0.0517, it failed the gold standard of the reduced χ^2 test and the Cormap

test to test the randomness of the residuals and indicate location of non-fit for the envelope.

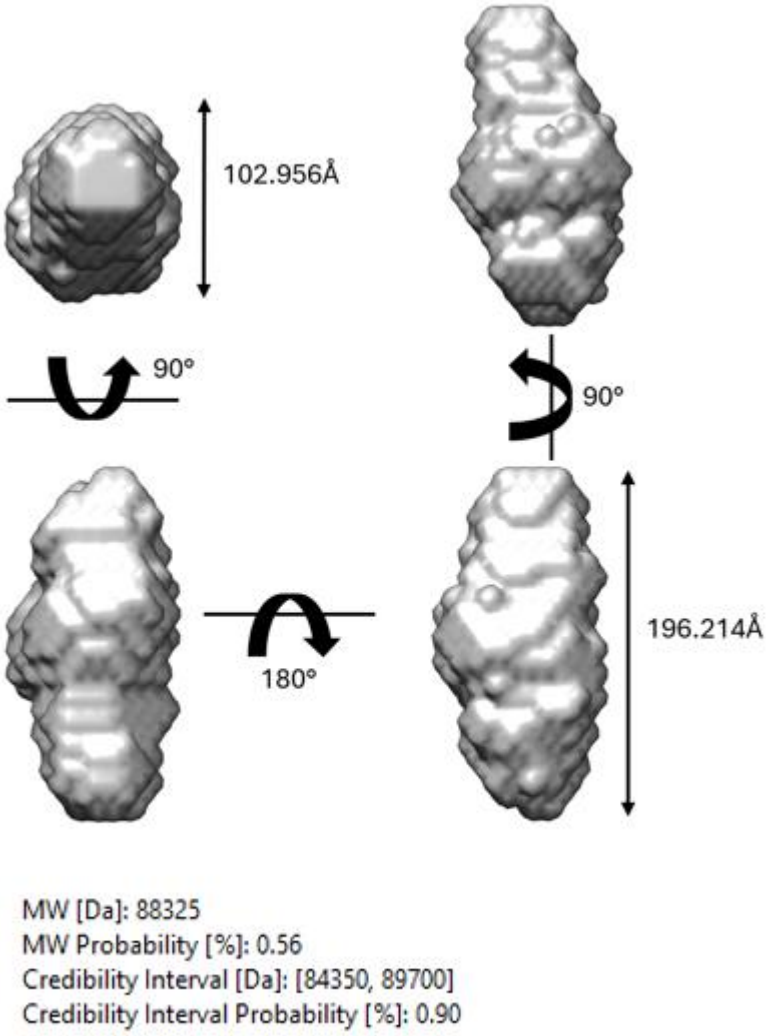


Figure 5.31 TRIM25 CC-PRYSPRY protein molecule envelope predicted from SAXS data. Data trimmed and refined in SCATTER prior to processing in ATSAS and PRIMUS from the CHROMIX suite. Dummy atom modelling processing performed to generate 10 fits which were averaged into a DAMSTART global fit and imported into UCSF Chimera 1.18 for visualisation. Envelope generated using the command molmap #0 15 and molecule distances calculated between selected atoms using in-built program functions. A. Shows just the surface of the envelope and overall shape with height/width of envelope depicted in angstroms. B. A structure for TRIM25 CC-PRYSPRY has been solved by X-ray crystallography (PDB 6FLN) (Koliopoulos et al., 2018) and this was imported into UCSF Chimera 1.18 where a map fit was performed to match the predicted protein shape to the envelope generated from SAXS data. C. This depicts the envelope matched to the surface structure of the solved TRIM25 CC-PRYSPRY domain. The surface of the protein has been generated using UCSF Chimera 1.18. D. Molecular weight predictions from CHROMIX PRIMUS in Daltons.

The molecular weight calculated is higher than that of the TRIM25 CC-PRYSPRY monomeric form (50.22kDa), the value is almost double suggesting potential dimerisation of the TRIM25 CC-PRYSPRY fragment. Due to this the solved crystal structure for the TRIM25 CC-PRYSPRY dimer was chosen for the protein to envelope map comparison (PDB: 6fln) (Koliopoulos et al., 2018). This comparison of the envelope from SEC-SAXS and the solved TRIM25 CC-PRYSPRY dimer is shown in Figure 5.32.

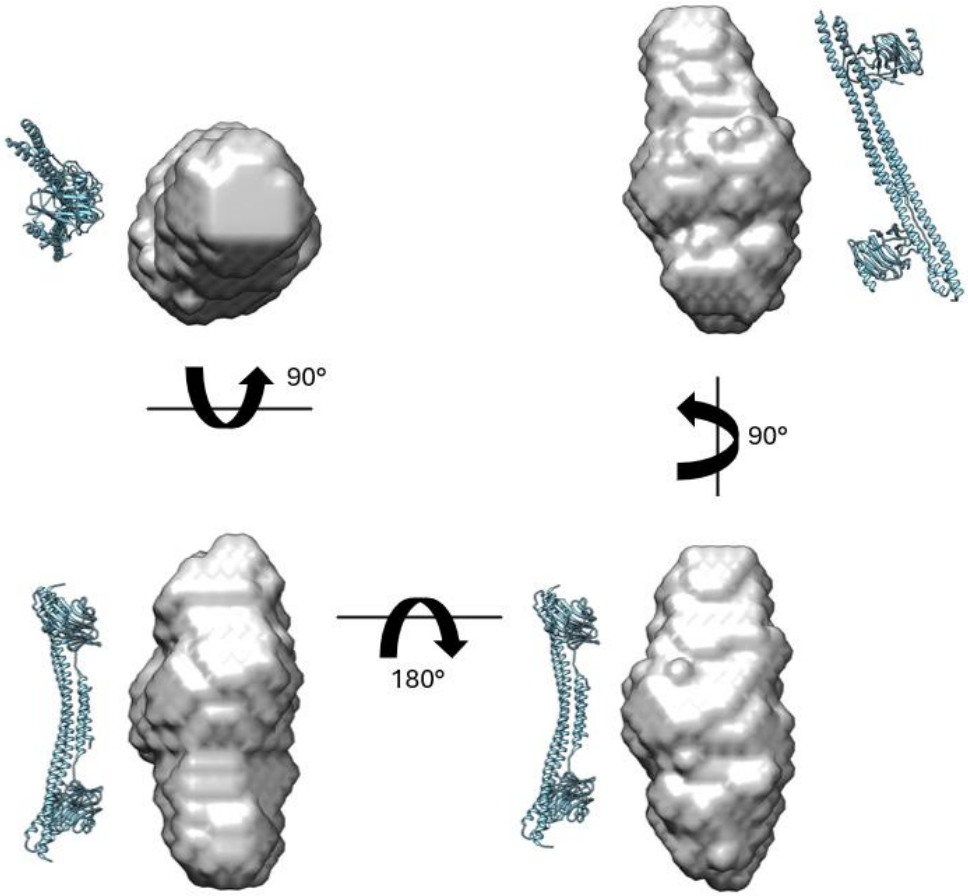


Figure 5.32 Side-by-side view of the solved crystal structure for the TRIM25 CC-PRYSPRY dimer (PDB: 6fln) (Koliopoulos et al., 2018) next to the SEC-SAXS envelope.

5.2.3 Complex formation

Analysis of complex formation between the TRIM25 CC-PRYSPRY domain and the arenavirus NP N-terminal core domain was not possible from the SEC-SAXS data. The SEC traces displayed multiple peaks, none of which appeared to correspond the expected elution volume for the complex. It was possible to extract data on individual proteins such as the LASV, LUJV, and MOPV NP N-terminal core domains from the complex data but none of the analysis passed any statistical test requirements nor produce any shape that were in agreement with the Alphafold3 prediction, which in itself can be a limiting factor. Alphafold3 works by using machine learning to predict a protein's 3D structure based upon previously identified structures; if there are structural aspects to the TRIM25-NP complex which are novel, Alphafold3 is unlikely to be able to predict this as it would not fit previous models it has been trained on. Therefore, the predicted structure from Alphafold3 may itself be inaccurate.

5.3 Discussion

This chapter has described the purification strategy of the LASV, LUV, and MOPV N-terminal core domains and subsequent analysis by SEC-SAXS to identify the protein envelopes formed by these proteins in solution. The nature of how these protein fragments behave in isolation has been further characterised; the SEC-SAXS data shows that the LASV NP fragment forms a large oligomer which supports published data that the full length wild type LASV NP can form oligomers as large as hexamers and can also elute from SEC columns as monomer and trimer (Hastie et al., 2011c, Brunotte et al., 2011a). It is interesting that the NP core domain appears to form even larger oligomers based on our SAXS data. For the LASV NP fragment, further SAXS analysis is required to validate these findings. The LUJV and MOPV NP fragments underwent two separate SEC-SAXS runs, which both led to two similar results in term of structure envelopes. This validation indicates that

the data for the LUJV NP fragment suggesting it stays as a single monomer is correct. It is curious whether this difference in behaviour to the LASV NP fragment might have implications in differences in pathogenicity between the viruses. Information on how the full-length wild type LUJV NP including the ExoN domain behaves is required for further analysis and would be an aim of any future structural research into arenavirus NP, as would investigation into the full length MOPV NP . Whether these observed differences in NP domain behaviour affect interactions such as RNP formation would be an interesting avenue for future research. The MOPV NP N-terminal core domain fragment, similar to LASV NP, formed a larger structure envelope than expected for a single monomer. Both the SCATTER and ATSAS analysis of the SEC-SAXS data indicate that this MOPV NP fragment forms either a trimer or a tetramer.

This chapter also described the purification of the TRIM25 CC-PRYSPRY domain which proved difficult despite numerous attempts at optimising factors such as the vector the fragment was in; both pET15b and pET28b were tested with pET15b both transforming and expressing better in culture. Different buffer compositions for lysis, wash, elution, dialysis, and AKTA buffers were tested. This protein was recalcitrant to several attempts to purify it from more than 2.5g of pelleted bacterial culture at a time meaning that each purification yielded low concentrations of the construct. As SEC-SAXS performs better with higher protein concentrations this was not ideal for further analysis. Much of the protein was often present in wash stages during purification and what was successfully eluted contained contaminants. As this was the pivotal protein for all of the complexes this meant that successful analysis of any TRIM25-NP complex was unable to be achieved. Despite this, SEC-SAXS analysis did yield a shape which roughly fitted the published structure for a TRIM25 CC-PRYSPRY dimer, although statistical analysis suggests this may be an improper fit likely due to the sample impurities preventing robust analysis. It is important to note that any data input into the SEC-SAXS analysis, regardless of quality, will

generate an envelope. Future research would focus on further improving and optimising purification of this TRIM25 CC-PRYSPRY domain, including exclusion at the first SEC of any oligomer. Changes and optimisations were made to the bacterial expression protocols, in which both small scale cultures and larger scales fermentation were tested. Different buffer sets were also tested, none of which performed as well as Bugbuster, indicating that the protein is sensitive to mechanical lysis. The IMAC protocol was adjusted to see if incubation with resin overnight or increasing the volume of resin used made a difference to the purification process to improve recovery. None of these optimisations were able to fully eliminate contaminants or completely prevent loss of target protein during IMAC wash stages. A major difference in protocol to consider would be utilising an insect expression system similar to the successfully published crystal structure (Koliopoulos et al., 2018), rather than the bacterial expression system used here. Whilst bacterial systems provide high protein yields, and do so in the shortest amount of time, insect expression systems are often more suitable for expressing intracellular proteins and multi-protein complexes and allow proteins to maintain their functional activity (Verma et al., 1998). Because the success of SEC-SAXS demands monodisperse (high purity) samples at high concentrations, (6-10mg/mL), ensuring the purity of the TRIM25 protein is paramount for any future attempts at characterising the biophysical properties and interactions of this TRIM25 CC-PRYSPRY domain.

Interestingly, Alphafold3 predictions for full length wild type TRIM25 and LASV NP showed a different result than that for the TRIM25 CC-PRYSPRY fragment complexed with LASV NP N-terminal core domain. Whilst the NP N-terminal core domain is still implicated, a different region is shown to be involved in the interaction; the RNA binding region does not appear to play any role at the binding interface between the two full length proteins. This contradicts the fragment complex prediction in which the RNA binding domain of LASV NP appears to be the crucial for the binding of LASV NP to

the pocket created by the TRIM25 residues labelled in figure 5.9D. Importantly the main difference between the two predicted complex models is the TRIM25 domain implicated, with the full length complex predicting a CC domain interaction and the fragment predicting the PRYSPRY domain. This difference in predicted model would not have affected the design of the structural study as all domains predicted to be involved were included in fragments (NP N-terminal core domain and TRIM25 CC-PRYSPRY). Currently there is no way to directly computationally analyse the likelihood of each predicted complex and the only way to infer which complex may be more accurate is in comparing the confidence Alphafold3 has in predicted conformation of the structures involved in the interface regions. A comparison between the fragment prediction (figure 5.7) and the full length prediction (figure 5.10) indicates that the structures involved in interaction between fragments has lower confidence of the predicted conformation C-terminal end of the NP fragment which is the section binding within the TRIM25 PRYSPRY pocket; whereas in the full length prediction, all residues and conformations involved along the NP N-terminal core domain and the TRIM25 CC domain are predicted with high confidence. The results of the co-immunoprecipitation between MOPV NP and TRIM25 mutants (figure 4.22) and the infection assay (figure 4.23) do support the fragment model over the full-length model as the SPRY domain is heavily implicated in the TRIM25-NP interaction as shown by the recovery of viral replication for the ΔSPRY mutant. Although further research would need to include a ΔCC domain mutant if structurally viable. The inability to gauge which of the predicted complex conformations has higher accuracy and confidence further highlights the need for high resolution crystal structures for the TRIM25-NP complex.

The next step in this investigation would be to utilise X-ray crystallography of these NP and TRIM25 protein fragments to obtain high resolution complex structures to determine the protein-protein interface. This would eliminate the need of utilising predictive modelling such as those from

Alphafold3, with their intrinsic limitation, especially for protein complexes, given the limited representation in the training databases. This information would also render the need for SEC-SAXS data redundant although high resolution structural data would allow for validation of SEC-SAXS data.

Chapter 6

Thesis conclusion

6 Thesis Conclusion

The main objective of this thesis was to identify host restriction factors of Old-World group arenavirus nucleoproteins through the use of BioID-derived proximity labelling and investigate their effectiveness as anti-arenaviral proteins. For this to be achieved, three main research objectives outlined in section 1.5 were established. Of these objectives, two were successfully completed. The third research objective focussing on structural and biophysical interactions between TRIM25 and arenavirus NP, whilst not achieved in totality, this research did yield successful and informative results and optimised the experimental and analytical processes involved to form a strong foundation for future research attempts.

Identification of the host interactome of Old-World arenavirus Nucleoproteins

Old world pathogenic arenaviruses such as LASV display a broad range of symptoms in individuals infected. Symptoms can range from completely asymptomatic to fatal viral haemorrhagic fever. Of the four proteins comprising arenaviruses, NP is the most abundant and the most multifunctional. NP plays critical roles in viral pathogenicity, including immune evasion and immune suppression, making it a key determinant of the virus's ability to establish and sustain infection. The interactions that arenavirus NP has with the infected host are poorly understood and previously, no proteomic data of the host interactomes of LASV, LUJV, or MOPV NP existed. Likewise, there is little research into the differences between human pathogenic arenaviruses such as LASV and LUJV and non-human pathogenic such as MOPV. Research elucidating these differences is key to further understanding of the mechanisms behind arenavirus pathogenicity. The first step in the research comprising this thesis was successfully utilising BioID2 based proteomic screens in tandem with mass

spectrometry analysis to identify those host proteins in close proximity with LASV, LUJV, and MOPV arenavirus NP. This data was then used to build host interactomes for each arenavirus NP investigated and identify the key immunological networks enriched by host interaction with NP. Although many protein networks and biological functions were identified for each interactome, the main proteins that stood out for potential future research were zinc antiviral protein (ZAP) and TRIM25. These two proteins were conserved across host interactomes for all three of the OW group arenavirus NPs of interest and had already been identified as antivirals for EBOV and Influenza.

Validation of host proteins interacting with arenavirus NP as potential host restriction factors of arenavirus infection.

For this research it was necessary to understand how the proteins involved behave within the cell. To achieve this immunofluorescence was used to show the localisations of LASV, LUJV, and MOPV NP within the cell and comparisons were able to be made. LASV and MOPV NP localise similarly as is expected due to the genetic similarity although LASV exhibits a more diffuse cytoplasmic distribution whilst MOPV displays some perinuclear speckling. LUJV exhibits more intense perinuclear speckling and also displays the same cytoplasmic distribution seen in the two other NPs. The localisations of endogenous TRIM25 and ZAP were shown, both displaying diffuse cytoplasmic distribution, and that they co-localise was also shown. Co-transfection of TRIM25 and either ZAP-Long or ZAP-Short was used to show that TRIM25 changes its normal localisation to match that of ZAP. This interaction between both ZAP isoforms and TRIM25 was experimentally confirmed through co-immunoprecipitation, validating previously published data that ZAP and TRIM25 are co-factors. Interactions between endogenous ZAP and TRIM25 with LASV, LUJV, and MOPV NP were also shown by co-immunoprecipitation.

Infection assays using MOPV showed that both ZAP and TRIM25 play crucial roles in the host innate immune response against arenavirus infection and

both are potent inhibitors of arenavirus replication. TRIM25 has been identified as a more potent host restriction factor compared to ZAP in the context of arenaviruses. Due to the decreased frequency of CpG dinucleotides in the arenavirus genome, it was hypothesised that ZAP-Short would be the more important isoform for ZAP-mediated arenavirus inhibition. The Long isoform utilises the PARP-like domain to target CpG rich regions on viral RNA to mark them for degradation. Due to the lack of these regions in arenavirus sequences, ZAP-Long ought not to be as effective and the ZAP-Short isoform, displaying different cellular localisation and thus suggesting an ability to reach different pools of viral RNA to target might be more involved. MOPV infection preferentially induced ZAP-Short isoform expression although infection assays did not show a significant difference in the effectiveness of either isoform as arenavirus restriction factors, with both displaying equal levels of potency.

Through the use of TRIM25 mutants, the SPRY domain was implicated in the TRIM25-NP interaction. When TRIM25 was missing the SPRY domain, the ΔSPRY mutant, MOPV replication was significantly enhanced, almost to the same extent as seen in total TRIM25 knockout cells.

Experimental and theoretical biophysical and structural methods to elucidate the molecular details of the interactions of NP with identified host factors

In the absence of X-ray crystallography data to provide crystal structures for individual proteins and complexes, the use of predictive modelling software such as AlphaFold3, combined with experimental techniques like SEC-SAXS, offers increasingly valuable insights. Given the constraints of time and challenges in achieving successful protein purification, these methods were employed as alternative approaches to characterize the structural and biophysical properties of the studied proteins.

The focus of this research was on both the individual and complex-formation SEC-SAXS analysis of LASV, LUJV, and MOPV NP N-terminal core

domain fragments, and the TRIM25 CC-PRYSPRY fragment. From the SEC-SAXS which generated an envelope for a very large globular protein of about 318kDa, it has been shown that the LASV NP fragment is likely forming large oligomers, with the maximum likely being that of an octamer. This research supports previous findings that wild type LASV NP forms hexamers (Hastie et al., 2011c).

For the LUJV NP fragment, SEC-SAXS suggested that this remained as a single monomeric unit. The protein envelope generated from the research was a near perfect fit for the Alphafold3 predicted LUJV NP fragment structure. Although an experimentally determined crystal structure for the N-terminal core domain of LUJV NP is not yet available, the strong agreement between the Alphafold3 predicted structure and the SEC-SAXS envelope supports not only the accuracy of the prediction but the power of combining both of these theoretical and experimental techniques.

The MOPV NP fragment behaved in a similar way to the LASV NP fragment in that SEC-SAXS analysis generated a larger than expected envelope. The size and shape of the envelope suggests that the MOPV NP fragment forms either trimers or tetramers. The differences in behaviour observed between the different NP N-terminal core domains raises questions about the mechanisms underpinning the observed differences in pathogenicity of these viruses. Research into how NP domain behaviour influences factors such as RNP formation, or accessibility to different pools of host RNA may expand upon current understanding of what affects arenavirus pathogenicity.

The difficulties faced in successfully purifying the TRIM25 CC-PRYSPRY fragment hindered research into the nature of the interaction between TRIM25 and arenavirus NP. Some of the purification challenges encountered included multiple attempts with different expression systems such as pET15b and pET28b vectors, standard small-scale culture, and large-scale fermentation, different buffer compositions for lysis and IMCA, and AKTA settings. Despite these attempts at optimisation, the TRIM25 CC-

PRYSPRY fragment yielded low protein concentrations containing contaminants. Whilst SEC-SAXS analysis generated a shape resembling a dimer, robustness was lacking due to initial sample impurities and so the confidence in the accuracy of this data is low. Unfortunately, this also rendered any of the data produced for the TRIM25 CC-PRYSPRY complex formation with arenavirus NP N-terminal core domains impossible to analyse. Future research will have to first focus on improving the purification protocols to achieve monodisperse, highly concentrated samples for SEC-SAXS or X-ray crystallography. Insect expression systems could be a consideration due to their proven success in purifying this fragment (Koliopoulos et al., 2018).

X-ray crystallography was an ambition of this research, time and experimental difficulties meant this was not achievable. In future, high-resolution structural studies are needed to determine the exact protein-protein interfaces between TRIM25 and arenavirus NP, and also to not only validate SEC-SAXS findings, but also eliminate the reliance on predictive models such as Alphafold3. To ensure this is possible, high-purity samples are required to improve the accuracy and robustness of biophysical and interaction analyses.

In conclusion, the research results presented in this thesis have advanced the understanding of host interactions with arenavirus NP and have also identified a number of potential targets for antiviral strategies and future research. The host interactomes identified through the use of BioID2 proximity labelling have highlighted key immunological pathways functions that NP targets and to better understand the mechanisms behind the broad multifunctionality. Host proteins which have highly conserved interaction across multiple arenavirus NPs (including New World) have been identified such as DDX1, TRIM25, and ZAP, providing targets for potential therapeutics. This work also confirmed that TRIM25 and ZAP play pivotal roles in restricting arenavirus replication as demonstrated through infection assays. Experimental results provided significant insight into the

localisation and interaction dynamics of ZAP and TRIM25 with arenavirus NPs and the SPRY domain of TRIM25 was identified as critical to interaction with NP. Structural studies employing the use of SEC-SAXS have provided preliminary insights into the behaviour of NP N-terminal core domain fragments, revealing interesting differences in oligomerisation states across LASV, LUJV, and MOPV NP. Although challenges in protein purification prevented full characterisation of the TRIM25 CC-PRYSPRY and NP interaction, optimisation efforts and initial SEC-SAXS results have laid the foundations for future structural and biophysical investigations.

The research comprising this thesis successfully identified for the first time that TRIM25 and ZAP inhibit arenavirus replication, with both having well established roles in the human innate immune response to infection. Much needed therapeutic strategies for arenavirus infection could leverage this knowledge by enhancing the activity or expression of TRIM25 and ZAP. This could be achieved through using small molecules/drug compounds that stabilise these proteins or boost their antiviral functions. Additionally, understanding the mechanisms by which arenaviruses evade TRIM25, and ZAP could guide the design of viral inhibitors that block evasion. This host-targeted approach may reduce the likelihood of resistance, making it a promising avenue for developing broad-spectrum antivirals against arenaviruses and other similar pathogens.

Chapter 7
Supplementary Data

7 Supplementary Data

7.1 List of proteins identified as interacting with MOPV NP from the University of Birmingham AMSF

CYP51A1	*Lanosterol 14-alpha demethylase*
AAAS	*Aladin*
ABCB7	*ATP-binding cassette sub-family B member 7, mitochondrial*
ACTBL2	*Beta-actin-like protein 2*
ACTR1A	*Alpha-centractin*
ACTR2	*Actin-related protein 2*
ADRM1	*Proteasomal ubiquitin receptor ADRM1*
AKAP1	*A-kinase anchor protein 1, mitochondrial*
ALYREF	*THO complex subunit 4*
ARPC3	*Actin-related protein 2/3 complex subunit 3*
ATXN2L	*Ataxin-2-like protein*
BCAS2	*BCAS2, pre-mRNA processing factor; Pre-mRNA-splicing factor SPF27*
BLVRA	*Biliverdin reductase A*
BMS1	*Ribosome biogenesis protein BMS1 homolog*
BZW1	*Basic leucine zipper and W2 domain-containing protein 1*
C2orf44	*WD repeat and coiled-coil-containing protein*
CCAR2	*Cell cycle and apoptosis regulator protein 2*
CDK4	*Cyclin-dependent kinase 4*
CDKN2A	*Cyclin-dependent kinase inhibitor 2A*
CENPF	*Centromere protein F*

CENPM	*Centromere protein M*
CEP68	*Centrosomal protein of 68 kDa*
CERS6	*Sphingoid base n-palmitoyltransferase*
CHERP	*Calcium homeostasis endoplasmic reticulum protein*
CHTOP	*Chromatin target of PRMT1 protein*
CIRH1A	*U3 small nucleolar RNA-associated protein 4 homologue*
CKAP5	*Cytoskeleton-associated protein 5*
CNP	*2',3'-cyclic-nucleotide 3'-phosphodiesterase*
COMMD4	*COMM domain-containing protein 4*
CPSF3L	*Integrator complex subunit 11*
CPSF7	*Cleavage and polyadenylation specificity factor subunit 7*
CUL3	*Cullin-3*
CXXC1	*CXXC-type zinc finger protein 1*
CYFIP2	*Cytoplasmic FMR1-interacting protein 2*
DCTN1	*Dynactin subunit 1*
DDB1	*DNA damage-binding protein 1*
DDX1	*ATP-dependent RNA helicase DDX1*
DDX47	*Probable ATP-dependent RNA helicase DDX47*
DDX50	*ATP-dependent RNA helicase DDX50*
DHFR	*Dihydrofolate reductase*
DLST	*2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide succinyltransferase)*
DNAJB1	*Dnaj heat shock protein family (hsp40) member b1*
DNTTIP2	*Deoxynucleotidyltransferase terminal-interacting protein 2*

DPM1	*Dolichol-phosphate mannosyltransferase subunit 1*
DUT	*Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial*
DVL2	*Segment polarity protein dishevelled homolog DVL-2*
DVL3	*Segment polarity protein dishevelled homolog DVL-3*
ECHS1	*Enoyl-CoA hydratase, mitochondrial*
EIF2B1	*Translation initiation factor eIF-2B subunit alpha*
EIF2B2	*Translation initiation factor eIF-2B subunit beta*
EIF2S2	*Eukaryotic translation initiation factor 2 subunit 2*
EIF4B	*Eukaryotic translation initiation factor 4B*
EIF5	*Eukaryotic translation initiation factor 5*
ELP5	*Elongator complex protein 5;*
ETF1	*Eukaryotic peptide chain release factor subunit 1*
FAM162A	*Family with sequence similarity 162 member a*
FAM50A	*Family with sequence similarity 50 member a*
FANCI	*Fanconi anemia group I protein*
FHL2	*Four and a half LIM domains protein 2*
FITM2	*Fat storage-inducing transmembrane protein 2*
G6PD	*Glucose-6-phosphate 1-dehydrogenase*
GEMIN2	*Gem-associated protein 2*
GFM2	*Ribosome-releasing factor 2, mitochondrial*
GINS2	*DNA replication complex GINS protein PSF2*
GLMN	*Glomulin, fkbp associated protein*
GTF3C1	*General transcription factor 3C polypeptide 1*

H1F0	*H1 histone family member 0*
HADHB	*hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit beta*
HDLBP	*High density lipoprotein binding protein*
HGH1	*Protein HGH1 homologue*
HIST1H1D	*Histone cluster 1 h1 family member d*
HSDL1	*Inactive hydroxysteroid dehydrogenase-like protein 1*
HSPA1L	*Heat shock 70 kDa protein 1-like*
HSPA2	*Heat shock-related 70 kDa protein 2*
HSPB1	*Heat shock protein family b (small) member 1*
HSPE1	*Heat shock protein family e (hsp10) member 1*
ILVBL	*Acetolactate synthase-like protein*
INTS7	*Integrator complex subunit 7*
IPO7	*Importin-7*
IPO9	*Importin-9*
IRF2BP1	*Interferon regulatory factor 2-binding protein 1*
KRT15	*Keratin, type I cytoskeletal 15*
LANCL2	*LanC-like protein 2*
LARP4B	*La-related protein 4B*
LIG3	*DNA ligase 3*
LONP1	*Lon protease homolog, mitochondrial*
LYPLA1	*Acyl-protein thioesterase 1*
MAVS	*Mitochondrial antiviral-signaling protein*
MCM2	*Minichromosome maintenance complex component 2*

MRPS14	*28S ribosomal protein S14, mitochondrial*
MRPS2	*28S ribosomal protein S2, mitochondrial*
MRPS9	*28S ribosomal protein S9, mitochondrial*
MSN	*Moesin*
MT-CO2	*Mitochondrially encoded cytochrome c oxidase ii*
MT-ND5	*NADH-ubiquinone oxidoreductase chain 5*
NAP1L4	*Nucleosome assembly protein 1-like 4*
NCBP1	*Nuclear cap-binding protein subunit 1*
NDRG1	*N-myc downstream regulated 1*
NDUFA13	*NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13*
NDUFB10	*NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10*
NIF3L1	*NGG1 interacting factor 3 like 1*
NOP56	*Nucleolar protein 56*
NT5DC2	*5'-nucleotidase domain-containing protein 2*
NUP133	*Nuclear pore complex protein Nup133*
OSBPL6	*Oxysterol-binding protein-related protein 6*
OTUB1	*Ubiquitin thioesterase OTUB1*
PATL1	*Protein PAT1 homolog 1*
PCBP3	*Poly(rc)-binding protein 3/4*
PDHB	*Pyruvate dehydrogenase E1 component subunit beta, mitochondrial*
PGRMC2	*Membrane-associated progesterone receptor component 2*
PHF6	*PHD finger protein 6*
PICALM	*Phosphatidylinositol-binding clathrin assembly protein*

PITPNB	*Phosphatidylinositol transfer protein beta isoform*
PLD3	*Phospholipase D3*
POLR2B	*DNA-directed RNA polymerase II subunit RPB2*
PPP1CB	*Serine/threonine-protein phosphatase PP1-beta catalytic subunit*
PRKAG1	*5'-AMP-activated protein kinase subunit gamma-1*
PSAT1	*Phosphoserine aminotransferase*
PSMA1	*Proteasome subunit alpha type-1*
PSMA4	*Proteasome subunit alpha type-4;*
PSMB2	*Proteasome subunit beta type-2*
PSMB4	*Proteasome subunit beta type-4*
PSMC2	*26S proteasome regulatory subunit 7*
PSPC1	*Paraspeckle component 1*
PUS10	*Putative tRNA pseudouridine synthase Pus10*
PYCR1	*Pyrroline-5-carboxylate reductase 1, mitochondrial*
PYGB	*Glycogen phosphorylase, brain form*
PYGL	*Glycogen phosphorylase, liver form*
RAB35	*RAB35, member RAS oncogene family; Ras-related protein Rab-35*
RAB3C	*Ras-related protein Rab-3C*
RANBP10	*Ran-binding protein 10*
RAP1B	*Ras-related protein Rap-1b*
RBM27	*RNA-binding protein 27*
RDX	*Radixin*
RELA	*Rela proto-oncogene, nf-kb subunit; Transcription factor p65*

RNMTL1	*16S rRNA (guanosine(1370)-2'-O)-methyltransferase*
RUVBL1	*RuvB-like 1*
RUVBL2	*RuvB-like 2*
SCD5	*Stearoyl-coa desaturase (delta-9 desaturase)*
SEC24B	*Protein transport protein Sec24B*
SF1	*Splicing factor 1*
SKIV2L2	*Superkiller viralicidic activity 2-like 2*
SLC25A1	*Tricarboxylate transport protein, mitochondrial*
SLC7A5	*Large neutral amino acids transporter small subunit 1*
SMC1A	*Structural maintenance of chromosomes protein 1A*
SMC2	*Structural maintenance of chromosomes protein 2*
SNRPA	*U1 small nuclear ribonucleoprotein A*
SNX27	*Sorting nexin family member 27*
SPEN	*Msx2-interacting protein*
SSR1	*Translocon-associated protein subunit alpha*
STAG2	*Cohesin subunit SA-2*
STAU1	*Double-stranded RNA-binding protein Staufen homolog 1*
STIP1	*Stress-induced-phosphoprotein 1*
SYNGR2	*Synaptogyrin-2*
TACO1	*Translational activator of cytochrome c oxidase 1*
TCOF1	*Treacle ribosome biogenesis factor 1*
TEN1-CDK3	*TEN1-CDK3 readthrough (NMD candidate)*
THG1L	*Trna-histidine guanylyltransferase 1 like*

TIMM44	*Mitochondrial import inner membrane translocase subunit TIM44*
TJP1	*Tight junction protein ZO-1*
TJP2	*Tight junction protein ZO-2*
TK1	*Thymidine kinase, cytosolic*
TMEM167A	*Transmembrane protein 167a*
TMEM263	*Transmembrane protein 263*
TOMM20	*Mitochondrial import receptor subunit TOM20 homolog*
TOMM7	*Mitochondrial import receptor subunit TOM7 homolog*
TP53	*Cellular tumor antigen p53*
TRIM25	*E3 ubiquitin/ISG15 ligase TRIM25*
TRIP13	*Pachytene checkpoint protein 2 homolog*
TSPO	*Translocator protein*
TSR3	*TSR3, acp transferase ribosome maturation factor*
TTF2	*Transcription termination factor 2;*
TTLL12	*Tubulin--tyrosine ligase-like protein 12*
UBR4	*E3 ubiquitin-protein ligase UBR4*
UCHL3	*Ubiquitin carboxyl-terminal hydrolase isozyme L3*
UGP2	*UTP--glucose-1-phosphate uridylyltransferase*
VCL	*Vinculin*
VPS33A	*Vacuolar protein sorting-associated protein 33A*
VPS35	*Vacuolar protein sorting-associated protein 35*
YES1	*Yes proto-oncogene 1, src family tyrosine kinase*
YTHDC1	*YTH domain-containing protein 1*

YWHAE	*14-3-3 protein epsilon*
YWHAG	*14-3-3 protein gamma*
ZC3HAV1	*Zinc finger CCCH-type antiviral protein 1*
ZNF326	*DBIRD complex subunit ZNF326*
ZNF512	*Zinc finger protein 512;*

Table 7.1 List of proteins identified as interacting with MOPV NP from the University of Birmingham AMSF

Chapter 8
Bibliography

8 Bibliography

ABERNATHY, E. & GLAUNSINGER, B. 2015. Emerging roles for RNA degradation in viral replication and antiviral defense. *Virology,* 479-480**,** 600-8.

ABRAMSON, J., ADLER, J., DUNGER, J., EVANS, R., GREEN, T., PRITZEL, A., RONNEBERGER, O., WILLMORE, L., BALLARD, A. J., BAMBRICK, J., BODENSTEIN, S. W., EVANS, D. A., HUNG, C.-C., O'NEILL, M., REIMAN, D., TUNYASUVUNAKOOL, K., WU, Z., ŽEMGULYTĖ, A., ARVANITI, E., BEATTIE, C., BERTOLLI, O., BRIDGLAND, A., CHEREPANOV, A., CONGREVE, M., COWEN-RIVERS, A. I., COWIE, A., FIGURNOV, M., FUCHS, F. B., GLADMAN, H., JAIN, R., KHAN, Y. A., LOW, C. M. R., PERLIN, K., POTAPENKO, A., SAVY, P., SINGH, S., STECULA, A., THILLAISUNDARAM, A., TONG, C., YAKNEEN, S., ZHONG, E. D., ZIELINSKI, M., ŽÍDEK, A., BAPST, V., KOHLI, P., JADERBERG, M., HASSABIS, D. & JUMPER, J. M. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature,* 630**,** 493-500.

ABUDUREXITI, A., ADKINS, S., ALIOTO, D., ALKHOVSKIY, S., ZUPANC, T., BALLINGER, M., BENTE, D., BEER, M., BERGERON, E., BLAIR, C., BRIESE, T., BUCHMEIER, M., BURT, F., CALISHER, C., CHÁNG, C., CHARREL, R., CHOI, I.-R., CLEGG, J., TORRE, J. & KUHN, J. 2019. Taxonomy of the order Bunyavirales: update 2019. *Archives of Virology,* 164.

AFRASIABI, A., ALINEJAD-ROKNY, H., KHOSH, A., RAHNAMA, M., LOVELL, N., XU, Z. & EBRAHIMI, D. 2022. The low abundance of CpG in the SARS-CoV-2 genome is not an evolutionarily signature of ZAP. *Scientific Reports,* 12.

AGNIHOTHRAM, S. S., YORK, J. & NUNBERG, J. H. 2006. Role of the Stable Signal Peptide and Cytoplasmic Domain of G2 in Regulating Intracellular Transport of the Junín Virus Envelope Glycoprotein Complex. *Journal of Virology,* 80**,** 5189-5198.

AKHIWU, H., YILTOK, E., EBONYI, A., GOMEREP, S., SHEHU, N., AMAECHI, E., ONUKAK, A., IDUH, P., OYAGBEMI, A. & OMAME, G. 2018. Lassa fever outbreak in adolescents in North Central Nigeria: report of cases. Elsevier.

AKHUEMOKHAN, O. C., EWAH-ODIASE, R. O., AKPEDE, N., EHIMUAN, J., ADOMEH, D. I., ODIA, I., OLOMU, S. C., PAHLMANN, M., BECKER-ZIAJA, B. & HAPPI, C. T. 2017. Prevalence of Lassa Virus Disease (LVD) in Nigerian children with fever or fever and convulsions in an endemic area. *PLoS neglected tropical diseases,* 11**,** e0005711.

AKPEDE, G. O., ASOGUN, D. A., OKOGBENIN, S. A. & OKOKHERE, P. O. 2018. Lassa fever outbreaks in Nigeria. *Expert review of anti-infective therapy,* 16**,** 663-666.

ÁLVAREZ, L., HAUBRICH, K., ISELIN, L., GILLIOZ, L., RUSCICA, V., LAPOUGE, K., AUGSTEN, S., HUPPERTZ, I., CHOUDHURY, N. R., SIMON, B., MASIEWICZ, P., LETHIER, M., CUSACK, S., RITTINGER, K., GABEL, F., LEITNER, A., MICHLEWSKI, G., HENTZE, M. W., ALLAIN, F. H. T., CASTELLO, A. & HENNIG, J. 2024. The molecular dissection of TRIM25's RNA-binding mechanism provides key insights into its antiviral activity. *Nature Communications,* 15**,** 8485.

AMARA, A. & MERCER, J. 2015. Viral apoptotic mimicry. *Nature Reviews Microbiology,* 13**,** 461-469.

ANDERSEN, K. G., SHAPIRO, B. J., MATRANGA, C. B., SEALFON, R., LIN, A. E., MOSES, L. M., FOLARIN, O. A., GOBA, A., ODIA, I., EHIANE, P. E.,

MOMOH, M., ENGLAND, E. M., WINNICKI, S., BRANCO, L. M., GIRE, S. K., PHELAN, E., TARIYAL, R., TEWHEY, R., OMONIWA, O., FULLAH, M., FONNIE, R., FONNIE, M., KANNEH, L., JALLOH, S., GBAKIE, M., SAFFA, S., KARBO, K., GLADDEN, A. D., QU, J., STREMLAU, M., NEKOUI, M., FINUCANE, H. K., TABRIZI, S., VITTI, J. J., BIRREN, B., FITZGERALD, M., MCCOWAN, C., IRELAND, A., BERLIN, A. M., BOCHICCHIO, J., TAZON-VEGA, B., LENNON, N. J., RYAN, E. M., BJORNSON, Z., MILNER, D. A., JR., LUKENS, A. K., BROODIE, N., ROWLAND, M., HEINRICH, M., AKDAG, M., SCHIEFFELIN, J. S., LEVY, D., AKPAN, H., BAUSCH, D. G., RUBINS, K., MCCORMICK, J. B., LANDER, E. S., GUNTHER, S., HENSLEY, L., OKOGBENIN, S., VIRAL HEMORRHAGIC FEVER, C., SCHAFFNER, S. F., OKOKHERE, P. O., KHAN, S. H., GRANT, D. S., AKPEDE, G. O., ASOGUN, D. A., GNIRKE, A., LEVIN, J. Z., HAPPI, C. T., GARRY, R. F. & SABETI, P. C. 2015. Clinical Sequencing Uncovers Origins and Evolution of Lassa Virus. *Cell,* 162**,** 738-50.

ANTZIN-ANDUETZA, I., MAHIET, C., GRANGER, L. A., ODENDALL, C. & SWANSON, C. M. 2017. Increasing the CpG dinucleotide abundance in the HIV-1 genomic RNA inhibits viral replication. *Retrovirology,* 14**,** 49.

ARCHER, A. M. & RICO-HESSE, R. 2002. High genetic divergence and recombination in Arenaviruses from the Americas. *Virology,* 304 2**,** 274-81.

ARIAS, J. F., IWABU, Y. & TOKUNAGA, K. 2011. Structural Basis for the Antiviral Activity of BST-2/Tetherin and Its Viral Antagonism. *Front Microbiol,* 2**,** 250.

ARMSTRONG, C. & LILLIE, R. D. 1934. Experimental Lymphocytic Choriomeningitis of Monkeys and Mice Produced by a Virus Encountered in Studies of the 1933 St. Louis Encephalitis Epidemic. *Public Health Reports (1896-1970),* 49**,** 1019.

ARMSTRONG, C. & PAUL, F. D. 1935. Benign Lymphocytic Choriomeningitis (Acute Aseptic Meningitis): A New Disease Entity. *Public Health Reports (1896-1970),* 50**,** 831-842.

ATKINSON, N. J., WITTEVELDT, J., EVANS, D. J. & SIMMONDS, P. 2014. The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. *Nucleic Acids Res,* 42**,** 4527-45.

BAGGIO, F., HETZEL, U., NUFER, L., KIPAR, A. & HEPOJOKI, J. 2021. A subpopulation of arenavirus nucleoprotein localizes to mitochondria. *Scientific Reports,* 11**,** 21048.

BAIRD, N. L., YORK, J. & NUNBERG, J. H. 2012. Arenavirus infection induces discrete cytosolic structures for RNA replication. *J Virol,* 86**,** 11301-10.

BAN, J., LEE, N.-R., LEE, N.-J., LEE, J. K., QUAN, F.-S. & INN, K.-S. 2018. Human Respiratory Syncytial Virus NS 1 Targets TRIM25 to Suppress RIG-I Ubiquitination and Subsequent RIG-I-Mediated Antiviral Signaling. *Viruses,* 10**,** 716.

BARRESI, R. & CAMPBELL, K. P. 2006. Dystroglycan: from biosynthesis to pathogenesis of human disease. *J Cell Sci,* 119**,** 199-207.

BARTON, L. L. & HYNDMAN, N. J. 2000. Lymphocytic choriomeningitis virus: reemerging central nervous system pathogen. *Pediatrics,* 105**,** E35.

BARTON, L. L. & METS, M. B. 2001. Congenital lymphocytic choriomeningitis virus infection: decade of rediscovery. *Clin Infect Dis,* 33**,** 370-4.

BENEJ, M., DANCHENKO, M., OVECKOVA, I., CERVENAK, F., TOMASKA, L., GROSSMANNOVA, K., POLCICOVA, K., GOLIAS, T. & TOMASKOVA, J. 2019. Quantitative Proteomics Reveal Peroxiredoxin Perturbation Upon Persistent Lymphocytic Choriomeningitis Virus Infection in Human Cells. *Frontiers in Microbiology,* 10.

BENTIM GÓES, L. G., FISCHER, C., ALMEIDA CAMPOS, A. C., DE CARVALHO, C., MOREIRA-SOTO, A., AMBAR, G., RUCKERT DA ROSA, A., DE OLIVEIRA, D. C., JO, W. K., CRUZ-NETO, A. P., PEDRO, W. A., QUEIROZ, L. H., MINOPRIO, P., DURIGON, E. L. & DREXLER, J. F. 2022. Highly Diverse Arenaviruses in Neotropical Bats, Brazil. *Emerg Infect Dis,* 28**,** 2528-2533.

BHATTACHARYYA, S. 2014. Can't RIDD off viruses. *Frontiers in Microbiology,* 5.

BICK, M. J., CARROLL, J. W., GAO, G., GOFF, S. P., RICE, C. M. & MACDONALD, M. R. 2003. Expression of the zinc-finger antiviral protein inhibits alphavirus replication. *J Virol,* 77**,** 11555-62.

BODEWES, R., KIK, M. J. L., RAJ, V. S., SCHAPENDONK, C. M. E., HAAGMANS, B. L., SMITS, S. L. & OSTERHAUS, A. 2013. Detection of novel divergent arenaviruses in boid snakes with inclusion body disease in The Netherlands. *J Gen Virol,* 94**,** 1206-1210.

BONTHIUS, D. J. 2012. Lymphocytic choriomeningitis virus: an underrecognized cause of neurologic disease in the fetus, child, and adult. *Semin Pediatr Neurol,* 19**,** 89-95.

BORROW, P., MARTÍNEZ-SOBRIDO, L. & DE LA TORRE, J. 2010. Inhibition of the Type I Interferon Antiviral Response During Arenavirus Infection. *Viruses,* 2**,** 2443-2480.

BORTZ, E., WESTERA, L., MAAMARY, J., STEEL, J., ALBRECHT, R. A., MANICASSAMY, B., CHASE, G., MARTÍNEZ-SOBRIDO, L., SCHWEMMLE, M. & GARCÍA-SASTRE, A. 2011. Host- and strain-specific regulation of influenza virus polymerase activity by interacting cellular proteins. *mBio,* 2.

BOSCH, J. A., CHEN, C. L. & PERRIMON, N. 2021. Proximity-dependent labeling methods for proteomic profiling in living cells: An update. *WIREs Developmental Biology,* 10**,** e392.

BOUSSIF, O., LEZOUALC'H, F., ZANTA, M. A., MERGNY, M. D., SCHERMAN, D., DEMENEIX, B. & BEHR, J. P. 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proceedings of the National Academy of Sciences,* 92**,** 7297-7301.

BOWEN, M. D., PETERS, C. J. & NICHOL, S. T. 1997. Phylogenetic analysis of the Arenaviridae: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. *Mol Phylogenet Evol,* 8**,** 301-16.

BOWEN, M. D., ROLLIN, P. E., KSIAZEK, T. G., HUSTAD, H. L., BAUSCH, D. G., DEMBY, A. H., BAJANI, M. D., PETERS, C. J. & NICHOL, S. T. 2000. Genetic diversity among Lassa virus strains. *Journal of virology,* 74**,** 6992-7004.

BOWICK, G. C., SPRATT, H. M., HOGG, A. E., ENDSLEY, J. J., WIKTOROWICZ, J. E., KUROSKY, A., LUXON, B. A., GORENSTEIN, D. G. & HERZOG, N. K. 2009. Analysis of the differential host cell nuclear proteome induced by attenuated and virulent hemorrhagic arenavirus infection. *J Virol,* 83**,** 687-700.

BOWIE, A. G. & UNTERHOLZNER, L. 2008. Viral evasion and subversion of pattern-recognition receptor signalling. *Nature Reviews Immunology,* 8**,** 911-922.

BRANCO, L. M., BOISEN, M. L., ANDERSEN, K. G., GROVE, J. N., MOSES, L. M., MUNCY, I. J., HENDERSON, L. A., SCHIEFFELLIN, J. S., ROBINSON, J. E. & BANGURA, J. J. 2011. Lassa hemorrhagic fever in a late term pregnancy from northern Sierra Leone with a positive maternal outcome: case report. *Virology Journal,* 8**,** 1-14.

BRUNOTTE, L., KERBER, R., SHANG, W., HAUER, F., HASS, M., GABRIEL, M., LELKE, M., BUSCH, C., STARK, H., SVERGUN, D. I., BETZEL, C., PERBANDT, M. & GÜNTHER, S. 2011a. Structure of the Lassa virus nucleoprotein revealed by X-ray crystallography, small-angle X-ray scattering, and electron microscopy. *J Biol Chem,* 286**,** 38748-38756.

BRUNOTTE, L., LELKE, M., HASS, M., KLEINSTEUBER, K., BECKER-ZIAJA, B. & GÜNTHER, S. 2011b. Domain Structure of Lassa Virus L Protein. *Journal of Virology,* 85**,** 324-333.

BUBA, M. I., DALHAT, M. M., NGUKU, P. M., WAZIRI, N., MOHAMMAD, J. O., BOMOI, I. M., ONYIAH, A. P., ONWUJEI, J., BALOGUN, M. S. & BASHORUN, A. T. 2018. Mortality among confirmed Lassa fever cases during the 2015–2016 outbreak in Nigeria. *American journal of public health,* 108**,** 262-264.

BUCHMEIER, M. 2007. Arenaviridae: the viruses and their replication. *Fields virology***,** 1792-1827.

BUCHMEIER, M. J. 2002. Arenaviruses: Protein Structure and Function. *In:* OLDSTONE, M. B. A. (ed.) *Arenaviruses I: The Epidemiology, Molecular and Cell Biology of Arenaviruses.* Berlin, Heidelberg: Springer Berlin Heidelberg.

BUCHMEIER, M. J., SOUTHERN, P. J., PAREKH, B. S., WOODDELL, M. K. & OLDSTONE, M. B. 1987. Site-specific antibodies define a cleavage site conserved among arenavirus GP-C glycoproteins. *Journal of Virology,* 61**,** 982-985.

BUCKLEY, S. M. & CASALS, J. 1970. Lassa fever, a new virus disease of man from West Africa. 3. Isolation and characterization of the virus. *Am J Trop Med Hyg,* 19**,** 680-91.

BUCKMASTER, M. V. & GOFF, S. P. 2022. Riplet Binds the Zinc Finger Antiviral Protein (ZAP) and Augments ZAP-Mediated Restriction of HIV-1. *Journal of Virology,* 96**,** e00526-22.

BURRI, D. J., PASQUAL, G., ROCHAT, C., SEIDAH, N. G., PASQUATO, A. & KUNZ, S. 2012. Molecular characterization of the processing of arenavirus envelope glycoprotein precursors by subtilisin kexin isozyme-1/site-1 protease. *J Virol,* 86**,** 4935-46.

CADENA, C., AHMAD, S., XAVIER, A., WILLEMSEN, J., PARK, S., PARK, J. W., OH, S.-W., FUJITA, T., HOU, F., BINDER, M. & HUR, S. 2019. Ubiquitin-Dependent and -Independent Roles of E3 Ligase RIPLET in Innate Immunity. *Cell,* 177**,** 1187-1200.e16.

CAPUL, A. A., PEREZ, M., BURKE, E., KUNZ, S., BUCHMEIER, M. J. & TORRE, J. C. D. L. 2007. Arenavirus Z-Glycoprotein Association Requires Z Myristoylation but Not Functional RING or Late Domains. *Journal of Virology,* 81**,** 9451-9460.

CARNEC, X., BAIZE, S., REYNARD, S., DIANCOURT, L., CARO, V., TORDO, N. & BOULOY, M. 2011. Lassa Virus Nucleoprotein Mutants Generated by Reverse Genetics Induce a Robust Type I Interferon Response in Human Dendritic Cells and Macrophages. *Journal of Virology,* 85**,** 12093-12097.

CASABONA, J. C., LEVINGSTON MACLEOD, J. M., LOUREIRO, M. E., GOMEZ, G. A. & LOPEZ, N. 2009. The RING domain and the L79 residue of Z protein are

involved in both the rescue of nucleocapsids and the incorporation of glycoproteins into infectious chimeric arenavirus-like particles. *J Virol,* 83**,** 7029-39.

CASTELLO, A., FISCHER, B., EICHELBAUM, K., HOROS, R., BECKMANN, BENEDIKT M., STREIN, C., DAVEY, NORMAN E., HUMPHREYS, DAVID T., PREISS, T., STEINMETZ, LARS M., KRIJGSVELD, J. & HENTZE, MATTHIAS W. 2012. Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell,* 149**,** 1393-1406.

CAUDILL, V. R., QIN, S., WINSTEAD, R., KAUR, J., TISTHAMMER, K., PINEDA, E. G., SOLIS, C., COBEY, S., BEDFORD, T., CARJA, O., EGGO, R. M., KOELLE, K., LYTHGOE, K., REGOES, R., ROY, S., ALLEN, N., AVILES, M., BAKER, B. A., BAUER, W., BERMUDEZ, S., CARLSON, C., CASTELLANOS, E., CATALAN, F. L., CHEMEL, A. K., ELLIOT, J., EVANS, D., FIUTEK, N., FRYER, E., GOODFELLOW, S. M., HECHT, M., HOPP, K., HOPSON, E. D., JR., JABERI, A., KINNEY, C., LAO, D., LE, A., LO, J., LOPEZ, A. G., LÓPEZ, A., LORENZO, F. G., LUU, G. T., MAHONEY, A. R., MELTON, R. L., NASCIMENTO, G. D., PRADHANANGA, A., RODRIGUES, N. S., SHIEH, A., SIMS, J., SINGH, R., SULAEMAN, H., THU, R., TRAN, K., TRAN, L., WINTERS, E. J., WONG, A. & PENNINGS, P. S. 2020. CpG-creating mutations are costly in many human viruses. *Evol Ecol,* 34**,** 339-359.

CHANG, L.-W. & JACOBSON, E. R. 2010. Inclusion Body Disease, A Worldwide Infectious Disease of Boid Snakes: A Review. *Journal of Exotic Pet Medicine,* 19**,** 216-225.

CHARREL, R. N. & DE LAMBALLERIE, X. 2003. Arenaviruses other than Lassa virus. *Antiviral Res,* 57**,** 89-100.

CHARREL, R. N., FELDMANN, H., FULHORST, C. F., KHELIFA, R., DE CHESSE, R. & DE LAMBALLERIE, X. N. 2002. Phylogeny of New World arenaviruses based on the complete coding sequences of the small genomic segment identified an evolutionary lineage produced by intrasegmental recombination. *Biochemical and biophysical research communications,* 296 5**,** 1118-24.

CHARREL, R. N., LEMASSON, J.-J., GARBUTT, M., KHELIFA, R., MICCO, P. D., FELDMANN, H. & LAMBALLERIE, X. D. 2003. New insights into the evolutionary relationships between arenaviruses provided by comparative analysis of small and large segment sequences. *Virology,* 317**,** 191-196.

CHARRON, G., LI, M. M. H., MACDONALD, M. R. & HANG, H. C. 2013. Prenylome profiling reveals S-farnesylation is crucial for membrane targeting and antiviral activity of ZAP long-isoform. *Proceedings of the National Academy of Sciences,* 110**,** 11085-11090.

CHE, Y. & KHAVARI, P. A. 2017. Research techniques made simple: emerging methods to elucidate protein interactions through spatial proximity. *Journal of Investigative Dermatology,* 137**,** e197-e203.

CHEN, G., GUO, X., LV, F., XU, Y. & GAO, G. 2008. p72 DEAD box RNA helicase is required for optimal function of the zinc-finger antiviral protein. *Proceedings of the National Academy of Sciences,* 105**,** 4352-4357.

CHEN, H.-C., LIN, W.-C., TSAY, Y.-G., LEE, S.-C. & CHANG, C.-J. 2002. An RNA Helicase, DDX1, Interacting with Poly(A) RNA and Heterogeneous Nuclear Ribonucleoprotein K *. *Journal of Biological Chemistry,* 277**,** 40403-40409.

CHEN, S., XU, Y., ZHANG, K., WANG, X., SUN, J., GAO, G. & LIU, Y. 2012. Structure of N-terminal domain of ZAP indicates how a zinc-finger protein recognizes complex RNA. *Nat Struct Mol Biol,* 19**,** 430-5.

CHEN, Y.-M., SADIQ, S., TIAN, J.-H., CHEN, X., LIN, X.-D., SHEN, J.-J., CHEN, H., HAO, Z.-Y., YANG, W.-D., ZHOU, Z.-C., WU, J., LI, F., WANG, H.-W., XU, Q.-Y., WANG, W., GAO, W.-H., HOLMES, E. C. & ZHANG, Y.-Z. 2021. RNA Virome Composition Is Shaped by Sampling Ecotype. *SSRN Electronic Journal*.

CHIU, H.-P., CHIU, H., YANG, C.-F., LEE, Y.-L., CHIU, F.-L., KUO, H.-C., LIN, R.-J. & LIN, Y.-L. 2018. Inhibition of Japanese encephalitis virus infection by the host zinc-finger antiviral protein. *PLOS Pathogens,* 14**,** e1007166.

CHOI, H., LARSEN, B., LIN, Z.-Y., BREITKREUTZ, A., MELLACHERUVU, D., FERMIN, D., QIN, Z. S., TYERS, M., GINGRAS, A.-C. & NESVIZHSKII, A. I. 2011. SAINT: probabilistic scoring of affinity purification–mass spectrometry data. *Nature Methods,* 8**,** 70-73.

CHOUDHURY, N. R., HEIKEL, G., TRUBITSYNA, M., KUBIK, P., NOWAK, J. S., WEBB, S., GRANNEMAN, S., SPANOS, C., RAPPSILBER, J., CASTELLO, A. & MICHLEWSKI, G. 2017. RNA-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is required for ubiquitination. *BMC Biology,* 15**,** 105.

CHOUDHURY, N. R., TRUS, I., HEIKEL, G., WOLCZYK, M., SZYMANSKI, J., BOLEMBACH, A., DOS SANTOS PINTO, R. M., SMITH, N., TRUBITSYNA, M., GAUNT, E., DIGARD, P. & MICHLEWSKI, G. 2022. TRIM25 inhibits influenza A virus infection, destabilizes viral mRNA, but is redundant for activating the RIG-I pathway. *Nucleic Acids Res,* 50**,** 7097-7114.

CHOW, K. T., GALE, M., JR. & LOO, Y. M. 2018. RIG-I and Other RNA Sensors in Antiviral Immunity. *Annu Rev Immunol,* 36**,** 667-694.

COELHO, A. C. & GARCÍA DÍEZ, J. 2015. Biological Risks and Laboratory-Acquired Infections: A Reality That Cannot be Ignored in Health Biotechnology. *Front Bioeng Biotechnol,* 3**,** 56.

COGSWELL-HAWKINSON, A., BOWEN, R., JAMES, S., GARDINER, D., CALISHER, C. H., ADAMS, R. & SCHOUNTZ, T. 2012. Tacaribe virus causes fatal infection of an ostensible reservoir host, the Jamaican fruit bat. *J Virol,* 86**,** 5791-9.

COHEN-DVASHI, H., ISRAELI, H., SHANI, O., KATZ, A., DISKIN, R. & ROSS, S. R. 2016. Role of LAMP1 Binding and pH Sensing by the Spike Complex of Lassa Virus. *Journal of Virology,* 90**,** 10329-10338.

COOPER, D. N. & GERBER-HUBER, S. 1985. DNA methylation and CpG suppression. *Cell Differ,* 17**,** 199-205.

CORDIN, O., BANROQUES, J., TANNER, N. K. & LINDER, P. 2006. The DEAD-box protein family of RNA helicases. *Gene,* 367**,** 17-37.

CORNU, T. I. & 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. *J Virol,* 75**,** 9415-26.

CORNU, T. I. & DE LA TORRE, J. C. 2002. Characterization of the arenavirus RING finger Z protein regions required for Z-mediated inhibition of viral RNA synthesis. *J Virol,* 76**,** 6678-88.

COUROUBLE, V. V., DEY, S. K., YADAV, R., TIMM, J., HARRISON, J. J. E., RUIZ, F. X., ARNOLD, E. & GRIFFIN, P. R. 2021. Resolving the dynamic motions of SARS-CoV-2 nsp7 and nsp8 proteins using structural proteomics. *bioRxiv*.

COUTERMARSH-OTT, S., EDEN, K. & ALLEN, I. C. 2016. Beyond the inflammasome: regulatory NOD-like receptor modulation of the host immune response following virus exposure. *J Gen Virol,* 97**,** 825-838.

CROSS, R. W., WOOLSEY, C., PRASAD, A. N., BORISEVICH, V., AGANS, K. N., DEER, D. J., GEISBERT, J. B., DOBIAS, N. S., FENTON, K. A. & GEISBERT, T.

W. 2022. A recombinant VSV-vectored vaccine rapidly protects nonhuman primates against heterologous lethal Lassa fever. *Cell Reports,* 40.

CROSS, R. W., XU, R., MATASSOV, D., HAMM, S., LATHAM, T. E., GERARDI, C. S., NOWAK, R. M., GEISBERT, J. B., OTA-SETLIK, A., AGANS, K. N., LUCKAY, A., WITKO, S. E., SOUKIEH, L., DEER, D. J., MIRE, C. E., FELDMANN, H., HAPPI, C., FENTON, K. A., ELDRIDGE, J. H. & GEISBERT, T. W. 2020. Quadrivalent VesiculoVax vaccine protects nonhuman primates from viral-induced hemorrhagic fever and death. *J Clin Invest,* 130**,** 539-551.

CROSSE, K. M., MONSON, E. A., BEARD, M. R. & HELBIG, K. J. 2017. Interferon-Stimulated Genes as Enhancers of Antiviral Innate Immune Signaling. *Journal of Innate Immunity,* 10**,** 85-93.

CUEVAS, C. D., LAVANYA, M., WANG, E. & ROSS, S. R. 2011. Junín Virus Infects Mouse Cells and Induces Innate Immune Responses. *Journal of Virology,* 85**,** 11058-11068.

CUI, X., FAN, K., LIANG, X., GONG, W., CHEN, W., HE, B., CHEN, X., WANG, H., WANG, X., ZHANG, P., LU, X., CHEN, R., LIN, K., LIU, J., ZHAI, J., LIU, D. X., SHAN, F., LI, Y., CHEN, R. A., MENG, H., LI, X., MI, S., JIANG, J., ZHOU, N., CHEN, Z., ZOU, J. J., GE, D., YANG, Q., HE, K., CHEN, T., WU, Y. J., LU, H., IRWIN, D. M., SHEN, X., HU, Y., LU, X., DING, C., GUAN, Y., TU, C. & SHEN, Y. 2023. Virus diversity, wildlife-domestic animal circulation and potential zoonotic viruses of small mammals, pangolins and zoo animals. *Nat Commun,* 14**,** 2488.

D'ANTUONO, A., LOUREIRO, M. E., FOSCALDI, S., MARINO-BUSLJE, C. & LOPEZ, N. 2014. Differential contributions of tacaribe arenavirus nucleoprotein N-terminal and C-terminal residues to nucleocapsid functional activity. *J Virol,* 88**,** 6492-505.

D'CRUZ, A. A., KERSHAW, N. J., CHIANG, J. J., WANG, M. K., NICOLA, N. A., BABON, J. J., GACK, M. U. & NICHOLSON, S. E. 2013. Crystal structure of the TRIM25 B30.2 (PRYSPRY) domain: a key component of antiviral signalling. *Biochem J,* 456**,** 231-40.

D'CRUZ, A. A., KERSHAW, N. J., HAYMAN, T. J., LINOSSI, E. M., CHIANG, J. J., WANG, M. K., DAGLEY, L. F., KOLESNIK, T. B., ZHANG, J. G., MASTERS, S. L., GRIFFIN, M. D. W., GACK, M. U., MURPHY, J. M., NICOLA, N. A., BABON, J. J. & NICHOLSON, S. E. 2018. Identification of a second binding site on the TRIM25 B30.2 domain. *Biochem J,* 475**,** 429-440.

DE ANDRADE, K. Q. & CIRNE-SANTOS, C. C. 2023. Antiviral Activity of Zinc Finger Antiviral Protein (ZAP) in Different Virus Families. *Pathogens,* 12**,** 1461.

DE JONG, L., DE KONING, E. A., ROSEBOOM, W., BUNCHERD, H., WANNER, M. J., DAPIC, I., JANSEN, P. J., VAN MAARSEVEEN, J. H., CORTHALS, G. L. & LEWIS, P. J. 2017. In-culture cross-linking of bacterial cells reveals large-scale dynamic protein–protein interactions at the peptide level. *Journal of proteome research,* 16**,** 2457-2471.

DOWNS, W. G., ANDERSON, C. R., SPENCE, L., AITKEN, T. H. & GREENHALL, A. H. 1963. Tacaribe virus, a new agent isolated from Artibeus bats and mosquitoes in Trinidad, West Indies. *Am J Trop Med Hyg,* 12**,** 640-6.

EHICHIOYA, D. U., ASOGUN, D. A., EHIMUAN, J., OKOKHERE, P. O., PAHLMANN, M., ÖLSCHLÄGER, S., BECKER-ZIAJA, B., GÜNTHER, S. & OMILABU, S. A. 2012. Hospital-based surveillance for Lassa fever in Edo State, Nigeria, 2005–2008. *Tropical Medicine & International Health,* 17**,** 1001-1004.

EHICHIOYA, D. U., DELLICOUR, S., PAHLMANN, M., RIEGER, T., OESTEREICH, L., BECKER-ZIAJA, B., CADAR, D., IGHODALO, Y., OLOKOR, T. & OMOMOH, E.

2019. Phylogeography of Lassa virus in Nigeria. *Journal of virology,* 93**,** 10.1128/jvi. 00929-19.
EICHLER, R., LENZ, O., GARTEN, W. & STRECKER, T. 2006. The role of single N-glycans in proteolytic processing and cell surface transport of the Lassa virus glycoprotein GP-C. *Virology Journal,* 3**,** 41.
ELLENBERG, P., EDREIRA, M., LOZANO, M. & SCOLARO, L. 2002. Synthesis and expression of viral antigens in Vero cells persistently infected with Junin virus. *Arch Virol,* 147**,** 1543-57.
ENRIA, D. A. & BARRERA ORO, J. G. 2002. Junin virus vaccines. *Curr Top Microbiol Immunol,* 263**,** 239-61.
ESCHLI, B., QUIRIN, K., WEPF, A., WEBER, J., ZINKERNAGEL, R. & HENGARTNER, H. 2006. Identification of an N-Terminal Trimeric Coiled-Coil Core within Arenavirus Glycoprotein 2 Permits Assignment to Class I Viral Fusion Proteins. *Journal of Virology,* 80**,** 5897-5907.
FAGERBERG, L., HALLSTRÖM, B. M., OKSVOLD, P., KAMPF, C., DJUREINOVIC, D., ODEBERG, J., HABUKA, M., TAHMASEBPOOR, S., DANIELSSON, A., EDLUND, K., ASPLUND, A., SJÖSTEDT, E., LUNDBERG, E., SZIGYARTO, C. A.-K., SKOGS, M., TAKANEN, J. O., BERLING, H., TEGEL, H., MULDER, J., NILSSON, P., SCHWENK, J. M., LINDSKOG, C., DANIELSSON, F., MARDINOGLU, A., SIVERTSSON, Å., VON FEILITZEN, K., FORSBERG, M., ZWAHLEN, M., OLSSON, I., NAVANI, S., HUSS, M., NIELSEN, J., PONTEN, F. & UHLÉN, M. 2014. Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics*. *Molecular & Cellular Proteomics,* 13**,** 397-406.
FAN, L., BRIESE, T. & LIPKIN, W. I. 2010. Z Proteins of New World Arenaviruses Bind RIG-I and Interfere with Type I Interferon Induction. *Journal of Virology,* 84**,** 1785-1791.
FANG, J., PIETZSCH, C., WITWIT, H., TSAPRAILIS, G., CRYNEN, G., CHO, K. F., TING, A. Y., BUKREYEV, A., SAPHIRE, E. O. & DE LA TORRE, J. C. 2022. Proximity interactome analysis of Lassa polymerase reveals eRF3a/GSPT1 as a druggable target for host-directed antivirals. *Proc Natl Acad Sci U S A,* 119**,** e2201208119.
FEDELI, C., TORRIANI, G., GALAN-NAVARRO, C., MORAZ, M.-L., MORENO, H., GEROLD, G., KUNZ, S. & DERMODY, T. S. 2018. Axl Can Serve as Entry Factor for Lassa Virus Depending on the Functional Glycosylation of Dystroglycan. *Journal of Virology,* 92**,** e01613-17.
FEHLING, S. K., LENNARTZ, F. & STRECKER, T. 2012. Multifunctional nature of the arenavirus RING finger protein Z. *Viruses,* 4**,** 2973-3011.
FEIGLIN, A., ASHKENAZI, S., SCHLESSINGER, A., ROST, B. & OFRAN, Y. 2014. Co-expression and co-localization of hub proteins and their partners are encoded in protein sequence. *Molecular BioSystems,* 10**,** 787-794.
FENNER, F. 1976. Classification and nomenclature of viruses. Second report of the International Committee on Taxonomy of Viruses. *Intervirology,* 7**,** 1-115.
FERNANDES, J., GUTERRES, A., DE OLIVEIRA, R. C., CHAMBERLAIN, J., LEWANDOWSKI, K., TEIXEIRA, B. R., COELHO, T. A., CRISÓSTOMO, C. F., BONVICINO, C. R., D'ANDREA, P. S., HEWSON, R. & DE LEMOS, E. R. S. 2018. Xapuri virus, a novel mammarenavirus: natural reassortment and increased diversity between New World viruses. *Emerg Microbes Infect,* 7**,** 120.
FICARELLI, M., ANTZIN-ANDUETZA, I., HUGH-WHITE, R., FIRTH, A. E., SERTKAYA, H., WILSON, H., NEIL, S. J. D., SCHULZ, R. & SWANSON, C. M. 2020. CpG

Dinucleotides Inhibit HIV-1 Replication through Zinc Finger Antiviral Protein (ZAP)-Dependent and -Independent Mechanisms. *J Virol,* 94.
FICARELLI, M., WILSON, H., PEDRO GALÃO, R., MAZZON, M., ANTZIN-ANDUETZA, I., MARSH, M., NEIL, S. J. & SWANSON, C. M. 2019. KHNYN is essential for the zinc finger antiviral protein (ZAP) to restrict HIV-1 containing clustered CpG dinucleotides. *Elife,* 8.
FICENEC, S. C., PERCAK, J., ARGUELLO, S., BAYS, A., GOBA, A., GBAKIE, M., SHAFFER, J. G., EMMETT, S. D., SCHIEFFELIN, J. S. & BAUSCH, D. 2020. Lassa Fever Induced Hearing Loss: The Neglected Disability of Hemorrhagic Fever. *International Journal of Infectious Diseases,* 100**,** 82-87.
FICENEC, S. C., SCHIEFFELIN, J. S. & EMMETT, S. D. 2019. A Review of Hearing Loss Associated with Zika, Ebola, and Lassa Fever. *Am J Trop Med Hyg,* 101**,** 484-490.
FINDLAY, G. M., ALCOCK, N. S. & STERN, R. 1936. THE VIRUS AETIOLOGY OF ONE FORM OF LYMPHOCYTIC MENINGITIS. *The Lancet,* 227**,** 650-654.
FISHER-HOCH, S. P., TOMORI, O., NASIDI, A., PEREZ-ORONOZ, G. I., FAKILE, Y., HUTWAGNER, L. & MCCORMICK, J. B. 1995. Review of cases of nosocomial Lassa fever in Nigeria: the high price of poor medical practice. *Bmj,* 311**,** 857-9.
FLANNERY, B., CAIN, M. & LY, H. 2023. Recent Discoveries of Novel Mammarenaviruses Infecting Humans and Other Mammals in Asia and Southeast Asia. *Virulence,* 14**,** 2231392.
FRAME, J. D., BALDWIN, J. M., JR., GOCKE, D. J. & TROUP, J. M. 1970. Lassa fever, a new virus disease of man from West Africa. I. Clinical description and pathological findings. *Am J Trop Med Hyg,* 19**,** 670-6.
FRANCIS, S. J. & SOUTHERN, P. J. 1988. Deleted Viral RNAs and Lymphocytic Choriomeningitis Virus Persistence in vitro. *Journal of General Virology,* 69**,** 1893-1902.
FROESCHKE, M., BASLER, M., GROETTRUP, M. & DOBBERSTEIN, B. 2003. Long-lived signal peptide of lymphocytic choriomeningitis virus glycoprotein pGP-C. *J Biol Chem,* 278**,** 41914-20.
FROS, J. J., DIETRICH, I., ALSHAIKHAHMED, K., PASSCHIER, T. C., EVANS, D. J. & SIMMONDS, P. 2017. CpG and UpA dinucleotides in both coding and non-coding regions of echovirus 7 inhibit replication initiation post-entry. *Elife,* 6.
FROS, J. J., VISSER, I., TANG, B., YAN, K., NAKAYAMA, E., VISSER, T. M., KOENRAADT, C. J. M., VAN OERS, M. M., PIJLMAN, G. P., SUHRBIER, A. & SIMMONDS, P. 2021. The dinucleotide composition of the Zika virus genome is shaped by conflicting evolutionary pressures in mammalian hosts and mosquito vectors. *PLoS Biol,* 19**,** e3001201.
FULLER-PACE, F. V. 2006. DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Research,* 34**,** 4206-4215.
FURUTA, Y., KOMENO, T. & NAKAMURA, T. 2017. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. *Proceedings of the Japan Academy, Series B,* 93**,** 449-463.
GACK, M. U., ALBRECHT, R. A., URANO, T., INN, K.-S., HUANG, I. C., CARNERO, E., FARZAN, M., INOUE, S., JUNG, J. U. & GARCÍA-SASTRE, A. 2009. Influenza A Virus NS1 Targets the Ubiquitin Ligase TRIM25 to Evade Recognition by the Host Viral RNA Sensor RIG-I. *Cell Host & Microbe,* 5**,** 439-449.

GACK, M. U., SHIN, Y. C., JOO, C.-H., URANO, T., LIANG, C., SUN, L., TAKEUCHI, O., AKIRA, S., CHEN, Z., INOUE, S. & JUNG, J. U. 2007. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature,* 446**,** 916-920.

GAIDAMOVICH, S. Y., BUTENKO, A. M. & LESCHINSKAYA, H. V. 2016. Human Laboratory Acquired ARBO-, ARENA-, and Hantavirus Infections. *Journal of the American Biological Safety Association,* 5**,** 5-11.

GALÃO, R. P., WILSON, H., SCHIERHORN, K. L., DEBELJAK, F., BODMER, B. S., GOLDHILL, D., HOENEN, T., WILSON, S. J., SWANSON, C. M. & NEIL, S. J. D. 2022. TRIM25 and ZAP target the Ebola virus ribonucleoprotein complex to mediate interferon-induced restriction. *PLOS Pathogens,* 18**,** e1010530.

GAO, G., GUO, X. & GOFF, S. P. 2002. Inhibition of Retroviral RNA Production by ZAP, a CCCH-Type Zinc Finger Protein. *Science,* 297**,** 1703-1706.

GARRUS, J. E., VON SCHWEDLER, U. K., PORNILLOS, O. W., MORHAM, S. G., ZAVITZ, K. H., WANG, H. E., WETTSTEIN, D. A., STRAY, K. M., CÔTÉ, M., RICH, R. L., MYSZKA, D. G. & SUNDQUIST, W. I. 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell,* 107**,** 55-65.

GARRY, C. E. & GARRY, R. F. 2019. Proteomics Computational Analyses Suggest that the Antennavirus Glycoprotein Complex Includes a Class I Viral Fusion Protein (α-Penetrene) with an Internal Zinc-Binding Domain and a Stable Signal Peptide. *Viruses,* 11**,** 750.

GARRY, R. F. 2023. Lassa fever — the road ahead. *Nature Reviews Microbiology,* 21**,** 87-96.

GAUNT, E., WISE, H. M., ZHANG, H., LEE, L. N., ATKINSON, N. J., NICOL, M. Q., HIGHTON, A. J., KLENERMAN, P., BEARD, P. M., DUTIA, B. M., DIGARD, P. & SIMMONDS, P. 2016. Elevation of CpG frequencies in influenza A genome attenuates pathogenicity but enhances host response to infection. *Elife,* 5**,** e12735.

GIARD, D. J., AARONSON, S. A., TODARO, G. J., ARNSTEIN, P., KERSEY, J. H., DOSIK, H. & PARKS, W. P. 1973. In Vitro Cultivation of Human Tumors: Establishment of Cell Lines Derived From a Series of Solid Tumors2. *JNCI: Journal of the National Cancer Institute,* 51**,** 1417-1423.

GINGRAS, A.-C., ABE, K. T. & RAUGHT, B. 2019. Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles. *Current opinion in chemical biology,* 48**,** 44-54.

GOH, P. Y., TAN, Y. J., LIM, S. P., TAN, Y. H., LIM, S. G., FULLER-PACE, F. & HONG, W. 2004. Cellular RNA helicase p68 relocalization and interaction with the hepatitis C virus (HCV) NS5B protein and the potential role of p68 in HCV RNA replication. *J Virol,* 78**,** 5288-98.

GONÇALVES-CARNEIRO, D., TAKATA, M. A., ONG, H., SHILTON, A. & BIENIASZ, P. D. 2021. Origin and evolution of the zinc finger antiviral protein. *PLOS Pathogens,* 17**,** e1009545.

GONZÁLEZ, P. H., COSSIO, P. M., ARANA, R., MAIZTEGUI, J. I. & LAGUENS, R. P. 1980. Lymphatic tissue in Argentine hemorrhagic fever. Pathologic features. *Archives of pathology & laboratory medicine,* 104**,** 250-254.

GOODIER, J. L., PEREIRA, G. C., CHEUNG, L. E., ROSE, R. J. & KAZAZIAN, H. H., JR. 2015. The Broad-Spectrum Antiviral Protein ZAP Restricts Human Retrotransposition. *PLoS Genet,* 11**,** e1005252.

GORI SAVELLINI, G., ANICHINI, G., GANDOLFO, C. & CUSI, M. G. 2021. SARS-CoV-2 N Protein Targets TRIM25-Mediated RIG-I Activation to Suppress Innate Immunity. *Viruses,* 13.

GRAHAM, F. L., RUSSELL, W. C., SMILEY, J. & NAIRN, R. 1977. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *Journal of General Virology,* 36**,** 59-72.

GRANT, T. D., LUFT, J. R., CARTER, L. G., MATSUI, T., WEISS, T. M., MARTEL, A. & SNELL, E. H. 2015. The accurate assessment of small-angle X-ray scattering data. *Acta Crystallogr D Biol Crystallogr,* 71**,** 45-56.

GRÄWERT, M. & SVERGUN, D. 2020. A beginner's guide to solution small-angle X-ray scattering (SAXS). *The Biochemist,* 42**,** 36-42.

GREENBAUM, B. D., LEVINE, A. J., BHANOT, G. & RABADAN, R. 2008. Patterns of evolution and host gene mimicry in influenza and other RNA viruses. *PLoS Pathog,* 4**,** e1000079.

GUO, X., CARROLL, J.-W. N., MACDONALD, M. R., GOFF, S. P. & GAO, G. 2004. The Zinc Finger Antiviral Protein Directly Binds to Specific Viral mRNAs through the CCCH Zinc Finger Motifs. *Journal of Virology,* 78**,** 12781-12787.

GUO, X., MA, J., SUN, J. & GAO, G. 2007. The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proc Natl Acad Sci U S A,* 104**,** 151-6.

HAMMONDS, J., WANG, J. J., YI, H. & SPEARMAN, P. 2010. Immunoelectron microscopic evidence for Tetherin/BST2 as the physical bridge between HIV-1 virions and the plasma membrane. *PLoS Pathog,* 6**,** e1000749.

HARNISH, D. G., LEUNG, W. C. & RAWLS, W. E. 1981. Characterization of polypeptides immunoprecipitable from Pichinde virus-infected BHK-21 cells. *Journal of Virology,* 38**,** 840-848.

HASS, M., GÖLNITZ, U., MÜLLER, S., BECKER-ZIAJA, B. & GÜNTHER, S. 2004. Replicon System for Lassa Virus. *Journal of Virology,* 78**,** 13793-13803.

HASS, M., LELKE, M., BUSCH, C., BECKER-ZIAJA, B. & GÜNTHER, S. 2008. Mutational Evidence for a Structural Model of the Lassa Virus RNA Polymerase Domain and Identification of Two Residues, Gly1394 and Asp1395, That Are Critical for Transcription but Not Replication of the Genome. *Journal of Virology,* 82**,** 10207-10217.

HASTIE, K. M., IGONET, S., SULLIVAN, B. M., LEGRAND, P., ZANDONATTI, M. A., ROBINSON, J. E., GARRY, R. F., REY, F. A., OLDSTONE, M. B. & SAPHIRE, E. O. 2016a. Crystal structure of the prefusion surface glycoprotein of the prototypic arenavirus LCMV. *Nat Struct Mol Biol,* 23**,** 513-521.

HASTIE, K. M., KIMBERLIN, C. R., ZANDONATTI, M. A., MACRAE, I. J. & SAPHIRE, E. O. 2011a. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. *Proceedings of the National Academy of Sciences,* 108**,** 2396-2401.

HASTIE, K. M., KIMBERLIN, C. R., ZANDONATTI, M. A., MACRAE, I. J. & SAPHIRE, E. O. 2011b. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. *Proc Natl Acad Sci U S A,* 108**,** 2396-401.

HASTIE, K. M., KING, L. B., ZANDONATTI, M. A. & SAPHIRE, E. O. 2012. Structural Basis for the dsRNA Specificity of the Lassa Virus NP Exonuclease. *PLOS ONE,* 7**,** e44211.

HASTIE, K. M., LIU, T., LI, S., KING, L. B., NGO, N., ZANDONATTI, M. A., WOODS, V. L., JR., DE LA TORRE, J. C. & SAPHIRE, E. O. 2011c. Crystal structure of

the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding. *Proc Natl Acad Sci U S A,* 108**,** 19365-70.
HASTIE, K. M. & SAPHIRE, E. O. 2018. Lassa virus glycoprotein: stopping a moving target. *Curr Opin Virol,* 31**,** 52-58.
HASTIE, K. M., ZANDONATTI, M., LIU, T., LI, S., WOODS, V. L., JR. & SAPHIRE, E. O. 2016b. Crystal Structure of the Oligomeric Form of Lassa Virus Matrix Protein Z. *J Virol,* 90**,** 4556-62.
HASTIE, K. M., ZANDONATTI, M. A., KLEINFELTER, L. M., HEINRICH, M. L., ROWLAND, M. M., CHANDRAN, K., BRANCO, L. M., ROBINSON, J. E., GARRY, R. F. & SAPHIRE, E. O. 2017. Structural basis for antibody-mediated neutralization of Lassa virus. *Science,* 356**,** 923-928.
HAYAKAWA, S., SHIRATORI, S., YAMATO, H., KAMEYAMA, T., KITATSUJI, C., KASHIGI, F., GOTO, S., KAMEOKA, S., FUJIKURA, D., YAMADA, T., MIZUTANI, T., KAZUMATA, M., SATO, M., TANAKA, J., ASAKA, M., OHBA, Y., MIYAZAKI, T., IMAMURA, M. & TAKAOKA, A. 2011. ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses. *Nat Immunol,* 12**,** 37-44.
HAYDON, D. T., CLEAVELAND, S., TAYLOR, L. H. & LAURENSON, M. K. 2002. Identifying reservoirs of infection: a conceptual and practical challenge. *Emerg Infect Dis,* 8**,** 1468-73.
HAYES, M. W., CARRION, R., NUNNELEY, J., MEDVEDEV, A. E., SALVATO, M. S. & LUKASHEVICH, I. S. 2012. Pathogenic Old World Arenaviruses Inhibit TLR2/Mal-Dependent Proinflammatory Cytokines In Vitro. *Journal of Virology,* 86**,** 7216-7226.
HEBERLE, H., MEIRELLES, G. V., DA SILVA, F. R., TELLES, G. P. & MINGHIM, R. 2015. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics,* 16.
HELGUERA, G., JEMIELITY, S., ABRAHAM, J., CORDO, S. M., MARTINEZ, M. G., RODRÍGUEZ, J. A., BREGNI, C., WANG, J. J., FARZAN, M., PENICHET, M. L., CANDURRA, N. A. & CHOE, H. 2012. An Antibody Recognizing the Apical Domain of Human Transferrin Receptor 1 Efficiently Inhibits the Entry of All New World Hemorrhagic Fever Arenaviruses. *Journal of Virology,* 86**,** 4024-4028.
HEPOJOKI, J., HEPOJOKI, S., SMURA, T., SZIROVICZA, L., DERVAS, E., PRÄHAUSER, B., NUFER, L., SCHRANER, E. M., VAPALAHTI, O., KIPAR, A. & HETZEL, U. 2018. Characterization of Haartman Institute snake virus-1 (HISV-1) and HISV-like viruses—The representatives of genus Hartmanivirus, family Arenaviridae. *PLOS Pathogens,* 14**,** e1007415.
HETZEL, U., SIRONEN, T., LAURINMÄKI, P., LILJEROOS, L., PATJAS, A., HENTTONEN, H., VAHERI, A., ARTELT, A., KIPAR, A., BUTCHER, S. J., VAPALAHTI, O. & HEPOJOKI, J. 2013. Isolation, identification, and characterization of novel arenaviruses, the etiological agents of boid inclusion body disease. *J Virol,* 87**,** 10918-35.
HINZ, A., MIGUET, N., NATRAJAN, G., USAMI, Y., YAMANAKA, H., RENESTO, P., HARTLIEB, B., MCCARTHY, A. A., SIMORRE, J. P., GÖTTLINGER, H. & WEISSENHORN, W. 2010. Structural basis of HIV-1 tethering to membranes by the BST-2/tetherin ectodomain. *Cell Host Microbe,* 7**,** 314-323.
HIRSCH, M. S., MOELLERING, R. C., JR., POPE, H. G. & POSKANZER, D. C. 1974. Lymphocytic-choriomeningitis virus infection traced to a pet hamster. *N Engl J Med,* 291**,** 610-2.

HOLLIDAY, R. & GRIGG, G. W. 1993. DNA methylation and mutation. *Mutat Res,* 285**,** 61-7.
HONDA, K., YANAI, H., NEGISHI, H., ASAGIRI, M., SATO, M., MIZUTANI, T., SHIMADA, N., OHBA, Y., TAKAOKA, A., YOSHIDA, N. & TANIGUCHI, T. 2005. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature,* 434**,** 772-777.
HORNUNG, V., ELLEGAST, J., KIM, S., BRZÓZKA, K., JUNG, A., KATO, H., POECK, H., AKIRA, S., CONZELMANN, K.-K., SCHLEE, M., ENDRES, S. & HARTMANN, G. 2006a. 5'-Triphosphate RNA Is the Ligand for RIG-I. *Science,* 314**,** 994-997.
HORNUNG, V., ELLEGAST, J., KIM, S., BRZOZKA, K., JUNG, A., KATO, H., POECK, H., AKIRA, S., CONZELMANN, K. K., SCHLEE, M., ENDRES, S. & HARTMANN, G. 2006b. 5'-Triphosphate RNA Is the Ligand for RIG-I. *Science,* 314**,** 994-997.
HOWLEY, P. M. & KNIPE, D. M. 2020. *Fields Virology: Emerging Viruses*, Wolters Kluwer Health.
HU, Y., LI, W., GAO, T., CUI, Y., JIN, Y., LI, P., MA, Q., LIU, X. & CAO, C. 2017a. The Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Inhibits Type I Interferon Production by Interfering with TRIM25-Mediated RIG-I Ubiquitination. *Journal of Virology,* 91**,** JVI.02143-16.
HU, Y., LI, W., GAO, T., CUI, Y., JIN, Y., LI, P., MA, Q., LIU, X. & CAO, C. 2017b. The Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Inhibits Type I Interferon Production by Interfering with TRIM25-Mediated RIG-I Ubiquitination. *Journal of Virology,* 91**,** 10.1128/jvi.02143-16.
HUANG, C., KOLOKOLTSOVA, O. A., YUN, N. E., SEREGIN, A. V., RONCA, S., KOMA, T. & PAESSLER, S. 2015a. Highly Pathogenic New World and Old World Human Arenaviruses Induce Distinct Interferon Responses in Human Cells. *J Virol,* 89**,** 7079-88.
HUANG, Q., SHAO, J., LAN, S., ZHOU, Y., XING, J., DONG, C., LIANG, Y. & LY, H. 2015b. In vitro and in vivo characterizations of pichinde viral nucleoprotein exoribonuclease functions. *J Virol,* 89**,** 6595-607.
HUANG, Z., WANG, X. & GAO, G. 2010. Analyses of SELEX-derived ZAP-binding RNA aptamers suggest that the binding specificity is determined by both structure and sequence of the RNA. *Protein & Cell,* 1**,** 752-759.
HUGOT, J. P., GONZALEZ, J. P. & DENYS, C. 2001. Evolution of the Old World Arenaviridae and their rodent hosts: generalized host-transfer or association by descent? *Infect Genet Evol,* 1**,** 13-20.
HUNG, V., LAM, S. S., UDESHI, N. D., SVINKINA, T., GUZMAN, G., MOOTHA, V. K., CARR, S. A. & TING, A. Y. 2017. Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *elife,* 6**,** e24463.
IGONET, S., VANEY, M.-C., VONRHEIN, C., BRICOGNE, G., STURA, E. A., HENGARTNER, H., ESCHLI, B. & REY, F. A. 2011. X-ray structure of the arenavirus glycoprotein GP2 in its postfusion hairpin conformation. *Proceedings of the National Academy of Sciences,* 108**,** 19967-19972.
ILORI, E. A., FURUSE, Y., IPADEOLA, O. B., DAN-NWAFOR, C. C., ABUBAKAR, A., WOMI-ETENG, O. E., OGBAINI-EMOVON, E., OKOGBENIN, S., UNIGWE, U., OGAH, E., AYODEJI, O., ABEJEGAH, C., LIASU, A. A., MUSA, E. O., WOLDETSADIK, S. F., LASUBA, C. L. P., ALEMU, W., IHEKWEAZU, C. & NIGERIA LASSA FEVER NATIONAL RESPONSE, T. 2019. Epidemiologic and Clinical Features of Lassa Fever Outbreak in Nigeria, January 1-May 6, 2018. *Emerg Infect Dis,* 25**,** 1066-1074.

IMPERIALI, M., SPÖRRI, R., HEWITT, J. & OXENIUS, A. 2008. Post-translational modification of α-dystroglycan is not critical for lymphocytic choriomeningitis virus receptor function in vivo. *Journal of General Virology,* 89**,** 2713-2722.

INAMORI, K., YOSHIDA-MORIGUCHI, T., HARA, Y., ANDERSON, M. E., YU, L. & CAMPBELL, K. P. 2012. Dystroglycan function requires xylosyl- and glucuronyltransferase activities of LARGE. *Science,* 335**,** 93-6.

INN, K.-S., GACK, M. U., TOKUNAGA, F., SHI, M., WONG, L.-Y., IWAI, K. & JUNG, J. U. 2011. Linear Ubiquitin Assembly Complex Negatively Regulates RIG-I- and TRIM25-Mediated Type I Interferon Induction. *Molecular Cell,* 41**,** 354-365.

INOUE, S., ORIMO, A., HOSOI, T., KONDO, S., TOYOSHIMA, H., KONDO, T., IKEGAMI, A., OUCHI, Y., ORIMO, H. & MURAMATSU, M. 1993. Genomic binding-site cloning reveals an estrogen-responsive gene that encodes a RING finger protein. *Proceedings of the National Academy of Sciences,* 90**,** 11117-11121.

ISAAC, A. B., KAROLINA, W., TEMITOPE, A. A., ANUSKA, R., JOANNE, E., DEBORAH, A., BIANCA, O. C., FILIP, T., ZOFIA, P., OLUWASEGUN, O. I., OLUWAFERANMI, O. & GRACE, B. T. 2022. PROSPECTS OF LASSA FEVER CANDIDATE VACCINES. *Afr J Infect Dis,* 16**,** 46-58.

IWASAKI, M., MINDER, P., CAI, Y., KUHN, J. H., YATES, J. R., 3RD, TORBETT, B. E. & DE LA TORRE, J. C. 2018. Interactome analysis of the lymphocytic choriomeningitis virus nucleoprotein in infected cells reveals ATPase Na+/K+ transporting subunit Alpha 1 and prohibitin as host-cell factors involved in the life cycle of mammarenaviruses. *PLoS Pathog,* 14**,** e1006892.

IWASAKI, M., NGO, N., CUBITT, B. & DE LA TORRE, J. C. 2015. Efficient Interaction between Arenavirus Nucleoprotein (NP) and RNA-Dependent RNA Polymerase (L) Is Mediated by the Virus Nucleocapsid (NP-RNA) Template. *J Virol,* 89**,** 5734-8.

JÁCAMO, R., LÓPEZ, N., WILDA, M. & FRANZE-FERNÁNDEZ, M. T. 2003. Tacaribe virus Z protein interacts with the L polymerase protein to inhibit viral RNA synthesis. *J Virol,* 77**,** 10383-93.

JAE, L. T., RAABEN, M., HERBERT, A. S., KUEHNE, A. I., WIRCHNIANSKI, A. S., SOH, T. K., STUBBS, S. H., JANSSEN, H., DAMME, M., SAFTIG, P., WHELAN, S. P., DYE, J. M. & BRUMMELKAMP, T. R. 2014. Virus entry. Lassa virus entry requires a trigger-induced receptor switch. *Science,* 344**,** 1506-10.

JAVANBAKHT, H., YUAN, W., YEUNG, D. F., SONG, B., DIAZ-GRIFFERO, F., LI, Y., LI, X., STREMLAU, M. & SODROSKI, J. 2006. Characterization of TRIM5alpha trimerization and its contribution to human immunodeficiency virus capsid binding. *Virology,* 353**,** 234-46.

JEMIELITY, S., WANG, J. J., CHAN, Y. K., AHMED, A. A., LI, W., MONAHAN, S., BU, X., FARZAN, M., FREEMAN, G. J., UMETSU, D. T., DEKRUYFF, R. H. & CHOE, H. 2013. TIM-family Proteins Promote Infection of Multiple Enveloped Viruses through Virion-associated Phosphatidylserine. *PLoS Pathogens,* 9**,** e1003232.

JI, W. T. & LIU, H. J. 2008. PI3K-Akt signaling and viral infection. *Recent Pat Biotechnol,* 2**,** 218-26.

JIANG, X., HUANG, Q., WANG, W., DONG, H., LY, H., LIANG, Y. & DONG, C. 2013. Structures of arenaviral nucleoproteins with triphosphate dsRNA reveal a

unique mechanism of immune suppression. *J Biol Chem,* 288**,** 16949-16959.
JOHNSON, K. M. 1965. Epidemiology of Machupo virus infection. 3. Significance of virological observations in man and animals. *Am J Trop Med Hyg,* 14**,** 816-8.
JOHNSON, K. M., KUNS, M. L., MACKENZIE, R. B., WEBB, P. A. & YUNKER, C. E. 1966. Isolation of Machupo virus from wild rodent Calomys callosus. *Am J Trop Med Hyg,* 15**,** 103-6.
JOHNSON, K. M., WIEBENGA, N. H., MACKENZIE, R. B., KUNS, M. L., TAURASO, N. M., SHELOKOV, A., WEBB, P. A., JUSTINES, G. & BEYE, H. K. 1965. VIRUS ISOLATIONS FROM HUMAN CASES OF HEMORRHAGIC FEVER IN BOLIVIA. *Proc Soc Exp Biol Med,* 118**,** 113-8.
KARLIN, S., DOERFLER, W. & CARDON, L. R. 1994. Why is CpG suppressed in the genomes of virtually all small eukaryotic viruses but not in those of large eukaryotic viruses? *J Virol,* 68**,** 2889-97.
KATOH, K. & STANDLEY, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol,* 30**,** 772-80.
KATZ, M. & DISKIN, R. 2024. The underlying mechanisms of arenaviral entry through matriglycan. *Frontiers in Molecular Biosciences,* 11.
KATZ, M., WEINSTEIN, J., EILON-ASHKENAZY, M., GEHRING, K., COHEN-DVASHI, H., ELAD, N., FLEISHMAN, S. J. & DISKIN, R. 2022. Structure and receptor recognition by the Lassa virus spike complex. *Nature,* 603**,** 174-179.
KEENLYSIDE, R. A., MCCORMICK, J. B., WEBB, P. A., SMITH, E., ELLIOTT, L. & JOHNSON, K. M. 1983. Case-control study of Mastomys natalensis and humans in Lassa virus-infected households in Sierra Leone. *Am J Trop Med Hyg,* 32**,** 829-37.
KENTSIS, A., GORDON, R. E. & BORDEN, K. L. B. 2002a. Control of biochemical reactions through supramolecular RING domain self-assembly. *Proceedings of the National Academy of Sciences,* 99**,** 15404-15409.
KENTSIS, A., GORDON, R. E. & BORDEN, K. L. B. 2002b. Self-assembly properties of a model RING domain. *Proceedings of the National Academy of Sciences,* 99**,** 667-672.
KERNÉIS, S., KOIVOGUI, L., MAGASSOUBA, N. F., KOULEMOU, K., LEWIS, R., APLOGAN, A., GRAIS, R. F., GUERIN, P. J. & FICHET-CALVET, E. 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**,** e548.
KERNS, J. A., EMERMAN, M. & MALIK, H. S. 2008. Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. *PLoS Genet,* 4**,** e21.
KERSCHER, O., FELBERBAUM, R. & HOCHSTRASSER, M. 2006. Modification of Proteins by Ubiquitin and Ubiquitin-Like Proteins. *Annual Review of Cell and Developmental Biology,* 22**,** 159-180.
KHAMINA, K., LERCHER, A., CALDERA, M., SCHLIEHE, C., VILAGOS, B., SAHIN, M., KOSACK, L., BHATTACHARYA, A., MÁJEK, P., STUKALOV, A., SACCO, R., JAMES, L. C., PINSCHEWER, D. D., BENNETT, K. L., MENCHE, J. & BERGTHALER, A. 2017. Characterization of host proteins interacting with the lymphocytic choriomeningitis virus L protein. *PLOS Pathogens,* 13**,** e1006758.

KIM, D. I., JENSEN, S. C., NOBLE, K. A., KC, B., ROUX, K. H., MOTAMEDCHABOKI, K. & ROUX, K. J. 2016. An improved smaller biotin ligase for BioID proximity labeling. *Molecular biology of the cell,* 27**,** 1188-1196.

KIM, D. I., KC, B., ZHU, W., MOTAMEDCHABOKI, K., DOYE, V. & ROUX, K. J. 2014. Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proceedings of the National Academy of Sciences,* 111**,** E2453-E2461.

KING, B. R., HERSHKOWITZ, D., EISENHAUER, P. L., WEIR, M. E., ZIEGLER, C. M., RUSSO, J., BRUCE, E. A., BALLIF, B. A. & BOTTEN, J. 2017a. A Map of the Arenavirus Nucleoprotein-Host Protein Interactome Reveals that Junin Virus Selectively Impairs the Antiviral Activity of Double-Stranded RNA-Activated Protein Kinase (PKR). *J Virol,* 91.

KING, B. R., HERSHKOWITZ, D., EISENHAUER, P. L., WEIR, M. E., ZIEGLER, C. M., RUSSO, J., BRUCE, E. A., BALLIF, B. A. & BOTTEN, J. 2017b. A Map of the Arenavirus Nucleoprotein-Host Protein Interactome Reveals that Junín Virus Selectively Impairs the Antiviral Activity of Double-Stranded RNA-Activated Protein Kinase (PKR). *Journal of Virology,* 91**,** JVI.00763-17.

KING, B. R., SAMACOITS, A., EISENHAUER, P. L., ZIEGLER, C. M., BRUCE, E. A., ZENKLUSEN, D., ZIMMER, C., MUELLER, F. & BOTTEN, J. 2018. Visualization of Arenavirus RNA Species in Individual Cells by Single-Molecule Fluorescence In Situ Hybridization Suggests a Model of Cyclical Infection and Clearance during Persistence. *J Virol,* 92.

KLEINE, H., POREBA, E., LESNIEWICZ, K., HASSA, P. O., HOTTIGER, M. O., LITCHFIELD, D. W., SHILTON, B. H. & LÜSCHER, B. 2008. Substrate-assisted catalysis by PARP10 limits its activity to mono-ADP-ribosylation. *Mol Cell,* 32**,** 57-69.

KMIEC, D., LISTA-BROTOS, M.-J., FICARELLI, M., SWANSON, C. M. & NEIL, S. J. 2021a. The C-terminal PARP domain of the long ZAP isoform contributes essential effector functions for CpG-directed antiviral activity. Cold Spring Harbor Laboratory.

KMIEC, D., LISTA, M. J., FICARELLI, M., SWANSON, C. M. & NEIL, S. J. D. 2021b. S-farnesylation is essential for antiviral activity of the long ZAP isoform against RNA viruses with diverse replication strategies. *PLoS Pathog,* 17**,** e1009726.

KMIEC, D., NCHIOUA, R., SHERRILL-MIX, S., STÜRZEL, C. M., HEUSINGER, E., BRAUN, E., GONDIM, M. V. P., HOTTER, D., SPARRER, K. M. J., HAHN, B. H., SAUTER, D. & KIRCHHOFF, F. 2020. CpG Frequency in the 5' Third of the env Gene Determines Sensitivity of Primary HIV-1 Strains to the Zinc-Finger Antiviral Protein. *mBio,* 11.

KNOBLOCH, J., MCCORMICK, J., WEBB, P., DIETRICH, M., SCHUMACHER, H. & DENNIS, E. 1980. Clinical observations in 42 patients with Lassa fever. *Tropenmedizin und Parasitologie,* 31**,** 389-398.

KNOPP, K. A., NGO, T., GERSHON, P. D. & BUCHMEIER, M. J. 2015. Single nucleoprotein residue modulates arenavirus replication complex formation. *mBio,* 6**,** e00524-15.

KOLIOPOULOS, M. G., LETHIER, M., VAN DER VEEN, A. G., HAUBRICH, K., HENNIG, J., KOWALINSKI, E., STEVENS, R. V., MARTIN, S. R., REIS E SOUSA, C., CUSACK, S. & RITTINGER, K. 2018. Molecular mechanism of influenza A NS1-mediated TRIM25 recognition and inhibition. *Nature Communications,* 9**,** 1820.

KOMANDER, D. & RAPE, M. 2012. The Ubiquitin Code. *Annual Review of Biochemistry,* 81**,** 203-229.

KRANZUSCH, P. J. & WHELAN, S. P. 2011a. Arenavirus Z protein controls viral RNA synthesis by locking a polymerase-promoter complex. *Proc Natl Acad Sci U S A,* 108**,** 19743-8.
KRANZUSCH, P. J. & WHELAN, S. P. J. 2011b. Arenavirus Z protein controls viral RNA synthesis by locking a polymerase-promoter complex. *Proceedings of the National Academy of Sciences,* 108**,** 19743-19748.
KRZYWINSKI, M. & ALTMAN, N. 2014a. Comparing samples--part I: robustly comparing pairs of independent or related samples requires different approaches to the t-test. *Nature methods,* 11**,** 215-217.
KRZYWINSKI, M. & ALTMAN, N. 2014b. Points of significance: comparing samples--part II. *nature methods,* 11**,** 355.
KUHN, J. H., BROWN, K., ADKINS, S., TORRE, J. C. D. L., DIGIARO, M., ERGÜNAY, K., FIRTH, A. E., HUGHES, H. R., JUNGLEN, S., LAMBERT, A. J., MAES, P., MARKLEWITZ, M., PALACIOS, G., SASAYA, T., SHI, M., ZHANG, Y.-Z., WOLF, Y. I. & TURINA, M. 2024. Promotion of order Bunyavirales to class Bunyaviricetes to accommodate a rapidly increasing number of related polyploviricotine viruses. *Journal of Virology,* 98**,** e01069-24.
KUMAR, R., KHANDELWAL, N., THACHAMVALLY, R., TRIPATHI, B. N., BARUA, S., KASHYAP, S. K., MAHERCHANDANI, S. & KUMAR, N. 2018. Role of MAPK/MNK1 signaling in virus replication. *Virus Res,* 253**,** 48-61.
KUNZ, S., ROJEK, J. M., KANAGAWA, M., SPIROPOULOU, C. F., BARRESI, R., CAMPBELL, K. P. & OLDSTONE, M. B. A. 2005. Posttranslational Modification of α-Dystroglycan, the Cellular Receptor for Arenaviruses, by the Glycosyltransferase LARGE Is Critical for Virus Binding. *Journal of Virology,* 79**,** 14282-14296.
KWON, S. C., YI, H., EICHELBAUM, K., FÖHR, S., FISCHER, B., YOU, K. T., CASTELLO, A., KRIJGSVELD, J., HENTZE, M. W. & KIM, V. N. 2013. The RNA-binding protein repertoire of embryonic stem cells. *Nature Structural & Molecular Biology,* 20**,** 1122-1130.
LACKIE, J. 2010. *A Dictionary of Biomedicine*, Oxford University Press.
LAPOŠOVÁ, K., PASTOREKOVÁ, S. & TOMÁŠKOVÁ, J. 2013. Lymphocytic choriomeningitis virus: invisible but not innocent. *Acta Virol,* 57**,** 160-70.
LARSON, R. A., DAI, D., HOSACK, V. T., TAN, Y., BOLKEN, T. C., HRUBY, D. E. & AMBERG, S. M. 2008. Identification of a Broad-Spectrum Arenavirus Entry Inhibitor. *Journal of Virology,* 82**,** 10768-10775.
LAW, L. M. J., RAZOOKY, B. S., LI, M. M. H., YOU, S., JURADO, A., RICE, C. M. & MACDONALD, M. R. 2019. ZAP's stress granule localization is correlated with its antiviral activity and induced by virus replication. *PLoS Pathog,* 15**,** e1007798.
LE SAGE, V., CINTI, A., AMORIM, R. & MOULAND, A. J. 2016. Adapting the Stress Response: Viral Subversion of the mTOR Signaling Pathway. *Viruses,* 8.
LE TORTOREC, A., WILLEY, S. & NEIL, S. J. 2011. Antiviral inhibition of enveloped virus release by tetherin/BST-2: action and counteraction. *Viruses,* 3**,** 520-40.
LEE, H., KOMANO, J., SAITOH, Y., YAMAOKA, S., KOZAKI, T., MISAWA, T., TAKAHAMA, M., SATOH, T., TAKEUCHI, O., YAMAMOTO, N., MATSUURA, Y., SAITOH, T. & AKIRA, S. 2013. Zinc-finger antiviral protein mediates retinoic acid inducible gene I-like receptor-independent antiviral response to murine leukemia virus. *Proc Natl Acad Sci U S A,* 110**,** 12379-84.

LEE, N.-R., KIM, H.-I., CHOI, M.-S., YI, C.-M. & INN, K.-S. 2015. Regulation of MDA5-MAVS Antiviral Signaling Axis by TRIM25 through TRAF6-Mediated NF-κB Activation. *Molecules and Cells,* 38**,** 759-764.
LEHMANN, M., PAHLMANN, M., JÉRÔME, H., BUSCH, C., LELKE, M. & GÜNTHER, S. 2014. Role of the C Terminus of Lassa Virus L Protein in Viral mRNA Synthesis. *Journal of Virology,* 88**,** 8713-8717.
LEIFER, E., GOCKE, D. J. & BOURNE, H. 1970. Lassa fever, a new virus disease of man from West Africa. II. Report of a laboratory-acquired infection treated with plasma from a person recently recovered from the disease. *Am J Trop Med Hyg,* 19**,** 677-9.
LELKE, M., BRUNOTTE, L., BUSCH, C. & GÜNTHER, S. 2010. An N-terminal region of Lassa virus L protein plays a critical role in transcription but not replication of the virus genome. *J Virol,* 84**,** 1934-44.
LEMKE, G. & ROTHLIN, C. V. 2008. Immunobiology of the TAM receptors. *Nat Rev Immunol,* 8**,** 327-36.
LENNARTZ, F., HOENEN, T., LEHMANN, M., GROSETH, A. & GARTEN, W. 2013. The role of oligomerization for the biological functions of the arenavirus nucleoprotein. *Archives of Virology,* 158**,** 1895-1905.
LEVINGSTON MACLEOD, J. M., D'ANTUONO, A., LOUREIRO, M. E., CASABONA, J. C., GOMEZ, G. A. & LOPEZ, N. 2011. Identification of two functional domains within the arenavirus nucleoprotein. *J Virol,* 85**,** 2012-23.
LI, M., YAN, K., WEI, L., YANG, J., LU, C., XIONG, F., ZHENG, C. & XU, W. 2015. Zinc finger antiviral protein inhibits coxsackievirus B3 virus replication and protects against viral myocarditis. *Antiviral Res,* 123**,** 50-61.
LI, M. M., LAU, Z., CHEUNG, P., AGUILAR, E. G., SCHNEIDER, W. M., BOZZACCO, L., MOLINA, H., BUEHLER, E., TAKAOKA, A., RICE, C. M., FELSENFELD, D. P. & MACDONALD, M. R. 2017. TRIM25 Enhances the Antiviral Action of Zinc-Finger Antiviral Protein (ZAP). *PLoS Pathog,* 13**,** e1006145.
LI, M. M. H., AGUILAR, E. G., MICHAILIDIS, E., PABON, J., PARK, P., WU, X., DE JONG, Y. P., SCHNEIDER, W. M., MOLINA, H., RICE, C. M. & MACDONALD, M. R. 2019. Characterization of Novel Splice Variants of Zinc Finger Antiviral Protein (ZAP). *Journal of Virology,* 93.
LI, S., SUN, Z., PRYCE, R., PARSY, M.-L., FEHLING, S. K., SCHLIE, K., SIEBERT, C. A., GARTEN, W., BOWDEN, T. A., STRECKER, T. & HUISKONEN, J. T. 2016. Acidic pH-Induced Conformations and LAMP1 Binding of the Lassa Virus Glycoprotein Spike. *PLOS Pathogens,* 12**,** e1005418.
LINERO, F., WELNOWSKA, E., CARRASCO, L. & SCOLARO, L. 2013. Participation of eIF4F complex in Junin virus infection: blockage of eIF4E does not impair virus replication. *Cell Microbiol,* 15**,** 1766-82.
LIU, C.-H., CHIEN, M.-J., CHANG, Y.-C., CHENG, Y.-H., LI, F.-A. & MOU, K. Y. 2020. Combining proximity labeling and cross-linking mass spectrometry for proteomic dissection of nuclear envelope interactome. *Journal of proteome research,* 19**,** 1109-1118.
LIU, F. & HECK, A. J. 2015. Interrogating the architecture of protein assemblies and protein interaction networks by cross-linking mass spectrometry. *Current opinion in structural biology,* 35**,** 100-108.
LIU, L., TIAN, J., NAN, H., TIAN, M., LI, Y., XU, X., HUANG, B., ZHOU, E., HISCOX, J. A. & CHEN, H. 2016. Porcine Reproductive and Respiratory Syndrome Virus Nucleocapsid Protein Interacts with Nsp9 and Cellular DHX9 To Regulate Viral RNA Synthesis. *Journal of Virology,* 90**,** 5384-5398.
LIU, L., WANG, P., LIU, A., ZHANG, L., YAN, L., GUO, Y., XIAO, G., RAO, Z. & LOU, Z. 2023. Structure basis for allosteric regulation of lymphocytic

choriomeningitis virus polymerase function by Z matrix protein. *Protein & Cell,* 14**,** 703-707.
LÓPEZ, N., JÁCAMO, R. & FRANZE-FERNÁNDEZ, M. A. T. 2001. Transcription and RNA Replication of Tacaribe Virus Genome and Antigenome Analogs Require N and L Proteins: Z Protein Is an Inhibitor of These Processes. *Journal of Virology,* 75**,** 12241-12251.
LOUREIRO, M. E., WILDA, M., LEVINGSTON MACLEOD, J. M., D'ANTUONO, A., FOSCALDI, S., MARINO BUSLJE, C. & LOPEZ, N. 2011. Molecular determinants of arenavirus Z protein homo-oligomerization and L polymerase binding. *J Virol,* 85**,** 12304-14.
LOUREIRO, M. E., ZORZETTO-FERNANDES, A. L., RADOSHITZKY, S., CHI, X., DALLARI, S., MAROOKI, N., LEGER, P., FOSCALDI, S., HARJONO, V., SHARMA, S., ZID, B. M., LOPEZ, N., DE LA TORRE, J. C., BAVARI, S. & ZUNIGA, E. 2018. DDX3 suppresses type I interferons and favors viral replication during Arenavirus infection. *PLoS Pathog,* 14**,** e1007125.
LOW, T. Y., SYAFRUDDIN, S. E., MOHTAR, M. A., VELLAICHAMY, A., A RAHMAN, N. S., PUNG, Y.-F. & TAN, C. S. H. 2021. Recent progress in mass spectrometry-based strategies for elucidating protein–protein interactions. *Cellular and Molecular Life Sciences,* 78**,** 5325-5339.
LUM, K. K. & CRISTEA, I. M. 2016. Proteomic approaches to uncovering virus-host protein interactions during the progression of viral infection. *Expert Rev Proteomics,* 13**,** 325-40.
LUO, X., WANG, X., GAO, Y., ZHU, J., LIU, S., GAO, G. & GAO, P. 2020. Molecular Mechanism of RNA Recognition by Zinc-Finger Antiviral Protein. *Cell Rep,* 30**,** 46-52.e4.
LUO, X. L., LU, S., QIN, C., SHI, M., LU, X. B., WANG, L., GA, S., JIN, D., MA, X. L., YANG, J., DAI, Y., BAO, L. L., CHENG, Y. P., GE, Y. J., BAI, Y. B., ZHU, W. T., PU, J., SUN, H., HUANG, Y. Y., XU, M. C., LEI, W. J., DONG, K., YANG, C. X., JIAO, Y. F., LV, Q., LI, F. D. & XU, J. 2023. Emergence of an ancient and pathogenic mammarenavirus. *Emerg Microbes Infect,* 12**,** e2192816.
MA, D. Y. & SUTHAR, M. S. 2015. Mechanisms of innate immune evasion in re-emerging RNA viruses. *Current Opinion in Virology,* 12**,** 26-37.
MACDONALD, M. R., MACHLIN, E. S., ALBIN, O. R. & LEVY, D. E. 2007. The Zinc Finger Antiviral Protein Acts Synergistically with an Interferon-Induced Factor for Maximal Activity against Alphaviruses. *Journal of Virology,* 81**,** 13509-13518.
MAES, P., ADKINS, S., ALKHOVSKY, S. V., AVŠIČ-ŽUPANC, T., BALLINGER, M. J., BENTE, D. A., BEER, M., BERGERON, É., BLAIR, C. D., BRIESE, T., BUCHMEIER, M. J., BURT, F. J., CALISHER, C. H., CHARREL, R. N., CHOI, I. R., CLEGG, J. C. S., DE LA TORRE, J. C., DE LAMBALLERIE, X., DERISI, J. L., DIGIARO, M., DREBOT, M., EBIHARA, H., ELBEAINO, T., ERGÜNAY, K., FULHORST, C. F., GARRISON, A. R., GĀO, G. F., GONZALEZ, J.-P. J., GROSCHUP, M. H., GÜNTHER, S., HAENNI, A.-L., HALL, R. A., HEWSON, R., HUGHES, H. R., JAIN, R. K., JONSON, M. G., JUNGLEN, S., KLEMPA, B., KLINGSTRÖM, J., KORMELINK, R., LAMBERT, A. J., LANGEVIN, S. A., LUKASHEVICH, I. S., MARKLEWITZ, M., MARTELLI, G. P., MIELKE-EHRET, N., MIRAZIMI, A., MÜHLBACH, H.-P., NAIDU, R., NUNES, M. R. T., PALACIOS, G., PAPA, A., PAWĘSKA, J. T., PETERS, C. J., PLYUSNIN, A., RADOSHITZKY, S. R., RESENDE, R. O., ROMANOWSKI, V., SALL, A. A., SALVATO, M. S., SASAYA, T., SCHMALJOHN, C., SHÍ, X., SHIRAKO, Y., SIMMONDS, P., SIRONI, M., SONG, J.-W., SPENGLER, J. R., STENGLEIN, M. D., TESH, R. B., TURINA, M., WÈI, T., WHITFIELD, A. E., YEH, S.-D.,

ZERBINI, F. M., ZHANG, Y.-Z., ZHOU, X. & KUHN, J. H. 2019a. Taxonomy of the order Bunyavirales: second update 2018. *Archives of Virology,* 164**,** 927-941.

MAES, P., ADKINS, S., ALKHOVSKY, S. V., AVSIC-ZUPANC, T., BALLINGER, M. J., BENTE, D. A., BEER, M., BERGERON, E., BLAIR, C. D., BRIESE, T., BUCHMEIER, M. J., BURT, F. J., CALISHER, C. H., CHARREL, R. N., CHOI, I. R., CLEGG, J. C. S., DE LA TORRE, J. C., DE LAMBALLERIE, X., DERISI, J. L., DIGIARO, M., DREBOT, M., EBIHARA, H., ELBEAINO, T., ERGUNAY, K., FULHORST, C. F., GARRISON, A. R., GAO, G. F., GONZALEZ, J. J., GROSCHUP, M. H., GUNTHER, S., HAENNI, A. L., HALL, R. A., HEWSON, R., HUGHES, H. R., JAIN, R. K., JONSON, M. G., JUNGLEN, S., KLEMPA, B., KLINGSTROM, J., KORMELINK, R., LAMBERT, A. J., LANGEVIN, S. A., LUKASHEVICH, I. S., MARKLEWITZ, M., MARTELLI, G. P., MIELKE-EHRET, N., MIRAZIMI, A., MUHLBACH, H. P., NAIDU, R., NUNES, M. R. T., PALACIOS, G., PAPA, A., PAWESKA, J. T., PETERS, C. J., PLYUSNIN, A., RADOSHITZKY, S. R., RESENDE, R. O., ROMANOWSKI, V., SALL, A. A., SALVATO, M. S., SASAYA, T., SCHMALJOHN, C., SHI, X., SHIRAKO, Y., SIMMONDS, P., SIRONI, M., SONG, J. W., SPENGLER, J. R., STENGLEIN, M. D., TESH, R. B., TURINA, M., WEI, T., WHITFIELD, A. E., YEH, S. D., ZERBINI, F. M., ZHANG, Y. Z., ZHOU, X. & KUHN, J. H. 2019b. Taxonomy of the order Bunyavirales: second update 2018. *Arch Virol,* 164**,** 927-941.

MALLAM, A. L., SAE-LEE, W., SCHAUB, J. M., TU, F., BATTENHOUSE, A., JANG, Y. J., KIM, J., WALLINGFORD, J. B., FINKELSTEIN, I. J. & MARCOTTE, E. M. 2019. Systematic discovery of endogenous human ribonucleoprotein complexes. *Cell reports,* 29**,** 1351-1368. e5.

MALMLOV, A., SEETAHAL, J., CARRINGTON, C., RAMKISSON, V., FOSTER, J., MIAZGOWICZ, K. L., QUACKENBUSH, S., ROVNAK, J., NEGRETE, O., MUNSTER, V. & SCHOUNTZ, T. 2017. Serological evidence of arenavirus circulation among fruit bats in Trinidad. *PLOS ONE,* 12**,** e0185308.

MANNING, J. T., FORRESTER, N. & PAESSLER, S. 2015. Lassa mammarenavirus isolates from Mali and the Ivory Coast represent an emerging fifth lineage. *Frontiers in microbiology,* 6**,** 1037.

MANOKARAN, G., FINOL, E., WANG, C., GUNARATNE, J., BAHL, J., ONG, E. Z., TAN, H. C., SESSIONS, O. M., WARD, A. M., GUBLER, D. J., HARRIS, E., GARCIA-BLANCO, M. A. & OOI, E. E. 2015. Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science,* 350**,** 217-221.

MANTLO, PAESSLER & HUANG 2019. Differential Immune Responses to Hemorrhagic Fever-Causing Arenaviruses. *Vaccines,* 7**,** 138.

MARIËN, J., BORREMANS, B., GRYSEELS, S., SOROPOGUI, B., DE BRUYN, L., BONGO, G. N., BECKER-ZIAJA, B., DE BELLOCQ, J. G., GÜNTHER, S., MAGASSOUBA, N. F., LEIRS, H. & FICHET-CALVET, E. 2017. No measurable adverse effects of Lassa, Morogoro and Gairo arenaviruses on their rodent reservoir host in natural conditions. *Parasites & Vectors,* 10.

MARIËN, J., BORREMANS, B., VERHAEREN, C., KIRKPATRICK, L., GRYSEELS, S., GOÜY DE BELLOCQ, J., GÜNTHER, S., SABUNI, C. A., MASSAWE, A. W., REIJNIERS, J. & LEIRS, H. 2020. Density dependence and persistence of Morogoro arenavirus transmission in a fluctuating population of its reservoir host. *Journal of Animal Ecology,* 89**,** 506-518.

MARTIN-SERRANO, J. & NEIL, S. J. 2011. Host factors involved in retroviral budding and release. *Nat Rev Microbiol,* 9**,** 519-31.

MARTÍNEZ-SOBRIDO, L., EMONET, S., GIANNAKAS, P., CUBITT, B., GARCÍA-SASTRE, A. & DE LA TORRE, J. C. 2009. Identification of amino acid residues critical for the anti-interferon activity of the nucleoprotein of the prototypic arenavirus lymphocytic choriomeningitis virus. *J Virol,* 83**,** 11330-40.

MARTÍNEZ-SOBRIDO, L., GIANNAKAS, P., CUBITT, B., GARCÍA-SASTRE, A. & TORRE, J. C. D. L. 2007. Differential Inhibition of Type I Interferon Induction by Arenavirus Nucleoproteins. *Journal of Virology,* 81**,** 12696-12703.

MATEER, E. J., MARUYAMA, J., CARD, G. E., PAESSLER, S., HUANG, C. & HEISE, M. T. 2020. Lassa Virus, but Not Highly Pathogenic New World Arenaviruses, Restricts Immunostimulatory Double-Stranded RNA Accumulation during Infection. *Journal of Virology,* 94**,** e02006-19.

MATEO, M., REYNARD, S., JOURNEAUX, A., GERMAIN, C., HORTION, J., CARNEC, X., PICARD, C., BAILLET, N., BORGES-CARDOSO, V., MERABET, O., VALLVE, A., BARRON, S., JOURJON, O., LACROIX, O., DUTHEY, A., DIRHEIMER, M., JOUVION, G., MOREAU, P.-H., FELLMANN, L., CARBONNELLE, C., RAOUL, H., TANGY, F. & BAIZE, S. 2021. A single-shot Lassa vaccine induces long-term immunity and protects cynomolgus monkeys against heterologous strains. *Science Translational Medicine,* 13**,** eabf6348.

MATKOVIC, R., BERNARD, E., FONTANEL, S., ELDIN, P., CHAZAL, N., HERSI, D. H., MERITS, A., PÉLOPONÈSE, J.-M. & BRIANT, L. 2019. The Host DHX9 DExH-Box Helicase Is Recruited to Chikungunya Virus Replication Complexes for Optimal Genomic RNA Translation. *Journal of Virology,* 93**,** 10.1128/jvi.01764-18.

MATTHEWS, R. E. 1979. Third report of the International Committee on Taxonomy of Viruses. Classification and nomenclature of viruses. *Intervirology,* 12**,** 129-296.

MAXMEN, A. 2018. Deadly Lassa-fever outbreak tests Nigeria's revamped health agency. *Nature,* 555**,** 421-423.

MAZZOLA, L. T. & KELLY-CIRINO, C. 2019. Diagnostics for Lassa fever virus: a genetically diverse pathogen found in low-resource settings. *BMJ Global Health,* 4**,** e001116.

MCCORMICK, J. & FISHER-HOCH, S. 2002. Lassa fever. *Arenaviruses I: the epidemiology, molecular and cell biology of arenaviruses***,** 75-109.

MCCORMICK, J. B., KING, I. J., WEBB, P. A., SCRIBNER, C. L., CRAVEN, R. B., JOHNSON, K. M., ELLIOTT, L. H. & BELMONT-WILLIAMS, R. 1986. Lassa Fever. *New England Journal of Medicine,* 314**,** 20-26.

MCKEE, J., KELLY T., BARRERA ORO, J. G., KUEHNE, A. I., SPISSO, J. A. & MAHLANDT, B. G. 2008. Candid No. 1 Argentine Hemorrhagic Fever Vaccine Protects against Lethal Junin Virus Challenge in Rhesus Macaques. *Intervirology,* 34**,** 154-163.

MCLAY, L., LIANG, Y. & LY, H. 2014. Comparative analysis of disease pathogenesis and molecular mechanisms of New World and Old World arenavirus infections. *J Gen Virol,* 95**,** 1-15.

MEAGHER, J. L., TAKATA, M., GONÇALVES-CARNEIRO, D., KEANE, S. C., REBENDENNE, A., ONG, H., ORR, V. K., MACDONALD, M. R., STUCKEY, J. A., BIENIASZ, P. D. & SMITH, J. L. 2019. Structure of the zinc-finger antiviral protein in complex with RNA reveals a mechanism for selective targeting of CG-rich viral sequences. *Proc Natl Acad Sci U S A,* 116**,** 24303-24309.

MEDZHITOV, R. 2007. Recognition of microorganisms and activation of the immune response. *Nature,* 449**,** 819-826.
METTLER, N. E., CASALS, J. & SHOPE, R. E. 1963. Study of the Antigenic Relationships between Junín Virus, the Etiological Agent of Argentinian Hemorrhagic Fever, and other Arthropod-Borne Viruses. *American Journal of Tropical Medicine and Hygiene,* 12**,** 647-52.
MEYER, B. & LY, H. 2016a. Inhibition of Innate Immune Responses Is Key to Pathogenesis by Arenaviruses. *Journal of Virology,* 90**,** 3810-3818.
MEYER, B. & LY, H. 2016b. Inhibition of Innate Immune Responses Is Key to Pathogenesis by Arenaviruses. *J Virol,* 90**,** 3810-3818.
MEYER, B. J., DE LA TORRE, J. C. & SOUTHERN, P. J. 2002. Arenaviruses: Genomic RNAs, Transcription, and Replication. *In:* OLDSTONE, M. B. A. (ed.) *Arenaviruses I: The Epidemiology, Molecular and Cell Biology of Arenaviruses*. Berlin, Heidelberg: Springer Berlin Heidelberg.
MEYERSON, N. R., ZHOU, L., GUO, Y. R., ZHAO, C., TAO, Y. J., KRUG, R. M. & SAWYER, S. L. 2017. Nuclear TRIM25 Specifically Targets Influenza Virus Ribonucleoproteins to Block the Onset of RNA Chain Elongation. *Cell Host & Microbe,* 22**,** 627-638.e7.
MIBAYASHI, M., MARTÍNEZ-SOBRIDO, L., LOO, Y. M., CÁRDENAS, W. B., GALE, M., JR. & GARCÍA-SASTRE, A. 2007. Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza A virus. *J Virol,* 81**,** 514-24.
MILAZZO, M. L., CAJIMAT, M. N., DUNO, G., DUNO, F., UTRERA, A. & FULHORST, C. F. 2011. Transmission of Guanarito and Pirital viruses among wild rodents, Venezuela. *Emerg Infect Dis,* 17**,** 2209-15.
MILLS, J. N., ELLIS, B. A., MCKEE, K. T., JR., CALDERON, G. E., MAIZTEGUI, J. I., NELSON, G. O., KSIAZEK, T. G., PETERS, C. J. & CHILDS, J. E. 1992. A longitudinal study of Junin virus activity in the rodent reservoir of Argentine hemorrhagic fever. *Am J Trop Med Hyg,* 47**,** 749-63.
MOFOLORUNSHO, K. C. 2016. Outbreak of lassa fever in nigeria: Measures for prevention and control. *Pan African Medical Journal,* 23.
MORDECAI, G. J., MILLER, K. M., DI CICCO, E., SCHULZE, A. D., KAUKINEN, K. H., MING, T. J., LI, S., TABATA, A., TEFFER, A., PATTERSON, D. A., FERGUSON, H. W. & SUTTLE, C. A. 2019. Endangered wild salmon infected by newly discovered viruses. *Elife,* 8.
MORIN, B., COUTARD, B., LELKE, M., FERRON, F., KERBER, R., JAMAL, S., FRANGEUL, A., BARONTI, C., CHARREL, R., DE LAMBALLERIE, X., VONRHEIN, C., LESCAR, J., BRICOGNE, G., GÜNTHER, S. & CANARD, B. 2010. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. *PLoS Pathog,* 6**,** e1001038.
MORIN, B., KRANZUSCH, P. J., RAHMEH, A. A. & WHELAN, S. P. 2013. The polymerase of negative-stranded RNA viruses. *Curr Opin Virol,* 3**,** 103-10.
MURPHY, F. A., WEBB, P. A., JOHNSON, K. M. & WHITFIELD, S. G. 1969. Morphological comparison of Machupo with lymphocytic choriomeningitis virus: basis for a new taxonomic group. *J Virol,* 4**,** 535-41.
NEIL, S. J., ZANG, T. & BIENIASZ, P. D. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature,* 451**,** 425-30.
NGUYEN, N. T. H., NOW, H., KIM, W.-J., KIM, N. & YOO, J.-Y. 2016. Ubiquitin-like modifier FAT10 attenuates RIG-I mediated antiviral signaling by

segregating activated RIG-I from its signaling platform. *Scientific Reports,* 6**,** 23377.
NGUYEN, T. H. V., YEKWA, E., SELISKO, B., CANARD, B., ALVAREZ, K. & FERRON, F. 2022. Inhibition of Arenaviridae nucleoprotein exonuclease by bisphosphonate. *IUCrJ,* 9**,** 468-479.
NUNBERG, J. H. & YORK, J. 2012. The curious case of arenavirus entry, and its inhibition. *Viruses,* 4**,** 83-101.
ODON, V., FROS, J. J., GOONAWARDANE, N., DIETRICH, I., IBRAHIM, A., ALSHAIKHAHMED, K., NGUYEN, D. & SIMMONDS, P. 2019. The role of ZAP and OAS3/RNAseL pathways in the attenuation of an RNA virus with elevated frequencies of CpG and UpA dinucleotides. *Nucleic Acids Research,* 47**,** 8061-8083.
OESTEREICH, L., RIEGER, T., LÜDTKE, A., RUIBAL, P., WURR, S., PALLASCH, E., BOCKHOLT, S., KRASEMANN, S., MUÑOZ-FONTELA, C. & GÜNTHER, S. 2015. Efficacy of Favipiravir Alone and in Combination With Ribavirin in a Lethal, Immunocompetent Mouse Model of Lassa Fever. *The Journal of Infectious Diseases,* 213**,** 934-938.
OGBU, O., AJULUCHUKWU, E. & UNEKE, C. J. 2007. Lassa fever in West African sub-region: an overview. *J Vector Borne Dis,* 44**,** 1-11.
OKOGBENIN, S., OKOEGUALE, J., AKPEDE, G., COLUBRI, A., BARNES, K. G., MEHTA, S., EIFEDIYI, R., OKOGBO, F., EIGBEFOH, J., MOMOH, M., RAFIU, M., ADOMEH, D., ODIA, I., AIRE, C., ATAFO, R., OKONOFUA, M., PAHLMAN, M., BECKER-ZIAJA, B., ASOGUN, D., OKOKHERE, P., HAPPI, C., GÜNTHER, S., SABETI, P. C. & OGBAINI-EMOVON, E. 2019. Retrospective Cohort Study of Lassa Fever in Pregnancy, Southern Nigeria. *Emerg Infect Dis,* 25**,** 1494-500.
OKOKHERE, P., COLUBRI, A., AZUBIKE, C., IRUOLAGBE, C., OSAZUWA, O., TABRIZI, S., CHIN, E., ASAD, S., EDIALE, E. & RAFIU, M. 2018. Clinical and laboratory predictors of Lassa fever outcome in a dedicated treatment facility in Nigeria: a retrospective, observational cohort study. *The Lancet Infectious Diseases,* 18**,** 684-695.
OKUMURA, A., LU, G., PITHA-ROWE, I. & PITHA, P. M. 2006. Innate antiviral response targets HIV-1 release by the induction of ubiquitin-like protein ISG15. *Proc Natl Acad Sci U S A,* 103**,** 1440-5.
OKUMURA, A., PITHA, P. M. & HARTY, R. N. 2008. ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. *Proc Natl Acad Sci U S A,* 105**,** 3974-9.
OLSCHEWSKI, S., CUSACK, S. & ROSENTHAL, M. 2020. The Cap-Snatching Mechanism of Bunyaviruses. *Trends in Microbiology,* 28**,** 293-303.
OPPLIGER, J., TORRIANI, G., HERRADOR, A., KUNZ, S. & DERMODY, T. S. 2016. Lassa Virus Cell Entry via Dystroglycan Involves an Unusual Pathway of Macropinocytosis. *Journal of Virology,* 90**,** 6412-6429.
ORTIZ-RIAÑO, E., CHENG, B. Y., DE LA TORRE, J. C. & MARTÍNEZ-SOBRIDO, L. 2012a. Self-association of lymphocytic choriomeningitis virus nucleoprotein is mediated by its N-terminal region and is not required for its anti-interferon function. *J Virol,* 86**,** 3307-17.
ORTIZ-RIAÑO, E., CHENG, B. Y. H., TORRE, J. C. D. L. & MARTÍNEZ-SOBRIDO, L. 2012b. Self-Association of Lymphocytic Choriomeningitis Virus Nucleoprotein Is Mediated by Its N-Terminal Region and Is Not Required for Its Anti-Interferon Function. *Journal of Virology,* 86**,** 3307-3317.
OSADA, N., KOHARA, A., YAMAJI, T., HIRAYAMA, N., KASAI, F., SEKIZUKA, T., KURODA, M. & HANADA, K. 2014. The Genome Landscape of the African

Green Monkey Kidney-Derived Vero Cell Line. *DNA Research,* 21**,** 673-683.
OSHIUMI, H., MIYASHITA, M., MATSUMOTO, M. & SEYA, T. 2013. A Distinct Role of Riplet-Mediated K63-Linked Polyubiquitination of the RIG-I Repressor Domain in Human Antiviral Innate Immune Responses. *PLoS Pathogens,* 9**,** e1003533.
PANNETIER, D., REYNARD, S., RUSSIER, M., JOURNEAUX, A., TORDO, N., DEUBEL, V. & BAIZE, S. 2011. Human Dendritic Cells Infected with the Nonpathogenic Mopeia Virus Induce Stronger T-Cell Responses than Those Infected with Lassa Virus. *Journal of Virology,* 85**,** 8293-8306.
PARISI, G., ECHAVE, J., GHIRINGHELLI, D. & ROMANOWSKI, V. 1996. Computational characterisation of potential RNA-binding sites in arenavirus nucleocapsid proteins. *Virus Genes,* 13**,** 247-54.
PARODI, A. S., GREENWAY, D. J., RUGIERO, H. R., FRIGERIO, M., DE LA BARRERA, J. M., METTLER, N., GARZON, F., BOXACA, M., GUERRERO, L. & NOTA, N. 1958. [Concerning the epidemic outbreak in Junin]. *Dia Med,* 30**,** 2300-1.
PARODI, A. S., RUGIERO, H. R., GREENWAY, D. J., METTLER, N., MARTINEZ, A., BOXACA, M. & DE LA BARRERA, J. M. 1959. [Isolation of the Junin virus (epidemic hemorrhagic fever) from the mites of the epidemic area (Echinolaelaps echidninus, Barlese)]. *Prensa Med Argent,* 46**,** 2242-4.
PASQUAL, G., ROJEK, J. M., MASIN, M., CHATTON, J.-Y. & KUNZ, S. 2011. Old World Arenaviruses Enter the Host Cell via the Multivesicular Body and Depend on the Endosomal Sorting Complex Required for Transport. *PLoS Pathogens,* 7**,** e1002232.
PATTIS, J. G. & MAY, E. R. 2020. Markov State Model of Lassa Virus Nucleoprotein Reveals Large Structural Changes during the Trimer to Monomer Transition. *Structure,* 28**,** 548-554.e3.
PAWESKA, J. T., SEWLALL, N. H., KSIAZEK, T. G., BLUMBERG, L. H., HALE, M. J., LIPKIN, W. I., WEYER, J., NICHOL, S. T., ROLLIN, P. E., MCMULLAN, L. K., PADDOCK, C. D., BRIESE, T., MNYALUZA, J., DINH, T. H., MUKONKA, V., CHING, P., DUSE, A., RICHARDS, G., DE JONG, G., COHEN, C., IKALAFENG, B., MUGERO, C., ASOMUGHA, C., MALOTLE, M. M., NTEO, D. M., MISIANI, E., SWANEPOEL, R., ZAKI, S. R., OUTBREAK, C. & INVESTIGATION, T. 2009. Nosocomial outbreak of novel arenavirus infection, southern Africa. *Emerg Infect Dis,* 15**,** 1598-602.
PEDERSEN, I. R. 1971. Lymphocytic Choriomeningitis Virus RNAs. *Nature New Biology,* 234**,** 112-114.
PENG, R., XU, X., JING, J., WANG, M., PENG, Q., LIU, S., WU, Y., BAO, X., WANG, P., QI, J., GAO, G. F. & SHI, Y. 2020. Structural insight into arenavirus replication machinery. *Nature,* 579**,** 615-619.
PEREZ, J. T., VARBLE, A., SACHIDANANDAM, R., ZLATEV, I., MANOHARAN, M., GARCÍA-SASTRE, A. & TENOEVER, B. R. 2010. Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *Proc Natl Acad Sci U S A,* 107**,** 11525-30.
PINSCHEWER, D. D., PEREZ, M. & TORRE, J. C. D. L. 2005. Dual Role of the Lymphocytic Choriomeningitis Virus Intergenic Region in Transcription Termination and Virus Propagation. *Journal of Virology,* 79**,** 4519-4526.
PONTREMOLI, C., FORNI, D., CAGLIANI, R. & SIRONI, M. 2018. Analysis of Reptarenavirus genomes indicates different selective forces acting on the S and L segments and recent expansion of common genotypes. *Infect Genet Evol,* 64**,** 212-218.

PONTREMOLI, C., FORNI, D. & SIRONI, M. 2019. Arenavirus genomics: novel insights into viral diversity, origin, and evolution. *Curr Opin Virol,* 34**,** 18-28.

PRICE, M. E., FISHER-HOCH, S. P., CRAVEN, R. B. & MCCORMICK, J. B. 1988. A prospective study of maternal and fetal outcome in acute Lassa fever infection during pregnancy. *Bmj,* 297**,** 584-7.

PUTNAM, C. D., HAMMEL, M., HURA, G. L. & TAINER, J. A. 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics,* 40**,** 191-285.

QI, X., LAN, S., WANG, W., SCHELDE, L. M., DONG, H., WALLAT, G. D., LY, H., LIANG, Y. & DONG, C. 2010a. Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature,* 468**,** 779-783.

QI, X., LAN, S., WANG, W., SCHELDE, L. M., DONG, H., WALLAT, G. D., LY, H., LIANG, Y. & DONG, C. 2010b. Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature,* 468**,** 779-83.

RAABE, V. & KOEHLER, J. 2017. Laboratory Diagnosis of Lassa Fever. *J Clin Microbiol,* 55**,** 1629-1637.

RAABE, V. N., KANN, G., RIBNER, B. S., MORALES, A., VARKEY, J. B., MEHTA, A. K., LYON, G. M., VANAIRSDALE, S., FABER, K., BECKER, S., EICKMANN, M., STRECKER, T., BROWN, S., PATEL, K., DE LEUW, P., SCHUETTFORT, G., STEPHAN, C., RABENAU, H., KLENA, J. D., ROLLIN, P. E., MCELROY, A., STRÖHER, U., NICHOL, S., KRAFT, C. S., WOLF, T. & UNIT, F. T. E. S. C. D. 2017. Favipiravir and Ribavirin Treatment of Epidemiologically Linked Cases of Lassa Fever. *Clinical Infectious Diseases,* 65**,** 855-859.

RAABEN, M., JAE, L. T., HERBERT, A. S., KUEHNE, A. I., STUBBS, S. H., CHOU, Y.-Y., BLOMEN, V. A., KIRCHHAUSEN, T., DYE, J. M., BRUMMELKAMP, T. R. & WHELAN, S. P. 2017. NRP2 and CD63 Are Host Factors for Lujo Virus Cell Entry. *Cell Host & Microbe,* 22**,** 688-696.e5.

RADOSHITZKY, S., KUHN, J., JAHRLING, P. & BAVARI, S. 2018. HEMORRHAGIC FEVER-CAUSING MAMMARENAVIRUSES.

RADOSHITZKY, S. R., BÀO, Y., BUCHMEIER, M. J., CHARREL, R. N., CLAWSON, A. N., CLEGG, C. S., DERISI, J. L., EMONET, S., GONZALEZ, J.-P., KUHN, J. H., LUKASHEVICH, I. S., PETERS, C. J., ROMANOWSKI, V., SALVATO, M. S., STENGLEIN, M. D. & DE LA TORRE, J. C. 2015a. Past, present, and future of arenavirus taxonomy. *Archives of Virology,* 160**,** 1851-1874.

RADOSHITZKY, S. R., BAO, Y., BUCHMEIER, M. J., CHARREL, R. N., CLAWSON, A. N., CLEGG, C. S., DERISI, J. L., EMONET, S., GONZALEZ, J. P., KUHN, J. H., LUKASHEVICH, I. S., PETERS, C. J., ROMANOWSKI, V., SALVATO, M. S., STENGLEIN, M. D. & DE LA TORRE, J. C. 2015b. Past, present, and future of arenavirus taxonomy. *Arch Virol,* 160**,** 1851-74.

RADOSHITZKY, S. R., BUCHMEIER, M. J., CHARREL, R. N., GONZALEZ, J.-P. J., GÜNTHER, S., HEPOJOKI, J., KUHN, J. H., LUKASHEVICH, I. S., ROMANOWSKI, V., SALVATO, M. S., SIRONI, M., STENGLEIN, M. D. & TORRE, J. C. D. L. 2023. ICTV Virus Taxonomy Profile: Arenaviridae 2023. *Journal of General Virology,* 104.

RADOSHITZKY, S. R., DONG, L., CHI, X., CLESTER, J. C., RETTERER, C., SPURGERS, K., KUHN, J. H., SANDWICK, S., RUTHEL, G., KOTA, K., BOLTZ, D., WARREN, T., KRANZUSCH, P. J., WHELAN, S. P. & BAVARI, S. 2010. Infectious Lassa virus, but not filoviruses, is restricted by BST-2/tetherin. *J Virol,* 84**,** 10569-80.

REGUERA, J., GERLACH, P., ROSENTHAL, M., GAUDON, S., COSCIA, F., GÜNTHER, S. & CUSACK, S. 2016. Comparative Structural and Functional Analysis of Bunyavirus and Arenavirus Cap-Snatching Endonucleases. *PLOS Pathogens,* 12**,** e1005636.

REHWINKEL, J. & GACK, M. U. 2020. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nature Reviews Immunology,* 20**,** 537-551.

REUTER, G., BOROS, Á., TAKÁTS, K., MÁTICS, R. & PANKOVICS, P. 2023. A novel mammarenavirus (family Arenaviridae) in hedgehogs (Erinaceus roumanicus) in Europe. *Arch Virol,* 168**,** 174.

RHEE, H.-W., ZOU, P., UDESHI, N. D., MARTELL, J. D., MOOTHA, V. K., CARR, S. A. & TING, A. Y. 2013. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science,* 339**,** 1328-1331.

RICHMOND, J. K. & BAGLOLE, D. J. 2003. Lassa fever: epidemiology, clinical features, and social consequences. *Bmj,* 327**,** 1271-1275.

RIVERS, T. M. & MCNAIR SCOTT, T. F. 1935. MENINGITIS IN MAN CAUSED BY A FILTERABLE VIRUS. *Science,* 81**,** 439-440.

ROBINSON, J. E., HASTIE, K. M., CROSS, R. W., YENNI, R. E., ELLIOTT, D. H., ROUELLE, J. A., KANNADKA, C. B., SMIRA, A. A., GARRY, C. E., BRADLEY, B. T., YU, H., SHAFFER, J. G., BOISEN, M. L., HARTNETT, J. N., ZANDONATTI, M. A., ROWLAND, M. M., HEINRICH, M. L., MARTÍNEZ-SOBRIDO, L., CHENG, B., DE LA TORRE, J. C., ANDERSEN, K. G., GOBA, A., MOMOH, M., FULLAH, M., GBAKIE, M., KANNEH, L., KOROMA, V. J., FONNIE, R., JALLOH, S. C., KARGBO, B., VANDI, M. A., GBETUWA, M., IKPONMWOSA, O., ASOGUN, D. A., OKOKHERE, P. O., FOLLARIN, O. A., SCHIEFFELIN, J. S., PITTS, K. R., GEISBERT, J. B., KULAKOSKI, P. C., WILSON, R. B., HAPPI, C. T., SABETI, P. C., GEVAO, S. M., KHAN, S. H., GRANT, D. S., GEISBERT, T. W., SAPHIRE, E. O., BRANCO, L. M. & GARRY, R. F. 2016. Most neutralizing human monoclonal antibodies target novel epitopes requiring both Lassa virus glycoprotein subunits. *Nature Communications,* 7**,** 11544.

RODRÍGUEZ, M. E., BRUNETTI, J. E., WACHSMAN, M. B., SCOLARO, L. A. & CASTILLA, V. 2014. Raf/MEK/ERK pathway activation is required for Junin virus replication. *Journal of General Virology,* 95**,** 799-805.

ROJEK, J. M., SANCHEZ, A. B., NGUYEN, N. T., TORRE, J.-C. D. L. & KUNZ, S. 2008. Different Mechanisms of Cell Entry by Human-Pathogenic Old World and New World Arenaviruses. *Journal of Virology,* 82**,** 7677-7687.

ROSENKE, K., FELDMANN, H., WESTOVER, J., HANLEY, P. W., MARTELLARO, C., FELDMANN, F., SATURDAY, G., LOVAGLIO, J., SCOTT, D., FURUTA, Y., KOMENO, T., GOWEN, B. & SAFRONETZ, D. 2018. Use of Favipiravir to Treat Lassa Virus Infection in Macaques. *Emerging Infectious Disease journal,* 24**,** 1696.

ROUX, K. J., KIM, D. I., RAIDA, M. & BURKE, B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of cell biology,* 196**,** 801-810.

ROWE, W. P., MURPHY, F. A., BERGOLD, G. H., CASALS, J., HOTCHIN, J., JOHNSON, K. M., LEHMANN-GRUBE, F., MIMS, C. A., TRAUB, E. & WEBB, P. A. 1970a. Arenoviruses: proposed name for a newly defined virus group. *J Virol,* 5**,** 651-2.

ROWE, W. P., PUGH, W. E., WEBB, P. A. & PETERS, C. J. 1970b. Serological relationship of the Tacaribe complex of viruses to lymphocytic choriomeningitis virus. *J Virol,* 5**,** 289-92.

RUGIERO, H. R., PARODI, A. S., GOTTA, H., BOXACA, M., OLIVARI, A. J. & GONZALEZ, E. 1962. [Epidemic hemorrhagic fever. Laboratory infection and inter-human passage]. *Rev Asoc Med Argent,* 76**,** 413-7.

SAFRONETZ, D., ROSENKE, K., WESTOVER, J. B., MARTELLARO, C., OKUMURA, A., FURUTA, Y., GEISBERT, J., SATURDAY, G., KOMENO, T., GEISBERT, T. W., FELDMANN, H. & GOWEN, B. B. 2015. The broad-spectrum antiviral favipiravir protects guinea pigs from lethal Lassa virus infection post-disease onset. *Scientific Reports,* 5**,** 14775.

SAKUMA, T., NODA, T., URATA, S., KAWAOKA, Y. & YASUDA, J. 2009. Inhibition of Lassa and Marburg virus production by tetherin. *J Virol,* 83**,** 2382-5.

SALAMI, K., GOUGLAS, D., SCHMALJOHN, C., SAVILLE, M. & TORNIEPORTH, N. 2019. A review of Lassa fever vaccine candidates. *Curr Opin Virol,* 37**,** 105-111.

SALKELD, D., HOPKINS, S., HAYMAN, D., SALKELD, D., HOPKINS, S. & HAYMAN, D. 2023. Reservoir hosts. *Emerging Zoonotic and Wildlife Pathogens: Disease Ecology, Epidemiology, and Conservation*. Oxford University Press.

SALVATO, M., CJ, B. M., CHARREL, R., GONZALES, J. & LUKASHEVICH, I. 2005. Family Arenaviridae. Virus Taxonomy Eighth report of the International Committee on Taxonomy of Viruses. Elsevier, Academic Press.

SÁNCHEZ-APARICIO, M. T., FEINMAN, L. J., GARCÍA-SASTRE, A. & SHAW, M. L. 2018. Paramyxovirus V Proteins Interact with the RIG-I/TRIM25 Regulatory Complex and Inhibit RIG-I Signaling. *Journal of Virology,* 92**,** 10.1128/jvi.01960-17.

SANCHEZ, JACINT G., CHIANG, JESSICA J., SPARRER, KONSTANTIN M. J., ALAM, STEVEN L., CHI, M., ROGANOWICZ, MARCIN D., SANKARAN, B., GACK, MICHAELA U. & PORNILLOS, O. 2016. Mechanism of TRIM25 Catalytic Activation in the Antiviral RIG-I Pathway. *Cell Reports,* 16**,** 1315-1325.

SANCHEZ, J. G., OKREGLICKA, K., CHANDRASEKARAN, V., WELKER, J. M., SUNDQUIST, W. I. & PORNILLOS, O. 2014. The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer. *Proceedings of the National Academy of Sciences,* 111**,** 2494-2499.

SANCHEZ, J. G., SPARRER, K. M. J., CHIANG, C., REIS, R. A., CHIANG, J. J., ZURENSKI, M. A., WAN, Y., GACK, M. U. & PORNILLOS, O. 2018. TRIM25 Binds RNA to Modulate Cellular Anti-viral Defense. *Journal of Molecular Biology,* 430**,** 5280-5293.

SÄNGER, L., WILLIAMS, H. M., YU, D., VOGEL, D., KOSINSKI, J., ROSENTHAL, M. & UETRECHT, C. 2023. RNA to Rule Them All: Critical Steps in Lassa Virus Ribonucleoparticle Assembly and Recruitment. *Journal of the American Chemical Society,* 145**,** 27958-27974.

SARDIELLO, M., CAIRO, S., FONTANELLA, B., BALLABIO, A. & MERONI, G. 2008. Genomic analysis of the TRIM family reveals two groups of genes with distinct evolutionary properties. *BMC Evol Biol,* 8**,** 225.

SARUTE, N. & ROSS, S. R. 2017. New World Arenavirus Biology. *Annu Rev Virol,* 4**,** 141-158.

SAYLER, K. A., BARBET, A. F., CHAMBERLAIN, C., CLAPP, W. L., ALLEMAN, R., LOEB, J. C. & LEDNICKY, J. A. 2014. Isolation of Tacaribe virus, a Caribbean arenavirus, from host-seeking Amblyomma americanum ticks in Florida. *PLoS One,* 9**,** e115769.

SCHILLIGER, L., SELLERI, P. & FRYE, F. L. 2011. Lymphoblastic lymphoma and leukemic blood profile in a red-tail boa (Boa constrictor constrictor) with concurrent inclusion body disease. *J Vet Diagn Invest,* 23**,** 159-62.

SCHLIE, K., MAISA, A., FREIBERG, F., GROSETH, A., STRECKER, T. & GARTEN, W. 2010. Viral protein determinants of Lassa virus entry and release from polarized epithelial cells. *J Virol,* 84**,** 3178-88.

SCHOGGINS, J. W. 2019. Interferon-Stimulated Genes: What Do They All Do? *Annu Rev Virol,* 6**,** 567-584.

SCHWARTZ, S. L., PARK, E. N., VACHON, V. K., DANZY, S., LOWEN, ANICE C. & CONN, G. L. 2020. Human OAS1 activation is highly dependent on both RNA sequence and context of activating RNA motifs. *Nucleic Acids Research,* 48**,** 7520-7531.

SCHWERK, J., SOVEG, F. W., RYAN, A. P., THOMAS, K. R., HATFIELD, L. D., OZARKAR, S., FORERO, A., KELL, A. M., ROBY, J. A., SO, L., HYDE, J. L., GALE, M., JR., DAUGHERTY, M. D. & SAVAN, R. 2019. RNA-binding protein isoforms ZAP-S and ZAP-L have distinct antiviral and immune resolution functions. *Nat Immunol,* 20**,** 1610-1620.

SEO, E. J. & LEIS, J. 2012. Budding of Enveloped Viruses: Interferon-Induced ISG15-Antivirus Mechanisms Targeting the Release Process. *Adv Virol,* 2012**,** 532723.

SETH, R. B., SUN, L., EA, C.-K. & CHEN, Z. J. 2005. Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF-κB and IRF3. *Cell,* 122**,** 669-682.

SEVILLA, N., KUNZ, S., HOLZ, A., LEWICKI, H., HOMANN, D., YAMADA, H., CAMPBELL, K. P., DE LA TORRE, J. C. & OLDSTONE, M. B. A. 2000. Immunosuppression and Resultant Viral Persistence by Specific Viral Targeting of Dendritic Cells. *Journal of Experimental Medicine,* 192**,** 1249-1260.

SHAFFER, J. G., GRANT, D. S., SCHIEFFELIN, J. S., BOISEN, M. L., GOBA, A., HARTNETT, J. N., LEVY, D. C., YENNI, R. E., MOSES, L. M., FULLAH, M., MOMOH, M., FONNIE, M., FONNIE, R., KANNEH, L., KOROMA, V. J., KARGBO, K., OTTOMASSATHIEN, D., MUNCY, I. J., JONES, A. B., ILLICK, M. M., KULAKOSKY, P. C., HAISLIP, A. M., BISHOP, C. M., ELLIOT, D. H., BROWN, B. L., ZHU, H., HASTIE, K. M., ANDERSEN, K. G., GIRE, S. K., TABRIZI, S., TARIYAL, R., STREMLAU, M., MATSCHINER, A., SAMPEY, D. B., SPENCE, J. S., CROSS, R. W., GEISBERT, J. B., FOLARIN, O. A., HAPPI, C. T., PITTS, K. R., GESKE, F. J., GEISBERT, T. W., SAPHIRE, E. O., ROBINSON, J. E., WILSON, R. B., SABETI, P. C., HENDERSON, L. A., KHAN, S. H., BAUSCH, D. G., BRANCO, L. M. & GARRY, R. F. 2014. Lassa Fever in Post-Conflict Sierra Leone. *PLoS Neglected Tropical Diseases,* 8**,** e2748.

SHAO, J., LIANG, Y. & LY, H. 2018. Roles of Arenavirus Z Protein in Mediating Virion Budding, Viral Transcription-Inhibition and Interferon-Beta Suppression. *Methods Mol Biol,* 1604**,** 217-227.

SHAW, A. B., TSE, H. N., BYFORD, O., PLAHE, G., MOON-WALKER, A., HOVER, S. E., SAPHIRE, E. O., WHELAN, S. P. J., MANKOURI, J., FONTANA, J. & BARR, J. N. 2024. Cellular endosomal potassium ion flux regulates arenavirus uncoating during virus entry. *mBio,* 15**,** e0168423.

SHI, M., LIN, X.-D., CHEN, X., TIAN, J.-H., CHEN, L.-J., LI, K., WANG, W., EDEN, J.-S., SHEN, J.-J., LIU, L., HOLMES, E. C. & ZHANG, Y.-Z. 2018a. The evolutionary history of vertebrate RNA viruses. *Nature,* 556**,** 197-202.

SHI, M., LIN, X. D., CHEN, X., TIAN, J. H., CHEN, L. J., LI, K., WANG, W., EDEN, J. S., SHEN, J. J., LIU, L., HOLMES, E. C. & ZHANG, Y. Z. 2018b. The evolutionary history of vertebrate RNA viruses. *Nature,* 556**,** 197-202.

SHIMOJIMA, M., STRÖHER, U., EBIHARA, H., FELDMANN, H. & KAWAOKA, Y. 2012. Identification of Cell Surface Molecules Involved in Dystroglycan-Independent Lassa Virus Cell Entry. *Journal of Virology,* 86**,** 2067-2078.

SHIMOJIMA, M., TAKADA, A., EBIHARA, H., NEUMANN, G., FUJIOKA, K., IRIMURA, T., JONES, S., FELDMANN, H. & KAWAOKA, Y. 2006. Tyro3 family-mediated cell entry of Ebola and Marburg viruses. *J Virol,* 80**,** 10109-16.

SHRESTHA, N., BAHNAN, W., WILEY, D. J., BARBER, G., FIELDS, K. A. & SCHESSER, K. 2012. Eukaryotic initiation factor 2 (eIF2) signaling regulates proinflammatory cytokine expression and bacterial invasion. *J Biol Chem,* 287**,** 28738-44.

SHUPP, A., CASIMIRO, M. C. & PESTELL, R. G. 2017. Biological functions of CDK5 and potential CDK5 targeted clinical treatments. *Oncotarget,* 8**,** 17373-17382.

SILVERMAN, R. H. & WEISS, S. R. 2014. Viral Phosphodiesterases That Antagonize Double-Stranded RNA Signaling to RNase L by Degrading 2-5A. *Journal of Interferon & Cytokine Research,* 34**,** 455-463.

SIMMONDS, P. 2012. SSE: a nucleotide and amino acid sequence analysis platform. *BMC Research Notes,* 5**,** 50.

SIMMONDS, P., XIA, W., BAILLIE, J. K. & MCKINNON, K. 2013. Modelling mutational and selection pressures on dinucleotides in eukaryotic phyla--selection against CpG and UpA in cytoplasmically expressed RNA and in RNA viruses. *BMC Genomics,* 14**,** 610.

SMELT, S. C., BORROW, P., KUNZ, S., CAO, W., TISHON, A., LEWICKI, H., CAMPBELL, K. P. & OLDSTONE, M. B. A. 2001. Differences in Affinity of Binding of Lymphocytic Choriomeningitis Virus Strains to the Cellular Receptor α-Dystroglycan Correlate with Viral Tropism and Disease Kinetics. *Journal of Virology,* 75**,** 448-457.

SOMMERSTEIN, R., FLATZ, L., REMY, M. M., MALINGE, P., MAGISTRELLI, G., FISCHER, N., SAHIN, M., BERGTHALER, A., IGONET, S., TER MEULEN, J., RIGO, D., MEDA, P., RABAH, N., COUTARD, B., BOWDEN, T. A., LAMBERT, P.-H., SIEGRIST, C.-A. & PINSCHEWER, D. D. 2015. Arenavirus Glycan Shield Promotes Neutralizing Antibody Evasion and Protracted Infection. *PLOS Pathogens,* 11**,** e1005276.

SONAWANE, N. D., SZOKA, F. C., JR. & VERKMAN, A. S. 2003. Chloride accumulation and swelling in endosomes enhances DNA transfer by polyamine-DNA polyplexes. *J Biol Chem,* 278**,** 44826-31.

SPEIR, R. W., WOOD, O., LIEBHABER, H. & BUCKLEY, S. M. 1970. Lassa fever, a new virus disease of man from West Africa. IV. Electron microscopy of Vero cell cultures infected with Lassa virus. *Am J Trop Med Hyg,* 19**,** 692-4.

SPIROPOULOU, C. F., KUNZ, S., ROLLIN, P. E., CAMPBELL, K. P. & OLDSTONE, M. B. A. 2002. New World Arenavirus Clade C, but Not Clade A and B Viruses, Utilizes α-Dystroglycan as Its Major Receptor. *Journal of Virology,* 76**,** 5140-5146.

STEIGENBERGER, B., ALBANESE, P., HECK, A. & SCHELTEMA, R. 2019a. To cleave or not to cleave in XL-MS? *Journal of the American Society for Mass Spectrometry,* 31**,** 196-206.

STEIGENBERGER, B., PIETERS, R. J., HECK, A. J. & SCHELTEMA, R. A. 2019b. PhoX: an IMAC-enrichable cross-linking reagent. *ACS central science,* 5**,** 1514-1522.

STEITZ, T. A. & STEITZ, J. A. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A,* 90**,** 6498-502.

STENGLEIN, M. D., SANDERS, C., KISTLER, A. L., RUBY, J. G., FRANCO, J. Y., REAVILL, D. R., DUNKER, F. & DERISI, J. L. 2012a. 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**,** 10.1128/mbio.00180-12.

STENGLEIN, M. D., SANDERS, C., KISTLER, A. L., RUBY, J. G., FRANCO, J. Y., REAVILL, D. R., DUNKER, F., DERISI, J. L. & BUCHMEIER, M. J. 2012b. 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**,** e00180-12.

STOTT, R. J., STRECKER, T. & FOSTER, T. L. 2020. Distinct Molecular Mechanisms of Host Immune Response Modulation by Arenavirus NP and Z Proteins. *Viruses,* 12.

STRECKER, T., MAISA, A., DAFFIS, S., EICHLER, R., LENZ, O. & GARTEN, W. 2006. The role of myristoylation in the membrane association of the Lassa virus matrix protein Z. *Virol J,* 3**,** 93.

SWIECKI, M., SCHEAFFER, S. M., ALLAIRE, M., FREMONT, D. H., COLONNA, M. & BRETT, T. J. 2011. Structural and biophysical analysis of BST-2/tetherin ectodomains reveals an evolutionary conserved design to inhibit virus release. *J Biol Chem,* 286**,** 2987-97.

TAGLIAPIETRA, V., ROSÀ, R., HAUFFE, H. C., LAAKKONEN, J., VOUTILAINEN, L., VAPALAHTI, O., VAHERI, A., HENTTONEN, H. & RIZZOLI, A. 2009. Spatial and temporal dynamics of lymphocytic choriomeningitis virus in wild rodents, northern Italy. *Emerg Infect Dis,* 15**,** 1019-25.

TAKATA, M. A., GONÇALVES-CARNEIRO, D., ZANG, T. M., SOLL, S. J., YORK, A., BLANCO-MELO, D. & BIENIASZ, P. D. 2017. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature,* 550**,** 124-127.

TAKEUCHI, O. & AKIRA, S. 2009. Innate immunity to virus infection. *Immunological Reviews,* 227**,** 75-86.

TAN, C. S. H., GO, K. D., BISTEAU, X., DAI, L., YONG, C. H., PRABHU, N., OZTURK, M. B., LIM, Y. T., SREEKUMAR, L. & LENGQVIST, J. 2018. Thermal proximity coaggregation for system-wide profiling of protein complex dynamics in cells. *Science,* 359**,** 1170-1177.

TAN, D., LI, Q., ZHANG, M.-J., LIU, C., MA, C., ZHANG, P., DING, Y.-H., FAN, S.-B., TAO, L. & YANG, B. 2016. Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states. *Elife,* 5**,** e12509.

TANG, Q., WANG, X. & GAO, G. 2017. The Short Form of the Zinc Finger Antiviral Protein Inhibits Influenza A Virus Protein Expression and Is Antagonized by the Virus-Encoded NS1. *J Virol,* 91.

TANG, X., WIPPEL, H. H., CHAVEZ, J. D. & BRUCE, J. E. 2021. Crosslinking mass spectrometry: a link between structural biology and systems biology. *Protein Science,* 30**,** 773-784.

TAYLOR, S., PONZINI, M., WILSON, M. & KIM, K. 2022. Comparison of imputation and imputation-free methods for statistical analysis of mass spectrometry data with missing data. *Briefings in Bioinformatics,* 23.

TENG, M. N., BORROW, P., OLDSTONE, M. B. & TORRE, J. C. D. L. 1996. A single amino acid change in the glycoprotein of lymphocytic choriomeningitis virus is associated with the ability to cause growth hormone deficiency syndrome. *Journal of Virology,* 70**,** 8438-8443.

TITECA, K., LEMMENS, I., TAVERNIER, J. & EYCKERMAN, S. 2019. Discovering cellular protein-protein interactions: Technological strategies and opportunities. *Mass spectrometry reviews,* 38**,** 79-111.
TORTORICI, M. A., GHIRINGHELLI, P. D., LOZANO, M. E., ALBARIÑO, C. G. & ROMANOWSKI, V. 2001. Zinc-binding properties of Junín virus nucleocapsid protein. *Journal of General Virology,* 82**,** 121-128.
TRAUB, E. 1935. A Filterable Virus Recovered from White Mice. *Science,* 81**,** 298-299.
TRAUB, E. 1936. AN EPIDEMIC IN A MOUSE COLONY DUE TO THE VIRUS OF ACUTE LYMPHOCYTIC CHORIOMENINGITIS. *Journal of Experimental Medicine,* 63**,** 533-546.
TRIFINOPOULOS, J., NGUYEN, L. T., VON HAESELER, A. & MINH, B. Q. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res,* 44**,** W232-5.
TRINKLE-MULCAHY, L. 2019. Recent advances in proximity-based labeling methods for interactome mapping. *F1000Research,* 8.
TULLOCH, F., ATKINSON, N. J., EVANS, D. J., RYAN, M. D. & SIMMONDS, P. 2014. RNA virus attenuation by codon pair deoptimisation is an artefact of increases in CpG/UpA dinucleotide frequencies. *Elife,* 3**,** e04531.
URATA, S. & YASUDA, J. 2012. Molecular Mechanism of Arenavirus Assembly and Budding. *Viruses,* 4**,** 2049-2079.
VASCONCELOS, P. F., TRAVASSOS DA ROSA, A. P., RODRIGUES, S. G., TESH, R., TRAVASSOS DA ROSA, J. F. & TRAVASSOS DA ROSA, E. S. 1993. [Laboratory-acquired human infection with SP H 114202 virus (Arenavirus: Arenaviridae family): clinical and laboratory aspects]. *Rev Inst Med Trop Sao Paulo,* 35**,** 521-5.
VERMA, R., BOLETI, E. & GEORGE, A. J. T. 1998. Antibody engineering: Comparison of bacterial, yeast, insect and mammalian expression systems. *Journal of Immunological Methods,* 216**,** 165-181.
VIDYA, M. K., KUMAR, V. G., SEJIAN, V., BAGATH, M., KRISHNAN, G. & BHATTA, R. 2018. Toll-like receptors: Significance, ligands, signaling pathways, and functions in mammals. *International Reviews of Immunology,* 37**,** 20-36.
VIETH, S., TORDA, A. E., ASPER, M., SCHMITZ, H. & GÜNTHER, S. 2004. Sequence analysis of L RNA of Lassa virus. *Virology,* 318**,** 153-168.
VOGEL, D., ROSENTHAL, M., GOGREFE, N., REINDL, S. & GÜNTHER, S. 2019. Biochemical characterization of the Lassa virus L protein. *J Biol Chem,* 294**,** 8088-8100.
VOLPON, L., OSBORNE, M. J., CAPUL, A. A., DE LA TORRE, J. C. & BORDEN, K. L. B. 2010. Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E. *Proceedings of the National Academy of Sciences,* 107**,** 5441-5446.
VON HEIJNE, G. 1985. Signal sequences. The limits of variation. *J Mol Biol,* 184**,** 99-105.
VYAS, S., CHESARONE-CATALDO, M., TODOROVA, T., HUANG, Y.-H. & CHANG, P. 2013. A systematic analysis of the PARP protein family identifies new functions critical for cell physiology. *Nature Communications,* 4**,** 2240.
WALKER-GRAY, R., STENGEL, F. & GOLD, M. G. 2017. Mechanisms for restraining cAMP-dependent protein kinase revealed by subunit quantitation and cross-linking approaches. *Proceedings of the National Academy of Sciences,* 114**,** 10414-10419.
WANG, N., DONG, Q., LI, J., JANGRA, R. K., FAN, M., BRASIER, A. R., LEMON, S. M., PFEFFER, L. M. & LI, K. 2010. Viral induction of the zinc finger antiviral

protein is IRF3-dependent but NF-kappaB-independent. *J Biol Chem,* 285**,** 6080-90.
WANG, W., GÖTTE, B., GUO, R. & PYLE, A. M. 2023. The E3 ligase Riplet promotes RIG-I signaling independent of RIG-I oligomerization. *Nature Communications,* 14**,** 7308.
WEBB, P. A., JUSTINES, G. & JOHNSON, K. M. 1975. Infection of wild and laboratory animals with Machupo and Latino viruses. *Bull World Health Organ,* 52**,** 493-9.
WEI, R., WANG, J., SU, M., JIA, E., CHEN, S., CHEN, T. & NI, Y. 2018. Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Scientific Reports,* 8.
WEST, B. R., HASTIE, K. M. & SAPHIRE, E. O. 2014a. Structure of the LCMV nucleoprotein provides a template for understanding arenavirus replication and immunosuppression. *Acta Crystallographica Section D Biological Crystallography,* 70**,** 1764-1769.
WEST, B. R., HASTIE, K. M. & SAPHIRE, E. O. 2014b. Structure of the LCMV nucleoprotein provides a template for understanding arenavirus replication and immunosuppression. *Acta Crystallogr D Biol Crystallogr,* 70**,** 1764-9.
WHITMER, S. L., STRECKER, T., CADAR, D., DIENES, H.-P., FABER, K., PATEL, K., BROWN, S. M., DAVIS, W. G., KLENA, J. D. & ROLLIN, P. E. 2018. New lineage of Lassa virus, Togo, 2016. *Emerging infectious diseases,* 24**,** 599.
WIEBENGA, N. H., SHELOKOV, A., GIBBS, C. J., JR. & MACKENZIE, R. B. 1964. EPIDEMIC HEMORRHAGIC FEVER IN BOLIVIA. II. DEMONSTRATION OF COMPLEMENT-FIXING ANTIBODY IN PATIENTS' SERA WITH JUN'IN VIRUS ANTIGEN. *Am J Trop Med Hyg,* 13**,** 626-8.
WILDA, M., LOPEZ, N., CASABONA, J. C. & FRANZE-FERNANDEZ, M. T. 2008a. 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**,** 11454-11460.
WILDA, M., LOPEZ, N., CASABONA, J. C. & FRANZE-FERNANDEZ, M. T. 2008b. 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. *J Virol,* 82**,** 11454-60.
WILSON, E. B., YAMADA, D. H., ELSAESSER, H., HERSKOVITZ, J., DENG, J., CHENG, G., ARONOW, B. J., KARP, C. L. & BROOKS, D. G. 2013. Blockade of Chronic Type I Interferon Signaling to Control Persistent LCMV Infection. *Science,* 340**,** 202-207.
WOLFF, S., BECKER, S. & GROSETH, A. 2013. Cleavage of the Junin virus nucleoprotein serves a decoy function to inhibit the induction of apoptosis during infection. *J Virol,* 87**,** 224-33.
WORTHAM, N. C., AHAMED, E., NICOL, S. M., THOMAS, R. S., PERIYASAMY, M., JIANG, J., OCHOCKA, A. M., SHOUSHA, S., HUSON, L., BRAY, S. E., COOMBES, R. C., ALI, S. & FULLER-PACE, F. V. 2009. The DEAD-box protein p72 regulates ERα-/oestrogen-dependent transcription and cell growth, and is associated with improved survival in ERα-positive breast cancer. *Oncogene,* 28**,** 4053-4064.
WRIGHT, R., JOHNSON, D., NEUMANN, M., KSIAZEK, T. G., ROLLIN, P., KEECH, R. V., BONTHIUS, D. J., HITCHON, P., GROSE, C. F., BELL, W. E. & BALE, J. F., JR. 1997. Congenital lymphocytic choriomeningitis virus syndrome: a

disease that mimics congenital toxoplasmosis or Cytomegalovirus infection. *Pediatrics,* 100**,** E9.
WU, Z., DU, J., LU, L., YANG, L., DONG, J., SUN, L., ZHU, Y., LIU, Q. & JIN, Q. 2018. Detection of Hantaviruses and Arenaviruzses in three-toed jerboas from the Inner Mongolia Autonomous Region, China. *Emerg Microbes Infect,* 7**,** 35.
XU, X., PENG, R., PENG, Q., WANG, M., XU, Y., LIU, S., TIAN, X., DENG, H., TONG, Y., HU, X., ZHONG, J., WANG, P., QI, J., GAO, G. F. & SHI, Y. 2021. Cryo-EM structures of Lassa and Machupo virus polymerases complexed with cognate regulatory Z proteins identify targets for antivirals. *Nature Microbiology,* 6**,** 921-931.
XUE, G., BRACZYK, K., GONÇALVES-CARNEIRO, D., DAWIDZIAK, D. M., SANCHEZ, K., ONG, H., WAN, Y., ZADROZNY, K. K., GANSER-PORNILLOS, B. K., BIENIASZ, P. D. & PORNILLOS, O. 2022. Poly(ADP-ribose) potentiates ZAP antiviral activity. *PLoS Pathog,* 18**,** e1009202.
YAN, Y., TANG, Y. D. & ZHENG, C. 2022. When cyclin-dependent kinases meet viral infections, including SARS-CoV-2. *J Med Virol,* 94**,** 2962-2968.
YANG, E. & LI, M. M. H. 2020. All About the RNA: Interferon-Stimulated Genes That Interfere With Viral RNA Processes. *Frontiers in Immunology,* 11.
YANG, E., NGUYEN, L. P., WISHEROP, C. A., KAN, R. L. & LI, M. M. H. 2022. The Role of ZAP and TRIM25 RNA Binding in Restricting Viral Translation. *Frontiers in Cellular and Infection Microbiology,* 12.
YANG, H., WANG, J., JIA, X., MCNATT, M. W., ZANG, T., PAN, B., MENG, W., WANG, H. W., BIENIASZ, P. D. & XIONG, Y. 2010. Structural insight into the mechanisms of enveloped virus tethering by tetherin. *Proc Natl Acad Sci U S A,* 107**,** 18428-32.
YE, Y. & RAPE, M. 2009. Building ubiquitin chains: E2 enzymes at work. *Nature Reviews Molecular Cell Biology,* 10**,** 755-764.
YEKWA, E., KHOURIEH, J., CANARD, B., PAPAGEORGIOU, N. & FERRON, F. 2017. Activity inhibition and crystal polymorphism induced by active-site metal swapping. *Acta Crystallographica Section D Structural Biology,* 73**,** 641-649.
YONEYAMA, M. & FUJITA, T. 2008a. Structural Mechanism of RNA Recognition by the RIG-I-like Receptors. *Immunity,* 29**,** 178-181.
YONEYAMA, M. & FUJITA, T. 2008b. Structural mechanism of RNA recognition by the RIG-I-like receptors. *Immunity,* 29**,** 178-81.
YONEYAMA, M., ONOMOTO, K., JOGI, M., AKABOSHI, T. & FUJITA, T. 2015. Viral RNA detection by RIG-I-like receptors. *Current Opinion in Immunology,* 32**,** 48-53.
YORK, J., ROMANOWSKI, V., LU, M. & NUNBERG, J. H. 2004. The Signal Peptide of the Junin Arenavirus Envelope Glycoprotein Is Myristoylated and Forms an Essential Subunit of the Mature G1-G2 Complex. *Journal of Virology,* 78**,** 10783-10792.
YOUNG, P. R., CHANAS, A. C., LEE, S. R., GOULD, E. A. & HOWARD, C. R. 1987. Localization of an arenavirus protein in the nuclei of infected cells. *J Gen Virol,* 68 (Pt 9)**,** 2465-70.
YU, C. & HUANG, L. 2018. Cross-linking mass spectrometry (XL-MS): An emerging technology for interactomics and structural biology. *Analytical chemistry,* 90**,** 144.
ZAPATA, J. C. & SALVATO, M. S. 2013. Arenavirus variations due to host-specific adaptation. *Viruses,* 5**,** 241-78.

ZHANG, Y., LI, L., LIU, X., DONG, S., WANG, W., HUO, T., GUO, Y., RAO, Z. & YANG, C. 2013. Crystal structure of Junin virus nucleoprotein. *Journal of General Virology,* 94**,** 2175-2183.

ZHENG, X., WANG, X., TU, F., WANG, Q., FAN, Z. & GAO, G. 2017. TRIM25 Is Required for the Antiviral Activity of Zinc Finger Antiviral Protein. *Journal of Virology,* 91**,** e00088-17.

ZHU, Y., CHEN, G., LV, F., WANG, X., JI, X., XU, Y., SUN, J., WU, L., ZHENG, Y. T. & GAO, G. 2011. Zinc-finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation. *Proc Natl Acad Sci U S A,* 108**,** 15834-9.

ZHU, Y. & GAO, G. 2008. ZAP-mediated mRNA degradation. *RNA Biol,* 5**,** 65-7.

ZHU, Y., WANG, X., GOFF, S. P. & GAO, G. 2012. Translational repression precedes and is required for ZAP-mediated mRNA decay. *The EMBO Journal,* 31**,** 4236-4246.

ZIEGLER, C. M., EISENHAUER, P., BRUCE, E. A., WEIR, M. E., KING, B. R., KLAUS, J. P., KREMENTSOV, D. N., SHIRLEY, D. J., BALLIF, B. A. & BOTTEN, J. 2016. The Lymphocytic Choriomeningitis Virus Matrix Protein PPXY Late Domain Drives the Production of Defective Interfering Particles. *PLOS Pathogens,* 12**,** e1005501.

ZUO, Y. & DEUTSCHER, M. P. 2001. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Research,* 29**,** 1017-1026.